

Werner Roddeck

Einführung in die Mechatronik

5. Auflage

Einführung in die Mechatronik

Werner Roddeck

Einführung in die Mechatronik

5., überarbeitete und erweiterte Auflage

 Springer Vieweg

Werner Roddeck
Hochschule Bochum
Bochum, Deutschland

ISBN 978-3-658-15843-9
DOI 10.1007/978-3-658-15844-6

ISBN 978-3-658-15844-6 (eBook)

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Vieweg

© Springer Fachmedien Wiesbaden GmbH 1997, 2003, 2006, 2012, 2016

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen.

Lektorat: Thomas Zipsner

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier.

Springer Vieweg ist Teil von Springer Nature

Die eingetragene Gesellschaft ist Springer Fachmedien Wiesbaden GmbH

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Strasse 46, 65189 Wiesbaden, Germany

Vorwort

Der Begriff Mechatronik (engl. mechatronics) ist ein Kunstwort und wurde vor ca. 45 Jahren in Japan von einem Entwickler aus dem Bereich der Robotertechnik geprägt. Er setzt sich aus den beiden Namen der bekannten Disziplinen der Ingenieurwissenschaften - Mechanik oder Maschinenwesen (engl. mechanics) und Elektronik (engl. electronics) - zusammen. In den letzten 25 Jahren ist er auch in Deutschland sehr verbreitet worden und vor 23 Jahren wurde der erste Studiengang in Deutschland mit dem Abschluss „Dipl.-Ing. Mechatronik“ an der Hochschule Bochum eröffnet. Diesem Beispiel sind inzwischen viele Hochschulen gefolgt, indem sie Studiengänge „Mechatronik“ oder Studienrichtungen mit entsprechenden Studienanteilen einrichteten. Seit vielen Jahren ist der „Mechatroniker“ auch ein anerkannter Ausbildungsberuf, was die Nachfrage der Fachdisziplin „Mechatronik“ in der Industrie dokumentiert.

Die Notwendigkeit für diese neue Disziplin der Ingenieurwissenschaften ergibt sich aus der immer weiter zunehmenden Durchdringung maschinenbaulicher Produkte mit Anteilen aus dem Bereich der Elektrotechnik und der Informatik. Dabei werden nicht nur einzelne Komponenten konventioneller Produkte ersetzt, sondern die Entwicklung folgt einem ganz neuen Denkansatz. Man versucht, das Gesamtsystem zu verstehen und zu modellieren und wählt dann für die verschiedenen Teilsysteme solche Komponenten und Methoden aus, die zu einfacheren, preiswerteren und funktionaleren Gesamtsystemen führen. Dies ist nur dann möglich, wenn Produktentwickler fachübergreifende Kenntnisse aus allen genannten Bereichen der Ingenieurwissenschaften besitzen. Im Prinzip sind alle Methoden und Komponenten, die in der Mechatronik eingesetzt werden, als Teilgebiete anderer Ingenieurdisziplinen bereits bekannt und es gibt darüber umfangreiche Spezialliteratur. Bei Erscheinen der 1. Auflage dieses Buches fehlte jedoch eine Gesamtschau der Mechatronik, die die Teilgebiete in einen Zusammenhang setzte. Diese Lücke will das vorliegende Buch schließen und ist dabei so gestaltet, dass es von vielen Technikern und Ingenieuren sowohl im Studium, als auch in der Berufspraxis verwendet werden kann. Es behandelt das Thema beginnend mit den erforderlichen Analyse- und Synthesemethoden von Systemen, über die Beschreibung einsetzbarer Systemkomponenten bis zur Darstellung ausgeführter Beispiele für mechatronische Systeme.

Das zu Beginn der Hochschulausbildung etablierte Curriculum der Mechatronik hat sich im Laufe der letzten 20 Jahre gewandelt. So bestand es zu Beginn noch aus einer Zusammenstellung bereits bekannter Teilgebiete aus Maschinenbau und Elektrotechnik, von denen man annahm, dass die „Mechatronik“ durch Zusammenschau der Teilgebiete in den Köpfen der Studierenden entstehen würde. Da Nomenklatur und Regelwerke in den genannten Wissenschaftsgebieten aber sehr heterogen sind, trat das nicht selbstverständlich ein. Folgerichtig sind zwischenzeitlich Studienfächer ins Curriculum aufgenommen worden, die auch den Begriff „Mechatronik“ im Namen führen und in denen versucht wird, die Stoffintegration zum Bestandteil der Ausbildung zu machen.

Leider fehlt noch immer ein gemeinsamer Leitgedanke, der zu mehr Vereinheitlichung der Teilgebiete führen könnte. Mit zu den wichtigsten Fachgebieten der Me-

chatronik gehören die Modellbildung und Simulation von Systemen. Ich hatte bereits in der 1. Auflage dieses Buches vor 20 Jahren die objektorientierte Modellbildung als einen wichtigen, die Teilgebiete integrierenden Aspekt berücksichtigt. Leider standen zu diesem Zeitpunkt noch keine eingeführten und einfach einsetzbaren Software-Tools zur Verfügung. Heute hat sich das geändert und inzwischen hat die Methodik der Modellierung von mechatronischen Systemen mit Bondgraphen einen Stand erreicht, der es erlaubt, diese Methodik als die gesuchte Klammer für eine Homogenisierung der sehr heterogenen Teilgebiete anzusehen. Es stehen nun auch Programmsysteme und Literatur zur Verfügung, die den Umgang mit dieser Methodik sehr einfach machen. Im Hinblick darauf ist in dieser Auflage die Modellbildung mit Bondgraphen umfassender behandelt worden.

Da die Mechatronik ein sehr umfangreiches Wissensgebiet umspannt, können natürlich im Rahmen eines solchen Buches nicht alle Grundlagen behandelt werden. So sollten dem Leser mathematische Methoden wie die Differential-, Integral- und Matrizenrechnung bekannt sein. Ebenso sind grundlegende Kenntnisse der Physik in den Bereichen Mechanik und Elektrotechnik erforderlich. Aufgrund der Stofffülle können manche Themen nur angerissen werden, so dass der Leser im Bedarfsfall das entsprechende Wissen anhand der angeführten Literatur vertiefen muss.

Die behandelte Themenstellung erstreckt sich über viele Fachgebiete. Daher kann ich natürlich nicht Spezialist für jedes einzelne Thema sein. Ich bitte deshalb schon jetzt die Fachleute einzelner Fachgebiete, denen die Darstellung ihres Spezialgebietes zu kurz oder zu oberflächlich erscheint, um Nachsicht, da ich zugunsten der Gesamtschau an einigen Stellen Kompromisse eingehen musste.

Das Buch wendet sich an Studierende der Ingenieurwissenschaften mit Mechatronik als Studienschwerpunkt an Hochschulen, Berufsakademien und Universitäten, sowie an Ingenieure in der Praxis.

Im Übrigen bin ich den Kollegen, Mitarbeitern und Studierenden an der Hochschule Bochum dankbar, die mich mit Anregungen, Beispielen und Material zu bestimmten Themen unterstützt haben. Durch die Diskussion mit ihnen und durch die gemeinschaftliche Weiterentwicklung mechatronischer Studiengänge an der Hochschule Bochum konnte auch diese 5. Auflage gegenüber den Vorgängern noch klarer strukturiert, aktualisiert und den Erfordernissen eines Mechatronik-Studiums angepasst werden.

Meiner Frau Renate, die mir durch Ihre Unterstützung die Zeit für die Fertigstellung dieses Buches freimachte, bin ich zu großem Dank verpflichtet. Herrn Thomas Zipsner vom Springer Vieweg Verlag danke ich für die gute Zusammenarbeit und die Anregungen zur Weiterentwicklung des Buches.

Witten, August 2016

Werner Roddeck

Inhalt

1	Einleitung	
1.1	Entwicklung von Maschinenbau und Elektrotechnik	1
1.2	Entwicklung der Technik am Beispiel der Werkzeugmaschine	4
1.3	Mechatronik als neues Bindeglied	6
1.4	Maschinenbau und Elektrotechnik - grundsätzlich verschieden?	8
1.5	Unterschiede zwischen Maschinenbau, Elektrotechnik und Mechatronik	16
1.6	Teilgebiete der Mechatronik	21
2	Modellbildung technischer Systeme	25
2.1	Systembegriff	25
2.2	Verfahren der Modellbildung	28
2.2.1	Theoretische Modellbildung	30
2.2.1.1	Allgemein bekannte Modellvorstellungen	30
2.2.1.2	Vorgehensweise bei der Modellbildung	31
2.3	Klassifizierung dynamischer Systeme	38
2.4	Modellierung von Geometrie und Körpereigenschaften	41
2.4.1	Mehrkörpersysteme	41
2.4.2	Systeme mit elastischen Elementen	43
2.5	Modellierung elektrischer Komponenten	45
3	Dynamik mechanischer Systeme	50
3.1	Kinematik des Massenpunktes	50
3.2	Kinematik des starren Körpers	53
3.2.1	Die ebene Bewegung des starren Körpers	53
3.2.2	Die ebene Relativbewegung eines Punktes	59
3.2.3	Die Bewegung des starren Körpers im Raum	61
3.2.3.1	Rotation im Raum	62
3.2.3.2	Relativbewegung eines Punktes des starren Körpers	64
3.2.3.3	Darstellung der Bewegung des starren Körpers in Matrizenschreibweise	66
3.3	Bindungen in Mehrkörpersystemen	77
3.4	Kinetik	88
3.4.1	Impuls-, Schwerpunkt- und Drallsatz	88
3.4.2	Energiesatz	97
3.4.3	Die Prinzipien der Mechanik	100
3.4.3.1	Prinzip der virtuellen Arbeit	100
3.4.3.2	Lagrange'sche Bewegungsgleichungen	105
4	Schwingungen	112
4.1	Einmassenschwinger	112
4.1.1	Freie ungedämpfte Schwingungen	114
4.1.2	Freie gedämpfte Schwingungen	116
4.1.3	Erzwungene Schwingungen	120
4.1.3.1	Nichtperiodische Erregung	121
4.1.3.2	Harmonische Erregung	122
4.1.4	Nichtlineare Schwinger	128
4.2	Mehrmassenschwinger	130
4.3	Schwingungsanalyse	138
4.3.1	Reelle Form der Fourier-Reihe	140
4.3.2	Komplexe Form der Fourier-Reihe	144
4.3.3	Fourier-Transformation nichtperiodischer Funktionen	146
4.3.4	Diskrete Fourier-Transformation zur Analyse von Abtastsignalen	148

5	Sensoren	157
5.1	Messtechnik	160
5.1.1	Messgrößen und Maßeinheiten	160
5.1.2	Messgrößenaufnehmer und Messwertwandler	162
5.1.2.1	Messwertanpassung	164
5.1.2.2	Analog-/Digital-Wandler	167
5.1.3	Kenngößen von Messeinrichtungen	169
5.1.3.1	Statische Kenngößen	169
5.1.3.2	Dynamische Kenngößen	171
5.1.3.3	Fehlerkenngößen	173
5.2	Messeffekte	177
5.2.1	Widerstandseffekte	180
5.2.1.1	Ohmsche Widerstandseffekte	180
5.2.1.2	Piezowiderstandseffekt	183
5.2.2	Magnetische Effekte	184
5.2.2.1	Induktionsprinzip	184
5.2.2.2	Galvanomagnetische Effekte	185
5.2.2.3	Magnetoelastische Effekte	187
5.2.3	Kapazitive Effekte	188
5.2.4	Piezo- und Pyroelektrische Effekte	189
5.2.5	Optische Effekte	190
5.3	Sensoren für mechatronische Systeme	193
5.3.1	Bewegungssensoren	194
5.3.1.1	Positionssensoren	194
5.3.1.2	Geschwindigkeitssensoren	204
5.3.1.3	Beschleunigungssensoren	206
5.3.2	Kraft- und Momentensensoren	209
6	Aktoren	212
6.1	Klassische Aktoren	214
6.1.1	Elektromotorische, rotierende Antriebe	214
6.1.1.1	Gleichstrommotoren	216
6.1.1.2	Drehfeldmotoren	227
6.1.1.3	Asynchronmotoren	229
6.1.1.4	Schrittmotoren	234
6.1.2	Elektromotorische Linearantriebe	236
6.1.3	Fluidische Aktoren	238
6.1.3.1	Pneumatische Aktoren	238
6.1.3.2	Hydraulische Aktoren	239
6.1.3.3	Geschwindigkeitsverstellung von hydraulischen Aktoren	242
6.2	Neuartige Aktoren	248
6.2.1	Piezoelektrische Aktoren	250
6.2.2	Aktoren aus Formgedächtnislegierungen	253
6.2.3	Dehnstoffaktoren	256
6.2.4	Mikrostrukturierte Aktoren	257
7	Automatisierungstechnik	260
7.1	Automatisierungskonzepte	260
7.1.1	Intelligente Maschinen	262
7.1.2	Steuerung und Regelung	264
7.1.3	Schlussfolgern und regelbasiertes Wissen	266
7.1.4	Autonome intelligente Agenten	267
7.1.5	Lernen und Mustererkennung	268
7.1.6	Architektur intelligenter Maschinen	269
7.1.6.1	Hierarchien	270
7.1.6.2	Netzwerke	271

7.1.6.3	Schichtarchitekturen	277
7.2	Steuerungstechnik	278
7.2.1	Boole'sche Algebra	281
7.2.1.1	Kombinatorische Steuerungen	285
7.2.1.2	Sequentielle Steuerungen	288
7.2.2	Probleme der Modellbildung digitaler Systeme	291
7.2.3	Mehrwertige und unscharfe Logik (Fuzzy Logic)	294
7.2.4	Neuronale Netzwerke	299
7.2.4.1	McCulloch-Pitts-Neuron	301
7.2.4.2	Perceptron	302
7.2.4.3	Backpropagation-Netzwerk	305
7.3	Regelungstechnik	307
7.3.1	Beschreibung und Analyse regelungstechnischer Systeme	308
7.3.1.1	Laplace-Transformation	310
7.3.1.2	Blockschaltbilder	318
7.3.1.3	Frequenzgang und Ortskurve	324
7.3.1.4	Verschiedenartige Übertragungssysteme	327
7.3.1.5	Frequenzkennlinien	333
7.3.1.6	Zustandsraumdarstellung	340
7.3.1.7	Regler	345
7.3.1.8	Stabilität von Regelkreisen	352
7.3.1.9	Systemidentifikation	365
7.3.2	Synthese von Regelkreisen	370
7.3.2.1	Spezifikationen	371
7.3.2.2	Einstellregeln	373
7.3.2.3	Mehrschleifige Regelkreise	376
7.4	Prozessdatenverarbeitung mit Mikrorechnern	380
7.4.1	Mikrorechner	381
7.4.1.1	Aufbau von Mikrorechnern	382
7.4.1.2	Software für Mikrorechner	392
7.4.2	Anwendungsspezifische Prozessoren und Bauelemente	393
8	Simulation	397
8.1	Numerische Integration	399
8.2	Modellbildung mit Bondgraphen	407
8.2.1	Elemente von Bondgraphen	408
8.2.2	1-Port Bauelemente	412
8.2.2.1	1-Port R-Elemente	412
8.2.2.2	1-Port C-Elemente	413
8.2.2.3	1-Port I-Elemente	414
8.2.2.4	1-Port Quellen	415
8.2.3	2-Port Bauelemente	416
8.2.3.1	2-Port Transformer	416
8.2.3.2	2-Port Gyrator	417
8.2.4	Multi-Ports	418
8.2.5	Erstellung von Modellen komplexerer Systeme	420
8.3	Simulationssysteme	422
8.3.1	Simulationssprachen	422
8.3.2	Simulation elektrischer Schaltungen	423
8.3.3	Simulation mechanischer Systeme	424
8.3.4	Modellbeschreibung mit Blockschaltbild-Editoren	425
8.3.5	Objektorientierte Modellbildung	430
8.3.5.1	20-Sim	431
8.3.5.2	CAMel-View	436
8.3.6	Hardware-in-the-Loop, Software-in-the-Loop	438
8.3.6.1	Hardware-in-the-Loop	439
8.3.6.2	Software-in-the-Loop	439

8.3.6.3	Kopplung von Modellen und Prototypen	439
8.3.7	Simulationssysteme für Industrieroboter	440
9	Mechatronische Systeme	443
9.1	Wann ist der Einsatz der Mechatronik sinnvoll?	443
9.2	Entwicklung mechatronischer Systeme	446
9.3	Mechatronische Teilsysteme	455
9.3.1	Magnetlager	456
9.3.2	Anti-Blockier-System (ABS)	461
9.3.3	Aktives Fahrwerk	463
9.3.3.1	Aktive Federung mit Hydrozylinder	465
9.3.3.2	Aktive Federung mit Hydrozylinder und aktivem Tilger	466
9.3.4	Mechatronische Anwendungen bei Industrierobotern	470
9.3.4.1	Nachführen eines Roboterarms an einer Freiformfläche	471
9.3.4.2	Zusätzliche Bewegungsachsen für Industrieroboter	476
9.4	Mechatronische Gesamtsysteme	483
9.4.1	Hexapodenkonzepte	483
9.4.2	Fahrrad mit aktiver Neigetechnik	486
	Literaturverzeichnis	493
	Sachwortverzeichnis	497

1 Einleitung

Der Begriff Mechatronik (engl. Mechatronics) ist ein Kunstwort und wurde vor ca. 40 Jahren in Japan von einem Entwickler aus dem Bereich der Robotertechnik geprägt. Es setzt sich aus den beiden Namen der bekannten Disziplinen der Ingenieurwissenschaften - Mechanik oder Maschinenwesen (engl. Mechanics) und Elektronik (engl. Electronics) - zusammen. Damit ist dieser Name bereits ein Programm und deutet an, dass die Mechatronik Inhalte der beiden oben genannten Disziplinen zusammenfügt. In Japan legt man den Begriff Mechatronics sehr weit aus. In Europa wurde eine eher enge Definition geprägt, die den Eindruck nahelegt, es handle sich bei dem Begriff um eine neue Wissenschaftsdisziplin. Diese Definition lautet:

Mechatronik ist ein interdisziplinäres Gebiet der Ingenieurwissenschaften, das auf den klassischen Disziplinen Maschinenbau, Elektrotechnik und Informatik aufbaut. Ein typisches mechatronisches System nimmt Signale auf, verarbeitet sie und gibt Signale aus, die es z. B. in Kräfte und Bewegungen umsetzt [1.1]. Eine ähnliche Definition wird auch in [1.2] angeführt:

Mechatronics is the synergetic integration of mechanical engineering with electronic and intelligent computer control in the design and manufacturing of industrial products and processes (IEEE/ASME Transactions on Mechatronics 1996).

Es scheint hinter dem Aufkommen dieses Begriffes jedoch nicht nur die Notwendigkeit zu stehen, für neuartige Produkte eine neue Beschreibungsmethode oder ein neues Denken zu schaffen. An dem Wort interdisziplinär erkennt man ein tieferes Bedürfnis nach Zusammenarbeit von Disziplinen in Wissenschaft, Forschung und Ausbildung, die heute noch meist strikt voneinander getrennt existieren. Die Ursachen für diese Trennung sind einerseits historisch bedingt und andererseits durch die rasante Entwicklung der Elektrotechnik geprägt. Viele technische Produkte des Maschinenbaus sind heute in so hohem Maße mit elektrotechnischen Komponenten ausgestattet, dass eine interdisziplinäre Zusammenarbeit der ingenieurwissenschaftlichen Gebiete Maschinenbau und Elektrotechnik geradezu zwingend erforderlich ist.

Warum besteht aber überhaupt die Notwendigkeit der Zusammenführung der Disziplinen? Wie war es dazu gekommen, dass ein Bedürfnis entstand, getrennte Disziplinen wieder näher zusammenzubringen?

1.1 Entwicklung von Maschinenbau und Elektrotechnik

In der Renaissance und bis ins 18. Jahrhundert hinein waren die Ingenieurwissenschaften durch Künstler-Ingenieure geprägt, Vertreter einer vorwiegend "höfischen" Technik. Sie war auf wenige Bereiche beschränkt und bei den von ihnen entwickelten Objekten bildeten Funktion und künstlerische Formgebung eine Einheit. Erst in der zweiten Hälfte des 18. Jahrhunderts, in der die Erfindung der selbstständig laufenden Dampfmaschine durch James Watt den Beginn der industriellen Revolution signalisiert, entstand eine "bürgerliche" Technik, deren Entwicklung durch Ökonomie und Funktionalität bestimmt wurde. Die durch die industrielle Revolution ausgelöste sprunghafte Zunahme technischen Wissens verlangte schon bald nach gut ausgebildeten Technikern vor allem im Bereich des Bergbaus, der Architektur aber auch im Militärwesen. Die ersten Bildungseinrichtungen, die sich ausschließlich der

Ausbildung von "Technikern" widmeten, waren die Polytechnischen Schulen und Bergakademien. In Deutschland und im deutschsprachigen Raum wurden solche Einrichtungen an folgenden Orten gegründet: Prag (1806), Wien (1815), Karlsruhe (1825), Braunschweig (1814/35), Clausthal (1810) und München (1827). Sie orientierten sich alle am Modell der französischen École Polytechnique in Paris (1794). Der Durchbruch für eine wissenschaftliche Technik erfolgte durch die Industrialisierung um die Mitte des 19. Jahrhunderts. Mit ihr ging eine Professionalisierung der technischen Berufe in entsprechend qualifizierenden Fachstudiengängen einher. Zu diesem Zeitpunkt mehrten sich die Stimmen, den Polytechnischen Anstalten Hochschulcharakter und Universitätsstrukturen zu geben. Eine wichtige Rolle spielte hierbei die Schweiz, die in Zürich 1855 die erste nationale und technische Hochschule schuf. Vorbild der neuen Eidgenössischen Polytechnischen Schule war diejenige in Karlsruhe; im Gegensatz zu dieser erhielt jedoch die Züricher die deutsche Universitätsstruktur mit Fakultäts-einteilung, Senats- und Rektoratsverfassung. Das Beispiel dieser seit 1911 den Hochschulnamen tragenden Anstalt führte in Deutschland zur Aufwertung der Polytechniken zu "Hochschulen", auch wenn sich die Bezeichnung "Technische Hochschule" erst allmählich durchsetzte. Von 1865 bis zum ersten Weltkrieg wurden elf Technische Hochschulen durch Umwandlung älterer Institutionen oder durch Neugründung errichtet: 1864 Graz, 1865/85 Karlsruhe, 1868/77 München, 1870/79 Aachen, 1872/79 Braunschweig, 1876/90 Stuttgart, 1877/95 Darmstadt, 1879 Berlin, 1880 Hannover, 1890 Dresden, 1904 Danzig, 1910 Breslau.

Typisch für deren Einteilung in Fakultäten, also voneinander unabhängigen Wissenschaftsdisziplinen, ist die der TH Aachen im Jahr 1880:

1. Abteilung für Architektur
2. Abteilung für Bauingenieurwesen
3. Abteilung für Maschineningenieurwesen
4. Abteilung für Bergbau- und Hüttenkunde und für Chemie
5. Abteilung für allgemeine Wissenschaften, insbesondere für Mathematik und Naturwissenschaften

Die Elektrotechnik, die einen wesentlichen Aufschwung durch die Erfindung der Dynamomaschine durch Werner Siemens im Jahre 1867 genommen hatte, war anfangs als wissenschaftliche Disziplin an den Technischen Hochschulen nicht vertreten. Bei der Herausbildung der Elektrotechnik als selbstständige technikwissenschaftliche Disziplin handelt es sich um einen längeren Prozess, der sich etwa zwischen 1880 und dem ersten Weltkrieg vollzog. Dabei kam es zuerst zu einer relativ kurzen Anlehnung an die Physik und anschließend zu einer stärkeren Anlehnung an den Maschinenbau.

Die akademische Etablierung der Elektrotechnik war ausschließlich mit der entstehenden Starkstromtechnik verbunden. Die Schwachstromseite, vertreten durch die Telegraphie und die Telefonie, spielte keine wesentliche Rolle, da die im Maschineningenieurwesen angesiedelte Elektrotechnik sich fast ausschließlich mit Elektromaschinen und der Übertragung elektrischer Energie befasste. Als einer der ersten hatte Werner Siemens in einem Vortrag am 27. Dezember 1881 die Einrichtung elektrotechnischer Lehrstühle angeregt. Nach den ersten Lehrstuhlgründungen (Aachen 1886, Berlin 1884, Darmstadt 1882, Wien 1884) kam es zu einer ausgedehnten Diskussion darüber, ob Elektrotechnik eigentlich ein eigenes Studium oder nur eine Zusatzqualifikation, vor allem für Maschinenbauer, sein sollte. In dieser

Diskussion bezog auch Werner Siemens Stellung und verdeutlichte seinen Standpunkt, dass Elektrotechnik nur als Zusatzqualifikation für alle Ingenieurstudenten dienen solle. Daneben stehe weiterhin als gleichwertiger Ausbildungsgang die Aneignung von Kenntnissen in der industriellen Praxis und durch Selbststudium. Andere Schlussfolgerungen - was die umstrittene Stellung der Elektrotechnik zwischen Physik und Maschinenbau verdeutlicht - zog der gerade nach Karlsruhe berufene Physiker Ferdinand Braun 1883. Braun argumentierte, dass die physikalischen Grundlagen der Elektrotechnik noch so ungesichert seien, dass man den Studenten keine zeitbeständigen Grundlagen für ihr Berufsleben vermitteln könne.

Diese Äußerungen kamen aber nach der Institutionalisierung der Elektrotechnik zu spät, um der Entwicklung eine andere Richtung zu geben. Die Eigendynamik, die mit der Einrichtung eines Faches verbunden ist, brachte es mit sich, dass Elektrotechnik mehr oder weniger schnell an allen Hochschulen als eigenständiges Studienfach etabliert wurde, das man mit dem Diplom abschließen konnte. Allerdings wies das Studium letztendlich große Überschneidungen mit dem Maschinenbaustudium auf. In der ersten Hälfte des Studiums wurde eine breite Maschinenbaugrundlage gelegt, bevor eine allmähliche Spezialisierung auf elektrotechnische Probleme einsetzte. Die Abteilungen für Elektrotechnik blieben meist noch lange in die Fakultät für Maschinenwesen eingegliedert.

In den meisten Fällen verblieb die Elektrotechnik bis zum Ende des zweiten Weltkriegs in der Fakultät für Maschinenwesen. Erst nach der Wiederaufnahme ihrer Lehrtätigkeit am Anfang der 50er Jahre wurden dann an den meisten Technischen Hochschulen eigene Fakultäten für Elektrotechnik eingerichtet (Berlin 1955).

An den meisten Maschinenbau- und Elektrotechnikfakultäten wurde aufgrund der technischen Entwicklung inzwischen ein bunter Strauß von verschiedenartig spezialisierten Ingenieuren ausgebildet. Die immer tiefer greifende Spezialisierung der technischen Wissensgebiete führte dann aber am Anfang der 60er Jahre zu der Tendenz, die Aufsplittung der Wissenschaften an Universitäten und Technischen Hochschulen generell zurückzuführen und im Bereich der Ingenieurwissenschaften mehr universell einsetzbare Ingenieure auszubilden. Die erste Universitätsneugründung in der Bundesrepublik Deutschland, die Ruhr-Universität Bochum, erhielt daher als erste Universität neben den traditionellen Fakultäten alle ingenieurwissenschaftlichen Abteilungen (Bauingenieurwesen, Elektrotechnik und Maschinenbau). Zwar waren auch hier die Abteilungen konsequent voneinander getrennt, aber innerhalb einer Abteilung wurde nur ein Ingenieurdiplom verliehen. Der Abteilung für Maschinenbau war ein Institut für Automatisierungstechnik angegliedert, das die immer stärker ausgeprägten Ansprüche an elektrische Komponenten im Maschinenbau abdeckte und damit die Rolle übernahm, die die Elektrotechnik in den alten Maschinenbaufakultäten gespielt hatte.

Die gleiche Entwicklung fand bei den Gesamthochschulen statt, in denen neben den traditionellen Universitätsfächern auch die Ingenieurwissenschaften vertreten sind.

Nach Gründung der Fachhochschulen in der Bundesrepublik, die aus den Vorläufereinstitutionen "Ingenieurschulen" hervorgingen, wurden an diesen teilweise sogar eigene Studienrichtungen "Automatisierungstechnik" eingeführt, die die beiden Zweige der Ingenieurwissenschaften Maschinenbau und Elektrotechnik wieder vereinigten.

Am Ende der 80er Jahre kam nun der Begriff der Mechatronik auf und hat an verschiedenen Hochschulen zur Gründung neuer Studienrichtungen oder Studien-

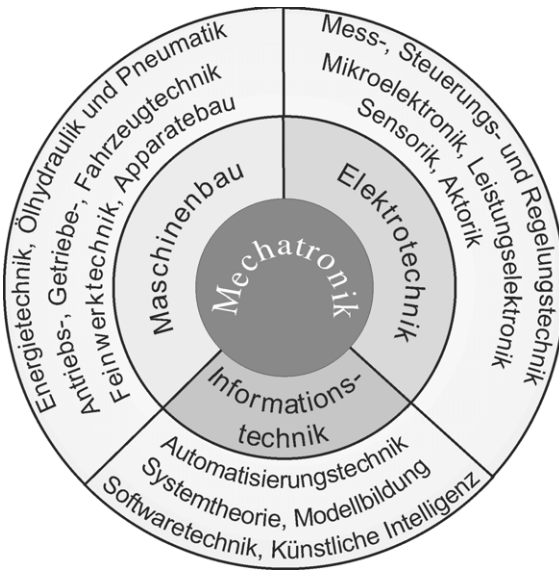


Bild 1-1: Mechatronik-Synergie verschiedener Disziplinen

schwerpunkte geführt. Schaut man sich die Studienpläne von Mechatronikstudiengängen an, so sieht man leicht, was dahintersteckt (Bild 1-1). Es gibt keine wesentlich neuen Studieninhalte, die nicht auch schon in anderen Hochschulstudiengängen vorher vorhanden waren. Aber man hat die wesentlichen Grunddisziplinen von Maschinenbau, Elektrotechnik und Informatik wieder zusammengeführt, so dass nicht Spezialisten für eine bestimmte Technologie, sondern Universalisten ausgebildet werden, die in der Lage sind, heutige hochkomplexe Produkte wie Industrieroboter, Werkzeugmaschinen, aktive Fahrwerkskomponenten und ähnliches in Entwicklung und Anwendung zu beherrschen. Damit scheint sich der Weg anzubahnen, das wieder zusammenzubringen, was durch die historische und technische Entwicklung getrennt wurde.

1.2 Entwicklung der Technik am Beispiel der Werkzeugmaschine

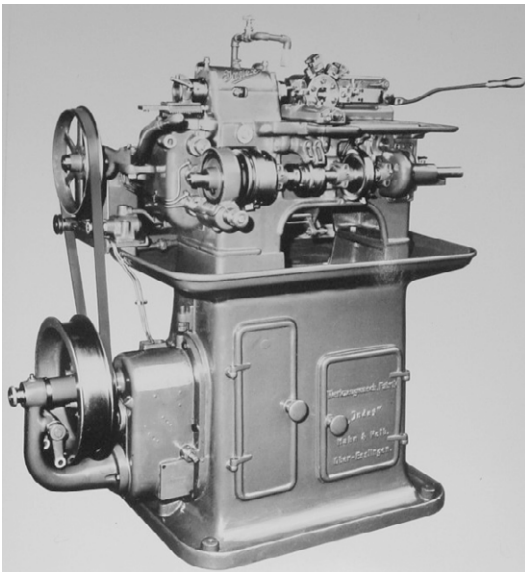


Bild 1-2: Drehautomat aus dem Jahre 1914 mit Riemenvorlege

Am Beispiel der Werkzeugmaschine lässt sich leicht aufzeigen, wie die technische Entwicklung, ausgehend von einem anfänglich rein maschinenbaulichen Produkt, über einen Zwitterzustand mit elektrischen Antrieben als einzige elektrische Komponenten, zu einem mechatronischen System führte.

In Bild 1-2 ist ein mechanischer Drehautomat aus dem Jahr 1914 mit Riemenvorlege der Fa. Index abgebildet. Der Hauptantrieb der Arbeitsspindel erfolgt über einen Treibriemen von einer Antriebsmaschine, die wahlweise noch eine Dampfmaschine oder ein Elektromotor sein konnten. Solche leistungsstarken Antriebsmaschinen versorgten gleichzeitig jeweils mehrere Werkzeugmaschinen über lange Transmissionswellen mit Antriebsenergie. Bild 1-3 zeigt einen Blick in eine

solche Dreherei der Werkzeugmaschinenfabrik Schieß in Düsseldorf. Drehzahlveränderungen der Arbeitsspindel waren nur durch Umlegen des von der Transmissionswelle kommenden Antriebsriemens möglich. Die Vorschubbewegung wird von der Arbeitsspindel über einen Wechselradsatz und ein mechanisches Schaltgetriebe durch Drehen der Vorschubspindel erzeugt, Zustellbewegungen werden über Kurvenscheiben vorgegeben. Eine solche Werkzeugmaschine enthält praktisch noch keine elektrotechnischen Komponenten, alle Bewegungen sind mechanisch automatisiert.

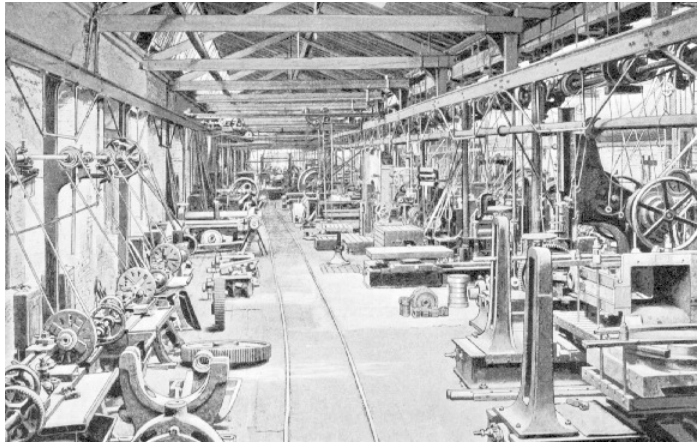


Bild 1-3: Dreherei der Werkzeugmaschinenfabrik Schieß in Düsseldorf

Der im Bild 1-4 dargestellte Drehautomat der Fa. Index aus dem Jahre 1952 hat sich rein äußerlich noch nicht sehr von ihrem Vorläufer aus dem Jahr 1914 entfernt. Wesentlicher Unterschied ist der nun in die Maschine eingebaute elektrische Hauptantrieb in Form eines Drehstrom-Asynchronmotors. Dieser ersetzt den Antrieb über die Transmissionswelle, so dass jetzt verschiedene Maschinen völlig unabhängig voneinander arbeiten können. Für die Motor- und Kupplungsschaltung gibt es einige wenige elektrische Komponenten, aber alle Bewegungen werden weiterhin über mechanische Schaltgetriebe und Kurvenscheiben erzeugt. Drehzahländerungen der Spindeln werden durch umschaltbare Zahnradsätze oder durch Polumschaltung des Hauptantriebsmotors erzeugt. Die elektrischen Komponenten dieser vollautomatisch arbeitenden Werkzeugmaschine dienen also im wesentlichen nur der Bereitstellung von Antriebsenergie.

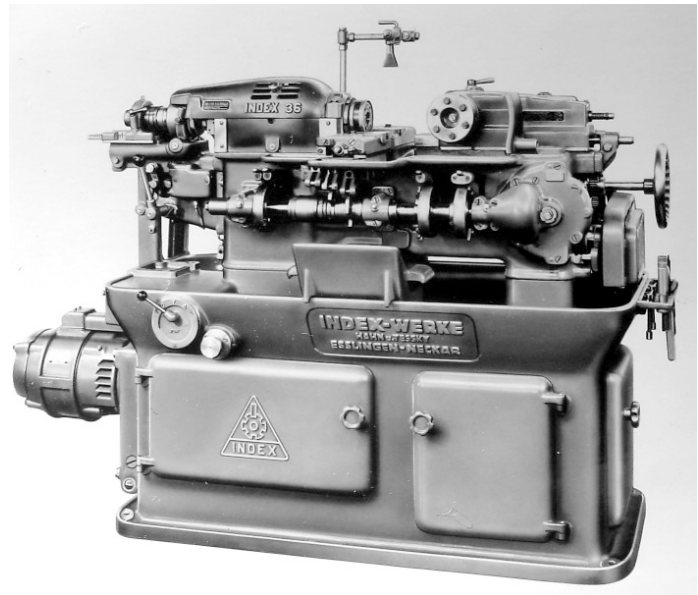


Bild 1-4: Drehautomat aus dem Jahre 1952 (Index)

Drehzahländerungen der Spindeln werden durch umschaltbare Zahnradsätze oder durch Polumschaltung des Hauptantriebsmotors erzeugt. Die elektrischen Komponenten dieser vollautomatisch arbeitenden Werkzeugmaschine dienen also im wesentlichen nur der Bereitstellung von Antriebsenergie.

In Bild 1-5 ist nun eine moderne NC-Drehmaschine dargestellt. Äußerlich unter-

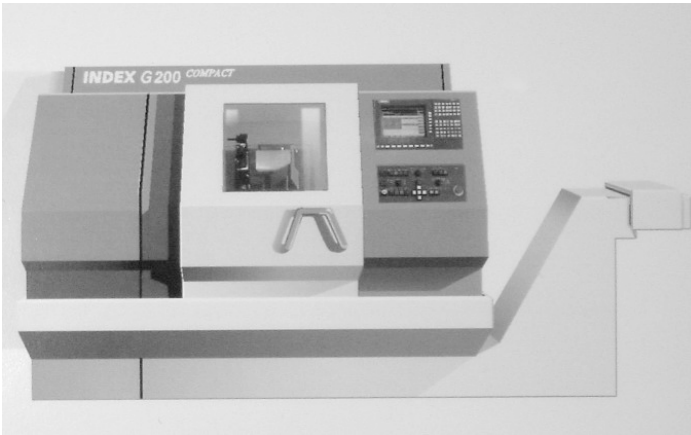


Bild 1-5: Moderne CNC-Drehmaschine (Index)

scheidet sie sich in ihrer vollgekapselten Form zwar stark von den beiden Vorläufermodellen, aber die eigentlichen, wesentlichen Unterschiede sind gar nicht sichtbar. Bei dieser Maschine kann in allen Achsen vollautomatisch und im Funktionszusammenhang positioniert werden. Dies setzt eine stufenlose Drehzahl- und Lageregelung aller Bewegungsachsen durch Servokreise vo-

raus. Auch die Hauptspindel ist stufenlos drehzahlregelbar und gegebenenfalls kann sie auch in ihrer Winkelstellung positioniert werden. Eine solche Maschine ist nun mit den verschiedenartigsten elektrischen und elektronischen Systemen (Computersteuerung, elektronisch drehzahlgeregelte elektrische Antriebe, elektrische Wegmesssysteme, u. v. m.) ausgestattet. Dies zeigt die dramatische Entwicklung der Elektrotechnik und ihren heutigen Anteil am Werkzeugmaschinenbau in den letzten 40 Jahren. Ein solches System kann nun in allen Phasen von der Konstruktion, Fertigung und Inbetriebnahme, Produktionsbetrieb bis zu Wartung und Instandhaltung nur noch von Personen mit entsprechender elektrotechnischer Ausbildung gehandhabt werden.

1.3 Mechatronik als neues Bindeglied

Es entsteht nun die Frage, ob eine NC-Drehmaschine, wie sie in Abschnitt 1.2 gezeigt wurde, schon ein mechatronisches System ist. Puristen verneinen dies und sagen, dass nur die über die "normalen" Fertigungsfunktionen hinausgehenden, sensorgeführten Möglichkeiten wie selbsteinstellende Werkzeuge oder adaptive Regelungen eine solche NC-Maschine zum mechatronischen System machen. Die oben aufgeführte Definition von Mechatronik umfasst jedoch auch schon die normale NC-Maschine, da hier die Sensoren "Wegmesssysteme" Lagemesswerte aufnehmen, die in der Steuerung verarbeitet werden (Soll- Istwertvergleich) und dadurch Positionierbewegungen durchgeführt werden können. Noch mehr gilt das für die weitaus komplexeren Universalwerkzeugmaschinen "Industrieroboter", für deren Bau und Einsatz heute unbedingt mechatronisch ausgebildetes Personal benötigt wird. Interessant ist hier, dass der Begriff **Mechatronik** gerade im Zusammenhang mit Industrierobotern in Japan entstanden ist, weil die Notwendigkeit der interdisziplinären Zusammenarbeit zwischen Maschinenbau, Elektrotechnik und Informatik sich besonders an solchen **mechatronischen** Geräten zeigte.

Durch eine breit angelegte, interdisziplinäre Mechatronikausbildung wird der Ingenieur in die Lage versetzt, konventionelle maschinenbauliche Lösungen zu überdenken und einfachere, meist kostengünstigere Kombinationslösungen von maschinenbaulichen und elektrotechnischen Komponenten einzusetzen.

Als ein Beispiel, wie man eine aufwendige **mechanische** Lösung durch eine **me-**

chatronische Lösung vom Aufwand und von den Kosten her optimieren kann, möge die im folgenden beschriebene Wälzstoßmaschine zur Zahnradfertigung dienen [1.3]. Bild 1-6 zeigt das Getriebschema einer solchen Wälzstoßmaschine, mit der man gerad- und schrägverzahnte Außen- und Innenverzahnungen herstellen kann. Mit einem Schneidrad bestimmter Zähnezahl als Verzahnungswerkzeug kann man auf einer solchen Maschine Werkräder beliebiger Zähnezahl vom gleichen Modul herstellen. Das Schneidrad hat dabei die Form eines Zahnrades und ist an der Unterseite schräg geschliffen und angeschärft, um in axialer Richtung schneidfähig zu sein. Der Fertigungsverfahren

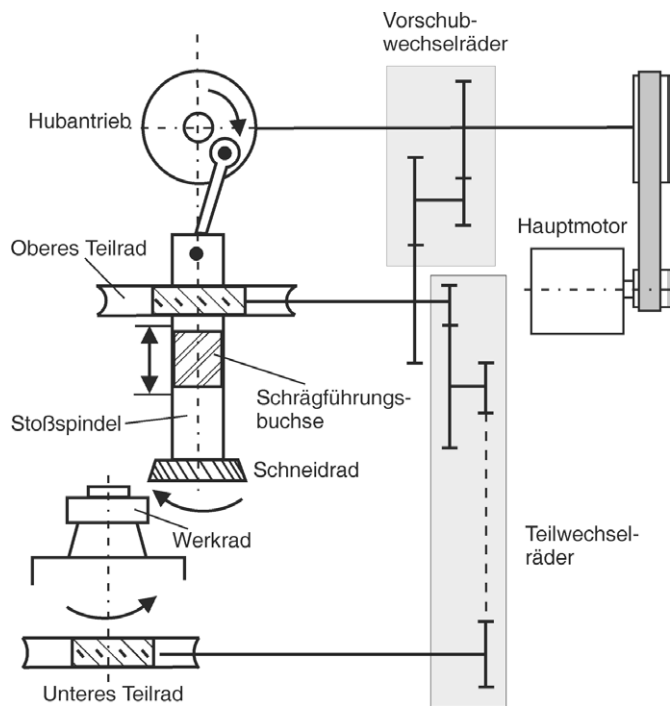


Bild 1-6: Getriebschema einer Wälzstoßmaschine

läuft so ab, dass Schneidrad und Werkrad durch einen mechanischen Getriebezug über Wechselräder synchronisiert aufeinander abrollen und das Schneidrad gleichzeitig noch für den Zerspanvorgang eine oszillierende Hubbewegung ausführt. Die erforderlichen Drehzahlen von Schneidrad und Werkrad können über Wechselräder vorgegeben werden und werden durch Schneckengetriebe, die sog. **Teilräder**, auf diese übertragen.

Die Komplexität des mechanischen Getriebes nimmt weiter zu, wenn die in Bild 1-7 gezeigte Ritzelwelle mit drei Verzahnungen gefertigt werden soll. Da sie in einer Aufspannung gefertigt werden muss und alle drei Verzahnungen verschiedene Zähnezahlen haben, muss im Getriebezug, der das obere und untere Teilrad miteinander synchronisiert, bei einer mechanischen Lösung für jede neue Verzahnung über eine Schaltkupplung jeweils ein anderer Teilwechselradsatz angewählt werden.

Denkt man nun mechatronisch, so kann man den ganzen Getriebezug mit den schaltbaren Wechselradsätzen und gegebenenfalls sogar die Teilräder weglassen (Bild 1-8). Statt dessen treibt man Schneid- und Werkrad mit jeweils einem elektrischen Lageregelkreis an, welche von der Steuerung die entsprechenden Drehzahlsoll-

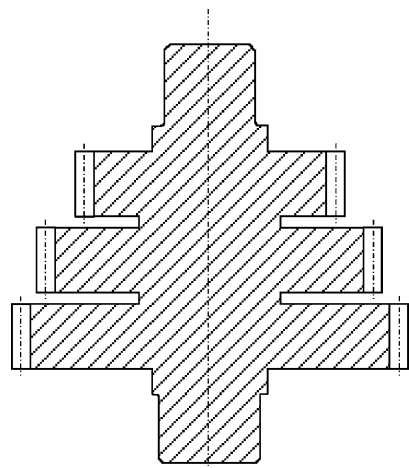


Bild 1-7: Ritzelwelle mit drei Verzahnungen

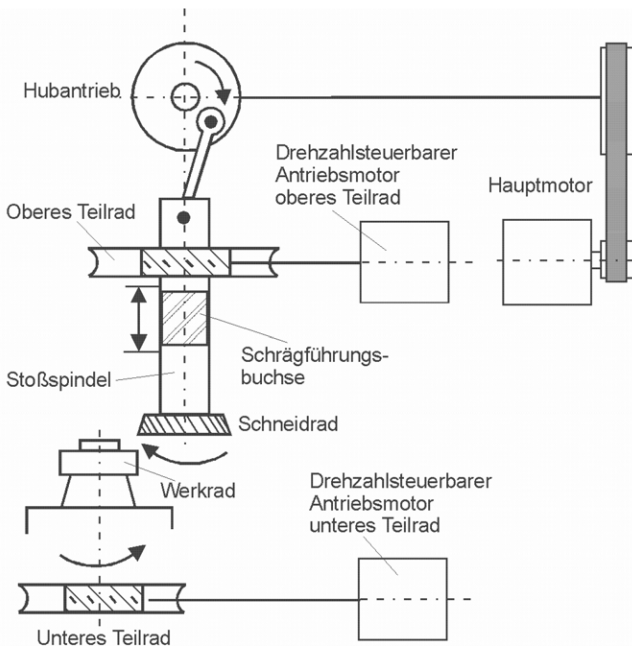


Bild 1-8: Getriebeschema einer Wälzstoßmaschine mit Einzelantrieben

werte, die zu den Zähnezahlen passen, vorgegeben bekommen. Zwar muss die Synchronsteuerung von Schneid- und Werkrad etwas anders aussehen als bei konventionellen Lageregelkreisen in NC-Maschinen, aber dies ist elektronisch relativ leicht zu realisieren. Außer der Einsparung des Getriebezuges zwischen den Teilrädern kann man auch auf die Bereithaltung eines größeren Vorrates verschiedener Wechsellräder zur Erzeugung beliebiger Zähnezahlen verzichten. Weiterhin können nun, da Schneid- und Werkrad zur Einhaltung einer gleichmäßigen Teilung nicht mehr zwangssynchronisiert sind, auf einer solchen Maschine auch Sonderverzahnungen mit un-

gleichförmiger Teilung erzeugt werden. Es wurde in diesem Beispiel also durch einen neuen Denkansatz unter Berücksichtigung elektrotechnischer Möglichkeiten eine kostengünstigere und vielseitigere Lösung gefunden. Die interdisziplinäre Mechatronikausbildung muss daher zum Ziel haben, Techniker und Ingenieure heranzubilden, die in dieser Weise in allen Richtungen offen sind, auch unkonventionelle Lösungen fachübergreifend einzusetzen.

1.4 Maschinenbau und Elektrotechnik - grundsätzlich verschieden?

Objekte, mit denen sich die Mechatronik befasst, sind in der Regel dynamische Systeme¹. Betrachtet man die Objekte, mit denen sich Maschinenbau und Elektrotechnik befassen rein phänomenologisch, d. h. nach dem Aussehen, so meint man eine deutliche Trennung zwischen diesen erkennen zu können. Beispiele für ein maschinenbauliches Objekt (Feder - Masse - Dämpfer - System) und ein elektrotechnisches Objekt (Kombination aus Kondensator - Spule - Widerstand) sind in Bild 1-9 dargestellt.

Erfolgt die Betrachtung der Objekte mehr **systemisch** (s. Kapitel 2), so verschwinden diese Unterschiede. Die Betrachtungsweise, die Eigenschaften eines dynamischen Systems nicht aufgrund des augenscheinlichen, körperlichen Aufbaus, sondern aufgrund der Eigenschaften eines mathematischen Modells des Systems zu beschreiben, hat sich vorrangig in der Regelungstechnik entwickelt. Man betrachtet

¹ Kennzeichen eines mechatronischen Systems.

jedes dynamische System als **Black Box** (Bild 1 - 10), in die Eingangsgrößen hineingehen (Ursachen für Veränderungen) und Ausgangsgrößen (Wirkungen von Veränderungen) herauskommen. Diese Größen können je nach System beliebige physikalische Größen sein, wie Kräfte, Drücke, Massenströme, elektrische Spannungen, elektrische Ströme. Selbst dynamische Systeme wie biologische Systeme oder Volkswirtschaften lassen sich mit Einschränkungen unter Verwendung der gleichen Methodik behandeln.

Das System, besser gesagt seine Übertragungseigenschaften vom Eingang zum Ausgang, wird durch ein mathematisches Modell abgebildet (Inhalt der Black Box). Dieses besteht in der Regel aus einer oder mehreren Differentialgleichungen und zugehörigen Rand- bzw. Anfangsbedingungen. Die mathematischen Modellierungen und Vorgehensweisen werden später noch ausführlich behandelt werden, um einen ersten Eindruck von der Themenstellung zu bekommen, reicht an dieser Stelle eine einfache Gegenüberstellung der mathematischen Zusammenhänge.

Viele einfache physikalische Systeme lassen sich durch lineare Differentialgleichungen der folgenden Art beschreiben:

$$b_n^{(n)} \cdot x_a^{(n)}(t) + b_{n-1}^{(n)} \cdot \dot{x}_a^{(n-1)}(t) + \dots + b_0^{(n)} \cdot x_a(t) = a_0^{(m)} \cdot x_e(t) + a_1^{(m)} \cdot \dot{x}_e(t) + \dots + a_m^{(m)} \cdot x_e^{(m)}(t)$$

Wendet man die so genannte **Laplace-Transformation** auf eine solche lineare DGL mit konstanten Koeffizienten an, so erhält man unter der Annahme verschwindender Anfangsbedingungen eine gebrochene rationale Funktion der komplexen Variablen $s = \delta + i\omega$, die sog. **Übertragungsfunktion**.

$$G(s) = \frac{X_a(s)}{X_e(s)} = \frac{(a_0 + a_1 \cdot s + \dots + a_m \cdot s^m)}{(b_0 + b_1 \cdot s + \dots + b_n \cdot s^n)}$$

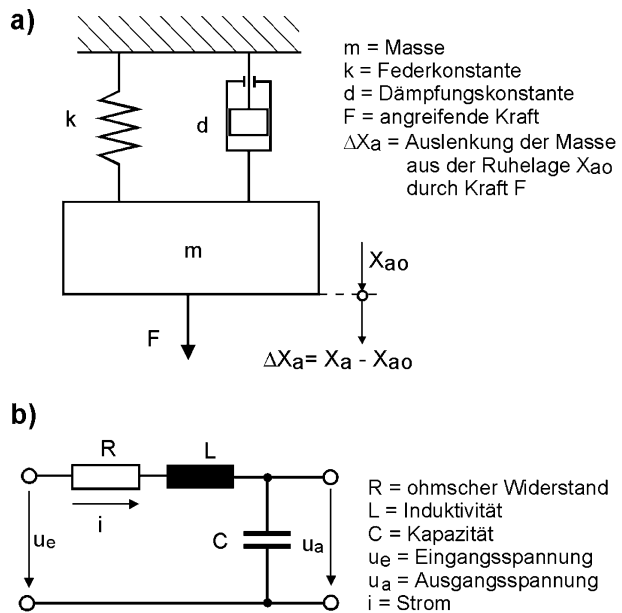


Bild 1-9: Schematische Darstellung a) eines mechanischen Systems und b) eines elektrischen Systems



Bild 1-10: Blockschaltbilddarstellung eines Systems

Genauer hierzu wird im Kapitel 7.3 ausführlich behandelt.

Die Übertragungsfunktion beschreibt das dynamische Übertragungsverhalten des betrachteten Systems zwischen Eingangs- und Ausgangsgröße.

Wendet man diese Methodik auf die oben angeführten Beispielsysteme an, so erhält man für das **mechanische** System unter der Annahme einer linearen Feder und der Annahme, dass alle Massen des Systems in der Masse m konzentriert sind, folgende Differentialgleichung und Übertragungsfunktion:

$$m \cdot \ddot{x}_a(t) + d \cdot \dot{x}_a(t) + k \cdot x_a(t) = F(t) \quad \textbf{Differentialgleichung}$$

$$G(s) = \frac{X_a(s)}{F(s)} = \frac{1}{\frac{m}{k}s^2 + \frac{d}{k}s + 1} \quad \textbf{Übertragungsfunktion}$$

Für das **elektrische** System erhält man unter ebenfalls idealisierenden Annahmen (keine dielektrischen Verluste im Kondensator, keine Streukapazität der Spule und keine Eigeninduktivität des Widerstandes) die folgende Differentialgleichung und Übertragungsfunktion:

$$LC\ddot{u}_a(t) + RC\dot{u}_a(t) + u_a(t) = u_e(t) \quad \textbf{Differentialgleichung}$$

$$G(s) = \frac{U_a(s)}{U_e(s)} = \frac{1}{LCs^2 + RCs + 1} \quad \textbf{Übertragungsfunktion}$$

Vergleicht man nun diese beiden Übertragungsfunktionen, so erkennt man, dass sie sich nur in den Koeffizienten der komplexen Variablen s unterscheiden. Das prinzipielle physikalische Übertragungsverhalten dieser beiden Systeme ist aber identisch. Man spricht bei solchen Systemen von Verzögerungsgliedern zweiter Ordnung. Vergleicht man die Koeffizienten miteinander, so findet man folgende Entsprechungen:

mechanisches System

elektrisches System

Masse m	$\hat{=}$	Induktivität L
Dämpfungskonstante d	$\hat{=}$	Widerstand R
Nachgiebigkeit $N = \frac{1}{k}$	$\hat{=}$	Kapazität C

Werden solche Systeme durch sprungförmige Änderungen der Eingangsgrößen angeregt und dann sich selbst überlassen, so hängt es von den Größen ab, die Einfluss auf die Dämpfung haben, wie das System am Ausgang reagiert. Die Sys-

Der **Piezoeffekt** beruht auf den Eigenschaften der Elementarzellen des Materialgefüges eines solchen Stoffes. Eine Elementarzelle ist die kleinste Systemeinheit des Materials, aus deren Vervielfachung der Aufbau des makroskopischen Kristalls möglich ist. Voraussetzung für das Auftreten des Piezoeffektes ist eine sehr geringe elektrische Leitfähigkeit und das Fehlen eines Symmetriezentrums in der Elementarzelle. Der Vorgang der Ausbildung des Piezoeffektes ist in Bild 1-11 am Beispiel des Quarzes gezeigt. Wird das Material durch äußeren Druck deformiert, so deformieren sich auch die Elementarzellen, wodurch die Schwerpunkte der positiven und negativen Ladungen verschoben werden. Dadurch bilden die Elementarzellen elektrische Dipole aus, wobei aus energetischen Gründen sich alle Dipole benachbarter Elementarzellen in gleicher Richtung orientieren und sog. Domänen bilden. Auf den äußeren Elektroden sammeln sich Ladungen an, so dass man zwischen ihnen eine Spannung messen kann.

Der **reziproke piezoelektrische Effekt** tritt auf, wenn man an die Elektroden eines solchen Elementes eine elektrische Spannung anlegt. Im elektrischen Feld verformen sich die Elementarzellen, so dass beispielsweise bei einer Scheibe dieses Stoffes eine Dickenänderung auftritt.

Piezoelektrische Stoffe werden in der Technik vielfältig eingesetzt, wobei sowohl der normale als auch der reziproke Effekt ausgenutzt werden. Ein solches Piezoelement kann man beispielsweise als Kraftmesssensor benutzen, da durch den piezoelektrischen Effekt an einem solchen Element durch Druck oder Zug elektrische Spannungen erzeugt werden, die der Größe der Kraft proportional sind. Man kann durch Nutzung des reziproken Effektes auch einen Aktor herstellen, den man für kurzhubige, genaue Stellbewegungen nutzen kann. Darüber hinaus kann man auch beide Effekte gleichzeitig in einem technischen Bauelement nutzen. Typisch hierfür ist die Anwendung in einem Ultraschallprüfkopf für die zerstörungsfreie Werkstoffprüfung mit Ultraschall. Weitere Verwendungsmöglichkeiten werden in den Kapiteln 5 (Sensoren) und 6 (Aktoren) behandelt werden.

Bei dieser Prüfung wird Ultraschall im Frequenzbereich zwischen 0,5 und 50 MHz in ein zu prüfendes Werkstück eingeleitet. Dieser Ultraschall wird an der Rückwand des Bauteils, aber auch an Fehlstellen im Material wie Lunkern, Einschlüssen oder Rissen reflektiert. Der reflektierte Ultraschall kann wieder aufgefangen und aufgrund der Laufzeit der aufgezeichneten Echos des Sendesignals kann auf die Größe und Tiefenlage von Fehlstellen geschlossen werden. Für das Senden und Empfangen des Ultraschalls benutzt man ein piezoelektrisches Element, das in einem Prüfkopf eingebaut ist und an die Oberfläche des zu prüfenden Bauteils angekoppelt wird. Diese Elemente sind meist scheibenförmig und die von ihnen abgestrahlte Ultraschallfrequenz hängt im wesentlichen von der Dicke der Scheibe ab, aber auch vom verwendeten Material.

Das Piezoelement wird in der Regel durch einen Hochspannungsimpuls zu Schwingungen in seiner Resonanzfrequenz angeregt. In dem durch den Impuls erzeugten elektrischen Feld zieht sich das Piezoelement zusammen. Nach Abklingen des Impulses bricht das Feld zusammen und das Piezoelement führt dann eine mehr oder weniger schnell abklingende Schwingung aus, die wegen der mechanischen Eigenschaften des Elementes im Ultraschallbereich liegt. Der abgestrahlte Ultraschall hat aufgrund des Erzeugungsvorgangs einen zeitlich begrenzten, impulsartigen Verlauf. Das im Prüfling reflektierte Signal kehrt zum Sendeelement zurück und wird dort durch den reziproken Effekt wieder in eine Spannung umgewandelt, die im Prüfgerät ausgewertet wird.

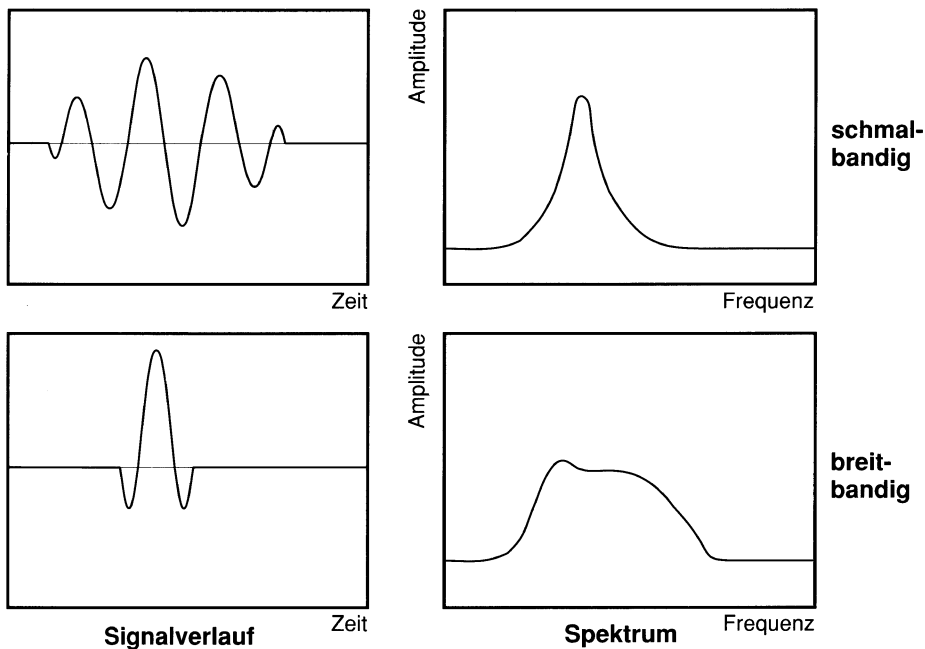


Bild 1-12: Signalverläufe und Spektren von schmal- und breitbandigen Ultraschallsignalen

Je nach Prüftechnik kann es wünschenswert sein, einen Impuls geringer oder hoher Bandbreite zu benutzen. Diese Eigenschaften des Spektrums des Signals sind mit entsprechenden Eigenschaften bezüglich Signalform und Frequenzgehalt des Prüfsignals verbunden (Bild 1-12, s. auch Kap. 4.3). Ein schmalbandiger Impuls enthält eine klar ausgeprägte Prüffrequenz, was die Beurteilung der Fehlergröße erleichtert, besitzt aber wegen der höheren Impulsdauer eine geringere Fehlerauflösung in Ausbreitungsrichtung des Signals. Breitbandige Impulse erhöhen demgegenüber das Auflösungsvermögen, vermindern aber die Möglichkeiten, die Fehlergröße abzuschätzen.

Die Signalform und damit die Bandbreite des Ultraschallsignals kann man nun sowohl elektrisch als auch mechanisch beeinflussen. Bild 1-13 zeigt einen Schnitt durch einen Ultraschallprüfkopf. Man erkennt außer dem am Ende des Gehäuses angebrachten piezoelektrischen Schwinger auf seiner der Schallabstrahlrichtung abgewandten Seite eine elektrische Spule und einen sog. Dämpfungskörper, der eine zusätzliche mechanische Dämpfung des Schwingers und eine Schallabsorption hervorruft. Dadurch wird zum einen verhindert, dass Echos von der Rück-

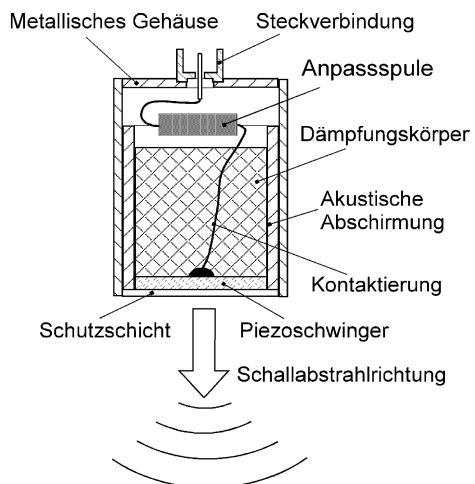


Bild 1-13: Querschnitt durch einen Ultraschallprüfkopf

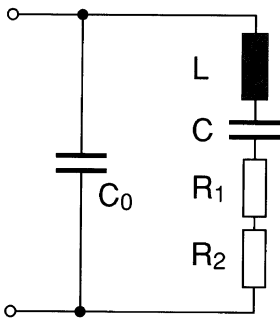


Bild 1-14: Ersatzschaltbild für einen piezoelektrischen Schwinger

seite des Dämpfungskörpers wieder zum Schwinger gelangen, zum anderen kann auch die Signalform des schwingenden Piezoelementes z. T. stark beeinflusst werden.

Will man das Schwingverhalten verstehen, so ist es wieder hilfreich, sich mit der oben festgestellten elektrischen Analogie für ein schwingungsfähiges mechanisches System vor Augen, ein elektrisches Ersatzschaltbild anzufertigen (Bild 1-14). Darin ist C_0 die elektrische Kapazität des kondensatorartigen Piezoelementes und der Serienschwingkreis aus L , C , R_1 und R_2 , wie wir ihn schon aus Bild 1-9 b kennen, steht für das mechanische System aus Bild 1-9 a, das den schwingenden Piezokristall repräsentiert. Darin steht L für die Masse m , C für die Nachgiebigkeit N , R_1 steht für die abgestrahlte Schallenergie und R_2 für die mechanischen Verluste im Piezoelement.

Wird mit einer parallel zum Schwinger angebrachten Induktivität ein zweiter elektrischer Parallelschwingkreis erzeugt, so kann man die Frequenz und die Bandbreite des abgestrahlten Ultraschallsignals in bestimmten Bereichen verändern und optimieren. Die mechanische Güte des schwingenden Systems die mit

$$Q_m = \frac{1}{R} \sqrt{\frac{L}{C}} : R = R_1 + R_2 \quad \text{mechanische Güte}$$

normalerweise recht hoch ist (R_2 ist klein), kann bei bestimmten Piezomaterialien² gezielt recht klein gestaltet werden. Der abgegebene Ultraschallimpuls ist dann wegen der hohen inneren Dämpfung breitbandig und hat keine ausgeprägte Eigenfrequenz (Bild 1-15), wie beim Material mit hoher mechanischer Güte. Erhöht man die mechanische Gesamtdämpfung durch einen entsprechenden Dämpfungskörper noch weiter, so tritt praktisch keine Schwingung mehr auf, sondern der aperiodische Grenzfall; der Impuls tritt im Wesentlichen nur noch in einer Polarität von der Nulllage auf (Bild 1-16). Hiermit hat man einen Prüfkopf mit maximaler Fehlerrückmeldung in Signalausbreitungsrichtung. Um dieses Ergebnis zu erzielen, kommt es

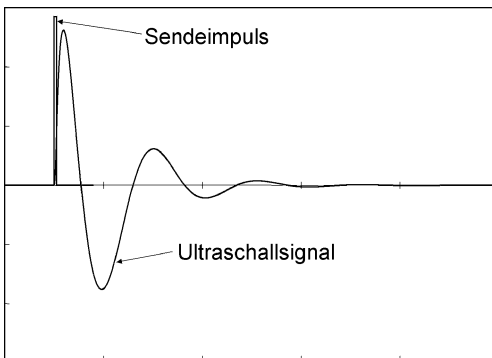


Bild 1-15: Ultraschallschwingung eines Piezoschwingers geringer Güte

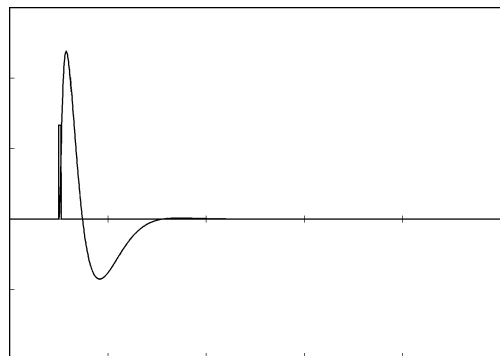


Bild 1-16: Breitbandige Ultraschallschwingung eines Piezoschwingers mit starker mechanischer Dämpfung

² Bleimetaniobat (PbNb_2O_6) hat geringe mechanische Güte, d. h. R_2 ist groß.

sehr stark auf die mechanischen Eigenschaften des Dämpfungskörpers an.

Da in einem solchen Schwinger mit extrem hoher Dämpfung der Widerstand im Serienschwingkreis sehr groß ist, treten dessen frequenzbestimmende Komponenten, nämlich die Spule mit Induktivität L (Masse des Schwingers) und der Kondensator mit der Kapazität C (Nachgiebigkeit des Schwingers), gar nicht mehr in Erscheinung. Man kann daher nun bei bekannter Kapazität C_0 des Piezoelementes durch Parallel- oder Reihenschaltung einer externen Induktivität einen solchen Prüfkopf schmalbandig in einem weiten Frequenzband auf eine feste Prüffrequenz abstimmen (Bild 1-17). Die Fehlerauflösung ist dann, wie bereits gesagt, nicht mehr optimal, aber die Fehlergrößenbewertung ist verbessert.

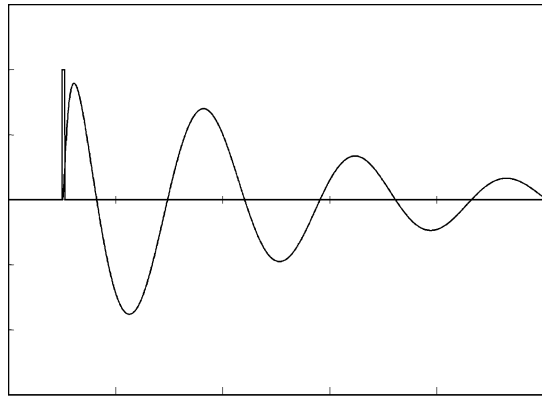


Bild 1-17: Ultraschallschwingung eines schmalbandig abgestimmten breitbandigen Piezoschwingers

Man sieht an diesem Beispiel, dass man die akustischen Eigenschaften eines solchen Systems, sowohl elektrisch als auch mechanisch, immer nach den gleichen Regeln beeinflussen und beide Eigenschaften formal gleich behandeln kann. Jemand der sowohl das mechanische als auch das elektrische Verhalten solcher Systeme gleichermaßen gut kennt, ist in der Lage, es für die verschiedensten Anwendungsfälle zu optimieren. Ein Denken in systemischen Denkweisen und nicht in entweder mechanischen oder elektrotechnischen Kategorien, verschafft tiefere Einblicke in das Systemverhalten und führt bei der Anwendung zu besseren Systemen. Das gilt nicht nur für dieses einfache Beispiel eines elektromechanischen Wandlers als einfaches mechatronisches System, sondern auch für komplexere Systeme, wie beispielsweise ein sich aktiv mit dem Fahr- und Federverhalten an die Umgebung anpassendes Fahrwerk eines Landfahrzeugs.

Eine weitere wichtige Disziplin des mechatronischen Arbeitens, nämlich die Informatik wurde zur Darstellung des Schwingverhaltens eines piezoelektrischen Wandlers in den Bildern 1-15 bis 1-17 benutzt. Anstatt entsprechende Wandler körperlich herzustellen und ihr Verhalten messtechnisch zu erfassen, wurden sie mit Hilfe eines rechnergestützten Simulationssystems modelliert und ihr Schwingverhalten getestet. Solche Systeme werden beim Entwurf von dynamischen Systemen in der Mechatronik vielfach benutzt, da man mit ihrer Hilfe sehr schnell zu Aussagen über ein System kommt und durch einfache Änderung der Parameter tiefgreifende Einblicke in das Systemverhalten erhält. Viele Probleme der Dynamik lassen sich überhaupt nur unter Einsatz von Digitalrechnern behandeln und beispielsweise die **Finite-Elemente-Methode** und die **Modalanalyse** lassen sich sinnvoll nur mit Rechnereinsatz durchführen. Auch die Entwurfsmethoden bei der Konstruktion mechanischer Komponenten und die Entwicklung elektrischer Schaltungen wird heute weitestgehend rechnergestützt durchgeführt. Die für solche Entwurfsprogramme üblichen abkürzenden Sammelbezeichnungen wie **CAD** und **CAE** sind jedermann, der sich mit Ingenieurwissenschaften beschäftigt, geläufig. Daher gehört der Digitalrechner zu einem der wichtigsten Handwerkzeuge des Mechatronikers und An-

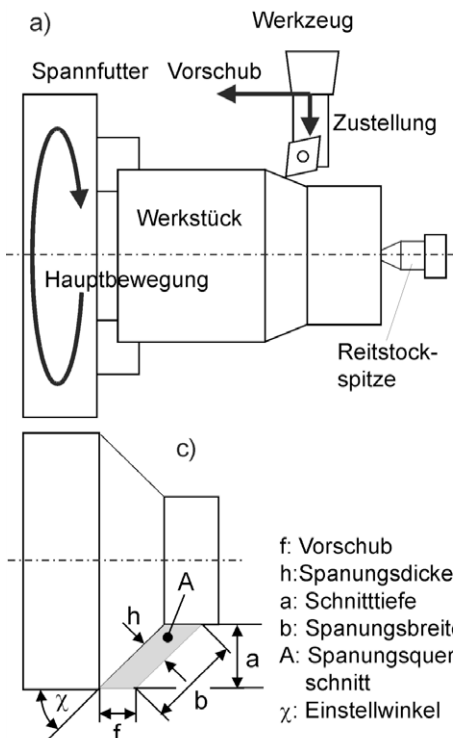
wendungsmöglichkeiten und entsprechende Programme werden in den folgenden Kapiteln entsprechend berücksichtigt.

Dieses Buch will dazu beitragen, das Grundlagenwissen, das bei der Beschreibung und beim Entwurf mechatronischer Systeme erforderlich ist, vorzustellen, ohne es natürlich für den Einzelaspekt zu sehr vertiefen zu können. Entsprechende Literaturangaben über Spezialliteratur zu Einzelthemen finden sich in den zugehörigen Kapiteln.

1.5 Unterschiede zwischen Maschinenbau, Elektrotechnik und Mechatronik

An einem weiteren Beispiel soll nun noch einmal die unterschiedliche Herangehensweise der drei Ingenieurdisziplinen an eine Aufgabenstellung schlaglichtartig verdeutlicht werden, wobei die Überzeichnung dazu dienen soll, die bestehenden Unterschiede klar hervortreten zu lassen.

Jeder der sich mit Zerspanungsprozessen auf Werkzeugmaschinen auskennt, kennt das Phänomen des **regenerativen Ratterns**. Die Bezeichnung wird für einen Schwingungsvorgang zwischen Werkzeug (Drehmeißel) und Werkstück verwendet, der unter ungünstigen Umständen während der Bearbeitung auftreten kann und sich als lautes Geräusch äußert und der sich auch in Form von Rattermarken auf der Oberfläche des Werkstückes abbildet.



a) Einspannsituation eines Drehteils in einer Drehmaschine und Bewegungen

b) Drehmaschine

c) Bei der Zerspanung auftretender Spanungsquerschnitt

Bild 1-18: Drehverfahren

In Bild 1-18 sind die Bearbeitungssituation und die unterschiedlichen Bewegungen beim Drehvorgang dargestellt. Das Werkstück ist im Spannfutter der Hauptspindel eingespannt und wird mit der Hauptbewegung gedreht. Gleichzeitig wird das Werkzeug am Werkstück durch eine Überlagerung von Vorschub- und Zustellbewegung entlanggeführt. Dabei wird ein Span abgetrennt, der einen durch die Vorschub- (Vorschub f) und Zustellbewegung (Schnitttiefe a) festgelegten Spanungsquerschnitt A besitzt. Die direkt am Spanungsquerschnitt messbaren Größen Spanungsdicke h und Spanungsbreite b ergeben sich aus den Maschineneinstellungen f und a durch den Einstellwinkel κ des Werkzeuges.

Wodurch kommt es nun zum Auftreten eines Rattervorgangs beim Zerspanen, der die Qualität des erzeugten Werkstückes negativ beeinflusst?

In Bild 1-19 ist die Hauptschnittkraft F_c dargestellt, die sich aufgrund der Hauptbewegung (Drehung des Werkstücks mit Schnittgeschwindigkeit v_c) ergibt:

$$F_c = k_c \cdot A .$$

Deren Größe ist proportional zum Spanungsquerschnitt A , wobei der Proportionalitätsfaktor k_c spezifische Schnittkraft genannt wird. Diese Größe ist vor allem vom Werkstoff, aber auch von weiteren Größen wie beispielsweise Schnittgeschwindigkeit v_c und Spanungsdicke h abhängig. Für den Spanungsquerschnitt gilt wiederum (Bild 1-18):

$$A = a \cdot f = b \cdot h .$$

Unter der Schnittkraft verformt sich nun das Werkzeug, das man in erster Näherung als Biegebalken betrachten kann, dessen Verformung dem Hookeschen Gesetz gehorcht:

$$F = k \cdot x .$$

Dabei ist k die Steifigkeit oder Federkonstante und x der Betrag, um den sich das Werkzeug in Richtung der Kraft verformt.

Aufgrund der Verformung (Bild 1-19) wird das Werkzeug aus dem Schnitt gedrängt. Das führt zu einer Verringerung der Schnitttiefe a , was wiederum zu einer Verkleinerung des Spanungsquerschnittes A führt. Da die Hauptschnittkraft dem Spanungsquerschnitt proportional ist, sinkt diese ab, wodurch wiederum die Verformung des Werkzeugs abnimmt und die Schnitttiefe erneut ansteigt. Dieser Vorgang wiederholt sich ständig. Entsprechend dem Hooke'

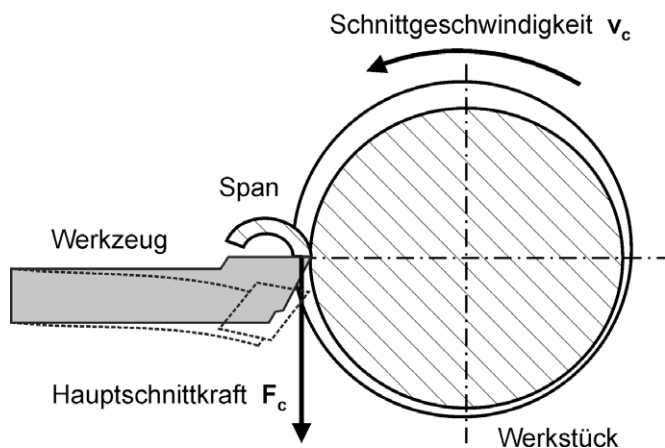


Bild 1-19: Verbiegung des Drehwerkzeugs unter der Hauptschnittkraft

schen Gesetz führt die ständige Veränderung der Kraft zu einer ständigen Veränderung der Verformung; es liegt eine Schwingung vor. Diese Schwingung kann, wie in Bild 1-19 gezeigt, durch eine sich im Laufe einer Werkstückumdrehung ändernde Schnitttiefe angefacht werden und sie kann sich weiter aufschaukeln, da bei nachfolgenden Umdrehungen die Schnitttiefenänderungen durch die davor liegenden Schwankungen phasenrichtig mit den aktuellen Schwankungen zusammentreffen können. Es liegt dann regeneratives Rattern vor.

Ist die Werkzeugmaschine manuell bedient, so kann der Maschinenbediener durch Variation der Einstelldaten (Schnittgeschwindigkeit, Vorschub, Schnitttiefe) versuchen die Ratterschwingung zu vermindern. Für automatisch arbeitende Maschinen muss man schon in der Konstruktions- und Entwicklungsphase Vorkehrungen treffen, um das Auftreten von Ratterschwingungen zu vermeiden oder zu beseitigen. Wie würden nun Konstrukteure der unterschiedlichen Ingenieurdisziplinen Maschinenbau, Elektrotechnik und Mechatronik vorgehen?

Eine der Ursachen für das Auftreten von Ratterschwingungen ist eine zu geringe Steifigkeit von Werkzeug und Werkzeughalterung. Dies entspricht im Hookeschen Gesetz einer zu kleinen Federkonstante k . Diese wiederum ist vom Werkstoff und von den Materialquerschnitten abhängig, sodass ein Maschinenbauingenieur an dieser Stelle Verbesserungen vornehmen würde und z. B. einen Drehmeißel mit größerem Schaftquerschnitt aus festerem Werkstoff wählen würde, der zusätzlich noch günstiger im Halter abgestützt wird.

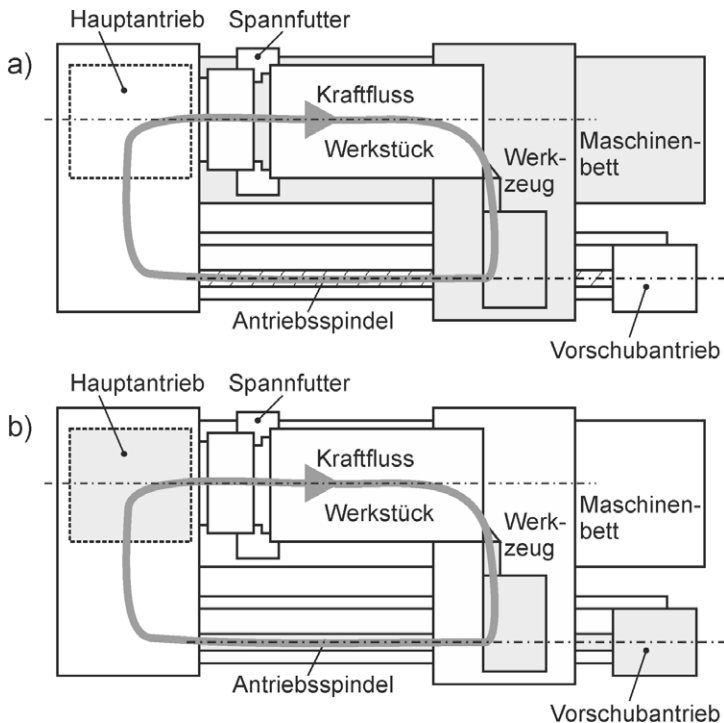


Bild 1-20: Schematische Draufsicht einer Drehmaschine mit Baugruppen im Kraftfluss a) grau unterlegt maschinenbauliche Maßnahmen b) grau unterlegt elektrotechnische Maßnahmen

In Bild 1-20 ist eine schematische Draufsicht auf eine Drehmaschine, mit den wichtigsten im Kraftfluss der Maschine liegenden Baugruppen dargestellt. Teilbild a) zeigt die Baugruppen grau unterlegt, die der Maschinenbauingenieur beeinflussen würde. Ob eine Schwingung angefacht wird, hängt auch von der Größe der Dämpfung in der Maschine, vor allem in den im Kraftfluss liegenden Bauteilen ab. Hier haben verschiedene Werkstoffe verschiedene Dämpfungseigenschaften. So hat beispielsweise Grauguss, aus dem häufig Gestellbauteile von Werkzeugmaschinen gefertigt werden, eine höhere innere Dämpfung als Stahl, spezieller im Werkzeugmaschinenbau eingesetzter Polymerbeton hat eine noch wesentlich höhere Dämpfung als Grauguss. Dies beruht auf der inhomogenen inneren Struktur des Werkstoffs. So kann der Maschinenbauingenieur die Eigenschaften der Maschine, hier speziell die Neigung zum Rattern, durch die Werkstoffwahl positiv beeinflussen. Schaut man sich im Bild 1-20 a die Baugruppen an, die grau unterlegt sind, so sieht man, dass es Baugruppen im Kraftfluss gibt, die nicht durch maschinenbauliche Maßnahmen beeinflusst werden.

Die Antriebe von Werkzeugmaschinen sind heute grundsätzlich elektrische Antriebe, die zusätzlich meist noch elektronisch drehzahlregelbar sind. Ein Elektroingenieur würde deshalb hier ansetzen um mögliche Ratterschwingungen zu bekämpfen.

So könnte man mit Hilfe von Sensoren, beispielsweise Beschleunigungssensoren, eventuell auftretende Schwingungen erfassen und mit gezielten Strategien über die elektronische Maschinensteuerung den Hauptantrieb und damit die Schnittgeschwindigkeit oder den Vorschubantrieb und damit die Vorschubgeschwindigkeit beeinflussen, um eine Ratterschwingung zu unterdrücken. Dies ist die Vorgehensweise, die ein Maschinenbediener im manuellen Betrieb auch anwenden würde. Auch hier sieht man in Bild 1-20 b, dass nur ein Teil der im Kraftfluss liegenden Baugruppen von diesen Maßnahmen betroffen ist.

Kombiniert man nun beides, so sind alle Bereiche der Maschine, die am Entstehen einer Ratterschwingung beteiligt sind, einbezogen. Ist dies dann schon die mechatronische Lösung?

Der Mechatronikingenieur untersucht diese Möglichkeiten und stellt fest, dass nun die Wahrscheinlichkeit des Auftretens einer Ratterschwingung minimiert worden ist, aber die eigentliche Ursache, nämlich eine **Modulation** der Kraft gar nicht direkt beeinflusst wird. Die Ursache der Schwingung (veränderliche Kraft $\tilde{F}$) bewirkt eine entsprechend veränderliche Verformung $\tilde{x}$:

$$\tilde{F} = k \cdot \tilde{x}$$

Will man nun erreichen, dass F überhaupt nicht mehr schwankt, so muss man die Steifigkeit k der im Kraftfluss liegenden Baugruppen genau mit einer Schwingung gleicher Frequenz modulieren, die gegenüber den Schwankungen der Kraft um 180° phasenverschoben ist. Ähnlich wie die Überlagerung von Lichtwellen an bestimmten Punkten durch Interferenz zur Auslöschung und damit Dunkelheit führen kann, sollten mechanische Schwingungen bei entsprechender Überlagerung ausgelöscht werden können. Bild 1-21 zeigt eine solche Auslöschung durch Interferenz. Der harmonischen Schwingung 1 mit bestimmter Frequenz f und einer Amplitude A wird eine Schwingung 2 überlagert d.h. dazu addiert, die gleiche Frequenz, am Anfang und am Ende abweichende Amplitude und eine Phasenverschiebung gegenüber Schwingung 1 von 180° besitzt. Wie man sieht, werden durch Interferenz dort, wo die beiden Schwingungen gleichen Absolutwert der Amplitude besitzen, diese sich gegenseitig komplett auslöschen. Dies stellt der durchgezogene Kurvenverlauf in Bild 1-21 dar. Der Effekt der Auslöschung wird

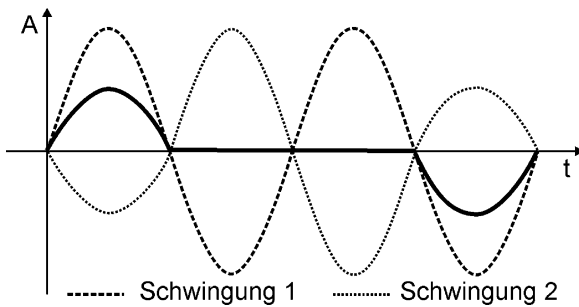


Bild 1-21: Interferenz zweier Schwingungen mit Auslöschung im mittleren Bereich

auch als **destruktive Interferenz** bezeichnet.

Wie kann man dieses Prinzip nun zur Schwingungsauslöschung bei dem behandelten Beispiel des regenerativen Rattern anwenden? Hier können auch wiederum piezoelektrische Elemente von Nutzen sein, die man in den Kraftfluss zwischen Werkzeug und Werkstück einbaut. Solche Piezoelemente können einerseits unter Ausnutzung des piezoe-

lektrischen Effektes als Sensor zur Registrierung solcher Schwingungen ausgenutzt werden, da sie ein kraftproportionales Spannungssignal liefern. Andererseits kann man durch Ausnutzung des reziproken piezoelektrischen Effektes Elemente in den Kraftfluss einbringen die zu einer Steifigkeitsmodulation im Kraftfluss führen, die gegenüber der registrierten Schwingung um 180° phasenverschoben verläuft. Die Beaufschlagung eines solchen Aktor-Elementes mit einer Wechselspannung führt zu einer Dickenänderung des Elementes, die bei einem dünnen scheibenförmigen Element im Bereich weniger Mikrometer liegt. Legt man es zwischen Werkzeug und Werkzeughalter und legt eine Wechselspannung an, so ändert sich die Gesamtsteifigkeit der Anordnung in Kraftrichtung. Die Steifigkeitsmodulation muss natürlich von einem Rechner aufgrund der Sensorsignale exakt gesteuert werden. Als Ergebnis ist eine solche Einrichtung in der Lage, die Ratterneigung komplett zu unterdrücken, während die anderen Lösungen nur Teilaspekte in Betracht zogen, ohne die eigentliche Ursache zu behandeln. In der Gleichung des Hooke'schen Gesetzes drückt sich dies so aus:

$$\tilde{F} = k \cdot x$$

Dies deutet an, dass durch gegenphasige Modulation des Steifigkeits- zu dem des Kraftverlaufs die Verformung konstant gehalten werden kann, d. h. die Schwingung verschwindet.

Durch den mechatronischen Denkansatz, konnte also eine viel generelle Lösung des Ratterproblems gefunden werden, die nicht nur einzelne Symptome behandelt und unter Umständen erheblich wirtschaftlicher arbeitet.

Eine Anwendung der Schallauslöschung durch destruktive Interferenz gibt es inzwischen zu kaufen und zwar in speziellen Kopfhörern [1.4]. Hört man Musik mit einem Kopfhörer bei gleichzeitig sehr lauten Außengeräuschen wie im Auto oder Flugzeug, so stört der Lärm den Musikgenuss erheblich. Dies kann man durch passive Dämpfung in der Kopfhörerabschirmung zu verbessern suchen, erreicht so aber nur eine Reduzierung des Lärnhintergrunds von 15 -25 dB.

Kunststoffschäume dämpfen Frequenzen ab 200 Hz, doch je langwelliger die Schwingung, desto dicker muss das Dämpfungsmaterial sein. Um den Tragekomfort des Kopfhörers dadurch nicht zu stark zu verschlechtern, ist eine aktive Geräuschreduktion wie in Bild 1-22 dargestellt von Vorteil. Die Muschel des Kopfhörers und die das Ohr umschließende Polsterung schwächen den hoch frequenteren Schallanteil des Geräusches. Geräusche tieferer Frequenz, die durch die Schalliso-

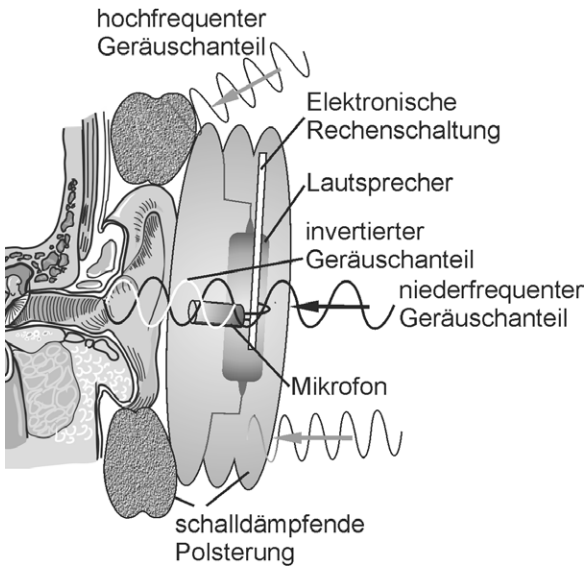


Bild 1-22: Kopfhörer mit Einbauten zur Geräuschminderung durch destruktive Interferenz

lierung dringen, werden von einem Mikrofon aufgenommen und in einer Rechenschaltung invertiert. Der Lautsprecher gibt das invertierte Geräusch dann mit an das Innenohr ab. Für die Invertierung wird das Nutzsignal (Sprache, Musik) aus der Mikrofonaufnahme herausgerechnet. Das invertierte und das Originalgeräusch heben sich dann im Ohr auf. Die Ergebnisse sind im Bild 1-23 dargestellt. Bis auf ein geringes Restgeräusch ist der Störhintergrund beim Musikhören verschwunden.

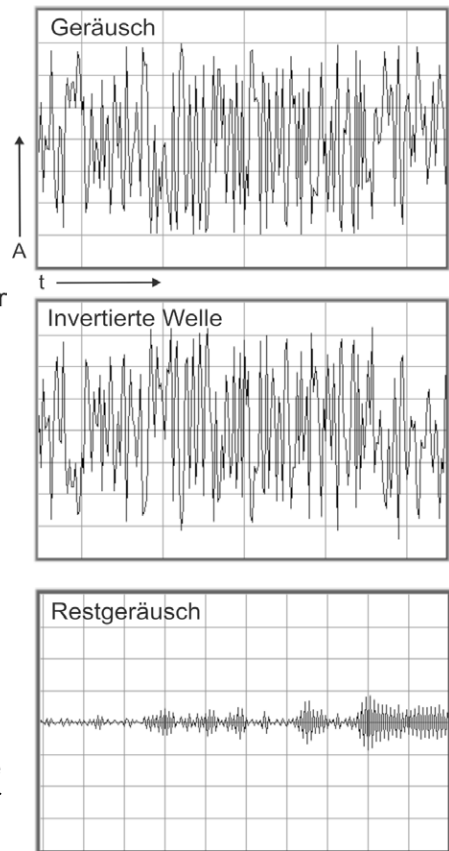


Bild 1-23: Destruktive Interferenz

oben: Geräusch
Mitte: invertiertes Geräusch
unten: Überlagerung

1.6 Teilgebiete der Mechatronik

Um einen Überblick darüber zu bekommen, welche Teilgebiete der Ingenieurwissenschaften in der Mechatronik von besonderer Bedeutung sind, kann man sich das in Bild 1-24 dargestellte Schema eines mechatronischen Systems zur Hilfe nehmen.

Den Kern eines solchen Systems bildet in der Regel ein **Digitalrechner**, meist ein Mikrorechner. Dieser bekommt Informationen über physikalische Größen aus der Außenwelt und über den Bewegungszustand oder die Lage des Stellsystems. Zwischen **Sensoren** und Digitalrechner befindet sich häufig noch eine Messwertverarbeitung zur Pegelanpassung und Vorverarbeitung der Messwerte, abgestimmt auf die Bedürfnisse des Digitalrechners. Auf diesem werden Algorithmen zur Steuerung und Regelung der Stellgrößen unter Berücksichtigung der Sensorsignale und Rückmeldungen durchgerechnet. Die Algorithmen ermitteln Stellgrößen, die noch an die Pegel und Signalformen (elektrisch, hydraulisch, pneumatisch) der **Aktoren** durch

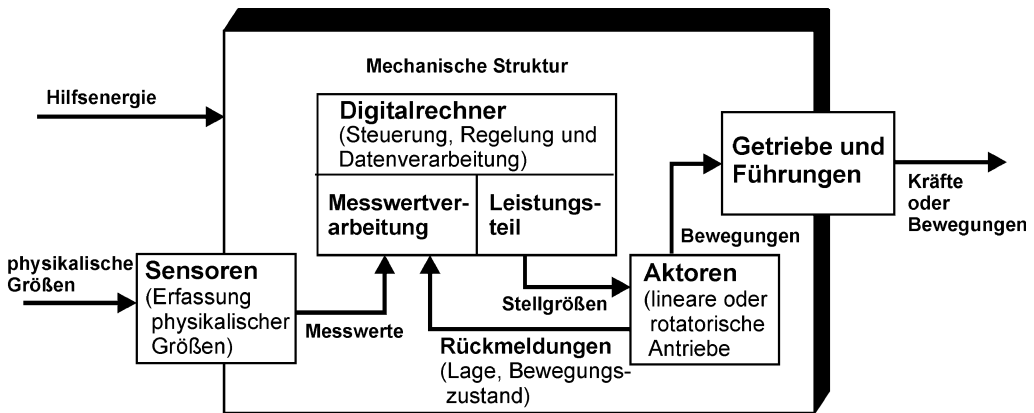


Bild 1-24: Grundstruktur eines mechatronischen Systems

das Leistungsteil angepasst werden. Die Ausgangsgrößen der Aktoren werden schließlich durch Getriebe (Zahnräder, Spindeln, Gelenke, Hebel, Ketten, Riemen) und Führungen in Bewegungen oder Kräfte umgesetzt. Alle Teilsysteme sind in der Regel durch eine mechanische Struktur (Gehäuse, Chassis, Träger) miteinander verbunden.

Aus diesem Überblick über ein **mechatronisches System** ergibt sich dann der Umfang und der Inhalt der folgenden Kapitel. Will man die Bewegungen eines mechatronischen Systems, das häufig aus einer Vielzahl mechanischer und elektrischer Bauteile besteht, beschreiben, so können die Methoden der **Kinematik** eingesetzt werden. Hierbei handelt es sich um die Lehre vom geometrischen und zeitlichen Ablauf von Bewegungen, ohne nach Ursachen (z. B. den Kräften) zu fragen [1.5].

Begriffe und Methoden der Kinematik benötigt man, um die verwickelten Bewegungen komplexer mechanischer Systeme wie beispielsweise von Industrierobotern beschreiben zu können. Die **Kinetik** untersucht die Wechselwirkungen zwischen Kräften und den Bewegungen von Massen. Die weiterhin gebräuchliche Bezeichnung **Dynamik** wiederum schließt die Kinetik und die **Statik**, also die Lehre vom Gleichgewicht ruhender Körper³, mit ein. In den letzten 30 Jahren hat sich auch der Begriff der **Strukturdynamik** eingebürgert, die sich mit der Dynamik und den Schwingungsvorgängen von komplizierten technischen Strukturen wie Fahrzeugen, Flugzeugen, Bauwerken und Maschinen auseinandersetzt [2.2]. Da mechatronische Systeme meist komplexe bewegte Strukturen sind, liegt ein Hauptaugenmerk auf der Strukturdynamik. Da zu jedem mechatronischen System eine elektronische Steuerung (engl. Controller) gehört, liegt ein weiterer Schwerpunkt bei der Mess-, Steuerungs- und Regelungstechnik. Erfassung und Erzeugung der Dynamik des Systems (Sensorik und Aktorik) sind weitere wichtige Bestandteile der Ausführungen in den Kapiteln 5 und 6. Außer Grundlagenwissen zu diesen Themen finden sich in Kapitel 9 beispielhafte Beschreibungen verschiedener ausgeführter mechatronischer Systeme.

Ein typisches Beispiel für ein solches System ist der heute in den meisten PKW's vorhandene Airbag (Bild 1-25). Ein Verzögerungssensor misst die jeweilige

³ Man kann den Zustand, den die Statik behandelt, als Sonderfall der Kinetik betrachten, d. h. eine Bewegung mit der Geschwindigkeit Null.

Auslösen des Airbags führen, sondern nur in bestimmten Fahrzuständen. Dies bewertet ein eingebauter Mikrorechner in Abhängigkeit weiterer Fahrzeuginformationen. Nur wenn alle Umstände auf einen Aufprall des Fahrzeugs auf ein Hindernis schließen lassen, wird der Gasgenerator (Aktor) ausgelöst, der den Airbag aufbläst.

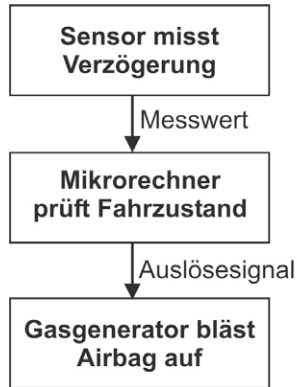


Bild 1-25: Auslösen eines Airbags in einem PKW

Eine andere nicht direkt an die strukturelle Darstellung mit körperlichen Baugruppen (Sensor, Aktor, Mikrorechner) angelehnte Form ist die Darstellung anhand der Funktionalitäten eines mechatronischen Systems [1.6]. Diese Darstellungsart findet sich in Bild 1-26.

Dies ist gleichzeitig die allgemeine Darstellung einer intelligenten Maschine, die Informationen aus der Umgebung aufnimmt (**Wahrnehmen**), um darauf entweder umgehend zu reagieren (**reaktives Verhalten**), oder aufgrund eines intelligenten Erkennungsapparates (**Erkennen**) sinnvoll und seiner Aufgabe entsprechend zu handeln (**zielorientiertes Verhalten**).

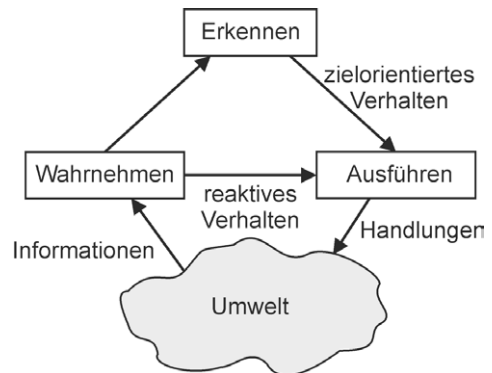


Bild 1-26: Modellstruktur eines mechatronischen Systems aufgrund von Funktionalitäten

Diese Art der Betrachtung eines mechatronischen Systems aufgrund seiner Funktionalitäten ist typisch für die Mechatronik. Um eine optimale Lösung für eine intelligente Maschine zu finden, betrachtet man in der Entwicklungsphase nicht physikalische Baugruppen, sondern Systemeigenschaften. So kann beispielsweise die Funktionalität "Wahrnehmen" körperlich in mehreren Bauteilen realisiert sein, etwa bei einem Bildverarbeitungssystem. Dort findet die Wahrnehmung eines bestimmten Gegenstandes in der Außenwelt durch eine Kamera (Sensor) in Verbindung mit einem Digitalrechner statt, der die Bilderkennungssoftware enthält. Die Funktion ist also nicht nur in einer körperlichen Baueinheit konzentriert.

Das Denken in Systemen und die Ermittlung von Modellen sowie die Entwicklung von Algorithmen, mit denen solche Systeme quasi intelligentes Verhalten (zielorientiertes Verhalten) entwickeln, sind demnach Kernaufgaben der Mechatronik. An dieser Art der Modellstruktur kann man auch eine Abgrenzung zu allgemeinen automatisierten Systemen erkennen. Wird beispielsweise die Wahrnehmungsfunktion von einer **Lichtschranke** realisiert, die durchlaufende Objekte registriert und dadurch eine Klappe steuert, um solche Objekte auf einen anderen Förderweg zu schicken, so ist das rein reaktives Verhalten (Bild 1-26). Kann aber die Licht-

schranke beispielsweise verschiedenfarbige Objekte unterscheiden und daraus entsprechend einer Strategie verschiedene Objektweichen betätigen, so handelt es sich um zielorientiertes Verhalten. Einem solchen System, das mit Hilfe eines Wahrnehmungsprozesses mit anschließendem Erkennen verschiedener Situationen Handlungen aufgrund zielorientierten Verhaltens durchführt, würde man als mechatronisches System bezeichnen. Das vorher erwähnte System, das rein reaktives Verhalten zeigt, würde man nicht zu den mechatronischen Systemen zählen.

2 Modellbildung technischer Systeme

2.1 Systembegriff

In Kapitel 1 wurde bereits kurz auf die Modellbildung technischer Systeme am Beispiel eines einfachen mechanischen und eines elektrischen Systems eingegangen. Um die Kinematik und die Dynamik eines komplexen Systems anschließend behandeln zu können und darauf aufbauend ein Steuerungs- und Regelungskonzept des Systems zu entwickeln, ist immer zuerst eine solche Modellbildung erforderlich, d. h. letztendlich die Bildung eines Satzes mathematischer Beschreibungen des Systemverhaltens (Bild 2-1). Im ersten Schritt dieser Vorgehensweise beschreibt man das betrachtete Problem durch sprachliche Formulierungen. Ein Beispiel wäre:

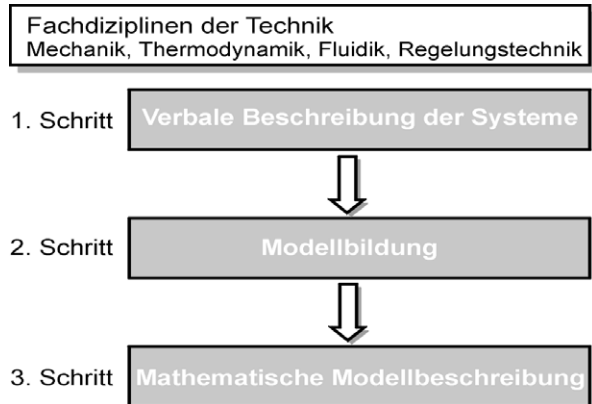


Bild 2-1: Vorgehensweise bei der Beschreibung physikalisch technischer Systeme

- Auf Gegenstände auf der Erde wirkt eine Kraft, so dass sie zu Boden fallen.

Die Beschreibung ist also rein qualitativ. Im nächsten Schritt sucht man nach Zusammenhängen zwischen den Objekten der Betrachtung, die eine quantitative Aussage beinhalten:

- Die wirkende Kraft ist zur Masse des Gegenstandes proportional.

$$F = m \cdot g$$

Im dritten Schritt formuliert man die gefundene Modellvorstellung durch mathematische Ausdrücke, die eine exakte Modelldefinition ermöglichen:

Dies wäre ein sehr einfaches mathematisches Modell, für die meisten technischen Systeme ergibt sich in der Regel ein Satz von Differentialgleichungen. Eine solche mathematische Beschreibung durch **Differentialgleichungen** und **Anfangsbedingungen** ist zwar exakt und lässt genaue Aussagen für das Modell zu, aber die Gleichungen gelten eben nicht für das reale Objekt der Betrachtung, sondern für sein Modell. Dies bedeutet, dass das Modell häufig nicht exakt das reale Verhalten beschreibt und meist auch gar nicht soll.

Um die Vorgehensweise bei der Modellbildung **von technischen Systemen** zu behandeln, muss man zuerst den Begriff des **Systems** genauer fassen. Der Begriff der Systemtechnik kam ebenfalls Ende der 1960er Jahre im anglo-amerikanischen

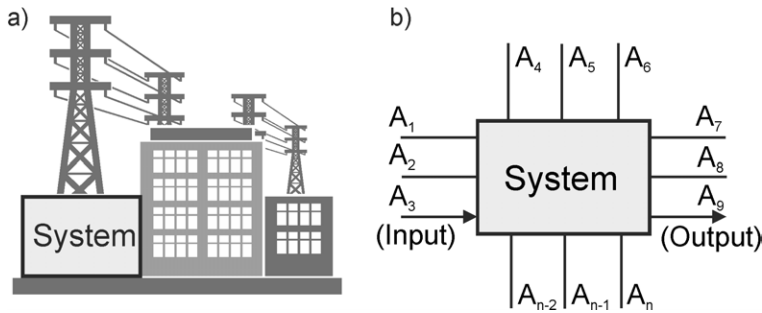


Bild 2-2: Systembegriff: a) Ein System ist ein abgegrenzter Teilbereich der Umgebung
b) System mit Systemgrenze und Attributen

Sprachbereich auf und wurde erstmals von G. Ropohl [2.1] im deutschen Sprachraum ausführlich behandelt.

Ein System ist ein von seiner Umgebung in irgend einer Weise abgegrenzter Gegenstand. Die Abgrenzung eines Systems ergibt sich jedoch nicht aus seinen physikalischen Grenzen, sondern aus der Fragestellung der Systembetrachtung. Ein wichtiger Bestandteil dieser Betrachtungsweise ist die Umgebung, wobei damit nicht automatisch die gesamte übrige Welt, sondern die für die Fragestellung der Systembetrachtung wichtigen Objekte außerhalb des Systems gemeint sind (Bild 2-2 a)).

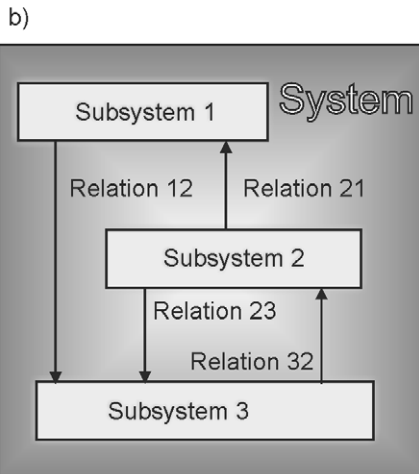
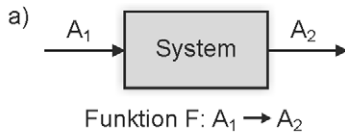


Bild 2-3: a) System mit Funktion F zwischen den Attributen
b) System mit Subsystemen und Relationen

Jedes System zeigt gegenüber der Umgebung gewisse Kennzeichen, Merkmale, Eigenschaften, die **Attribute** genannt werden. Attribute, die weder Eingangsgrößen (Input) noch Ausgangsgrößen (Output) sind, sondern die Verfassung des Systems beschreiben, werden **Zustände** genannt (Bild 2-2 b)). Zwischen den Attributen eines Systems bestehen Beziehungen in Form von Funktionen. Ist die **Funktion F** unbekannt, so bezeichnet man das System auch als **Black Box** (Bild 2-3 a)).

Wie Bild 2-3 b) zeigt, enthält ein System normalerweise **Subsysteme**, deren Beziehungen untereinander durch **Relationen R_{ij}** beschrieben werden. Die Menge der Relationen heißt **Struktur** des Systems. Es gibt also zweierlei Aspekte bei der Beschreibung von Systemen, nämlich einerseits die funktionalen Zusammenhänge zwischen den Attributen eines Systems und die strukturellen Zusammenhänge zwischen den Subsystemen eines größeren Gesamtsystems. Die Feststellung, ein System sei mehr als die Summe der Eigenschaften seiner Teile beruht gerade darauf, dass die

Relationen zwischen den Teilen dem Gesamtsystem eine zusätzliche Qualität ver-

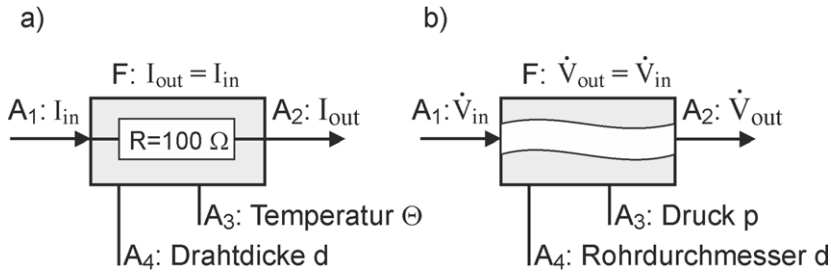


Bild 2-4: Beispiele für einfache Systeme a) elektrisches System b) hydraulisches System

leihen, die nicht aus den Eigenschaften der Teilsysteme gefolgert werden kann. Die Modellbildung eines Systems muss sich daher einerseits mit der Ermittlung der Funktionen zwischen den Attributen der Teilsysteme und andererseits mit der Ermittlung der Systemstruktur (Summe der Relationen) eines Gesamtsystems befassen.

Bild 2-4 zeigt zwei sehr einfache Beispiele für Systeme und zwar im Teilbild a) das elektrische System "ohmscher Widerstand" und in Teilbild b) das hydraulische System "Rohr". Die Funktionen F , die das Verhältnis zwischen den Eingängen und den Ausgängen im **stationären Zustand**¹ beschreibt, ist hier besonders einfach und entspricht der **Kontinuitätsgleichung**. Sie besagt, dass der ins System hineinfließende Strom (Volumenstrom) gleich dem aus dem System herausfließenden Strom ist. Weitere angedeutete Attribute dieser Systeme sind geometrische Daten oder Zustandsdaten wie Temperatur oder Druck. Dabei ist die Anzahl der gezeigten Attribute nur eine kleine Auswahl der tatsächlich vorhandenen Attribute. Welche Attribute man in die Betrachtung anlässlich einer Modellbildung mit einbezieht, hängt wiederum von der im Zusammenhang mit dem System behandelten Fragestellung ab.

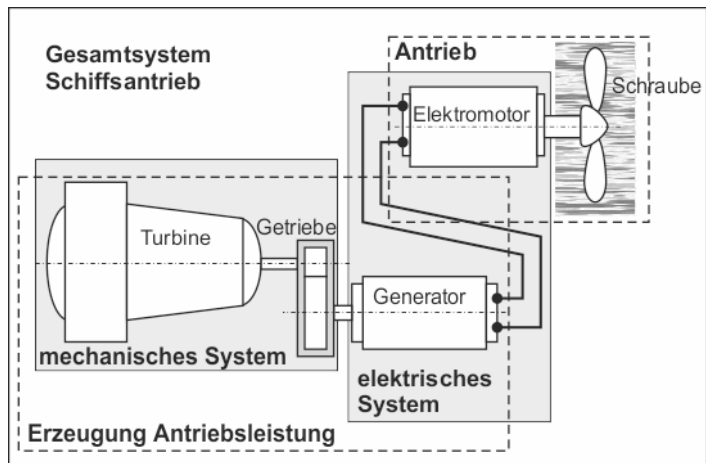


Bild 2-5: Gesamtsystem eines Schiffsantriebs mit unterschiedlichen Subsystemen

Wie bereits oben erwähnt, sind die Systemgrenzen nicht gleichbedeutend mit den physikalischen Grenzen von Objekten, sondern hängen von der Fragestellung ab, die mit Hilfe der Systembeschreibung behandelt werden soll. Bild 2-5 zeigt ein Sys-

¹ auch Beharrungszustand, Gleichgewichtszustand oder Ruhezustand genannt, dadurch gekennzeichnet, dass die zeitveränderlichen Größen des Systems konstant sind.

tem, wie es heute bei modernen Passagierschiffen als Antrieb verwendet wird. In älteren Schiffen war der Antriebsmotor irgendwo im Schiffsinnen und die Schraube wurde über eine lange Welle angetrieben. Diese lange Welle mit der Durchführung durch die Schiffsaußenhaut führt zu einer starken Geräuscentwicklung und es wird zur Steuerung des Schiffes als zusätzliches Bauelement ein Steuerruder benötigt. Das in Bild 2-5 dargestellte neuartige Konzept besitzt keine direkte mechanische Verbindung zwischen Primärtrieb (Turbine) und Schraube mehr. Die Turbine treibt einen elektrischen Generator an, der Strom erzeugt. Die Schraube sitzt direkt auf der Welle eines Elektromotors, der sich außerhalb des Schiffsrumpfes in einer Gondel befindet. Die Antriebsleistung für den Elektromotor wird über Kabel in diese Gondel übertragen. Die paarig, symmetrisch zu beiden Seiten des Schiffes angebrachten Gondeln können geschwenkt werden, so dass der Antrieb des Schiffes auch unter einem bestimmten Verdrehwinkel zur Mittelachse erfolgen kann. Deshalb benötigt ein solches Schiff für die Kurssteuerung kein Steuerruder mehr.

Innerhalb dieses Gesamtsystems kann man, je nach Fragestellung die behandelt werden soll, nun verschiedene Subsysteme herausgreifen, wobei einzelne Teilsysteme in unterschiedlichen Subsystemen auftauchen können. So können für die Behandlung eines mechanischen Systemanteils die Turbine und das Getriebe zum Subsystem "mechanisches System" zusammengefasst werden. Entsprechend können als elektrischer Systemanteil der Generator und der Elektromotor zum Subsystem "elektrisches System" zusammengefasst werden. Stehen mehr die Gesichtspunkte "Erzeugung der Antriebsleistung" oder der eigentliche "Antrieb" im Vordergrund, so liegt nun der Generator mit dem mechanischen System innerhalb eines Subsystems und der Elektromotor mit der Schraube zusammen in einem anderen Subsystem. Man sieht an diesen Beispielen, dass es keine natürliche, physikalische Zugehörigkeit bestimmter Objekte zu bestimmten Subsystemen gibt, sondern dass sie je nach Fragestellung zu dem einen oder auch zu einem anderen Subsystem gehören können.

2.2 Verfahren der Modellbildung

Modelle dienen zur Beschreibung der Eigenschaften und der Struktur eines Systems. Sie sind nie ein absolut vollständiges Abbild eines Systems. Je nachdem, welchen Zweck man mit der Modellbildung verfolgt, gibt es verschiedenartige Modelle mit unterschiedlichen Eigenschaften. In Bild 2-6 sind unterschiedliche Modelltypen aufgeführt. Dabei unterscheidet man **physikalische Modelle** und **mathematische Modelle**. Physikalische Modelle sind stets gegenständlich und maßstäblich, mathematische Modelle sind abstrakt und dienen einer formalen Beschreibung der Systemeigenschaften. Bei den physikalischen Modellen unterscheidet man folgende Arten:

- Prototypmodell
- Pilotmodell
- Ähnlichkeitsmodell

Das **Prototypmodell** ist 1:1-maßstäblich und besitzt höchste qualitative und quantitative Ähnlichkeit. Wie im Beispiel gezeigt, wird ein solcher Prototyp beispielsweise vor der Serienherstellung von einem PKW erstellt. Dies ist ein weitestgehend mit den Serieneigenschaften ausgestatteter Originalaufbau, an dem alle Eigenschaften des späteren Originals direkt und konkret getestet werden können. Nachteil eines solchen Prototypmodells ist, dass seine Herstellung aufwändig und teuer ist und nur geringe Flexibilität bei erforderlichen Änderungen besitzt. Die Erstellung eines solchen Modells wird daher nur der letzte Schritt vor Serienanlauf eines Massenproduktes sein.

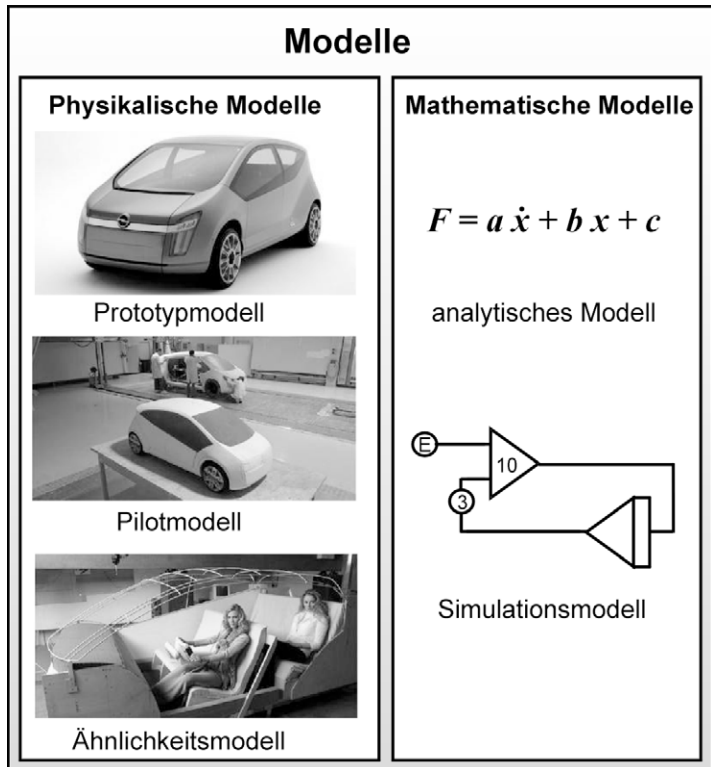


Bild 2-6: Unterschiedliche Arten von Modellen

Das **Pilotmodell** ist häufig maßstäblich unterschiedlich zum Original z. B. 1:10. Es bildet daher nur wesentliche Eigenschaften genau ab. Seine Herstellung ist in der Regel mit reduziertem Aufwand möglich und es lässt sich einfacher ändern. Häufig ist die Aufgabe eines solchen Modells nur die Visualisierung, um beispielsweise das Design beurteilen zu können.

Den geringsten Aufwand zur Herstellung eines physikalischen Modells tritt beim **Ähnlichkeitsmodell** auf. Es werden hier nur noch Teile des Systems hergestellt, an denen man ein eingeschränktes Spektrum von Untersuchungen vornehmen kann. So könnten unter Berücksichtigung der Ähnlichkeitsverhältnisse an einem solchen Ähnlichkeitsmodell Untersuchungen im Windkanal über das Strömungsverhalten der Karosserie gemacht werden, d. h. es handelt sich um Untersuchungen während des Entwicklungsprozesses.

Deutlich flexibler und mit geringem Aufwand herstellbar sind abstrakte mathematische Modelle. Für ein **analytisches Modell** muss man die analytischen Zusammenhänge zwischen den Attributen eines Systems bestimmen, was einen Satz von Gleichungen liefert, die eine geschlossene, analytische Lösung besitzen. Dies ist in der Regel aber ohne Rechneinsatz nur für sehr einfache Systeme möglich. Für einige einfache Systeme werden im folgenden die Vorgehensweise zur Erstellung

eines mathematischen Modells und die dabei auftretenden Probleme beschrieben.

Komplexere Systeme kann man mit Hilfe eines **Simulationsmodells** behandeln. Dieses Modell wird auf einem Digitalrechner erstellt und mit Hilfe numerischer Rechenverfahren gelöst. Dieses Thema wird ausführlich im Kapitel 8 behandelt.

2.2.1 Theoretische Modellbildung

Um ein mathematisches Modell eines realen Systems zu entwickeln, stehen zwei verschiedene Vorgehensweisen zur Verfügung. Liegen relativ genaue Kenntnisse der inneren Zusammenhänge eines System vor, so liefert eine theoretische Systemanalyse ein theoretisches Modell. Sind kaum Kenntnisse über die Beziehung der Attribute zueinander und über die Struktur des Systems bekannt, so muss man experimentelle Methoden anwenden, die sogenannten **Identifikationsverfahren**. Hierüber wird noch im Kapitel 7 berichtet.

2.2.1.1 Allgemein bekannte Modellvorstellungen

Eine praktisch jedem bekannte Modellvorstellung ist die des mathematischen Punktes. Diese Modellvorstellung ist jedem so selbstverständlich, dass man normalerweise gar nicht mehr wahrnimmt, dass es sich dabei nicht um die Beschreibung der Realität, sondern um ein Modell der Realität handelt. In Bild 2-7 a) ist dargestellt, was man unter einem "Punkt" versteht und wie man seine Eigenschaften beschreibt. Man geht davon aus, dass der Punkt keine räumliche Ausdehnung hat. Damit wird klar, dass der Punkt nichts Reales ist - denn reale Objekte haben immer

eine räumliche Ausdehnung - sondern dass es sich um eine Modellvorstellung handelt.

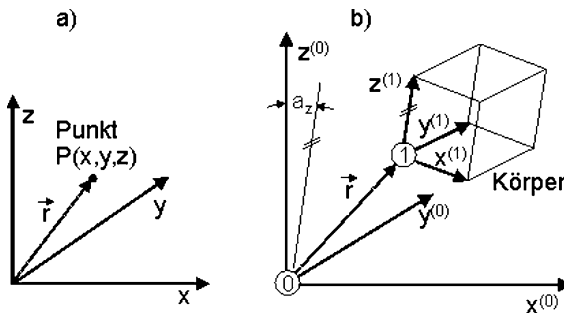


Bild 2-7: a) mathematischer Punkt in kartesischen Koordinaten b) Positionsbeschreibung eines realen Körpers

Man wählt eine solche Modellvorstellung, weil sich unter ihrer Verwendung Fragestellungen der Mathematik, insbesondere der Geometrie, leichter beschreiben lassen. So reichen für die eindeutige Positionsbestimmung eines Punktes im Raum bei Verwendung eines kartesischen Koordinatensystems² drei Koordinatenangaben x , y , z aus.

Reale Körper haben jedoch nicht 3 sondern 6 Freiheitsgrade im Raum. Ihre Lage im Raum kann durch einen dreidimensionalen Ortsvektor $\vec{r}$ im raumfesten Koordinatensystem (0) und durch drei Verdrehwinkel $\alpha_x, \alpha_y, \alpha_z$ eines körpereigenen Koordinatensystems (1) gegenüber dem raumfesten beschrieben werden (Bild 2-7 b). Bei-

2 Ein von drei paarweise senkrechten Einheitsvektoren mit gemeinsamen Ursprungspunkt aufgespanntes Koordinatensystem; benannt nach dem französischen Mathematiker René Descartes (Renatus Cartesius), 1596-1650.

spielsweise benötigt man für die Beschreibung der Positionsänderungen bei Handhabungsaufgaben von Objekten Angaben über die 6 Freiheitsgrade, die Modellvorstellung "mathematischer Punkt" hilft hier nicht weiter.

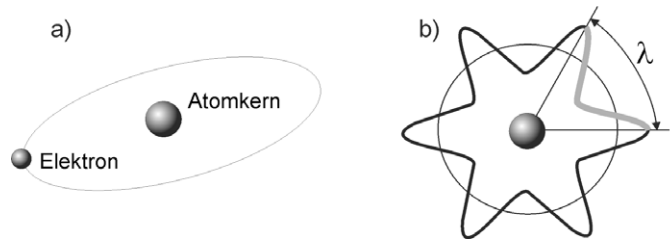


Bild 2-8: a) einfaches Atommodell mit Elektron als Kugel

b) Elektron als stehende Welle mit Wellenlänge λ

Eine andere aus der Kernphysik bekannte einfache

Modellvorstellung ist die des Elektrons als kompakte Kugel, die um den Atomkern kreist (Bild 2-8 a), oder sich frei in einem Kontinuum bewegen kann. Dieses einfache Modell reicht aus, wenn man die Geschwindigkeit oder den Ort des Elektrons beschreiben will. Hier existiert jedoch ein Problem, das durch die **Heisenberg'sche Unschärferelation** beschrieben wird. Sie besagt, dass Ort und Geschwindigkeit eines Elektrons nicht gleichzeitig beliebig genau bestimmt werden können. Dies gibt einen Hinweis darauf, dass die einfache Modellvorstellung nicht uneingeschränkt die Realität beschreibt.

Einen weiteren Einwand gegen dieses Modell liefert die Überlegung, dass ein um ein Atom kreisender geladener Körper (das Elektron) elektromagnetische Wellen abstrahlen müsste und dadurch Energie verlieren würde. Die Folge davon wäre, dass das Elektron nach einiger Zeit in den Kern stürzen würde.

Eine Lösung dieses Modellproblems liefert die Interpretation des Elektrons als stehende Welle mit der Wellenlänge λ (Bild 2-8 b)). Dabei sind nur Bahnen mit ganzzahligem Vielfachen von λ möglich, was die Existenz einzelner Elektronen auf unterschiedlichen Bahnen erklärt.

An diesen Beispielen sieht man, dass Modelle nicht die Realität in allen Details wiedergeben müssen, sondern dass diese in der Regel die Verhältnisse stark vereinfachen. Dadurch werden bestimmte Fragestellungen übersichtlicher und sind einfacher zu beantworten. Außerdem wird durch diese Vorgehensweise der mathematische Aufwand für die Modellbeschreibung gering gehalten.

2.2.1.2 Vorgehensweise bei der Modellbildung

Das Verhalten der häufig behandelten kontinuierlichen Systeme lässt sich durch wenige physikalische Grundgesetze beschreiben. Solche Gesetze sind beispielsweise:

- Newton'sche Axiome der Mechanik
- Hebelgesetze
- Hauptsätze der Thermodynamik
- Ohm'sches Gesetz und Kirchhoff'sche Regeln

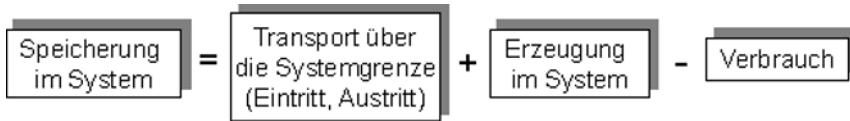


Bild 2-9: Bilanzgleichung zur Erstellung eines mathematischen Modells

Häufig lassen sich mit Hilfe dieser Grundgesetze Bilanzgleichungen für gespeicherte Energien, Massen und Impulse herleiten (Bild 2-9), deren Formulierung in der Regel auf Differentialgleichungen führt, d. h. die behandelten Größen treten in der Gleichung auch in Form ihrer Ableitungen auf. Hängen die Zustandsgrößen des behandelten Systems nur von der Zeit t ab, so kann man die Systeme durch **gewöhnliche Differentialgleichungen** (s. a. Kapitel 4) beschreiben. Man spricht dann auch von Systemen mit **konzentrierten Parametern**.

Hängen die Zustandsgrößen außer von der Zeit t auch noch von anderen Größen wie beispielsweise dem Ort x oder dem Druck p ab, so sind für die mathematische Modellbeschreibung **partielle Differentialgleichungen** erforderlich, d. h. die Zustandsgrößen müssen partiell nach mehreren Variablen abgeleitet werden. Man spricht dann von Systemen mit **verteilten Parametern**.

Um die Bilanzgleichung nicht zu kompliziert werden zu lassen, führt man häufig Randbedingungen und Einschränkungen ein, die einerseits eine mathematische Lösung ermöglichen, aber andererseits die Gültigkeit des Modells auf bestimmte Aspekte und Fälle beschränken.

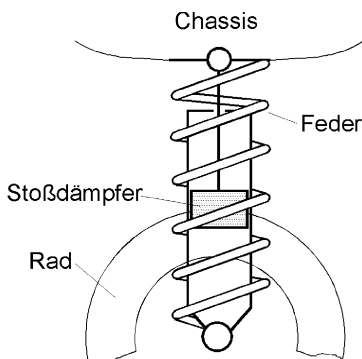


Bild 2-10: PKW-Federbein

Dies wird am Beispiel des mechanischen Einmassenschwingers aus Bild 1-9 deutlich, das jeder Technikinteressierte schon einmal im Physikunterricht kennen gelernt hat. Für welches technische System könnte ein solches Modell stehen? Eine Möglichkeit wäre die Aufhängung eines PKW-Rades, die aus einer Feder/Dämpfer Kombination aus Schraubenfeder und Stoßdämpfer (Bild 2-10) besteht. Betrachtet man dieses reale System, so kann man schnell erkennen, dass das Modell des viskos gedämpften Einmassenschwingers nur durch Vernachlässigung einer Anzahl realer Einflüsse für dieses System Gültigkeit besitzt.

So ist eine wichtige Einschränkung des Modells, dass es nur einen Freiheitsgrad besitzt, da es nur lineare Bewegungen in Richtung der Koordinate x (Bild 1-9) zulässt. Im realen System ist der Stoßdämpfer an der Karosserie drehbar aufgehängt, wodurch Drehbewegungen des Gesamtsystems um die Aufhängung möglich sind. Diese treten natürlich auch auf, da die äußere Zwangskraft $F(t)$ nicht nur in Richtung von x als Reaktionskraft zwischen Reifen und Untergrund auftritt. Die Feder wird also nicht nur in x -Richtung verformt, sondern auch seitlich belastet. Außerdem wurde ein lineares Dehnungsverhalten der Feder im ganzen Arbeitsbereich vorausgesetzt. Schlägt diese bei extremen Stößen jedoch durch, so verhält sie sich wegen der dann auftretenden Begrenzung stark **nichtlinear**.

Für den pneumatischen Stoßdämpfer wurde eine **viskose Dämpfung** angenommen, die geschwindigkeitsproportional ($F = d \cdot \dot{x}$) ist. Dies gilt für die hauptsächlich auftretende Dämpfung durch das Komprimieren und Abströmen der Luft im Dämpfer, jedoch nicht für die Reibung der Dichtung an der Außenwand. Hier liegt trockene oder **Coulomb-Reibung** vor ($F = \mu \cdot F_N$). Das Abklingverhalten von Schwingungsvorgängen ist in Abhängigkeit von diesen Reibungstypen mit dem entsprechenden Reibverhalten unterschiedlich. Bei trockener Reibung klingt die gedämpfte Schwingung linear ab (Bild 2-12), bei viskoser Reibung folgt das Abklingverhalten einer Exponentialfunktion (Bild 2-11).

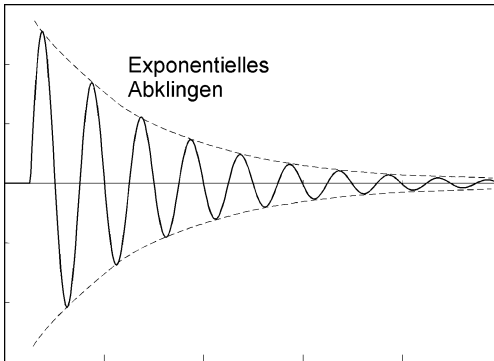


Bild 2-11: Abklingen der Schwingung eines Einmassenschwingers mit Dämpfung durch viskose Reibung

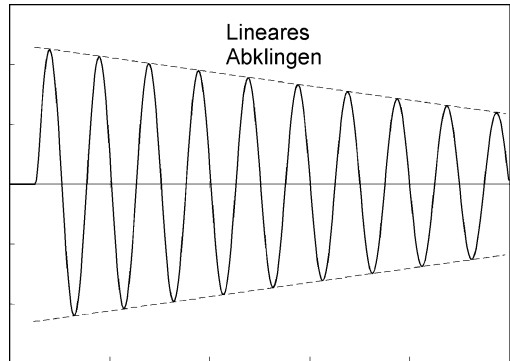


Bild 2-12: Abklingen der Schwingung eines Einmassenschwingers mit Dämpfung durch coulombsche Reibung

Weiterhin wurden bei der Modellbildung alle Teilmassen zu einer Masse m zusammengefasst und in einem Punkt konzentriert angenommen, um den Angriffspunkt der Massenkräfte eindeutig festzulegen. Im realen System sind die Massen über das ganze System verteilt, weshalb der Schwerpunkt nur schwer zu bestimmen ist und seine Lage verändert sich auch noch. Schließlich wurden untergeordnete Kräfte wie beispielsweise der Luftwiderstand des Rades oder des Stoßdämpfers weggelassen.

Auf den ersten Blick scheint daher das reale System nur wenig mit dem einfachen Modell des Einmassenschwingers zu tun zu haben. Weshalb taucht er trotzdem in jedem Physikbuch und in allen Abhandlungen über Kinetik immer wieder auf? Ein historisch bedingter Grund ist leicht einsehbar. Um mit klassischen Methoden, d. h. mit Papier, Bleistift und Logarithmentafel das Problem der Ermittlung des Bewegungszustandes des Systems zu beliebigen Zeitpunkten vornehmen zu können, muss das Modell so einfach gestaltet werden. Erst die Verfügbarkeit leistungsfähiger Digitalrechner lässt heute das Durchrechnen komplexerer Modelle zu, die das reale Verhalten von Systemen noch besser und auch in Grenzbereichen beschreiben.

Zum anderen kann man auch schon aus dem einfachen Modell, mit einer in der Technik hinreichenden Genauigkeit, bestimmte Kenngrößen ermitteln und das reale System dimensionieren. Größere Fehler treten ja nur auf, wenn die vernachlässigten Kräfte oder die vereinfachenden Annahmen durch Extremsituationen in solchen

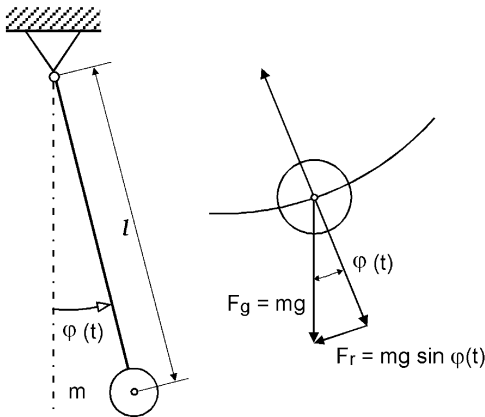


Bild 2-13: Pendel und angreifende Kräfte

Bereichen liegen, in denen sie nicht mehr so ohne weiteres vernachlässigt werden können.

Ein Beispiel für die Modellierung der Dynamik eines einfachen Systems, bei dem das Verlassen des **Gültigkeitsbereichs** der Modellannahmen zu starken Abweichungen zwischen Modell und Realität führt, ist das jedem bekannte mathematische Pendel. Das Bild 2-13 zeigt das Schema eines Pendels und außerdem die an der Masse angreifenden Kräfte. Wird die Masse aus der Ruhelage um den Winkel $\varphi = \varphi(t)$ ausgelenkt, so wirkt auf sie infolge der Massenkraft $F_g = m \cdot g$

in der zur Auslenkung entgegengesetzten Richtung die Rückstellkraft $F_r = m \cdot g \cdot \sin \varphi$. Die Bogenlänge beträgt dabei $l \cdot \varphi$, die Beschleunigung $l \cdot \ddot{\varphi}$. Durch Einsetzen in das Newton'sche Bewegungsgesetz erhält man:

$$m \cdot l \cdot \ddot{\varphi}(t) = -m \cdot g \cdot \sin \varphi(t)$$

oder

$$m \cdot l \cdot \ddot{\varphi}(t) + m \cdot g \cdot \sin \varphi(t) = 0 \quad (2.1)$$

Dies ist eine Differentialgleichung, zu der noch zusätzlich die Anfangsbedingungen festgelegt werden müssen:

$$\varphi(t=0) = \varphi_0 \quad \ddot{\varphi}(t=0) = \ddot{\varphi}_0 = 0$$

Bei dieser Modellierung wurden natürlich wieder vereinfachende Annahmen getroffen, nämlich dass der Faden masselos ist und alle Masse in einem Punkt - dem **Schwerpunkt** - konzentriert ist. Außerdem wurden Kräfte durch Luftwiderstand und Lagerreibung vernachlässigt.

Aus der Bewegungsgleichung und den Anfangsbedingungen lässt sich eine Lösung gewinnen, die die freien Schwingungen des Pendels beschreibt. Jedoch handelt es sich bei der Bewegungsgleichung, da φ sowohl als 2. Ableitung als auch als Argument der Sinusfunktion auftaucht, um eine **nichtlineare Differentialgleichung**, deren Lösung schwierig ist. Daher wird gerne der Fall behandelt, dass das Pendel nur sehr kleine Ausschläge macht, d. h. . Unter dieser Voraussetzung gilt nämlich

$$\sin \varphi(t) \approx \varphi(t) \quad (2.2)$$

Damit kann die Differentialgleichung folgendermaßen linearisiert werden, wodurch sie leichter lösbar ist:

$$m \cdot l \cdot \ddot{\varphi}(t) + m \cdot g \cdot \varphi(t) = 0 \quad (2.3)$$

Dass dieses mathematische Modell für das Pendel aber nur sehr eingeschränkt gilt, kann man leicht an der Bildfolge in Bild 2-14 erkennen, die durch Simulation des

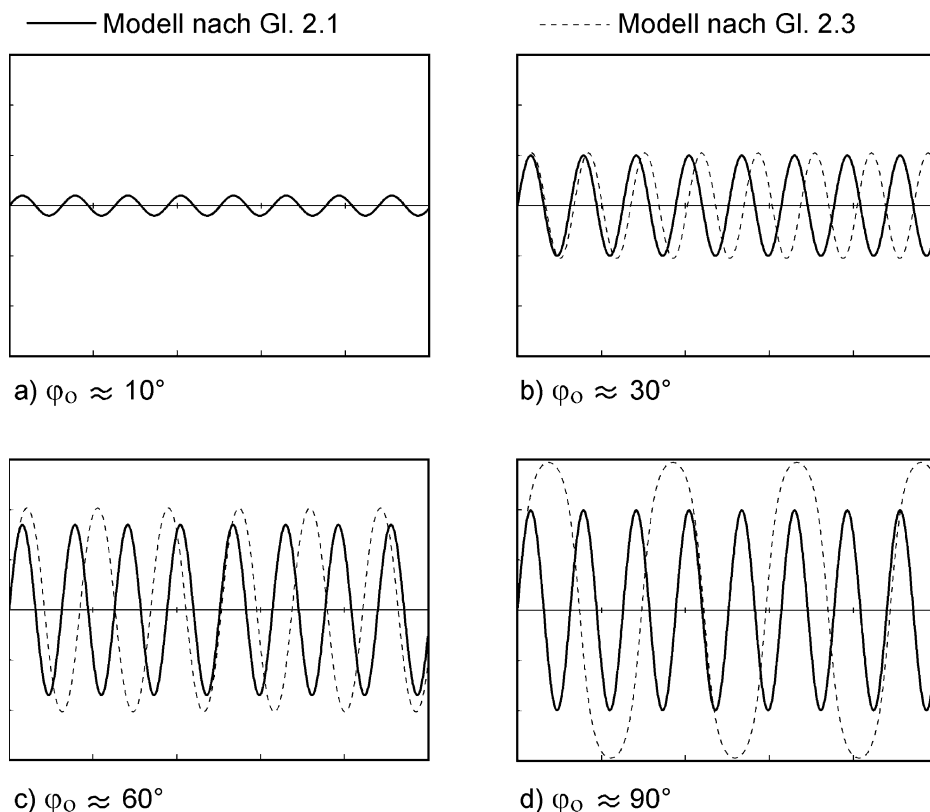


Bild 2-14: Simulation des Schwingungsverlaufs verschiedener Modelle eines Pendels für unterschiedliche Anfangsauslenkungen

Modells mit einem Simulationssystem erstellt wurde. Sie zeigt in jedem Teilbild die Schwingungsverläufe nach Loslassen aus einer ausgelenkten Stellung für beide Modellgleichungen (2.1) und (2.3). Alle Teilbilder haben den gleichen Amplituden- und Zeitmaßstab. Im Teilbild a) ist der Winkel noch sehr klein ($\varphi_0 \approx 10^\circ$), so dass Gl. (2.2) gilt und die beiden Schwingungsverläufe der unterschiedlichen Modelle kaum zu unterscheiden sind. In den Teilbildern b) - d) wird der Auslenkungswinkel schrittweise bis auf $\varphi_0 \approx 90^\circ$ vergrößert. Bei $\varphi_0 \approx 30^\circ$ weichen die beiden Modelle erst nach mehreren Schwingungen deutlich voneinander ab, bei $\varphi_0 \approx 60^\circ$ wird die Abweichung in der Frequenz schon nach einer Schwingung sichtbar, bei $\varphi_0 \approx 90^\circ$ tritt sofort eine starke Abweichung in Frequenz und Winkel auf. Das vereinfachte Modell nach Gl. (2.1) liefert also nur für den kleinsten Bereich der möglichen Anfangsauslenkungen φ_0 den richtigen Wert für $\varphi(t)$. Trotzdem wäre der Aufwand für die rechnerische Behandlung des Modells nach Gl. (2.3) unnötig hoch, wenn man ein technisches System untersuchen würde, in dem ein Pendel vorkommt, das nur Ausschläge geringer Amplitude ausführt und daher zur Beschreibung auch das Modell nach Gl. (2.1) ausreicht.

Für die Erstellung des Modells eines technischen Systems gelten drei allgemeine Anforderungen:

- Die Modellelemente müssen klar definiert, eindeutig beschreibbar und in sich widerspruchsfrei sein (**physikalische Transparenz**).
- Die Folgerungen über das Verhalten, die man aus den Verknüpfungen der Modellelemente zu einem Gesamtmodell ziehen kann, müssen im Rahmen des Modellzwecks (**Gültigkeitsbereich**) dem realen Systemverhalten entsprechen (**Modellgültigkeit**).
- Gibt es verschiedene Möglichkeiten zur Darstellung des Systems, die alle den ersten beiden Forderungen genügen, so sollte man die einfachst mögliche auswählen (**Effizienz**).

Für die Herleitung eines einfachen, effizienten und gültigen Modells gibt es keine in allgemeingültige Regeln fassbare Vorgehensweise.

Das Modell eines mechanischen Systems, das beispielweise alle nur denkbaren Bewegungsmöglichkeiten berücksichtigt, ist zwar physikalisch richtig, aber für die praktische Anwendung unübersichtlich, unhandlich und verliert für die meisten Fälle die physikalische Überschaubarkeit. Die Kunst bei der Modellbildung besteht daher darin, das Modell so einfach wie möglich zu gestalten, um es mit technisch und wirtschaftlich vertretbarem Aufwand untersuchen zu können. Dabei dürfen aber keine unzulässigen, das Systemverhalten zu stark verfälschenden Annahmen getroffen werden. Albert Einstein drückte diese Maxime einmal treffend so aus:

Alles sollte so einfach wie möglich gemacht werden, aber nicht einfacher.

Bei mechatronischen Systemen steht häufig die mechanische Struktur des Bewegungssystems bei der Modellbildung im Vordergrund. Innerhalb des Bewegungssystems werden dann meist auch elektrische Komponenten aus dem Bereich der Aktoren (z. B. Elektromotoren) und der Sensoren in die Modellbildung einbezogen.

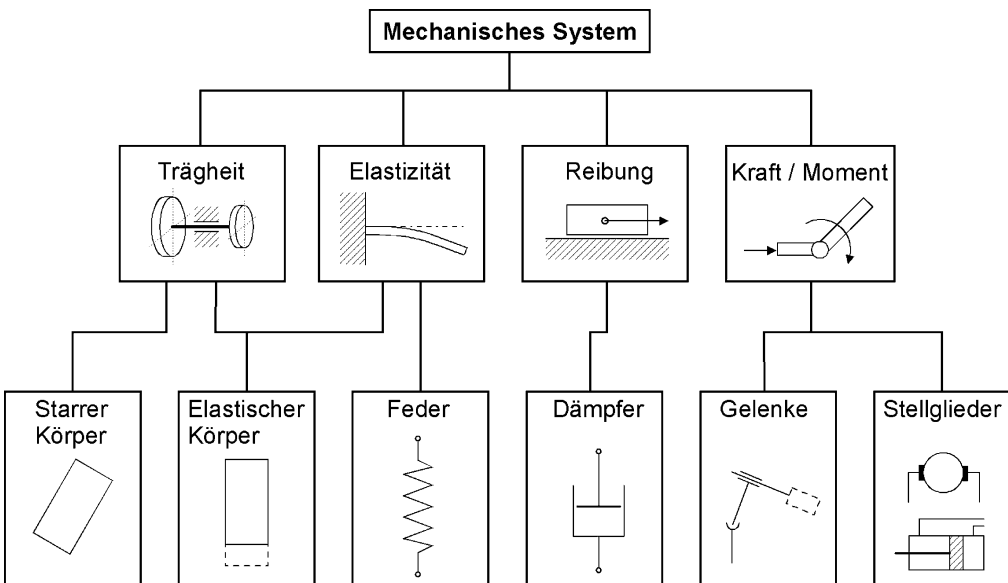


Bild 2-15: Elemente für Modelle von mechatronischen Systemen

Die mechanischen Eigenschaften von mechatronischen Systemen sind im wesentlichen durch Trägheit, Elastizität und Reibungsvorgänge gekennzeichnet, die den durch äußere Kräfte und Stellkräfte bzw. Momente hervorgerufenen Bewegungszustand beeinflussen. Diese Merkmale der Bauelemente des Systems werden durch idealisierte Modelle repräsentiert. Dabei werden Körpermodelle in der Regel als mit Masse und Trägheit behaftet angenommen, jedoch solche Elemente wie Federn und Dämpfer als masse- und trägheitslos. Sich translatorisch oder rotatorisch zueinander bewegend Körper sind durch Gelenke oder Führungen miteinander verbunden, wobei die Einschränkung des Freiheitsgrades der Bewegung von Körpern Reaktionskräfte und -momente zur Folge hat. Bild 2 - 15 zeigt eine Zusammenstellung wichtiger verwendeter Ersatzmodelle und die ihnen zugeordneten Eigenschaften.

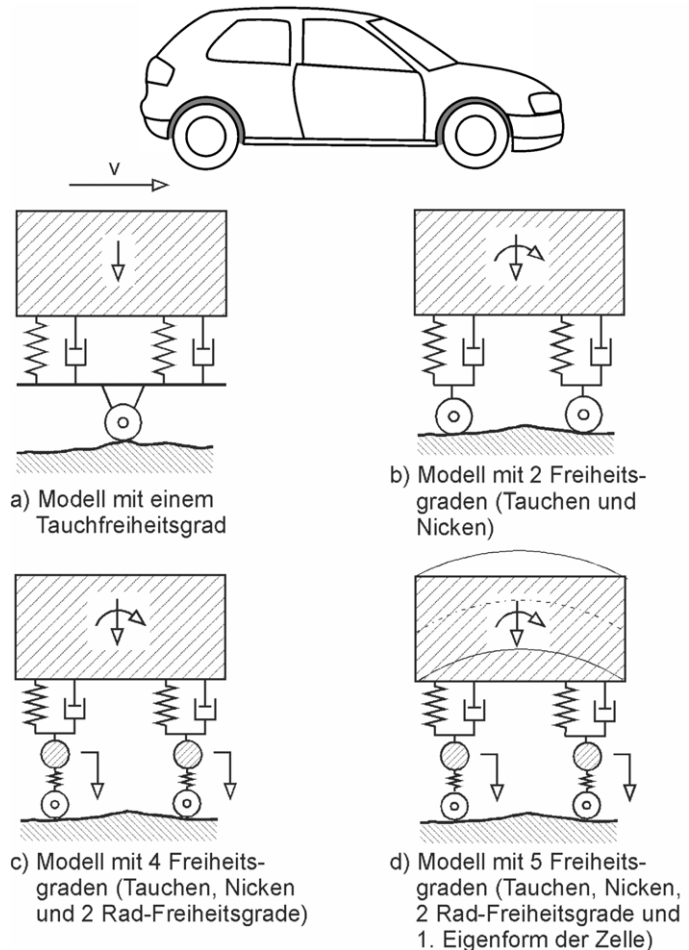


Bild 2-16: Ebene mechanische Modelle mit unterschiedlicher Anzahl von Freiheitsgraden für einen Personenkraftwagen

Zur Bildung des Modells eines mechatronischen Systems versucht man, die einzelnen realen Objekte durch die in Bild 2-15 dargestellten Ersatzmodelle zu beschreiben und damit eine kinematische Struktur aufzubauen. Dabei muss man immer beachten, welche Fragen man mit dem Modell beantworten will. Ein kompliziertes technisches System wie beispielsweise ein Personenkraftwagen auf welliger Straße hat nicht einfach eine bestimmte Zahl von **Freiheitsgraden**, sondern die Anzahl der Freiheitsgrade, die man notwendigerweise einführen muss, hängt davon ab, welche Informationen man benötigt. In Bild 2-16 ist ein Beispiel für die Modellierung eines Personenkraftwagens gegeben [2.2], wobei von Stufe zu Stufe immer mehr Freiheitsgrade eingeführt werden.

Im sehr einfachen Modell in Bild 2 - 16 a) mit einem Freiheitsgrad wurde die Reifenfederung und -dämpfung mit der Federung und Dämpfung zwischen Rad und Auf-

bau zusammengefasst. Das Rad ist als starrer Körper idealisiert. Dieses Modell liefert bezüglich des **Tauch-Freiheitsgrades** vernünftige Aussagen für die Abstimmung des Systems, die im allgemeinen so erfolgt, dass die Taucheigenfrequenz bei etwa 1 bis 2 Hz und der Dämpfungsgrad bei 0,2 bis 0,3 liegt. Für eine genauere Untersuchung des Fahrkomforts muss zumindest der **Nick-Freiheitsgrad** wie in Bild 2-16 b) mit einbezogen werden. Erst durch ihn kommt der Zeitunterschied zur Geltung, der zwischen Vorder- und Hinterrad beim Überfahren einer Bodenwelle auftritt. Dieses Modell gibt aber nur unzureichend Auskunft darüber, ob beim Überfahren von Hindernissen Radentlastungen bis hin zu kurzzeitigem Abheben auftreten. Darüber kann erst das in Bild 2-16 c) dargestellte Modell Aussagen machen, das die **Vertikal-Freiheitsgrade** der Achsmassen berücksichtigt. Mit diesem Modell erfasst man den Frequenzbereich bis 15 Hz schon sehr gut. Ein Modell, das bis 25 Hz gute Aussagen liefert, muss die Annahme einer starren Karosserie aufgeben und als zusätzlichen Freiheitsgrad die 1. Biegeschwingungseigenform der Karosserie einbeziehen (Bild 2-16 d)).

Das letzte benutzte Modell ist aber natürlich immer noch kein allgemeingültiges Modell des realen Systems, da es zweidimensional ist und nur die Untersuchung von Vertikalschwingungen zulässt. Ein entsprechendes räumliches Modell wird noch über erheblich mehr Freiheitsgrade verfügen müssen. Man sieht an diesem Beispiel jedoch gut, dass die Komplexität des Modells nicht unabhängig von der Fragestellung an das Modell ist.

2.3 Klassifizierung dynamischer Systeme

Wie wir im vorherigen Abschnitt gesehen haben, können bei der Erstellung eines gültigen und effizienten Modells eines dynamischen Systems verschiedene Probleme auftreten. Dies liegt natürlich an der physikalischen Beschaffenheit und der Komplexität eines Systems. Verschiedenartige physikalische Systeme haben bestimmte globale Eigenschaften, die sie für die Modellbildung entweder leicht handhabbar machen oder bewirken, dass eine mathematische Beschreibung mit großen Schwierigkeiten verbunden ist.

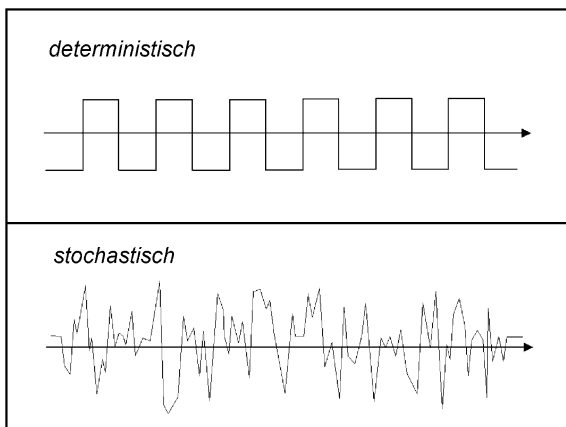


Bild 2-17: Deterministische und stochastische Signale

Bei den meisten dynamischen Systemen in der Technik kann man davon ausgehen, dass die in den Systemen auftretenden physikalischen Größen bzw. die sie darstellenden Signale nicht regellos schwanken. Vielmehr können sie analytisch beschrieben werden, weshalb man von **determinierten** Größen und Systemen spricht. Dabei kann die analytische Beschreibung entweder durch Differentialgleichungen erfolgen und man spricht dann von **parametrischen Modellen**, oder der Zu-

sammenhang zwischen Ein- und Ausgangsgrößen des Systems liegt in Form von Wertetabellen oder Kurvenverläufen vor, wobei man von **nichtparametrischen Modellen** spricht.

Im Gegensatz zu den determinierten Systemen stehen **stochastische** Systeme, deren Zustand nicht durch analytische Beschreibungen vorherbestimmt werden kann. Beispiele für stochastische Größen sind Rauschsignale, hervorgerufen durch thermische Bewegungen der Moleküle. Solche Signale können höchstens mit den Methoden der Statistik behandelt werden. Bild 2-17 zeigt stochastische Signale und ein deterministisches Signal. Ferner kann man bei vielen Systemen davon ausgehen, dass die Systemgrößen innerhalb gewisser Grenzen jeden beliebigen Wert annehmen können. Man spricht dann von **kontinuierlichen Systemen**. Determinierte, kontinuierliche Systeme können durch partielle Differentialgleichungen beschrieben werden. Vielfach kann man diese durch örtliche Diskretisierung zu gewöhnlichen Differentialgleichungen vereinfachen. Solche Systeme, die als unabhängige Variable nur noch von der Zeit abhängen, heißen Systeme mit **konzentrierten Parametern** im Gegensatz zu Systemen mit **verteilten Parametern**, die durch partielle Differentialgleichungen beschrieben werden.

Eine Systemeigenschaft, die mathematische Behandlung von Systemen besonders vereinfacht, ist die **Linearität**. Für solche Systeme und ihre Systemgrößen gelten folgende zwei Prinzipien:

a) Das **Verstärkungsprinzip**

- Wenn der Zusammenhang zwischen der Ausgangsgröße $x_a(t)$ und der Eingangsgröße $x_e(t)$ durch eine allgemeine lineare Operation $x_a(t) = \text{Op}[x_e(t)]$ beschrieben werden kann, dann gilt für eine beliebige konstante Größe c :

$$x_a(t) = \text{Op}[c \cdot x_e(t)] = c \cdot \text{Op}[x_e(t)]$$

b) Das **Superpositionsprinzip** (Überlagerungsprinzip)

- Führt die Anwendung der linearen Operation Op auf die Summe der Eingangsgrößen $x_{e1}(t)$ und $x_{e2}(t)$ zur Ausgangsgröße, so gilt:

$$x_a(t) = \text{Op}[x_{e1}(t) + x_{e2}(t)] = \text{Op}[x_{e1}(t)] + \text{Op}[x_{e2}(t)]$$

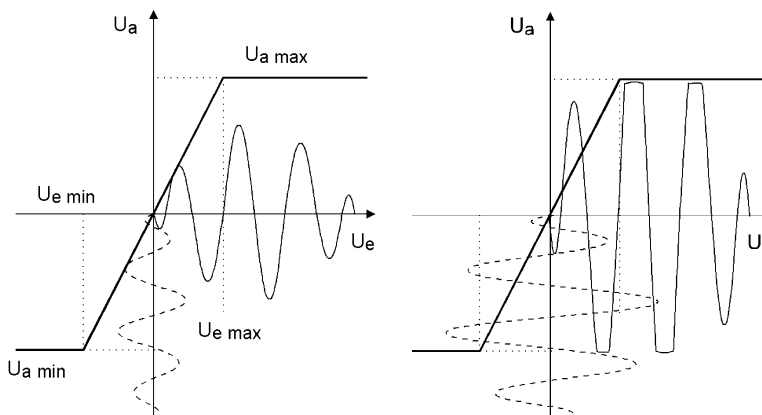


Bild 2-18: Verstärkung eines sinusförmigen Eingangssignals mit und ohne Begrenzung

Nichtlineare Systeme besitzen diese Eigenschaften nicht, was man leicht an dem einfachen Beispiel eines elektrischen Verstärkers verdeutlichen kann. Die in Bild 2-18 dargestellte Kennlinie besagt, dass das Ausgangssignal sich zwischen den Werten $U_{a \max}$ und $U_{a \min}$ ändern kann. Gibt man eine sinusförmige Spannung der Größe

$$U_e < U_{e \max} - U_{e \min}$$

auf den Eingang des Verstärkers, der eine Verstärkung um den Faktor c aufweist, so verstärkt er diese auf die Ausgangsspannung $U_a = c \cdot U_e$.

Zwischen Ein- und Ausgang herrscht ein linearer Zusammenhang und es gelten die oben genannten Prinzipien. Erhöht man nun die Eingangsspannung auf den doppelten Wert, so sieht man, dass das Ausgangssignal nicht im ganzen Bereich den Wert annimmt. Das Ausgangssignal wird auf die maximale Ausgangsamplitude begrenzt und stark verzerrt. Die beiden oben genannten Prinzipien gelten nicht mehr, es liegt eine Nichtlinearität vor. Lässt man als Wertebereich des Eingangssignals jedoch nur solche Werte zu, die höchstens zur maximalen Ausgangsamplitude führen, so hat man das prinzipiell nichtlineare System linearisiert und damit seine mathematische Behandlung stark vereinfacht.

Eine weitere wichtige Systemeigenschaft ist die **Zeitvarianz**. Ändert ein System oder eine Größe ihre Übertragungseigenschaften nicht in Abhängigkeit von der Zeit, so spricht man von **zeitinvarianten** Systemen und das Systemmodell enthält keine von der Zeit abhängigen Parameter. Demgegenüber kann ein **zeitvariantes** System unterschiedliches Verhalten aufweisen, je nachdem, zu welchem Zeitpunkt man das System betrachtet. So sind beispielsweise Flugzeuge meist zeitvariante Systeme, da sich ihre Masse während des Fluges durch den Verbrauch des Treibstoffs stark ändert. Bei zeitvarianten Systemen spielt einerseits die Geschwindigkeit, mit der sich Systemparameter ändern, eine Rolle und andererseits auch die Art des Änderungseffektes. So kann sich ein Parameter durch Driften (langsame Veränderung) oder auch sprunghaft ändern.

Zu den am schwersten zugänglichen Systemen gehören solche mit verteilten Parametern. Bei diesen lässt sich die partielle Differentialgleichung nicht in eine gewöhnliche überführen, da alle oder einzelne Zustandsgrößen des Systems nicht nur von der Zeit, sondern zusätzlich von anderen Größen, wie etwa Ortskoordinaten abhängig sind. Zu dieser Klasse gehören im wesentlichen Strömungs- und Wärmevorgänge.

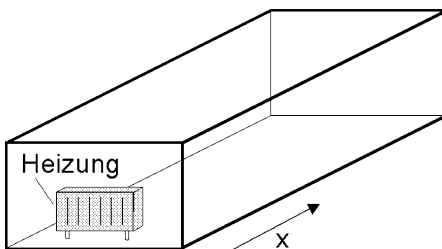


Bild 2-19: System mit verteilten Parametern

Betrachtet man beispielsweise einen langgestreckten Raum (Bild 2-19), an dessen einem Ende sich eine Wärmequelle befindet, so wird eine Temperaturerhöhung zuerst in unmittelbarer Nähe der Wärmequelle (kleine Ortskoordinate in x -Richtung) spürbar sein. Am anderen Ende des Raumes (große Ortskoordinate x) wird eine Temperaturerhöhung erst nach einiger Zeit spürbar sein. Die für das System interessante Zustandsgröße Temperatur hängt also nicht nur von der Zeit, sondern auch von der Ortskoordinate

ab. Sind auch Höhe und Breite des Raumes nicht zu vernachlässigen, so kommen noch zwei weitere Ortskoordinaten hinzu.

2.4 Modellierung von Geometrie und Körpereigenschaften

Schaut man sich ein Lehrbuch der *Technischen Mechanik* an, so beginnt die **Kinematik**, also die Lehre vom geometrischen und zeitlichen Ablauf von Bewegungen immer mit der Bewegung des Massenpunktes. Dieses sehr einfache Modell einer bewegten Masse kann für einen Körper stehen, bei dem die Drehträgeit für eine Drehung um den Körperschwerpunkt vernachlässigbar ist [2.2]. Das sind vergleichsweise kleine Körper oder solche, deren Eigendrehung von den anderen Bewegungen entkoppelt ist und nicht in die Betrachtung eingeht. Solche einfachen Modelle lassen sich zwar zu Ausbildungszwecken recht leicht behandeln, haben aber wenig mit realen Systemen gemein. Bei der Fiktion des Massenpunktes handelt es sich praktisch um den Schwerpunktsatz, der die Behandlung der Eigenschaften des räumlich ausgedehnten Körpers durch die Behandlung des Schwerpunktes des Körpers ersetzt, in dem man sich die Gesamtmasse des Körpers im Schwerpunkt konzentriert denkt.

Das nächst komplexere Modell einer bewegten Masse ist das des **starrten Körpers**. Ein starrer Körper besteht als Modell eines realen Körpers aus einer Vielzahl infinitesimal kleiner Masseelemente dm , deren relative Lage im Körper zueinander stets erhalten bleibt. Für viele massive Bauelemente in mechanischen Systemen wie Bolzen, Balken, Platten u. ä. geben sie ein recht brauchbares Modell ab. Voraussetzung für den starren Körper ist, dass er sich unter äußeren und inneren Kräften nicht verformt, d. h. die Elastizität wird vernachlässigt.

Lässt man im Modell Verformungen des Körpers durch Kräfte zu, so spricht man von einem **elastischen Körper**. Diese Verformungen sind in den meisten Fällen klein. Geht die Verformung im unbelasteten Zustand vollständig zurück, so spricht man von **elastischer Verformung**, andernfalls von **plastischer Verformung**.

Im Gesamtmodell eines mechatronischen Systems sind solche Körper entweder starr oder über Koppellemente miteinander verbunden, die einem bestimmten Kraftgesetz gehorchen und **eingeprägte** Kräfte erzeugen. Dies sind Federn und Dämpfer sowie Stellantriebe. In der Regel werden diese Koppellemente als masselos angesehen.

2.4.1 Mehrkörpersysteme

Ein mechatronisches System besteht in der Regel aus mehr als einem bewegten Körper. Typische mechatronische Systeme wie beispielsweise Industrieroboter bestehen aus mehreren relativ steifen Gliedern, die über Dreh- und Schubachsen miteinander verbunden sind. Solche Systeme kann man sehr gut durch das Konzept des Mehrkörpersystems (MKS) modellieren. Diese Art der Modellbildung wurde in den 70er Jahren eingeführt, als die ersten Programmsysteme für die numerische Berechnung auf Digitalrechnern zur Verfügung standen. Im einfachsten Fall besteht ein Mehrkörpersystem aus einer offenen Kette starrer Körper, die durch starre Ge-

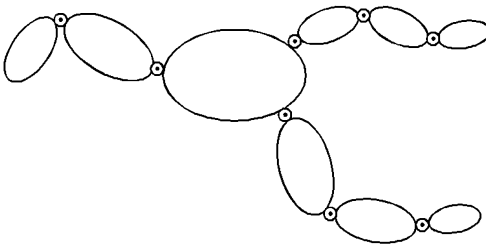


Bild 2-20: Mehrkörpersystem in Form einer offenen Kette starrer Körper

lenke miteinander verbunden sind (Bild 2-20). Diese Starrkörper repräsentieren die Eigenschaften der Masse und der Trägheit. In einer verbesserten Form können diese Systeme auch elastische Körper beinhalten und den Gelenken elastische Eigenschaften zugeordnet werden. Bei sehr gut eigensteifen Systemen wie Industrierobotern tritt der größte Anteil an Elastizität in den Gelenken und Antrieben auf und wird durch Feder- und Dämpferelemente modelliert (Bild 2-15).

In Anlehnung an die Methode der **Finiten-Elemente**, bei der die mechanische Struktur eines Systems aus einer Vielzahl von Einzelementen wie Balken, Scheiben, Platten und Schalenelementen zusammengesetzt wird, spricht man bei dieser Methode auch von der **Finite-Segment-Methode (Finite-Segment-System FSS)**.

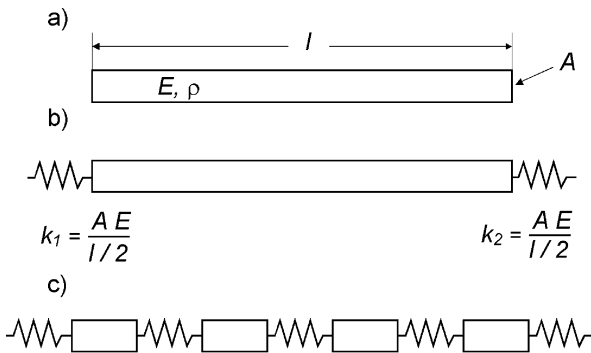


Bild 2-21: Finites Segment Modell eines Stabes der Länge l a) Stab als Starrkörper
b) Modellierung der Steifigkeit
c) Modellierung mit mehreren Segmenten

Mit dieser Methode können die wesentlichen Einflüsse großer Starrkörperbewegungen auf die Dynamik des Systems sehr gut wiedergegeben werden. Da die Anzahl der aus dem Modell hergeleiteten Bewegungsgleichungen vergleichsweise gering ist, eignen sich Mehrkörpersysteme auch gut für die Auslegung der Lagereger bewegter Strukturen und können mit relativ geringem Aufwand numerisch simuliert werden.

Erfordern größere Verformungen der einzelnen Glieder der Mehr-

körperkette eine Modellierung der Elastizität, so kann man im einfachsten Fall beispielsweise einen als Starrkörper modellierten Balken durch eine Kette örtlich verteilter Punktmassen ersetzen, die durch masselose Federn miteinander verbunden sind (**lumped mass system**). Die dadurch eingeführten zusätzlichen Bewegungsmöglichkeiten führen zu einer Erhöhung der Zahl der Freiheitsgrade des Modells. Ein genaueres Ersatzsystem erhält man, wenn man den Balken, wie in Bild 2-21 dargestellt, durch mehrere Starrkörper, die mit Federn untereinander gekoppelt sind, modelliert. Dabei wird die Steifigkeit eines stabförmigen Elementes mit der Querschnittsfläche A und der Länge l ,

$$k = \frac{A \cdot E}{l} \quad \text{Federsteifigkeit}$$

durch zwei Federn der doppelten Steifigkeit an den Stabenden ersetzt, die in Reihe geschaltet wieder die Steifigkeit des Elementes ergeben. Aus einer Vielzahl solcher

Tabelle 2-1: Vergleich der ersten 7 Eigenfrequenzen eines kontinuierlichen Modells und zweier Finite Segment Modelle eines Stabes ($E=120\text{N/mm}^2$, $r=1\text{g/cm}^3$, $l=1\text{m}$, $A=196\text{mm}^2$

Schwingungsmoden	Kontinuierliches Modell [f] = Hz	FSS aus 10 Segmenten [f] = Hz	Abweichung der Modelle in %	FSS aus 30 Segmenten [f] = Hz	Abweichung der Modelle in %
1	108,8	108,4	0,4	108,6	0,0
2	217,6	214,1	1,6	217,3	0,1
3	326,5	314,5	3,7	325,2	0,2
4	435,3	407,2	6,4	432,1	0,7
5	544,1	489,9	9,9	537,9	1,1
6	652,9	560,5	14,2	642,3	1,6
7	761,6	617,3	18,9	744,9	2,2

in Reihe geschalteter Elemente kann man dann recht genau das elastische Verhalten des ursprünglichen Stabes modellieren. Dabei hängt die Genauigkeit des Modells natürlich von der Anzahl der verwendeten Segmente des Finite - Segmente - Systems ab. Tabelle 2-1 zeigt die Ergebnisse einer Untersuchung aus [2.4], wo für den Fall der Anregung einer Längsschwingung im Stab die ersten 7 Eigenfrequenzen rechnerisch ermittelt wurden. Dafür wurden verschiedene Finite-Segmente-Modelle und ein kontinuierliches Modell durchgerechnet. Man sieht, dass die Grundschwingungen bei beiden Finite-Segment-Modellen recht gut mit der wirklichen Resonanzfrequenz des Stabes übereinstimmen, jedoch bei Oberschwingungen höherer Ordnung beim Ersatzmodell aus 10 Starrkörpern eine Abweichung von ca. 19% auftritt. Die Verfeinerung auf 30 Starrkörper vermindert die Abweichung aber bereits auf ca. 2%.

2.4.2 Systeme mit elastischen Elementen

Muss bei Körpern oder Gelenken das elastische Verhalten modelliert werden, so lässt sich das Verformungsproblem eines Bauteils mit den Grundgleichungen der Elastomechanik beschreiben. Elastische Körper einfacher Geometrie, deren Masse - und Elastizitätsverteilung exakt beschrieben werden können, können auch durch **kontinuierliche Systeme (KOS)** modelliert werden. Die mathematische Formulierung des Modells führt auf partielle Differentialgleichungen, die nur für sehr einfache Geometrien exakt gelöst werden können. Um auch für komplexe Bauteile Beschreibungen des Verformungsverhaltens zu ermöglichen greift man auf **Näherungsverfahren** zurück. Bei solchen Näherungsverfahren ist zu unterscheiden, ob die entsprechende Lösung das Verhalten des Bauteils wie bei der exakten Lösung kontinuierlich beschreibt oder lediglich an vorgegebenen diskreten Punkten. Ein Näherungsverfahren mit kontinuierlicher Beschreibung ist beispielsweise die **Methode nach Ritz**, die durch funktionsanalytische Verfahren eine geschlossene Näherung ermöglicht. Diese Näherung kann in einfachen Fällen wieder mit der exakten Lösung übereinstimmen.

Andere Näherungsmethoden liefern diskrete Lösungen, d. h. Lösungen, die nur an bestimmten Punkten des Bauteils gelten. Die beiden bekanntesten Verfahren sind das **Differenzenverfahren** und die **Finite-Elemente-Methode (FEM)**. Beim Differenzenverfahren werden die in den Gleichungen und Randbedingungen vorkommenden Differentialquotienten durch Differenzenquotienten ersetzt. Das dadurch entstehende System algebraischer Gleichungen wird mit den aus der numerischen Mathematik bekannten Verfahren gelöst.

Eine Standardmethode zur Modellierung elastischer Strukturelemente ist die Finite-Elemente-Methode. Sie basiert auf der Idee, dass das zu berechnende Kontinuum in eine große Anzahl einfach berandeter, endlich großer Elemente zerlegt werden kann, die der Berechnung auf dem Digitalrechner leicht zugänglich sind. Aus den Lösungen für die Einzelemente wird dann unter Berücksichtigung von Kontinuitäts- und Gleichgewichtsbedingungen eine Lösung für das Gesamtsystem konstruiert. Die Bedingungen werden dabei nur an einer endlichen Zahl von Punkten, den sogenannten Knoten, formuliert. Sie führen auf ein Gleichungssystem, dessen Lösung im allgemeinen eine Näherungslösung für das behandelte System ist. Diese Methodik, die Ende der 60er Jahre aufkam, war die erste größere numerische Simulationsmethode und war quasi, obwohl schon früher gedanklich vorgezeichnet, in der Praxis an das Aufkommen leistungsfähiger Digitalrechner gebunden. Heute ist die Methode sicher das am meisten benutzte Verfahren, um naturwissenschaftliche und technische Probleme mit Hilfe des Computers zu lösen.

Für die Finite-Elemente-Methode gibt es eine Reihe von gebräuchlichen Verfahren zur Modellierung komplexer Strukturen. Die Modellierung verläuft dabei in drei Stufen:

- Zerlegung der Struktur in möglichst gleichartige Elemente (Balken, Scheiben, Platten, Schalen), so dass durch die Verschiebungsfreiheitsgrade in den Knoten alle geometrischen Rand- und Übergangsbedingungen erfüllt werden können.
- Untersuchung der einzelnen Elemente, wobei das Verhalten ausschließlich durch die Verschiebung in den Knotenpunkten beschrieben wird.
- Zusammenfassung der Elemente zum Gesamtsystem.

In Bild 2-22 ist beispielhaft die netzartige Struktur des mit der FEM-Methode model-

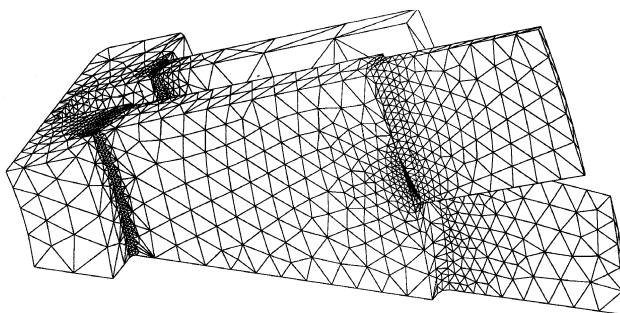


Bild 2-22: FEM Struktur eines Schaufelsegmentes in einem Turboverdichter

lierten Bauteils eines Finite - Elemente - Systems (FES) dargestellt. Solche FES sind zur Untersuchung statischer Lastfälle und dynamischer Vorgänge mit kleinen Bewegungen, relativ zu einem Inertialsystem, sehr gut geeignet. Bei Bewegungssystemen mit großen Starrkörperbewegungen, wie etwa bei Industrierobotern, sind diesen Führungsbewegungen kleine elastische

Bewegungen überlagert. Die Behandlung beider Bewegungstypen innerhalb eines Modells bereitet aber Schwierigkeiten. Daher werden in solchen Fällen eher Mehrkörpermodelle vorgezogen, wobei zur Dimensionierung einzelner Bauteile des Systems wieder sehr gut die FEM-Methode herangezogen werden kann.

2.5 Modellierung elektrischer Komponenten

Bei elektrischen Systemen scheint auf den ersten Blick die Modellbildung einfacher vorstatten zu gehen, als bei mechanischen Systemen. Betrachtet man den elektrischen Vierpol in Bild 2-23 a) aus zwei Widerständen, einem Kondensator und einer Spule, so entspricht diese Darstellung exakt den körperlich vorhandenen Bauteilen und ihren Verbindungen [2.5]. Die Bauteile selber können durch einfache bekannte elektrische Grundgleichungen beschrieben werden. Man darf sich aber nicht darüber täuschen lassen, dass auch diese Darstellung nicht einfach ein Lageplan (Schaltplan) der elektrischen Komponenten ist, sondern ein Modell des realen Systems. Dieses Modell hat nämlich nur Gültigkeit, solange Eingangssignale in das System niedrige Frequenz (0 - 10 MHz) besitzen. Bei mittleren Frequenzen (10 - 1000 MHz) kann ein aussagefähiges System nicht mehr die Einflüsse gewisser Eigenschaften der Bauteile und vor allem der Verbindungen vernachlässigen. So besitzen die Drahtverbindungen Koppelkapazitäten und Leitungsinduktivitäten, die Spule Windungskapazitäten und der Kondensator dielektrische Verluste oder auch Eigeninduktivität. Ein gültiges Modell des gleichen Vierpols muss daher wie in Bild 2-23 b) dargestellt aussehen. Hier sind die unerwünschten, parasitären Eigenschaften elektrischer Bauelemente als zusätzliche parallel und in Reihe geschaltete konzentrierte Bauelemente eingezeichnet. Die daraus folgenden Modellgleichungen sind entsprechend komplizierter.

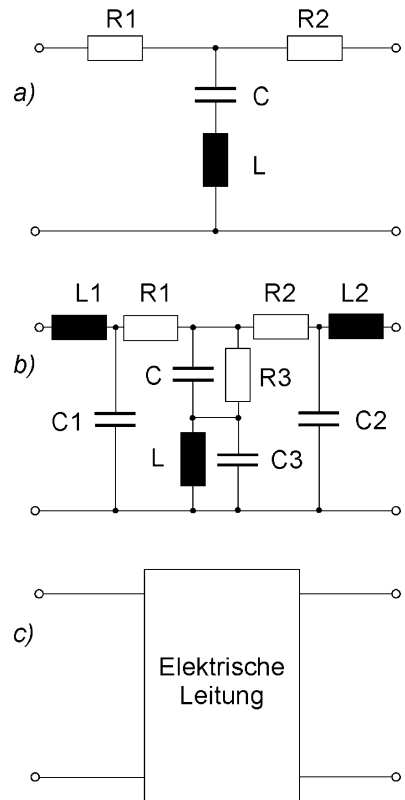


Bild 2-23: Verschiedene Modelle eines elektrischen Vierpols
a) für niedrige
b) für mittlere
c) für hohe Frequenzen

Für sehr hohe Frequenzen (> 1 GHz) reicht dann ein solches Modell aus konzentrierten Elementen auch nicht mehr aus. Der Wechselstromwiderstand der konzentrierten Bauelemente ist sehr klein und als Modell muss ein Leitungsstück genommen werden (Bild 2-23 c)).

Für die Modellbildung mechatronischer Systeme sind vor allem die elektromotorischen Antriebe von Bedeutung. Bei ihrer Modellbildung kommt wieder das Zusam-

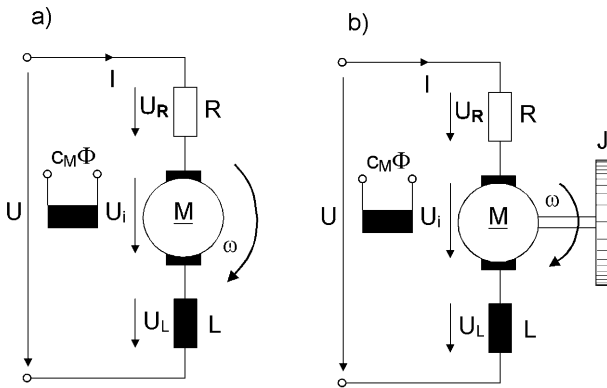


Bild 2-24: Ersatzmodell eines fremderregten Gleichstrommotors a) ohne und b) unter Berücksichtigung des Massenträgheitsmomentes des Ankers

menwirken mechanischer und elektrischer Einflüsse zum Tragen, da das dynamische Verhalten des Motors nur durch Kenntnisse der Gesetze der Kinetik und der Elektrodynamik erklärt werden kann. Eine häufig in mechatronischen Systemen, wie beispielsweise Industrierobotern, verwendete Antriebsmaschine ist der Gleichstrommotor. Bei dieser elektrischen Maschine wird sowohl das Erregungsfeld im Stator als auch das Magnetfeld im Rotor durch Gleichstrom oder durch einen Permanentmagnet

erzeugt. Daher muss zur Erzeugung einer fortschreitenden Drehung die Ankerspannung durch einen elektromechanischen Schalter, den sog. **Kommutator**, synchron mit der Drehung des Ankers ständig umgepolt werden.

Das einfachste elektrische Ersatzschaltbild für den Gleichstrommotor ist in Bild 2-24 a) dargestellt. Es benutzt für die Beschreibung des elektrischen Verhaltens die Ankerinduktivität L , den Ankerwiderstand R und die elektromotorische Kraft (EMK) U_i als interne Spannungsquelle. Energiespeicher im System werden dabei vernachlässigt. So wird der rotierende Anker, der eigentlich kinetische Energie speichert, als masselos angesehen und die Speicherfähigkeit der Induktivität für elektrische Energie nicht berücksichtigt. Verluste in realen Motoren wie Reibung, Eisenverluste, Leitungswiderstände werden ebenso vernachlässigt. Dieses einfache Modell liefert nur dann eine Aussage über das dynamische Verhalten des Systems, wenn man ein äußeres Lastmoment, das am Rotor angreift, annimmt. Für die Klemmenspannung U am Motor gilt mit $U_i = c_M \cdot \Phi \cdot \omega$ und den bekannten Zusammenhängen zwischen Strom und Spannung an ohmschen und induktiven Widerständen aufgrund des 1. Kirchhoffschen Gesetzes folgende Differentialgleichung:

$$U = U_L + U_R + U_i = L \cdot \dot{I} + R \cdot I + c_M \cdot \Phi \cdot \omega \quad (2.4)$$

Daraus folgt die sog. Drehzahlgleichung, mit der man die Drehzahl des Motors bestimmen kann:

$$\omega = \frac{U - R \cdot I - L \cdot \dot{I}}{c_M \cdot \Phi} \quad (2.5)$$

wobei c_M eine von der Motorbauart abhängige Konstante und Φ der magnetische Fluss im Erregerfeld sind. Nach dieser Gleichung hängt die Winkelgeschwindigkeit des Motors von der Klemmenspannung U , vom magnetischen Fluss Φ und von der Belastung des Motors ab. Diese Abhängigkeit ergibt sich über den Motorstrom I , der das vom Motor gelieferte innere Moment bestimmt:

$$M_i = c_M \cdot \Phi \cdot I \quad (2.6)$$

Dieses steht zur Beschleunigung der äußeren Last zur Verfügung.

Ein Modell zur Beschreibung der Eigendynamik des Motors liefert die Berücksichtigung des Trägheitsmomentes des Ankers wie in Bild 2-24 b). Auch ohne äußeres Lastmoment ist jetzt sein dynamisches Drehzahlverhalten beschreibbar, das innere Motormoment liefert das Beschleunigungsmoment zur Beschleunigung des Ankers. Für dieses vom Motor gelieferte Beschleunigungsmoment gilt:

$$M_B = J \cdot \dot{\omega} = M_i - M_L \quad (2.7)$$

mit dem Trägheitsmoment J der rotierenden Massen (Anker, Welle). Das innere Motormoment und das Beschleunigungsmoment sind gleich, wenn an der Motorwelle kein äußeres Lastmoment angreift ($M_L = 0 \Rightarrow M_i = J \cdot \dot{\omega}$). Damit kann man aus Gl. (2.6) die Werte von I und durch Bildung der zeitlichen Ableitung von $\dot{I}$ bestimmen. Eingesetzt in Gl. (2.4) erhält man die Differentialgleichung dieses Modells für sprungförmige Änderungen der Ankerspannung von ΔU :

$$\frac{J \cdot L}{\Phi \cdot c_M} \Delta \ddot{\omega} + \frac{J \cdot R}{\Phi \cdot c_M} \Delta \dot{\omega} + \Phi \cdot c_M \cdot \Delta \omega = \Delta U \quad (2.8)$$

Dies ist die Gleichung eines Systems 2. Ordnung, wie wir es schon in Abschnitt 1.4 kennengelernt haben. Bei sprungförmigen Änderungen der Ankerspannung von ΔU folgt dieses System mit Verzögerung, wobei abhängig vom Dämpfungsmaß D die Winkelgeschwindigkeit ω sich dem Endwert mit mehr oder weniger starkem Überschwingen nähert. Problematisch wird dieses Modell zur Erfassung des eingeschwungenen Zustandes, da hier das Beschleunigungsmoment Null wird. Dies würde bedeuten, dass auch das Motormoment Null und damit der Strom durch den Anker Null werden müsste.

Dieses Problem in der Modellaussage kann man beheben, wenn man beispielsweise die Lagerreibung im Motor mitberücksichtigt, die man als proportional zur Winkelgeschwindigkeit annehmen kann. Das Reibmoment $M_R = d \cdot \omega$ vermindert das zur Verfügung stehende Beschleunigungsmoment, so dass gilt:

$$M_B = M_i - M_R - M_L \quad (2.9)$$

Damit lautet die zugehörige Differentialgleichung:

$$\frac{J \cdot L}{\Phi \cdot c_M} \Delta \ddot{\omega} + \frac{(J \cdot R + d \cdot L)}{\Phi \cdot c_M} \Delta \dot{\omega} + (\Phi \cdot c_M + d \cdot R) \Delta \omega = \Delta U \quad (2.10)$$

In diesem Modell kann nun auch der Zustand gleichbleibender Drehzahl beschrieben werden, da bei verschwindendem Beschleunigungsmoment das Motormoment immer noch das Reibmoment aufbringen muss und daher der Ankerstrom nicht Null werden kann.

Bei der grafischen Darstellung des Modells eines solchen elektromechanischen Systems wie im Bild 2-24 gezeigt, oder auch bei den Modelldarstellungen als Differentialgleichung, werden die Zusammenhänge und Einflüsse der verschiedenen

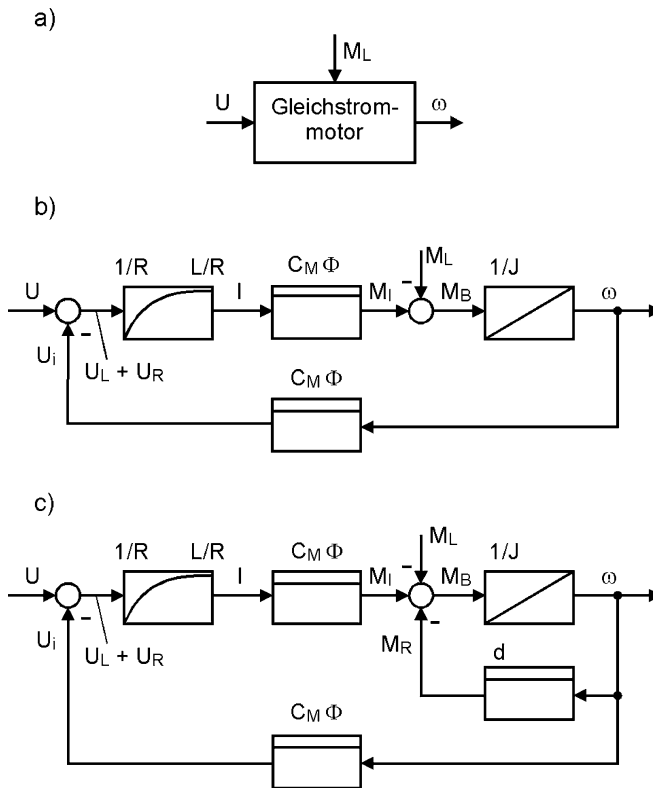


Bild 2-25: Modellbildung des Gleichstrommotors als
 Blockschaltbild a) grob b) detailliert nach Gl.
 (2.8) c) detailliert nach Gl. (2.10)

Größen auf das Systemverhalten nicht besonders deutlich. Sehr viel anschaulicher wird das Modellgeschehen, wenn man die vor allem in der Regelungstechnik gebräuchliche Darstellung als Blockschaltbild wählt, wie sie im Kapitel 1.4 bereits kurz vorgestellt wurde. Solche Blockschaltbilder kann man in verschiedenen Detaillierungsgraden verwenden. Betrachtet man das ganze System erst einmal nur als einen Block, so erhält man eine Darstellung entsprechend Bild 2-25 a). Der Block besitzt als Eingangsgröße die Ankerspannung U und als Ausgangsgröße die Winkelgeschwindigkeit ω . Das gesamte Systemverhalten ist im Block dazwischen enthalten und wird durch eine der Modellgleichungen, die Verzögerungsglieder zweiter Ord-

nung darstellen, mathematisch beschrieben. Das dynamische Verhalten des Systems nach der Modellgleichung (2.4) wird aber, wie bereits oben festgestellt wurde, nur bei Vorhandensein eines äußeren Lastmomentes verständlich (Pfeil von oben).

Man kann nun das Systemverhalten nach Gl. (2.8) auch detaillierter darstellen, indem man die Zusammenhänge der einzelnen Gleichungen (2.4), (2.6) und (2.7) als eigene Blöcke darstellt (Bild 2-25 b)). Kreise mit Pfeilen, die hinein- und hinausführen, sind sog. Summationsstellen, in denen Eingangsgrößen (Pfeile, die in die Summationsstelle hineinführen) zu Ausgangsgrößen (Pfeile, die aus der Summationsstelle hinausführen) aufsummiert werden. So bedeutet die Summationsstelle auf der linken Bildseite, dass die Differenz von U und U_i als Ergebnis die Summe von U_L und U_R ergibt. Dies ist gleichbedeutend mit der Gl. (2.4). Die rechte Summationsstelle steht im Bild 2-25 b) für die Gl. (2.7) und in Bild 2-25 c) für die Gl. (2.9). In den Einzelblöcken wird symbolisch durch einen Graphen das zugehörige Zeitverhalten dargestellt. So gibt der auf der linken Seite in der oberen Reihe stehende Block des Bildes 2-25 b) das Übertragungsverhalten der Ankerinduktivität ($L \cdot \dot{I} + R \cdot I$) aus Gl. (2.4) wieder, der zweite Block von links den Zusammenhang der Gl. (2.6), der dritte Block den Zusammenhang der Gl. (2.7) und der Block in der unteren Reihe den Term für die elektromotorische Kraft ($c_M \cdot \Phi \cdot \omega$) aus Gl. (2.4). Das Zusammenwirken und die Einflüsse aller Bereiche des Modells werden durch diese

Darstellung sehr deutlich. Im Bild 2-25 c) kommt noch ein weiterer Block für das Reibverhalten entsprechend Gl. (2.10) dazu.

Da die Methoden der Systembeschreibung, der Modellbildung und der mathematischen Beschreibung und Untersuchung schon seit langem von der **Regelungstechnik** benutzt werden, wird dieses Thema dort (Kap. 7.3) nochmals aufgenommen und vertieft.

3 Dynamik mechanischer Systeme

Hat man ein Modell für das mechatronische System aufgestellt, so muss man die in ihm ablaufenden Bewegungen mathematisch beschreiben. Dazu gehören die Beschreibung der Lage und der Orientierung der einzelnen Körper zueinander und die Bestimmung der Geschwindigkeiten und Beschleunigungen. Die Kinematik beschäftigt sich mit den Bewegungen, ohne die Ursachen der Bewegungen, nämlich Kräfte oder Momente, in die Betrachtung mit einzubeziehen. Die Kinetik wiederum untersucht die Wechselwirkung zwischen den Ursachen und den Bewegungen von Massen. Der Begriff Dynamik ist noch umfassender und ist manchmal mit dem Newton'schen Grundgesetz gleichgesetzt worden [2.4]:

$$m \cdot \vec{a} = \sum \vec{F}$$

In dieser Gleichung präsentiert die Kraft $\vec{F}$ die **Kinetik** des Systems, die Masse m die **Trägheit** und die Beschleunigung $\vec{a}$ die **Kinematik**.

Wie bereits oben festgestellt wurde, lassen sich viele mechatronische Systeme als Mehrkörpersysteme (MKS) modellieren. Bevor aber auf die Kinematik von MKS eingegangen werden soll, werden einige grundlegende Begriffe und Gesetzmäßigkeiten über die Kinematik des Massenpunktes und des starren Körpers wiederholt. Anschließend werden dann die Kinetik und die Aufstellung der Bewegungsgleichungen solcher Systeme behandelt.

Dieses Kapitel des Buches behandelt wie auch die weiteren Kapitel Thematiken der Ingenieurwissenschaften, über die es bereits viel Literatur gibt und können daher in einer Einführung in die Mechatronik nicht in der Ausführlichkeit behandelt werden, wie in solchen Büchern, die sich ausschließlich einem Thema oder einer eng gefassten Themenpalette widmen. Zur Vertiefung des jeweiligen Themas sollte daher auf die zitierte Spezialliteratur zurückgegriffen werden, der vielfach auch Anregungen zu den behandelten Themen entnommen wurden.

3.1 Kinematik des Massenpunktes

Die Bewegung eines realen Systems ist erst dann vollständig beschrieben, wenn der Ortsvektor $\vec{r}(t)$ eines jeden zum System gehörenden Teiles bestimmt ist. Für die meisten technischen Systeme ist diese Aufgabe nur näherungsweise zu lösen. Wie wir gehört haben, ersetzt man daher die tatsächlichen Verhältnisse durch Modellvorstellungen. Das einfachste mechanische Modell eines Körpers ist der Massenpunkt. Der Begriff Massenpunkt bezieht sich auf ein kleines Objekt und zwar so klein, dass seine Abmessungen infinitesimal klein und daher zu vernachlässigen sind. Von allen den Körper kennzeichnenden Eigenschaften verbleibt einzig die Gesamtmasse, die in einem Punkt, meist dem Massenmittelpunkt, konzentriert gedacht wird. Ausgehend von diesem Modell eines festen Körpers baute Newton seine Punktmechanik auf, für die das Gesetz gilt:

$$\vec{F} = m \cdot \frac{d^2 \vec{r}}{dt^2} = m \cdot \vec{\ddot{r}} \quad \text{Kraft} = \text{Masse} \times \text{Beschleunigung.}$$

Dabei ist $\vec{F}$ der resultierende Vektor aller am Körper angreifenden Kräfte und ist $\vec{r}$ die Beschleunigung. Die Beschleunigung ist die zweite Ableitung des Ortsvektors nach der Zeit [3.1], [3.2].

Um die Bewegung eines Körpers im Raum zu beschreiben, benötigt man ein Bezugssystem. In der Strukturdynamik werden dazu meist kartesische Koordinaten verwendet, d. h. drei senkrecht aufeinander stehende, rechtsorientierte Koordinatenachsen. Ein Punkt hat im Raum drei Freiheitsgrade und jedem dieser Freiheitsgrade entspricht eine lineare Verschiebung in Richtung einer der Achsen des Koordinatensystems. Dies ist ein weiterer Vorteil des Massenpunktmodells, weil allgemein ein starrer Körper sechs Freiheitsgrade im Raum hat, die drei bereits erwähnten und drei weitere in Form von Drehungen um die Koordinatenachsen. Die aktuelle Position eines Massenpunktes zu einem Zeitpunkt t ist dann durch folgenden Ortsvektor gegeben:

$$\vec{r}(t) = x(t) \cdot \vec{e}_x + y(t) \cdot \vec{e}_y + z(t) \cdot \vec{e}_z \quad .$$

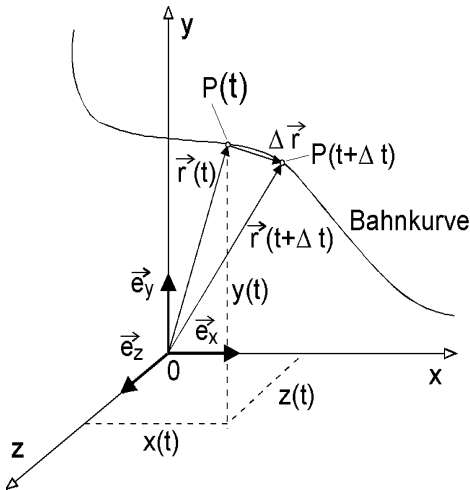


Bild 3-1: Bewegung eines Massenpunktes P auf einer Bahnkurve innerhalb eines kartesischen Koordinatensystems

Die aktuelle Position auf der Bahnkurve, die der Massenpunkt im Raum beschreibt, ergibt sich zu einem Zeitpunkt $t + \Delta t$ als der Ortsvektor $\vec{r}(t + \Delta t)$, der gegenüber $\vec{r}(t)$ einen Zuwachs um $\Delta \vec{r}(t)$ aufweist (Bild 3-1). $\vec{e}_x, \vec{e}_y, \vec{e}_z$ sind dabei drei senkrecht aufeinander stehende Einheitsvektoren. Bei dieser Betrachtung wird das Koordinatensystem als im Raum ruhend oder geradlinig gleichförmig bewegt angenommen, es handelt sich also um ein **Inertialsystem**. In einem **Inertialsystem** sind die Einheitsvektoren von der Zeit unabhängig, deren Ableitungen nach der Zeit sind gleich Null.

In Matrixschreibweise lässt sich der Ortsvektor ausdrücken als

$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} .$$

Durch Ableitung des Ortsvektors nach der Zeit erhält man die Geschwindigkeit des Massenpunktes P auf der Bahnkurve

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \dot{\vec{r}} = \dot{x}(t) \cdot \vec{e}_x + \dot{y}(t) \cdot \vec{e}_y + \dot{z}(t) \cdot \vec{e}_z = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{pmatrix}$$

und durch erneute Ableitung die Beschleunigung

$$\vec{a}(t) = \frac{d^2 \vec{r}}{dt^2} = \ddot{\vec{r}} = \ddot{x}(t) \cdot \vec{e}_x + \ddot{y}(t) \cdot \vec{e}_y + \ddot{z}(t) \cdot \vec{e}_z = \begin{pmatrix} \ddot{x}(t) \\ \ddot{y}(t) \\ \ddot{z}(t) \end{pmatrix}$$

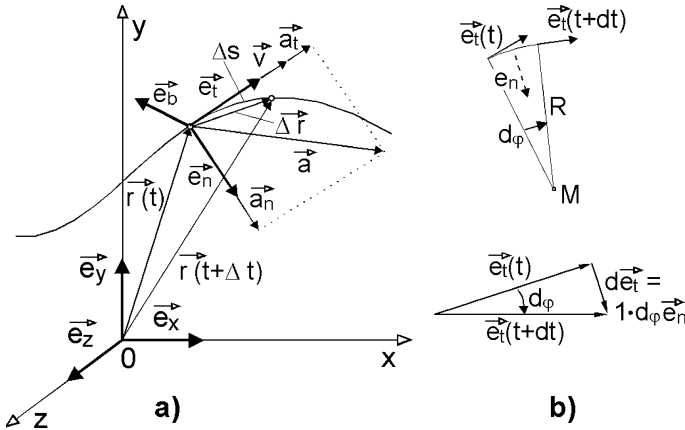


Bild 3-2: a) Mitbewegtes, natürliches Koordinatensystem
b) Differentiation des Tangenteneinheitsvektors

Der Geschwindigkeitsvektor tangiert stets die Bahnkurve und kann auch in einem sog. natürlichen Koordinatensystem dargestellt werden (Bild 3-2 a)). Dieses **natürliche Koordinatensystem** ist ein den Massenpunkt begleitendes Dreiein mit den Koordinatenachsen t , n , b , die die sog. **Schmiegungsebene** festlegen. Hierbei ist t die **Tangentenrichtung** in der Schmiegungsebene, n die **Normalenrichtung** in der Schmiegungsebene und b die **Binormalenrichtung**, die senkrecht zu t und n ist. In diesem System gilt

$$\vec{v}(t) = \frac{d \vec{r}(t)}{dt} = \frac{d \vec{r}}{ds} \cdot \frac{ds}{dt} = \vec{e}_t \cdot v$$

Dabei ist $\vec{e}_t$ der Tangenteneinheitsvektor. Der Betrag der Geschwindigkeit ist

$$|\vec{v}| = v = \frac{ds}{dt} = \dot{s} = \sqrt{v_x^2 + v_y^2 + v_z^2} = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$$

In natürlichen Koordinaten ausgedrückt, beträgt die Beschleunigung

$$\vec{a}(t) = \frac{d}{dt} (\vec{e}_t \cdot v) = \vec{e}_t \cdot \frac{dv}{dt} + \frac{d \vec{e}_t}{dt} \cdot v$$

Aus Bild 3-2 b) kann man entnehmen

$$\frac{d \vec{e}_t}{dt} = \frac{d \vec{e}_t}{ds} \cdot \frac{ds}{dt} = \frac{d \vec{e}_t}{ds} \cdot v = \frac{d\phi}{ds} \cdot \vec{e}_n \cdot v = \frac{1}{R} \cdot \vec{e}_n \cdot v$$

woraus folgt

$$\vec{a}(t) = \dot{v} \cdot \vec{e}_t + \frac{v}{R} \cdot \vec{e}_n = \vec{a}_t + \vec{a}_n$$

Der Beschleunigungsvektor $\vec{a}(t)$ liegt immer in der Schmiegungsebene (Bild 3-2 b)).

Er besitzt keine Komponente in der Binormalenrichtung und seine Komponenten in Tangential- und Normalenrichtung heißen **Tangential-** und **Normalbeschleunigung**:

$$a_t = \frac{dv}{dt} = \dot{v}(t) = \ddot{s}(t) \quad \text{und} \quad a_n = \frac{v^2}{R}$$

wobei R der **Krümmungsradius** der Bahnkurve ist. Die Normalbeschleunigung ist stets zum **Krümmungsmittelpunkt** M gerichtet, also immer eine Zentripetalbeschleunigung. Für den Betrag der Beschleunigung gilt:

$$|\vec{a}| = a = \sqrt{a_x^2 + a_y^2 + a_z^2} = \sqrt{a_t^2 + a_n^2} \quad .$$

3.2 Kinematik des starren Körpers

Ein **starrer Körper** besteht aus einer großen Anzahl von Massenpunkten, die ihre relative Lage zueinander nicht ändern, d. h. es treten keine Verformungen des Körpers auf. Dies ist natürlich auch ein fiktives Gebilde, das aber als Modell zur Untersuchung von Bewegungsvorgängen in der Regel völlig ausreicht. Um die Gesetzmäßigkeiten der Bewegung eines Starrkörpers zu beschreiben, soll zuerst nur die ebene Bewegung eines solchen Körpers betrachtet werden.

3.2.1 Die ebene Bewegung des starren Körpers

Um die ebene Bewegung des starren Körpers zu beschreiben, würde es genügen, wenn man die Bewegung zweier Punkte $P1$ und $P2$ des Starrkörpers beschreibt. Dazu wären für jeden Punkt zwei Koordinatenangaben, d. h. insgesamt 4 Koordinatenangaben erforderlich. Da aber der Abstand d dieser beiden Punkte stets konstant bleibt und bekannt ist, reichen die beiden Koordinaten des Punktes $P1$ und eine Koordinate des Punktes $P2$, oder die Angabe der beiden Koordinaten des Punktes $P1$ und eine zusätzliche Winkelkoordinate φ aus (Bild 3-3). Kann sich daher der Körper in der Ebene frei bewegen, ohne durch äußere Bindungen behindert zu werden, so besitzt er drei **Freiheitsgrade**. Dies können je eine Translation (geradlinige Verschiebung) in Richtung der beiden Koordinatenachsen sein, wodurch sich jeweils die x - bzw. y -Koordinate des Punktes $P1$ ändert und eine Rotation um einen Punkt des Körpers, wodurch sich der Winkel φ ändert. Durch zusätzliche Bindungen (z. B. Gelenke, Führungen) kann die Anzahl der Freiheitsgrade eingeschränkt werden.

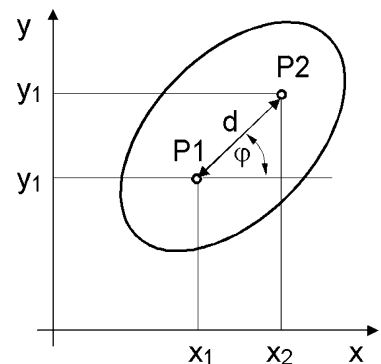


Bild 3-3: Festlegung der Lage eines ebenen Starrkörpers durch zwei Punkte

Das bedeutet, dass sich eine beliebige ebene Bewegung eines Starrkörpers durch eine Translation eines Punktes des Körpers und eine Rotation des Körpers um diesen Punkt beschreiben lässt. Zu die-

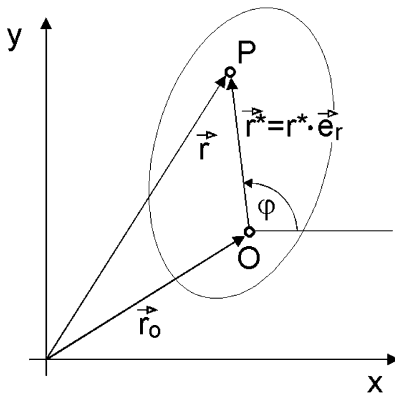


Bild 3-4: Beschreibung der Bewegung eines ebenen Starrkörpers durch Translation und Rotation

sem Zweck wird in Bild 3-4 ein ausgewählter Punkt O in einem raumfesten, kartesischen Koordinatensystem mit dem Ortsvektor $\vec{r}_0$ verfolgt, der die Translation beschreibt. Die Rotation wird als Bewegung eines beliebigen Punktes P des Starrkörpers um den Punkt O beschrieben. Dazu werden **Polarkoordinaten** verwendet, wobei der Ursprung dieses Koordinatensystems in O liegt. Die Lage des Punktes P in Bezug auf O wird durch den Vektor $\vec{r}^*$ und den Winkel φ beschrieben. Weil für den starren Körper der Abstand r^* der Punkte O und P konstant bleibt, ist in dem Vektor $\vec{r}^*$, der die Lage von P relativ zu O beschreibt, nur der Einheitsvektor $\vec{e}_r$ des Polarkoordinatensystems von der Zeit abhängig. Für den Punkt P , dessen Lage im ortsfesten Koordinatensystem durch

$$\vec{r}(t) = \vec{r}_0(t) + \vec{r}^*(t) = \vec{r}_0(t) + r^* \cdot \vec{e}_r(t) \quad (3.1)$$

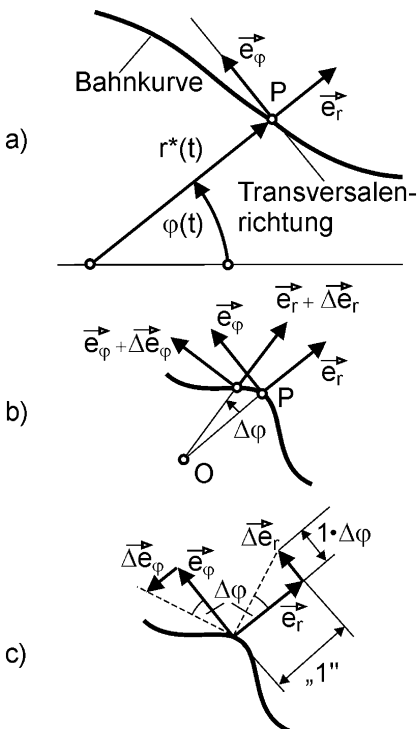


Bild 3-5: Beschreibung einer Bahnkurve durch Polarkoordinaten
a) Ortskurve b) Zeitliche Ableitung des Ortsvektors c) Änderung der Einheitsvektoren

beschrieben wird, kann der Geschwindigkeits- und der Beschleunigungsvektor durch ein- bzw. zweimaliges Ableiten der Gl. (3.1) nach der Zeit ermittelt werden.

Zur Herleitung der Ableitung von $\vec{e}_r$ nach der Zeit kann man die Betrachtung aus Bild 3-5 heranziehen. Hier wird die Lage eines auf einer Bahn bewegten Punktes P durch Polarkoordinaten beschrieben. Sein Abstand vom Koordinatenursprung O beträgt dabei $r^*(t)$, die Winkellage zu einer Bezugsgeraden beträgt $\varphi(t)$. Mit dem Einheitsvektor $\vec{e}_r$, der stets die Richtung des Ortsvektors $\vec{r}^*(t)$ hat, kann dieser dann geschrieben werden

$$\vec{r}^*(t) = r^* \cdot \vec{e}_r(t) \quad (3.2)$$

Die Koordinate φ steckt dabei in dem Einheitsvektor $\vec{e}_r$. Da der Geschwindigkeits- und der Beschleunigungsvektor des Punktes P nicht die Richtung des Einheitsvektors besitzen, muss zu ihrer Beschreibung noch der Einheitsvektor $\vec{e}_\varphi$ eingeführt werden. Zur Herleitung der zeitlichen Ableitung des Einheitsvektors wird in Bild 3-5 b) die Lage des Punktes P in zwei eng benachbarten Positionen betrachtet, die sich durch

Änderung des Winkels φ um $\Delta\varphi$ unterscheiden. Dadurch ändert sich $\vec{e}_r$ um $\Delta\vec{e}_r$, ein Vektor in Richtung $\vec{e}_\varphi$, der senkrecht auf $\vec{e}_r$ steht. Die Länge von $\Delta\vec{e}_r$ kann für sehr kleine Winkel $\Delta\varphi$ gleich $\Delta\varphi$ gesetzt werden ($\tan \Delta\varphi \approx \Delta\varphi$). Daher gilt für die Änderungen der beiden aufeinander senkrecht stehenden Einheitsvektoren (Bild 3-5 c))

$$\Delta\vec{e}_r = \Delta\varphi \cdot \vec{e}_\varphi$$

$$\Delta\vec{e}_\varphi = -\Delta\varphi \cdot \vec{e}_r$$

Dividiert man diese Änderungen durch $\Delta\varphi$ und macht den Grenzübergang $\Delta\varphi \rightarrow 0$, so erhält man die Ableitungen der Einheitsvektoren nach φ und nach der Zeit:

$$\frac{d\vec{e}_r}{d\varphi} = \vec{e}_\varphi \quad \frac{d\vec{e}_\varphi}{d\varphi} = -\vec{e}_r \quad (3.3)$$

$$\frac{d\vec{e}_r}{dt} = \frac{d\vec{e}_r}{d\varphi} \cdot \frac{d\varphi}{dt} = \vec{e}_\varphi \cdot \dot{\varphi} \quad \frac{d\vec{e}_\varphi}{dt} = \frac{d\vec{e}_\varphi}{d\varphi} \cdot \frac{d\varphi}{dt} = -\vec{e}_r \cdot \dot{\varphi} \quad (3.4)$$

Unter Verwendung der Gleichungen (3.3) und (3.4) ergeben sich dann die zeitlichen Ableitungen des Ortsvektors, also die Geschwindigkeit und die Beschleunigung, nach Gl. 3.2 in Polarkoordinaten als

$$\vec{v}(t) = \frac{d\vec{r}^*}{dt} = \dot{r}^* \cdot \vec{e}_r + r^* \cdot \dot{\varphi} \cdot \vec{e}_\varphi \quad (3.5)$$

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = (\ddot{r}^* - \dot{\varphi}^2 \cdot r^*) \vec{e}_r + (r^* \cdot \ddot{\varphi} + 2\dot{r}^* \cdot \dot{\varphi}) \vec{e}_\varphi \quad (3.6)$$

Mit diesen Beziehungen können nun die zeitlichen Ableitungen des Ortsvektors im ortsfesten Koordinatensystem nach Gl. (3.1) gebildet werden:

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \frac{d\vec{r}_0}{dt} + \frac{d\vec{r}^*}{dt} = \frac{d\vec{r}_0}{dt} + r^* \cdot \frac{d\vec{e}_r}{dt} = \frac{d\vec{r}_0}{dt} + r^* \cdot \dot{\varphi} \cdot \vec{e}_\varphi \quad (3.7)$$

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}_0}{dt^2} + r^* \cdot \ddot{\varphi} \cdot \vec{e}_\varphi + r^* \cdot \dot{\varphi} \cdot \frac{d\vec{e}_\varphi}{dt} = \frac{d^2\vec{r}_0}{dt^2} + r^* \cdot \ddot{\varphi} \cdot \vec{e}_\varphi - r^* \cdot \dot{\varphi}^2 \cdot \vec{e}_r \quad (3.8)$$

Unter Verwendung dieser Beziehungen kann man nun die Geschwindigkeit und die Beschleunigung in natürlichen Koordinaten herleiten. Da die Ableitung des Winkels φ nach der Zeit die Winkelgeschwindigkeit ω ist, mit der der Punkt P sich um O bewegt und da gilt (Bild 3-6) $\vec{e}_\varphi = \vec{e}_t$, sowie $\vec{e}_r = -\vec{e}_n$, erhält man für die Gesamtgeschwindigkeit und die Gesamtbeschleunigung eines beliebigen Punktes P des starren Körpers:

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \frac{d\vec{r}_0}{dt} + r^* \cdot \omega \cdot \vec{e}_t = \vec{v}_{trans} + \vec{v}_{rot} \quad (3.9)$$

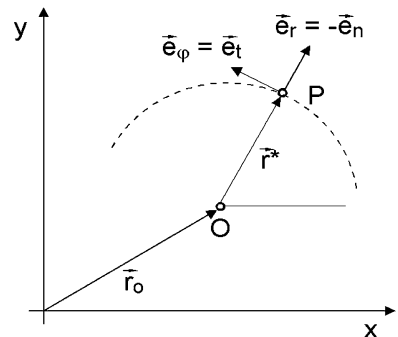


Bild 3-6: Zusammenhang zwischen Polarkoordinaten und natürlichen Koordinaten

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}_0}{dt^2} + r^* \cdot \dot{\omega} \cdot \vec{e}_t + r^* \cdot \omega^2 \cdot \vec{e}_n = \vec{a}_{trans} + \underbrace{\vec{a}_{t,rot} + \vec{a}_{n,rot}}_{\vec{a}_{rot}} \quad (3.10)$$

Zusammenfassend gilt daher in nicht vektorieller Schreibweise [1.5]:

- Die **ebene Bewegung des starren Körpers** wird durch die **Translation** eines beliebigen Bezugspunktes O beschrieben, der eine **Rotation** um den Bezugspunkt überlagert ist. Die Translation des Bezugspunktes kann auf einer beliebigen Bahnkurve erfolgen und wird nach den Regeln der **Kinematik des Massenpunktes** beschrieben. Die Rotation um den Bezugspunkt erfolgt mit der Winkelgeschwindigkeit ω , die in der Regel zeitlich veränderlich ist. Jeder Punkt des starren Körpers führt dabei eine Kreisbewegung um den Bezugspunkt aus.
- Die **Geschwindigkeit eines beliebigen Punktes P** eines starren Körpers setzt sich aus einem translatorischen Anteil v_{trans} , der Bahngeschwindigkeit des Bezugspunktes O und dem rotatorischen Anteil $v_{rot} = r^* \cdot \omega$, der Bahngeschwindigkeit der Kreisbewegung von P auf einem Kreis mit dem Radius r^* um den Bezugspunkt O , zusammen. Die beiden Geschwindigkeitsanteile lassen sich vektoriell zur Geschwindigkeit v überlagern.
- Die **Beschleunigung eines beliebigen Punktes P** setzt sich aus einem translatorischen Anteil a_{trans} und einem rotatorischen Anteil a_{rot} zusammen. Dabei ist die Gesamtbeschleunigung des Bezugspunktes O , die sich im allgemeinen aus der tangentialen Bahnbeschleunigung des Bezugspunktes $a_{t,trans}$ und seiner Normalbeschleunigung $a_{n,trans} = v_{trans}^2 / \rho$ (ρ : Krümmungsradius der Kreisbewegung des Punktes P) zusammensetzt. Die Beschleunigung a_{rot} , die sich im allgemeinen aus den zwei Anteilen $a_{t,rot}$ und $a_{n,rot}$ zusammensetzt, ist die Beschleunigung der Kreisbewegung des Punktes P um den Bezugspunkt O . Dabei ist:

$$a_{t,rot} = r^* \cdot \dot{\omega} \quad \text{und} \quad a_{n,rot} = r^* \cdot \omega^2$$

- Im allgemeinen Fall ist daher die Gesamtbeschleunigung des Punktes P aus vier Anteilen vektoriell zusammengesetzt.

Führt der starre Körper ausschließlich eine translatorische Bewegung aus, so haben alle seine Punkte die Geschwindigkeit $\vec{v}$. Bei einer reinen Rotation mit der Winkelgeschwindigkeit $\vec{\omega}$ bleibt der Drehpunkt des Körpers in Ruhe, alle anderen Punkte bewegen sich auf Kreisbahnen um den Drehpunkt mit einer Bahngeschwindigkeit $\vec{v}_{rot} = \vec{r} \times \vec{\omega}$, d. h. die Geschwindigkeit nimmt mit dem Abstand zum Drehpunkt

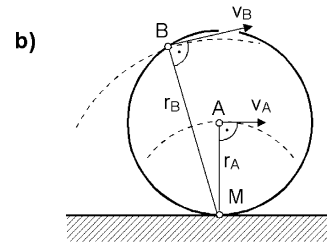
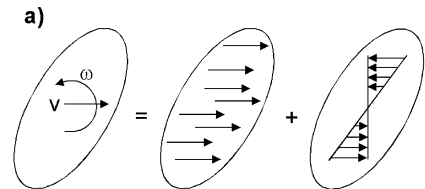


Bild 3-7:

Geschwindigkeitsverteilung beim beliebig bewegten Starrkörper a) Zusammensetzung aus Translation und Rotation b) Momentanpol

linear zu. Führt der Körper gleichzeitig eine Translation und eine Rotation aus, so können die Geschwindigkeiten aller Punkte durch Überlagerung der beiden Anteile ermittelt werden. Eine solche Überlagerung wird im Bild 3-7 a vorgenommen. Man sieht, dass bei vektorieller Addition der Rotations- und Translationsgeschwindigkeit die zugehörigen Vektoren in dem Teil des Körpers, der oberhalb des Drehpunktes liegt, entgegengesetzt orientiert und im unteren Teil gleichgerichtet sind. Es muss daher in der Ebene genau einen Punkt geben, für den die Vektorsumme der Geschwindigkeiten Null ergibt, das heißt, im Bezugssystem ist dieser Punkt momentan in Ruhe. Er wird daher auch **Momentanpol** genannt und mit M bezeichnet. Wenn dieser Punkt als Bezugspunkt für die Bewegung gewählt wird, entfällt dessen translatorischer Anteil und man kann den Geschwindigkeitszustand wie bei einer momentanen reinen Rotation um den Momentanpol beschreiben.

Betrachtet man beispielsweise das in Bild 3-7 b dargestellte, auf einer Unterlage abrollende Rad, so ist der Berührungspunkt zwischen Rad und Unterlage der Momentanpol M der Bewegung. Der Geschwindigkeitsvektor eines beliebigen Punktes B des Rades steht dann senkrecht auf der Verbindungslinie MB . Von einem abrollenden Rad ist meist die horizontale Geschwindigkeit der Radachse bekannt, die Geschwindigkeit des Punktes A . Die momentane Winkelgeschwindigkeit des Körpers kann nach dem oben gesagten als

$$\omega = v_A / r_A$$

geschrieben werden, da man den Punkt A als mit der Bahngeschwindigkeit v_A im Abstand r_A um den Momentanpol M rotierend betrachten kann. Da für jeden Punkt des Körpers das gleiche bezüglich M gelten muss, kann die Geschwindigkeit des Punktes B nach

$$v_B = r_B \cdot \omega$$

berechnet werden. Entsprechend diesem Beispiel gilt auch allgemein

$$\frac{v_A}{v_B} = \frac{r_A}{r_B}, \quad (3.11)$$

d. h. die Geschwindigkeiten zweier Punkte A und B eines starren Körpers verhalten sich wie ihre Abstände vom Momentanpol.

Zum Abschluss soll die Berechnung des Bewegungszustandes eines in der Ebene translatorisch und rotatorisch bewegten starren Körpers an einem Beispiel gezeigt werden. Der Bewegung eines Planetenrades um das Sonnenrad eines Planetengetriebes [1.5] (Bild 3-8) erscheint auf den ersten Blick als eine aus zwei Rotationen zusammengesetzte Bewegung. Da im oben gesagten aber die allgemeine Bewegung des starren Körpers in der Ebene als zusammengesetzte Bewegung aus der Translation

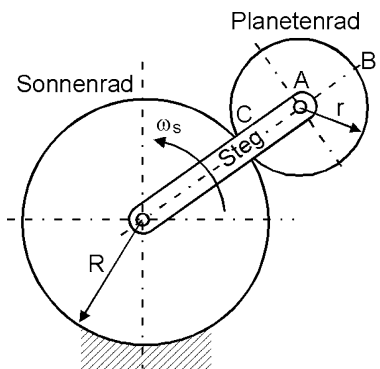


Bild 3-8: Sonnen- und Planetenrad eines Planetengetriebes

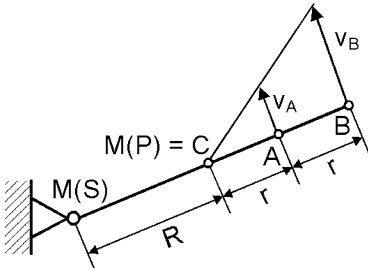


Bild 3-9: Berechnung der Geschwindigkeit als Rotation um den Momentanpol

eines Punktes und einer Rotation um diesen Punkt hergeleitet wurde, muss man im Beispiel die Bewegung des Steges (Punkt A) als Sonderfall der Translation (Krümmungsradius der Bahn ist konstant) betrachten. Das Planetenrad mit dem Radius r rollt auf einem feststehenden Sonnenrad mit dem Radius R ab. Der Steg dreht sich mit der konstanten Winkelgeschwindigkeit ω_S . Der Bewegungszustand des Planetenrades soll durch Angabe von Geschwindigkeit und Beschleunigung des Außenpunktes B und des Berührungspunktes mit dem Sonnenrad C ermittelt werden.

Infolge der Drehbewegung des Steges mit konstanter Winkelgeschwindigkeit bewegt sich der Punkt A, der zum Steg und zum Planetenrad gehört, mit der Geschwindigkeit

$$v_A = (R + r) \cdot \omega_S \quad .$$

Die Geschwindigkeit der Punkte B und C kann auf zwei verschiedenen Wegen ermittelt werden.

a) Der Momentanpol $M(P)$ des Planetenrades ist der Berührungspunkt C mit dem feststehenden Sonnenrad. Die Geschwindigkeit des Punktes A ist bekannt, weshalb nach Gl. (3.11) gilt (Bild 3-9):

$$v_B = 2 \cdot v_A = 2(R + r) \cdot \omega_S \quad , \quad v_C = 0 \quad .$$

Für die Winkelgeschwindigkeit des Planetenrades gilt:

$$\omega_P = \frac{v_A}{r} = \left(1 + \frac{R}{r}\right) \cdot \omega_S$$

b) Für die **Überlagerung einer Translation mit einer Rotation** nach Gl. (3.9) wird der Mittelpunkt A des Planetenrades als Bezugspunkt gewählt (Bild 3-10). Allen Punkten des Planetenrades wird die translatorische Geschwindigkeit v_A zugeordnet, und der Translation wird die Rotation um den Punkt A mit der Winkelgeschwindigkeit ω_P überlagert. Die Überlagerung liefert:

$$v_B = v_A + r \cdot \omega_P = 2(R + r) \cdot \omega_S \quad , \quad v_C = v_A - r \cdot \omega_P = 0 \quad .$$

Die Beschleunigungen werden durch Überlagerung nach Gl. (3.10) ermittelt. Da die Translation im Beispiel eine Kreisbewegung mit konstanter Winkelgeschwindigkeit ist, ist der translatorische Anteil eine Normalbeschleunigung; eine Tangentialbeschleunigung tritt nicht auf. Dieser Beschleunigungsanteil ist für alle Punkte des Planetenrades gleich. Auch die Rotation erfolgt mit konstanter Winkelgeschwindigkeit ω_P , so dass der rotatorische Anteil ebenfalls eine Normalbeschleunigung ohne Tangentialkomponente ist. Sie ist immer vom betrachteten Punkt zum Bezugspunkt A gerichtet und hängt vom Abstand des betrachteten Punktes vom Punkt A ab. Da die Beschleunigungsanteile von Translation und Rotation in den Punkten B und C

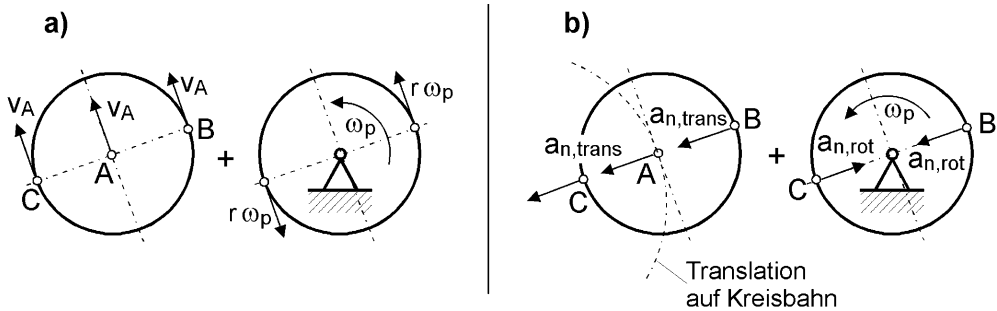


Bild 3-10: Bestimmung des Bewegungszustandes eines Starrkörpers durch Überlagerung von Translation und Rotation a) Geschwindigkeiten b) Beschleunigungen

gleich oder entgegengesetzt ausgerichtet sind, können sie skalar zusammengefasst werden:

$$a_B = a_{n,trans} + a_{n,rot} = (R + r) \cdot \omega_S^2 + r \cdot \omega_P^2 = (3R + 2r + \frac{R^2}{r}) \cdot \omega_S^2 \quad ,$$

$$a_C = a_{n,trans} - a_{n,rot} = (R + r) \cdot \omega_S^2 - r \cdot \omega_P^2 = -R \cdot (1 + \frac{R}{r}) \cdot \omega_S^2 \quad .$$

3.2.2 Ebene Relativbewegung eines Punktes

In vielen technischen Systemen tritt der Fall auf, dass sich auf dem starren Körper ein weiterer starrer Körper oder im einfachsten Fall ein Massenpunkt P befindet, der relativ zur Bewegung des Starrkörpers eine weitere Bewegung ausführt. Diese Bewegung erfolgt mit der Geschwindigkeit v_{rel} . Um den Bewegungszustand des Massenpunktes P , der sich relativ zum Starrkörper bewegt, welcher sich wiederum in einem festen Bezugssystem bewegt, beschreiben zu können, gibt es prinzipiell zwei Möglichkeiten:

a) Die Bewegung wird mit einem Ortsvektor bezüglich eines festen Koordinatensystems beschrieben. Dann lassen sich in Bezug auf das ruhende System Geschwindigkeit und Beschleunigung nach den Regeln der Kinematik des Massenpunktes durch Differenzieren ermitteln. Dieser Weg ist in vielen Fällen sehr aufwendig.

b) Die **Führungsbewegung** des starren Körpers bezüglich eines festen Koordinatensystems und die Relativbewegung werden gesondert betrachtet. Dann können die **Absolutgeschwindigkeit** und die **Absolutbeschleunigung** bezogen auf das ruhende System durch **Überlagerung** ermittelt werden.

Bei der Möglichkeit b) geht man folgendermaßen vor (Bild 3-11):

Die Führungsbewegung des starren Körpers wird in einem festen Koordinatensystem beschrieben. Für den Ort auf dem Starrkörper, in dem sich der Massenpunkt P , der die Relativbewegung ausführt, gerade befindet, werden nach Gl. (3.9) die **Führungsgeschwindigkeit** $\vec{v}_f$ und nach Gl. (3.10) die **Führungsbeschleunigung** $\vec{a}_f$ ermittelt.

Für die Untersuchung der **Relativbewegung** des Punktes P kann man das Füh-

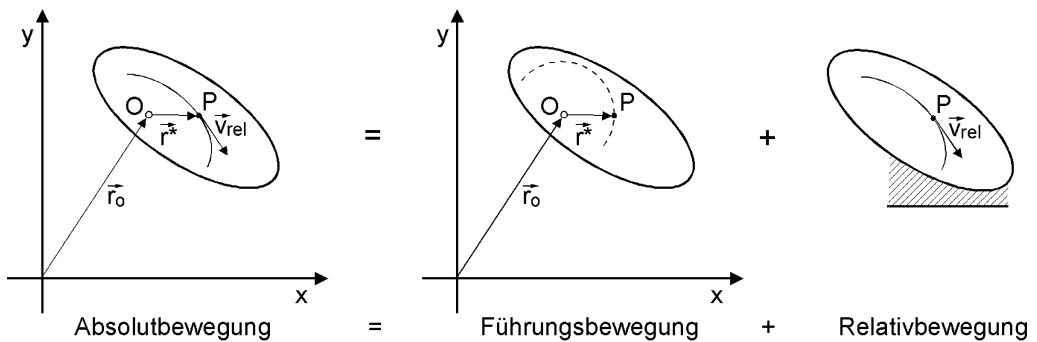


Bild 3-11: Ermittlung des Bewegungszustandes eines Starrkörpers durch Überlagerung von Führungs- und Relativbewegung

runssystem als in Ruhe befindlich ansehen. Nach den Regeln der Kinematik des Massenpunktes werden dann die Relativgeschwindigkeit $\vec{v}_{rel}$ und die Relativbeschleunigung $\vec{a}_{rel}$ ermittelt.

Die Absolutgeschwindigkeit des Punktes P bezüglich des festen Koordinatensystems ergibt sich dann durch vektorielle Addition von Führungs- und Relativgeschwindigkeit:

$$\vec{v} = \vec{v}_f + \vec{v}_{rel} \quad . \quad (3.12)$$

Die **Absolutbeschleunigung** des Punktes P bezüglich des festen Koordinatensystems ergibt sich durch Addition von Führungs- und Relativbeschleunigung und einer dritten Komponente, der so genannten **Coriolisbeschleunigung** :

$$\vec{a} = \vec{a}_f + \vec{a}_{rel} + \vec{a}_C \quad . \quad (3.13)$$

Diese zusätzliche Beschleunigungskomponente beruht darauf, dass im allgemeinen durch die Relativbewegung gegenüber dem Führungssystem, sich P vom Zentrum des Rotationsanteils der Führungsbewegung entfernt oder sich ihm nähert. Daher gelangt der Punkt in einen Bereich höherer oder niedrigerer Führungsgeschwindigkeit, d. h. er wird beschleunigt oder verzögert. Der Vektor der Coriolisbeschleunigung hat den Wert

$$\vec{a}_C = 2\vec{\omega}_f \times \vec{v}_{rel} \quad , \quad (3.14)$$

wobei $\vec{\omega}_f$ die Winkelgeschwindigkeit der Führungsbewegung ist.

Im folgenden Beispiel sollen die resultierenden Geschwindigkeiten und Beschleunigungen bei Auftreten einer Führungs- und einer Relativbewegung berechnet werden. Wie in Bild 3-12 a) gezeigt, bewegt sich der Mittelpunkt einer an einer Stange gelenkig geführten starren Scheibe mit der Winkelgeschwindigkeit ω_0 und der Winkelbeschleunigung α_0 auf einer Kreisbahn mit dem Radius R . Die Scheibe wiederum dreht sich um ihren Mittelpunkt mit ω_1 und α_1 , bezogen auf den sich auf der Kreisbahn bewegenden Endpunkt der Stange. In der kreisförmigen Führungsrinne mit Radius r bewegt sich relativ zur Scheibe ein Punkt P mit v_{rel} und a_{rel} . Zum betrachteten Zeitpunkt hat P den Abstand r^* vom Mittelpunkt der Scheibe.

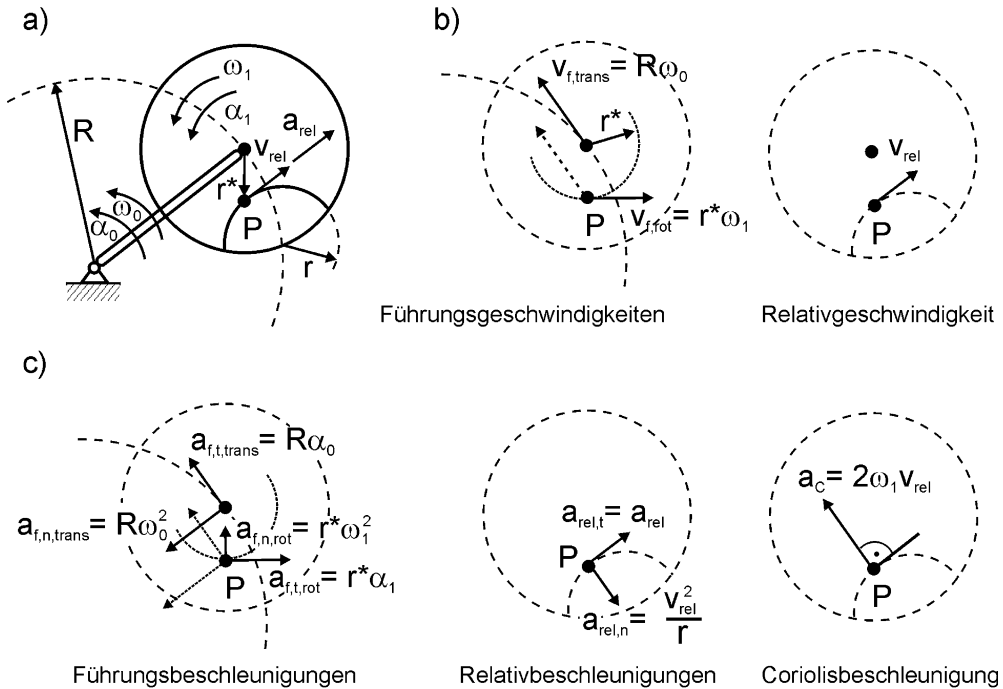


Bild 3-12: Absolutgeschwindigkeit und -beschleunigung bei einer Relativbewegung

In diesem Beispiel kommen alle Geschwindigkeits- und Beschleunigungsanteile vor, die für den Punkt P überhaupt möglich sind. Die Führungsbewegung, die P erfährt, wird als Translation mit dem Scheibenmittelpunkt als Bezugspunkt (Bewegung auf einer Kreisbahn mit Radius R) und einer Rotation um diesen Punkt (Radius r^*) betrachtet. Die Relativbewegung ist eine Bewegung auf einer Kreisbahn mit Radius r . Der translatorische Anteil der Führungsgeschwindigkeit gilt für alle Scheibenpunkte und damit auch für P (Bild 3-12 b)). Hinzu kommt die Bahngeschwindigkeit der Rotation um den Scheibenmittelpunkt. Die Absolutgeschwindigkeit ergibt sich dann, wenn man die beiden Komponenten der Führungsgeschwindigkeit mit der Relativgeschwindigkeit vektoriell addiert.

Die Führungsbeschleunigung besteht aus vier Anteilen: Der Scheibenmittelpunkt erfährt eine Bahnbeschleunigung und eine Normalbeschleunigung, die auch für P gelten (Bild 3-12 c)). Die Rotation um den Scheibenmittelpunkt liefert ebenfalls zwei Anteile. Die Relativbewegung steuert noch einmal zwei Anteile bei. Die ebenfalls auftretende Coriolisbeschleunigung, die senkrecht zu $\mathbf{v}_{rel}$ in Richtung des Drehsinns von ω_1 gerichtet ist, liefert schließlich eine siebte Komponente, die mit den anderen vektoriell zur Absolutbeschleunigung aufaddiert werden kann.

3.2.3 Die Bewegung des starren Körpers im Raum

Um die räumliche Bewegung eines starren Körpers beschreiben zu können, kann man, ähnlich wie beim ebenen Starrkörper, die Lage zweier beliebiger Punkte des starren Körpers im Raum verfolgen. Die Lage zweier Punkte P_1 und P_2 sei daher

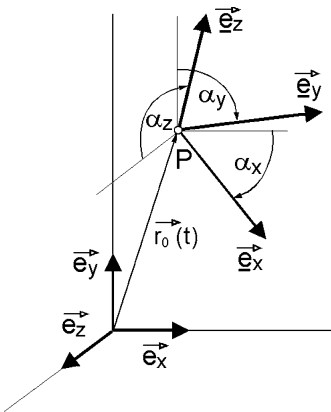


Bild 3-13: Position und Orientierung eines Punktes des starren Körpers

durch zwei Ortsvektoren $\vec{r}_1$ und $\vec{r}_2$ beschrieben. Diese beiden Punkte erfordern sechs Koordinatenangaben, aber die beiden Vektoren sind nicht unabhängig voneinander. Dies ist leicht einsehbar, da der feste Abstand der beiden Punkte des starren Körpers in jeder Lage eingehalten wird. Das hieße, dass die Anzahl der Freiheitsgrade des Körpers kleiner sechs wäre, aber bei Vorgabe von fünf Koordinatenwerten und dem Abstand der beiden Punkte wäre die Lage des starren Körpers noch nicht eindeutig beschrieben. Er könnte bei gleichen Koordinaten der beiden Punkte noch eine Drehung um die Verbindungslinie der beiden Punkte ausführen. Der Körper besitzt also sechs Freiheitsgrade im Raum, da man für seine eindeutige Bestimmung beispielsweise fünf Koordinatenangaben zweier Punkte und eine Winkelkoordinate benötigt. Allgemein kann man sagen, dass die Lage des starren Körpers im Raum durch die Angabe von sechs geeigneten Koordinaten bestimmt werden kann. Dazu verwendet man beispielsweise häufig die drei Koordinaten eines Punktes des Starrkörpers in einem Bezugssystem und drei Winkelangaben. Die Winkel geben die Verdrehung der Achsen eines **körpereigenen Koordinatensystems**, zu den Achsen des **ortsfesten Bezugssystems** an, und zwar für den betrachteten Punkt (Bild 3-13). Man bezeichnet die drei kartesischen Koordinaten des Punktes P im Bezugssystem auch als **Position** des Punktes und die drei Winkelkoordinaten als **Orientierung** des Punktes des starren Körpers.

Führt der starre Körper nur eine reine Translation im Raum aus, so verändert sich nur die Position des Punktes P und die Bewegung kann mit den im Kapitel 3.1 hergeleiteten Gleichungen beschrieben werden. Die Beschreibung der Bewegung des Starrkörpers wird schwieriger, wenn sich die Orientierung eines Punktes P des starren Körpers durch Rotationen ändert.

3.2.3.1 Rotation im Raum

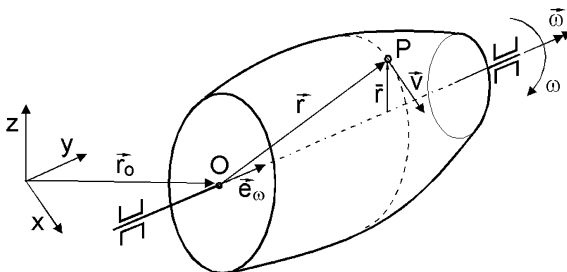


Bild 3-14: Rotation eines starren Körpers um eine feste Achse

Zur Herleitung der allgemeinen Bewegung des Starrkörpers im Raum soll zuerst der einfachere Fall der reinen Rotation des Körpers um eine Achse betrachtet werden, ein Fall der in technischen Systemen häufig auftritt. Wie in Bild 3-14 gezeigt, sei der Punkt O des Starrkörpers, dessen Position im Raum durch den Ortsvektor $\vec{r}_0$ in einem ruhenden Bezugssystem festgelegt

ist, durch eine starre Bindung so fixiert, dass der Körper nur um eine Achse, die durch den Punkt O läuft, rotieren kann. Die Richtung dieser Achse wird durch einen Einheitsvektor $\vec{e}_\omega$ festgelegt. Diese Rotation um eine feste Achse kann durch die

Position des Punktes O , den Einheitsvektor $\vec{e}_\omega$ und eine skalare Winkelgeschwindigkeit $\omega(t)$ eindeutig beschrieben werden. Der **Vektor der Winkelgeschwindigkeit**, der die Richtung von $\vec{e}_\omega$ hat, ist dann folgendermaßen definiert:

$$\vec{\omega}(t) = \vec{e}_\omega \cdot \omega(t)$$

Betrachtet man nun einen beliebigen Punkt P des Körpers, dessen Position relativ zu O durch den Ortsvektor $\vec{r}$ beschrieben wird und der einen Abstand $\bar{r}$ zur Drehachse des Körpers hat, so kann seine Geschwindigkeit in Bezug auf O folgendermaßen beschrieben werden:

$$\vec{v} = \vec{\omega} \times \vec{r} = \omega(t) \cdot \vec{e}_\omega \times \vec{r}(t) \quad (3.15)$$

Die Bahngeschwindigkeit $\vec{v}$ des Punktes P relativ zum Punkt O ist also das Vektorprodukt aus dem Vektor der Winkelgeschwindigkeit $\vec{\omega}$ und dem Ortsvektor $\vec{r}$ des Punktes. Der Betrag der Bahngeschwindigkeit ist entsprechend den Regeln für das Vektorprodukt $\omega \cdot |\vec{r}| \cdot \sin \alpha$, wobei α der von $\vec{r}$ und $\vec{e}_\omega$ eingeschlossene Winkel ist. Da der senkrechte Abstand $\bar{r}$ des Punktes P von der Drehachse ist, hat der Ergebnisvektor $\vec{v}$ den Betrag der Bahngeschwindigkeit $\bar{r} \cdot \omega$. Außerdem liegt der Ergebnisvektor senkrecht zu den beiden Vektoren des Kreuzproduktes, so dass dieser die Kreisbahn des Punktes P bei der Rotation tangieren muss.

Die Ableitung nach der Zeit liefert dann unter Verwendung der Produktregel den Beschleunigungsvektor:

$$\vec{a} = \frac{d\vec{v}}{dt} = \dot{\omega}(t) \cdot \vec{e}_\omega \times \vec{r}(t) + \omega(t) \cdot \vec{e}_\omega \times \frac{d\vec{r}(t)}{dt}$$

In diesem Ausdruck taucht im zweiten additiven Term die Bahngeschwindigkeit des Punktes P auf:

$$\frac{d\vec{r}}{dt} = \vec{v} = \vec{\omega} \times \vec{r} = \omega(t) \cdot \vec{e}_\omega \times \vec{r}(t)$$

Damit ergibt sich der **Vektor der Beschleunigung bei einer Rotation des starren Körpers um eine feste Achse** zu:

$$\vec{a} = \dot{\omega}(t) \cdot \vec{e}_\omega \times \vec{r}(t) + \omega^2(t) \cdot \vec{e}_\omega \times (\vec{e}_\omega \times \vec{r}(t)) = \vec{a}_t + \vec{a}_n \quad (3.16)$$

Der erste additive Term ist ein Vektor, der wie der Geschwindigkeitsvektor des Punktes P tangential zur Bahnkurve gerichtet ist, d. h. dies ist der Vektor der Tangentialbeschleunigung $\vec{a}_t$ oder auch der Bahnbeschleunigung, der den Betrag $\bar{r} \cdot \dot{\omega}$ besitzt. Der zweite additive Term, der ein zweifaches Vektorprodukt enthält, liefert den Vektor der Normalbeschleunigung $\vec{a}_n$, der zur Drehachse weist und den Betrag $\bar{r} \cdot \omega^2$ besitzt.

Ist die Drehachse des Körpers nicht mehr wie in Bild 3-14 dargestellt in zwei Punkten fixiert, sondern wird nur noch im Punkt O festgehalten, so hat diese keine raum-feste Lage mehr und kann sich selbst um den Punkt O drehen. In diesem allgemeinen Fall der Rotation ändert sich auch die Richtung des Vektors $\vec{e}_\omega$, wodurch er

zeitlich veränderlich wird, d. h. der Einheitsvektor muss nun als $\vec{e}_\omega(t)$ bezeichnet werden. Bei der Herleitung des Geschwindigkeits- und Beschleunigungsvektors des Punktes P durch ein- bzw. zweimaliges Ableiten des Ortsvektors $\vec{r}$ nach der Zeit muss daher auch der Einheitsvektor $\vec{e}_\omega(t)$ abgeleitet werden und man erhält:

$$\vec{v} = \vec{\omega} \times \vec{r} = \omega \cdot \vec{e}_\omega \times \vec{r} \quad (3.17)$$

$$\vec{a} = \frac{d\vec{v}}{dt} = \omega \cdot \frac{d\vec{e}_\omega}{dt} \times \vec{r} + \dot{\omega} \cdot \vec{e}_\omega + \omega^2 \cdot \vec{e}_\omega \times (\vec{e}_\omega \times \vec{r}) \quad (3.18)$$

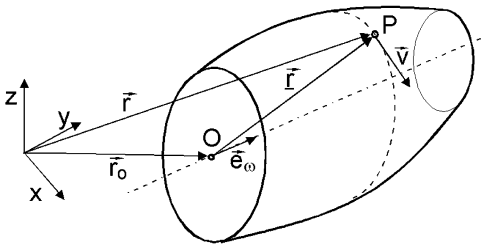


Bild 3-15: Allgemeine Bewegung eines starren Körpers zusammengesetzt aus Translation und

Lässt man jetzt zusätzlich noch eine Translation des Punktes O im Bezugssystem zu (Bild 3-15), wobei der starre Körper um O rotiert, so erhält man die Beschreibung der **allgemeinen Bewegung des starren Körpers** durch folgende drei Gleichungen:

$$\vec{r} = \vec{r}_0 + \vec{r} \quad (3.19)$$

$$\vec{v} = \frac{d\vec{r}_0}{dt} + \vec{\omega} \times \vec{r} = \frac{d\vec{r}_0}{dt} + \omega \cdot \vec{e}_\omega \times \vec{r} \quad (3.20)$$

$$\begin{aligned} \vec{a} &= \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}_0}{dt^2} + \vec{\omega} \times \vec{r} + \vec{\omega} \times (\vec{\omega} \times \vec{r}) \\ &= \frac{d^2\vec{r}_0}{dt^2} + \omega \cdot \frac{d\vec{e}_\omega}{dt} \times \vec{r} + \dot{\omega} \cdot \vec{e}_\omega \times \vec{r} + \omega^2 \cdot \vec{e}_\omega \times (\vec{e}_\omega \times \vec{r}) \quad (3.21) \end{aligned}$$

3.2.3.2 Relativbewegung eines Punktes des starren Körpers

Bereits im Abschnitt 3.2.2 war die Relativbewegung eines Punktes des ebenen Starrkörpers zur Führungsbewegung des Starrkörpers so behandelt worden, dass die Führungsbewegung in einem ruhenden Bezugssystem beschrieben wurde und die Relativbewegung in einem mit dem Starrkörper mitgeführten Bezugssystem. Diese Art der Beschreibung der Bewegung ist in Bild 3-16 dargestellt. Dabei bedeuten:

$\vec{e}_x, \vec{e}_y, \vec{e}_z$: Einheitsvektoren des Inertialsystems

$\vec{e}_x^*, \vec{e}_y^*, \vec{e}_z^*$: Einheitsvektoren des bewegten Bezugssystems

$\vec{r}_0(t)$: Ortsvektor der Bewegung des Ursprungs O des bewegten Systems gegenüber dem ruhenden System

$\vec{\omega}_0(t)$: Drehung des bewegten Systems gegenüber dem ruhenden System

$\vec{r}^*(t)$: Ortsvektor der Bahnkurve des Punktes P im bewegten System

$\vec{r}(t)$: Ortsvektor der Bahnkurve des Punktes P im ruhenden System

Zur Bestimmung der Bahnkurve des Punktes P im Inertialsystem wird zuerst der Ortsvektor im bewegten System ermittelt und anschließend auf das Inertialsystem zurückgerechnet. Es gilt:

$$\vec{r}(t) = \vec{r}_0(t) + \vec{r}^*(t) \quad \text{oder kurz} \quad \vec{r} = \vec{r}_0 + \vec{r}^* ,$$

da alle Größen in diesem allgemeinen Fall der räumlichen Bewegung des Starrkörpers von der Zeit abhängig sind.

Die Führungsgeschwindigkeit im Inertialsystem ist nach Gl. (3.20):

$$\vec{v}_f = \frac{d\vec{r}_0}{dt} + \vec{\omega}_0 \times \vec{r}^*$$

Die Relativgeschwindigkeit $\vec{v}_{rel}$ ist für den Beobachter im bewegten Bezugssystem die Änderung des Vektors $\vec{r}^*$:

$$\vec{v}_{rel} = \frac{d^*\vec{r}^*}{dt} ,$$

wobei das Symbol d^*/dt die Ableitung im bewegten System bedeutet, d. h. die Änderung von $\vec{r}^*$ durch die Bewegung des Punktes O des Starrkörpers im Inertialsystem wird bei dieser Ableitung nicht berücksichtigt. Die Absolutgeschwindigkeit erhält man durch Addition von Führungs- und Relativgeschwindigkeit:

$$\vec{v} = \vec{v}_f + \vec{v}_{rel} = \frac{d\vec{r}_0}{dt} + \vec{\omega}_0 \times \vec{r}^* + \frac{d^*\vec{r}^*}{dt} . \quad (3.22)$$

Eine andere Möglichkeit zur Ermittlung der Absolutgeschwindigkeit ist die Ableitung des Ortsvektors $\vec{r}$ im ruhenden System:

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{d\vec{r}_0}{dt} + \frac{d\vec{r}^*}{dt} .$$

Der Vergleich der beiden letzten Gleichungen liefert den Zusammenhang zwischen der Ableitung eines Vektors im ruhenden und der Ableitung im bewegten System:

$$\frac{d\vec{r}^*}{dt} = \vec{\omega}_0 \times \vec{r}^* + \frac{d^*\vec{r}^*}{dt} . \quad (3.23)$$

Entsprechend gilt daher auch

$$\frac{d\vec{r}_{rel}}{dt} = \vec{\omega}_0 \times \vec{v}_{rel} + \frac{d^*\vec{v}_{rel}}{dt} .$$

Um zur Absolutbeschleunigung im ruhenden System zu kommen, muss die Gl. (3.22) nach der Zeit abgeleitet werden:

$$\vec{a} = \frac{d\vec{v}_f}{dt} + \frac{d\vec{v}_{rel}}{dt} = \frac{d^2\vec{r}_0}{dt^2} + \frac{d\vec{\omega}_0}{dt} \times \vec{r}^* + \vec{\omega}_0 \times \frac{d\vec{r}^*}{dt} + \frac{d^*\vec{v}_{rel}}{dt} . \quad (3.24)$$

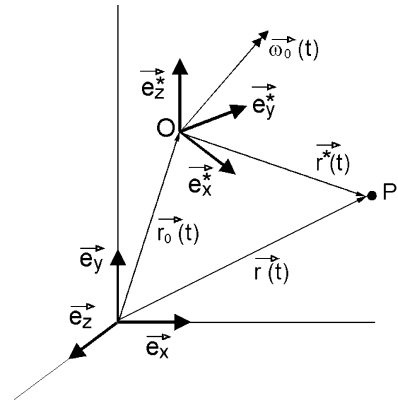


Bild 3-16: Massenpunkt im bewegten Koordinatensystem

$$\begin{aligned}
 \vec{a} &= \frac{d^2 \vec{r}_0}{dt^2} + \frac{d\vec{\omega}_0}{dt} \times \vec{r}^* + \vec{\omega}_0 \times (\vec{\omega}_0 \times \vec{r}^* + \frac{d^* \vec{r}^*}{dt}) + \vec{\omega}_0 \times \vec{v}_{rel} + \frac{d^* \vec{v}_{rel}}{dt} \\
 &= \frac{d^2 \vec{r}_0}{dt^2} + \frac{d\vec{\omega}_0}{dt} \times \vec{r}^* + \vec{\omega}_0 \times (\vec{\omega}_0 \times \vec{r}^*) + 2\vec{\omega}_0 \times \vec{v}_{rel} + \frac{d^* \vec{v}_{rel}}{dt} \\
 &= \vec{a}_f + \vec{a}_C + \vec{a}_{rel}
 \end{aligned}$$

Die ersten drei additiven Terme stellen die **Führungsbeschleunigung** des Punktes P des starren Körpers dar und entsprechen der Beschleunigung in Gl. (3.21). Der letzte Term ist die **Relativbeschleunigung** im bewegten System, wobei der starre Körper als in Ruhe befindlich betrachtet werden kann. Der dritte Anteil ist die **Coriolisbeschleunigung**, die wir bereits bei der ebenen Bewegung nach Gl. (3.14) kennengelernt haben.

3.2.3.3 Darstellung der Bewegung des starren Körpers in Matrixschreibweise

Um die einzelnen Komponenten eines Ortsvektors und seiner Ableitungen bei der Beschreibung des Bewegungszustandes (Lage, Geschwindigkeit, Beschleunigung) eines starren Körpers im Raum behandeln zu können, verwendet man, wie auch schon bei der Bewegung des Massenpunktes im Raum, die Matrixschreibweise. Es wird für die Beschreibung wie in den vorherigen Kapiteln ein raumfestes Inertialsystem, gekennzeichnet durch die Schreibweise $^{(0)}$, und ein körperfestes Koordinatensystem, gekennzeichnet durch die Schreibweise $^{(1)}$, verwendet (Bild 3-17) [2.2].

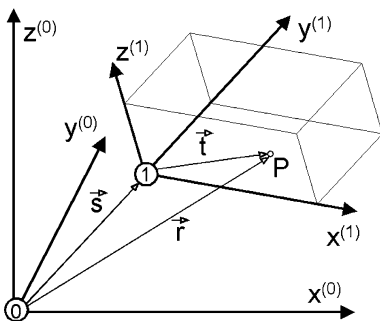


Bild 3-17: Beschreibung des Bewegungszustandes eines Punktes P des starren Körpers

Die Lage des Ursprungs des körperfesten Koordinatensystems wird durch den Vektor $\vec{s}$, die Lage des Punktes P des starren Körpers im Inertialsystem durch den Vektor $\vec{r}$ und die Lage des Punktes P des starren Körpers im körperfesten System durch den Vektor $\vec{t}$ beschrieben. Es gilt dann:

$$\vec{r} = \vec{s} + \vec{t} \quad (3.25)$$

Die Darstellung dieser Vektorgleichung in Matrixschreibweise lautet:

$$\begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix}^{(0)} = \begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix}^{(0)} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(0)}$$

Dies kann man auch abgekürzt als

$$\mathbf{r}^{(0)} = \mathbf{s}^{(0)} + \mathbf{t}^{(0)} \quad (3.26)$$

schreiben, wobei fett gedruckte Buchstaben ($\mathbf{r}$, $\mathbf{t}$, $\mathbf{M}$, $\mathbf{T}$) ohne Pfeil Vektoren und Matrizen darstellen.

Die Lage eines Punktes P des starren Körpers wird in der Regel im körperfesten Koordinatensystem angegeben werden können. Der entsprechende Ortsvektor in

der transponierten Schreibweise lautet dann:

$$\mathbf{t}^{(1)T} = (\mathbf{t}_x^{(1)}, \mathbf{t}_y^{(1)}, \mathbf{t}_y^{(1)}) .$$

Den Vektor $\mathbf{t}^{(0)}$ erhält man daraus durch Umrechnung vom körperfesten auf das raumfeste Koordinatensystem mit Hilfe einer Transformationsmatrix $\mathbf{T}$:

$$\mathbf{t}^{(0)} = \mathbf{T} \cdot \mathbf{t}^{(1)} \quad (3.27)$$

Dadurch kann die Gl. (3.26) folgendermaßen geschrieben werden:

$$\mathbf{r}^{(0)} = \mathbf{s}^{(0)} + \mathbf{T} \cdot \mathbf{t}^{(1)} \quad (3.28)$$

Der physikalische Vektor $\vec{r}$ lässt sich demnach je nach Bezugssystem durch $\mathbf{t}^{(0)}$ oder durch $\mathbf{t}^{(1)}$ darstellen. Den Zusammenhang zwischen den beiden Koordinatensystemen mit den Kennzeichnungen $^{(0)}$ und $^{(1)}$ ist durch die Matrix $\mathbf{T}$ gegeben. Die notwendigen Operationen um die beiden Koordinatensysteme ineinander zu überführen sind lineare **Verschiebungen** und **Verdrehungen**.

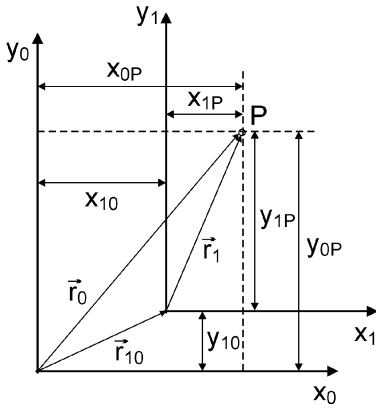


Bild 3-18: Koordinaten eines Punktes P in der Ursprungslage (0) und nach der Verschiebung (1)

Die Form der Transformationsmatrix für Verschiebungen und Verdrehungen soll nun für den zweidimensionalen Fall hergeleitet werden. Für die Verschiebung eines Koordinatensystems aus der Lage $^{(0)}$ in die Lage $^{(1)}$ gilt das in Bild 3-18 dargestellte. Hierin wird die Lage eines Punktes P durch Ortsvektoren beschrieben:

$$\vec{r}_1 = \vec{r}_0 - \vec{r}_{10} .$$

In Komponentenschreibweise lautet die gleiche Beziehung:

$$x_{1P} = x_{0P} - x_{10}$$

$$y_{1P} = y_{0P} - y_{10}$$

Um daraus eine Matrixschreibweise zu gewinnen, werden die beiden linearen Gleichungen um eine dritte Gleichung ergänzt, die die Identität $1 = 1$ enthält.

$$1 = 1$$

$$x_{1P} = x_{0P} - x_{10}$$

$$y_{1P} = y_{0P} - y_{10}$$

Daraus kann man nun die folgende Matrixschreibweise ableiten:

$$\begin{pmatrix} 1 \\ x_{1P} \\ y_{1P} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -x_{10} & 1 & 0 \\ -y_{10} & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x_{0P} \\ y_{0P} \end{pmatrix}$$

oder kurz

$$\mathbf{r}^{(1)} = \mathbf{T}_{10}^S \cdot \mathbf{r}^{(0)} \quad (3.29)$$

In Bild 3-19 ist die Verdrehung des Koordinatensystems um den Winkel φ_{10} aus der Lage $^{(0)}$ in die Lage $^{(1)}$ dargestellt. Um wieder wie bei der Verschiebung eine entsprechende Komponentendarstellung zu erhalten, werden die Hilfsmaße a, b, c und d eingeführt. Es gilt:

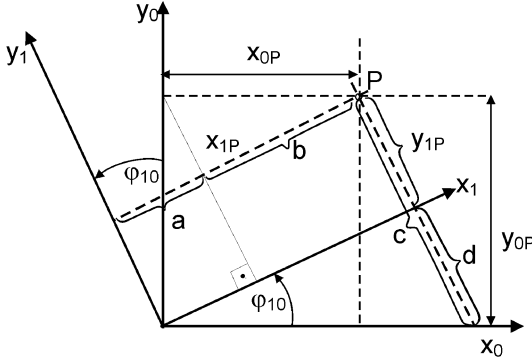


Bild 3-19: Koordinaten eines Punktes P in der Ausgangslage $^{(0)}$ und nach Verdrehung $^{(1)}$

$$x_{1P} = a + b \quad y_{1P} = c - d$$

Die Hilfsmaße kann man folgendermaßen darstellen:

$$a = y_{0P} \sin \varphi_{10} \quad b = x_{0P} \cos \varphi_{10}$$

$$c = y_{0P} \cdot 1 / \cos \varphi_{10} \quad d = x_{1P} \tan \varphi_{10}$$

Daraus kann man direkt die erste Komponentengleichung aufschreiben:

$$x_{1P} = x_{0P} \cos \varphi_{10} + y_{0P} \sin \varphi_{10} \quad (3.30)$$

Für die zweite Komponentengleichung findet man:

$$y_{1P} = y_{0P} \cdot 1 / \cos \varphi_{10} - x_{1P} \tan \varphi_{10}$$

Nach einigen Umformungen mit Hilfe trigonometrischer Beziehungen findet man dann:

$$y_{1P} = y_{0P} \cos \varphi_{10} - x_{0P} \sin \varphi_{10} \quad (3.31)$$

Nun kann man aus den Gleichungen (3.30, 3.31) und aus der Identität $1 = 1$ wieder ein lineares Gleichungssystem bilden, das in Matrixschreibweise folgendermaßen aussieht.

$$\begin{pmatrix} 1 \\ x_{1P} \\ y_{1P} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_{10} & \sin \varphi_{10} \\ 0 & -\sin \varphi_{10} & \cos \varphi_{10} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x_{0P} \\ y_{0P} \end{pmatrix}$$

Die Kurzschreibweise für diese Matrixdarstellung ist die gleiche wie in Gl. (3.29). Die Transformationsmatrizen unterscheiden sich nur durch den Index für die Veränderungsgröße φ :

$$\mathbf{r}^{(1)} = \mathbf{T}_{10}^\varphi \cdot \mathbf{r}^{(0)} \quad (3.32)$$

Ist das Koordinatensystem sowohl verschoben als auch verdreht, so ergibt sich eine Gesamttransformationsmatrix, die das Produkt der Einzeltransformationsmatrizen ist:

$$\begin{pmatrix} 1 \\ x_{1P} \\ y_{1P} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -x_{10} & \cos \varphi_{10} & \sin \varphi_{10} \\ -y_{10} & -\sin \varphi_{10} & \cos \varphi_{10} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x_{0P} \\ y_{0P} \end{pmatrix}$$

$$\mathbf{r}^{(1)} = \mathbf{T}_{10}^s \mathbf{T}_{10}^o \cdot \mathbf{r}^{(0)} = \mathbf{T}_{10} \cdot \mathbf{r}^{(0)}$$

Nun kann, wie in Bild 3-17 dargestellt, das körperfeste System gegenüber dem raumfesten System um alle drei Achsen des Koordinatensystems verdreht sein. Für jede dieser Einzeldrehungen um α_1 , α_2 und α_3 gibt es eine zugehörige Transformationsmatrix, die den Vektor $\mathbf{t}^{(0)}$ über zwei Zwischenstufen in den Vektor $\mathbf{t}^{(1)}$ überführt (Bild 3-20). Die einzelnen Zwischenschritte werden folgendermaßen geschrieben:

$$\mathbf{t}^{(I)} = \mathbf{T}_1^T(\alpha_1) \cdot \mathbf{t}^{(0)} \quad (3.33a)$$

$$\mathbf{t}^{(II)} = \mathbf{T}_2^T(\alpha_2) \cdot \mathbf{t}^{(I)} \quad (3.33b)$$

$$\mathbf{t}^{(1)} = \mathbf{t}^{(III)} = \mathbf{T}_3^T(\alpha_3) \cdot \mathbf{t}^{(II)} \quad (3.33c)$$

Dabei bezeichnet (α_1) die Abhängigkeit der Matrix $\mathbf{T}_1^T$ vom Winkel α_1 . Die drei Gleichungen kann man folgendermaßen zusammenfassen:

$$\mathbf{t}^{(1)} = \mathbf{T}^T(\alpha_1, \alpha_2, \alpha_3) \cdot \mathbf{t}^{(0)} \quad (3.34)$$

Dabei bedeutet

$$\mathbf{T}^T(\alpha_1, \alpha_2, \alpha_3) = \mathbf{T}_3^T(\alpha_3) \cdot \mathbf{T}_2^T(\alpha_2) \cdot \mathbf{T}_1^T(\alpha_1) \quad (3.35)$$

Es ist zu beachten, dass bei anderer Wahl der Drehreihenfolge sich eine andere Endlage des Vektors $\mathbf{t}$ ergibt. Ist die Reihenfolge wie im obigen Beispiel α_1 , α_2 , α_3 , d. h. Drehung in der Reihenfolge um die x - y - z - Achsen, so bezeichnet man die Winkel als **Kardan-Winkel**, ein Name, der ebenso wie das Prinzip der **kardanischen Lagerung** auf den italienischen Arzt und Mathematiker Cardanus (1501-1576) zurückgeht. Wählt man eine Drehreihenfolge um die z - x - z - Achsen, so spricht man von **Euler-Winkeln**. Diese Form der Reihenfolge der Drehungen verwendet man aus Gründen der Anschaulichkeit und es sind noch weitere Drehreihenfolgen üblich [3.3].

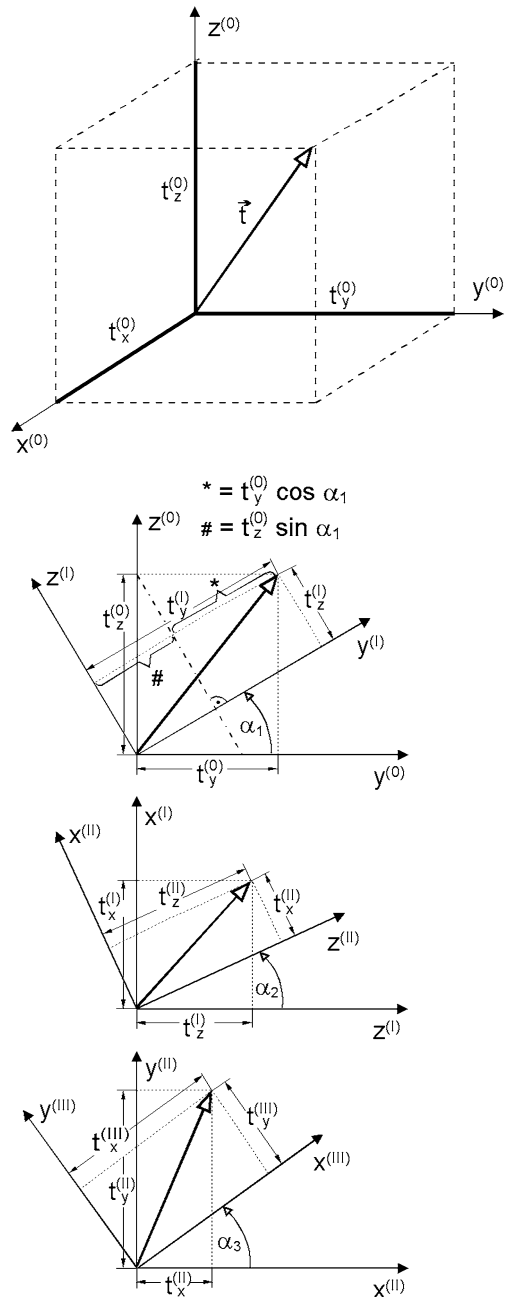


Bild 3-20: Transformation eines Vektors aus dem raumfesten in ein körperfestes Koordinatensystem

Drehung um die $x^{(0)}$ - Achse

$$\begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(I)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_1 & \sin \alpha_1 \\ 0 & -\sin \alpha_1 & \cos \alpha_1 \end{pmatrix} \cdot \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(0)}$$

Drehung um die $y^{(I)}$ - Achse

$$\begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(II)} = \begin{pmatrix} \cos \alpha_2 & 0 & -\sin \alpha_2 \\ 0 & 1 & 0 \\ \sin \alpha_2 & 0 & \cos \alpha_2 \end{pmatrix} \cdot \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(I)}$$

Drehung um die $z^{(II)}$ - Achse

$$\mathbf{t}^{(1)} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(III)} = \begin{pmatrix} \cos \alpha_3 & \sin \alpha_3 & 0 \\ -\sin \alpha_3 & \cos \alpha_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(II)}$$

Will man umgekehrt einen im körperfesten Koordinatensystem $^{(1)}$ gegebenen Vektor ins raumfeste Koordinatensystem $^{(0)}$ überführen, so benötigt man die inverse Transformationsmatrix:

$$\mathbf{t}^{(0)} = (\mathbf{T}^T)^{-1} \cdot \mathbf{t}^{(1)} .$$

Eine Transformationsmatrix, die den Übergang zwischen kartesischen Koordinatensystemen bestimmt, wird als **orthogonal** bezeichnet [3.4]. Für eine solche orthogonale Koordinatentransformation gilt, dass die inverse Transformation durch die transponierte Matrix der ursprünglichen Transformationsmatrix bewirkt wird. Wendet man dies auf die letzte Gleichung an, so ergibt sich wegen

$$(\mathbf{T}^T)^{-1} = \mathbf{T} \quad (3.36)$$

Folgendes:

$$\mathbf{t}^{(0)} = \mathbf{T} \cdot \mathbf{t}^{(1)} .$$

Die hier und in Gl. (3.28) benötigte Transformationsmatrix $\mathbf{T}$ erhält man durch Lösung der Gl. (3.35) durch Multiplikation der Matrizen für die Einzeldrehungen und durch anschließende Transposition. Dabei muss darauf geachtet werden, dass die Reihenfolge der Matrizenmultiplikationen

$$\mathbf{T}_3^T(\alpha_3) \cdot \mathbf{T}_2^T(\alpha_2) \cdot \mathbf{T}_1^T(\alpha_1)$$

nicht vertauscht wird, da das einer Vertauschung der Drehreihenfolge entsprechen würde.

$$\mathbf{T} = (\mathbf{T}^T(\alpha_1, \alpha_2, \alpha_3)) = \begin{pmatrix} \cos \alpha_2 \cos \alpha_3 & -\cos \alpha_2 \sin \alpha_3 & \sin \alpha_2 \\ \sin \alpha_1 \sin \alpha_2 \cos \alpha_3 + \cos \alpha_1 \sin \alpha_3 & -\sin \alpha_1 \sin \alpha_2 \sin \alpha_3 + \cos \alpha_1 \cos \alpha_3 & -\sin \alpha_1 \cos \alpha_2 \\ -\cos \alpha_1 \sin \alpha_2 \cos \alpha_3 + \sin \alpha_1 \sin \alpha_3 & \cos \alpha_1 \sin \alpha_2 \sin \alpha_3 + \sin \alpha_1 \cos \alpha_3 & \cos \alpha_1 \cos \alpha_2 \end{pmatrix}$$

Als Beispiel für eine solche Koordinatentransformation aus dem raumfesten Koordinatensystem $^{(0)}$ in ein körperfestes Koordinatensystem $^{(1)}$ soll eine Verdrehung des in Bild 3-21 dargestellten Einheitswürfels dienen. Seine Ecken haben in der Lage die folgenden Koordinaten:

$P_1(0,0,0)$; $P_2(1,0,0)$; $P_3(0,1,0)$

$P_4(0,0,1)$; $P_5(1,1,1)$; $P_6(0,1,1)$

$P_7(1,0,1)$; $P_8(1,1,0)$

Die Verdrehungen sollen um die x-Achse ($\alpha_1 = 30^\circ$) und um die y-Achse ($\alpha_2 = 45^\circ$) vorgenommen werden. Die für die Transformation erforderliche Matrix $\mathbf{T}$ bzw. $\mathbf{T}^T$ erhält man durch Einsetzen der folgenden Werte in die oben abgebildete Matrix:

$$\begin{aligned} \sin \alpha_1 &= \frac{1}{2} & \sin \alpha_2 &= \frac{\sqrt{2}}{2} & \sin \alpha_3 &= 0 \\ \cos \alpha_1 &= \frac{\sqrt{3}}{2} & \cos \alpha_2 &= \frac{\sqrt{2}}{2} & \cos \alpha_3 &= 1 \end{aligned}$$

Die Matrix aus Gl. (3.36) ergibt sich damit zu:

$$\mathbf{T} = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{4} & \frac{\sqrt{3}}{2} & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{6}}{4} & \frac{1}{2} & \frac{\sqrt{6}}{4} \end{pmatrix}$$

Die Transformation findet nach Gl. (3.34) statt und ergibt folgende Gleichung:

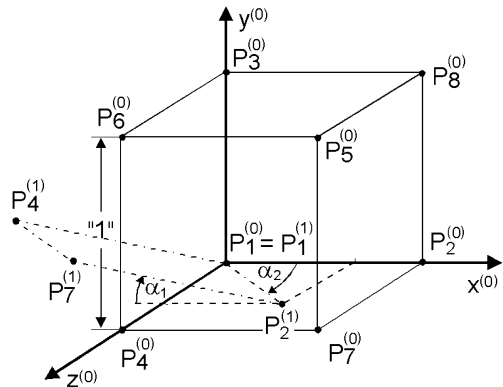


Bild 3-21: Beispiel für die Transformation eines Einheitswürfels aus dem körperfesten in das raumfeste Koordinatensystem

$$\begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(1)} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{4} & -\frac{\sqrt{6}}{4} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{4} & \frac{\sqrt{6}}{4} \end{pmatrix} \cdot \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}^{(0)}$$

Hieraus kann man die einzelnen Transformationsgleichungen für die acht Punkte entnehmen:

$$t_x^{(1)} = \frac{\sqrt{2}}{2} t_x^{(0)} + \frac{\sqrt{2}}{4} t_y^{(0)} - \frac{\sqrt{6}}{4} t_z^{(0)}$$

$$t_y^{(1)} = 0 \cdot t_x^{(0)} + \frac{\sqrt{3}}{2} t_y^{(0)} + \frac{1}{2} t_z^{(0)}$$

$$t_z^{(1)} = \frac{\sqrt{2}}{2} t_x^{(0)} - \frac{\sqrt{2}}{4} t_y^{(0)} + \frac{\sqrt{6}}{4} t_z^{(0)}$$

Diese Transformationsgleichungen liefern für die Koordinaten der Eckpunkte in der Lage ⁽¹⁾ folgende Werte:

$$\begin{aligned} P_1 &= (0,0,0) & P_5 &= \left(\frac{1}{4}(3 \cdot \sqrt{2} - \sqrt{6}), \frac{1}{2}(\sqrt{3} + 1), \frac{1}{4}(\sqrt{2} + \sqrt{6}) \right) \\ P_2 &= \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) & P_6 &= \left(\frac{1}{4}(\sqrt{2} - \sqrt{6}), \frac{1}{2}(\sqrt{3} + 1), \frac{1}{4}(\sqrt{6} - \sqrt{2}) \right) \\ P_3 &= \left(\frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2}, \frac{\sqrt{2}}{4} \right) & P_7 &= \left(\frac{1}{4}(2 \cdot \sqrt{2} - \sqrt{6}), \frac{1}{2}, \frac{1}{4}(2 \cdot \sqrt{2} + \sqrt{6}) \right) \\ P_4 &= \left(\frac{\sqrt{6}}{4}, \frac{1}{2}, \frac{\sqrt{6}}{4} \right) & P_8 &= \left(\frac{3 \cdot \sqrt{2}}{4}, \frac{\sqrt{3}}{2}, \frac{\sqrt{2}}{4} \right) \end{aligned}$$

Kennt man die Lage eines Punktes P im raumfesten Inertialsystem, so erhält man die **Geschwindigkeit des Punktes P im Inertialsystem** durch Ableitung des Ortsvektors $\mathbf{r}^{(0)}$ nach der Zeit.

$$\mathbf{v}^{(0)} = \dot{\mathbf{r}}^{(0)} = \dot{\mathbf{s}}^{(0)} + \dot{\mathbf{T}}\mathbf{t}^{(1)} + \mathbf{T}\dot{\mathbf{t}}^{(1)} \quad (3.37a)$$

Der dritte Term der Ableitung entfällt, da bei einem starren Körper die Geschwindigkeit des Punktes im körpereigenen System $\mathbf{t}^{(1)}$ gleich Null ist. Die Absolutgeschwindigkeit $\mathbf{v}^{(0)}$ lässt sich nach den o. a. Regeln für Vektoren ins körpereigene Koordinatensystem transformieren:

$$\mathbf{v}^{(1)} = \mathbf{T}^T \cdot \mathbf{v}^{(0)} \quad (3.37b)$$

Da bei orthogonalen Matrizen die inverse gleich der transponierten Matrix ist und wegen $\mathbf{T}\mathbf{T}^T = \mathbf{E}$, d. h. das Produkt einer Matrix mit ihrer Inversen ist gleich der Einheitsmatrix und außerdem gilt $\mathbf{T}\mathbf{T}^T = \mathbf{E}$, kann man die Gl. (3.37a) folgendermaßen

schreiben:

$$\mathbf{v}^{(0)} = \dot{\mathbf{s}}^{(0)} + \dot{\mathbf{T}} \cdot \mathbf{T}^T \cdot \mathbf{T} \cdot \mathbf{t}^{(1)} \quad (3.37c)$$

Setzt man Gl. (3.27) in Gl. (3.37b) ein, so erhält man

$$\mathbf{v}^{(0)} = \dot{\mathbf{s}}^{(0)} + \dot{\mathbf{T}} \cdot \mathbf{T}^T \cdot \mathbf{t}^{(0)} \quad (3.37d)$$

Die zeitliche Ableitung der Matrix $\mathbf{T}$ erhält man durch folgende partielle Differentiation:

$$\dot{\mathbf{T}} = \left(\frac{\partial \mathbf{T}}{\partial \alpha_1} \cdot \dot{\alpha}_1 + \frac{\partial \mathbf{T}}{\partial \alpha_2} \cdot \dot{\alpha}_2 + \frac{\partial \mathbf{T}}{\partial \alpha_3} \cdot \dot{\alpha}_3 \right) \quad (3.37e)$$

Damit erhält man für das Matrixprodukt aus Gl. (3.37d)

$$\dot{\mathbf{T}} \cdot \mathbf{T}^T = \begin{pmatrix} 0 & -\omega_z^{(0)} & \omega_y^{(0)} \\ \omega_z^{(0)} & 0 & -\omega_x^{(0)} \\ -\omega_y^{(0)} & \omega_x^{(0)} & 0 \end{pmatrix} = \langle \omega^{(0)} \rangle$$

wobei $\langle \rangle$ für eine **antisymmetrisch** besetzte Matrix steht und folgende Abkürzungen gelten:

$$\omega_x^{(0)} = \dot{\alpha}_1 + \sin \alpha_2 \dot{\alpha}_3 \quad (3.38a)$$

$$\omega_y^{(0)} = \dot{\alpha}_2 \cos \alpha_1 - \dot{\alpha}_3 \cos \alpha_2 \sin \alpha_1 \quad (3.38b)$$

$$\omega_z^{(0)} = \dot{\alpha}_2 \sin \alpha_1 - \dot{\alpha}_3 \cos \alpha_2 \cos \alpha_1 \quad (3.38c)$$

Für die Absolutgeschwindigkeit aus den Gl. (3.37c) und (3.37d) lässt sich damit auch abgekürzt schreiben

$$\mathbf{v}^{(0)} = \dot{\mathbf{s}}^{(0)} + \langle \omega^{(0)} \rangle \mathbf{T} \cdot \mathbf{t}^{(1)} \quad (3.39a)$$

$$\mathbf{v}^{(0)} = \dot{\mathbf{s}}^{(0)} + \langle \omega^{(0)} \rangle \mathbf{t}^{(0)} \quad (3.39b)$$

Die Komponenten der Matrix $\langle \omega^{(0)} \rangle$ sind die Projektionen der **momentanen Winkelgeschwindigkeiten** der bewegten Achsen des körperfesten Koordinatensystems im Inertialsystem. Wie in Bild 3-22 dargestellt, stehen die momentanen Winkelgeschwindigkeiten $\omega_x^{(0)}, \omega_y^{(0)}, \omega_z^{(0)}$ senkrecht aufeinander. Das Bild zeigt ebenfalls, wie diese Winkelgeschwindigkeiten aus den Kardanwinkelgeschwindigkeiten $\dot{\alpha}_1, \dot{\alpha}_2, \dot{\alpha}_3$ entstehen, welche wiederum nicht senkrecht aufeinander stehen. Zur Ermittlung der Projektionen kann man wieder die Transformationsmatrizen $\mathbf{T}_1(\alpha_1), \mathbf{T}_2(\alpha_2)$ verwenden. Die in Bild 3-22 vorge-

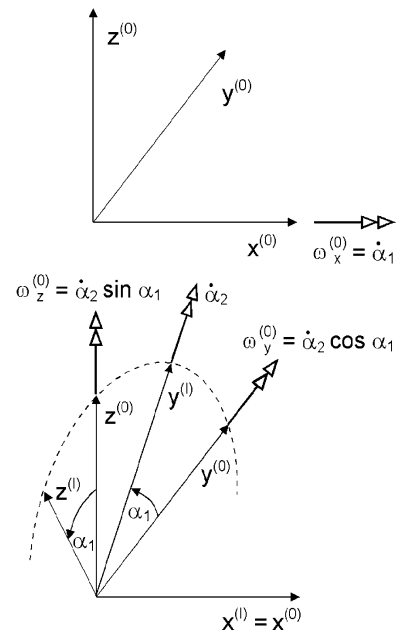


Bild 3-22: Transformation der Winkelgeschwindigkeiten

nommenen Drehungen kann man auch als Vektorgleichung wie folgt in Gl. (3.40) zusammenfassen:

$$\begin{pmatrix} \omega_x^{(0)} \\ \omega_y^{(0)} \\ \omega_z^{(0)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \sin\alpha_2 \\ 0 & \cos\alpha_1 & -\cos\alpha_2 \sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_2 \cos\alpha_1 \end{pmatrix} \cdot \begin{pmatrix} \dot{\alpha}_1 \\ \dot{\alpha}_2 \\ \dot{\alpha}_3 \end{pmatrix},$$

$$\dot{\omega}^{(0)} = \mathbf{C}_0(\alpha_1, \alpha_2) \dot{\alpha} \quad (3.40)$$

Projektion von $\dot{\alpha}_1$

$$\begin{pmatrix} \omega_x^{(0)} \\ \omega_y^{(0)} \\ \omega_z^{(0)} \end{pmatrix} = \begin{pmatrix} \dot{\alpha}_1 \\ 0 \\ 0 \end{pmatrix}$$

Projektion von $\dot{\alpha}_2$

$$\begin{pmatrix} \omega_x^{(0)} \\ \omega_y^{(0)} \\ \omega_z^{(0)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \dot{\alpha}_2 \\ 0 \end{pmatrix}$$

$\underbrace{\hspace{10em}}_{\mathbf{T}_1(\alpha_1)}$

Projektion von $\dot{\alpha}_3$

$$\begin{pmatrix} \omega_x^{(0)} \\ \omega_y^{(0)} \\ \omega_z^{(0)} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_1 \end{pmatrix}}_{\mathbf{T}_1(\alpha_1)} \cdot \underbrace{\begin{pmatrix} \cos\alpha_2 & 0 & \sin\alpha_2 \\ 0 & 1 & 0 \\ -\sin\alpha_2 & 0 & \cos\alpha_2 \end{pmatrix}}_{\mathbf{T}_2(\alpha_2)} \cdot \begin{pmatrix} 0 \\ 0 \\ \dot{\alpha}_3 \end{pmatrix} \quad (\text{s. Bild 3-22})$$

Zwischen den Komponenten des Vektors der Winkelgeschwindigkeit $\omega^{(0)}$ (Gl. (3.40)) und den **Kardanwinkelgeschwindigkeiten** besteht demnach ein linearer Zusammenhang; in der Transformationsmatrix $\mathbf{C}_0$ treten die Winkel α_1 und α_2 allerdings in nichtlinearer Form auf.

Will man die Absolutgeschwindigkeit des Punktes P nicht im Inertialsystem sondern im körperfesten Koordinatensystem darstellen, kann man die Gl. (3.37b) heranziehen. Setzt man in dieser Gleichung für $\mathbf{v}^{(0)}$ den Ausdruck aus Gl. (3.37c) ein, so erhält man:

$$\mathbf{v}^{(1)} = \mathbf{T}^T \cdot \mathbf{v}^{(0)} = \mathbf{T}^T \cdot \dot{\mathbf{s}}^{(0)} + \dot{\mathbf{T}} \cdot \mathbf{T}^T \cdot \mathbf{t}^{(1)}$$

oder abgekürzt

$$\mathbf{v}^{(1)} = \mathbf{T}^T \cdot \mathbf{v}^{(0)} = \mathbf{T}^T \cdot \dot{\mathbf{s}}^{(0)} + \langle \omega^{(1)} \rangle \cdot \mathbf{t}^{(1)} \quad (3.41)$$

Die Abkürzung

$$\langle \omega^{(1)} \rangle = \mathbf{T}^T \cdot \dot{\mathbf{T}}$$

kann man durch Multiplikation der transponierten Transformationsmatrix mit der zeitlichen Ableitung der Transformationsmatrix nach Gl. (3.37e) gewinnen.

$$\langle \omega^{(1)} \rangle = \begin{pmatrix} 0 & -\omega_z^{(1)} & \omega_y^{(1)} \\ \omega_z^{(1)} & 0 & -\omega_x^{(1)} \\ -\omega_y^{(1)} & \omega_x^{(1)} & 0 \end{pmatrix}$$

In dieser antisymmetrischen Matrix stehen die Winkelgeschwindigkeiten für folgende Ausdrücke:

$$\omega_x^{(1)} = \dot{\alpha}_1 \cos \alpha_2 \cos \alpha_3 + \dot{\alpha}_2 \sin \alpha_3 \quad (3.42a)$$

$$\omega_y^{(1)} = -\dot{\alpha}_1 \cos \alpha_2 \sin \alpha_3 + \dot{\alpha}_2 \cos \alpha_3 \quad (3.42b)$$

$$\omega_z^{(1)} = -\dot{\alpha}_1 \sin \alpha_2 + \dot{\alpha}_3 \quad (3.42c)$$

Dies sind die Winkelgeschwindigkeiten des körperfesten Dreibeins, dargestellt im körperfesten Koordinatensystem. Mit diesen Winkelgeschwindigkeiten im körperfesten Koordinatensystem kann man entsprechend Gl. (3.40) wieder den Vektor der Winkelgeschwindigkeit $\omega^{(0)}$ definieren:

$$\begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}^{(1)} = \begin{pmatrix} \cos \alpha_2 \cos \alpha_3 & \sin \alpha_3 & 0 \\ -\cos \alpha_2 \sin \alpha_3 & \cos \alpha_3 & 0 \\ -\sin \alpha_2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \dot{\alpha}_1 \\ \dot{\alpha}_2 \\ \dot{\alpha}_3 \end{pmatrix},$$

$$\omega^{(1)} = \mathbf{C}_1(\alpha_2, \alpha_3) \dot{\alpha} \quad (3.43)$$

Den Zusammenhang zwischen $\langle \omega^{(1)} \rangle$ und $\langle \omega^{(0)} \rangle$ kann man aus den Gleichungen (3.40) und (3.43) gewinnen:

$$\langle \omega^{(1)} \rangle = \mathbf{T}^T \langle \omega^{(0)} \rangle \mathbf{T}, \quad (3.44a)$$

$$\langle \omega^{(0)} \rangle = \mathbf{T} \langle \omega^{(1)} \rangle \mathbf{T}^T \quad (3.44b)$$

Auch die Matrizen $\mathbf{C}_0$ und $\mathbf{C}_1$ können ineinander umgerechnet werden:

$$\mathbf{C}_0 = \mathbf{T} \cdot \mathbf{C}_1 \quad \text{und} \quad \mathbf{C}_1 = \mathbf{T}^T \cdot \mathbf{C}_0 \quad (3.45)$$

Vergleicht man den in Gl. (3.39b) gefundenen Ausdruck für die Absolutgeschwindigkeit im Inertialsystem

$$\mathbf{v}^{(0)} = \dot{\mathbf{r}}^{(0)} = \dot{\mathbf{s}}^{(0)} + \langle \omega^{(0)} \rangle \cdot \mathbf{t}^{(0)}$$

mit dem im Abschnitt 3.2.3.1 gefundenen Ausdruck aus Gl. (3.20) für die allgemeine Bewegung des starren Körpers im Raum

$$\vec{v} = \frac{d \vec{r}_0}{dt} + \vec{\omega} \times \vec{r},$$

so erkennt man die Korrespondenzen der Matrixschreibweise und der koordinatenfreien Schreibweise unter Verwendung physikalischer Vektoren.

Um die **Absolutbeschleunigung** des Punktes P im Inertialsystem zu erhalten, muss man den Ausdruck für die Geschwindigkeit in Gl. (3.37a) nach der Zeit ableiten:

$$\begin{aligned}
\mathbf{a}^{(0)} &= \dot{\mathbf{v}}^{(0)} = \ddot{\mathbf{r}}^{(0)} = \ddot{\mathbf{s}}^{(0)} + \ddot{\mathbf{T}} \cdot \mathbf{t}^{(1)} \\
&= \ddot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(0)} \rangle \mathbf{T} \cdot \mathbf{t}^{(1)} + \langle \omega^{(0)} \rangle \dot{\mathbf{T}} \cdot \mathbf{t}^{(1)} \\
&= \ddot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(0)} \rangle \mathbf{T} \cdot \mathbf{t}^{(1)} + \langle \omega^{(0)} \rangle \langle \omega^{(0)} \rangle \mathbf{T} \cdot \mathbf{t}^{(1)}
\end{aligned} \quad (3.46)$$

Für die Darstellung der Absolutbeschleunigung im körperfesten Koordinatensystem muss $\mathbf{a}^{(0)}$ mit $\mathbf{T}^T$ multipliziert werden

$$\mathbf{a}^{(1)} = \mathbf{T}^T \cdot \mathbf{a}^{(0)} .$$

Die Anwendung dieser Multiplikation auf Gl. (3.46) liefert:

$$\mathbf{a}^{(1)} = \mathbf{T}^T \cdot \ddot{\mathbf{s}}^{(0)} + \mathbf{T}^T \langle \dot{\omega}^{(0)} \rangle \mathbf{T} \cdot \mathbf{t}^{(1)} + \mathbf{T}^T \langle \omega^{(0)} \rangle \langle \omega^{(0)} \rangle \mathbf{T} \cdot \mathbf{t}^{(1)} \quad (3.47)$$

Fügt man zwischen den beiden Matrizen $\langle \omega^{(0)} \rangle$ die Einheitsmatrix $\mathbf{T}^T \mathbf{T} = \mathbf{E}$ ein und verwendet den schon vorher gefundenen Zusammenhang zwischen den Winkelgeschwindigkeiten in den Systemen $^{(0)}$ und $^{(1)}$ aus Gl. (3.44a), sowie den analog gültigen Zusammenhang

$$\langle \dot{\omega}^{(1)} \rangle = \mathbf{T}^T \langle \dot{\omega}^{(0)} \rangle \mathbf{T} , \quad (3.48)$$

so ergibt sich für die Beschleunigung im körperfesten Koordinatensystem

$$\mathbf{a}^{(1)} = \mathbf{T}^T \cdot \ddot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(1)} \rangle \cdot \mathbf{t}^{(1)} + \langle \omega^{(1)} \rangle \langle \omega^{(1)} \rangle \cdot \mathbf{t}^{(1)} \quad (3.49)$$

Die Besetzung der beiden antisymmetrischen Matrizen der Winkelbeschleunigung $\langle \dot{\omega}^{(0)} \rangle$ und $\langle \dot{\omega}^{(1)} \rangle$ kann man durch Ableitung der Vektoren für die Winkelgeschwindigkeiten aus den Gl. (3.40) und (3.43) gewinnen:

$$\langle \dot{\omega}^{(0)} \rangle = \dot{\mathbf{C}}_0(\alpha_1, \alpha_2, \alpha_3) \dot{\alpha} + \mathbf{C}_0(\alpha_1, \alpha_2) \ddot{\alpha} \quad (3.50)$$

$$\langle \dot{\omega}^{(1)} \rangle = \dot{\mathbf{C}}_1(\alpha_1, \alpha_2, \alpha_3) \dot{\alpha} + \mathbf{C}_1(\alpha_2, \alpha_3) \ddot{\alpha} \quad (3.51)$$

Tab. 3-1: Zusammenstellung wichtiger kinematischer Größen

	Matrizenschreibweise		Physikalische Vektorschreibweise
	Inertialsystem (0)	Körperfestes System (1)	
Lagevektor	$\mathbf{r}^{(0)} = \mathbf{s}^{(0)} + \mathbf{T} \mathbf{t}^{(1)}$ $= \mathbf{s}^{(0)} + \mathbf{t}^{(0)}$	$\mathbf{r}^{(1)} = \mathbf{T}^T \mathbf{r}^{(0)}$ $= \mathbf{T}^T \mathbf{s}^{(0)} + \mathbf{t}^{(1)}$	$\vec{r} = \vec{s} + \vec{t}$
Absolute Geschwindigkeit	$\mathbf{v}^{(0)} = \dot{\mathbf{r}}^{(0)}$ $= \dot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(0)} \rangle \mathbf{T} \mathbf{t}^{(1)}$ $= \dot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(0)} \rangle \mathbf{t}^{(0)}$	$\mathbf{v}^{(1)} = \mathbf{T}^T \mathbf{v}^{(0)}$ $= \mathbf{T}^T \dot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(1)} \rangle \mathbf{t}^{(1)}$	$\vec{v} = \dot{\vec{r}}$ $= \dot{\vec{s}} + \vec{\omega} \times \vec{t}$
Absolute Beschleunigung	$\mathbf{a}^{(0)} = \ddot{\mathbf{r}}^{(0)}$ $= \ddot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(0)} \rangle \mathbf{T} \mathbf{t}^{(1)}$ $+ \langle \omega^{(0)} \rangle \langle \omega^{(0)} \rangle \mathbf{T} \mathbf{t}^{(1)}$ $= \ddot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(0)} \rangle \mathbf{t}^{(0)}$ $+ \langle \omega^{(0)} \rangle \langle \omega^{(0)} \rangle \mathbf{t}^{(0)}$	$\mathbf{a}^{(1)} = \mathbf{T}^T \mathbf{a}^{(0)}$ $= \mathbf{T}^T \ddot{\mathbf{s}}^{(0)} + \langle \dot{\omega}^{(1)} \rangle \mathbf{t}^{(1)}$ $+ \langle \omega^{(1)} \rangle \langle \omega^{(1)} \rangle \mathbf{t}^{(1)}$	$\vec{a} = \dot{\vec{v}} = \ddot{\vec{r}}$ $= \ddot{\vec{s}} + \dot{\vec{\omega}} \times \vec{t}$ $+ \vec{\omega} \times (\vec{\omega} \times \vec{t})$

In der Tabelle 3-1 sind nochmals alle Gleichungen über die allgemeine Bewegung eines starren Körpers im Raum in Inertialsystemen und in körperfesten Systemen für die verschiedenen Darstellungsarten zusammengestellt.

3.3 Bindungen in Mehrkörpersystemen

Der durch keine Bindungen gefesselte starre Körper hat im Raum 6 Freiheitsgrade und in der Ebene 3 Freiheitsgrade. Um seine Lage eindeutig zu beschreiben sind daher 6 bzw. 3 Koordinatenangaben erforderlich.

Häufig sind die Bewegungsmöglichkeiten von einer Kette von starren Körpern, wie sie typischerweise in technischen Systemen auftreten, durch Bindungen an vorgegebene Bahnen oder durch Fixierung einzelner Punkte der Kette eingeschränkt. So hat etwa ein Körper, der an einem Punkt festgehalten wird, in der Ebene nur noch einen Freiheitsgrad und im Raum noch drei Freiheitsgrade. Ist der festgehaltene Punkt etwa eine Achse, so kann sich der Körper nur noch um diese drehen, wodurch die Bewegungsmöglichkeiten des Körpers auf eine Ebene festgelegt werden. Ist der festgehaltene Punkt ein Kugelgelenk, so kann der Körper Drehungen um die drei Achsen des Koordinatensystems ausführen. Typische Beispiele für durch Bindungen eingeschränkte Bewegungsmöglichkeiten zweier starrer Körper zeigt Bild 3-23.

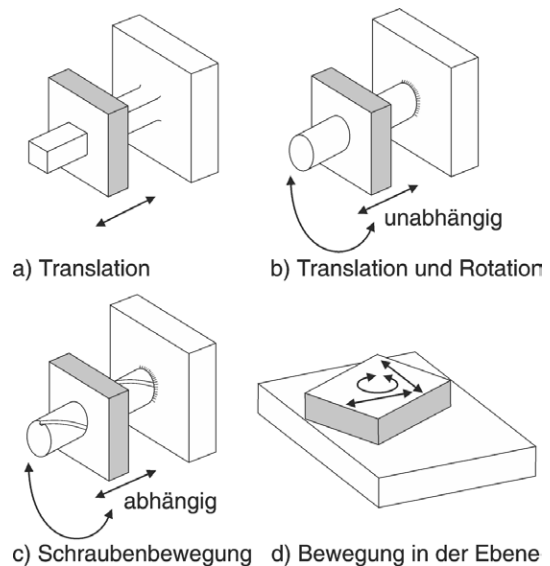


Bild 3-23: Beispiele für eine durch Bindungen reduzierte Anzahl von Freiheitsgraden

Die in Teilbild a) dargestellte Verbindung zweier Starrkörper schränkt die Bewegungsmöglichkeit der auf einer prismatischen Führungsstange beweglichen Platte auf genau einen translatorischen Freiheitsgrad ein, rotatorische Bewegungen werden durch physikalischen Kontakt der beiden Elemente verhindert. Die Anordnung in Teilbild b) besitzt zwei Freiheitsgrade, einen rotatorischen und einen translatorischen. Die Anordnung in Teilbild c) erlaubt zwar die gleichen Bewegungen wie in Teilbild b), die nun aber nicht unabhängig voneinander ausgeführt werden können, da die Schraubenführung eine Zwangskopplung der beiden Bewegungen bewirkt. Deshalb liegt in dieser Anordnung auch nur ein Freiheitsgrad vor. In Teilbild d) wird ein ebener Körper auf einer ebenen Fläche geführt. Er kann daher translatorische Bewegungen in zwei Koordinatenrichtungen in der Ebene ausführen und sich um die dritte Achse eines Dreibeins drehen. Es liegen also drei Freiheitsgrade der Bewegung vor.

Bei einem aus verschiedenen starren Körpern bestehenden **Mehrkörpersystem** kann jeder einzelne Körper solchen Bindungen unterworfen sein, außerdem können

sie untereinander gekoppelt sein. Diese Kopplungen können starr (Gelenke, Stäbe) oder nicht starr (elastische Federn) sein.

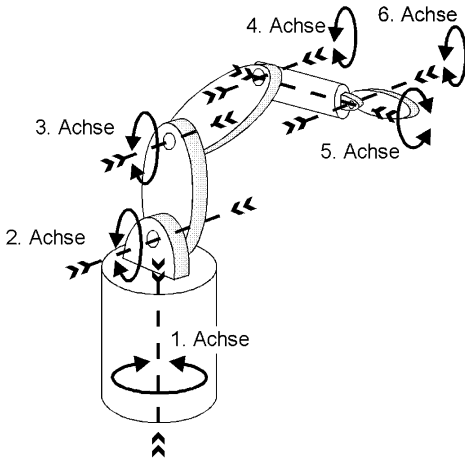


Bild 3-24: 6-achsiger Knickarmroboter mit 6 Freiheitsgraden

Kopplungen, die nicht starr sind, schränken die Anzahl der Freiheitsgrade des Einzelkörpers nicht ein, es wirken aber über die Kopplungselemente Kräfte zwischen den Körpern, die bei Problemen in der Kinetik berücksichtigt werden müssen. Starre Kopplungen (**kinematische Kopplungen**) schränken die Anzahl der Freiheitsgrade ein, da zwischen den Koordinaten, die die Lage der Körper beschreiben, feste Beziehungen, sog. **Zwangsbedingungen**, bestehen. Die Anzahl der Koordinaten, die dann mindestens erforderlich ist, um die Lage eines Systems starrer Körper zu beschreiben, entspricht der Anzahl der Freiheitsgrade des Systems. Besteht, wie in Bild 3-24 dargestellt, ein Roboterarm aus

einer kinematischen Kette von Einzelkörpern (Gliedern), die jeweils durch ein Drehgelenk miteinander gekoppelt sind, so hat jeder Einzelkörper aufgrund der Einschränkung der Bewegungsmöglichkeit auf eine Rotation nur einen Freiheitsgrad. Die gesamte kinematische Kette hat dann zusammen wieder sechs Freiheitsgrade, wodurch der Roboterarm ein Objekt innerhalb seines Arbeitsbereichs an eine beliebige Position mit beliebiger Orientierung im Raum bewegen kann.

Bindungen zwischen den Körpern eines Systems lassen sich in der Regel als implizite Vektorfunktion der Lagekoordinaten und der Zeit darstellen:

$$\Phi(x, y, z, \varphi_x, \varphi_y, \varphi_z, t) \quad (3.52)$$

Die entsprechende Vektorfunktion für ein Fadenpendel mit einer Punktmasse m und der Fadenlänge r , das sich nur in der (x, y) -Ebene auf einer Kreisbahn bewegen kann und keine Drehung um eine Koordinatenachse ausführen kann, wäre dann:

$$\Phi = \begin{pmatrix} x^2 + y^2 - r^2 \\ z \end{pmatrix} = 0 \quad .$$

Bindungen in mechanischen Systemen werden nach ihren Eigenschaften eingeteilt. Ist die Bindung eines Systems wie oben gezeigt darstellbar, so spricht man von **holonomen** Bindungen (gr.: ganz gesetzlich). Da in dieser Vektorfunktion die Zeit nicht vorkommt, wird diese Art der Bindung zusätzlich als **skleronom** (gr.: starr) bezeichnet. Ist bei dem erwähnten Beispiel des Pendels die Fadenlänge nach einem bekannten Zeitgesetz von der Zeit abhängig $r = r(t)$, so tritt in der Bindungsgleichung wieder die Zeit auf und man spricht von **rheonomer** (gr.: fließend) Bindung.

Nichtholonome Bindungen gelten z. B. für ein auf einer Ebene rollendes Rad, bei dem die Bindungen nur in nicht integrierbarer Form angebar sind. Seine Bindungsgleichung ist außer von den Koordinaten und der Zeit auch noch von den Ab-

leitungen der Koordinaten nach der Zeit, also von Geschwindigkeiten, abhängig. Hier sind die Bindungen nur für infinitesimal kleine Zuwächse der Koordinaten gültig, aber nicht für die ganze Bewegung.

Hat man eine kinematische Kette von n Starrkörpern, so kann man den Bewegungszustand des n -ten Starrkörpers so beschreiben, dass man entsprechend der Vorgehensweise in Abschnitt 3.2.3 jedem Teilkörper ein eigenes Koordinatensystem zuordnet, und dann rekursiv von der Basis der Körperkette beginnend den Bewegungszustand vom körperfesten Koordinatensystem ins Inertialsystem transformiert. Da die Bindungen in technischen, kinematischen Ketten nur bestimmte Bewegungsmöglichkeiten zulassen, gibt es bestimmte Verfahren, die den Transformationsaufwand minimieren. So wurde von Denavit und Hartenberg 1955 eine Methode vorgeschlagen, die für die bei Roboterarmen typische Aufgabe der Transformation der Koordinaten des aktiven Punktes des Greifsystems (**TCP**)¹ in die so genannten **Weltkoordinaten** verwendet wird [3.5]. Dabei handelt es sich um ein Inertialsystem, das man in der Regel in der Basis der kinematischen Kette, wo der Arm mit dem Boden verbunden ist, anordnet (Bild 3-25). Die Position des TCP erhält man dann (siehe 3.2.3.3) durch einen Ortsvektor $\vec{r}$, der die Verschiebung des körperfesten ⁽¹⁾ gegenüber dem ortsfesten Koordinatensystem ⁽⁰⁾ angibt und eine Transformationsmatrix **T**, die seine Orientierung, d. h. die Drehungen des körperfesten gegenüber dem ortsfesten Koordinatensystem enthält.

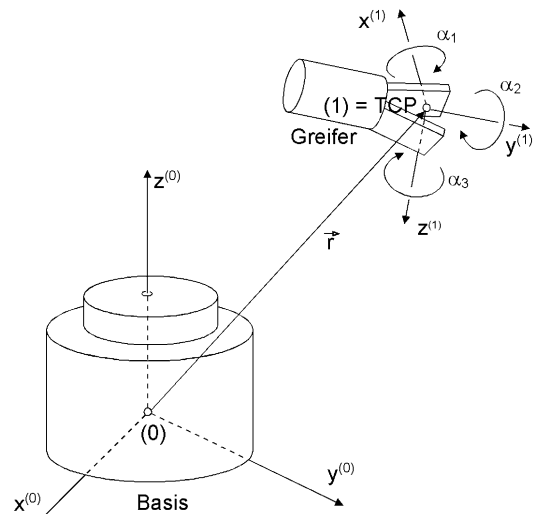


Bild 3-25: Angabe von Position und Orientierung des Greifers eines Roboterarms

Die einzelnen Bewegungsachsen, die zwischen der Basis und dem Greifer des Roboterarms liegen, sind in der Regel unabhängig voneinander und besitzen entweder einen rotatorischen oder einen translatorischen Freiheitsgrad, dem eine **ge-lenknatürliche** oder **verallgemeinerte** Koordinate zugeordnet wird. Diese verallgemeinerten Koordinaten werden zur mathematischen Darstellung in einem Lagevektor zusammengefasst. Die Anzahl der Komponenten dieses Lagevektors entspricht der Anzahl der Freiheitsgrade des Systems. Besitzt der Roboterarm wie in Bild 3-24 dargestellt 6 hintereinanderliegende Rotationsachsen, so enthält der Lagevektor als Komponenten 6 Winkel:

$$\vec{y} = (\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6)$$

Bringt man im Massenmittelpunkt jedes Teilkörpers i ein körperfestes Koordinatensystem an, so wird die Lage eines Teilkörpers i des Roboterarms durch den Ortsvektor

1 TCP = Tool Center Point (Werkzeugmittelpunkt).

$$\vec{r}_i = \vec{r}_i(\vec{y}, t)$$

zum Massenmittelpunkt und durch eine Drehungsmatrix

$$\mathbf{T}_i = \mathbf{T}_i(\vec{y}, t)$$

eindeutig festgelegt. Durch $\mathbf{T}_i$ wird der Zusammenhang zwischen dem Inertialsystem und dem jeweiligen körperfesten System beschrieben.

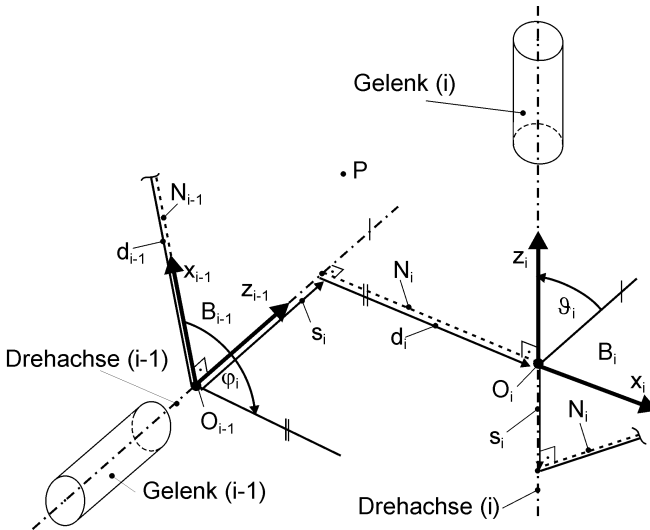


Bild 3-26: Basissysteme und Beschreibungsgrößen in der kinematischen Kette für die Denavit-Hartenberg-Methode

Bei der von **Denavit - Hartenberg** vorgeschlagenen Methode zur Bestimmung der Lage von Teilkörpern eines Roboterarms in Weltkoordinaten (Inertialsystem), geht man von einer besonderen Anordnung und Orientierung der Koordinatensysteme der Teilkörper aus [3.6], [3.7], [3.8]. In Bild 3-26 ist ein Ausschnitt aus einer kinematischen Kette mit zwei beliebig windschief angeordneten Gelenken dargestellt. Die in der Regel bei technischen kinematischen Ketten wie Industrierobotern vorwiegend anzutreffende rechtwinklig gekreuzte oder parallele Anordnung von

Achsen ist in dieser Annahme enthalten. Die Gelenke können Dreh- oder auch Schubachsen sein, der Übersichtlichkeit halber sind nur Drehachsen dargestellt. Für die Beschreibung der kinematischen Abmessungen müssen, wie in Bild 3-26 gezeigt, die Kreuzungswinkel φ und ϑ , sowie die Kreuzungsabstände s und d bekannt sein. Drei dieser Maße sind dabei in der Regel durch die konstruktiv bedingte geometrische Anordnung vorgegeben, das vierte Maß wird durch Messung mit einem Wegmesssystem ermittelt.

Die mit Gelenken verbundenen Teilkörper werden von der Basis aus beginnend bis zum n -ten Teilkörper durchnummeriert. Der kürzeste Abstand zwischen den Drehachsen $i-1$ und i ist die **Verbindungsnormale** N_i . Der i -te Teilkörper der kinematischen Kette wird mit einem körperfesten, rechtsdrehenden Koordinatensystem in folgender Anordnung und Orientierung versehen (Bild 3-26) :

- 1.) Der Ursprung O_i der Basis B_i ist der Schnittpunkt der Verbindungsnormalen N_i mit der Drehachse i .
- 2.) Die z -Achse der Basis B_i liegt in der Drehachse i (Orientierung beliebig).
- 3.) Die x -Achse der Basis B_i liegt in der Verbindungsnormalen N_i (Orientierung beliebig).

4.) Die y -Achse ist so zu wählen, dass B_i ein Rechtssystem bildet.

In Bild 3-26 ist die y -Achse aus Gründen der Übersichtlichkeit weggelassen. Hat man zwei Drehachsen $i-1$ und i , so gibt das Maß s_i den Abstand zwischen dem Ursprung O_{i-1} und dem Fußpunkt der Verbindungsnormalen N_i , gemessen in Richtung von z_{i-1} , an. Das Maß d_i ist der Abstand der beiden Drehachsen $i-1$ und i , gemessen entlang der Verbindungsnormalen N_i . Der Winkel φ_i ist der Winkel zwischen den Richtungen der x -Achsen in den Lagen $i-1$ und i . Dabei wird φ positiv gemessen, wenn x_{i-1} in die neue Lage x_i um z_{i-1} im Sinne einer Rechtschraube gedreht wird. Der Winkel ϑ_i ist der Winkel zwischen den Richtungen der z -Achsen in den Lagen $i-1$ und i . Dabei wird ϑ positiv gemessen, wenn in die neue Lage z_i um x_i im Sinne einer Rechtschraube gedreht wird.

Man kann sich den Vorgang der Transformation auch als Verlagerung der Basis B_{i-1} in die Lage von B_i und umgekehrt durch einen vierschrittigen Vorgang vorstellen. Zuerst wird die Basis B_{i-1} entlang der Drehachse $i-1$ in Richtung von z_{i-1} um s_i bis zur Verbindungsnormalen N_i verschoben. Dann wird die Basis B_{i-1} um den Winkel φ_i so um z_{i-1} gedreht, dass die x -Achse in Richtung der Verbindungsnormalen N_i zeigt. Danach wird der Ursprung O_{i-1} der Basis entlang der Verbindungsnormalen um das Maß d_i bis zum Ursprung O_i der Basis B_i verschoben. Die abschließende Drehung der Basis B_{i-1} um ihre x -Achse um den Winkel ϑ_i überführt diese in die gleiche Lage wie die Basis B_i .

Mit Hilfe der vier Größen s , φ , d , ϑ lassen sich Bewegungen von Punkten, deren Koordinaten im System i gegeben sind, im Koordinatensystem $i-1$ beschreiben und umgekehrt. Die Vorgehensweise bei der Koordinatentransformation soll hier am Beispiel einer Transformation vom System (2) in das System (1) dargestellt werden (Bild 3-27).

Der Punkt P habe im System (2) den Ortsvektor $\vec{r}_{2P}$ mit den bekannten Koordinaten x_{2P} , y_{2P} , z_{2P} und es wird der Ortsvektor $\vec{r}_{1P}$ mit den Koordinaten x_{1P} , y_{1P} , z_{1P} im System (1) gesucht. Die Koordinaten können als lineares Gleichungssystem geschrieben werden, dem aus Gründen der Homogenisierung noch die Identitätsgleichung $1 = 1$ vorangestellt wird. Der erste Schritt der Transformation, die Parallelverschiebung von B_1 mit dem Ursprung O_1 um den Abstand s_2 in die neue Lage der Basis mit O_{21} als Ursprung, liefert das Gleichungssystem Gl. (3.53) oder die Matrixschreibweise Gl. (3.54).

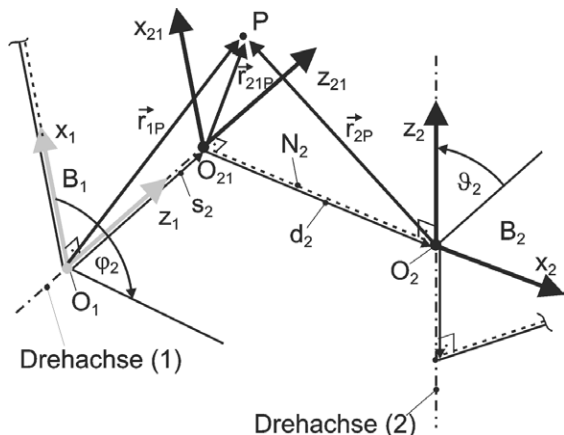


Bild 3-27: Verschiebung der Basis B_1 um das Maß s_2

$$1 = 1$$

$$x_1 = x_{21} \quad (3.53)$$

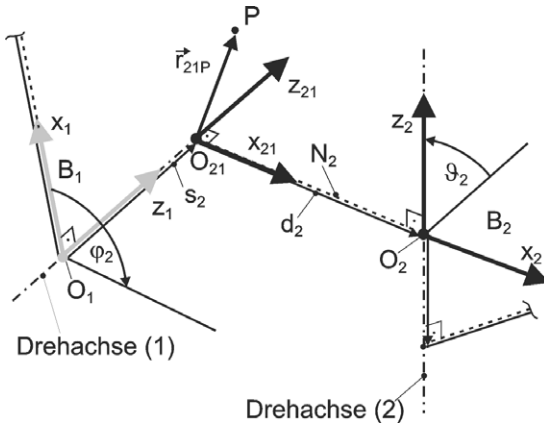
$$y_1 = y_{21}$$

$$z_1 = z_{21} + s_2$$

$$\begin{pmatrix} 1 \\ x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ s_2 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x_{21} \\ y_{21} \\ z_{21} \end{pmatrix} \quad (3.54)$$

Die Vektorschreibweise erhält man aus den Ortsvektoren $\vec{r}_{1P}$ des Punktes P bezogen auf B_1 und $\vec{r}_{21P}$ bezogen auf die um s_2 verschobene Basis B_{21} , sowie der Koeffizientenmatrix der Verschiebung $\mathbf{T}_{21}^s$:

$$\vec{r}_{1P} = \mathbf{T}_{21}^s \cdot \vec{r}_{21P} \quad (3.55)$$



Im nächsten Schritt erhält man für die Drehung der Basis B_{21} um die z -Achse im Sinne einer Rechtsschraube um den Winkel φ_2 (Bild 3-28) die Koeffizientenmatrix $\mathbf{T}_{21}^\varphi$ entsprechend Gl. (3.32):

$$\mathbf{T}_{21}^\varphi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \varphi_2 & -\sin \varphi_2 & 0 \\ 0 & \sin \varphi_2 & \cos \varphi_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.56)$$

Bild 3-28: Drehung der Basis B_{21} um den Winkel φ_2

Die sich anschließende Parallelverschiebung des Basisursprungs O_{21} entlang der Verbindungsnormalen N_2

nach O_2 um das Maß d_2 liefert:

$$\mathbf{T}_{21}^d = \begin{pmatrix} 1 & 0 & 0 & 0 \\ d_2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.57)$$

und schließlich die Drehung der Basis in O_2 um den Winkel ϑ_2 im Sinne einer Rechtsschraube um die x -Achse in die endgültige Lage der Basis B_2 :

$$\mathbf{T}_{21}^\vartheta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \vartheta_2 & -\sin \vartheta_2 \\ 0 & 0 & \sin \vartheta_2 & \cos \vartheta_2 \end{pmatrix} \quad (3.58)$$

Durch Multiplikation aller vier Koeffizientenmatrizen entsteht die Transformationsmatrix $\mathbf{T}_{21}$ für die Gesamttransformation der Basis B_2 in die Basis B_1 :

$$\mathbf{T}_{21} = \mathbf{T}_{21}^s \cdot \mathbf{T}_{21}^\varphi \cdot \mathbf{T}_{21}^d \cdot \mathbf{T}_{21}^\vartheta = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline d_2 \cos \varphi_2 & \cos \varphi_2 & -\sin \varphi_2 \cos \vartheta_2 & \sin \varphi_2 \sin \vartheta_2 \\ d_2 \sin \varphi_2 & \sin \varphi_2 & \cos \varphi_2 \cos \vartheta_2 & -\cos \varphi_2 \sin \vartheta_2 \\ s_2 & 0 & \sin \vartheta_2 & \cos \vartheta_2 \end{array} \right) \quad (3.59)$$

Diese Transformationsmatrix enthält in der ersten Spalte die Information über die Verschiebung der Basis um die Maße s und d und im übrigen Bereich die Verdrehung der Basis um die beiden Eulerwinkel φ und ϑ . Damit lässt sich die Transformation des Ortsvektors des Punktes P folgendermaßen schreiben:

$$\vec{r}_{1P} = \mathbf{T}_{21} \cdot \vec{r}_{2P} \quad (3.60)$$

Zur Rücktransformation

$$\vec{r}_{2P} = \mathbf{T}_{21}^{-1} \cdot \vec{r}_{1P} \quad (3.61)$$

ist die inverse Matrix $\mathbf{T}_{21}^{-1}$ erforderlich, für die gilt:

$$\mathbf{T}_{21} \cdot \mathbf{T}_{21}^{-1} = \mathbf{E} \quad .$$

Die inverse Matrix hat folgendes Aussehen:

$$\mathbf{T}_{21}^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline -d_2 & \cos \varphi_2 & \sin \varphi_2 & 0 \\ -s_2 \sin \vartheta_2 & -\sin \varphi_2 \cos \vartheta_2 & \cos \varphi_2 \cos \vartheta_2 & \sin \vartheta_2 \\ -s_2 \cos \vartheta_2 & \sin \varphi_2 \sin \vartheta_2 & -\cos \varphi_2 \sin \vartheta_2 & \cos \vartheta_2 \end{array} \right) \quad (3.62)$$

Hat man eine Kette von Teilkörpern, aus denen der Roboterarm besteht, so erfolgt die Transformation durch Multiplikation der einzelnen Transformationsmatrizen, mit denen man Koordinaten von einem Teilkörper zum anderen transformieren kann. Beispielsweise kann man den TCP, der bei einem Arm aus fünf aneinandergekoppelten Teilkörpern mit dem Ortsvektor $\vec{r}_{5P}$ im zugehörigen körperfesten Koordinatensystem (5) gegeben ist, durch folgende Operation in das raumfeste Basissystem (1) transformieren:

$$\vec{r}_{1P} = \mathbf{T}_{21} \cdot \mathbf{T}_{32} \cdot \mathbf{T}_{43} \cdot \mathbf{T}_{54} \cdot \vec{r}_{5P}$$

Da die Bestimmung des Matrizenproduktes sehr rechenaufwendig ist, sollte man dieses sinnvollerweise nur mit einer rechnergestützten Methode berechnen.

Entsprechend der oben dargestellten Vorgehensweise und wie in Abschnitt 3.2.3.3 dargestellt, lassen sich dann auch Geschwindigkeiten, Winkelgeschwindigkeiten und Beschleunigungen transformieren.

Abschließend soll die Methode der Denavit-Hartenberg Transformation nochmals an einer einfachen Roboterkinematik aus drei Teilkörpern mit drei Drehachsen beispielhaft erläutert werden (Bild 3-29).

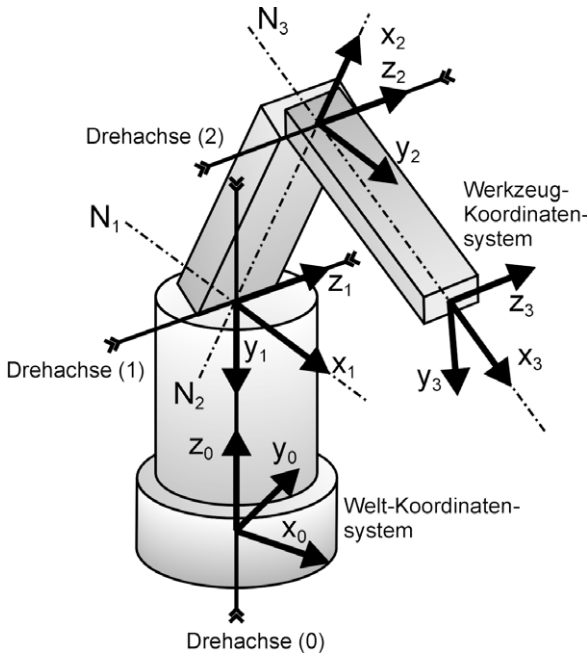


Bild 3-29: 3-achsige Roboterkinematik mit den nach der Denavit-Hartenberg Methode orientierten Koordinatensystemen

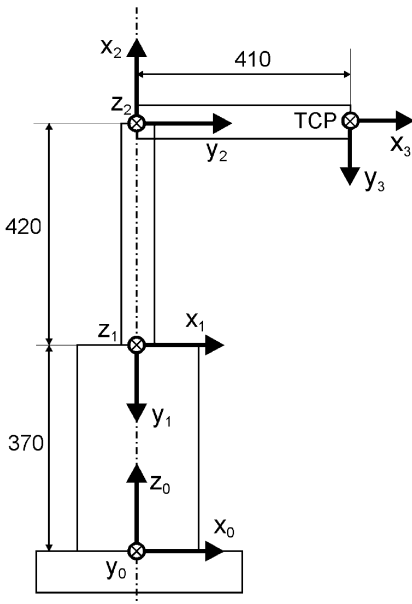


Bild 3-30: Spezielle Anordnung eines Roboterarms mit drei Drehachsen

Der Roboter besteht aus einem Fuß und einer Säule mit senkrecht stehender Drehachse (0). Daran schließen sich zwei prismaische Armteile an, deren Drehachsen (1) und (2) senkrecht zur Drehachse (0) stehen und waagrecht zur Aufstandsfläche angeordnet sind. Um die zu jeder Drehachse gehörende Basis für die Denavit-Hartenberg Methode (DHM) einzeichnen zu können, muss man zuerst die Verbindungsnormalen der Drehachsen festlegen. Die Verbindungsnormale N_1 zwischen Drehachse (0) und Drehachse (1) geht beispielsweise durch den Schnittpunkt der beiden Drehachsen und steht auf beiden Drehachsen senkrecht. Da die beiden Drehachsen (1) und (2) parallel sind und sich an den jeweiligen Enden des ersten prismaischen Gliedes befinden, läuft die Verbindungsnormale N_2 parallel

zu den Längskanten des Bauteils. Da sich am Anfang und am Ende der kinematischen Kette kein Glied davor oder dahinter anschließt und es daher dorthin auch keine Verbindungsnormale gibt, ist die Lage der zugehörigen Basen dort nicht eindeutig festgelegt. Die Basis B_0 zwischen Fuß und Säule, das sogenannte Welt-Koordinatensystem, hat entsprechend den Regeln der DHM eine Orientierung der z-Achse in Richtung der Drehachse (0). Danach müsste die x-Achse in Richtung der Verbindungsnormalen zur davorliegenden Drehachse orientiert werden, ist aber wegen des Fehlens frei wählbar. Die y-Achse wird dann entsprechend den Regeln für ein rechtsdrehendes Koordinatensystem dazu fügt.

Die Basis B_1 hat die z-Achse in Richtung der Drehachse (1) (Orientierung beliebig), die x-Achse zeigt in Richtung der Verbindungsnormalen N_1 . Bei Drehachse (2) erfolgt die Lagebestimmung entsprechend. Am Ende der kinematischen Kette befindet sich das Werkzeug- oder

Tool-Koordinatensystem mit dem Ursprung im TCP. Da sich hier keine Drehachse befindet, kann die Richtung der z-Achse frei gewählt werden, wodurch die Richtung der Verbindungsnormalen zur Drehachse (2) festgelegt wird. Dadurch ist dann auch die Richtung der x-Achse mit beliebiger Orientierung festgelegt und somit auch die Richtung der y-Achse.

Es soll nun mit Hilfe der Transformationsgleichungen Gl. (3.60) und Gl. (3.61) für eine bestimmte Lage des Roboterarms (Bild 3-30) die Aufgabenstellung der Vorwärts- und der Rückwärtstransformation durchgeführt werden. In dieser Lage ist das letzte prismatische Glied um 90° gegenüber dem davorliegenden Glied abgewinkelt, Drehachse (0) und Verbindungsnormale N_2 fluchten und der Winkel zwischen x_0 und x_3 ist 0° .

Die **Vorwärtstransformation** wird benötigt, um die Lage des TCP im Weltkoordinatensystem angeben zu können. Dies wird bei einem Industrieroboter beispielsweise für die Positionsanzeige des Gerätes benötigt.

Die Koordinatensysteme sind wie in Bild 3-29 orientiert. Der Ortsvektor des TCP bezogen auf die Basis B_3 beträgt:

$$\vec{r}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

da der TCP sich definitionsgemäß im Nullpunkt der Basis B_3 am Ende der kinematischen Kette befindet.

Gesucht wird der Ortsvektor des TCP bezogen auf die Basis B_0 :

$$\vec{r}_0 = \begin{pmatrix} 1 \\ x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 410 \\ 0 \\ 790 \end{pmatrix}.$$

Wegen der sehr einfachen räumlichen Anordnung kann man diesen Ortsvektor natürlich auch wie oben ausgeführt direkt bestimmen.

Nach der DHM muss unter Verwendung der Transformationsmatrix aus Gl. (3.59) der Ortsvektor von Basis zu Basis transformiert werden. Zur Berechnung der jeweiligen Transformationsmatrizen benötigt man die Maße des Roboterarms, die man Bild 3-30 entnehmen kann. Diese betragen für den ersten Transformationsschritt von B_3 nach B_2 :

$$s_3 = 0 \text{ mm} ; \varphi_3 = 90^\circ ; d_3 = 410 \text{ mm} ; \vartheta_3 = 0^\circ.$$

Dabei hat s_3 den Wert Null, weil die Verbindungsnormale N_3 durch die Basis B_2 geht, d_3 ist der Abstand von B_2 nach B_3 gemessen entlang der Verbindungsnormalen, der Wert von φ_3 ergibt sich bei Drehung von x_2 in die Lage von x_3 im Sinne einer Rechtsschraube um die Achse z_2 und ϑ_3 ist Null, da z_2 und z_3 parallel und

gleichorientiert sind. Setzt man diese Werte in die Transformationsmatrix ein, so erhält man folgende Gleichung für den Ortsvektor des TCP bezogen auf die Basis B_2 :

$$\vec{r}_2 = \mathbf{T}_{32} \cdot \vec{r}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 410 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 410 \\ 0 \end{pmatrix}.$$

Von der Richtigkeit des Ortsvektors $\vec{r}_2$, in dem nur die y-Komponente von Null verschieden und gleich 410 mm ist, kann man sich leicht anhand von Bild 3-30 überzeugen.

Für den nächsten Transformationsschritt betragen die Maße in der Transformationsmatrix:

$$s_2 = 0 \text{ mm} ; \varphi_2 = 270^\circ ; d_2 = 420 \text{ mm} ; \vartheta_2 = 0^\circ.$$

Damit gilt:

$$\vec{r}_1 = \mathbf{T}_{21} \cdot \vec{r}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -420 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 410 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 410 \\ -420 \\ 0 \end{pmatrix}.$$

Für den letzten Transformationsschritt betragen die Maße in der Transformationsmatrix:

$$s_1 = 370 \text{ mm} ; \varphi_1 = 0^\circ ; d_1 = 0 \text{ mm} ; \vartheta_1 = 270^\circ.$$

Dies ergibt:

$$\vec{r}_0 = \mathbf{T}_{10} \cdot \vec{r}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 370 & 0 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 410 \\ -420 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 410 \\ 0 \\ 790 \end{pmatrix}.$$

Der letzte Transformationsschritt liefert also tatsächlich die Position des TCP im Welt-Koordinatensystem, die auch schon aus der einfachen Geometrie der Anordnung gefolgert werden konnte.

Die inverse oder **Rücktransformation** wird benötigt, wenn die Position des TCP im Weltkoordinatensystem vorgegeben wird und für das Anfahren der Position durch den Roboterarm die Stellungen der einzelnen Gelenke ermittelt werden müssen. Bei dieser Aufgabenstellung ist also der Ortsvektor $\vec{r}_0$ des TCP im Weltkoordinatensystem B_0 bekannt.

Gesucht wird jedoch nicht der Ortsvektor $\vec{r}_3$ des TCP im Bezug auf die Basis B_3 , sondern diejenigen Gelenkwinkel φ_1 , φ_2 , φ_3 , die die Drehachsen (0), (1) und (2) annehmen müssen, damit der TCP die geforderte Position im Raum erreicht. Dazu

muss wieder in drei Schritten mit Hilfe der inversen Transformationsmatrix Gl. (3.62) der Ortsvektor des TCP von Basis nach Basis rücktransformiert werden:

$$\vec{r}_1 = \mathbf{T}_{10}^{-1} \cdot \vec{r}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi_1 & \sin\varphi_1 & 0 \\ 370 & 0 & 0 & -1 \\ 0 & -\sin\varphi_1 & \cos\varphi_1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 410 \\ 0 \\ 790 \end{pmatrix} = \begin{pmatrix} 1 \\ 410\cos\varphi_1 \\ -420 \\ -410\sin\varphi_1 \end{pmatrix}.$$

Dieser Ortsvektor ist gleich dem Ortsvektor $\vec{r}_1$ bei der Vorwärtstransformation für den Fall $\varphi_1 = 0^\circ$. Entsprechend erhält man den Ortsvektor $\vec{r}_2$:

$$\begin{aligned} \vec{r}_2 = \mathbf{T}_{21}^{-1} \cdot \vec{r}_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ -420 & \cos\varphi_2 & \sin\varphi_2 & 0 \\ 0 & -\sin\varphi_2 & \cos\varphi_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 410\cos\varphi_1 \\ -420 \\ -410\sin\varphi_1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ -420 + 410\cos\varphi_1\cos\varphi_2 - 420\sin\varphi_2 \\ -410\cos\varphi_1\sin\varphi_2 - 420\cos\varphi_2 \\ -410\sin\varphi_1 \end{pmatrix}. \end{aligned}$$

Auch dieser Ortsvektor ist gleich dem Ortsvektor $\vec{r}_2$ der Vorwärtstransformation für den Fall $\varphi_1 = 0^\circ$, $\varphi_2 = 270^\circ$. Abschließend erhält man den Ortsvektor $\vec{r}_3$:

$$\begin{aligned} \vec{r}_3 = \mathbf{T}_{32}^{-1} \cdot \vec{r}_2 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ -420 & \cos\varphi_2 & \sin\varphi_2 & 0 \\ 0 & -\sin\varphi_2 & \cos\varphi_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -420 + 410\cos\varphi_1\cos\varphi_2 - 420\sin\varphi_2 \\ -410\cos\varphi_1\sin\varphi_2 - 420\cos\varphi_2 \\ -410\sin\varphi_1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ -410 - 420c\varphi_3 + 410c\varphi_1c\varphi_2c\varphi_3 - 420s\varphi_2c\varphi_3 - 410c\varphi_1s\varphi_2s\varphi_3 - 420c\varphi_2s\varphi_3 \\ 420s\varphi_3 - 410c\varphi_1c\varphi_2s\varphi_3 + 420s\varphi_2s\varphi_3 - 410c\varphi_1s\varphi_2c\varphi_3 - 420c\varphi_2c\varphi_3 \\ -410\sin\varphi_1 \end{pmatrix} \end{aligned}$$

In der Schreibweise des letzten Vektors wird aus Platzgründen s anstelle von sin und c anstelle von cos verwendet. Dies ist der gleiche Ortsvektor wie bei der Vorwärtstransformation für den Fall $\varphi_1 = 0^\circ$, $\varphi_2 = 270^\circ$, $\varphi_3 = 90^\circ$. Da die x,y,z-Komponenten des Ortsvektors des TCP gleich Null sind, liefert die letzte Gleichung ein Gleichungssystem mit drei Gleichungen in den drei Unbekannten φ_1 , φ_2 , φ_3 . Dieses ist lösbar aber nicht eindeutig, da wegen der Periodizität trigonometrischer Funktionen Mehrdeutigkeiten existieren. Betrachtet man wie in Bild 3-31 dargestellt den Fall $\varphi_1 = 0^\circ$, so erhält man ein Gleichungssystem aus 2 Gleichungen für die Winkel und φ_2 und φ_3 :

$$\begin{aligned} 1 + \frac{420}{410}\cos\varphi_3 - \cos\varphi_2\cos\varphi_3 + \frac{420}{410}\sin\varphi_2\cos\varphi_3 + \sin\varphi_2\sin\varphi_3 + \frac{420}{410}\cos\varphi_2\sin\varphi_3 &= 0 \\ -\frac{420}{410}\sin\varphi_3 + \frac{420}{410}\cos\varphi_2\cos\varphi_3 + \sin\varphi_2\cos\varphi_3 - \frac{420}{410}\sin\varphi_2\sin\varphi_3 + \cos\varphi_2\sin\varphi_3 &= 0 \end{aligned}$$

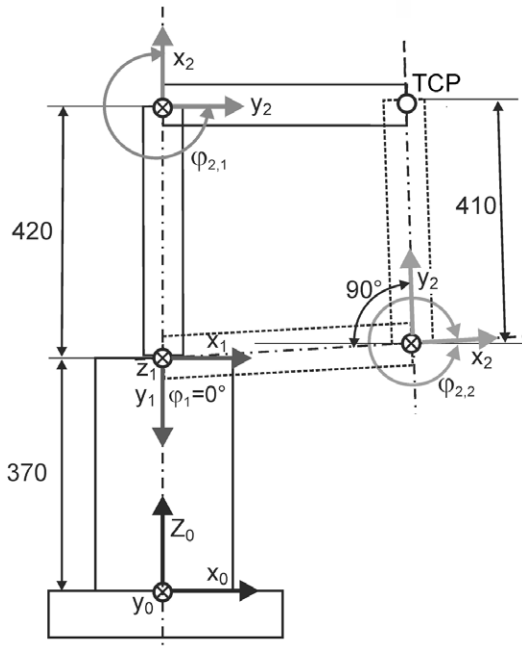


Bild 3-31: Roboterarm mit zwei möglichen Armstellungen bei gleicher Position des TCP

Diese beiden Gleichungen haben, wie in Bild 3-31 dargestellt, zwei Lösungen für die Winkel φ_2 und φ_3 . Davon ist die eine die schon behandelte Lösung $\varphi_{2,1} = 270^\circ$, $\varphi_{3,1} = 90^\circ$ und die andere $\varphi_{2,2} \approx 358^\circ$, $\varphi_{3,2} \approx 272^\circ$, d. h. das erste prismatische Armglied liegt annähernd waagerecht und das zweite steht annähernd senkrecht. Ohne zusätzliche Informationen ist daher mit Hilfe der Rücktransformation die Aufgabenstellung, die Gelenkwinkel für eine bestimmte Position des TCP zu ermitteln, nicht eindeutig lösbar. Bei Roboterarmen wird diese Information bei einer wie in Bild 3-31 gezeigten Anordnung in Form der **Konfiguration** mitgespeichert, wenn die Position dem Roboter durch Teachin mitgeteilt wird. Bei dem gezeigten Vertikal-Knickarm ist dies eine 1-Bit Information der Art "Flipped" oder "Nonflipped". Beim Fahren zwischen verschiedenen Positionen wird dann versucht diese Konfiguration beizubehalten. Wird daher den Servokrei-

sen die Information

$$(\varphi_1, \varphi_2, \varphi_3, [\text{conf}]) = (0^\circ, 270^\circ, 90^\circ, [\text{NF}])$$

vorgegeben, so fährt der Roboter in die im Bild 3-31 durchgezogen dargestellte Stellung, ansonsten in die gestrichelt dargestellte Stellung.

3.4 Kinetik

Während die Kinematik die Bewegungsabläufe ohne Frage nach ihren Ursachen (Kräfte, Momente) untersucht, wird in der Kinetik der Zusammenhang zwischen den kinematischen Größen (Weg, Zeit, Geschwindigkeit, Winkelgeschwindigkeit, Beschleunigung) und den Kräften bzw. Momenten behandelt. Die Weg-Zeit- und Geschwindigkeits-Zeit-Gesetze für bewegte Körper werden durch verschiedene Methoden aus den meist als Differentialgleichung gegebenen Bewegungsbedingungen hergeleitet. Je nach Komplexitätsgrad des Bewegungsproblems sind verschiedene im folgenden behandelte Ansätze möglich.

3.4.1 Impuls-, Schwerpunkt- und Drallsatz

Die Anfänge der Mechanik gründen sich auf die drei Newton'schen Axiome, die der englische Mathematiker Isaac Newton (1642-1727) als erster in dieser Form formu-

lierte. Sie lauten:

- 1.) Die Bewegungsgröße (Impuls) ändert sich nicht, wenn keine Kraft auf eine Masse einwirkt.
- 2.) Die zeitliche Änderung der Bewegungsgröße ist der einwirkenden Kraft proportional und geschieht in Richtung dieser Kraft.
- 3.) Zu jeder Aktion gehört eine gleich große Reaktion (actio = reactio), oder: Die gegenseitigen Wirkungen zweier Körper aufeinander sind stets gleich groß und von entgegengesetzter Wirkung.

Während das 3. Axiom schon für die hier nicht behandelte Statik von Bedeutung ist und das 1. Axiom einen Sonderfall des 2. Axioms darstellt, hat das 2. Axiom für die Kinetik die entscheidende Bedeutung. Betrachtet man den einfachen Fall, dass nur Kräfte und keine Momente auf den Körper einwirken, so lautet dieses Axiom als Gleichung ausgedrückt:

$$\frac{d}{dt}(m \cdot \vec{v}) = \vec{F} \quad (3.63)$$

oder als Zeitintegral der Kraft

$$\int_{t_0}^{t_1} \vec{F} dt = (m \cdot \vec{v})_1 - (m \cdot \vec{v})_0 = \vec{p}_1 - \vec{p}_0$$

mit dem Vektor des **Impulses**

$$\vec{p} = m \cdot \vec{v}$$

Dies bedeutet, dass das Integral über der Kraftwirkung die Impulsänderung des Körpers durch die einwirkende Kraft ergibt, weshalb das Axiom auch als **Impulsatz** bezeichnet wird.

Wenn die Masse m zeitlich unveränderlich ist, vereinfacht sich die Gl. (3.63) zu der gebräuchlichen Form des Newton'schen Axioms

$$\vec{F} = m \cdot \frac{d(\vec{v})}{dt} = m \cdot \vec{a} = m \cdot \ddot{\vec{r}} \quad (3.64)$$

Häufig wird dieses Gesetz auch folgendermaßen formuliert:

$$\vec{F} - m \cdot \ddot{\vec{r}} = 0 \quad \text{mit} \quad -m \cdot \ddot{\vec{r}} = \vec{F}_T \quad \text{d'Alembert'sche Trägheitskraft} \quad (3.65)$$

Den Term $-m \cdot \ddot{\vec{r}}$ kann man als Kraft (Trägheitskraft) deuten. Dann stellt die Gl. (3.65) das Kräftegleichgewicht für die Masse m dar:

$$\sum \vec{F} = \vec{F} + \vec{F}_T = 0$$

Die Darstellung des 2. Newtonschen Axioms in dieser Form nennt man **Prinzip von d'Alembert** (nach Jean Le Rond d'Alembert, 1717-1783). Dadurch kann die Dynamik auf statische Betrachtungen zurückgeführt werden.

In den Grundlagen der Mechanik wird die Aufstellung von Bewegungsgleichungen für einzelne Massenpunkte, für Systeme von Massenpunkten, so genannte Mas-

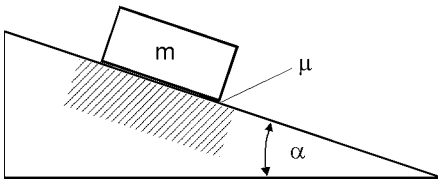
senpunkthaften und schließlich für ausgedehnte Körper aufgezeigt. Ein wichtiges Werkzeug ist dabei das **Schnittprinzip**. Die auf einen Massenpunkt wirkenden Kräfte werden durch gedachtes Freischneiden des Massenelementes aus den Bindungen an den Körper freigelegt. Einen starren Körper zerlegt man auf diese Weise in einzelne Massenpunkte dm (Bild 3-33). Um dann die Bewegungsgleichung aufschreiben zu können, muss man nach dem Freischneiden die dadurch wegfallenden Verbindungen zu den übrigen Teilen des Körpers durch Kräfte ersetzen. Danach muss man alle am bewegten Massenpunkt angreifenden Kräfte antragen. Dies sind im einzelnen:

- alle eingepägten Kräfte (Gewicht, Antriebskräfte, Magnetkräfte, usw.) F_e
- alle Zwangskräfte (Reaktionskräfte) infolge äußerer Bindungen und Führungen F_z
- die Bewegungswiderstände (Reibung, Luftwiderstand) F_r
- die Massen- oder Trägheitskräfte $-m \cdot \ddot{r}$ (d'Alembertsche Kräfte).

Danach können die Gleichgewichtsbedingungen wie in der Statik aufgeschrieben werden:

$$\vec{F}_e + \vec{F}_z + \vec{F}_r + \vec{F}_T = 0$$

a)



b)

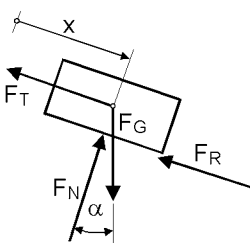


Bild 3-32: Masse auf schiefer Ebene mit angreifenden Kräften

Bei einfachen Körpern und Bewegungsverhältnissen kann man diese Vorgehensweise auch auf den Starrkörper als Ganzes anwenden. Ein einfaches Beispiel für die Ermittlung der Bewegungsgesetze nach dem d'Alembert'schen Prinzip ist der in Bild 3-32 a) dargestellte Fall einer ebenen, translatorischen Bewegung eines Körpers. Auf einer schiefen Ebene beginnt eine Masse m zum Zeitpunkt $t = 0$ aus der Ruhelage heraus zu rutschen. Zur Formulierung der Gleichgewichtsbedingung nach Gl. (3.65) muss man sich zuerst überlegen, welche Kräfte für die Beschreibung der Bewegung berücksichtigt werden müssen (Wahl des Modells). So wird man in diesem Fall die Gleitreibung zwischen rutschendem Körper und dem Untergrund berücksichtigen müssen, während man den Luftwiderstand oder gar die Corioliskraft aufgrund der Drehung der Erde sicher vernachlässigen kann. Gewisse Minimalkenntnisse über das Problem sind natürlich erforderlich, um überhaupt Aussagen machen zu können. So möge für dieses Problem die Größe der Masse m , der Neigungswinkel α der schiefen Ebene und der Reibungskoeffizient μ bekannt sein.

Zur Verfolgung der Bewegung der Masse wird eine Koordinate x vom Startpunkt aus eingeführt. Dies reicht als Koordinatenangabe aus, da die Masse aufgrund der

Formulierung des Problems nur einen Freiheitsgrad besitzt. In Teilbild 3-32 b) ist nun die Masse als von den Bindungen an die Ebene freigeschnittener Körper dargestellt, an den die am Körper angreifenden Kräfte angetragen sind. Dies sind die Gewichtskraft $F_G = m \cdot g$ (eingeprägte Kraft), die Normalkraft F_N (Zwangskraft, die die Masse an die Bahn bindet), die Reibkraft $F_R = \mu \cdot F_N$ im Gleitspalt zwischen Masse und Ebene, sowie die d'Alembert'sche Kraft $F_T = -m \cdot \ddot{x}$, die entgegen der Bewegungsrichtung angetragen wird.

Nun kann man zwei Gleichgewichtsbedingungen für die Kräfte in Richtung der schiefen Ebene und senkrecht dazu formulieren, indem man die Wirkung der Unterlage auf den Körper durch die entsprechenden Kräfte ersetzt:

$$\begin{aligned} m \cdot \ddot{x} + \mu \cdot F_N - mg \sin \alpha &= 0 && \text{in Hangrichtung} \\ F_N - mg \cos \alpha &= 0 && \text{senkrecht zur Hangrichtung} \end{aligned}$$

Da in beiden Gleichungen die für die Bewegung nicht interessierende Kraft F_N vorkommt, kann sie eliminiert werden. Es bleibt dann ein Ausdruck für die Beschleunigung übrig, den man zur Ermittlung des Weg-Zeit-Gesetzes zweimal integrieren muss:

$$\ddot{x} = (\sin \alpha - \mu \cos \alpha) \cdot g = a \quad , \quad \dot{x} = a \cdot t + C_1 \quad , \quad x = \frac{1}{2} a \cdot t^2 + C_1 \cdot t + C_2 \quad .$$

Für die Berechnung der Integrationskonstanten stehen zwei Anfangsbedingungen zur Verfügung. Die eine ergibt sich aus der Art der Problemstellung (keine Anfangsgeschwindigkeit), die andere wurde durch die Wahl des Koordinatensystems vorherbestimmt:

$$x(t=0) = x_0 \rightarrow C_2 = 0$$

$$\dot{x}(t=0) = v_0 \rightarrow C_1 = 0$$

Damit lauten die gesuchten Bewegungsgesetze:

$$x(t) = \frac{1}{2} (\sin \alpha - \mu \cos \alpha) \cdot g t^2$$

$$\dot{x}(t) = (\sin \alpha - \mu \cos \alpha) \cdot g t \quad .$$

Diese Bewegungsgleichungen gelten nur für Geschwindigkeiten $v \geq 0$, weil dies bei der Annahme der Richtung der Reibkraft vorausgesetzt wurde.

Beschreibungen der **allgemeinen Bewegung eines starren Körpers** kann man herleiten, indem man sich den Körper erst einmal in viele kleine, freigeschnittene Massenelemente zerlegt vorstellt (Bild 3-33). Nach dem 2. Newton'schen Axiom gilt:

$$d\vec{F} = \vec{r} \cdot dm$$

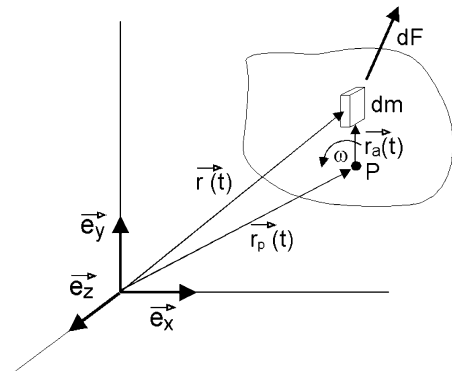


Bild 3-33: Massenpunkt eines starren Körpers

Durch Integration über die Gesamtmasse des Körpers erhält man daraus

$$\vec{F} = \int_m \vec{f} dm \quad (3.66)$$

Nun kann man die allgemeine Bewegung eines starren Körpers, wie in Abschnitt 3.2 behandelt, aus einer Translationsbewegung eines beliebigen körperfesten Punktes P und einer Rotation um diesen Punkt zusammensetzen. Es gilt:

$$\vec{r}(t) = \vec{r}_p(t) + \vec{r}_a(t) \quad \text{mit} \quad |\vec{r}_a| = \text{const.} \quad (3.67)$$

Als Differentiationsregel für Vektoren gilt:

$$\dot{\vec{r}} = \dot{\vec{r}}_p + \vec{\omega} \times \vec{r}_a \quad (3.68)$$

Dabei ist $\vec{\omega}$ die Winkelgeschwindigkeit, mit der sich der Vektor $\vec{r}_a$ um den Punkt P dreht. Für die Beschleunigung erhält man ebenso

$$\ddot{\vec{r}} = \ddot{\vec{r}}_p + \dot{\vec{\omega}} \times \vec{r}_a + \vec{\omega} \times (\vec{\omega} \times \vec{r}_a) \quad .$$

Unter Verwendung der Identität für dreifache Vektorprodukte, die aus dem Entwicklungssatz folgt

$$\vec{\omega} \times (\vec{\omega} \times \vec{r}_a) = (\vec{\omega} \cdot \vec{r}_a) \cdot \vec{\omega} - (\vec{\omega} \cdot \vec{\omega}) \cdot \vec{r}_a \quad , \quad \text{mit} \cdot \text{ für das Skalarprodukt}$$

erhält man für die Beschleunigung

$$\ddot{\vec{r}} = \ddot{\vec{r}}_p + \dot{\vec{\omega}} \times \vec{r}_a + (\vec{\omega} \cdot \vec{r}_a) \cdot \vec{\omega} - \omega^2 \cdot \vec{r}_a \quad .$$

Durch Einsetzen in die Bewegungsgleichung (3.66) findet man:

$$\begin{aligned} \vec{F} &= \int_m (\ddot{\vec{r}}_p + \dot{\vec{\omega}} \times \vec{r}_a + (\vec{\omega} \cdot \vec{r}_a) \cdot \vec{\omega} - \omega^2 \cdot \vec{r}_a) dm \\ &= \ddot{\vec{r}}_p \cdot m + (\dot{\vec{\omega}} \times \vec{r}_{ps}) \cdot m - (\omega^2 \cdot \vec{r}_{ps}) \cdot m \quad , \end{aligned} \quad (3.69)$$

da $\vec{\omega}$ und $\vec{r}_a$ senkrecht aufeinander stehen. Der Vektor $\vec{r}_{ps}$ ist dabei der Vektor vom Punkt P zum **Massenmittelpunkt S** oder **Schwerpunkt** des starren Körpers:

$$\vec{r}_{ps} = \frac{\int \vec{r}_a dm}{m} \quad .$$

Wählt man nun den Punkt P so, dass er mit dem Massenmittelpunkt zusammenfällt, wodurch der Vektor $\vec{r}_{ps}$ zu Null wird, so erhält man den sogenannten Schwerpunktsatz:

$$\vec{F} = m \cdot \ddot{\vec{r}}_s \quad , \quad (3.70)$$

wobei $\vec{r}_s$ der Vektor zum Massenmittelpunkt (Schwerpunkt) des Körpers ist. Dieser besagt demnach, dass man sich die Gesamtmasse m im Schwerpunkt des Körpers konzentriert denken darf, an dem alle äußeren Kräfte angreifen. Zumindest für den rein translatorischen Anteil an der Bewegung ist die Situation also genauso, wie für

einen Massenpunkt. Führt der Körper jedoch eine allgemeine Bewegung mit rotatorischem Anteil aus, so ist für die Behandlung dieses Bewegungsanteils nach dem d'Alembert'schen Prinzip (Gleichgewicht der Momente) ein weiterer Satz, der sog. **Drall- oder Momentensatz** von Bedeutung.

Aus Gl. (3.66) wird durch vektorielles Multiplizieren mit dem Ortsvektor $\vec{r}$

$$\vec{M} = \vec{r} \times \vec{F} = \int_m (\vec{r} \times \vec{f}) dm \quad (3.71)$$

$\vec{M} = \vec{r} \times \vec{F}$ ist das resultierende äußere Moment, das am Körper angreift. Analog zum Impuls $m \cdot \vec{v}$ der translatorischen Bewegung kann man nun den **Drehimpuls** oder **Drall** $\Phi \cdot \vec{\omega}$, mit Φ als **Drehträgeit** oder **Trägheitsmoment** definieren. Den Drall des Körpers bezogen auf den Ursprung des raumfesten Koordinatensystems definiert folgender Ausdruck:

$$\vec{D} = \int_m (\vec{r} \times \vec{r}) dm \quad (3.72)$$

Durch Einsetzen der Gl. (3.67) und (3.68) in diese Gleichung erhält man:

$$\begin{aligned} \vec{D} &= \int_m (\vec{r}_p + \vec{r}_a) \times (\vec{r}_p + \vec{\omega} \times \vec{r}_a) dm \\ &= \int_m (\vec{r}_p \times \vec{r}_p) dm + \int_m (\vec{r}_p \times (\vec{\omega} \times \vec{r}_a)) dm + \int_m (\vec{r}_a \times \vec{r}_p) dm + \int_m (\vec{r}_a \times (\vec{\omega} \times \vec{r}_a)) dm. \end{aligned} \quad (3.73)$$

Hieraus folgt durch Integration und unter Verwendung des Vektors $\vec{r}_{ps}$ (Vektor vom Punkt P zum Massenmittelpunkt S des starren Körpers):

$$\vec{D} = \int_m (\vec{r}_a \times (\vec{\omega} \times \vec{r}_a)) dm + m(\vec{r}_p \times (\vec{v}_p + (\vec{\omega} \times \vec{r}_{ps}))) + m(\vec{r}_{ps} \times \vec{v}_p) \quad (3.74)$$

Wählt man den Punkt P wieder so, dass er mit dem Massenmittelpunkt S übereinstimmt ($\vec{r}_{ps} = 0$), dann vereinfacht sich die Formel für den Drall zu

$$\vec{D} = \int_m (\vec{r}_a \times (\vec{\omega} \times \vec{r}_a)) dm + m(\vec{r}_s \times \vec{v}_s) \quad (3.74)$$

Das Integral in Gl. (3.74) stellt den Drall $\vec{D}_p$ des Körpers bezogen auf den körperfesten Punkt P oder S dar. Diesen Ausdruck kann man so umformen, dass im Ergebnis nur noch Integrale vorkommen, die von der Massenverteilung des Körpers, nicht aber vom Bewegungszustand abhängen. Es gilt

$$\vec{D}_p = \int_m (\vec{r}_a \times (\vec{\omega} \times \vec{r}_a)) dm = \int_m (\vec{\omega} \cdot r_a^2 - \vec{r}_a (\vec{\omega} \cdot \vec{r}_a)) dm$$

Dies kann man nun in Matrixenschreibweise folgendermaßen formulieren:

$$\vec{D}_p = \begin{pmatrix} \Theta_{xx} & -\Theta_{xy} & -\Theta_{xz} \\ -\Theta_{yx} & \Theta_{yy} & -\Theta_{yz} \\ -\Theta_{zx} & -\Theta_{zy} & \Theta_{zz} \end{pmatrix} \cdot \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \Theta \cdot \vec{\omega} \quad (3.75)$$

Darin bedeuten die Elemente der Matrix die Drehträgeiten oder Massenträgheitsmomente des Körpers bezüglich eines körperfesten Koordinatensystems. Dabei

sind

$$\Theta_{xx} = \int_m (y^2 + z^2) dm = \int_m r_x^2 dm$$

$$\Theta_{yy} = \int_m (x^2 + z^2) dm = \int_m r_y^2 dm$$

$$\Theta_{zz} = \int_m (x^2 + y^2) dm = \int_m r_z^2 dm$$

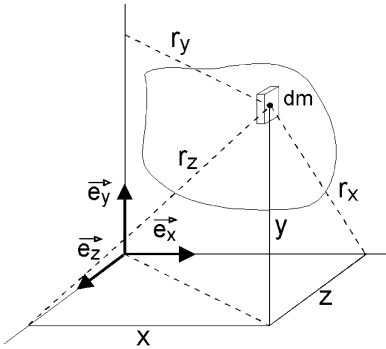


Bild 3-34: Massenträgheitsmomente

die **axialen Massenträgheitsmomente** für Drehungen des Massenelementes dm um die entsprechenden Achsen des Koordinatensystems mit dem zugehörigen Drehradius $r_{x,y,z}$ (Bild 3-34).

Die Größen

$$\Theta_{xy} = \int_m (x \cdot y) dm$$

$$\Theta_{yz} = \int_m (y \cdot z) dm$$

$$\Theta_{xz} = \int_m (x \cdot z) dm$$

werden als **Deviations- oder Zentrifugalmomente** bezeichnet. Die Größe der Massenträgheits- und der Deviationsmomente ist von der Lage des Koordinatenursprungs und den Richtungen der Achsen abhängig. Für jeden Koordinatenursprung gibt es mindestens drei Achsen, für die die Deviationsmomente Null werden. Dies sind die **Hauptachsen** für den gewählten Koordinatenursprung.

Die Integrale über die Masse m werden zur Berechnung durch Dreifachintegrale über die drei Richtungen der Koordinatenachsen gebildet:

$$\Theta_{xx} = \int_m r_x^2 dm = \iiint \rho (y^2 + z^2) dx dy dz$$

Je nach Körperform verwendet man auch Zylinder- oder Kugelkoordinaten. So kann man das Massenträgheitsmoment für einen Zylinder (Bild 3-35) folgendermaßen berechnen:

$$\begin{aligned} \Theta_{xx} &= \int_{r=0}^R \int_{\varphi=0}^{2\pi} \int_{z=-L/2}^{L/2} \rho r^2 (dr (r \cdot d\varphi) dz) = \int_{r=0}^R \int_{\varphi=0}^{2\pi} \int_{z=-L/2}^{L/2} \rho r^3 (dr d\varphi dz) \\ &= \rho \left(\frac{R^4}{4} \right) 2\pi L = m \frac{R^2}{2} \quad \text{mit} \quad m = \rho \pi R^2 L \end{aligned}$$

Bild 3-35 zeigt als Beispiel die axialen Massenträgheitsmomente von zwei symmetrischen, homogenen Körpern.

Der **Momentensatz** für die Drehbewegung lautet dann :

$$\vec{M} = \frac{d\vec{D}}{dt} = \frac{d(\Theta \vec{\omega})}{dt} \quad (3.76)$$

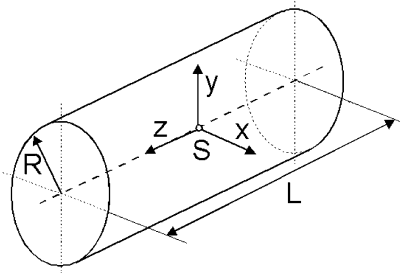
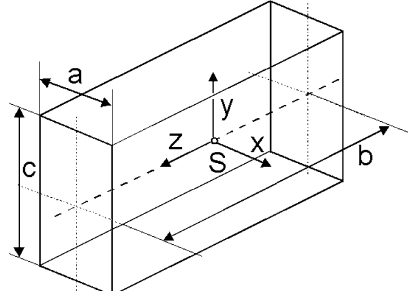
	Kreiszyylinder: $\Theta_{zz} = \frac{1}{2} m R^2$ $\Theta_{xx} = \Theta_{yy} = \frac{1}{12} m (3 R^2 + L^2)$
	Quader: $\Theta_{xx} = \frac{1}{12} m (b^2 + c^2)$ $\Theta_{yy} = \frac{1}{12} m (a^2 + b^2)$ $\Theta_{zz} = \frac{1}{12} m (a^2 + c^2)$

Bild 3-35: Axiale Massenträgheitsmomente symmetrischer, homogener Körper

Man kann also für die Drehbewegung entsprechend dem 2. Newtonschen Axiom folgendermaßen formulieren:

Die zeitliche Änderung der Bewegungsgröße (Drallvektor) entspricht der Wirkung des äußeren Momentes. Als Bezugspunkt für den Drall sind der beliebig bewegte Schwerpunkt oder ein ruhender Punkt des Körpers zugelassen.

Man kann diesen Satz für einfache Probleme (siehe folgendes Beispiel), in denen der sich drehende Körper geometrisch einfach (z.B. symmetrisch) beschreiben lässt, verwenden. Hierbei gelingt die Bestimmung der Trägheitsmomente ohne größere Schwierigkeiten. Auch bei Problemen, in denen durch entsprechende Führung die Bewegungen überschaubar sind, lässt sich der Satz anwenden.

Im folgenden Beispiel [1.5] wird eine aus Translation und Rotation zusammengesetzte, ebene Bewegung behandelt (Bild 3-36 a). Darin wird eine Winde mit einem konstanten Moment angetrieben und zieht über ein dehnstarres, masseloses (Modellannahmen zur Vereinfachung des Problems) Seil eine Walze auf einer schiefen Ebene aufwärts. Die Bewegung soll zum Zeitpunkt $t = 0$ aus der Ruhe heraus beginnen. Die Walze führt eine reine Rollbewegung aus. Unter Verwendung der im Bild gegebenen Größen sollen die Beschleunigung des Walzenmittelpunktes, sowie die Kräfte im Seil und die Lagerkräfte bei A bestimmt werden.

Da die Walze sich drehen kann und sich dabei ihr Mittelpunkt translatorisch verschiebt, werden für sie als Bewegungskordinaten x_1 und φ_1 gewählt. Die reine Drehung der Winde wird mit der Bewegungskordinate φ_2 verfolgt. Für die Walze soll die Rollbedingung gelten $\dot{\varphi}_1 = \dot{x}_1 / R_1$. Ihr Momentanpol ist der Berührungspunkt mit der schiefen Ebene. Der obere Punkt, an dem das Seil angreift, hat die Ge-

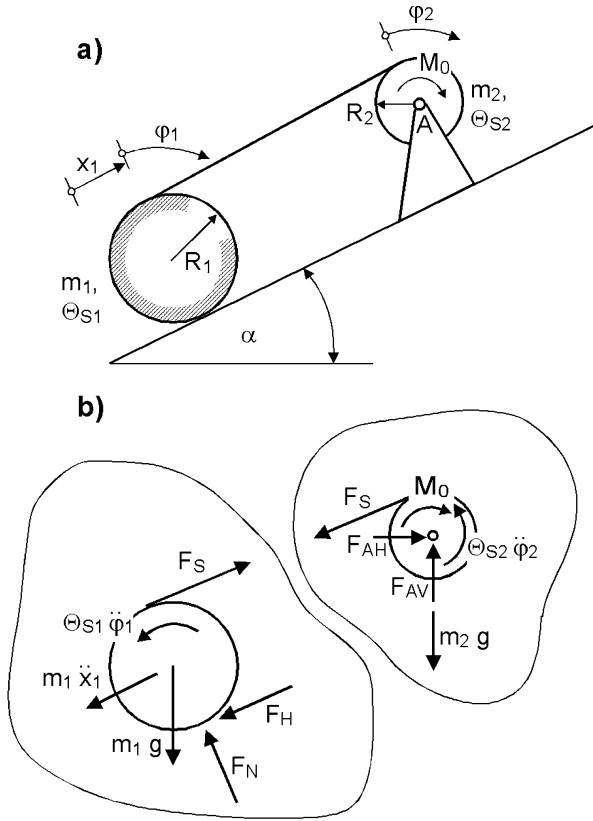


Bild 3-36: Beispiel für die Aufstellung der Bewegungsgleichung mit Hilfe des Drallsatzes

geschwindigkeit $2\dot{x}_1$, weil er doppelt so weit vom Momentanpol entfernt ist wie der Walzenmittelpunkt, der die Geschwindigkeit $\dot{x}_1$ besitzt. Dies ist dann auch die Geschwindigkeit des Seils und damit die Bahngeschwindigkeit am Außenradius der Winde, für deren Winkelgeschwindigkeit demnach gilt: $\dot{\varphi}_2 = 2\dot{x}_1/R_2$. Alle Koordinaten sollen bei $t = 0$ ebenfalls Null sein, so dass die genannten Zwangsbedingungen in gleicher Form für Weg, Geschwindigkeit und Beschleunigung gelten.

Für die Anwendung des Prinzips von d'Alembert zeigt Bild 3-36 b) die freigeschnittenen Massen mit den angreifenden Kräften und Momenten. Für fünf unbekannte Kräfte und die unbekannte Beschleunigung stehen sechs Gleichgewichtsbedingungen zur Verfügung. Da im Beispiel jedoch die Auflagerkräfte F_V und F_H der Walze nicht benötigt werden, wird für die Walze nur eine von drei

möglichen Gleichgewichtsbedingungen aufgeschrieben. Dies ist die Momentengleichgewichtsbedingung mit dem Angriffspunkt der Auflagerkräfte als Bezugspunkt (Momentanpol: ruhender Punkt des Körpers).

$$F_S \cdot 2R_1 - m_1 \ddot{x}_1 R_1 - \Theta_{S1} \ddot{\varphi}_1 - m_1 g R_1 \sin \alpha = 0$$

$$M_0 - \Theta_{S2} \ddot{\varphi}_2 - F_S R_2 = 0, \quad F_{AH} - F_S \cos \alpha = 0, \quad F_{AV} - m_2 g - F_S \sin \alpha = 0$$

In den beiden Gleichungen für das Momentengleichgewicht werden die Winkelbeschleunigungen $\ddot{\varphi}_1$ und $\ddot{\varphi}_2$ durch die Beschleunigung $\ddot{x}_1$ des Walzenmittelpunktes ersetzt, die man aus den oben genannten Zwangsbedingungen herleiten kann.

$$\left(m_1 R_1 + \frac{\Theta_{S1}}{R_1} \right) \ddot{x}_1 = 2F_S R_1 - m_1 g R_1 \sin \alpha \Rightarrow \left(m_1 + \frac{\Theta_{S1}}{R_1^2} \right) \ddot{x}_1 = 2F_S - m_1 g \sin \alpha$$

$$2 \frac{\Theta_{S2}}{R_2} \ddot{x}_1 = M_0 - F_S R_2 \Rightarrow 2 \frac{\Theta_{S2}}{R_2^2} \ddot{x}_1 = \frac{M_0}{R_2} - F_S$$

Durch Eliminierung von F_S aus den beiden Gleichungen erhält man die Beschleunigung:

$$\ddot{x}_1 = \frac{2 \frac{M_0}{R_2} - m_1 g \sin \alpha}{m_1 + \frac{\Theta_{S1}}{R_1^2} + 4 \frac{\Theta_{S2}}{R_2^2}}$$

Aus den drei übrigen Gleichgewichtsbedingungen kann man dann die gesuchten Seil- und Auflagerkräfte gewinnen:

$$F_S = \frac{M_0}{R_2} - 2 \frac{\Theta_{S2}}{R_2^2} \ddot{x}_1, \quad F_{AH} = F_S \cos \alpha, \quad F_{AV} = m_2 g + F_S \sin \alpha$$

Die Beschleunigung ist nach der gefundenen Gleichung konstant, so dass sich für diesen Fall das Geschwindigkeits-Zeit-Gesetz und das Weg-Zeit-Gesetz problemlos durch Integration finden lassen. Der Ausdruck für die Beschleunigung zeigt einen typischen Aufbau: Im Nenner stehen alle **trägen Massen**, die ausnahmslos positiv sind, im Zähler ist das Moment M_0 als antreibende Größe positiv, während das Eigengewicht der Walze als bremsende Größe negatives Vorzeichen trägt.

Hat man aber nun als Bewegungsproblem unsymmetrische Körper in allgemeiner räumlicher Bewegung, so ist der Drallsatz für eine Berechnung schwierig zu handhaben. Allerdings lassen sich viele technische Fragestellungen auf die Untersuchung ebener Bewegungen beschränken.

3.4.2 Energiesatz

Die Beziehungen für Impuls und Drall reichen aus, um die Bewegungen beliebiger mechanischer Systeme zu beschreiben [1.5], [3.9]. Bei gewissen Problemen mit komplizierten Bewegungen erlaubt jedoch der **Energiesatz** spezielle Aussagen über die Bewegungen. Insbesondere wenn nur so genannte **konservative** Kräfte auf den Körper einwirken, enthält dieser Energiesatz nur Zustandsgrößen zweier ausgewählter Zeitpunkte der Bewegung. Der Energiesatz vergleicht gewissermaßen den Systemzustand zu zwei verschiedenen Zeitpunkten, ohne den Verlauf der Bewegung zu verfolgen. Konservative Kräfte sind dabei Kräfte, die ein Potential besitzen, so dass die von ihnen verrichtete Arbeit nur vom Anfangs- und Endpunkt der Bewegung, aber nicht vom zurückgelegten Weg abhängen. Eine typische Kraft dieses Typs ist die Gewichtskraft $F_G = m \cdot g$. Bewegt sich ein Massenpunkt im Potentialfeld der Schwerkraft, so ist für die Masse m die **Arbeit** ihrer Gewichtskraft vom Weg unabhängig und hängt nur von der Höhendifferenz $h = h_1 - h_2$ des Anfangs- und Endpunktes der Bewegung ab:

$$W = mgh = mg \cdot (h_1 - h_2) = U_1 - U_2 \quad (3.77)$$

Die Energie U , die die Masse im Potentialfeld aufgrund der Wirkung der Gewichtskraft aufnimmt oder abgibt, bezeichnet man daher auch als **potentielle Energie**. Die in Gl. (3.77) gemachte Aussage findet man aus der allgemeinen Definition für

die mechanische Arbeit, die eine Kraft an einem Massenpunkt bei einer Bewegung entlang einer Bahn verrichtet (Bild 3-37):

$$W = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F} \cdot d\vec{r} = \int_{s_1}^{s_2} F_s ds \quad , \quad (3.78)$$

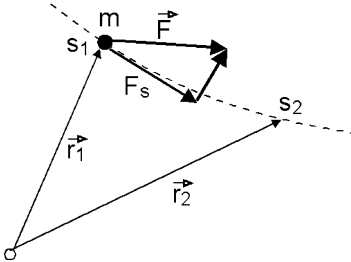


Bild 3-37: Massenpunkt auf einer Bahn

wobei F_s die tangential zur Bahnrichtung wirkende Kraftkomponente ist. Dies gilt, weil das Skalarprodukt des Kraftvektors und des Bahnvektors gerade den Betrag des tangentialen Anteils des Bahnvektors besitzt. Setzt man in Gl. (3.78) für die Kraft den Ausdruck aus Gl. (3.64) ein, so erhält man unter der Voraussetzung, dass die Masse m konstant ist:

$$W = m \int_{\vec{r}_1}^{\vec{r}_2} \frac{d\vec{v}}{dt} d\vec{r} = m \int_{\vec{v}_1}^{\vec{v}_2} \frac{d\vec{r}}{dt} d\vec{v} = m \int_{\vec{v}_1}^{\vec{v}_2} \vec{v} d\vec{v} \quad ,$$

so dass mit Bezug auf Gl. (3.78) gilt:

$$\int_{s_1}^{s_2} F_s ds = \frac{m}{2} v_2^2 - \frac{m}{2} v_1^2 = T_2 - T_1 \quad . \quad (3.79)$$

Dieser so genannte **Arbeitssatz** besagt, dass die Arbeit, die von der die Bahnkurve tangierenden äußeren Kraft F_s längs des Weges s verrichtet wird, gleich der Differenz der kinetischen Energie des Massenpunktes m zwischen Anfangs- und Endpunkt der Bewegung ist. Die **kinetische Energie**

$$T = \frac{m}{2} v^2 \quad (3.80)$$

ist das Arbeitsvermögen des Massenpunktes mit der Masse m infolge seiner Bewegung mit der Geschwindigkeit v .

Wirken nun am Massenpunkt ausschließlich **konservative** Kräfte (Reibkräfte sind beispielweise keine konservativen Kräfte), so kann man das Integral auf der linken Seite der Gl. (3.78) durch den Ausdruck für die potentielle Energie aus Gl. (3.77) ersetzen, woraus die einfachste Form des sogenannten **Energiesatzes** oder **Energieerhaltungssatzes** für den Massenpunkt folgt:

$$mg \cdot (h_1 - h_2) = \frac{m}{2} v_2^2 - \frac{m}{2} v_1^2 \quad \text{oder} \quad mgh_1 + \frac{m}{2} v_1^2 = mgh_2 + \frac{m}{2} v_2^2 \quad . \quad (3.81)$$

In vereinfachter Form kann dieser auch folgendermaßen formuliert werden:

$$U_1 + T_1 = U_2 + T_2 = \text{konstant} \quad ,$$

d. h. unter der Voraussetzung, dass nur konservative Kräfte auf den Massenpunkt wirken, ist die Summe aus potentieller und kinetischer Energie konstant. Dieser Satz gilt natürlich nicht nur für die Bewegung im Schwerfeld der Erde, sondern auch für die Wirkung anderer Potentialkräfte. Ein weiteres Beispiel für eine solche Potentialkraft ist die Kraft, die eine gespannte Feder ausübt.

Im folgenden Beispiel soll nun der Bewegungszustand zweier Massen, auf die das Potentialfeld einer gespannten Feder einwirkt, mit Hilfe des Energiesatzes ermittelt werden. Wie im Bild 3-38 dargestellt, schleudert eine um den Weg x_1 vorgespannte Feder mit der Federsteifigkeit k zum Zeitpunkt t_1 zwei Massen m_1 und m_2 auseinander. Unter Vernachlässigung von Reibungskräften soll die Geschwindigkeit der Massen zum Zeitpunkt t_2 nach der Entspannung der Feder ermittelt werden.

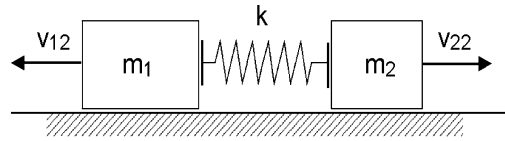


Bild 3-38: Zwei Massen im Potentialfeld einer Feder

Der Impulssatz (Gl. 3.63) angewendet auf mehrere Körper besagt, dass das Integral über die auf die Körper wirkenden äußeren Kräfte gleich der Differenz aller Impulse bzw. gleich der Differenz der Schwerpunktimpulse ist. Wirken keine äußeren Kräfte, so bleibt der Gesamtimpuls konstant. Außerdem gilt nach dem 3. Newton'schen Axiom, dass ein Impuls einen gleichgroßen entgegengesetzt gerichteten Impuls zur Folge hat, d. h. $m_1 v_{12}$ (Impuls von m_1 zum Zeitpunkt t_2) = $m_2 v_{22}$ (Impuls von m_2 zum Zeitpunkt t_2). Vor der Entspannung der Federn (Zeitpunkt t_1) gilt weiterhin: $v_{11} = v_{21} = 0$. Damit gilt für die kinetische Energie der Massen: $T_{11} = T_{21} = 0$. Die kinetische Energie zum Zeitpunkt t_2 erhält man aus dem Integral über den Impuls:

$$T_{12} = m_1 \int_0^{v_{12}} v dv = \frac{m_1}{2} v_{12}^2 \quad T_{22} = m_2 \int_0^{v_{22}} v dv = \frac{m_2}{2} v_{22}^2 \quad .$$

Diese beiden kinetischen Energien müssen wegen dem oben Gesagten gleich groß sein. Für die Potentialkraft der Feder gilt nach dem Hooke'schen Gesetz:

$$F = k \cdot x$$

Für die potentielle Energie der Feder gilt:

$$W = U_1 - U_2 = \int_{x_1}^{x_2} F dx = \int_{x_1}^{x_2} k \cdot x dx = \frac{k}{2} (x_2^2 - x_1^2) \quad .$$

Die potentielle Energie U_2 der Feder zum Zeitpunkt t_2 ist gleich Null, da die Feder dann entspannt ist und damit $x_2 = 0$ gilt.

Der Energiesatz für ein System, auf das ausschließlich Potentialkräfte einwirken, besagt, dass die Summe aus potentieller und kinetischer Energie des Systems zu jedem Zeitpunkt gleich und konstant ist. Demnach gilt für das Feder-Masse-System:

$$\frac{k}{2} x_1^2 = \frac{m_1}{2} v_{12}^2 + \frac{m_2}{2} v_{22}^2 \quad .$$

Hieraus gewinnt man mit den obigen Feststellungen:

$$v_{12} = \sqrt{\frac{kx_1^2}{\left(m_1 + \frac{m_2^2}{m_2}\right)}} \quad v_{22} = \sqrt{\frac{kx_1^2}{\left(m_2 + \frac{m_1^2}{m_1}\right)}}$$

Für die allgemeine ebene Bewegung starrer Körper muss man zur Berechnung der kinetischen Energie noch berücksichtigen, dass sich eine solche Bewegung aus der Translation eines ausgezeichneten Bezugspunktes des Körpers und der Rotation des Körpers um diesen Punkt zusammensetzen lässt. Wählt man einen geeigneten Punkt P , etwa den Schwerpunkt des Körpers oder einen momentan in Ruhe befindlichen Punkt, so lässt sich die kinetische Energie folgendermaßen aufschreiben:

$$T = T_{trans} + T_{rot} = \frac{1}{2} m \cdot v_P^2 + \frac{1}{2} \Theta_P \cdot \dot{\varphi}^2 = \frac{1}{2} m \cdot v_P^2 + \frac{1}{2} \Theta_P \cdot \omega^2 \quad . \quad (3.82)$$

3.4.3 Die Prinzipien der Mechanik

Die Mechanik kann vollständig auf wenigen Axiomen aufgebaut werden. Dazu gehören die Axiome der Statik und die Newton'schen Axiome. Als **Prinzipien der Mechanik** werden Aussagen bezeichnet, die die klassischen Axiome ersetzen können. Dies sind Prinzipien, die zwar keine neuen Erkenntnisse für die Mechanik liefern, aber nicht zu ihnen im Widerspruch stehen und aus den klassischen Axiomen herleitbar sind. Die analytische Mechanik beschäftigt sich mit den Prinzipien der Mechanik und stellt für spezielle Problemstellungen, die mit den klassischen Axiomen nur schwer oder gar nicht lösbar sind, die mathematische Formulierung zur Verfügung. Hierunter fallen vor allem Bewegungen von Mehrkörpersystemen. Die wichtigsten Prinzipien der Mechanik sind:

- Prinzip der virtuellen Arbeit
- Prinzip vom Minimum des Potentials
- Prinzip von Hamilton

Das erste Prinzip, formuliert in verschiedenen Ausprägungen, soll anschließend behandelt werden.

3.4.3.1 Prinzip der virtuellen Arbeit

In Abschnitt 3.4.2 wurde die Arbeit als Produkt aus einer Verschiebung $d\vec{r}$ und der in Richtung der Verschiebung wirkenden Kraftkomponente definiert. Wir haben im Abschnitt 3.4.1 gesehen, dass das Prinzip von d'Alembert die Dynamik auf statische Betrachtungen zurückführt. Um den Arbeitsbegriff auch auf Probleme der Statik anwenden zu können, bei denen definitionsgemäß keine Verschiebungen auftreten, kann man **virtuelle Verschiebungen** $\delta\vec{r}$ definieren. Diese virtuellen (möglichen) Verschiebungen haben dabei folgende Eigenschaften:

- Virtuelle Verschiebungen oder Verdrehungen sind infinitesimal klein und können daher wie Differentiale behandelt werden.
- Virtuelle Verschiebungen oder Verdrehungen müssen mit den geometrischen Bindungen des Systems verträglich sein, d. h. virtuelle Verschiebungen oder Verdrehungen sind nur in Richtung vorhandener Freiheitsgrade möglich.

Die virtuelle Arbeit einer Kraft oder eines Momentes ist das Skalarprodukt aus der Kraft oder dem Moment und der virtuellen Verschiebung oder Verdrehung des Kraftangriffspunktes:

$$\delta W = \vec{F} \cdot \delta \vec{r} \quad , \quad \delta W = \vec{M} \cdot \delta \vec{\varphi} \quad . \quad (3.83)$$

Will man dieses Prinzip auf ein einfaches, statisches System wie den in Bild 3-39 dargestellten Balken mit einem Festlager im Punkt A und einem Loslager in Punkt B anwenden, so muss man an der Lagerstelle B, die einen Freiheitsgrad besitzt, die Bindung lösen und dafür eine Lagerkraft F_B einführen. Der Balken kann dann durch die Lagerkraft um den Punkt A um einen virtuellen Verdrehwinkel $\delta\varphi$ verdreht werden. Weil diese Verdrehung infinitesimal klein ist, betragen die virtuellen Verschiebungen der Kraftangriffspunkte der Lastkraft F und der Auflagerkraft F_B :

$$x(F) = a \cdot \delta\varphi \quad , \quad x(F_B) = l \cdot \delta\varphi$$

Die bei dieser virtuellen Verdrehung durch die beiden entgegengesetzt gerichteten Kräfte verrichtete virtuelle Arbeit beträgt:

$$\delta W = F_B \cdot l \cdot \delta\varphi - F \cdot a \cdot \delta\varphi = (F_B \cdot l - F \cdot a) \cdot \delta\varphi$$

Der Ausdruck in der Klammer stellt das aus der Statik bekannte Momentengleichgewicht um den Punkt A dar, das den Wert Null haben muss, da sich das System in Ruhe (im Gleichgewicht) befindet. Damit muss für die verrichtete virtuelle Arbeit gelten $\delta W = 0$. Dies gilt als Prinzip für alle mechanischen Systeme und wird als **Prinzip der virtuellen Arbeit** bezeichnet:

$$\delta W = \sum_{i=1}^n F_{i,e} \cdot \delta \vec{r}_i + \sum_{i=1}^m M_{i,e} \cdot \delta \vec{\varphi}_i = 0 \quad (3.84)$$

Die Gl. (3.84) [1.5] besagt, dass sich ein mechanisches System im Gleichgewicht befindet, wenn bei einer virtuellen Verschiebung oder Verdrehung aus der Gleichgewichtslage heraus die dabei von den äußeren (eingepägten: Index e) Kräften und Momenten verrichtete virtuelle Arbeit gleich Null ist. Dieses Prinzip ist gleichwertig mit dem Gleichgewichtsaxiom der Statik. Im Beispiel war der Punkt A ein fester Punkt, der keinen virtuellen Verschiebungen unterliegt. Die Zwangskräfte aus der festen Einspannung leisten daher keine virtuelle Arbeit. Daher lässt sich das Prinzip der virtuellen Arbeit besonders gut auf komplizierte Systeme mit vielfachen Bindungen anwenden. Zur Ermittlung der gesuchten Größen löst man nur so viele Bindungen, wie man Gleichungen zur Ermittlung der Größen benötigt.

Das Prinzip lässt sich aber nicht nur auf statische Systeme anwenden, sondern kann auch gut zur Ermittlung unbekannter Gleichgewichtslagen von bewegten Sys-

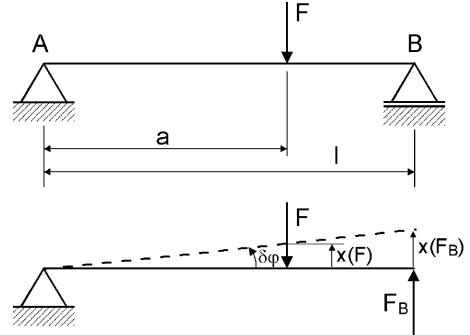


Bild 3-39: Virtuelle Verschiebung eines statischen Systems mit einem Freiheitsgrad

temen mit mehreren Freiheitsgraden dienen.

Die Lage eines Systems starrer Körper mit f Freiheitsgraden kann immer durch f voneinander unabhängige Koordinaten $q_1, \dots, q_f$ eindeutig beschrieben werden. Solche Koordinaten (Verschiebungskoordinaten für Translationen und Verdrehwinkel für Rotationen) bezeichnet man auch als **generalisierte** Koordinaten. Wird die Lage eines Punktes des Systems, beispielsweise eines Kraftangriffspunktes, durch den Ortsvektor $\vec{r}$ beschrieben, so ist seine virtuelle Verschiebung $\delta\vec{r}$. Der Ortsvektor hängt von allen f generalisierten Koordinaten ab, da diese nach Definition unabhängig voneinander sind. Da die virtuelle Verschiebung als Differential behandelt werden kann, erhält man nach den Regeln der Differentialrechnung folgende Schreibweise:

$$\vec{r} = \vec{r}(q_1, q_2, \dots, q_f) \Rightarrow \delta\vec{r} = \frac{\partial\vec{r}}{\partial q_1} \delta q_1 + \frac{\partial\vec{r}}{\partial q_2} \delta q_2 + \dots + \frac{\partial\vec{r}}{\partial q_f} \delta q_f \quad (3.85)$$

Da die virtuellen Koordinaten δq_i sowohl Verschiebungen als auch Verdrehungen sein können, lässt sich Gl. (3.84) für die virtuelle Arbeit folgendermaßen schreiben:

$$\delta W = (\dots) \cdot \delta q_1 + (\dots) \cdot \delta q_2 + \dots + (\dots) \cdot \delta q_f \quad (3.86)$$

Da die q_i unabhängig voneinander sind und daher die δq_i beliebige Werte annehmen können, kann Gl. (3.86) nur erfüllt sein, wenn alle Klammerausdrücke einzeln den Wert Null annehmen. Man kann sie daher Null setzen und erhält damit f Bestimmungsgleichungen für f unbekannte Größen.

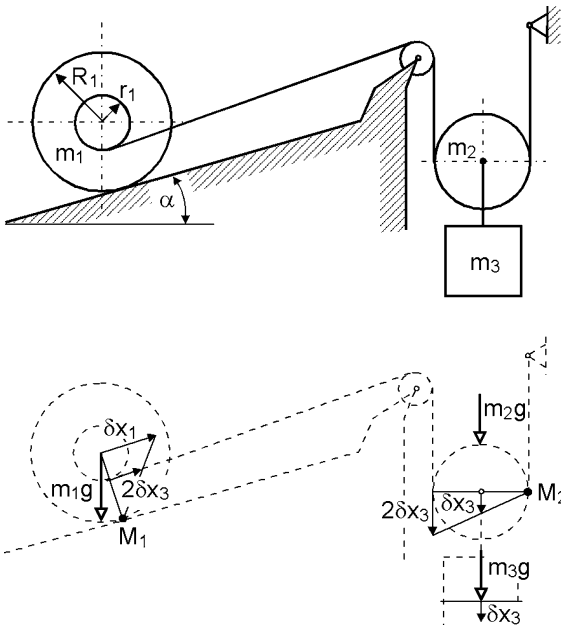


Bild 3-40: Untersuchung eines Gleichgewichtszustandes mit dem Prinzip der virtuellen Arbeit

Als Beispiel für die oben angeführte Vorgehensweise soll das in Bild 3-40 a) [1.5] dargestellte Gleichgewichtsproblem dienen. Hierin wird eine Walze auf einer schiefen Ebene über ein starres Seil von einem Gegengewicht im Gleichgewicht gehalten. Bei bekannter Geometrie und Masse der beteiligten Körper soll die Masse m_3 des Gegengewichtes so bestimmt werden, dass Gleichgewicht herrscht.

Da das System nur einen Freiheitsgrad besitzt, können alle Verschiebungen der beteiligten Körper durch eine Koordinate ausgedrückt werden. Die kinematischen Zusammenhänge können über die im unteren Teilbild dargestellten Momentenpole M_1 und M_2 der Bewegungen von Walze bzw. Seilrolle einfach analysiert werden. Wenn die Masse m_3 um die virtu-

elle Verschiebung δx_3 verrückt wird, bewegt sich auch der Mittelpunkt der Seilrolle m_2 um diesen Betrag. Der Punkt auf der gegenüber dem Momentanpol M_2 liegenden Seite der Rolle m_2 muss sich um den doppelten Betrag bewegen, wodurch über das Seil die Walze m_1 mit dem Seilangriffspunkt ebenfalls um $2\delta x_3$ bewegt wird. Die Mittelpunktsverschiebung δx_1 der Walze ergibt sich aus der Verrückung des Seilangriffspunktes und der Kenntnis des Momentanpols M_1 aus dem Strahlensatz:

$$\frac{\delta x_1}{2\delta x_3} = \frac{R_1}{R_1 - r_1} \Rightarrow \delta x_1 = \frac{2R_1}{R_1 - r_1} \delta x_3 \quad .$$

Während sich die Kraftangriffspunkte der Gewichtskräfte m_2g und m_3g jeweils um δx_3 bewegen, wirkt von der Gewichtskraft m_1g nur die Komponente parallel zur schiefen Ebene $m_1g \sin \alpha$ in Richtung von δx_1 . Das Prinzip der virtuellen Arbeit liefert dann folgende Gleichung:

$$\begin{aligned} \delta W &= m_3g \delta x_3 + m_2g \delta x_3 - m_1g \delta x_1 \sin \alpha \\ &= \left(m_3g + m_2g - m_1g \frac{2R_1}{R_1 - r_1} \sin \alpha \right) \cdot \delta x_3 = 0 \quad . \end{aligned}$$

Die virtuelle Arbeit kann für beliebige virtuelle Verschiebungen nur Null sein, wenn der Ausdruck in der Klammer verschwindet. Daher ist das System für folgenden Wert von m_3 des Gegengewichtes im Gleichgewicht:

$$m_3 = 2m_1 \frac{R_1}{R_1 - r_1} \sin \alpha - m_2 \quad .$$

Das Beispiel zeigt den Vorteil des Prinzips der virtuellen Arbeit gegenüber dem Prinzip von d'Alembert in dem speziellen Fall der vorliegenden Fragestellung. Weil das System nicht freigeschnitten werden musste, gehen die Zwangskräfte (Normalkraft und Haftkraft zwischen Walze und schiefer Ebene, Seilkraft,...) nicht in die Rechnung ein. Das äußere Gleichgewicht von Systemen starrer Körper kann dadurch recht einfach analysiert werden. Sind gerade die Zwangskräfte gefragt, so liefert die Gleichgewichtsbedingung nach d'Alembert diese sehr viel einfacher.

Man kann natürlich auch beide Prinzipien miteinander kombinieren. Dazu soll der Fall der Gleichgewichtsgleichung von d'Alembert für den Fall betrachtet werden, dass auf einen Massenpunkt mit der Masse m nur eingeprägte Kräfte, Zwangskräfte und die Trägheitskraft (d'Alembert'sche Kraft $m\vec{a}$), aber keine Bewegungswiderstände (Reibung) wirken:

$$\vec{F}_e + \vec{F}_z - m\vec{a} = 0$$

Da nach Gl. (3.78) nur tangential zur Bahn einer bewegten Masse wirkende Kräfte Arbeit verrichten, die Zwangskräfte einer Führung aber gerade normal zur Bahn stehen, gilt für ihre virtuelle Arbeit:

$$\vec{F}_z \cdot \delta \vec{r} = 0$$

Multipliziert man nun den Rest der Gleichung, die aus dem d'Alembert'schen Prinzip folgt, mit der virtuellen Verschiebung $\delta \vec{r}$, so geht diese in die auf Joseph Louis

Comte de Lagrange (1736 - 1813) zurückgehende Form über:

$$\delta W = (\vec{F}_e - m\vec{a}) \cdot \delta \vec{r} = 0 \quad (3.87)$$

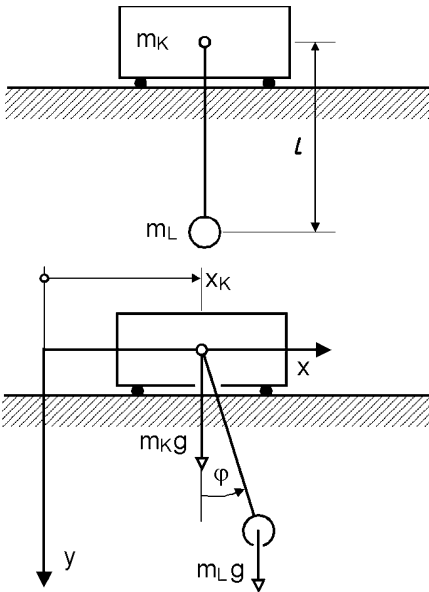
Dies bedeutet, dass ein Massenpunkt sich immer so bewegt, dass bei einer virtuellen Verschiebung die Summe der von den eingepprägten Kräften und den d'Alembertschen Kräften geleisteten Arbeit verschwindet.

Für ein System von n Massenpunkten, die starr miteinander verbunden sind, kann man Gl. (3.87) folgendermaßen schreiben:

$$\delta W = \sum_{i=1}^n (\vec{F}_{i,e} - m_i \vec{a}_i) \cdot \delta \vec{r}_i = 0 \quad (3.88)$$

weil auch die Kräfte in den starren Verbindungen keinen Beitrag zur virtuellen Arbeit leisten. Dies gilt allerdings nicht für Systeme mit elastischen Verbindungen (Federn), da dort nicht wie bei starren Verbindungen die Kraftangriffspunkte an den Enden der Verbindungsglieder den gleichen virtuellen Verschiebungen ausgesetzt sind, so dass im elastischen Fall die Kräfte in den Verbindungen virtuelle Arbeit leisten.

Die Aussagen der Gl. (3.87) und (3.88) können sinngemäß auf starre Körper übertragen werden. Hier kommen zu den Kräften (eingepprägte Kräfte, d'Alembert'sche Kräfte) noch Momente (eingepprägte Momente, d'Alembert'sche Momente) hinzu, deren virtuelle Arbeit analog zu Gl. (3.84) berechnet werden kann. Bei Verwendung generalisierter Koordinaten erhält man wieder eine Gleichung vom Typ der Gl. (3.86), die bei f Freiheitsgraden f Bewegungsgleichungen liefert.



Als Beispiel für die Aufstellung der Bewegungsgleichungen für zwei starr miteinander gekoppelte Massenpunkte möge das in Bild 3-41 dargestellte System aus einer Masse m_L dienen, die über ein masseloses, dehnstarres Seil der Länge l mit einer Laufkatze der Masse m_K verbunden ist [1.5]. Fährt die Laufkatze an, oder bremst sie, so führt die Last Pendelschwingungen aus. Da das System zwei Freiheitsgrade besitzt, werden zur Beschreibung der Position der beiden Massen in der (x,y) -Ebene die generalisierten Koordinaten x_K (horizontale Verschiebung der Laufkatze) und φ (Pendelwinkel der Last) verwendet. Bei Systemen mit mehreren Freiheitsgraden empfiehlt sich das Aufschreiben der Ortsvektoren in Matrixschreibweise bezüglich eines ortsfesten Koordinatensystems:

Bild 3-41: Ermittlung der Bewegungsgleichung für zwei über eine starre Stange gekoppelte Massen

$$\vec{r}_K = \begin{pmatrix} x_K \\ 0 \end{pmatrix} \Rightarrow \delta \vec{r}_K = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \delta x_K, \quad \vec{a}_K = \frac{d^2 \vec{r}_K}{dt^2} = \begin{pmatrix} \ddot{x}_K \\ 0 \end{pmatrix},$$

$$\vec{r}_L = \begin{pmatrix} x_K + l \sin \varphi \\ l \cos \varphi \end{pmatrix} \Rightarrow \delta \vec{r}_L = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \delta x_K + \begin{pmatrix} l \cos \varphi \\ -l \sin \varphi \end{pmatrix} \cdot \delta \varphi,$$

$$\vec{a}_L = \frac{d^2 \vec{r}_L}{dt^2} = \begin{pmatrix} \ddot{x}_K - l \dot{\varphi}^2 \sin \varphi + l \ddot{\varphi} \cos \varphi \\ -l \dot{\varphi}^2 \cos \varphi - l \ddot{\varphi} \sin \varphi \end{pmatrix}.$$

Setzt man diese Vektoren in Gl. (3.88) ein, so ergibt sich:

$$(\vec{F}_{K,e} - m_K \vec{a}_K) \cdot \delta \vec{r}_K + (\vec{F}_{L,e} - m_L \vec{a}_L) \cdot \delta \vec{r}_L = 0,$$

$$\text{mit } \vec{F}_{K,e} = \begin{pmatrix} 0 \\ m_K g \end{pmatrix}, \quad \vec{F}_{L,e} = \begin{pmatrix} 0 \\ m_L g \end{pmatrix}.$$

Nach Einsetzen der Kraftvektoren erhält man:

$$\begin{aligned} & [-(m_K + m_L) \ddot{x}_K - m_L l \ddot{\varphi} \cos \varphi + m_L l \dot{\varphi}^2 \sin \varphi] \cdot \delta x_K \\ & + [-m_L l \ddot{x}_K \cos \varphi - m_L l^2 \ddot{\varphi} - m_L g l \sin \varphi] \cdot \delta \varphi = 0. \end{aligned}$$

Für beliebige virtuelle Verrückungen δx_K und $\delta \varphi$ kann diese Gleichung nur erfüllt sein, wenn die beiden in eckigen Klammern stehenden Ausdrücke gleichzeitig Null sind. Daraus folgt:

$$\begin{aligned} \ddot{x}_K (m_K + m_L) + m_L l \ddot{\varphi} \cos \varphi &= m_L l \dot{\varphi}^2 \sin \varphi \\ \ddot{x}_K \cos \varphi + l \ddot{\varphi} &= -g \sin \varphi. \end{aligned}$$

Dieses nichtlineare Gleichungssystem, das sogar eine Koppelung in den Beschleunigungen besitzt, ist für beliebige Anfangsbedingungen nicht geschlossen lösbar und kann nur durch numerische Näherungsverfahren berechnet werden.

Das Prinzip von d'Alembert in der Lagrange'schen Fassung kann sehr gut zum Aufstellen der Bewegungsgleichungen von Systemen starrer Körper, die durch starre Verbindungen miteinander gekoppelt sind, verwendet werden. Sollen auch nicht-starre Verbindungen erfasst werden können, so kann man das im folgenden Abschnitt beschriebene Verfahren verwenden.

3.4.3.2 Lagrange'sche Bewegungsgleichungen

Zuerst soll ein System von n Massenpunkten m_i betrachtet werden, die untereinander durch nicht-starre Verbindungen gekoppelt sind. Ihre Lage kann in einem kartesischen Koordinatensystem durch die Ortsvektoren $\vec{r}_i$ beschrieben werden. Das Massenpunktsystem möge f Freiheitsgrade besitzen, so dass seine Lage eindeutig durch f generalisierte Koordinaten q_j beschrieben werden kann. Um das Prinzip von d'Alembert nach Gl. (3.88) anwenden zu können, muss man sich die nicht-starren Verbindungen aufgetrennt denken, so dass man die in ihnen wirkenden Kräfte als äußere Kräfte behandeln kann und es gilt:

$$\delta W = \sum_{i=1}^n (\vec{F}_{i,e} - m_i \vec{a}_i) \cdot \delta \vec{r}_i = \sum_{i=1}^n \vec{F}_{i,e} \cdot \delta \vec{r}_i - \sum_{i=1}^n m_i \vec{a}_i \cdot \delta \vec{r}_i = \delta W_e - \delta W_m = 0. \quad (3.89)$$

Der erste Summenterm der Gleichung beschreibt die virtuelle Arbeit der eingepprägten Kräfte δW_e , der zweite Term die virtuelle Arbeit der Massenkräfte δW_m . Diese sollen nun weiter untersucht werden. In Gl. (3.85) wurde ein Ausdruck für die virtuellen Verschiebungen $\delta \vec{r}_i$ in Abhängigkeit von den generalisierten Koordinaten q_j ermittelt. Setzt man dies in den Ausdruck für die virtuelle Arbeit der eingepprägten Kräfte δW_e ein, so erhält man folgendes:

$$\begin{aligned}
 \delta W_e &= \sum_{i=1}^n \vec{F}_{i,e} \cdot \delta \vec{r}_i = \sum_{i=1}^n \vec{F}_{i,e} \cdot \left(\frac{\partial \vec{r}_i}{\partial q_1} \delta q_1 + \dots + \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j + \dots + \frac{\partial \vec{r}_i}{\partial q_f} \delta q_f \right) \\
 &= \sum_{i=1}^n \vec{F}_{i,e} \cdot \frac{\partial \vec{r}_i}{\partial q_1} \delta q_1 + \dots + \sum_{i=1}^n \vec{F}_{i,e} \cdot \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j + \dots + \sum_{i=1}^n \vec{F}_{i,e} \cdot \frac{\partial \vec{r}_i}{\partial q_f} \delta q_f \quad (3.90) \\
 &= Q_1 \cdot \delta q_1 + \dots + Q_j \cdot \delta q_j + \dots + Q_f \cdot \delta q_f \\
 &= \sum_{j=1}^f Q_j \cdot \delta q_j \quad \text{mit} \quad Q_j = \sum_{i=1}^n \vec{F}_{i,e} \cdot \frac{\partial \vec{r}_i}{\partial q_j}
 \end{aligned}$$

Die Größen Q_j , deren Produkt mit den virtuellen Verschiebungen die von den eingepprägten Kräften verrichtete virtuelle Arbeit ergeben, bezeichnet man als **generalisierte Kräfte**.

Sind die auf eine Masse einwirkenden Kräfte ausschließlich Potentialkräfte, so ist die Zunahme der virtuellen Arbeit durch eine virtuelle Verschiebung unter Einfluss einer Kraft gleich der Abnahme der potentiellen Energie der Masse:

$$\delta W_e = -\delta U = -\frac{\partial U}{\partial q_1} \delta q_1 - \dots - \frac{\partial U}{\partial q_j} \delta q_j - \dots - \frac{\partial U}{\partial q_f} \delta q_f \quad .$$

Vergleicht man dies mit Gl. (3.90) so sieht man, dass sich in diesem Fall die generalisierten Kräfte relativ einfach aus der potentiellen Energie berechnen lassen:

$$Q_j = \frac{\partial U}{\partial q_j} \quad . \quad (3.91)$$

Um die virtuelle Arbeit der Massenkräfte δW_m aus Gl. (3.89) zu ermitteln, setzt man wieder den Ausdruck aus Gl. (3.85) für die virtuellen Verschiebungen $\delta \vec{r}_i$ ein:

$$\begin{aligned}
 \delta W_m &= \sum_{i=1}^n m_i \vec{a}_i \cdot \delta \vec{r}_i = \sum_{i=1}^n m_i \vec{a}_i \cdot \left(\frac{\partial \vec{r}_i}{\partial q_1} \delta q_1 + \dots + \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j + \dots + \frac{\partial \vec{r}_i}{\partial q_f} \delta q_f \right) \\
 &= \sum_{i=1}^n \left(m_i \vec{a}_i \sum_{j=1}^f \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \right) = \sum_{j=1}^f \sum_{i=1}^n m_i \vec{a}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \quad . \quad (3.92)
 \end{aligned}$$

Man kann nun den Zusammenhang zwischen der virtuellen Arbeit der Massenkräfte des Massenpunktsystems und der kinetischen Energie ableiten, für die nach Gl. (3.80) gilt:

$$T = \frac{1}{2} \sum_{i=1}^n m_i \vec{v}_i^2 \quad .$$

Für die Geschwindigkeiten der Massenpunkte gilt:

$$\vec{v}_i = \frac{d\vec{r}_i}{dt} = \frac{\partial \vec{r}_i}{\partial q_1} \dot{q}_1 + \dots + \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j + \dots + \frac{\partial \vec{r}_i}{\partial q_f} \dot{q}_f \quad (3.93)$$

Die Ableitungen der generalisierten Koordinaten nach der Zeit sind die **generalisierten Geschwindigkeiten** $\dot{q}_j$.

Zur Herleitung des gesuchten Zusammenhanges zwischen virtueller Arbeit und kinetischer Energie wird diese zunächst partiell nach q_j und $\dot{q}_j$ abgeleitet:

$$\frac{\partial T}{\partial q_j} = \sum_{i=1}^n m_i \vec{v}_i \frac{\partial \vec{v}_i}{\partial q_j} \quad , \quad \frac{\partial T}{\partial \dot{q}_j} = \sum_{i=1}^n m_i \vec{v}_i \frac{\partial \vec{v}_i}{\partial \dot{q}_j} = \sum_{i=1}^n m_i \vec{v}_i \frac{\partial \vec{r}_i}{\partial q_j}$$

Das letzte Gleichheitszeichen ist gültig, weil die Ortsvektoren nicht von den generalisierten Geschwindigkeiten abhängig sind und daher folgende Gleichung gilt:

$$\frac{\partial \vec{v}_i}{\partial \dot{q}_j} = \frac{\partial \vec{r}_i}{\partial q_j} \quad .$$

Differenziert man jetzt die partielle Ableitung $\frac{\partial T}{\partial \dot{q}_j}$ nochmals nach der Zeit, so erhält man:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) = \sum_{i=1}^n m_i \frac{d\vec{v}_i}{dt} \frac{\partial \vec{r}_i}{\partial q_j} + \sum_{i=1}^n m_i \vec{v}_i \frac{d}{dt} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right) = \sum_{i=1}^n m_i \vec{a}_i \frac{\partial \vec{r}_i}{\partial q_j} + \underbrace{\sum_{i=1}^n m_i \vec{v}_i \frac{\partial \vec{v}_i}{\partial q_j}}_{\frac{\partial T}{\partial q_j}}$$

Diese Gleichung kann man auch folgendermaßen schreiben:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = \sum_{i=1}^n m_i \vec{a}_i \frac{\partial \vec{r}_i}{\partial q_j} \quad (3.94)$$

Die rechte Seite der Gl. (3.94) entspricht genau der inneren Summe des Ausdruckes, der in Gl. (3.92) für die virtuelle Arbeit der Massenkräfte gefunden wurde. Man kann diese virtuelle Arbeit demnach folgendermaßen formulieren:

$$\delta W_m = \sum_{j=1}^f \sum_{i=1}^n m_i \vec{a}_i \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j = \sum_{j=1}^f \left[\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] \delta q_j \quad (3.95)$$

Ausgangspunkt der Betrachtung war die Gl. (3.89) mit der Formulierung des d'Alembert'schen Prinzips in der Fassung von Lagrange für einen Massenpunkthaufen mit nicht-starren Verbindungen. In der obigen Betrachtung wurden nun Ausdrücke für die virtuelle Arbeit der eingepprägten Kräfte und für die virtuelle Arbeit der Massenkräfte abgeleitet. Setzt man die Ergebnisse aus Gl. (3.90) und Gl. (3.95) in Gl. (3.89) ein, so erhält man:

$$\delta W = \delta W_e - \delta W_m = \sum_{j=1}^f \left\{ Q_j - \left[\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] \right\} \delta q_j \quad (3.96)$$

Da die f virtuellen Verschiebungen δq_j von einander unabhängig sind, kann Gl. (3.96) nur erfüllt sein, wenn der Ausdruck in der geschweiften Klammer für jeden einzelnen Wert von j gleich Null ist. Die aus dieser Tatsache ableitbaren Beziehungen heißen **Lagrange'sche Gleichungen der 2. Art** :

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = Q_j \quad , \quad j = 1, 2, \dots, f \quad . \quad (3.97)$$

Handelt es sich bei dem Massesystem um ein konservatives System, in dem nur Potentialkräfte wirken, so kann man für die Q_j die Ausdrücke aus Gl. (3.91) einsetzen:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} + \frac{\partial U}{\partial q_j} = 0 \quad \Rightarrow \quad \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial (T - U)}{\partial q_j} = 0 \quad .$$

Da die potentielle Energie U von den generalisierten Geschwindigkeiten $\partial \dot{q}_j$ unabhängig ist gilt außerdem:

$$\frac{\partial (T - U)}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} \quad .$$

Damit erhält man die Lagrangeschen Gleichungen 2. Art für konservative Systeme:

$$\frac{d}{dt} \left(\frac{\partial (T - U)}{\partial \dot{q}_j} \right) - \frac{\partial (T - U)}{\partial q_j} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = 0 \quad j = 1, 2, \dots, f \quad (3.98)$$

wobei $L = T - U$ die **Lagrange'sche Funktion** heißt. Je nach Anzahl der Freiheitsgrade f des Systems erhält man f Bewegungsgleichungen zur Ermittlung des Bewegungszustandes des Systems.

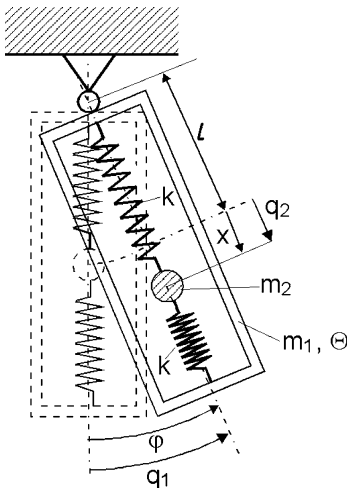


Bild 3-42: Zusammengesetztes Pendel (Rohrpendel)

Die Lagrangeschen Gleichungen 2. Art (Gl. (3.97) und Gl. (3.98)) wurden hier für Massensystemen hergeleitet, gelten aber auch für Systeme starrer Körper. Ein Teil der generalisierten Koordinaten können dann natürlich auch Winkel sein, so dass in den Gl. (3.90) und (3.91) statt generalisierter Kräfte **generalisierte Momente** auftreten. In die kinetische Energie der Körper fließen dann auch rotatorische Anteile ein.

Die Anwendung der Lagrange'schen Gleichungen 2. Art soll nun an dem Beispiel eines Pendels verdeutlicht werden, das sich aus mehreren Teilmassen und Federn zusammensetzt [3.2]. Es besteht, wie in Bild 3-42 dargestellt, aus einer Einzelmasse, die zwischen zwei Federn mit der Federsteifigkeit k in einem Rohr eingespannt ist. Das Rohr ist um einen Aufhängepunkt drehbar gelagert. Die im Rohr befindliche Masse verlagert sich aus der Ruhelage im Rohr, wenn dieses in Schwingungen versetzt wird. Seine Position

in der (x,y) - Ebene kann durch zwei generalisierte Koordinaten beschrieben werden: $\varphi = q_1$, $x = q_2$. Dabei ist φ die Verdrehung des Rohres mit der Masse m_1 und dem Trägheitsmoment Θ gegenüber der Ruhelage des Pendels (gestrichelt) und x die Verschiebung der Punktmasse m_2 . Das Trägheitsmoment des Rohres bezieht sich auf den Aufhängepunkt.

In der Ruhelage wird das Potential der Massen zu Null gesetzt. Um die Lagrange'sche Funktion aufstellen zu können, müssen zuerst die potentielle und kinetische Energie des Systems in Abhängigkeit von q_1 und q_2 bestimmt werden. Für die potentielle Energie der Federn, die wegen der Parallelschaltung² die Gesamtfederkonstante $2k$ besitzen, gilt:

$$U = U_1 - U_2 = \int_0^{q_2} F dq = \int_0^{q_2} (2k \cdot x) dq = 2 \frac{k}{2} q_2^2 = k q_2^2.$$

Die Verschiebung der Schwerpunkte S_1 des Rohres und des Schwerpunktes S_2 der Punktmasse im Schwerfeld der Erde erfolgt um die in Bild 3-43 dargestellten Strecken, so dass deren potentielle Energie sich um folgende Beträge ändert:

$$U(S_1) = l(1 - \cos q_1) m_1 g$$

$$U(S_2) = (l - (l + q_2) - \cos q_1) m_2 g$$

Die gesamte potentielle Energie des Systems in Abhängigkeit von q_1 und q_2 beträgt demnach:

$$U = k q_2^2 + l \cdot g (m_1 + m_2) (1 - \cos q_1) - m_2 g q_2 \cos q_1$$

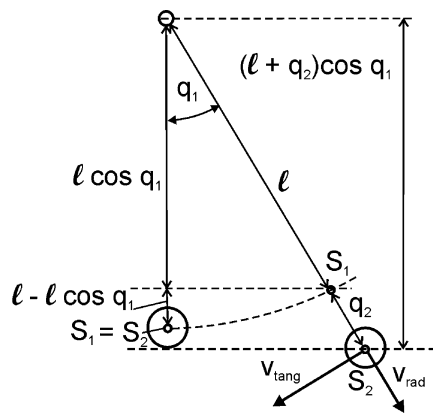


Bild 3-43: Schwerpunktverlagerung beim Rohrpendel

Das Rohr mit der Masse m_1 führt eine reine Rotationsbewegung um den Aufhängepunkt aus. Seine kinetische Energie ist daher:

$$T(m_1) = \frac{1}{2} \Theta \dot{q}_1^2.$$

Die Punktmasse m_2 hat aufgrund der Pendelbewegung eine tangentielle Geschwindigkeitskomponente v_{tang} und aufgrund der Dehnung der Federn eine radiale Geschwindigkeitskomponente v_{rot} . Es gilt:

$$v_{\text{tang}} = (l + q_2) \dot{q}_1, \quad v_{\text{rad}} = \dot{q}_2 \Rightarrow T(m_2) = \frac{1}{2} m_2 (l + q_2)^2 \dot{q}_1^2 + \frac{1}{2} m_2 \dot{q}_2^2.$$

Damit beträgt die kinetische Energie des Pendels:

$$T = \frac{1}{2} \Theta \dot{q}_1^2 + \frac{1}{2} m_2 (l + q_2)^2 \dot{q}_1^2 + \frac{1}{2} m_2 \dot{q}_2^2.$$

2 Die Federn scheinen geometrisch in Reihe geschaltet zu sein, sind aber tatsächlich kräftemäßig parallel geschaltet.

Für die Lagrange'sche Funktion ergibt sich demnach:

$$L = T - U = \frac{1}{2} \Theta \dot{q}_1^2 + \frac{1}{2} m_2 (l + q_2)^2 \dot{q}_1^2 + \frac{1}{2} m_2 \dot{q}_2^2 - k q_2^2 - l \cdot g (m_1 + m_2) + l \cdot g (m_1 + m_2) \cos q_1 + m_2 g q_2 \cos q_1 .$$

Zur Aufstellung der Lagrange'schen Bewegungsgleichung müssen nun verschiedene Ableitungen der Lagrange'schen Funktion gebildet werden:

$$\frac{\partial L}{\partial \dot{q}_1} = \Theta \dot{q}_1 + m_2 \dot{q}_1 (l + q_2)^2$$

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_1} \right) = \Theta \ddot{q}_1 + m_2 \ddot{q}_1 (l + q_2)^2 + m_2 \dot{q}_1 2(l + q_2) \dot{q}_2$$

$$\frac{\partial L}{\partial \dot{q}_2} = m_2 \dot{q}_2$$

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_2} \right) = m_2 \ddot{q}_2$$

$$\frac{\partial L}{\partial q_1} = -l \cdot g (m_1 + m_2) \sin q_1 - m_2 g q_2 \sin q_1$$

$$\frac{\partial L}{\partial q_2} = m_2 \dot{q}_1^2 (l + q_2) - 2k q_2 + m_2 g \cos q_1 .$$

Setzt man die oben hergeleiteten Ableitungen in die Gl. (3.98) ein, so erhält man die Bewegungsgleichungen des Systems:

$$(\Theta + m_2 \dot{q}_1 (l + q_2)^2) \cdot \ddot{q}_1 + 2m_2 (l + q_2) \dot{q}_1 \dot{q}_2 + (l \cdot g (m_1 + m_2) + m_2 g q_2) \sin q_1 = 0 \quad (3.99)$$

$$m_2 \ddot{q}_2 - m_2 \dot{q}_1^2 (l + q_2) + 2k q_2 - m_2 g \cos q_1 = 0 \quad (3.99a)$$

Dieses nichtlineare Gleichungssystem, das einige Koppelungen besitzt, ist für beliebige Anfangsbedingungen nicht geschlossen lösbar und kann nur durch numerische Näherungsverfahren berechnet werden.

Über die bis hierher kennengelernten Verfahren und Methoden hinaus gibt es noch weitere, die in bestimmten Fällen das Aufstellen der Bewegungsgleichungen vereinfachen oder erst ermöglichen. Hierfür sei auf die bereits genannte Spezialliteratur verwiesen. Mit den in den vorigen Abschnitten des Kapitels 3 behandelten Methoden der Dynamik können nun aber bereits die Bewegungen der mechanischen Teile vieler mechatronischer Systeme beschrieben werden.

Der nächste Schritt muss nun sein, aus den gefundenen Bewegungsgleichungen Lösungen für den Weg, die Geschwindigkeit und die Beschleunigung zu beliebigen Zeitpunkten der Bewegung zu ermitteln. Ohne auf Digitalrechnern durchgeführte Näherungsverfahren, sind nur wenige Differentialgleichungen für beliebige Anfangsbedingungen geschlossen lösbar.

Viele Bewegungsvorgänge in der Mechanik, insbesondere Schwingungen, kann man jedoch bei geschickter Modellbildung durch lineare Differentialgleichungen be-

schreiben, für die es Lösungsverfahren gibt. Andere Bewegungsvorgänge, deren Bewegungsgleichungen Nichtlinearitäten enthalten, kann man durch entsprechende Einschränkungen des Bewegungsbereiches (Arbeitspunkt) linearisieren (s. Beispiel des Pendels in Abschnitt 2.2 und invertiertes Pendel in Abschnitt 9.2). Dies wird vor allem im Zusammenhang mit der Regelung eines Bewegungsvorgangs der Fall sein, bei dem häufig die Aufgabe darin besteht, das System in einem bestimmten Systemzustand zu halten.

Eine weitere häufig auftretende Form dynamischer Veränderungen in mechanischen Systemen ist die Schwingung, ein mehr oder weniger periodischer Prozess der Systembewegung. Entsprechend der in Kapitel 1 behandelten Analogie des Systemverhaltens mechanischer und elektrischer Systeme, kann das zu beschreibende dynamische Verhalten in den verschiedensten Bereichen mechatronischer Systeme auftreten. Im folgenden sollen nun Schwingungen, die Aufstellung der Bewegungsgleichungen und deren Lösung behandelt werden. Aufgrund der oben angeführten Analogie, können die Aussagen über Schwingungen in mechanischen Systemen auf vergleichbare elektrische Systeme übertragen werden.

4 Schwingungen

4.1 Einmassenschwinger

Für viele mechanische Strukturen ist das einfachste Modell der Einmassenschwinger, so wie er in Bild 4-1 dargestellt ist. Er besteht aus einer Masse, einer Feder und einem Dämpfungselement. Dieses Modell geht davon aus, dass alle Massen der Anordnung in der Masse konzentriert sind und somit die Feder und der Dämpfer masselos sind. Die aus Gründen der besseren Darstellbarkeit räumlich gezeichnete Masse muss man sich als Punktmasse denken. Weiterhin wird in der Regel für die Feder die Gültigkeit des Hooke'schen Gesetzes (Robert Hooke: engl. Naturphilosoph 1635 - 1703) angenommen. Es besagt, dass die Verformung eines elastischen Körpers der verformenden Kraft proportional ist.

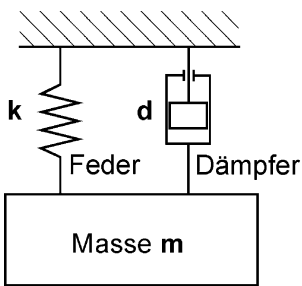


Bild 4-1: Einmassenschwinger

Dies gilt natürlich nur im so genannten elastischen Bereich des Körpers; bei sehr großen Kräften tritt plastische Verformung auf, ein nichtlineares Verhalten. Die Feder wird daher als ein Element angenommen, bei dem ein linearer Zusammenhang zwischen Belastung und Dehnung besteht. Das Dämpfungselement ist ein viskoser Dämpfer, für das ein linearer Zusammenhang zwischen der Dämpfungskraft und der Relativgeschwindigkeit der sich bewegenden Dämpferteile besteht. Diese Gesetzmäßigkeit wird auch als Newton'sches Reibungsgesetz bezeichnet. Die Masse selber hat nur einen Freiheitsgrad, kann sich also nur in vertikaler Richtung in der Bildebene bewegen. Um ein solches System in Schwingungen zu versetzen, muss eine äußere Erregerkraft auf die Masse einwirken.

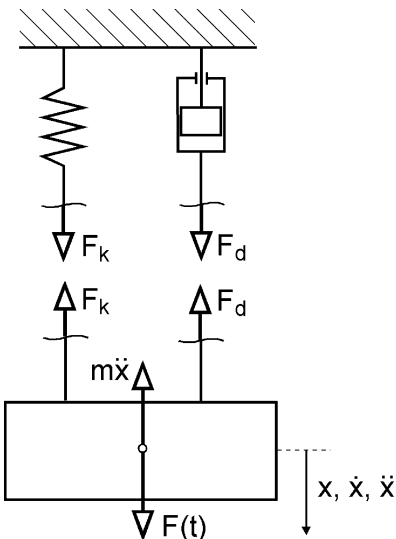


Bild 4-2: Anwendung des Schnittprinzips auf einen Einmassenschwinger

Ein solcher **Einmassenschwinger** kann als Modell für einfache technische Strukturen wie Wellen, Balken, Stäbe, Platten, aber auch in erster Näherung für kompliziertere Systeme wie beispielsweise das PKW-Federbein aus Bild 2-10 stehen. Da es nur einen Freiheitsgrad haben soll, reicht eine Koordinate zur Beschreibung des Bewegungsverhaltens aus. Um dieses Modell mathematisch beschreiben zu können, muss die Bewegungsgleichung für die Punktmasse des Systems formuliert werden. Dies geht sehr einfach durch Anwendung des in Kapitel 3 beschriebenen d'Alembert'schen Prinzips. Dazu muss zuerst das Schnittprinzip angewendet werden, d. h. alle an der Masse angreifenden Verbindungen werden aufgeschnitten und dafür an der Masse angreifende Kräfte angetragen (Bild 4-2). Diese Kräfte sind:

$$F_k = k \cdot x$$

Federkraft

$F_d = d \cdot \dot{x}$	Dämpfungskraft
$F(t) :$	Erregerkraft
$m \cdot \ddot{x} :$	Trägheitskraft

Nach dem d'Alembert'schen Prinzip gilt:

$$F(t) - F_k - F_d - m\ddot{x} = 0$$

Bei einer Anordnung wie in Bild 4-2 geht man normalerweise davon aus, dass in der statischen Ruhelage das Gewicht der Masse gerade durch eine entsprechend große aber entgegengesetzt gerichtete Anfangskraft der Feder kompensiert wird, so dass die Gewichtskraft für die dynamischen Betrachtungen nicht berücksichtigt werden muss.

Nun kann man die bekannten Kräfte einsetzen und nach der Ordnung der Ableitungen des Weges sortieren:

$$m\ddot{x} + d\dot{x} + kx = F(t) \quad (4.1)$$

Dies ist die Differentialgleichung der Bewegung für einen Einmassenschwinger mit einem Freiheitsgrad. Es handelt sich um eine **lineare, gewöhnliche Differentialgleichung mit konstanten Koeffizienten**. Dabei bedeutet **linear**, dass die Variable x und ihre Ableitungen nicht durch nichtlineare Funktionen wie beispielsweise x^2 , $\sin x$, o. ä. beeinflusst sind, oder in Kombinationen wie $x \cdot \dot{x}$ vorkommen. Von **gewöhnlichen** Differentialgleichungen spricht man, wenn nur eine Variable (hier x) und ihre Ableitungen in der Gleichung auftreten. Bezeichnet man die Größe x als Ausgangsgröße $x_a(t)$ des Systems und ist $F(t)$ eine Funktion der Eingangsgröße $x_e(t)$, die sich wiederum als eine Summe von x_e und ihren Ableitungen mit konstanten Koeffizienten schreiben lässt, so lautet die allgemeine Form :

$$b_n x_a^{(n)}(t) + \dots + b_1 \dot{x}_a(t) + b_0 x_a(t) = a_0 x_e(t) + \dots + a_m x_e^{(m)}(t) \quad (4.2)$$

Für Gleichungen dieses Typs, die man auch als **Differentialgleichungen n-ter Ordnung** bezeichnet, gibt es gute analytische Lösungsmethoden, so dass man ohne Rechnerunterstützung eine geschlossene Lösung angeben kann. Man kann leicht erkennen, dass man die Lösung einer DGL n-ter Ordnung durch n-fache Integration der DGL erhält. Dabei entstehen n Integrationskonstanten C_1 bis C_n . Die Lösung einer DGL n-ter Ordnung enthält daher n willkürliche Parameter. Die Menge aller Lösungsfunktionen, die durch diese n Parameter gegeben sind, nennt man die **allgemeine Lösung** der DGL. Weist man, z. B. aufgrund bekannter Anfangsbedingungen, den Lösungsparametern bestimmte Werte zu, so erhält man eine **partikuläre Lösung**.

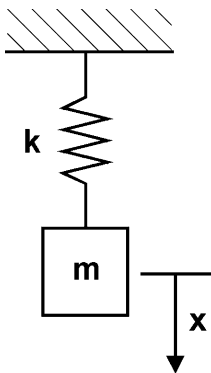
Um die analytische Lösung der Differentialgleichung (DGL) herzuleiten, betrachtet man meist zuerst einfachere Sonderfälle. Ein solcher Sonderfall liegt vor, wenn die rechte Seite der Gl. (4.2) verschwindet, wenn also keine äußere Erregung des Systems vorliegt. Die verbleibende DGL bezeichnet man als **homogene DGL**, die die sog. **Eigenvorgänge** des Systems beschreibt. Dies ist das dynamische Systemverhalten ohne äußere Krafteinflüsse. Die durch Integration auffindbare Lösung der homogenen DGL ist die homogene Lösung $x_{ah}(t)$. Es lässt sich zeigen, dass die all-

gemeine Lösung der DGL unter Verwendung einer partikulären Lösung $x_{ap}(t)$ für bestimmte Anfangsbedingungen folgendermaßen geschrieben werden kann:

$$x_a(t) = x_{ah}(t) + x_{ap}(t) \quad (4.3)$$

Für das obige Beispiel ist der einfachste zu behandelnde Fall die ungedämpfte (Dämpfer weggelassen) Schwingung ohne äußere Anregung, also der Eigenvorgang des ungedämpften Systems.

4.1.1 Freie ungedämpfte Schwingungen



Die Bewegungsgleichung für den nicht erregten Schwinger ohne Dämpfung aus Bild 4-3 ermittelt man nach der oben aufgezeigten Methode zu:

$$m\ddot{x} + kx = 0$$

Zur Lösung dieser DGL kann man den sog. **klassischen Ansatz** verwenden. Dieser homogene Lösungsansatz und seine Ableitungen lauten:

$$x(t) = \hat{x} \cdot e^{st}$$

$$\dot{x}(t) = \hat{x} \cdot s \cdot e^{st}$$

$$\ddot{x}(t) = \hat{x} \cdot s^2 \cdot e^{st} .$$

Bild 4-3:

Schwinger ohne Dämpfung und Erregung

Setzt man die entsprechenden Ausdrücke in die DGL ein, so erhält man:

$$m \cdot \hat{x} \cdot s^2 \cdot e^{st} + k \cdot \hat{x} \cdot e^{st} = 0 = m \cdot s^2 + k .$$

Den auf der rechten Seite der Gleichung gefundenen Ausdruck bezeichnet man als **charakteristische Gleichung**, deren Lösungen (Wurzeln des **charakteristischen Polynoms**) die **Eigenwerte** des Systems genannt werden. Die Bezeichnung **Eigenwerte** deutet an, dass die homogene DGL den Eigenvorgang des autonomen Systems beschreibt. Die Eigenwerte sind:

$$s_{1,2} = \pm \sqrt{-\frac{k}{m}} = \pm i \sqrt{\frac{k}{m}} \quad \text{mit} \quad i^2 = -1 .$$

Die Gesamtlösung der homogenen DGL mit diesen beiden Eigenwerten lautet :

$$x(t) = A \cdot e^{s_1 t} + B \cdot e^{s_2 t} ,$$

oder wegen $s_1 = -s_2$:

$$x(t) = A \cdot e^{s_1 t} + B \cdot e^{-s_1 t} .$$

Die Werte A und B sind die Integrationskonstanten, die noch aus den Anfangsbedingungen ermittelt werden müssen. Als Anfangsbedingungen sind die Anfangsauslenkung x_0 (hier Auslenkung der Feder aus statischer Ruhelage, welche durch die Gewichtskraft vorgegeben wird) und die Anfangsgeschwindigkeit $\dot{x}_0$ bekannt:

$$x(t=0) = x_0 \quad , \quad \dot{x}(t=0) = \dot{x}_0 \quad .$$

Es gilt dann:

$$x(0) = A + B = x_0 \quad , \quad \dot{x}(0) = A \cdot s_1 - B \cdot s_1 = \dot{x}_0 \quad .$$

Die Integrationskonstanten A und B haben damit folgende Werte:

$$A = \frac{s_1 x_0 + \dot{x}_0}{2s_1} \quad , \quad B = \frac{s_1 x_0 - \dot{x}_0}{2s_1} \quad .$$

Für den Wurzelausdruck in den beiden Eigenwerten benutzt man folgende Abkürzung:

$$\omega = \omega_0 = \sqrt{\frac{k}{m}} \quad \textbf{Eigenkreisfrequenz},$$

womit man diese als

$$s_{1,2} = \pm i\omega$$

schreiben kann, d. h. die beiden Eigenwerte sind konjugiert komplex und liegen auf der imaginären Achse. Damit lautet die Gesamtlösung der DGL:

$$x(t) = \frac{i\omega x_0 + \dot{x}_0}{2i\omega} e^{i\omega t} + \frac{i\omega x_0 - \dot{x}_0}{2i\omega} e^{-i\omega t} \quad .$$

Um das Bewegungsverhalten der Masse in Abhängigkeit von der Zeit darstellen zu können, muss man diese Gleichung in eine reelle Form bringen. Dazu bietet sich die sog. **Euler-Relation** an:

$$e^{i\omega t} = \cos \omega t + i \sin \omega t \quad .$$

Wendet man diese Relation auf die obige Lösung der DGL an, so erhält man

$$\begin{aligned} x(t) &= \left(\frac{x_0}{2} + i \frac{\dot{x}_0}{2\omega} \right) \cdot (\cos \omega t + i \sin \omega t) + \left(\frac{x_0}{2} - i \frac{\dot{x}_0}{2\omega} \right) \cdot (\cos \omega t - i \sin \omega t) \\ &= x_0 \cos \omega t + \frac{\dot{x}_0}{\omega} \sin \omega t \end{aligned}$$

Diese Überlagerung einer Cosinusschwingung mit der Kreisfrequenz ω durch eine Sinusschwingung gleicher Kreisfrequenz ist im Bild 4-4 dargestellt. Um diese Schwingung zeichnen zu können, verfährt man wie unten beschrieben. Eine allgemeine phasenverschobene Cosinusschwingung kann man folgendermaßen darstellen:

$$x(t) = \hat{x} \cos(\omega t + \varphi) \quad .$$

Durch Umformen mit Hilfe eines Additionstheorems der Trigonometrie erhält man daraus:

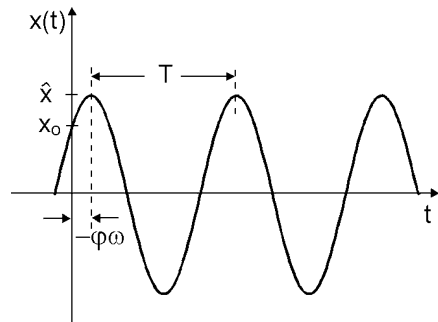


Bild 4-4: Schwingung des ungedämpften Einmassenschwingers

$$x(t) = \hat{x} \cos \varphi \cos \omega t - \hat{x} \sin \varphi \sin \omega t \quad .$$

Vergleicht man dies mit der obigen Lösung der DGL so findet man

$$\hat{x} \cos \varphi = x_0 \quad , \quad \hat{x} \sin \varphi = \frac{\dot{x}_0}{\omega} \quad ,$$

woraus sich $\hat{x}$ und φ berechnen lassen:

$$\hat{x}^2 \underbrace{(\cos^2 \varphi + \sin^2 \varphi)}_{=1} = x_0^2 + \frac{\dot{x}_0^2}{\omega^2} \quad ,$$

$$\frac{\sin \varphi}{\cos \varphi} = \tan \varphi = \frac{\dot{x}_0}{x_0 \omega} \quad ,$$

Die beiden für die Schwingung bestimmenden Größen betragen demnach

$$\hat{x} = \sqrt{x_0^2 + \frac{\dot{x}_0^2}{\omega^2}} \quad \textbf{Amplitude}$$

$$\varphi = \arctan\left(-\frac{\dot{x}_0}{x_0 \omega}\right) \quad \textbf{Anfangsphase}$$

Da es sich hier um eine Schwingung handelt, deren Frequenz nur durch den autonomen Eigenvorgang bestimmt wird, bezeichnet man ω auch als **Eigenkreisfrequenz** mit dem Symbol ω_0 . Man sieht aus Bild 4-4, dass es sich bei dem Schwingungsvorgang um eine harmonische Schwingung handelt, deren Periode

$$T = \frac{2\pi}{\omega}$$

beträgt und die eine **Eigenfrequenz** von

$$f_e = \frac{\omega}{2\pi} = \frac{1}{T}$$

besitzt. Sie weist eine Phasenverschiebung um den Winkel φ gegenüber der reinen Cosinusschwingung auf und hat zum Zeitpunkt $t = 0$ die Anfangsauslenkung x_0 ; die Amplitude beträgt $\hat{x}$.

4.1.2 Freie gedämpfte Schwingungen

Verwendet man das in Bild 4-2 dargestellte Modell des Einmassenschwingers mit viskoser Dämpfung, so erhält man aus der Gl. (4.1) ohne eine äußere Erregungskraft $F(t)$ folgende Bewegungsgleichung für die freie, gedämpfte Schwingung:

$$m\ddot{x} + d\dot{x} + kx = 0 \quad . \quad (4.4)$$

Verwendet man zur Lösung der DGL wieder den klassischen Ansatz

$$x(t) = \hat{x} \cdot e^{st}$$

so erhält man wieder die charakteristische Gleichung

$$ms^2 + ds + k = 0 \quad .$$

Aus dieser quadratischen Gleichung gewinnt man die Eigenwerte des Systems zu:

$$\begin{aligned} s_{1,2} &= -\frac{d}{2m} \pm \sqrt{\left(\frac{d}{2m}\right)^2 - \frac{k}{m}} \\ &= -\frac{d}{2m} \pm i \underbrace{\sqrt{\frac{k}{m} - \left(\frac{d}{2m}\right)^2}}_{\omega_0^2} = \delta \pm i\omega \quad . \end{aligned} \quad (4.5)$$

Durch Einfluss der Dämpfung verschieben sich die Eigenwerte gegenüber denen, die man für den ungedämpften Fall erhält. Man unterscheidet daher zwischen der Eigenkreisfrequenz des ungedämpften Systems ω_0 und der des gedämpften Systems ω (der gesamte Wurzelausdruck). Für die Eigenwerte muss man drei Fälle unterscheiden (siehe auch Bild 4-7), je nachdem ob der Radikand in Gl. (4.5) positiv, Null oder negativ ist. Gilt für den Ausdruck im Radikanden folgende Relation

$$\left(\frac{d}{2m}\right)^2 - \frac{k}{m} > 0 \quad . \quad (4.6)$$

so gibt es zwei reelle Eigenwerte $s_1 = \delta_1$, $s_2 = \delta_2$ und die Lösung der DGL lautet:

$$x(t) = A \cdot e^{\delta_1 t} + B \cdot e^{\delta_2 t} \quad .$$

Dies ist die Addition zweier Exponentialfunktionen, die eine monotone Funktion darstellt. Es liegt also keine Schwingung, sondern ein Kriechvorgang vor. Der Grenzfall zwischen einem Kriechvorgang und einer periodischen Schwingung tritt bei einer kritischen Dämpfung d_{krit} auf, für die der Ausdruck aus Gl. (4.6) gerade Null wird. Es gilt:

$$\left(\frac{d_{krit}}{2m}\right)^2 - \frac{k}{m} = 0 \quad \Rightarrow \quad d_{krit} = 2m \sqrt{\frac{k}{m}} = 2m\omega_0 = 2\sqrt{k \cdot m} \quad .$$

Für diesen Fall sind die beiden Eigenwerte gleich und zwar negativ und reell.

Gilt für den Ausdruck aus Gl. (4.5)

$$\left(\frac{d}{2m}\right)^2 - \frac{k}{m} < 0 \quad ,$$

so sind die Eigenwerte konjugiert komplex und ihr Wert ist

$$s_{1,2} = \delta + i\omega$$

Die Lösung der DGL lautet dann:

$$x(t) = A \cdot e^{s_1 t} + B \cdot e^{s_2 t} = e^{\delta t} (A \cdot e^{i\omega t} + B \cdot e^{-i\omega t}) \quad . \quad (4.7)$$

In der letzten Gleichung stellt der Faktor $e^{\delta t}$ eine Dämpfungsfunktion dar, während

der Ausdruck in der Klammer wieder einen periodischen Vorgang beschreibt, wie wir ihn schon in Abschnitt 4.1.1 kennengelernt haben. Die Integrationskonstanten kann man genauso wie in Abschnitt 4.1.1 gezeigt bestimmen und erhält dann aus Gl. (4.7) unter Anwendung der Euler-Relation:

$$x(t) = e^{\delta t} \left(x_0 \cos \omega t + \left(\frac{\dot{x}_0 - \delta \cdot x_0}{\omega} \right) \sin \omega t \right) . \quad (4.8)$$

Wie der generelle zeitliche Verlauf dieser Funktion aussieht, hängt von dem Wert der Variablen δ ab. Es gilt für:

- $\delta > 0$, die Schwingung wäre nicht gedämpft, sondern zu unendlichen Amplituden aufklingend. Da dies beim betrachteten System¹ physikalisch unmöglich ist, tritt dieser Fall nicht auf.
- $\delta = 0$, die Schwingung verläuft mit ungedämpfter Amplitude (der bereits in 4.1.1 behandelte Fall).
- $\delta < 0$, die Amplitude der Schwingung nimmt ab, es liegt also ein gedämpfter Eigenvorgang vor.

Der Dämpfungsfaktor δ ist eine mit Maßeinheiten behaftete Größe. Will man die oben behandelten verschiedenen Eigenvorgänge übersichtlich klassifizieren, so ist die Einführung eines dimensionslosen Dämpfungsparameters D des **Lehr'schen Dämpfungsmaßes** üblich:

$$D = \frac{d}{d_{krit}} = \frac{d}{2m\omega_0} = \frac{d}{2\sqrt{m \cdot k}} = \frac{|\delta|}{\omega_0} = \frac{|\delta|}{\sqrt{\delta^2 + \omega^2}} .$$

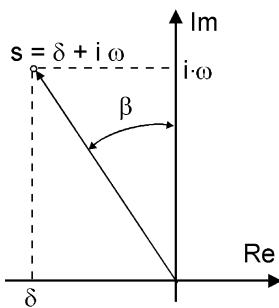


Bild 4-5: Eigenwert in der komplexen Ebene

Setzt man diesen Parameter in Gl. (4.5) ein, so lauten die Eigenwerte:

$$s_{1,2} = \omega_0 \left(-D \pm \sqrt{D^2 - 1} \right) .$$

Bild 4-5 zeigt die Darstellung eines Eigenwertes in der komplexen Ebene. Es ist der **Dämpfungswinkel** β und der Zusammenhang mit den Dämpfungsgrößen D und δ dargestellt. Es gilt:

$$D = \sin \beta \quad , \quad \delta = \tan \beta .$$

Für $D < 1$ sind die beiden Eigenwerte konjugiert komplex (Bild 4-7), der Schwingungsvorgang ist periodisch und gedämpft. Die Lösung der DGL lautet dann:

$$x(t) = e^{-\omega_0 D t} \left(x_0 \cos \omega t + \left(\frac{\dot{x}_0 + D \omega_0 \cdot x_0}{\omega} \right) \sin \omega t \right) . \quad (4.9)$$

1 In anderen realen Systemen kann ein solcher Fall jedoch auftreten. Wird ein Schwingungsvorgang durch äußere Kräfte so angefacht, dass er zu sehr hohen Amplituden aufschwingt, so wird die Werkstofffestigkeit überschritten und das System zerstört. Ein klassisches Beispiel für einen solchen Vorgang ist die Zerstörung der Tacoma Narrows Bridge in den USA im Jahre 1940 (s. Wikipedia, Tacoma_Narrows_Bridge_Destruction.ogg)

Für $D = 1$ gilt:

$$s_{1,2} = \omega_0 \left(-1 \pm \sqrt{1^2 - 1} \right) = -\omega_0$$

d. h. die beiden Eigenwerte sind gleich.

Für die Lösung der DGL gilt:

$$x(t) = e^{-\omega_0 t} (x_0 + (\dot{x}_0 + \omega_0 x_0) \cdot t) .$$

Man sieht, dass das Zeitverhalten des Systems keinen periodischen Charakter mehr besitzt und man spricht vom **aperiodischen Grenzfall**.

Ist $D > 1$, so liegt die sogenannte **überkritische Dämpfung** vor. Auch dieser Bewegungsvorgang ist nicht mehr periodisch, sondern es handelt sich um einen Kriechvorgang. Für die Lösung der DGL gilt:

$$x(t) = e^{-\omega_0 D t} \left(x_0 \cosh \hat{\omega} t + \left(\frac{\dot{x}_0 + D \omega_0 \cdot x_0}{\hat{\omega}} \right) \sinh \hat{\omega} t \right) , \quad \text{mit} \quad \hat{\omega} = \omega_0 \sqrt{D^2 - 1} .$$

In Bild 4-6 ist der Verlauf von Schwingungsvorgängen mit der Anfangsamplitude x_0 in Abhängigkeit von der Größe des Dämpfungsgrades D dargestellt. Im Bild sind verschiedene schwach gedämpfte Schwingungen, der aperiodische Grenzfall und die überkritische Dämpfung dargestellt.

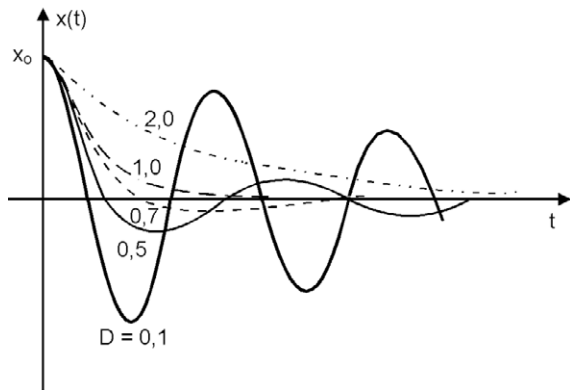


Bild 4-6: Gedämpfte Schwingungen in Abhängigkeit vom Dämpfungsgrad D

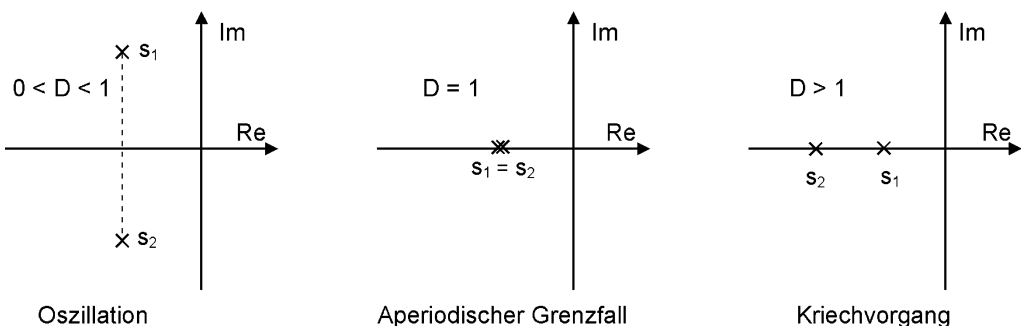


Bild 4-7: Lage der Eigenwerte in der komplexen Ebene bei verschiedenen Dämpfungsgraden

unterscheiden, wie sie in Bild 4-7 gezeigt sind. Bei Dämpfungsgraden D zwischen 0 und 1 gibt es zwei konjugiert komplexe Eigenwerte, im aperiodischen Grenzfall sind

die beiden Eigenwerte negativ und reell und fallen zusammen, bei Dämpfungsgraden größer als 1 sind die beiden Eigenwerte negativ und reell.

4.1.3 Erzwungene Schwingungen

Im Abschnitt 4.1 war schon die Differentialgleichung (Bild 4-2 und Gl. (4.1)) für die allgemeine Schwingung des Einmassenschwingers hergeleitet worden. Als Erregung des Systems war hier eine zeitlich veränderliche Kraft $F(t)$ angenommen worden, so dass sich eine inhomogene Bewegungs-DGL ergab.

Ist der zeitliche Verlauf der Erregungskraft bekannt, so spricht man von **deterministischer** Erregung. Sind hingegen nur gewisse statistische Merkmale der Erregung bekannt, so spricht man von **stochastischer** Erregung. Im folgenden sollen nur deterministische Fälle behandelt werden. Bei solchen deterministischen Anregungen unterscheidet man wiederum die drei in Bild 4-8 dargestellten Fälle: die **harmonische** Anregung mit Sinusschwingungen, die allgemeine **periodische** Anregung mit beispielsweise Rechteck- oder Dreieckschwingungen und die **transiente** Anregung mit nichtperiodischen, impulsartigen Schwingungen der Anregungskraft. Bei gleichem Systemverhalten (gleiche homogene DGL) ergeben sich dann, wie in Bild 4-8 dargestellt, unterschiedliche Bewegungsvorgänge $x(t)$.

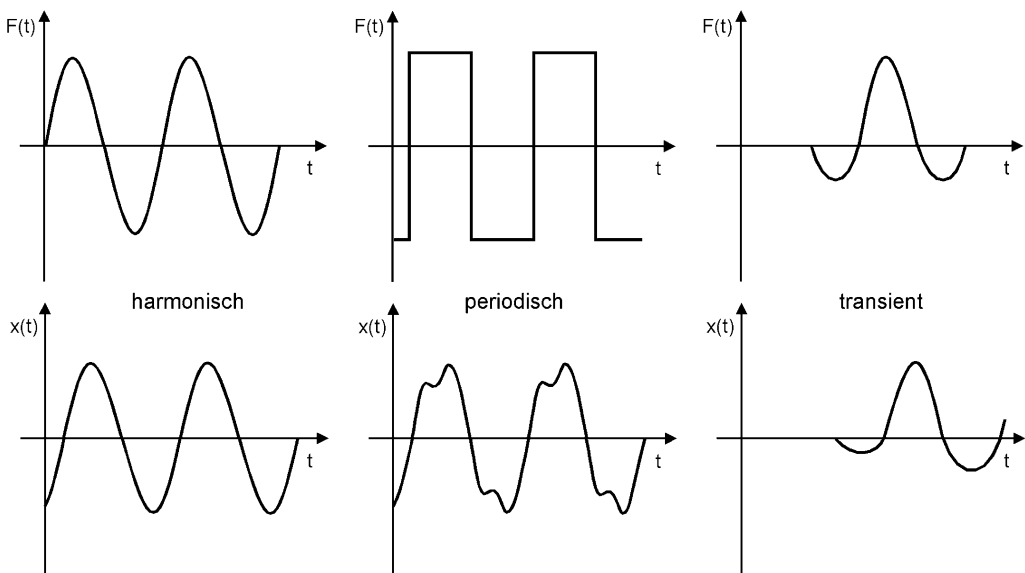


Bild 4-8: Unterschiedliche deterministische Anregungen von schwingungsfähigen Systemen

Wie bereits oben festgestellt, kann man die Lösung der inhomogenen DGL, die das Bewegungsverhalten des durch eine zeitlich veränderliche Kraft erregten Systems beschreibt, aus der Lösung der homogenen DGL und einer speziellen Lösung für die Anregungsfunktion finden. Diese spezielle Lösung heißt auch partikuläre Lösung. Es gilt:

$$x_a(t) = x_{ah}(t) + x_{ap}(t)$$

Zur Bestimmung der partikulären Lösung verwendet man Ansatzfunktionen, die eine möglichst große Ähnlichkeit mit den tatsächlichen Erregungsfunktionen haben und sich leicht beschreiben lassen. Dazu gehören Sinus- und Cosinusschwingungen und Überlagerungen von beiden. Aber auch viel einfachere Erregungsfunktionen, wie beispielsweise der Rechtecksprung, nähern wichtige Formen der Anregung schwingungsfähiger Systeme an. Ist die Anregungsfunktion sehr komplex oder der Schwinger nichtlinear, so ist meist ohnehin die numerische Simulation des Schwingungsverhaltens die einzig sinnvolle Art der Behandlung. Einfache Anregungen lassen sich aber auch gut algebraisch im **Zeitbereich** oder im **Frequenzbereich** behandeln. Dabei bedeutet Behandlung im Zeitbereich die Beschreibung des Bewegungsverhaltens in Abhängigkeit von der Zeit, Behandlung im Frequenzbereich die Darstellung des Bewegungsverhaltens in Abhängigkeit von der Frequenz der anregenden Schwingung. Im Kapitel 7.3 wird hierauf noch näher eingegangen.

4.1.3.1 Nichtperiodische Erregung

Als Beispiel für die Behandlung eines erregten Systems im Zeitbereich soll die Reaktion eines Einmassenschwingers auf einen Rechtecksprung behandelt werden. Diese nichtperiodische Anregungsfunktion ist in Bild 4-9 dargestellt. Die Funktion, die den Verlauf der Kraft $F(t)$ über die Zeit für einen Rechtecksprung beschreibt, kann folgendermaßen formuliert werden:

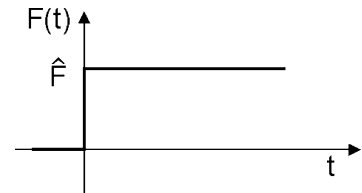


Bild 4-9: Sprungfunktion

$$F(t) = \begin{cases} 0 & \text{für } t < 0 \\ \hat{F} & \text{für } t > 0 \end{cases}$$

Da das Hooke'sche Gesetz

$$\hat{F} = k \cdot x_p$$

gilt, kann man als Lösungsansatz für die Partikularlösung folgendes wählen:

$$x_p = \frac{\hat{F}}{k} = \text{const}$$

Setzt man dies zusammen mit der bereits gefundenen Lösung der homogenen DGL in die Gesamtlösung ein, so ergibt sich:

$$x(t) = e^{-\omega_0 D t} (A \cos \omega t + B \sin \omega t) + \frac{\hat{F}}{k}, \quad \text{für } t \geq 0.$$

Die Integrationskonstanten A und B muss man dann wieder so wählen, dass die Anfangsbedingungen für den betrachteten Fall erfüllt sind. Man erkennt an der Gleichung, dass die Lösung der homogenen DGL die Eigenschwingungen beschreibt, die mit der Abklingkonstante $\omega_0 D$ über die Zeit verschwinden. Es verbleibt dann eine statische Auslenkung des Systems aufgrund der sprungförmigen Erregung.

In der jetzt gefundenen Form muss man jedoch für jede Art der Anfangsbedingungen die Konstanten A und B neu bestimmen. Um dies zu vermeiden, kann man die

Partikularlösung um einen Eigenschwungsanteil erweitern und in dieser Lösung die Konstanten für die Anfangsbedingungen $x_0 = 0$ und $\dot{x}_0 = 0$ (vor dem Auftreten des Sprungs sind Geschwindigkeit und Auslenkung gleich Null) ermitteln:

$$\tilde{x}_p(t) = e^{-\omega_0 D t} (A \cos \omega t + B \sin \omega t) + \frac{\hat{F}}{k} ,$$

$$\dot{\tilde{x}}_p(t) = -\omega_0 D e^{-\omega_0 D t} (A \cos \omega t + B \sin \omega t) + e^{-\omega_0 D t} (-A \omega \sin \omega t + B \omega \cos \omega t) .$$

Aus den Anfangsbedingungen gewinnt man

$$\tilde{x}_p(t=0) = 0 = A + \frac{\hat{F}}{k} ,$$

$$\dot{\tilde{x}}_p(t=0) = 0 = -\omega_0 D A + B \omega .$$

Hieraus kann man die Konstanten A und B berechnen:

$$A = -\frac{\hat{F}}{k} , \quad B = -\frac{\hat{F}}{k} \cdot \frac{D \omega_0}{\omega} .$$

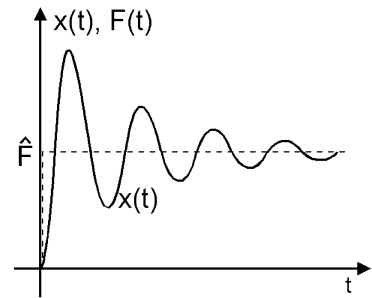


Bild 4-10: Durch Rechtecksprung erregte Schwingung

Setzt man diese Konstanten in die Gesamtlösung der DGL ein, so ergibt sich:

$$\begin{aligned} x(t) &= x_h(t) + x_p(t) \\ &= e^{-\omega_0 D t} \left(x_0 \cos \omega t + \left(\frac{\dot{x}_0 + D \omega_0 \cdot x_0}{\omega} \right) \sin \omega t \right) + \frac{\hat{F}}{k} \left(1 - e^{-\omega_0 D t} \left(\cos \omega t + \frac{D \omega_0}{\omega} \sin \omega t \right) \right) \end{aligned}$$

In dieser Lösung kommen jetzt nur noch die Systemgrößen, die Erregungsamplitude und die Anfangsbedingungen vor, so dass man den Bewegungsvorgang des Systems, wenn diese Größen bekannt sind, aufzeichnen kann (Bild 4-10).

4.1.3.2 Harmonische Erregung

Ein weiteres häufig behandeltes Beispiel ist die Reaktion eines Einmassenschwingers auf eine harmonische Anregung. Solche Anregungen treten oft in der Technik auf, weil sie beispielsweise durch Unwuchten rotierender Massen hervorgerufen werden. Es wird wieder die DGL

$$m\ddot{x} + d\dot{x} + kx = F(t) \quad \text{mit} \quad F(t) = \hat{F} \cos \Omega t$$

betrachtet. Der Einfachheit halber wird angenommen, dass die anregende Cosinusschwingung zum Zeitpunkt $t = 0$ gerade ihren Scheitelwert $\hat{F}$ aufweist. Zur Behandlung des Problems ist es günstig, die komplexe Schreibweise für die Darstellung der harmonischen Schwingung zu wählen.

Die allgemeine Sinusschwingung kann, wie in Bild 4-11 dargestellt,

$$x(t) = \hat{x} \cos(\omega t + \varphi_0) = \hat{x}_1 \sin \omega t + \hat{x}_2 \cos \omega t$$

als Zeitfunktion der horizontalen Projektion eines mit konstanter Winkelgeschwindigkeit ω rotierenden Radiusvektors vom Betrag $\hat{x}$ aufgefasst werden. Die Größe nennt man die Amplitude der Schwingung. Bei seiner Rotation erreicht der Radiusvektor jedesmal nach der Zeit T seine Ausgangslage. Die Winkelgeschwindigkeit dieser Drehung beträgt daher

$$\omega = \frac{2\pi}{T} = 2\pi \cdot f$$

und wird entsprechend als **Kreisfrequenz** bezeichnet. Die Winkelsumme

$$\varphi = \omega t + \varphi_0$$

wird **Phasenwinkel** genannt, der Phasenwinkel φ_0 zum Zeitpunkt $t = 0$ heißt **Phasenverschiebung**. Die Darstellung der harmonischen Schwingung kann zu einer Zeigerdarstellung in der komplexen Zahlenebene vereinfacht werden, in der nur noch der Radiusvektor in seiner Ausgangslage gezeichnet wird und die Kreisfrequenz angegeben wird (Bild 4-12).

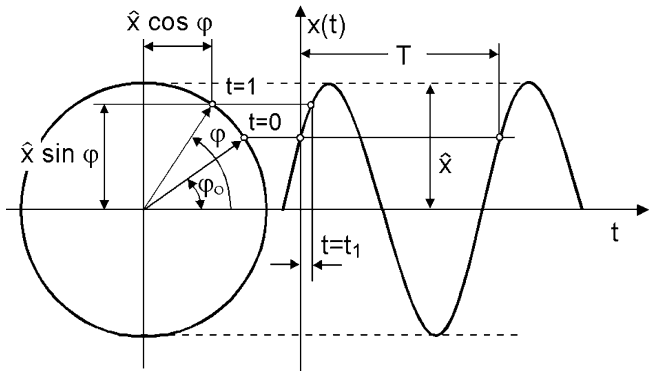


Bild 4-11: Erzeugung einer Sinusschwingung durch rotierenden Radiusvektor

In einer solchen Darstellung kann der Radiusvektor durch zwei Größen festgelegt werden, und zwar entweder durch den Betrag $\hat{x}$ und den Winkel φ oder durch Angabe der Koordinaten auf der reellen und der imaginären Achse des Koordinatensystems. In der komplexen Ebene ist der Radiusvektor der Ortsvektor einer komplexen Zahl z . Ihr Wert ist:

$$z = \hat{x}(\cos \varphi + i \sin \varphi) \quad .$$

Unter Verwendung der Euler'schen Relation

$$e^{i\varphi} = \cos \varphi + i \sin \varphi \quad ,$$

kann man eine komplexe Zahl als

$$z = \hat{x} \cdot e^{i\varphi}$$

schreiben. Ersetzt man in der letzten Gleichung den Phasenverschiebungswinkel φ durch den Phasenwinkel

$$\varphi = \omega t + \varphi_0$$

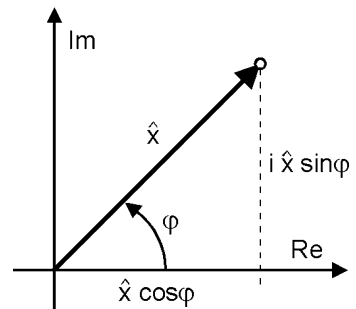


Bild 4-12: Darstellung einer komplexen Zahl

so entsteht daraus die Schreibweise des Radiusvektors der harmonischen Schwingung in Polarkoordinaten:

$$z = \hat{x} \cdot e^{i(\omega t + \varphi_0)} \quad , \quad (4.10a)$$

oder für die Phasenverschiebung $\varphi_0 = 0$

$$z = \hat{x} \cdot e^{i\omega t} \quad . \quad (4.10b)$$

Möchte man nun die Antwort des gedämpften Einmassenschwingers auf eine harmonische Anregung des Typs

$$F(t) = \hat{F} \cos \Omega t$$

berechnen, so kann man die komplexe Schreibweise

$$F(t) = \Re \left\{ \hat{F} \cdot e^{i\Omega t} \right\}$$

zur Vereinfachung der Darstellung verwenden. Um die Schreibweise noch einfacher zu gestalten, verzichtet man auch häufig auf die Kennzeichnung "Realteil $\Re$ " und schreibt kürzer aber ungenauer

$$F(t) = \hat{F} \cdot e^{i\Omega t} \quad .$$

Als partikuläre Lösung der DGL kann man dann gut die Funktion

$$x(t) = \Re \left\{ \hat{x} \cdot e^{i\Omega t} \right\} \quad \text{oder kürzer} \quad x(t) = \hat{x} \cdot e^{i\Omega t}$$

verwenden. In diesem Lösungsansatz kann die Variable $\hat{x}$ auch eine komplexe Größe sein. Durch Ableiten nach der Zeit erhält man die Geschwindigkeit und die Beschleunigung:

$$\dot{x}(t) = i\Omega \cdot \hat{x} \cdot e^{i\Omega t} \quad \text{und} \quad \ddot{x}(t) = -\Omega^2 \cdot \hat{x} \cdot e^{i\Omega t} \quad .$$

In die DGL eingesetzt ergibt sich dann:

$$(-\Omega^2 m + i\Omega d + k) \hat{x} \cdot e^{i\Omega t} = \hat{F} \cdot e^{i\Omega t}$$

Nach Division durch $e^{i\Omega t}$ ergibt sich

$$(-\Omega^2 m + i\Omega d + k) \hat{x} = \hat{F} \quad .$$

Hieraus erhält man die komplexe Amplitude zu

$$\hat{x} = \frac{\hat{F}}{(-\Omega^2 m + i\Omega d + k)} = \frac{\hat{F}}{k} \frac{1}{\left(-\frac{m}{k} \Omega^2 + i \frac{d}{k} \Omega + 1 \right)} \quad .$$

Unter Verwendung schon bekannter Größen, nämlich der **Eigenfrequenz** ω_0 des ungedämpften Systems, dem **Lehr'schen Dämpfungsmaß** D und dem **Frequenzverhältnis**

$$\eta = \frac{\Omega}{\omega_0} \quad , \quad \text{mit} \quad \omega_0 = \sqrt{\frac{k}{m}}$$

kann man die komplexe Amplitude auch folgendermaßen schreiben:

$$\hat{x} = \frac{\hat{F}}{k} \frac{1}{\left(1 - \eta^2 + i \frac{d}{k} \sqrt{\frac{k}{m}} \eta\right)} = x_{\text{stat}} \frac{1}{1 - \eta^2 + 2iD\eta} .$$

Hierin ist x_{stat} die statische Auslenkung durch die Kraft $\hat{F}$. Man erkennt, dass die Größe der komplexen Amplitude von der Erregerfrequenz Ω abhängt. Ihren Betrag errechnet man zu:

$$|\hat{x}| = x_{\text{stat}} \frac{1}{\sqrt{(1 - \eta^2)^2 + 4D^2\eta^2}} = x_{\text{stat}} \cdot V(\eta) .$$

Die Funktion $V(\eta)$ nennt man **Vergrößerungsfunktion**. Stellt man sie wie in Bild 4-13 in Abhängigkeit von η mit D als Parameter dar, so ergibt sich ein charakteristisches Bild. Diese Funktion nimmt an der Stelle $\eta = 1$ für eine Dämpfung $D = 0$ einen unendlichen Wert an. Das bedeutet, dass dann der Betrag der Amplitude unendlich groß werden würde. Dieser Wert der Vergrößerungsfunktion entspricht nämlich der **Resonanzstelle** des Systems, bei der die Erregerfrequenz Ω gleich der Eigenfrequenz ω_0 ist. Bei Erregung des Systems in der Resonanzfrequenz können demnach bei schwacher Dämpfung sehr große Amplitudenbeträge der Schwingung auftreten.

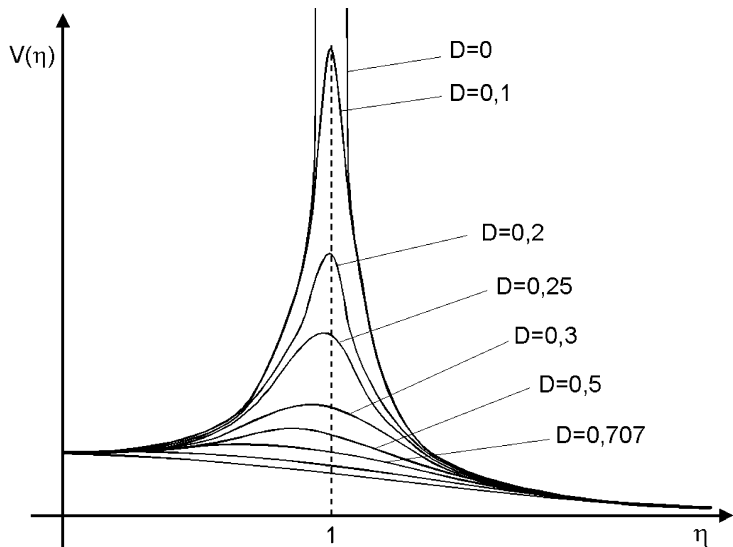


Bild 4-13: Vergrößerungsfunktion eines harmonisch erregten Einmassenschwingers bei unterschiedlichen Dämpfungsgraden

Das Verhältnis von $\hat{x}$ zu $\hat{F}$ kann man in Anlehnung an die Definition des Frequenzgangs in der Regelungstechnik (s. Kap. 7.3) als **Frequenzgang** des Einmassenschwingers definieren:

$$H(\Omega) = \frac{\hat{x}}{\hat{F}} = \frac{1}{k} \frac{1}{1 - \left(\frac{\Omega}{\omega_0}\right)^2 + 2iD\left(\frac{\Omega}{\omega_0}\right)} .$$

Als grafische Darstellung des Frequenzgangs ist in der Regelungstechnik die Dar-

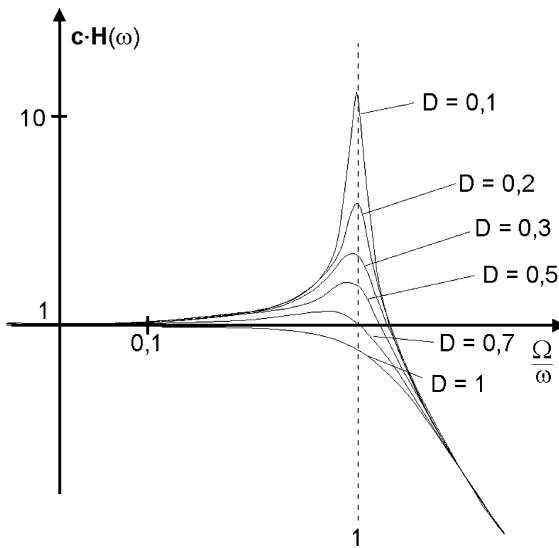


Bild 4-14: Amplitudenkennlinie des Frequenzganges eines Einmassenschwingers

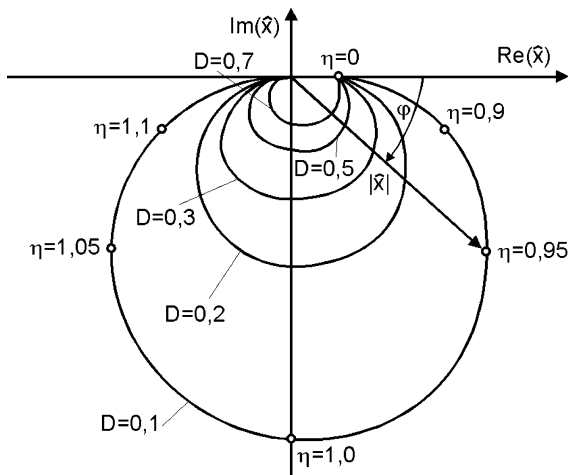


Bild 4-15: Ortskurven eines Einmassenschwingers

stellung in einem doppelt logarithmischen Koordinatensystem üblich. Diese Darstellung wird als **Frequenzkennlinie** (eigentlich zwei Kennlinien für die Amplitude und die Phase) oder auch als Bode-Diagramm bezeichnet. Für den Einmassenschwinger ist die **Amplitudenkennlinie** in Bild 4-14 dargestellt.

Auch an der Amplitudenkennlinie des Frequenzganges erkennt man, dass es an der Resonanzstelle abhängig von der Systemdämpfung zu einer starken Amplitudenüberhöhung gegenüber der statischen Auslenkung kommen kann. Stimmt bei einem technischen Bauteil die Erregerfrequenz Ω mit der Eigenfrequenz ω_0 , die im wesentlichen durch die Bauteilmasse und die Bauteilsteifigkeit bestimmt ist, überein, so liegt der Resonanzfall vor. Bei nicht ausreichender Dämpfung kann es dann zu starken Amplitudenüberhöhungen gegenüber den dynamischen Amplituden unterhalb der Resonanzfrequenz kommen, die zur Bauteilzerstörung führen können. Andererseits erkennt man auch, dass jenseits der Resonanzfrequenz die Amplituden unter die statische Auslenkung sinken und dass bei viel höheren Erregerfrequenzen das System der Erregerschwingung gar nicht mehr folgen kann.

Dies kann man in der Bauteilauslegung gezielt anwenden, indem Bauteilmasse und Bauteilsteifigkeit bei bekannter Anregungsfrequenz so wählt, dass die Bauteileigenfrequenz möglichst weit oberhalb oder unterhalb der Anregungsfrequenz liegt. Bei einer Bauteilabstimmung unterhalb der Anregungsfrequenz wird dann die dynamische Verformung in der Größenordnung der statischen Verformung, bei Abstimmung oberhalb der Anregungsfrequenz sogar noch deutlich darunter liegen.

Eine dritte Möglichkeit der grafischen Darstellung ist das Auftragen der komplexen Amplitude $\hat{x}$ in der komplexen Zahlenebene. Diese in der Regelungstechnik ebenfalls vielfach verwendete Darstellung wird als **Ortskurve** bezeichnet. Die in Bild 4-

15 gezeigte Darstellung erhält man, wenn man Real- und Imaginärteil für verschiedene Werte des Parameters η (proportional zur Erregerfrequenz Ω) als Ortskoordinaten für Punkte der komplexen Zahlenebene aufträgt. Die Verbindung dieser Punkte ist die Ortskurve. Variiert man noch den Parameter D , so erhält man eine Schar von Ortskurven. Diese Darstellung enthält mehr Informationen als die Amplitudenkennlinie, da wegen der Verwendung von Real- und Imaginärteil auch die Phaseninformation in der Ortskurve enthalten ist. Die Ortsvektoren der einzelnen Punkte der Ortskurve sind nämlich durch den Betrag der Amplitude (Länge des Ortsvektors) und durch die Phasenlage φ (Winkel zwischen reeller Achse und Ortsvektor) bestimmt.

Bei $\eta = 0$ beginnt die Ortskurve auf der reellen Achse (x_{stat}). Der Imaginärteil ist gleich Null und der Realteil entspricht der statischen Auslenkung ($\eta = 0 \Rightarrow \Omega = 0$). Beim Phasenwinkel 90° ist der Realteil Null; wegen $\eta = 1$ liegt der Resonanzfall vor und der Imaginärteil entspricht der maximalen dynamischen Amplitude. Für höhere Werte von η sinkt die Amplitude wieder und wird für $\eta = \infty$ zu Null.

Um nun die Gesamtlösung der DGL des harmonisch erregten Schwingers formulieren zu können, müssen wieder die Lösung der homogenen DGL und die partikuläre Lösung für die Anfangsbedingungen addiert werden. Die partikuläre Lösung

$$x_p(t) = \Re\{\hat{x} \cdot e^{i\Omega t}\}$$

mit der komplexen Variablen

$$\hat{x} = \Re\{\hat{x}\} + i\Im\{\hat{x}\}$$

kann man auch folgendermaßen schreiben:

$$x_p(t) = \Re\{\hat{x}\}\cos\Omega t - \Im\{\hat{x}\}\sin\Omega t \quad .$$

Für die Anfangsbedingungen der Anregeschwingung

$$x_p(t) = 0 \quad , \quad \dot{x}_p(t) = 0 \quad ,$$

lautet die partikuläre Lösung

$$\begin{aligned} x_p(t) = & -e^{-\omega_0 D t} \left(\Re\{\hat{x}\}\cos\omega t + \left(\frac{\Im\{\hat{x}\}\Omega + \Re\{\hat{x}\}D\omega_0}{\omega} \right) \sin\omega t \right) \\ & + \Re\{\hat{x}\}\cos\Omega t - \Im\{\hat{x}\}\sin\Omega t \quad . \end{aligned}$$

Die Gesamtlösung der DGL errechnet sich dann wie folgt:

$$\begin{aligned} x(t) = & x_h(t) + x_p(t) \\ = & e^{-\omega_0 D t} \left(x_0 \cos\omega t + \left(\frac{\dot{x}_0 + x_0 D \omega_0}{\omega} \right) \sin\omega t \right) \\ & - e^{-\omega_0 D t} \left(\Re\{\hat{x}\}\cos\omega t + \left(\frac{\Im\{\hat{x}\}\Omega + \Re\{\hat{x}\}D\omega_0}{\omega} \right) \sin\omega t \right) \\ & + \Re\{\hat{x}\}\cos\Omega t - \Im\{\hat{x}\}\sin\Omega t \quad . \end{aligned}$$

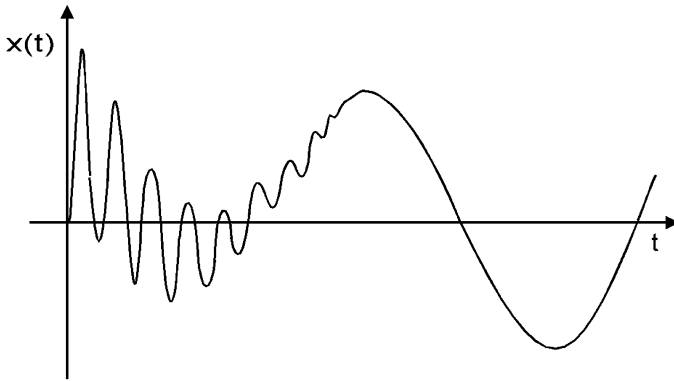


Bild 4-16: Einschwingvorgang einer harmonisch erregten Schwingung

Dieser Schwingungsvorgang wird für verschiedene Werte von η und unterschiedliche Anfangsbedingungen sehr unterschiedlich verlaufen. In Bild 4-16 ist ein solcher Schwingungsverlauf für den Fall $\omega > \Omega$ dargestellt. Man sieht, dass die anfänglichen Eigenvorgänge (ω) mit der Zeit abklingen und dass danach das System der Erreger-schwingung (Ω) folgt.

4.1.4 Nichtlineare Schwinger

Für nichtlineare Schwingungen gibt es keine so gut ausgebaute Theorie wie bei linearen Schwingungen. Das für lineare Differentialgleichungen geltende **Superpositionsprinzip** gilt hier nicht. Deshalb muss jedes Problem einzeln gelöst werden, und es ist sehr schwer, analytische Lösungen zu finden. Eine vollständige Integration der DGL zur Bestimmung des Bewegungsgesetzes ist meist nicht möglich, allenfalls kann man analytisch ein erstes Integral finden. Die Darstellung dieses ersten Integrals liefert aber häufig schon wichtige Aussagen über das Systemverhalten. In der Regel jedoch sind solche Probleme nur rechnergestützten numerischen Lösungsverfahren zugänglich.

Als Beispiel für ein System mit nichtlinearem Verhalten kann das mathematische Pendel, wie es bereits im Abschnitt 2.2 (Bild 2-13) behandelt wurde, herangezogen werden. Seine Bewegungsgleichung

$$\ddot{\varphi} = -\frac{g}{l} \sin \varphi \quad (4.11)$$

enthält als nichtlineare Funktion den Sinus der Bewegungsgröße. Im Abschnitt 2.2 war gezeigt worden, dass diese nichtlineare DGL für kleine Werte von φ mit ausreichender Genauigkeit linearisiert werden kann. Zur Bestimmung einer geschlossenen Lösung für beliebige Werte von φ müsste man die DGL integrieren. Um dies zu ermöglichen, muss man zuerst folgende Umformung vornehmen:

$$\ddot{\varphi} = \frac{d\dot{\varphi}}{dt} = \frac{d\dot{\varphi}}{dt} \cdot \frac{d\varphi}{d\varphi} = \frac{d\varphi}{dt} \cdot \frac{d\dot{\varphi}}{d\varphi} = \dot{\varphi} \frac{d\dot{\varphi}}{d\varphi}$$

Damit kann man Gl. (4.11) folgendermaßen schreiben:

$$\dot{\varphi} \frac{d\dot{\varphi}}{d\varphi} = -\frac{g}{l} \sin \varphi \quad .$$

Dies führt zur folgenden Variablentrennung:

$$\dot{\varphi} \cdot d\dot{\varphi} = -\frac{g}{l} \sin \varphi \cdot d\varphi \quad .$$

Integriert man diese Gleichung, so erhält man:

$$\frac{\dot{\varphi}^2}{2} - \frac{\dot{\varphi}_0^2}{2} = -\frac{g}{l} (-\cos \varphi + \cos \varphi_0) \quad .$$

Die Funktion $\dot{\varphi}(\varphi)$ heißt **Phasenfunktion**, ihr Graph wird als **Phasenkurve** bezeichnet. Für das mathematische Pendel ist

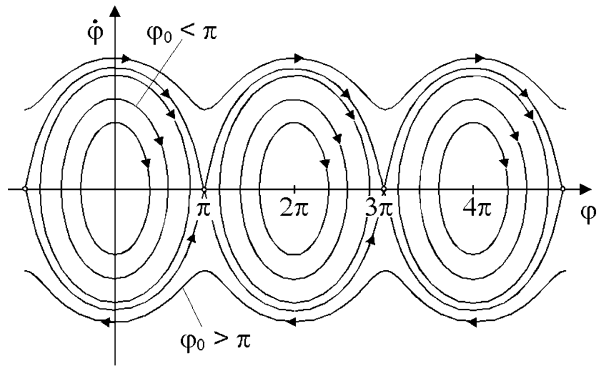


Bild 4-17: Phasendiagramm des mathematischen Pendels

dies die Darstellung der Winkelgeschwindigkeit in Abhängigkeit vom Winkel der Auslenkung aus der Ruhelage. Wie in Bild 4-17 dargestellt, ergeben sich für verschiedene Anfangsauslenkungen φ_0 verschiedene Verläufe der Phasenkurve.

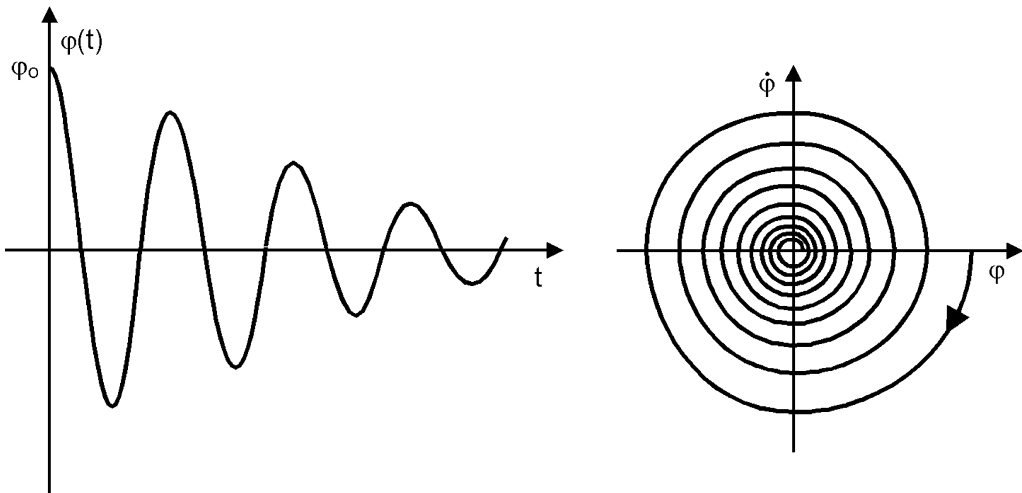


Bild 4-18: Zeitlicher Verlauf und Phasenkurve einer gedämpften harmonischen Schwingung

Für Anfangsauslenkungen $\varphi_0 < \pi$ sind die Phasenkurven geschlossene Linien. An einer solchen Kurve erkennt man beispielsweise, dass der Schwingungsvorgang ungedämpft abläuft, da sie immer wieder in ihren Anfangspunkt zurückkehrt. Ist eine Schwingung gedämpft, so ergibt sich eine Phasenkurve wie in Bild 4-18, beginnend bei einer Anfangsauslenkung $\varphi_0 < \pi$. Die Phasenkurve nähert sich dann in Form einer Spirale dem Nullpunkt des Phasendiagramms ($\varphi = 0, \dot{\varphi} = 0$).

Nähert sich die Anfangsauslenkung φ_0 dem Wert π (Pendelmasse erreicht den oberen Totpunkt), so ergibt sich in der Phasenkurve ein instabiler Wirbelpunkt (Bild 4-17). Dies gilt ebenso für Werte $\varphi = (2n + 1)\pi$, $\dot{\varphi} = 0$, wobei n ganze Zahlen sind. Im oberen Totpunkt ist die Lage instabil und der weitere Verlauf von φ hängt von

zufälligen Schwankungen ab. Bei Überschreiten der oberen Totpunktlage schwingt das Pendel über und φ wächst kontinuierlich. Im Phasendiagramm (Bild 4-17) sind die Phasenkurven dann keine in sich geschlossenen Linien mehr, sondern in Richtung von φ fortschreitende oszillierende Kurven.

Mit Hilfe der Phasenkurven lassen sich demnach für nichtlineare Systeme bereits einige Aussagen über das Systemverhalten treffen.

4.2 Mehrmassenschwinger

Der Einmassenschwinger stellt nur ein sehr einfaches Modell für technische Systeme dar. Viele mechanische Systeme kann man aus diskreten, einfachen Elementen wie Federn, Massen, Dämpfern und starren Hebeln zusammensetzen, um ähnlich wie beim Einmassenschwinger die Bewegungsgleichungen zu ermitteln. Bei solchen Systemen mit mehreren diskreten Massen handelt es sich um Mehr-Freiheitsgrad-Systeme, für deren Beschreibung mehr als eine Koordinate erforderlich ist. Das Aufstellen der Bewegungsgleichungen für solche Systeme soll hier anhand einfacher Beispiele aufgezeigt werden. Die Lösung des dabei entstehenden Gleichungssystems der Bewegungsgleichungen wird für ein einfaches mechanisches System durchgeführt, für kompliziertere Systeme wird man dies aber sinnvollerweise wieder einem Digitalrechner überlassen.

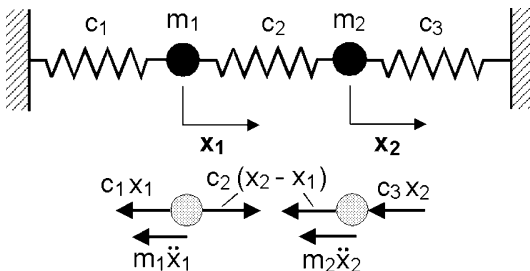


Bild 4-19: Mehrmassenschwinger und freigeschnittene Massen

In Bild 4-19 ist ein System aus zwei Punktmassen dargestellt. Nach Freischneiden der Massen entsprechend Bild 4-19 kann man nach dem d'Alembert'schen Prinzip die Bewegungsgleichungen aufstellen. Im freigeschnittenen Zustand ist der Fall angenommen, dass die Koordinate x_2 größer als x_1 ist, so dass die mittlere Feder verlängert wird. Nimmt man die Gültigkeit des Hooke'schen Gesetzes an, so ergeben sich folgende

lineare Differentialgleichungen für die Massen m_1 und m_2 :

$$\begin{aligned} m_1 \ddot{x}_1 + k_1 x_1 - k_2 (x_2 - x_1) &= 0 \\ m_2 \ddot{x}_2 + k_3 x_2 - k_2 (x_2 - x_1) &= 0 \end{aligned}$$

Als Lösungsansatz für dieses lineare, homogene Differentialgleichungssystem wählt man günstigerweise

$$x_1 = A_1 \cos(\omega t + \varphi) \quad , \quad x_2 = A_2 \cos(\omega t + \varphi) \quad ,$$

weil sich dann die Cosinus - Anteile herausheben, wenn man die Lösungen in die DGL einsetzt. Es ergibt sich:

$$\begin{aligned} (k_1 + k_2 - m_1 \omega^2) A_1 &- k_2 A_2 = 0 \\ -k_2 A_1 + (k_2 + k_3 - m_2 \omega^2) A_2 &= 0 \end{aligned}$$

Dies ist ein lineares homogenes Gleichungssystem für die unbekannten Ansatzparameter A_1 und A_2 . Die beiden Ansatzparameter können nicht unabhängig voneinander sein, da folgendes aus der ersten der beiden letzten Gleichungen abgelesen werden kann:

$$A_2 = \frac{1}{k_2} (k_1 + k_2 - m_1 \omega^2) A_1 = \mu \cdot A_1 \quad .$$

Die Trivillösung $A_1 = A_2 = 0$ beschreibt den Ruhezustand des Systems. Nichttriviale Lösungen können nur existieren, wenn die Determinante der Koeffizienten der Lösungsparameter verschwindet:

$$\begin{vmatrix} k_1 + k_2 - m_1 \omega^2 & -k_2 \\ -k_2 & k_2 + k_3 - m_2 \omega^2 \end{vmatrix} = 0 \quad .$$

Hieraus kann man eine quadratische Gleichung für ω gewinnen:

$$m_1 m_2 \omega^4 - [(k_1 + k_2)m_2 + (k_2 + k_3)m_1] \omega^2 + (k_1 + k_2)(k_2 + k_3) - k_2^2 = 0$$

Diese Gleichung hat folgende zwei Lösungen:

$$\omega_{1,2}^2 = \frac{(k_1 + k_2)m_2 + (k_2 + k_3)m_1}{2m_1 m_2} \pm \sqrt{\left[\frac{(k_1 + k_2)m_2 + (k_2 + k_3)m_1}{2m_1 m_2} \right]^2 + \frac{k_2^2 - (k_1 + k_2)(k_2 + k_3)}{m_1 m_2}}$$

Da es zwei Lösungen ω_1 und ω_2 der quadratischen Gleichung gibt, kann man nun die Lösungsansätze als Summe von zwei linear unabhängigen Partikulärlösungen formulieren. Dies ist möglich, da es sich bei dem zu lösenden Gleichungssystem um lineare Differentialgleichungen handelt:

$$x_1 = A_{11} \cos(\omega_1 t + \varphi_1) + A_{12} \cos(\omega_2 t + \varphi_2)$$

$$x_2 = A_{21} \cos(\omega_1 t + \varphi_1) + A_{22} \cos(\omega_2 t + \varphi_2)$$

Benutzt man die oben festgestellte Tatsache, dass die Koeffizienten A_{ij} nicht unabhängig voneinander sind, so erhält man schließlich die allgemeine Lösung

$$x_1 = A_{11} \cos(\omega_1 t + \varphi_1) + A_{12} \cos(\omega_2 t + \varphi_2)$$

$$x_2 = \mu_1 A_{11} \cos(\omega_1 t + \varphi_1) + \mu_2 A_{12} \cos(\omega_2 t + \varphi_2)$$

Die vier Integrationskonstanten A_{11} , φ_1 , A_{12} , φ_2 können aus vier Anfangsbedingungen wie Lage und Geschwindigkeit der Massen zu Anfang der Bewegung bestimmt werden. Die beiden Koeffizienten μ_1 und μ_2 werden als **Modalkoeffizienten** bezeichnet.

Um den allgemeinen Lösungsweg für Schwingungsprobleme von komplizierten Mehrmassenschwingern aufzuzeigen, wird das Modell aus Bild 4-19 nochmals vereinfacht und zwar soll gelten:

$$m_1 = m_2 = m \quad , \quad k_1 = k_2 = k \quad .$$

In Matrizenschreibweise stellt sich dann das Differentialgleichungssystem des Modells wie folgt dar:

$$\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} + \begin{pmatrix} 2k & -k \\ -k & 2k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{oder kürzer} \quad \mathbf{M} \cdot \ddot{\vec{x}} + \mathbf{K} \cdot \vec{x} = \mathbf{0}$$

Der Lösungsansatz lautet in gleicher Schreibweise

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \cos \omega t + \varphi \quad \text{oder kurz} \quad \vec{x} = \vec{A} \cos \omega t + \varphi$$

Setzt man diesen Lösungsansatz in das Gleichungssystem ein, so erhält man:

$$\left[\begin{pmatrix} 2k & -k \\ -k & 2k \end{pmatrix} - \omega^2 \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \right] \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{oder kürzer} \quad (\mathbf{K} - \omega^2 \mathbf{M}) \vec{A} = \mathbf{0}$$

Dies ist ein **allgemeines Matrizen Eigenwertproblem** mit den beiden symmetrischen Matrizen $\mathbf{K}$ (Steifigkeitsmatrix) und $\mathbf{M}$ (Massenmatrix). Für dessen Lösung, auch bei viel komplizierteren Fällen, stehen in mathematischen Programmbibliotheken entsprechende Programme zur Verfügung (siehe dazu auch [3.2]). In dem oben aufgeführten einfachen Fall erhält man folgende Eigenkreisfrequenzen und Modalkoeffizienten:

$$\omega_1 = \sqrt{\frac{k}{m}}, \quad \omega_2 = \sqrt{3 \frac{k}{m}}, \quad \mu_1 = 1, \quad \mu_2 = -1$$

Die wesentliche Arbeit besteht daher im Aufstellen des Differentialgleichungssystems für das gegebene Schwingungssystem. Dies soll im folgenden für drei verschiedenartige Fälle an Beispielen aufgezeigt werden.

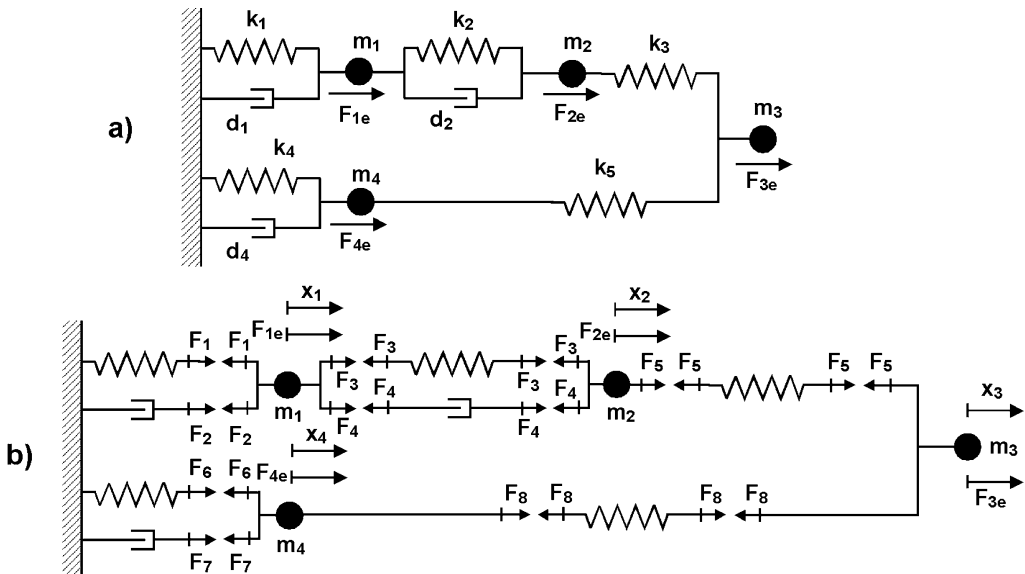


Bild 4-20: Diskretes Mehrmassen-Schwingersystem a) Systemaufbau b) Freischneiden der Massen

In Bild 4-20 Teilbild a) ist ein diskretes Schwingungssystem aus vier Punktmassen und einigen Feder- und Dämpferelementen dargestellt. An jeder Masse greift eine äußere Erregerkraft an.

Die Punktmassen sollen sich nur geradlinig in der Bildebene bewegen können, so dass sie jeweils nur einen Freiheitsgrad besitzen. Es müssen daher zur Aufstellung der Bewegungsgleichungen vier Wegkoordinaten für die vier Punktmassen eingeführt werden. Zur Anwendung des d'Alembert'schen Prinzips müssen alle Massen freigeschnitten werden (Bild 4-20 b)) und an den Schnittstellen ersatzweise Schnittkräfte $F_1 \dots F_8$ eingeführt werden. Danach können die Kräftebilanzen für die vier Punktmassen gebildet werden:

$$m_1 \ddot{x}_1 = -F_1 - F_2 + F_3 + F_4 + F_{1e}$$

$$m_2 \ddot{x}_2 = -F_3 - F_4 + F_5 + F_{2e}$$

$$m_3 \ddot{x}_3 = -F_5 - F_8 + F_{3e}$$

$$m_4 \ddot{x}_4 = -F_6 - F_7 + F_8 + F_{4e}$$

Die Schnittkräfte kann man unter der Annahme, dass das Hooke'sche Gesetz gilt

$$F_i = k_i \cdot x_i \quad (k_i: \text{Federsteifigkeit})$$

und für viskose Dämpfer mit geschwindigkeitsproportionaler Dämpfung

$$F_i = d_i \cdot \Delta \dot{x}_i \quad (d_i: \text{Dämpfungskonstante, } \Delta \dot{x}_i: \text{Relativgeschwindigkeit})$$

wie folgt bestimmen:

$$F_1 = k_1 \cdot x_1 \quad F_2 = d_1 \cdot \dot{x}_1$$

$$F_3 = k_2 \cdot (x_2 - x_1) \quad F_4 = d_2 \cdot (\dot{x}_2 - \dot{x}_1)$$

$$F_5 = k_3 \cdot (x_3 - x_2) \quad F_6 = k_4 \cdot x_4$$

$$F_7 = d_4 \cdot \dot{x}_4 \quad F_8 = k_5 \cdot (x_3 - x_4)$$

Setzt man die Schnittkräfte in die obigen Bewegungsgleichungen ein, so erhält man

$$m_1 \ddot{x}_1 + (d_1 + d_2) \dot{x}_1 - d_2 \dot{x}_2 + (k_1 + k_2) x_1 - k_2 x_2 = F_{1e}$$

$$m_2 \ddot{x}_2 - d_2 \dot{x}_1 + d_2 \dot{x}_2 - k_2 x_1 + (k_2 + k_3) x_2 - k_3 x_3 = F_{2e}$$

$$m_3 \ddot{x}_3 - k_3 x_2 + (k_3 + k_5) x_3 - k_5 x_4 = F_{3e}$$

$$m_4 \ddot{x}_4 + d_4 \dot{x}_4 - k_5 x_3 + (k_4 + k_5) x_4 = F_{4e}$$

Dieses System von Differentialgleichungen lässt sich auch in Matrizenform schreiben:

$$\begin{pmatrix} m_1 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 \\ 0 & 0 & m_3 & 0 \\ 0 & 0 & 0 & m_4 \end{pmatrix} \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \\ \ddot{x}_4 \end{pmatrix} + \begin{pmatrix} (d_1 + d_2) & -d_2 & 0 & 0 \\ -d_2 & d_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_4 \end{pmatrix} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} + \begin{pmatrix} (k_1 + k_2) & -k_2 & 0 & 0 \\ -k_2 & (k_1 + k_2) & -k_3 & 0 \\ 0 & -k_3 & (k_1 + k_2) & -k_5 \\ 0 & 0 & -k_5 & (k_1 + k_2) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} F_{1e} \\ F_{2e} \\ F_{3e} \\ F_{4e} \end{pmatrix}$$

Kürzt man die Massenmatrix mit **M**, die Dämpfungsmatrix mit **D**, die Steifigkeits-

matrix mit $\mathbf{K}$, sowie den Auslenkungsvektor mit $\vec{x}$ und den Kraftvektor mit $\vec{f}$ ab, so kann man das Gleichungssystem auch folgendermaßen schreiben:

$$\mathbf{M} \cdot \ddot{\vec{x}} + \mathbf{D} \cdot \dot{\vec{x}} + \mathbf{K} \cdot \vec{x} = \vec{f} \quad .$$

Die weitere Lösung dieses Gleichungssystems erfolgt als Lösung des Matrizenwertproblems wie oben ausgeführt auf einem Digitalrechner. Will man die Reaktion des Systems auf bestimmte Werte des Kraftvektors untersuchen, so ist es hilfreich, das System mit Hilfe eines **Simulationssystems** zu untersuchen, dessen Aufbau und Arbeitsweise im Kapitel 8 ausführlich behandelt wird.

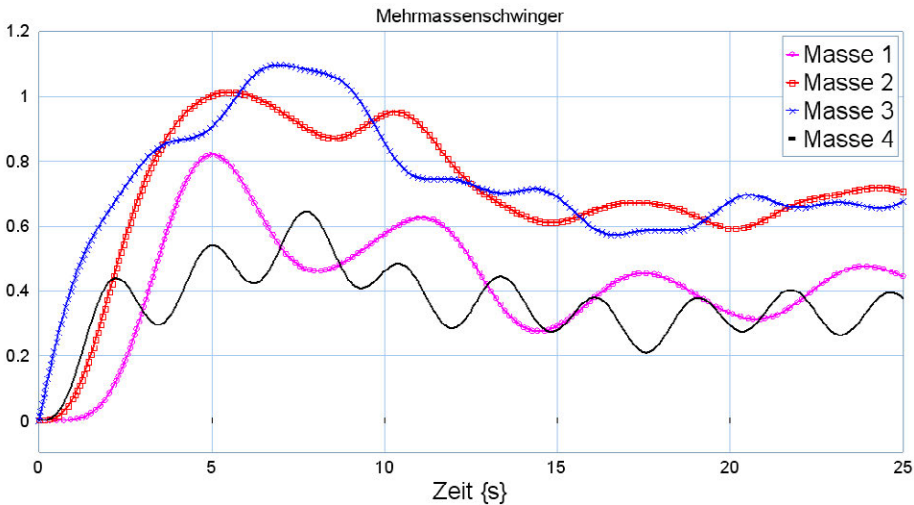


Bild 4-21: Geschwindigkeitsverläufe der Massen des Mehrmassenschwingers bei sprunghafter Erregung

Bild 4-21 zeigt ein Simulationsergebnis, das mit dem Simulationssystem „20sim“ erzielt wurde. Es zeigt die Geschwindigkeit der vier Massen, nachdem sich die Erregerkraft F_{3e} zum Zeitpunkt $t = 0$ sprunghaft geändert hat. Dabei wurden folgende Parameter verwendet: $m_1 = m_2 = m_4 = 0,5$, $m_3 = 1$ kg, $k_2 = k_3 = 0,5$, $k_1 = 0,5$, $k_4 = 1$ N/m, $d_1 = 2d_4 = 2$ Ns/m, $d_2 = 3d_4$. Man sieht, dass das Berechnen dieser Ergebnisse aus dem Gleichungssystem der Bewegungsgleichungen sehr aufwendig wäre, während man das Simulationsergebnis ohne großen Aufwand in kurzer Zeit erhält.

Das vorhergehende Beispiel behandelte ein System mit reinen Linearbewegungen. Im folgenden Beispiel (Bild 4-22) wird ein System aus Wellen und Zahnrädern behandelt, das ausschließlich rotatorische Bewegungen ausführen kann. Für die Verformung der Wellen unter den Torsionsmomenten gilt wieder das Hooke'sche Gesetz in der Form

$$M_i = k_i \cdot \varphi \quad \text{mit} \quad k_i = \left(\frac{G \cdot J_p}{l} \right)_i \quad .$$

Die Federsteifigkeit der Wellen hängt dabei vom Gleitmodul G , vom polaren Träg-

heitsmoment J_p und von der Wellenlänge l ab. Das Wendegetriebe hat die Übersetzung $\ddot{u}$, die beiden großen Zahnräder haben den Durchmesser D , das kleine Zahnrad den Durchmesser d .

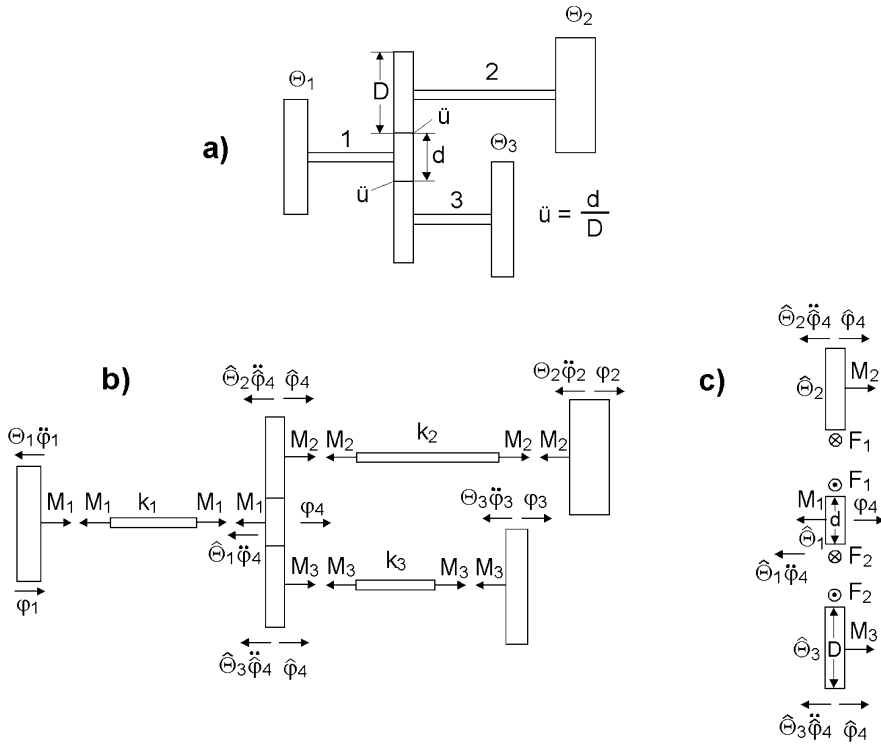


Bild 4-22: Schwingungsfähiges rotatorisches System a) Wellenstrang aus Einzelmassen und einem Übersetzungsgetriebe b) Freischnitten der Massen c) Kräfte und Momente am Zahnradgetriebe

Zur Beschreibung der Bewegungen sind vier Winkelkoordinaten erforderlich. Als Bewegungskordinaten der drei Räder des Übersetzungsgetriebes ist nur eine Koordinatenangabe erforderlich, da die der anderen Räder folgendermaßen bestimmt werden kann:

$$\varphi_4 \frac{d}{2} = -\hat{\varphi}_4 \frac{D}{2} \quad , \quad \hat{\varphi}_4 = -\frac{d}{D} \varphi_4 = -\ddot{u} \varphi_4 \quad .$$

Entsprechend dem Schnitt in Teilbild 4-22 c) kann man nach dem d'Alembert'schen Prinzip für die Zahnräder des Wendegetriebes folgende Bewegungsgleichungen aufstellen.

$$\hat{\Theta}_1 \hat{\varphi} = -M_1 + F_1 \frac{d}{2} + F_2 \frac{d}{2} \quad (4.13a)$$

$$\hat{\Theta}_2 \hat{\varphi}_4 = M_2 + F_1 \frac{D}{2} \quad (4.13b)$$

$$\hat{\Theta}_3 \hat{\ddot{\varphi}}_4 = M_3 + F_2 \frac{D}{2} \quad (4.13c)$$

Aus den Gl. (4.13 b) und (4.13 c) kann man die beiden Kräfte an den Zahneingriffen berechnen:

$$F_1 = -\hat{\Theta}_2 \frac{2d}{D^2} \ddot{\varphi}_4 - \frac{2M_2}{D}$$

$$F_2 = -\hat{\Theta}_3 \frac{2d}{D^2} \ddot{\varphi}_4 - \frac{2M_3}{D}$$

In die Gl. (4.13 a) eingesetzt ergibt dies:

$$\hat{\Theta}_1 \ddot{\varphi}_4 = -M_1 - \hat{\Theta}_2 \frac{d^2}{D^2} \ddot{\varphi}_4 - \frac{M_2 d}{D} - \hat{\Theta}_3 \frac{d^2}{D^2} \ddot{\varphi}_4 - \frac{M_3 d}{D} \quad (4.14)$$

Setzt man nun das Übersetzungsverhältnis in die Gl. (4.14) ein, so erhält man die Beziehung:

$$(\hat{\Theta}_1 + \ddot{u}^2 \cdot \hat{\Theta}_2 + \ddot{u}^2 \cdot \hat{\Theta}_3 \frac{d^2}{D^2}) \ddot{\varphi}_4 = -M_1 - \ddot{u} \cdot M_2 - \ddot{u} \cdot M_3 \quad .$$

Für die übrigen rotierenden Massen folgt aus dem d'Alembert'schen Prinzip:

$$\Theta_1 \ddot{\varphi}_1 = M_1 \quad , \quad \Theta_2 \ddot{\varphi}_2 = -M_2 \quad , \quad \Theta_3 \ddot{\varphi}_3 = -M_3 \quad .$$

Die Drehmomente haben nach dem Hooke'schen Gesetz folgende Werte:

$$M_1 = k_1(\varphi_4 - \varphi_1)$$

$$M_2 = k_2(\varphi_2 + \ddot{u} \varphi_4)$$

$$M_3 = k_3(\varphi_3 + \ddot{u} \varphi_4)$$

Mit folgenden Hilfsgrößen

$$\Theta_4 = -\hat{\Theta}_1 + \ddot{u}^2 \cdot \hat{\Theta}_2 + \ddot{u}^2 \cdot \hat{\Theta}_3$$

$$k_4 = k_1 + \ddot{u}^2 \cdot k_2 + \ddot{u}^2 \cdot k_3$$

kann man dann endgültig das System der Bewegungsgleichungen aufstellen:

$$\begin{aligned} \Theta_1 \ddot{\varphi}_1 + k_1 \varphi_1 - k_1 \varphi_4 &= 0 \\ \Theta_2 \ddot{\varphi}_2 + k_2 \varphi_2 + \ddot{u} k_2 \varphi_4 &= 0 \\ \Theta_3 \ddot{\varphi}_3 + k_3 \varphi_3 + \ddot{u} k_3 \varphi_4 &= 0 \\ \Theta_4 \ddot{\varphi}_4 - k_1 \varphi_1 + \ddot{u} k_2 \varphi_2 + \ddot{u} k_3 \varphi_3 + k_4 \varphi_4 &= 0 \quad . \end{aligned}$$

In Matrizenschreibweise lautet dieses Gleichungssystem:

$$\begin{pmatrix} \Theta_1 & 0 & 0 & 0 \\ 0 & \Theta_2 & 0 & 0 \\ 0 & 0 & \Theta_3 & 0 \\ 0 & 0 & 0 & \Theta_4 \end{pmatrix} \cdot \begin{pmatrix} \ddot{\varphi}_1 \\ \ddot{\varphi}_2 \\ \ddot{\varphi}_3 \\ \ddot{\varphi}_4 \end{pmatrix} + \begin{pmatrix} k_1 & 0 & 0 & -k_1 \\ 0 & k_2 & 0 & \ddot{u}k_2 \\ 0 & 0 & k_3 & \ddot{u}k_3 \\ -k_1 & \ddot{u}k_2 & \ddot{u}k_3 & k_4 \end{pmatrix} \cdot \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

oder in vereinfachter Form

$$\mathbf{M} \cdot \ddot{\vec{\varphi}} + \mathbf{K} \cdot \vec{\varphi} = \mathbf{0} \quad .$$

Im Kapitel 2 waren bereits verschiedene Modelle für das Fahrwerk eines PKW behandelt worden. Für das letzte Beispiel (Bild 4-23) wird dasjenige ebene Modell behandelt, das über vier Freiheitsgrade verfügt, nämlich **Tauchen** (Koordinate x_S) und

Nicken (Koordinate φ) der Karosserie, sowie zwei Radfreiheitsgrade (Koordinaten x_1 und x_2). Es treten also translatorische und rotatorische Bewegungsanteile auf. Nach dem Freischneiden der Massen kann man entsprechend dem d'Alembert'schen Prinzip die Bewegungsgleichungen formulieren:

$$M \cdot \ddot{x}_S = F_1 + F_2$$

$$m_1 \cdot \ddot{x}_1 = -F_1 + F_3$$

$$m_2 \cdot \ddot{x}_2 = -F_2 + F_4$$

$$\Theta_S \cdot \ddot{\varphi} = F_1 \cdot a - F_2 \cdot b$$

Die Kräfte kann man wieder, wie im ersten Beispiel, bei Gültigkeit des Hooke'schen Gesetzes und unter der Annahme viskoser Dämpfung wie folgt bestimmen:

$$F_1 = -d_1 \dot{x}_S - d_1 a \dot{\varphi} + d_1 \dot{x}_1 - k_1 x_S - k_1 a \varphi + k_1 x_1$$

$$F_2 = -d_2 \dot{x}_S + d_2 b \dot{\varphi} + d_2 \dot{x}_2 - k_2 x_S + k_2 b \varphi + k_2 x_2$$

$$F_3 = -k_3 (x_1 - u_1)$$

$$F_4 = -k_4 (x_2 - u_2)$$

In Matrizenschreibweise erhält man folgendes Gleichungssystem:

$$\begin{pmatrix} M & 0 & 0 & 0 \\ 0 & \Theta_S & 0 & 0 \\ 0 & 0 & m_1 & 0 \\ 0 & 0 & 0 & m_2 \end{pmatrix} \cdot \begin{pmatrix} \ddot{x}_S \\ \ddot{\varphi} \\ \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} + \begin{pmatrix} (d_1 + d_2) & (d_1 a - d_2 b) & -d_1 & -d_2 \\ (d_1 a - d_2 b) & (d_1 a^2 + d_2 b^2) & -d_1 a & d_2 b \\ -d_1 & -d_1 a & d_1 & 0 \\ -d_2 & d_2 b & 0 & d_2 \end{pmatrix} \cdot \begin{pmatrix} \dot{x}_S \\ \dot{\varphi} \\ \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} + \begin{pmatrix} (k_1 + k_2) & (k_1 a - k_2 b) & -k_1 & -k_2 \\ (k_1 a - k_2 b) & (k_1 a^2 + k_2 b^2) & -k_1 a & k_2 b \\ -k_1 & -k_1 a & (k_1 + k_3) & 0 \\ -k_2 & k_2 b & 0 & (k_2 + k_4) \end{pmatrix} \cdot \begin{pmatrix} x_S \\ \varphi \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ k_3 u_1 \\ k_4 u_2 \end{pmatrix}$$

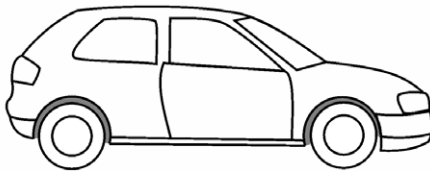


Bild 4-23: Mechanisches Modell eines PKW mit Anregung durch Fahrbahnunebenheiten

oder kurz

$$\mathbf{M} \cdot \ddot{\vec{x}} + \mathbf{D} \cdot \dot{\vec{x}} + \mathbf{K} \cdot \vec{x} = \vec{f} \quad .$$

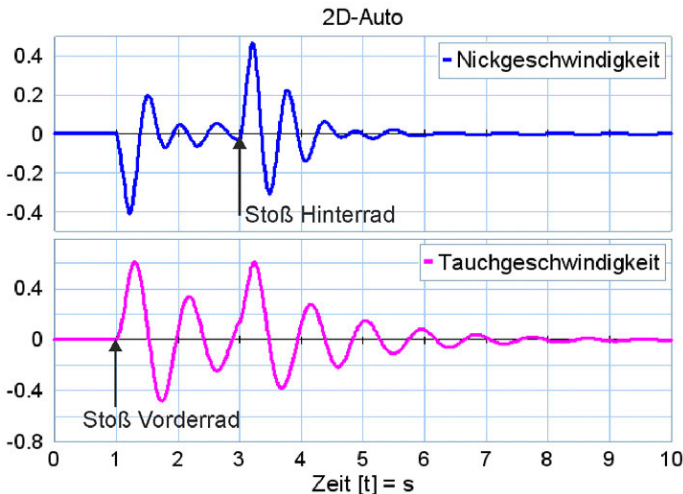


Bild 4-24: Nick- und Tauchgeschwindigkeit des Fahrzeugschwerpunktes beim Fahren über eine Bodenwelle

Wie das Schwingungsverhalten eines solchen Modells bei einer konkreten Anregung aussieht, lässt sich mit vertretbarem Aufwand nur auf einem Simulationssystem, wie es in Kapitel 8 besprochen wird, untersuchen. Da es sich um vier lineare DGL's 2. Ordnung handelt, bereitet die Simulation auch keine Probleme und kann für die verschiedenartigsten Anregungstypen durchgeführt werden. Bild 4-24 zeigt das Simulationsergebnis für das 2D-Simulationsmodell eines

PKW beim Fahren über eine Bodenwelle [4.1]. Diese erzeugt eine zeitlich versetzte Stoßanregung der beiden Räder, die zu entsprechendem Tauchen und Nicken führen.

4.3 Schwingungsanalyse

Schwingungsvorgänge in mechanischen Systemen sind dadurch gekennzeichnet, dass die Ausschläge bestimmter Bauteile gedämpft oder ungedämpft, aber im wesentlichen periodisch verlaufen. Liegen keine oder nur ungenügende Informationen über den Modellcharakter des Systems vor, so kann eine Analyse der gemessenen Schwingungsformen wichtige Informationen über das System liefern. Dies wird noch klarer, wenn man die Struktur durch genau definierte Testsignale zu Schwingungen anregt und dann die Messwerte des Schwingungsverlaufs analysiert. Man spricht dann von **Systemidentifikation**.

Die einfachste periodische Schwingung ist die harmonische Schwingung, wie sie in Teilbild 4-25 a) im sog. **Zeitbereich** dargestellt ist [3.2], [4.2]. Die Funktion, die die harmonische Schwingung im Zeitbereich beschreibt, lautet in allgemeiner Form

$$f(t) = \hat{a} \sin(\omega_0 t + \varphi_0) \quad . \quad (4.15)$$

Sie ist zu einem beliebigen Zeitpunkt t durch die Kenngrößen Amplitude $\hat{a}$ Kreisfrequenz ω_0 und die Phasenverschiebung φ_0 eindeutig festgelegt. Die Kreisfrequenz kann man auch durch die Grundfrequenz f_0 oder die Periodendauer T_0 ausdrücken:

$$\omega_0 = 2\pi \cdot f_0 = \frac{2\pi}{T_0}$$

Neben der Darstellung im Zeitbereich durch Angabe des Verlaufs der Zeitfunktion kann die harmonische Schwingung auch im sog. **Frequenzbereich** dargestellt werden. Die Darstellung im Frequenzbereich erfordert zwei Diagramme, nämlich eins für den Amplitudenverlauf (Amplitudendiagramm) und eins für den Phasenverlauf (Phasendiagramm).

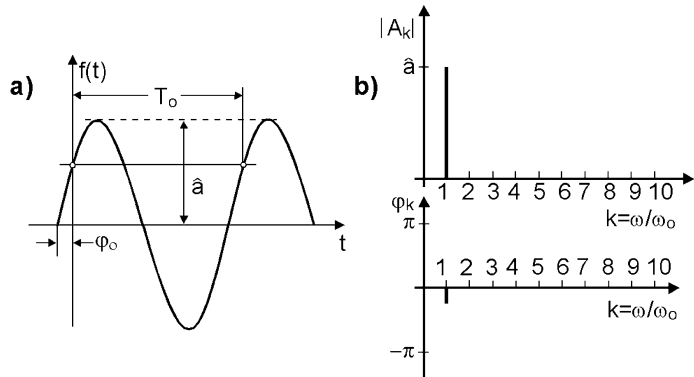


Bild 4-25: Darstellung einer harmonischen Schwingung a) im Zeitbereich b) im Frequenzbereich

In beiden Diagrammen (Teilbild 4-25 b)) ist auf der Abszisse die bezogene Kreisfrequenz $k = \omega/\omega_0$ aufgetragen. Die Darstellung der harmonischen Schwingung im Frequenzbereich besteht dann aus einer senkrechten Linie im Amplitudendiagramm vom Betrag $|A_k| = \hat{a}$ an der Stelle $k = 1$ und im Phasendiagramm aus einer senkrechten Linie vom Betrag $\varphi_k = \varphi_0$ an der Stelle $k = 1$.

Der französische Mathematiker Jean Baptiste Joseph Fourier wies bereits im Jahre 1822 darauf hin, dass man diese Art der Darstellung einer harmonischen Funktion allgemein auf jede **periodische** Zeitfunktion übertragen kann. Er zeigte, dass man jede periodische Funktion (Bild 4-26)

$$f(t) = f(t + nT_0) \quad -\infty < n < \infty$$

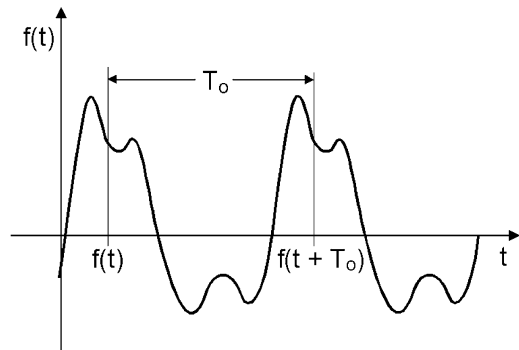


Bild 4-26: Periodische Zeitfunktion

in eine Reihe amplituden- und phasengewichteter, harmonischer Schwingungen zerlegen kann. Zwischen dem Verlauf der Zeitfunktion und den Gewichtungsfaktoren der Summenterme der Reihenentwicklung nach Fourier, die zu Ehren ihres Entdeckers **Fourierkoeffizienten** genannt werden, besteht ein fester Zusammenhang.

Die Erweiterung der Theorie auf **nichtperiodische** Vorgänge, sowie die Modifikation der Theoreme für die Anwendung auf **zeit- und wertdiskrete** Signalfunktionen machen die nach ihrem Entdecker benannten mathematischen Verfahren zur Überführung von Funktionen vom Zeit- in den Frequenzbereich und umgekehrt zu einem der wichtigsten Hilfsmittel der Signalanalyse.

4.3.1 Reelle Form der Fourier-Reihe

Periodische Funktionen wie in Bild 4-26 kann man in eine Fourier-Reihe zerlegen.

$$f(t) = \frac{1}{T_0} \left[\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(\omega_0 t) + b_k \sin(\omega_0 t)) \right], \quad \omega_0 = \frac{2\pi}{T_0} \quad (4.16)$$

Die Funktion $f(t)$ wird durch diese Reihenentwicklung in eine Konstante, den sogenannten **Gleichanteil** der Funktion und in eine Summe von Sinus- und Cosinusschwingungen zerlegt. Dies bedeutet praktisch, dass eine periodische, im allgemeinen nicht harmonische Zeitfunktion in **diskrete** harmonische Schwingungen der Grundfrequenz ω_0 und ganzzahliger Vielfacher hiervon zerlegt werden kann. Die Amplituden- und Phasenlagen der einzelnen diskreten Schwingungen sind dabei durch die so genannten Fourierkoeffizienten a_k und b_k gegeben. Die Koeffizienten kann man aus folgenden Gleichungen bestimmen:

$$a_k = 2 \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t) \cos(k\omega_0 t) dt \quad k = 0, 1, 2, \dots \quad (4.17a)$$

$$b_k = 2 \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t) \sin(k\omega_0 t) dt \quad k = 0, 1, 2, \dots \quad (4.17b)$$

Eine dem Ingenieur vertrautere Schreibweise für die Fourierreihe ist

$$f(t) = A_0 + \sum_{k=1}^{\infty} A_k \cos(\omega_0 t + \varphi_k) \quad , \quad A_0 = 0 \quad , \quad A_k = \frac{a_k}{\cos \varphi_k} = -\frac{b_k}{\sin \varphi_k} \quad .$$

sowie

$$A_k = \sqrt{a_k^2 + b_k^2} \quad \text{und} \quad \varphi_k = \arctan\left(\frac{a_k}{b_k}\right) \quad .$$

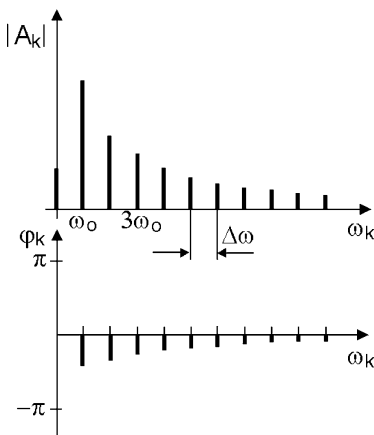


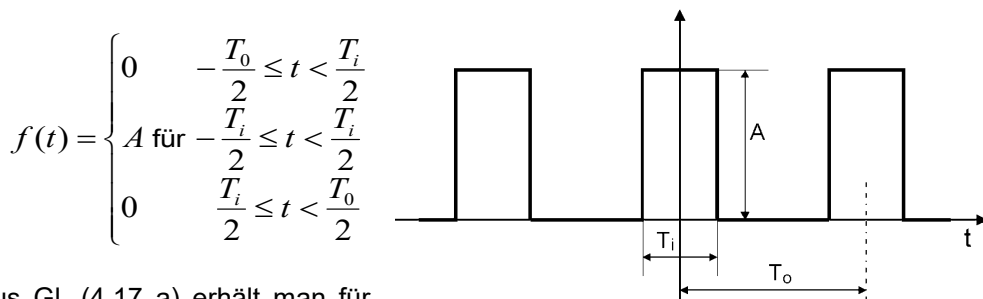
Bild 4-27: Linienspektrum einer periodischen Funktion

Die Darstellung der diskreten Amplitudenwerte A_k und der Phasenwerte φ_k der periodischen Zeitfunktion in Abhängigkeit von ω bezeichnet man als **Spektrum** der Funktion. Es besteht aus Linien an den Stellen $\omega_k = k \cdot \omega_0$, deren Länge durch die Zahlenwerte von A_k und φ_k (Bild 4-27) bestimmt werden. Man nennt solch ein Spektrum auch Linienspektrum. Der Abstand zweier benachbarter zulässiger Stellen für die Existenz derartiger Linien ist umgekehrt proportional zur Periodendauer T_0 , nämlich $\Delta\omega = 2\pi/T_0$. Das bedeutet, dass die Linien um so dichter aufeinander folgen, je größer die Grundperiode T_0 des Signals, bzw. je niedriger die Grundfrequenz ω_0 ist. Die Grundfrequenz ω_0 tritt im

Amplitudenspektrum deutlich mit der größten Amplitude A_k hervor.

Hat man beispielsweise bei einem Schwingungsvorgang einen periodischen Verlauf wie in Bild 4-26 registriert, so kann man durch Entwicklung der Funktion in eine Fourier-Reihe leicht auf die Grundfrequenz der Schwingung zurückschließen. Allerdings sind die Integrale zur Ermittlung der Fourierkoeffizienten nur für ganz einfache periodische Funktionen analytisch lösbar. Deshalb hat es schon immer Hilfsmittel gegeben, die Fourierkoeffizienten näherungsweise mit mechanischen Vorrichtungen oder numerisch in Tabellenform zu berechnen. Seitdem es leistungsfähige Digitalrechner gibt, auf denen man die Koeffizienten mit numerischen Näherungsverfahren berechnen kann, hat die Bedeutung der **Fourier-Analyse** stark zugenommen. Es gibt heute auch speziell für den Fall der Spektralanalyse entwickelte elektronische Geräte, bei denen sich der Benutzer nicht mehr um Berechnungen zu kümmern braucht.

Des besseren Verständnisses wegen sei hier noch einmal an einer mathematisch sehr einfach zu beschreibenden periodischen Funktion die Ermittlung der Fourierkoeffizienten mit Hilfe der Gl. (4.17 a,b) dargestellt. Ein solches Beispiel ist die Rechteckfunktion nach Bild 4-28, die dadurch gekennzeichnet ist, dass sie einen von Null verschiedenen Gleichanteil besitzt. Das Verhältnis von Impulsdauer T_i zur Impulswiederholperiode T_0 bezeichnet man als **Tastverhältnis** $\varepsilon = T_i/T_0$. Die mathematische Beschreibung der Funktion sieht folgendermaßen aus:



Aus Gl. (4.17 a) erhält man für $k = 0$ den Wert für den Fourierkoeffizienten a_0 , der den Gleichanteil beschreibt:

Bild 4-28: Periodische Rechteckfunktion

$$a_0 = 2 \left(2 \int_0^{T_i/2} f(t) dt \right) = 4 \int_0^{T_i/2} A dt = 2AT_i = 2A\varepsilon T_0$$

Die übrigen Koeffizienten gewinnt man entsprechend aus Gl. (4.17 a,b):

$$a_k = 2 \left(2 \int_0^{T_i/2} A \cos\left(\frac{2\pi}{T_0} kt\right) dt \right) = 2\varepsilon AT_0 \frac{\sin(k\pi\varepsilon)}{k\pi\varepsilon}$$

$$b_k = 2 \left(2 \int_0^{T_i/2} A \sin\left(\frac{2\pi}{T_0} kt\right) dt \right) = 0 \quad .$$

Die Fourier-Reihe für die gegebene Rechteckfunktion lautet daher entsprechend Gl.

(4.16) :

$$f(t) = \varepsilon A + 2\varepsilon A \sum_{k=1}^{\infty} \frac{\sin(k\pi\varepsilon)}{k\pi\varepsilon} \cos\left(\frac{2\pi}{T_0} kt\right)$$

$$= 2\varepsilon A \left[\frac{1}{2} + \sum_{k=1}^{\infty} \frac{\sin(k\pi\varepsilon)}{k\pi\varepsilon} \cos\left(\frac{2\pi}{T_0} kt\right) \right] .$$

Für den Fall des Tastverhältnisses $\varepsilon = 0,5$ ist im Bild 4-29 die Funktion $f(t)$ für $k=1-5$ dargestellt, sowie die Verläufe der Funktionen, die sich aus folgenden Fourierkoeffizienten berechnen lassen:

$$a_0 = 0,5A \quad a_1 = A \frac{\sin \pi/2}{\pi/2} = 0,637A \quad a_2 = A \frac{\sin \pi}{\pi} = 0$$

$$a_3 = A \frac{\sin 1,5\pi}{1,5\pi} = -0,212A \quad a_4 = A \frac{\sin 2\pi}{2\pi} = 0 \quad a_5 = A \frac{\sin 2,5\pi}{2,5\pi} = 0,127A$$

Die Teilfunktionen, die sich aus den Fourierkoeffizienten durch Multiplikation mit $\cos \omega_0 kt$ berechnen lassen, summieren sich dann zur Funktion $f(t)$ auf:

$$f(t) = a_0 + f(a_1) + f(a_3) + f(a_5) .$$

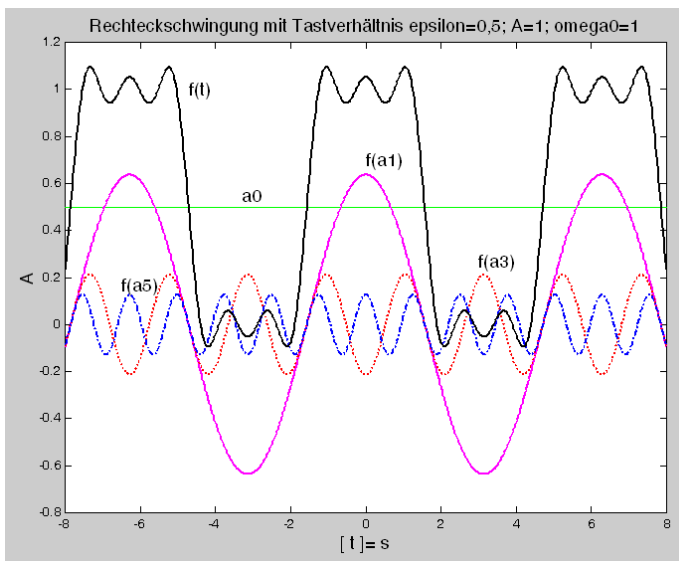


Bild 4-29: Rechteckschwingung approximiert Teilfunktionen der Fourier-Reihe

Am Bild kann man deutlich nachvollziehen, wie sich diese Summenfunktion aus einzelnen Cosinusanteilen zusammensetzt und deren Summe ein der Originalfunktion angenähertes Aussehen hat. Da nur die ersten 5 Spektralanteile verwendet wurden, ist die Übereinstimmung natürlich nur eine Annäherung an das Aussehen der Originalfunktion. Eine Hinzunahme zusätzlicher Fourierkoeffizienten würde die Approximation natürlich weiter verbessern.

Der Wert des Gleichanteils bei $\omega_0 = 0$ sowie der

Verlauf der Einhüllenden des diskreten Spektrums hängen stark vom Tastverhältnis ab. Dies wird in Bild 4-30 gezeigt, in dem die Rechteckfunktion für unterschiedliche Tastverhältnisse, sowie die zugehörigen Betragsspektren in normierter Darstellung abgebildet sind.

Die Höhe der Spektrallinien an den diskreten Stellen $\omega_k = k \cdot \omega_0$ ist proportional zum

jeweiligen Wert des Terms

$$\frac{\sin(k\pi\varepsilon)}{k\pi\varepsilon},$$

d. h. die Einhüllende des Linienspektrums verläuft nach einer Funktion vom Typ $\sin(x)/x$. Diese Funktion wird auch als **Spaltfunktion** bezeichnet.

Zu bemerken ist, dass wegen

$$\sin\left(\frac{k\pi}{2}\right) = 0, \quad \text{für gerade } k$$

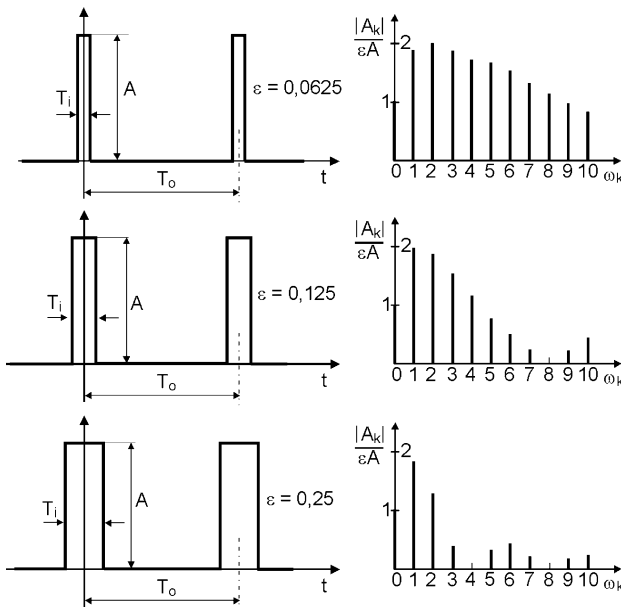


Bild 4-30: Zeitverlauf und normiertes Amplitudenspektrum periodischer Rechteckfunktionen mit unterschiedlichen Tastverhältnissen

für gerade k Spektralanteile verschwinden. Dies hängt vom Tastverhältnis ab und tritt für $\varepsilon = 1/8$ bei den Spektrallinien für $k = 8, 16, \dots$, für $\varepsilon = 1/4$ bei den Spektrallinien für $k = 4, 8, \dots$ auf.

Allgemein gilt, dass für alle **geraden** Funktionen $f(t) = f(-t)$, wie beispielsweise die Cosinus-Funktion, die Fourierkoeffizienten b_k verschwinden, während für alle **ungeraden** Funktionen $f(t) = -f(-t)$, wie beispielsweise die Sinus-Funktion, die Fourierkoeffizienten a_k verschwinden. Dies sieht man auch an den Fourier-Entwicklungen der Funktionen, die als Beispiele für technisch interessante Signale in Bild 4-31 dargestellt sind. Bei der geraden Funktion **Rechteck** fallen in

der Reihenentwicklung die Sinus-Terme heraus, während bei den beiden ungeraden Funktionen **Dreieck** und **Sägezahn** die Cosinus-Terme herausfallen.

An den Beispielen sieht man auch, dass die Spektrallinien höherer Ordnung ($k > 1$) im Amplitudenspektrum um so ausgeprägter auftreten, je unähnlicher die Funktion dem Verlauf einer harmonischen Schwingung ist, bei der es ja ausschließlich die Spektrallinie bei $k = 1$ gibt. Bei der stark von der Sinusform abweichenden Rechteckschwingung gibt es viele Spektrallinien höherer Ordnung mit relativ hohen Amplituden, während bei der dem Verlauf der Sinusschwingung recht ähnlichen Dreiecksschwingung Spektrallinien höherer Ordnung nur mit sehr kleinen Amplituden auftreten. Es ist weiterhin zu beachten, dass aus dem Amplitudenspektrum allein nicht auf den Verlauf der Zeitfunktion zurückgeschlossen werden kann, da bei

gleichen Amplituden- aber unterschiedlichen Phasenspektren sich unterschiedliche Zeitfunktionen ergeben.

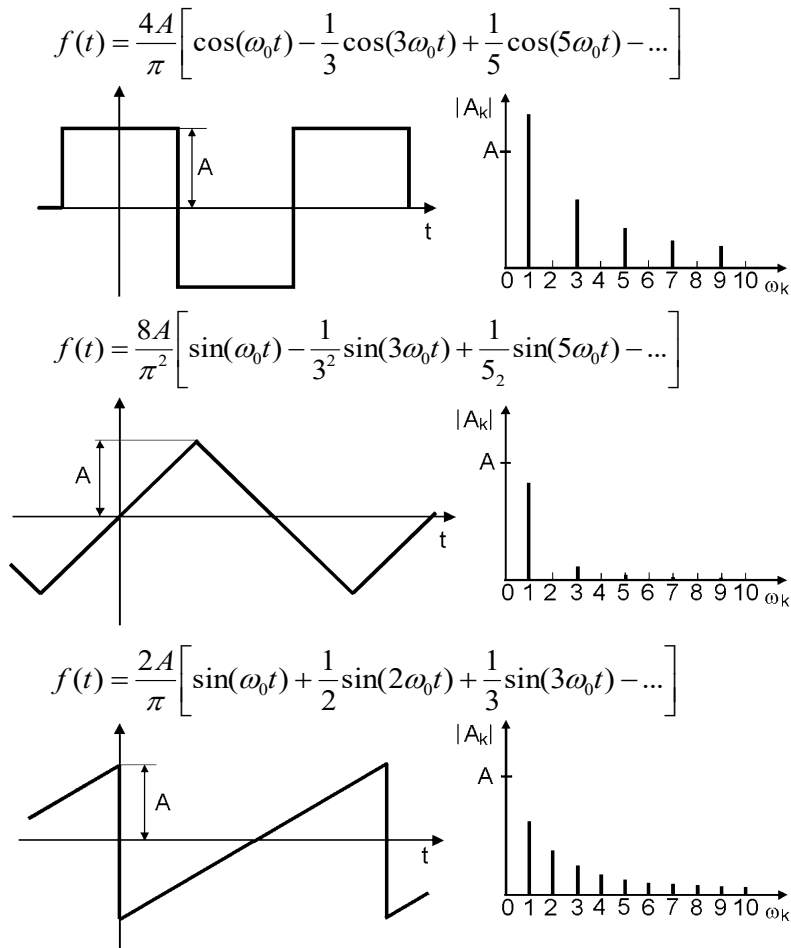


Bild 4-31: Zeitverläufe und Amplitudenspektren periodischer Funktionen

4.3.2 Komplexe Form der Fourier-Reihe

Mit Hilfe der bereits mehrfach angewendeten Euler-Relation

$$e^{ik\omega_0 t} = \cos(k\omega_0 t) + i \sin(k\omega_0 t) \quad ,$$

kann man die Gleichung der Fourier-Reihe in reeller Form (Gl. (4.16)) auch in komplexer Schreibweise formulieren:

$$f(t) = \frac{1}{T_0} \left[\frac{a_0}{2} + \frac{1}{2} \sum_{k=1}^{\infty} (a_k - ib_k) e^{ik\omega_0 t} + \frac{1}{2} \sum_{k=1}^{\infty} (a_k + ib_k) e^{-ik\omega_0 t} \right] . \quad (4.18)$$

Lässt man auch negative Werte für k zu, so gilt wegen der Symmetrie der Sinus- und Cosinus-Funktion

$$a_{-k} = a_k = 2 \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t) \cos(k\omega_0 t) dt$$

$$b_{-k} = b_k = 2 \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t) \sin(k\omega_0 t) dt \quad .$$

Hiermit kann man Gl. (4.18) folgendermaßen vereinfachen:

$$f(t) = \frac{1}{T_0} \sum_{k=1}^{\infty} \frac{1}{2} (a_k - ib_k) e^{ik\omega_0 t} = \sum_{k=1}^{\infty} F_k \cdot e^{ik\omega_0 t} \quad \text{mit} \quad (4.19)$$

$$F_k = \frac{1}{2T_0} (a_k - ib_k) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f(t) e^{-ik\omega_0 t} dt \quad .$$

F_k ist eine komplexe Zahl, die angibt, wie groß die Amplitude der Cosinus- und Sinus-Anteile ist. Da man auch negative k zugelassen hat, haben sich die Amplituden gegenüber der reellen Darstellung nun halbiert.

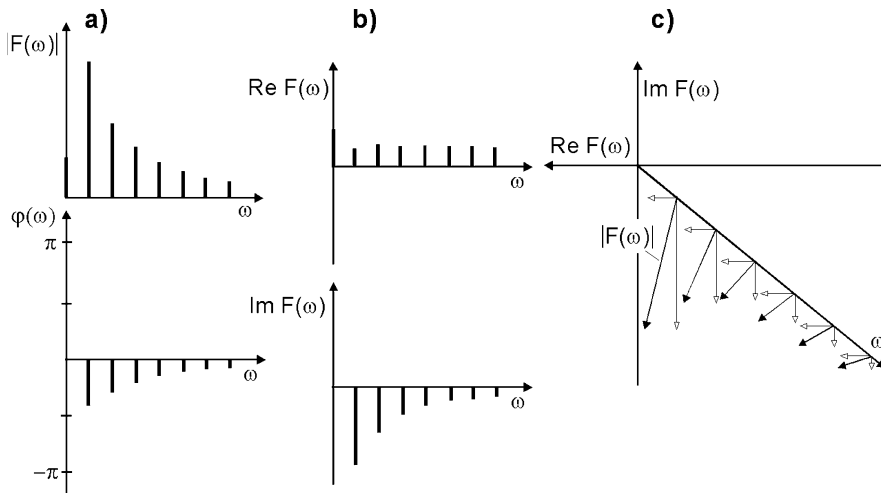


Bild 4-32: Spektrum einer periodischen Zeitfunktion a) Betrag und Phase b) Real- und Imaginärteil c) dreidimensionale Veranschaulichung des Zusammenhangs der beiden Darstellungen aus a) und b)

Im Bild 4-32 sind die beiden prinzipiell gleichwertigen Darstellungsarten des diskreten Spektrums einer periodischen Zeitfunktion gegenübergestellt. Im Teilbild a) ist das Spektrum durch Betrag und Phase und in Teilbild b) durch Real- und Imaginärteil dargestellt. Das Teilbild c) veranschaulicht in dreidimensionaler Darstellung den Zusammenhang zwischen Real-, Imaginärteil und Betragsspektrum.

4.3.3 Fourier-Transformation nichtperiodischer Funktionen

Eine Vielzahl interessanter Schwingungsvorgänge ist **nichtperiodisch**. Die Entwicklung einer Funktion in eine Fourier-Reihe setzt aber gerade Periodizität voraus. Daher benötigt man für die Analyse nichtperiodischer Funktionen andere Möglichkeiten. Man kann dazu die Integraltransformation benutzen, die ebenfalls nach J. B. J. Fourier benannt worden ist, die **Fourier-Transformation**. Transformationen dienen in der Mathematik dazu, Funktionen aus dem **Originalbereich** (hier Zeitbereich) in einen **Bildbereich** (hier Frequenzbereich) zu überführen, um im Bildbereich bestimmte Rechenprobleme leichter lösen zu können. Nach der Rücktransformation erhält man dann die Lösung im Originalbereich.

Zu den bekanntesten und am häufigsten angewendeten Funktionaltransformationen gehört die Logarithmusfunktion. Logarithmiert man beispielsweise eine Potenzfunktion, so wird diese im Bildbereich zu einem Produkt, das man leicht berechnen und dann wieder delogarithmieren kann. Eine weitere Funktionaltransformation ist die **Laplace-Transformation**, die in der Regelungstechnik häufig zur Transformation von Differentialgleichungen benutzt wird. In Kapitel 7.3 wird diese ausführlicher behandelt werden.

In einem Gedankenexperiment kann man den Übergang von periodischen zu nichtperiodischen Funktionen dadurch vollziehen, dass man die Periodendauer T_0 eines allgemeinen periodischen Signals gegen Unendlich gehen lässt. Da der Abstand der Spektrallinien $\Delta\omega$ im diskreten Spektrum umgekehrt proportional zur Periodendauer ist (Bild 4-27), wird beim Grenzübergang aus dem Linienspektrum ein kontinuierliches Spektrum und aus der Summenbildung der Fourier-Reihe ein Integral, das sogenannte Fourier-Integral.

Die Fourier-Transformation ist demnach folgendermaßen definiert:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \cdot e^{-i\omega t} dt \quad \text{Funktion im Bildbereich (Frequenzbereich)}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \cdot e^{i\omega t} d\omega \quad \text{Funktion im Originalbereich (Zeitbereich)}$$

Damit die Fouriertransformierte $F(\omega)$ einer Zeitfunktion $f(t)$ auch existiert, muss folgende Konvergenzbedingung erfüllt sein:

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty \quad .$$

Das Spektrum $F(\omega)$ ist im allgemeinen komplex. Man kann das Spektrum daher wieder durch Betrags- und Phasenfunktion in Abhängigkeit von der Frequenz ω , oder durch ihren Real- und Imaginärteil darstellen.

Ist die Zeitfunktion $f(t)$ eine **gerade** Funktion (s. Abschn. 4.3.1), dann ist $F(\omega)$ auch **gerade** und besitzt nur einen Realteil. Ist $f(t)$ eine **ungerade** Funktion (s. Abschn. 4.3.1), dann ist $F(\omega)$ auch **ungerade** und besitzt nur einen Imaginärteil. Der Realteil von $F(\omega)$ ist immer eine **gerade**, der Imaginärteil immer eine **ungerade** Funkti-

on.

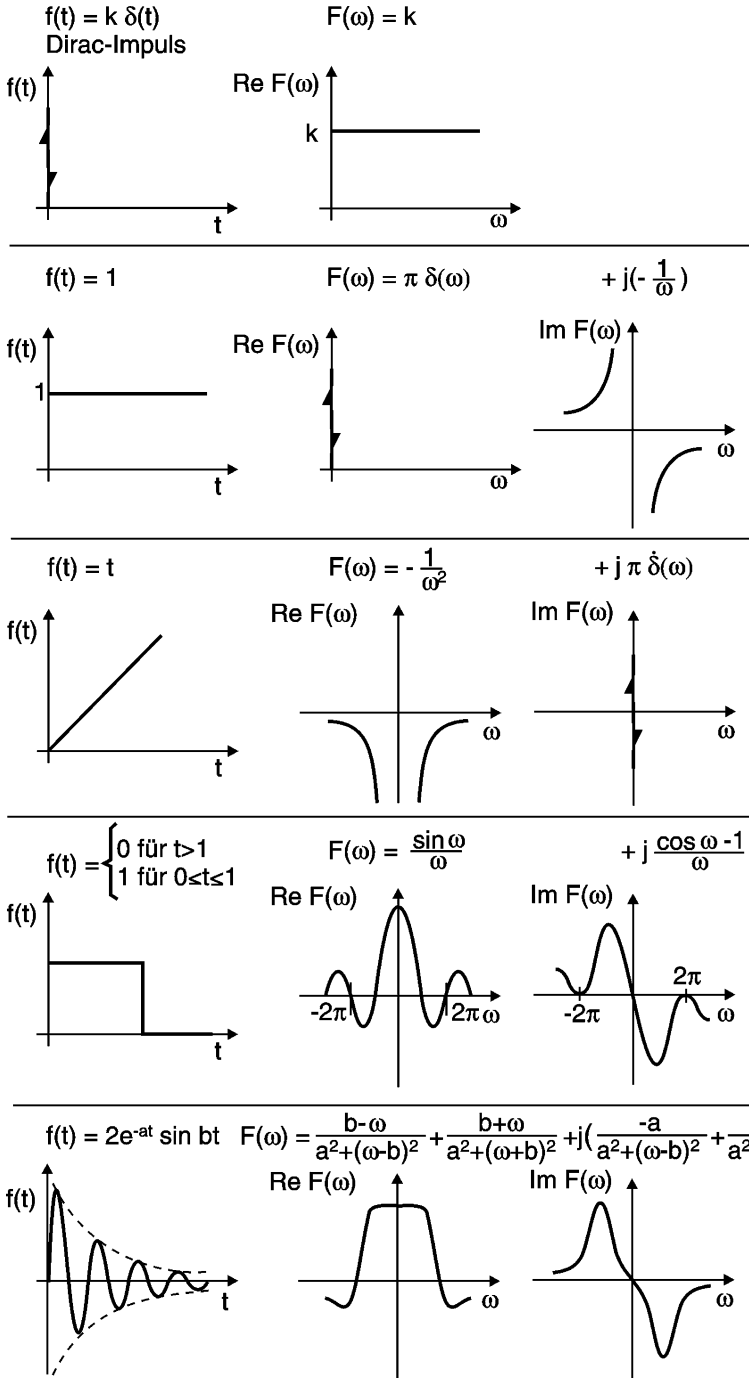


Bild 4-33: Verschiedene nichtperiodische Zeitfunktionen und Real- und Imaginärteil der Fourier-Transformierten. Alle Zeitfunktionen sind für $t \geq 0$ definiert

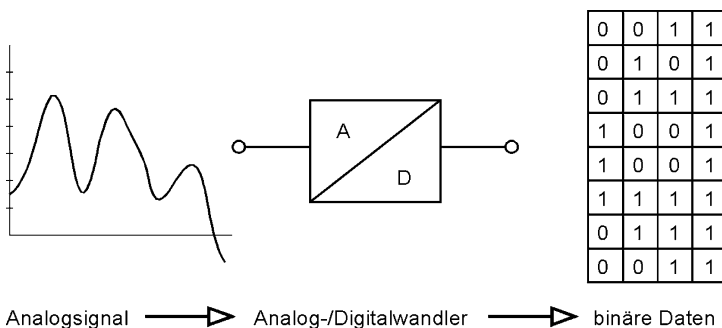
Für die Fourier-Transformation gelten ganz bestimmte Rechenregeln und für viele bekannte Funktionen existieren in der Literatur Korrespondenztabelle. Dies soll hier nicht weiter behandelt werden, da die Durchführung der Transformation heute in der Regel rechnergestützt mit Hilfe der **diskreten Fourier-Transformation** erfolgt. Im Bild 4-33 sind einige Beispiele für die Fourier-Transformierte von Zeitfunktionen als Real- und Imaginärteil dargestellt.

Für die Technik ist die Erkenntnis interessant, dass das Amplitudenspektrum kurzer impulsartiger Funktionsverläufe eine über große Frequenzbereiche konstante, waagrecht verlaufende Linie ist. Im Extremfall, wie beim **Dirac-Impuls** (unendlich kurzer Impuls unendlicher Amplitude) in Bild 4-33, ist das Amplitudenspektrum eine waagerechte Linie parallel zur Frequenzachse. Dies bedeutet, dass das Aufbringen eines solchen Impulses auf ein Bauteil alle Eigenfrequenzen in einem weiten Frequenzbereich anregen kann und daher das Gesamtsystem bei mangelhafter Dämpfung zu großen dynamischen Schwingungsamplituden anregt. Demgegenüber tritt die rein harmonische Anregung eines Bauteils in der Resonanzfrequenz (beispielsweise durch einen Unwuchterreger) sehr viel seltener auf.

4.3.4 Diskrete Fourier-Transformation zur Analyse von Abtastsignalen

Bei der Analyse von Schwingungsverläufen mit Hilfe eines Digitalrechners müssen die analogen Werte des gemessenen Schwingungssignals zur Verarbeitung zuerst einmal in eine binäre Datenstruktur umgesetzt werden. Dazu muss das Signal zu diskreten Zeitpunkten abgetastet, amplitudenquantisiert und in einen Speicher abgelegt oder direkt verarbeitet werden (Bild 4-34). Dadurch wird die Zeitfunktion in eine wert- und zeitdiskrete Datenmenge gewandelt, wobei natürlich der Datenumfang und die Genauigkeit der Einzeldaten endlich ist. Die Genauigkeit hängt von der Auflösung des Analog-/Digitalwandlers ab und die Länge des Datenfeldes von der zur Verfügung stehenden Speicherkapazität des signalverarbeitenden Systems.

Die Analyse solcher Datenfelder, die ein Abbild der ursprünglich wert- und zeitkontinuierlichen Signalfunktion sind, wird mit der **diskreten Fourier-Transformation** (DFT) durchgeführt.



Damit man eine DFT durchführen kann, muss man eine Signalfunktion, wie oben beschrieben, vorher durch **Abtastung** in kurzen Zeitintervallen Δt und anschließende A/D-Wandlung amplitudenquantisieren.

Um eine möglichst geringe Abweichung des Verlaufs der tat-

Bild 4-34: Umwandlung eines analogen Signals in wert- und zeitdiskrete Daten

sächlichen von der digitalisierten Amplitude zu erreichen und auch hochfrequente Signalanteile zu erfassen, sollte Δt möglichst klein oder die Abtastfrequenz $f_a = 1/\Delta t$ möglichst hoch sein. Dem sind natürlich hardwaremäßige Grenzen durch den A/D-Wandler und das Speichermedium gesetzt. Auch die Gesamtmessdauer ist zeitlich limitiert, so dass man Abtastwerte $x_i(t)$ des Signals an einer endlichen Anzahl N äquidistanter Stützstellen innerhalb einer Messdauer T erhält (Bild 4-35).

Bei der Abtastung eines veränderlichen Zeitsignals ist zu beachten, dass zwischen der Frequenz des Signals und der minimal möglichen Abtastfrequenz ein Zusammenhang besteht, der erstmals von Shannon als Theorem formuliert wurde. Das **Theorem von Shannon** besagt, dass die Abtastfrequenz mehr als doppelt so groß wie die höchste im Signal enthaltene Signalfrequenz gewählt werden muss, damit in den Abtastwerten des analogen Signals die ursprünglich in ihm enthaltene Information vollständig erhalten bleibt. Das Analogsignal kann dann aus den Abtastwerten gegebenenfalls rekonstruiert werden (Beispiel **Compact Disc** speichert digital analoge Tonsignale, die aus den Digitaldaten bei Wiedergabe rekonstruiert werden).

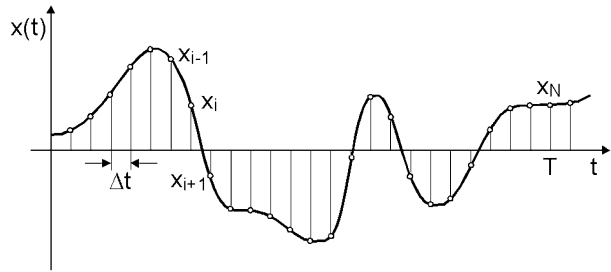


Bild 4-35: Abtastung eines nichtperiodischen Zeitsignals

Bezeichnet man die höchste im Signal enthaltene Frequenz mit $f_{\max}$, die Abtastfrequenz mit f_a und das Abtastintervall mit Δt , dann gilt:

$$f > 2f_{\max} \quad \text{bzw.} \quad \Delta t < \frac{1}{2f_{\max}}.$$

Wird das Abtast-Theorem verletzt, so tritt ein Effekt auf, der im englischen Sprachraum als **Aliasing** bezeichnet wird. Durch Unterabtastung werden im Signal spektrale Komponenten vorgetäuscht, die in Wirklichkeit gar nicht vorhanden sind. Der Grund hierfür ist an einem Beispiel in Bild 4-36 dargestellt, wo durch Abtastung mit zu niedriger Frequenz digitale Daten erzeugt werden, aus denen man ein Originalsignal viel zu niedriger Frequenz rekonstruieren würde.

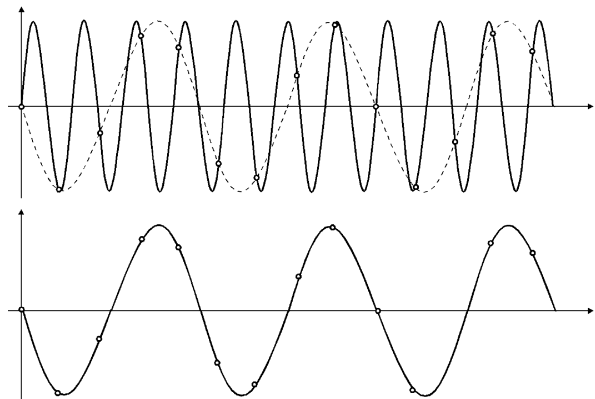


Bild 4-36: Aliasing Effekt: Durch zu geringe Abtastfrequenz wird aus den Abtastdaten eine Schwingung niedrigerer Frequenz als die der Originalschwingung rekonstruiert

Unter den oben genannten Bedingungen wird aus der Formel für

die kontinuierliche Fourier-Transformierte diejenige für die diskrete Fourier-Transformation:

$$F_n(\omega) = \Delta t \sum_{k=0}^{N-1} f_k(\Delta t) \cdot e^{\frac{-i2\pi k}{N}} \quad (n = 0, 1, 2, \dots, N-1) \quad (4.20)$$

$$f_k(\Delta t) = \frac{1}{T} \sum_{n=0}^{N-1} F_n(\omega) \cdot e^{\frac{i2\pi nk}{N}} \quad (k = 0, 1, 2, \dots, N-1)$$

Diese beiden Formeln kann man sich aus der kontinuierlichen Fourier-Transformation entstanden denken, indem man nur ein endliches Integrationsintervall $0 \leq t \leq T$ betrachtet. Das Spektrum für die DFT ist auch nicht mehr kontinuierlich, sondern ist nur für bestimmte Frequenzen

$$\omega_n = n \cdot \Delta\omega$$

definiert, d. h. die kontinuierliche Variable ω wird durch eine diskrete Variable ω_n ersetzt. Der Preis für die Beschränkung auf endlich viele Abtastwerte bei der Berechnung der Fourier-Transformierten ist, dass das ursprünglich kontinuierliche Spektrum nur noch für endlich viele voneinander unabhängige diskrete Frequenzwerte berechnet werden kann. Daher erhält man als Spektrum eine Folge diskreter Linien im Abstand

$$\Delta\omega = \frac{2\pi}{N\Delta t}.$$

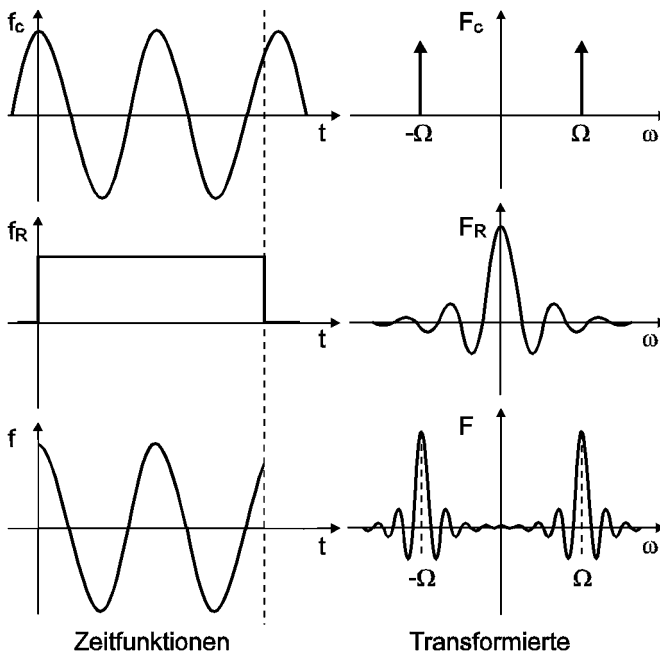


Bild 4-37: Zeitfunktion und Fourier-Transformierte für eine Cosinus-Funktion innerhalb eines Rechteck-Zeitfensters

Die zeitliche Begrenzung bei der Messung eines Signals, das transformiert werden soll, ist wie gesagt ja auch aus technischen Gründen erforderlich. Die Begrenzung der Zeitfunktion kann als Multiplikation der unbegrenzten Zeitfunktion mit einer rechteckförmigen Impulsfunktion, einer so genannten **Fensterfunktion** der Breite

$$T = N \cdot \Delta t$$

aufgefasst werden (Bild 4-37). Wie schon in Bild 4-33 gezeigt, ist für die Rechteckfunktion $f_R(t)$ der Realteil der Fourier-Transformierten eine **Spaltfunktion** $(\sin \omega T)/\omega$. Die im Bild 4-35 dargestellte Cosinusfunktion hat folgende Zeit-

funktion und Fourier-Transformierte:

$$f_C(t) = \cos \Omega t \quad F_C(\omega) = \pi(\delta(\omega - \Omega) + (\omega + \Omega)) \quad .$$

Der Operator δ steht wieder für den **Dirac-Impuls**. Diesem Dirac-Impuls an den Stellen $+\Omega$ und $-\Omega$ entsprach bei der Fourier-Reihe für periodische Funktionen die einzelne Spektrallinie im Betragsspektrum bei der Frequenz Ω der Schwingung. Verbindet man nun die beiden Zeitfunktionen Cosinus und die Fensterfunktion "Rechteck" miteinander, so gewinnt man die Transformierte durch die Rechenoperation **Faltung** der Fourier-Transformation, deren Bedeutung in der Spezialliteratur nachgelesen werden kann. Das Ergebnis dieser

Operation liefert einen Realteil, der zwei Spaltfunktionen an den Stellen $+\Omega$ und $-\Omega$ enthält. Durch die Fensterung der Zeitfunktion wird also ein Spektrum erzeugt, das sich periodisch mit $\Omega = 2\pi/\Delta t$ wiederholt. Außerdem tritt für eine harmonische Schwingung nicht nur eine einzelne Linie im Spektrum auf, sondern ein verbreiterter sinusförmiges Maximum (Spaltfunktion) an der Stelle der Frequenz der Schwingung. Die Darstellung eines solchen Spektrums erfolgt normalerweise als Überlagerung des negativen und positiven Frequenzbereichs im Bereich der positiven Frequenzachse, da ja nur dieser Teil in der Praxis interessiert. In Bild 4-38 ist ein solches überlagertes Betragsspektrum in logarithmischem Amplitudenmaßstab einer gefensternten, harmonischen Schwingung dargestellt.

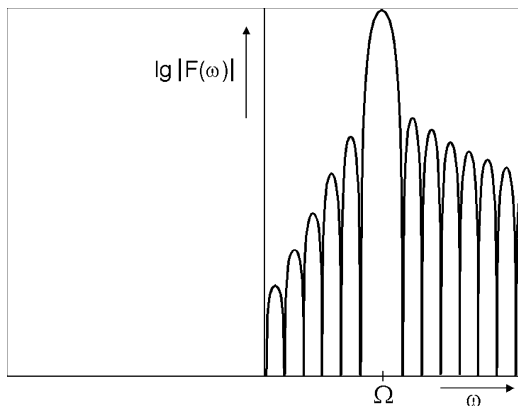


Bild 4-38: Betragsspektrum einer gefensternten harmonischen Schwingung

Von der Wahl der Fensterfunktion ist auch die Art des Spektrums der diskreten Fourier-Transformation abhängig. Bild 4-39 zeigt den Einfluss der Fensterbreite auf das Spektrum der DFT. Dieses Spektrum besteht nun natürlich aus einzelnen Spektrallinien im Abstand $\Delta\omega$ und ist nicht mehr kontinuierlich. Man sieht, dass mit kürzer werdendem Zeitfenster der Abstand der Spektrallinien immer größer wird, bzw. je länger man den Abtastzeitraum wählt, um so dichter rücken die Spektrallinien wieder zusammen, um im Grenzfall der unendlichen Messdauer wieder in das kontinuierliche Spektrum überzugehen. Im letzten Teilbild, in dem die Spektrallinien einen großen Abstand voneinander haben, ist die Einhüllende des DFT-Spektrums für sehr lange Abtastzeiten eingetragen. Man sieht, dass hier ein Messfehler der DFT bezüglich der Frequenz Ω auftritt, da die Spektrallinien der DFT nicht unbedingt mit dem Maximum der kontinuierlichen Fourier-Analyse zusammenfallen müssen. Dieser Fehler wird natürlich um so kleiner, je dichter die Spektrallinien zusammenliegen, d. h. je länger das Abtastintervall ist. Der maximale spektrale Fehler tritt auf, wenn die tatsächliche Frequenz der abgetasteten Schwingung genau zwischen zwei Spektrallinien der DFT fällt und führt bei einem Rechteckfenster zu einem maximalen Amplitudenfehler von ca. 4 dB. Außerdem beträgt der Amplitudenunterschied zwischen dem Hauptmaximum bei der Frequenz Ω und den durch die Fens-

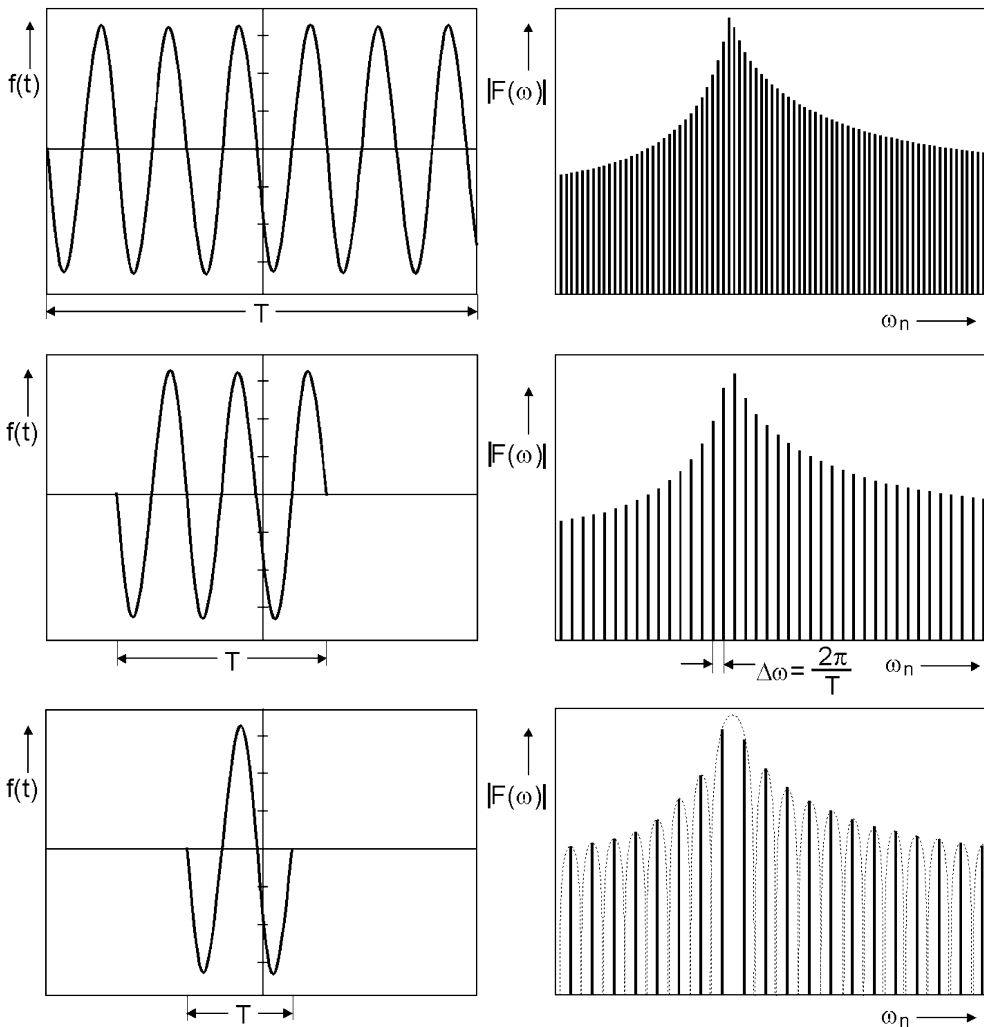


Bild 4-39: Veränderung des DFT-Spektrums einer harmonischen Schwingung bei unterschiedlicher Fensterbreite

terfunktion verursachten Nebenmaxima beim Rechteckfenster nur -13 dB.

Der Amplitudenfehler bei der Abtastung, aber auch der Abstand zu den Nebenmaxima und weitere Kenngrößen lassen sich durch die Form des Abtastfensters beeinflussen. Man verwendet dafür Funktionen wie das Dreieck oder auch kompliziertere Funktionen. Die Verwendung einer Dreiecksfunktion bedeutet, dass nicht wie beim Rechteck alle abgetasteten Amplituden innerhalb der Abtastdauer T mit dem Faktor 1 gewichtet werden, sondern dass bis zur Hälfte des Abtastintervalls die Abtastamplituden mit einem linear von 0 auf 1 ansteigenden Faktor gewichtet werden und in der zweiten Hälfte mit einem linear von 1 auf 0 abfallenden Faktor. Dadurch verändert sich der Amplitudenfehler für das Hauptmaximum auf 1,8 dB und der Abstand zu den Nebenmaxima auf -27 dB. Es gibt in der Literatur noch viele weitere

Fensterformen, mit denen man den Wert verschiedener Kenngrößen des DFT-Spektrums günstig beeinflussen kann. Eine weitere häufig verwendete Fensterfunktion ist das **Hann-Fenster**, das durch folgende Funktion beschrieben wird:

$$f_H(t) = \begin{cases} 0,5(1 + \cos(2\pi/T)) & |t| < T/2 \\ 0 & \text{sonst} \end{cases}$$

Hierdurch verringert sich der Amplitudenfehler auf 1,4 dB und die Dämpfung der Nebenmaxima steigt auf -32 dB. Die positive Auswirkung einer solchen Fensterfunktion bei der DFT-Analyse der periodischen Funktion

$$f_H(t) = 10 \cos \Omega t + 0,25 \cos 2\Omega t + 4 \cos 3\Omega t$$

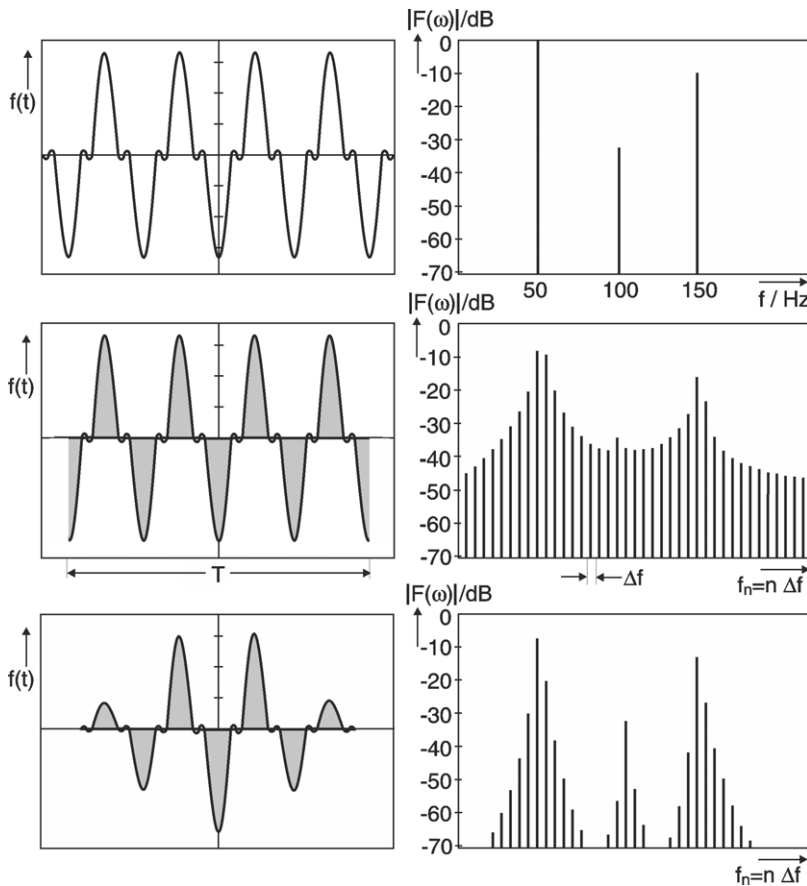


Bild 4-40: Vergleich zwischen Fourier-Reihenentwicklung und DFT-Analyse mit unterschiedlicher Fensterung a) Zeitfunktion und Fourier-Reihenentwicklung b) Rechteckfenster c) Hann-Fenster

verglichen mit dem Ergebnis der DFT-Analyse mit einem Rechteckfenster zeigt Bild 4-40. In Teilbild a) ist der Verlauf der Zeitfunktion und das Amplitudenspektrum einer Fourier-Reihenentwicklung in logarithmischem Maßstab abgebildet. Man erkennt deutlich, dass Ω den Wert von 50 Herz besitzt und dass die Spektrallinie bei

2Ω um den Faktor $1/40 = -32$ dB kleiner und dass die Spektrallinie bei 3Ω um den Faktor $1/2,5 = -8$ dB kleiner ist.

Im Teilbild b) ist die zeitlimitierte Abtastfunktion und das DFT-Spektrum bei Verwendung eines Rechteckfensters dargestellt. Das Hauptmaximum bei 2Ω hebt sich nun kaum noch aus den Nebenmaxima heraus; die Interpretierbarkeit des Spektrums hat deutlich gelitten.

Im Teilbild c) wird nun die Zeitfunktion mit einem Hann-Fenster abgetastet. Das DFT-Spektrum ähnelt jetzt viel deutlicher dem diskreten Spektrum in Teilbild a), der Abstand zwischen Haupt- und Nebenmaxima ist deutlich vergrößert.

Für die DFT-Analyse nach der Gl. (4.20) auf einem Digitalrechner benötigt man zur Berechnung eines vollständigen Satzes diskreter Komponenten N^2 komplexe Multiplikationen. Diese Rechenschritte beanspruchen auf einem Digitalrechner relativ viel Zeit, zumal die Multiplikanden u. a. noch trigonometrische Funktionen enthalten. Daher benötigt ein Rechenalgorithmus zur Berechnung der DFT relativ viel Zeit. Man hat deshalb in der Vergangenheit zur Verkürzung der Rechenzeit sog. **Fast-Fourier-Transformation-Algorithmen (FFT)** entwickelt. Bei dem klassischen von J.W. Cooley und J.W. Tuckey entwickelten FFT-Algorithmus reduziert sich die Anzahl der Multiplikationen auf $2(N/2)^2 + N$ Multiplikationsvorgänge. Ist N beispielsweise gleich 10^3 , so sinkt die Anzahl der Multiplikationen von 10^6 auf $5,01 \cdot 10^5$, d. h. die Anzahl der Multiplikationen ist nahezu halbiert. Andere FFT-Algorithmen reduzieren diese Anzahl sogar auf den Wert $2N \log_2 N$, so dass der Zahlenwert des Beispiels auf $0,993 \cdot 10^4$ sinkt. Dadurch kann erheblich an Rechenzeit für die Durchführung des Algorithmus gespart werden.

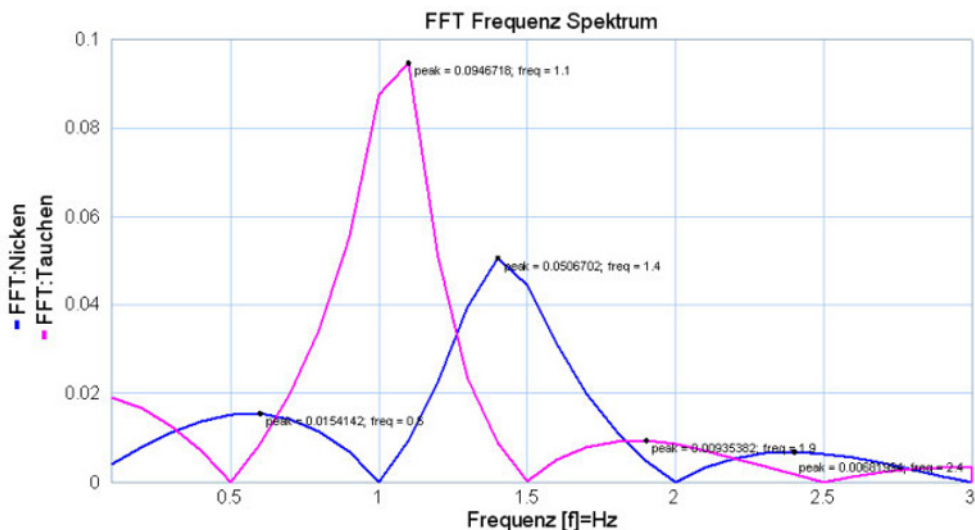


Bild 4-41: FFT-Analyse der Signalverläufe des Tauchens und Nickens beim Fahren über eine Bodenwelle

Die FFT-Analyse wird in der Praxis vor allem zur Identifikation und Messung von Spektralanteilen in Schwingungsvorgängen benutzt. In Bild 4-24 war das Simulationsergebnis des Überfahrens einer Bodenwelle durch ein Fahrzeugmodell gezeigt

worden. Dabei werden Tauch- und Nickschwingungen hervorgerufen. Bild 4-41 zeigt nun eine mit dem gleichen Simulationssystem durchgeführte FFT-Analyse der entsprechenden Schwingungen. Im FFT-Spektrum treten zwei Frequenzen deutlich hervor, die durch die Eigenfrequenzen des Tauch- und Nickvorgangs bestimmt sind. Deren Wert hängt für das Tauchen (Translation) von den gefederten Massen und den Federkonstanten ab. Beim Nicken (Rotation), hängt die Eigenfrequenz vom Trägheitsmoment der Karosserie um den Schwerpunkt, den Radabständen zum Schwerpunkt und den Federkonstanten ab. Durch die FFT-Analyse lassen sich diese Eigenfrequenzen, die wichtig für den Fahrkomfort sind direkt messen, ohne sie berechnen zu müssen. Dies ist natürlich vor allem für Messungen an konkreten Prototypen und nicht so sehr für das Simulationsmodell von Bedeutung.

In einem abschließenden Beispiel soll gezeigt werden, wie man die FFT-Analyse zur Untersuchung komplexer Signale benutzen kann. Dabei geht es um ein stark mit stochastischen Störungen verrauschtes Nutzsignal, das im Rauschuntergrund kaum zu identifizieren ist. Im unteren Bildteil von Bild 4-42 ist die FFT eines rein stochastischen Signals (oberer Bildteil) dargestellt, die mit dem Programmsystem MATLAB² erstellt wurde. Stochastische Signale besitzen ein kontinuierliches Frequenzspektrum. Die Beschreibung durch Real- und Imaginärteil, bzw. Betrag und Phase des Amplitudenspektrums, liefert

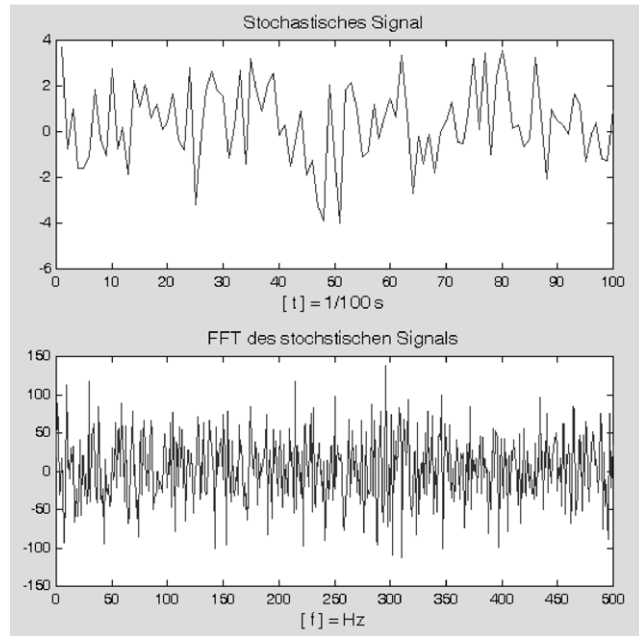


Bild 4-42: FFT eines rein stochastischen Signals

keine sinnvollen Informationen, da die Phase eines statistischen Prozesses statistisch schwankt. Zur Beschreibung des spektralen Verhaltens ist eine phasenlose Funktion wie das **Leistungsdichtespektrum** $W(\omega)$ besser geeignet. Für die Fourier-Transformierte $F(\omega)$ gilt:

$$\begin{aligned} W(\omega) &= [\Re\{F(\omega)\} + i \cdot \Im\{F(\omega)\}] \cdot [\Re\{F(\omega)\} - i \cdot \Im\{F(\omega)\}] \\ &= [\Re\{F(\omega)\}]^2 + [\Im\{F(\omega)\}]^2 \end{aligned}$$

Bild 4-43 zeigt im unteren Bildteil das Leistungsdichtespektrum des stochastischen Signals im oberen Bildteil. Da das analysierte Signal stochastisch ist, treten im Leistungsspektrum natürlich Peaks im ganzen Frequenzbereich auf. Anhand des Spektrums kann man hier nur die Aussage treffen, dass im Signal keinerlei periodische Anteile verborgen sind. Umgekehrt erlangt diese Spektralanalyse Bedeutung für

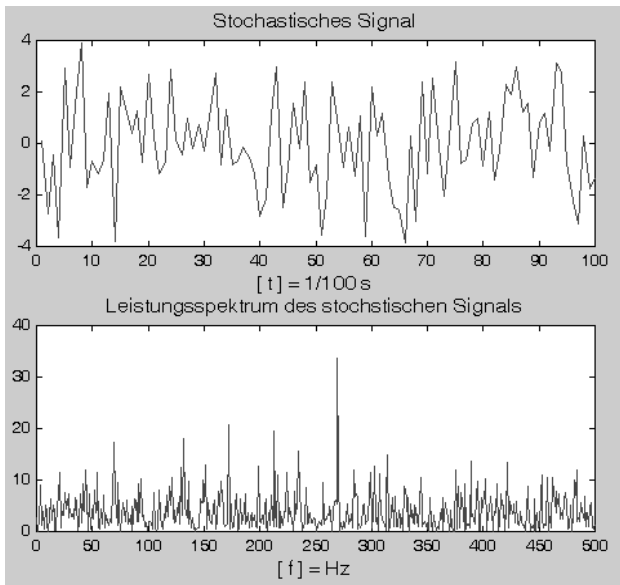


Bild 4-43: Leistungsdichtespektrum $W(\omega)$ eines stochastischen Signals

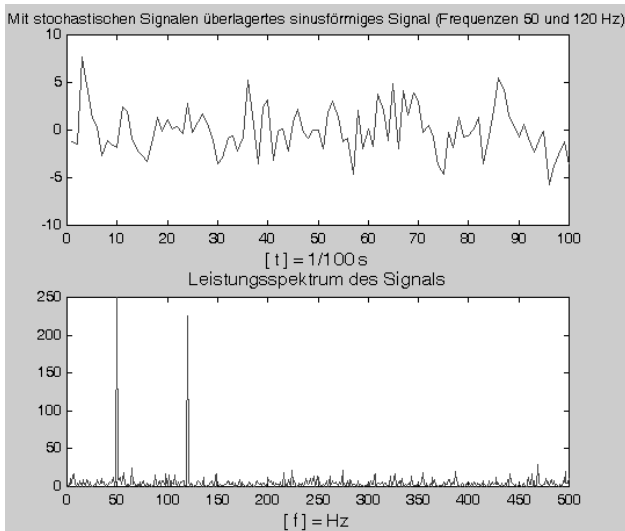


Bild 4-44: Sichtbarmachung eines durch Rauschen gestörten Nutzsignals im Leistungsdichtespektrum

periodische Signale, die durch ein Rauschsignal (stochastisches Signal) überlagert werden, oder für Frequenzgemische periodischer Signale.

Bild 4-44 zeigt im unteren Bildteil das Leistungsdichtespektrum eines Zeitsignals, das sich aus der Überlagerung zweier Sinussignale zusammensetzt und das mit einem Rauschsignal überlagert ist. Die beiden Sinussignalanteile haben die Frequenzen 50 Hz und 120 Hz. Im Zeitsignal des oberen Bildteils kann man die periodischen Signalanteile kaum identifizieren. Das Leistungsdichtespektrum jedoch zeigt sehr deutlich, dass diese Signalanteile im Frequenzgemisch dominierend vorhanden sind.

5 Sensoren

Um solche und andere Signale, wie sie im letzten Kapitel dargestellt wurden, erfassen zu können und daraus Informationen über die Umwelt und den inneren Zustand des mechatronischen Systems zu gewinnen, benötigt man Sensoren. Schauen wir uns das schon mehrfach behandelte Strukturbild eines mechatronischen Systems in Bild 1-24 an, so finden wir die in diesem Kapitel zu behandelnden **Sensoren** oder Messwertaufnehmer an zwei Stellen. Zum einen dienen sie dazu, dem System Informationen über physikalische Größen (Kräfte, Temperaturen, Magnetfelder usw.) aus der Umwelt zu verschaffen und zum anderen finden sie sich als Aufnehmer für Rückmeldungen aus den systemeigenen Bewegungssystemen. Bei der letzten Gruppe, die in jedem mechatronischen System in verschiedenartigem Umfang vorhanden ist, handelt es sich um Sensoren für die Messung von Wegen, Geschwindigkeiten und Beschleunigungen, die benötigt werden, um den Bewegungszustand des Systems zu erfassen und anschließend zu regeln. Die erstgenannte Gruppe kann praktisch alle denkbaren Messwerterfassungen beinhalten, wobei nicht ausschließlich Größen außerhalb des Systems, sondern möglicherweise auch innere Zustände wie beispielsweise Motor- oder Öltemperaturen erfasst werden müssen.

Es gibt Sensoren für mehr als 100 physikalische Größen [5.1]. Berücksichtigt man auch Sensoren für verschiedene chemische Substanzen, so geht die Zahl in die Hunderte. Man kann etwa 2000 grundlegende Sensortypen unterscheiden, die in ca. 100 000 Varianten weltweit kommerziell angeboten werden. Um hier nicht den Überblick zu verlieren, ist zuerst einmal eine Definition und Klassifizierung erforderlich.

Begriffe wie Messfühler, -wandler, -element, Geber u. a. sind seit langem bekannt und definiert. Demgegenüber ist der Begriff **Sensor** erst in den letzten 30 Jahren aufgekommen und reiht sich ein in die Kette schillernder Begriffe wie Mikroprozessor, Transputer, Fuzzy-Technik und andere, die als Zeichen für technische Innovation gelten. Neben dem Begriff Sensor tauchen auch Sensorelement, Sensorsystem, intelligenter oder smarter Sensor, Sensorik oder Sensortechnik auf. Allen Begriffen mangelt es an einer exakten Definition und Abgrenzung. Abgeleitet wurde der Begriff aus den englischsprachigen Begriffen "sense" = Sinnesorgan, bzw. "sensorium" mit der gleichen Bedeutung. Sinnesorgane haben die Aufgabe physikalische Eigenschaften der Umwelt dem Nervensystem (Gehirn) eines Lebewesens zugänglich zu machen. Dazu werden in der Regel nichtelektrische, physikalische Zustände in elektrische Signale umgewandelt, die das Gehirn aufnehmen und verarbeiten kann. Häufig werden daraus Muskelaktivitäten des Lebewesens abgeleitet. Diese biologische Analogie entspricht weitgehend von der Struktur her einem mechatronischen System.

Man sieht an dieser Analogie, dass die Aufgabe eines Sensors sehr komplex sein kann. So findet beispielsweise in einem menschlichen Auge nicht nur die Registrierung des Auftreffens von Photonen statt, wie es einem reinen Messfühler, etwa einer Fotodiode, entsprechen würde. Es gibt darüber hinaus im System "Auge" eigenintelligente Regelungen, die über die Krümmung der Augenlinse ein Bild scharf stellen. Weiterhin findet im Bereich der Netzhaut und des austretenden Sehnervs eine

Datenverrechnung und -kompression statt, um die Anzahl der vom Gehirn in der Sehrinde weiterzuverarbeitenden Informationen zu reduzieren. Es finden also auch weitgehende Messwertwandlungen und eigenintelligente Anpassungen im Bereich dieses Sensors statt.

Darauf Rücksicht nehmend könnte die Definition eines technischen Sensors in Anlehnung an [5.1] folgendermaßen lauten:

Ein Sensor wandelt die zu messende physikalische Größe und ihre Änderungen in elektrische Größen und ihre Änderungen um und verarbeitet diese so, dass sie leicht übertragen und weiterverarbeitet werden können.

Dies weicht von der Definition des **klassischen Messfühlers** ab, mit dem ja im Prinzip etwas ähnliches erreicht werden soll. Dessen Anforderungen an seine Eigenschaften und die Anforderungen an einen Sensor sind in Tabelle 5-1 gegenübergestellt [5.2].

Tabelle 5-1: Bedeutung verschiedener Anforderungen an Sensoren und klassische Messfühler

Eigenschaft	Klassische Messfühler	Sensor
Kosten	weniger entscheidend	möglichst gering
Höhe des Ausgangssignals	möglichst hoch	weniger wichtig
Signal-Rausch-Verhältnis	entscheidend	entscheidend
Störsicherheit	entscheidend	bei Digitalisierung weniger anfällig
Linearität	ausschlaggebend	weniger wichtig
Abhängigkeit von weiteren Parametern	kaum zulässig	korrigierbar
Integrationsfähigkeit	-	sehr erwünscht

Häufig werden Sensoren in Massenprodukten eingebaut, die dort natürlich nur geringe Kosten verursachen sollen. Messfühler findet man in Messgeräten oder Messeinrichtungen, die aufgrund ihrer hohen Präzision ohnehin teuer sind, wodurch deren Kosten nicht so ins Gewicht fallen. Der Forderung nach geringen Kosten beim Sensor kann häufig dadurch entsprochen werden, dass die Weiterverarbeitung des vom Fühler im Sensor erzeugten elektrischen Signals durch eingebaute integrierte Schaltungen in ein gut nutzbares Signal erfolgt. Die Verstärkung auf ein hohes Ausgangssignal verursacht innerhalb des Sensors keine Probleme und Nichtlinearitäten können ebenfalls korrigiert werden. Weiterhin ist eine Korrektur des Einflusses anderer Parameter auf die Messgröße möglich, die dann im Sensor durch weitere Fühler erfasst werden müssen.

Da somit häufig mehrere Fühler und eine Auswerteelektronik im Sensor zusammenwirken, ist der Aspekt der Integrierbarkeit aller dieser Komponenten von Bedeutung. In der Regel treten physikalische Größen als analoge Werte auf. Daher

liefert ein Messverfahren im Sensor, das die physikalische Größe nicht in eine analoge elektrische Größe, wie beispielsweise eine Amplitude (Spannung, Strom), sondern in eine einfache digitale Größe, wie eine Zählung von digitalen Referenzimpulsen umsetzt, weitere Vorteile.

Ein Beispiel hierfür ist die Schichtdickenmessung nichtmagnetischer Schichten auf ferromagnetischen Materialien, wie etwa ein Farbanstrich auf Eisen. Dazu wird eine Magnetspule als Fühler verwendet, die durch eine hochfrequente Wechselspannung gespeist wird. Das Messprinzip beruht darauf, dass der ferromagnetische Körper bei Annäherung der Magnetspule deren Induktivität beeinflusst. Dies wiederum verursacht eine Amplitudenänderung und eine Änderung der Phasenverschiebung gegenüber der Speisespannung (Bild 5-1). Die Änderungen stehen in einem stark nichtlinearen Zusammenhang mit dem Abstand der Spule zum ferromagnetischen Trägermaterial. Damit erhält man indirekt die Schichtdicke des Farbauftrags als Abstand vom Trägermaterial, wenn die Spule auf der Farbschicht aufliegt. Misst man wie früher üblich die Amplitudenänderung, so muss die Anzeige über eine entsprechend nichtlinear bemaßte Skala erfolgen. Das ganze System ist damit von der Messwerterfassung, der Messwertübertragung und der Darstellung mit vielen Unsicherheiten und Störungen behaftet.

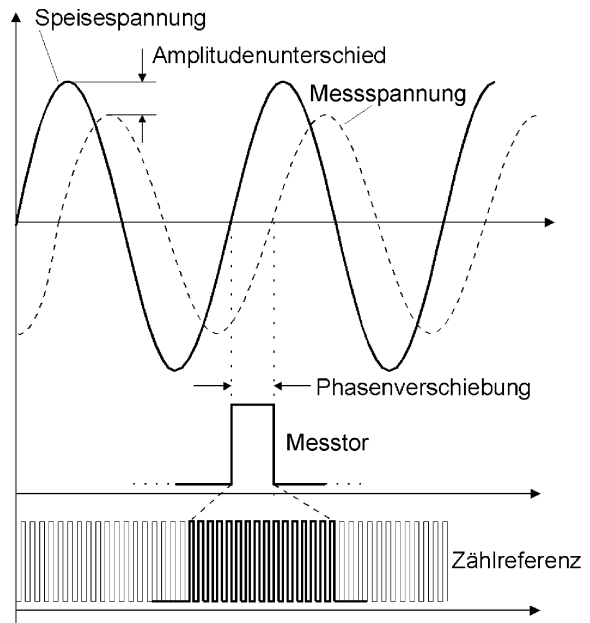


Bild 5-1: Messung der Phasenverschiebung zwischen Speise- und Messspannung durch Auszählen eines Messtors mit einem Referenzsignal

Man kann nun auch mit Hilfe einer Torschaltung, die durch Komparatoren gesteuert die Phasenverschiebung als Zeitdauer misst (Bild 5-1), den Messwert einfach als digitalen Zahlenwert gewinnen. Dazu muss man nur die Zeitdauer des Messtores mit einem genauen Referenzzählsignal ermitteln (Anzahl der gezählten Impulse im $T \times \text{Impulsdauer}$). Das Zählergebnis kann man in einem Mikrorechner beispielsweise tabellarisch linearisieren und erhält somit sofort einen exakten Messwert. Alle erforderlichen Bauelemente wie Spule, Signalgenerator, Torschaltung und Zähler kann man in einem Sensor integrieren, der nur noch eine Speisespannung erhält und direkt einen digitalen Messwert liefert. Dieser kann von einem nachfolgenden Mikrorechner ausgewertet und dargestellt werden.

Der eigentliche Unterschied zwischen klassischen **Messwertaufnehmern** und Sensoren liegt also im Schritt nach der Messwertaufnahme, nämlich in der elektrischen Signalaufbereitung und -verarbeitung. Natürlich wird es zwischen beiden immer

mehr oder weniger starke Überschneidungen geben.

In diesem Sinne spricht man dann auch häufig den eigentlichen Bereich der Messwertaufnahme als "Sensor" an, der zusammen mit der nachfolgenden Signalaufbereitung wie Verstärkung, Filterung, Analog-Digital-Wandlung oder Wertkorrektur ein Sensorsystem bildet. Je nach Integrationsgrad von Sensor und Aufbereitungselektronik im Sensorsystem, kann man dann noch wie in Bild 5-2 dargestellt, unter den Begriffen **Sensor**, **integrierter Sensor** und **intelligenter Sensor** unterscheiden.

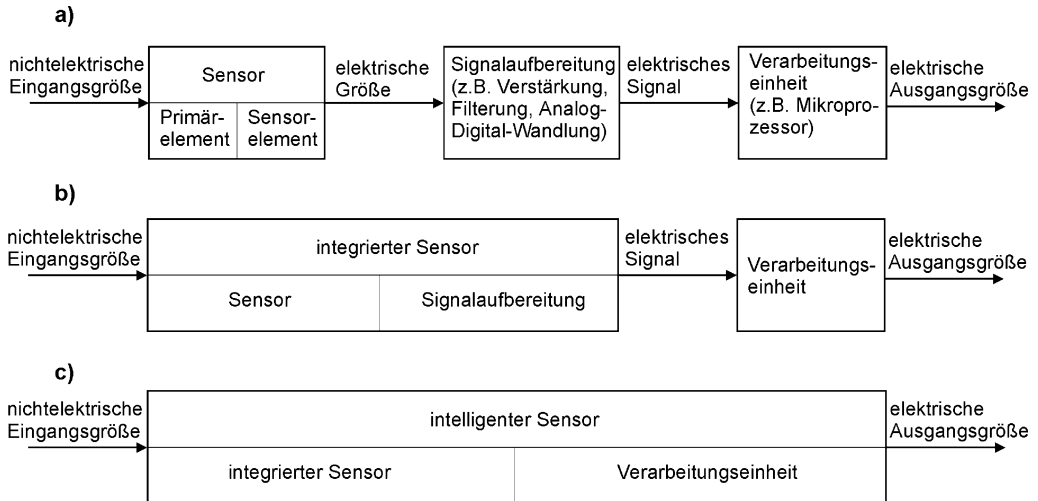


Bild 5-2: Die Arten möglicher Sensorsysteme: a) Sensorsystem mit diskretem Aufbau
b) Sensorsystem mit integriertem Sensor c) intelligentes Sensorsystem

In den nachfolgenden Abschnitten sollen zuerst einige Grundlagen der Messtechnik behandelt werden und anschließend die Messeffekte, die zur Erfassung von physikalischen Größen in Sensoren verwendet werden.

5.1 Messtechnik

Im Bereich der Messtechnik gibt es drei wichtige Teilgebiete, mit denen man sich näher befassen muss. Dies sind die Beschreibung der **Messgröße** und ihrer Maßeinheit, die Beschreibung der **technischen Messmittel** und ihrer Eigenschaften, sowie die Beschreibung und Bewertung des **Ergebnisses der Messung**. In Zusammenhang mit mechatronischen Systemen sind natürlich nur automatische Messvorgänge von Bedeutung, die in der Regel elektrische Messverfahren beinhalten. Im folgenden werden die allgemeine Prinzipien der Messtechnik behandelt. Die Begriffe der Messtechnik sind in VDI/VDE 2600 und in DIN 1319 festgelegt.

5.1.1 Messgrößen und Maßeinheiten

Die Messgröße ist die physikalische Größe, deren Wert durch die Messung ermittelt werden soll. Solche physikalischen Größen können beispielsweise die Temperatur,

der Druck oder ein Weg sein. Der Messvorgang besteht darin, die vor der Messung unbekannte Quantität der Messgröße zu bestimmen. Dazu vergleicht man die Quantität einer bereits durch Definition festgesetzten Maßeinheit mit der Messgröße. Das Ergebnis dieses Vergleichs ist eine Zahl, die angibt, wie oft die Maßeinheit in der Messgröße enthalten ist. Diese Zahl multipliziert mit der Maßeinheit, ergibt den **Messwert**.

Ein einfaches Beispiel für diesen Vorgang ist die Längenmessung mit einem Messschieber. Um etwa den Durchmesser einer zylindrischen Welle zu bestimmen, legt man die beiden Messschenkel an der Oberfläche des Zylinders an und erzeugt dadurch einen Abstand der Schenkel, der dem zu bestimmenden Durchmesser bis auf den nie vermeidbaren **Messfehler** entspricht. Der Messschieber enthält eine Skalierung, die in Untereinheiten der physikalischen Maßeinheit Meter für Längenmaße geeicht ist. Der Nullpunkt der Skalierung steht in einem örtlich feststehenden Bezug zu dem feststehenden Schenkel des Messschiebers, auf dem beweglichen Schenkel befindet sich eine in festem örtlichen Bezug zur Anlagefläche des Schenkels stehende Messmarke. Hat man nun durch den Messvorgang die beiden Schenkel in eine Lage gebracht, die der Messgröße entspricht, so kann man an der Skalierung durch Vergleich mit der Messstrecke feststellen, wie oft die Grundeinheit (hier Millimeter oder 1/10 Millimeter) in der Messgröße enthalten ist. Auf der Skalierung kann man den Zahlenwert ablesen, der mit der Maßeinheit multipliziert den Messwert ergibt. Dieser Messwert stimmt je nach Qualität des Messvorgangs oder abhängig von der theoretischen Messgenauigkeit, mehr oder weniger gut mit der Messgröße überein. Im Beispiel des Messschiebers wird der Messwert unter Verwendung eines Nonius höchstens auf 1/10 mm genau aufgelöst. Die Abweichung zwischen Messgröße und Messwert wird daher günstigstenfalls gleich dieser Auflösung sein. Den Schieber mit der Skalierung bezeichnet man als **Maßverkörperung**. Seine Skalierung beruht auf dem Vorbild eines **Messnormal**s, das in der Regel erheblich genauer als die zum Messen benutzte Maßverkörperung ist.

Um einen Messwert festlegen zu können benötigt man eine Maßeinheit. Weltweit ist ein einheitliches System von Maßeinheiten das **SI-System** (Systeme International) verbindlich vereinbart worden, das für sieben Basisgrößen die zugehörigen Einheiten definiert. Diese Basiseinheiten sind:

Basisgröße	Länge	Masse	Zeit	Stromstärke	Temperatur	Stoffmenge	Lichtstärke
Kurzzeichen	<i>l</i>	<i>m</i>	<i>t</i>	<i>i</i>	<i>T</i>	-	<i>I_T</i>
Basiseinheit	Meter	Kilogramm	Sekunde	Ampere	Kelvin	Mol	Candela
Einheitenzeichen	m	kg	s	A	K	mol	cd

Aus den Basiseinheiten kann man kohärente (abgeleitete) Einheiten entwickeln. Das sind Einheiten, die aus den Einheiten des SI-Systems mit einem Umrechnungsfaktor 1 abgeleitet werden können. Dazu gehören auch Potenzen und Potenzprodukte:

$$1N = 1 \frac{kg \cdot m}{s^2} \quad , \quad 1Hz = \frac{1}{s} \quad , \quad 1m^2 = 1m \cdot 1m \quad .$$

Inkohärente Einheiten sind solche, die durch einen von 1 abweichenden Zahlenfaktor an das SI-System angeschlossen sind:

$$1h = 3600s \quad , \quad 1Kt(Karat) = 0,2g \quad .$$

Tabelle 5-2 zeigt weitere kohärente Einheiten des SI-Systems.

Tabelle 5-2: Kohärente Einheiten des SI-Systems

	Kurzzeichen	Einheit	Einheitenzeichen	Ableitung aus Basisgrößen oder abgeleiteten Größen	Ableitung aus Basiseinheiten
Geschwindigkeit	v	-	-	$v = l \cdot t^{-1}$	$m \cdot s^{-1}$
Beschleunigung	a	-	-	$a = l \cdot t^{-2}$	$m \cdot s^{-2}$
Kraft	F	Newton	N	$F = m \cdot a$	$N = kg \cdot m \cdot s^{-2}$
Arbeit	W	Joule	J	$W = F \cdot l$	$J = kg \cdot m^2 \cdot s^{-2}$
Leistung	P	Watt	W	$P = W \cdot t^{-1}$	$W = kg \cdot m^2 \cdot s^{-3}$
Druck	p	Pascal	Pa	$pa = F \cdot l^2$	$Pa = kg \cdot m^{-1} \cdot s^{-2}$
elektr. Spannung	U	Volt	V	$U = W \cdot i^{-1}$	$V = kg \cdot m^2 \cdot s^{-3} \cdot A^{-1}$
elektr. Ladung	Q	Coulomb	C	$Q = i \cdot t$	$C = A \cdot s$
elektr. Kapazität	C	Farad	F	$C = Q \cdot U^{-1}$	$F = A^2 \cdot kg^{-1} \cdot m^{-2} \cdot s^4$
elektr. Widerstand	R	Ohm	Ω	$R = U \cdot i^{-1}$	$\Omega = kg \cdot m^2 \cdot s^{-3} \cdot A^{-2}$

5.1.2 Messgrößenaufnehmer und Messwertwandler

Im einfachsten Fall kann eine Messeinrichtung aus einer einzelnen Baugruppe bestehen. Häufig jedoch gliedert sie sich in mehrere Baugruppen auf, die Teilfunktionen der Messwertdarstellung übernehmen. Diese einzelnen Baugruppen sind wie in Bild 5-3 in Form einer offenen Wirkkette hintereinander geschaltet. Die Eingangsgröße x ist die zu messende physikalische Größe und die Ausgangsgröße y stellt den Messwert dar, der in der Regel einer elektrischen Signalgröße (**Messsignal**) als Parameter eingeprägt wird. Das Messsignal trägt damit die Information über die Quantität der Messgröße.

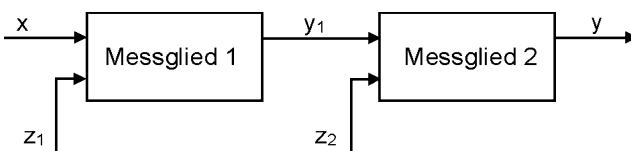


Bild 5-3: Signalflussplan einer Messanordnung

Zwischen den Messgliedern der **Messkette** treten Zwischensignale auf, die dann in den anschließenden Messgliedern nochmals in die endgültige Form gewandelt werden.

den. Ebenfalls eingezeichnet sind Störgrößen z wie etwa die Umgebungstemperatur, die auf jede der einzelnen Baugruppen einwirken und den Messwert verfälschen können.

Die Trennung der Messkette in einzelne Baugruppen findet schon häufig deswegen statt, weil am Messort meist nicht ausreichend Platz zur Anbringung der gesamten Messeinrichtung ist. Ein Grund dafür ist, dass der Vergleich des Messsignals mit der Maßverkörperung meist aufwendigere Schaltungen beinhaltet. Die Standardaufteilung innerhalb der Messkette ist die in **Messgrößenaufnehmer (Sensor, Messfühler)** und **Messwertwandler**.

Der Messgrößenaufnehmer beinhaltet ein **Messprinzip**, mit dessen Hilfe der Wert der Messgröße einer anderen physikalischen Größe aufgeprägt wird. Aus dieser Größe kann man durch den Messwertwandler das endgültige Messsignal erzeugen. Dazu verwendet man innerhalb des Messprinzips möglichst physikalische Zusammenhänge, die eindeutig und zeitlich konstant sind und möglichst gering durch Störungen beeinflusst werden.

Messwandler wandeln Messsignale zum Zweck der Signalübertragung oder der günstigeren Auswertung. Weitere Einrichtungen innerhalb einer Messkette sind häufig noch Signalquellen, deren Signalen die Messgrößen im Messfühler aufgeprägt werden.

Als Beispiel für eine Messkette möge die im folgenden beschriebene Einrichtung zur Positionsmessung dienen (Bild 5-4). Der Messfühler ist hier ein **Potentiometer**, ein ohmscher Schiebewiderstand, an dessen verschiebbaren Abgriff ein variabler Teilwiderstand abgegriffen werden kann. Der Gesamtwiderstand R repräsentiert den Messbereich, über den eine Positionsmessung möglich ist. Damit die Verschiebeposition des Potentiometers in eine Signalgröße umgesetzt werden kann, muss man an das Potentiometer eine Spannung U anlegen. Am Abgriff des Potentiometers mit dem Teilwiderstand R_x ist dann eine dem Weg x proportionale Spannung U_x messbar (Bild 5-4 a). Die Positionsinformation wird demnach durch das Potentiometer zuerst mechanisch in einen proportionalen ohmschen Widerstand umge-

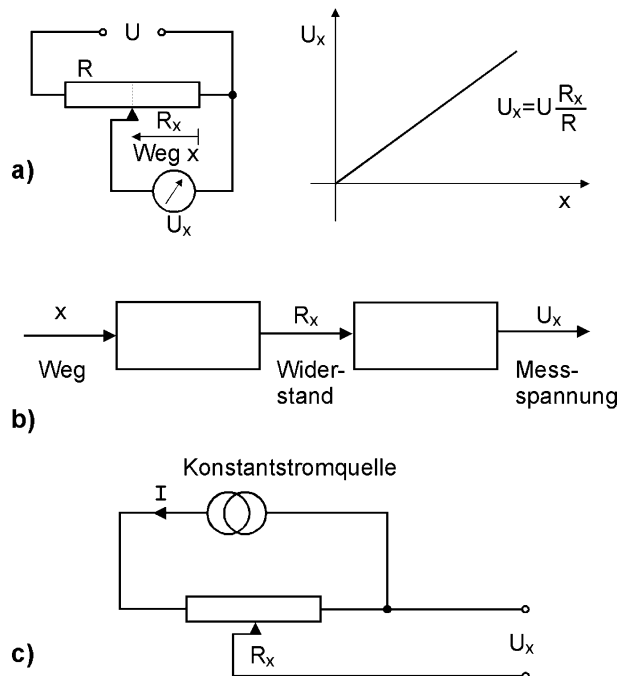


Bild 5-4: Positionsmessung mit Potentiometer

- a) Prinzip
- b) Messkette mit Signalfluss
- c) elektrisches Schaltbild

setzt und anschließend in ein proportionales elektrisches Spannungssignal gewandelt (Bild 5-4 b).

Würde man eine Konstantspannungsquelle zur Speisung des Messwiderstandes benutzen, so würden der Widerstand der Zuleitung und eventuelle Übergangswiderstände mit in den Lastwiderstand der Quelle eingehen und der Wert der Messspannung U_x wäre abhängig von diesen nicht vorhersehbaren Zusatzwiderständen. Man benutzt daher zur Speisung des Potentiometers eine Konstantstromquelle (Bild 5-4 c), so dass über dem Messwiderstand R_x immer die gleiche Messspannung U_x abfällt.

Hinter der beschriebenen Anordnung würde nun in der Messkette in der Regel ein weiterer Messwandler zur Signalaufbereitung folgen (Bild 5-2 a). In solchen Wandlern wird eine Änderung der Größenart, der Signalart, der Signalparameter, der Signalleistung oder des Wertebereichs vorgenommen. Dadurch erreicht man, dass nachfolgende Auswertegeräte mit Einheitssignalen gleicher Charakteristik angesteuert werden können. Für elektrische Größen sind solche üblichen Einheitsbereiche bei der Signalart Gleichstrom 0 - 5 mA, 0 - 20 mA und 4 - 20 mA. Bei der Signalart Gleichspannung ist 0 - 10 V üblich. Durch die Umformung der Messgrößen in Einheitssignale innerhalb der Messkette ergeben sich große Rationalisierungsvorteile bei der Weiterverarbeitung von Messwerten.

Außer der Ausgabe des Messwertes in einem solchen Einheitssignalebereich hat der Messwandler bei dem oben behandelten Beispiel noch die Aufgabe einer Widerstandsanpassung. Registriergeräte wie Schreiber oder Galvanometer haben einen relativ geringen Innenwiderstand. Würde man den Potentiometerausgang direkt mit dem Innenwiderstand R_i des Registriergerätes belasten (Parallelschaltung von R_x und R_i), so würde die Messspannung U_x unzulässig verfälscht werden. Ist R_i aber sehr hochohmig, so wird die Messwertverfälschung vernachlässigbar klein. Diese Anpassung von hochohmigem Eingangswiderstand auf niederohmigen Ausgangswiderstand übernimmt der Messwandler ebenfalls. In den beiden folgenden Abschnitten werden einige elektronische Schaltungen zur Messwertanpassung und -wandlung behandelt.

5.1.2.1 Messwertanpassung

Häufig liefern die physikalischen Messeffekte nur kleine Spannungen oder Ströme und eine zu starke Belastung der Signalquelle **Messfühler** würde zu einer Messwertverfälschung führen. Daher müssen die Messsignale für die Weiterverarbeitung und Fernübertragung verstärkt, eventuell in eine andere elektrische Größe umgesetzt und von störenden Signalanteilen befreit werden. Ein Beispiel hierfür ist in den beiden Bildern 5-5 und 5-6 dargestellt. Bild 5-5 zeigt das mit dem für Mess- und Automatisierungsaufgaben entwickelten Simulationssystem **LabVIEW**¹ aufgestellte Blockschaltbild. Es enthält eine Signalquelle, die ein Sinussignal mit überlagertem, höherfrequenten Gauß'schen Rauschen abgibt. Das Signal ist im Bild 5-6a) dargestellt. Dieses wird in der nachfolgenden Stufe 2fach verstärkt (Bild 5-6b)) und anschließend mit einem Tiefpass gefiltert. Das ursprünglich stark gestörte Messsignal

1 LabVIEW ist ein Produkt der Fa. National Instruments

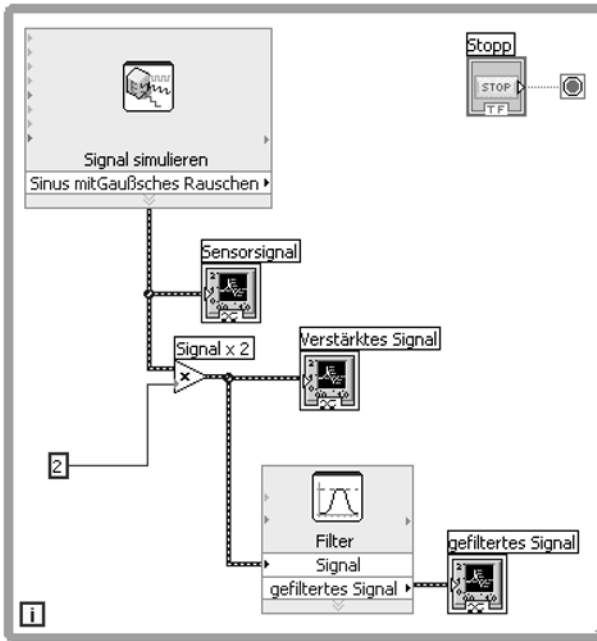


Bild 5-5: Blockschaltbild einer Messkette aus Signalquelle, Verstärker und Filter des Simulationssystems LabVIEW

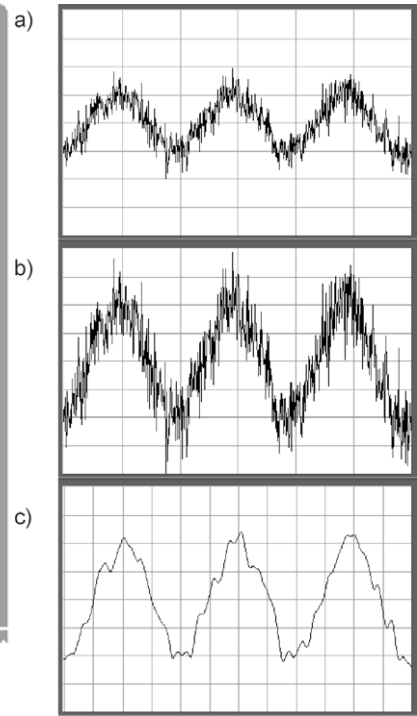


Bild 5-6: Signalanzeigen aus Bild 5-5 a) Nutzsignal
b) verstärktes
c) gefiltertes Nutzsignal

ist nun wieder klarer zu erkennen. Liegen die Frequenzen des Nutzsignals (Sinus) und des Rauschens genügend weit auseinander, so gelingt es mit einem hoch selektiven Filter, das ursprüngliche Nutzsignal als eindeutige Sinuskurve herauszufiltern.

Verstärker haben die Aufgabe ein Eingangssignal zu verstärken und eventuell in eine andere elektrische Größe umgewandelt am Ausgang auszugeben. In der Regel sind diese elektrischen Größen Spannungen oder Ströme, die **Verstärkung** K_V ist das Verhältnis von Ausgangsgröße zu Eingangsgröße (Bild 5-8). Es gibt als Verstärkertypen alle vier möglichen Kombinationen für Strom und Spannungsgrößen, d.h. U/U -Verstärker, U/I -Verstärker, I/U -Verstärker und I/I -Verstärker.

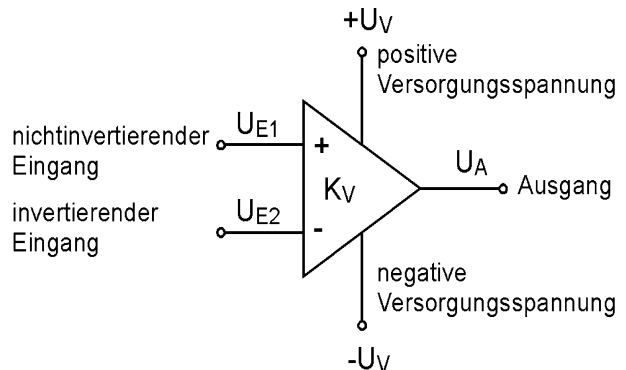


Bild 5-7: Schaltsymbol des Operationsverstärkers

Eine der wichtigsten integrierten Schaltungen zur Signalaufbereitung ist der **Operationsverstärker**. Wie in Bild 5-7 gezeigt, ist der Operationsverstärker ein Differenzverstärker, der die Spannungsdifferenz U_d zwischen den beiden Eingangsspannungen

gen U_{E1} und U_{E2} verstärkt. Der nichtinvertierende Eingang ruft eine gleichsinnige Spannungsänderung, der invertierende Eingang eine gegensinnige Spannungsänderung am Ausgang hervor. Die Spannungsdifferenz $U_d = U_E = U_{E1} - U_{E2}$ wird im Operationsverstärker, dessen innerer Aufbau hier nicht interessieren soll, um den Verstärkungsfaktor K_V (Spannungsverstärkung) verstärkt, der bei handelsüblichen Operationsverstärkern sehr hoch ist und im Bereich $K_V = 10^3 - 10^7$ liegt.

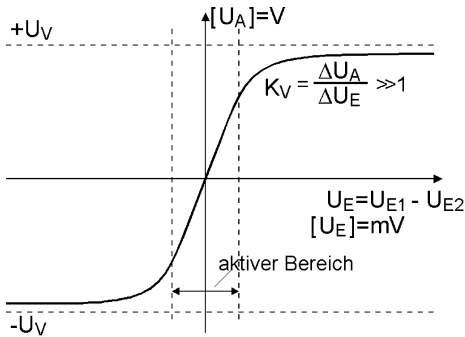


Bild 5-8: Übertragungskennlinie eines Operationsverstärkers

Im Bild 5-8 ist die Übertragungskennlinie eines Operationsverstärkers dargestellt. Aufgrund der hohen Spannungsverstärkung ist der Eingangsspannungsbereich (aktiver Bereich), in dem der Verstärker linear arbeiten kann, natürlich sehr klein. Beim Überschreiten dieses Bereichs gerät die Ausgangsspannung in die **Sättigung**, d. h. sie ändert sich bei weiterer Änderung der Eingangsspannung nicht mehr. Der Verstärker ist daher in dieser Form nicht als lineares Verstärkungselement, sondern nur zum Vergleich von Spannungen (Komparator), die an den beiden Differenzeingängen anlie-

gen, geeignet. Im Abschnitt 7.3.1.2 wird im Zusammenhang mit der Rückkoppelung dargestellt, wie man mit Hilfe eines rückgekoppelten Operationsverstärkers beliebige Verstärkungswerte einstellen kann. Verwendet man in der Rückführung eines solchen Verstärkers Bauelemente wie Spulen oder Kondensatoren, deren Widerstandswert frequenzabhängig ist, so kann man mit Hilfe solcher Operationsverstärker auch Filterfunktionen wie Hoch-, Tief- und Bandpässe erzeugen. Außerdem kann man mathematische Operationen mit Spannungen und Strömen ausführen wie Addition, Subtraktion, Differenzieren und Integrieren.

Weitere Aufgaben des Messwandlers sind heute dadurch gegeben, dass das Messsignal, das häufig als analoge Größe zur Verfügung steht, als digitale Größe zur Weiterverarbeitung in einem Digitalrechner benötigt wird. Er kann daher zusätzlich noch einen **Analog-/Digital-Wandler** enthalten, der den Messwert in der erforderlichen Auflösung (typisch sind Darstellungen als Dualzahl von 8, 10 oder 12 Bit) bereitstellt.

Weiterhin werden Sensoren heute häufig zusammen mit Aktoren und Rechnern über einen **Feldbus** untereinander verbunden. Ein Feldbus ist eine serielle Leitungsverbindung, wobei jedes Kettenglied die auf dem Bus kreisenden Informationen weiterreicht und sich die für ein einzelnes Glied bestimmte Information herausgreift (Aktor), oder eine neue Information einspeist (Sensor). Solche Feldbussysteme sind beispielsweise der CAN-Bus (**C**ontroller **A**rea **N**etwork), der Interbus oder der Profibus, die entweder herstellerspezifisch oder genormt sind. Für die direkte Anschlussmöglichkeit an einen solchen Feldbus muss der Messwandler ebenfalls die entsprechende Signalaufbereitung liefern.

5.1.2.2 Analog-/Digital-Wandler

Wichtige Kenngrößen von A/D-Wandlern sind die digitale Auflösung des Analogwertes und die maximal mögliche Umsetzungsgeschwindigkeit, mit der ein solcher Wandler arbeitet. Sie bestimmen die obere Grenzfrequenz dynamischer Messwerte, die noch gewandelt werden können. Da beide Eigenschaften vom Schaltungsaufwand des A/D-Wandlers her nicht gleichzeitig optimiert werden können, gibt es unterschiedliche Wandlungsverfahren.

Beim Verfahren der **Parallelumsetzung** werden als **Komparatoren**² wirkende Operationsverstärker eingesetzt (Bild 5-9 c)). In Bild 5-9 a) ist das Schaltsymbol und in Bild 5-9 b) die Kennlinie eines Komparators dargestellt. An die Differenzeingänge werden die beiden zu vergleichenden Spannungen angelegt. Je nach dem, ob die eine oder die andere größer ist, springt die Ausgangsspannung auf einen Minimalwert oder einen Maximalwert. Die Höhe des Spannungspegels U_A hängt von der Versorgungsspannung U_V ab. An den nicht invertierenden Eingang des Komparators wird in der Regel die unbekannte zu vergleichende Spannung U_x und an den invertierenden Eingang eine genau eingestellte Referenzspannung U_{Ref} angelegt. Überschreitet U_x den Wert von U_{Ref} , so springt U_A vom Minimalwert auf den Maximalwert. Dies stellt eine binäre Information dar, die mit den logischen Werten "0" und "1" bewertet werden kann.

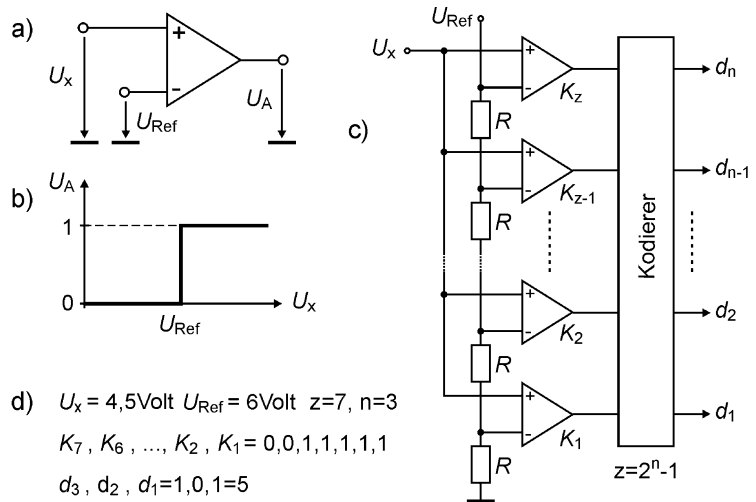


Bild 5-9: A/D-Wandler nach dem Parallelverfahren
a) Komparator b) Kennlinie c) Schaltbild
d) Beispiel für A/D-Wandlung

Der Parallel-A/D-Wandler besteht aus einer parallelen Anordnung von z Komparatoren (Bild 5-9 c)). Die Referenzspannung wird durch eine Kette von z gleichen Widerständen in z gleich große Spannungsintervalle geteilt. Die Spannungen des Spannungsteilers liegen an den jeweiligen invertierenden Eingängen der Komparatoren, während alle nicht invertierenden Eingänge mit der zu wandelnden Messspannung verbunden sind. Alle Komparatoren, deren Vergleichsspannung unterhalb der zu wandelnden Spannung liegen, kippen am Ausgang in den logischen Zustand "1", alle anderen führen den Zustand "0". Man bekommt dadurch eine binäre Darstellung der zu wandelnden Spannung, die jedoch meist durch einen nachfolgenden Kodierer in eine Dualzahl gewandelt wird. Bild 5-9 d) zeigt ein Beispiel mit

² Komparator: lat. Vergleichler

sieben Komparatoren und einer 3 Bit-Darstellung als Dualzahl. Man sieht, dass trotz des schon relativ hohen Aufwandes von sieben Komparatoren man nur einen recht ungenau gewandelten Wert erhält. Schon die Auflösung als 4-Bit Dualzahl erfordert 15, die als 8-Bit Dualzahl 255 Komparatoren. Der Vorteil dieses mit einem hohen Schaltungsaufwand verbundenen Verfahrens liegt in der hohen oberen Grenzfrequenz der Signale, die noch gewandelt werden können. Die Zeit zur Wandlung liegt in der Größenordnung der Zeitverzögerung der elektronischen Schaltungsteile des A/D-Wandlers. Aufgrund dieser sehr geringen Verzögerungszeit ($< 100 \text{ ns}$) werden diese auch Flash-Converter genannt. Übliche Stellenzahlen dieser Wandler liegen bei $n = 4 \dots 8$ Bit. Man kann mit solchen Wandlern Signale bis in den Bereich von einigen 100 MHz wandeln, hat aber in der Regel nur eine kleine Auflösung des Messwertes.

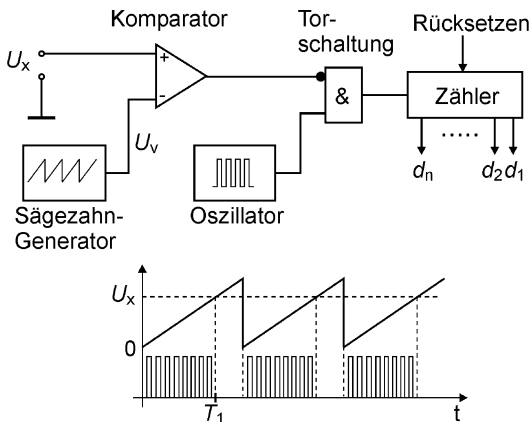


Bild 5-10: A/D-Wandler nach dem Rampenverfahren

Sind die zu wandelnden Signale deutlich niederfrequenter, so kann man Verfahren mit einer kleineren Wandlungsrate aber höherer Auflösung wie das **Rampenverfahren** verwenden. Ein Beispiel hierfür ist der Sägezahn-A/D-Wandler in Bild 5-10. Solange die rampenförmig ansteigende Sägezahnspannung U_V kleiner als die zu wandelnde Messspannung U_x ist, lässt die Torschaltung die Zählimpulse eines Oszillators durch, die dann im Zähler gezählt werden. Wenn U_x die Spannung U_V überschreitet, sperrt das Tor die Impulse und der dann anstehende Zählerstand stellt die digitalisierte Messspannung dar. Am Ende eines Zykluses wird der Zähler zurückgesetzt und die Spannung kann erneut gewandelt werden. Die Spannung U_x wird durch das Verfahren in eine proportionale Zeit T_1 gewandelt und mit Hilfe der Referenzfrequenz des Oszillators ausgezählt. Hohe Genauigkeit setzt natürlich einen stabilen Sägezahn und eine sehr stabile Oszillatorfrequenz voraus. Es ist unmittelbar ersichtlich, dass mit diesem Verfahren nur Signale gewandelt werden können, deren Periodendauer bei einem Vielfachen der Dauer eines Sägezahns liegen. Typische Wandelzeiten solcher A/D-Wandler liegen bei 10 ms. Der Vorteil dieses Wandlers ist, dass hohe Wortbreiten der Dualzahl > 14 Bit mit geringem Aufwand (erfordert Dualzähler höherer Wortbreite) erreicht werden können.

Es gibt weitere Verfahren, nach denen A/D-Wandler aufgebaut werden können, wie

- Halbparallelverfahren
- Spannungs-Frequenz-Umsetzer
- Wägeverfahren
- Delta-Sigma Wandler

Die unterschiedlichen Verfahren bieten unterschiedliche Kompromisse zwischen Schaltungsaufwand, Genauigkeit und Umsetzungsgeschwindigkeit. Genauer hierzu kann der Literatur entnommen werden [5.3], [5.4].

5.1.3 Kenngrößen von Messeinrichtungen

Bei der Auswahl der einzelnen Baugruppen einer Messkette muss man darauf achten, dass diese sinnvoll aneinander angepasst sind. Dabei sind folgende Gesichtspunkte zu berücksichtigen:

- In welchem Wertebereich ändert sich die Messgröße?
- Welche dynamischen Änderungen treten bei der Messgröße auf?
- Welche Genauigkeitsforderungen sind an die Messkette zu stellen?

Diese messtechnischen Eigenschaften der Baugruppen in der Messkette werden von Anbietern in Datenblättern durch Kenngrößen beschrieben. Sie haben den Vorteil, dass man unterschiedliche technische Lösungen miteinander vergleichen kann. Da hohe Anforderungen an die technischen Messeigenschaften und an die Genauigkeit in der Regel mit hohen Kosten verbunden sind³, sind solche Kenngrößen auch für die Optimierung des Preis-/Leistungs-Verhältnisses recht nützlich.

Man kann drei wesentliche Arten von Kenngrößen unterscheiden:

- Statische Kenngrößen, die das Übertragungsverhalten der Messeinrichtung für statische Eingangsgrößen beschreiben.
- Dynamische Kenngrößen, die das Übertragungsverhalten der Messeinrichtung für dynamische Eingangsgrößen beschreiben.
- Fehlerkenngrößen, die das nicht ideale Übertragungsverhalten der Messeinrichtung beschreiben.

Im allgemeinen kann man diese Kenngrößen durch Funktionen beschreiben, die man als Gleichung oder in einem Diagramm darstellen kann. Häufig werden solchen Kennfunktionen einzelne Kennwerte entnommen, die eine punktuelle Aussage über bestimmte Eigenschaften der Messeinrichtung erlauben und mit deren Hilfe man sich leichter und schneller orientieren kann. Daher findet man in Datenblättern häufig solche Kennwerte.

5.1.3.1 Statische Kenngrößen

Für jedes analog arbeitende Glied einer Messkette gibt es einen funktionalen Zusammenhang zwischen Ein- und Ausgangsgröße. Idealerweise sollte dieser Zusammenhang linear sein, d. h. die Übertragungsfunktion lautet

$$y = a \cdot x \quad \text{bzw.} \quad y = a \cdot x + b \quad .$$

³ Ab einer bestimmten durchschnittlichen Messgenauigkeit kann man als Faustformel damit rechnen, dass jede Erhöhung der Genauigkeit im gleichen Verhältnis den Preis des Messgerätes erhöht.

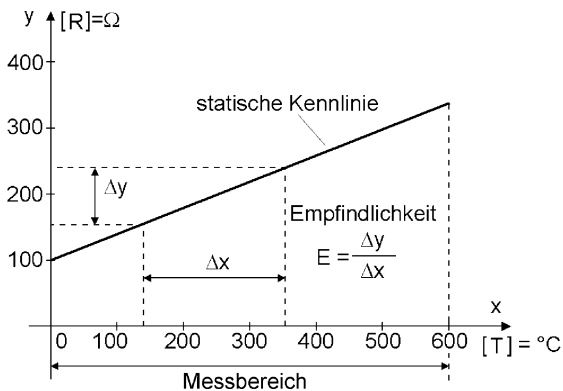


Bild 5-11: Statische Kennlinie eines Temperaturfühlers PT100

Diese Funktion kann man grafisch darstellen; der Funktionsgraph wird als **statische Kennlinie** bezeichnet. Beispielsweise gilt für einen Temperaturmessfühler in Form eines Widerstandsdrahtes vom Typ PT100 (100 Ohm Nennwiderstand bei 0° C) im Temperaturbereich 0° C bis 600° C in erster Näherung folgender funktionaler Zusammenhang zwischen dem aktuellen ohmschen Widerstand R und der Temperatur T :

$$R = R_0 \cdot (1 + \alpha T) \quad .$$

Dabei sind R_0 der Widerstand bei 0°

C und α der Temperaturkoeffizient, mit denen sich die Steigung der statischen Kennlinie ergibt. Diese Funktion ist in Bild 5-11 grafisch dargestellt. Der zulässige Änderungsbereich der Messgröße (hier 0° C - 600° C) wird als **Messbereich** bezeichnet. Die Steigung der Kennlinie

$$R_0 \cdot \alpha T = \frac{\Delta y}{\Delta x} = E \quad ,$$

die man aus einem beliebigen Steigungsdreieck an der Kennlinie ablesen kann, wird als **Empfindlichkeit** E bezeichnet. Dieser Name bezieht sich darauf, dass mit zunehmender Steilheit der Kennlinie, die Ausgangsgröße y immer empfindlicher auf Änderungen der Eingangsgröße x reagiert.

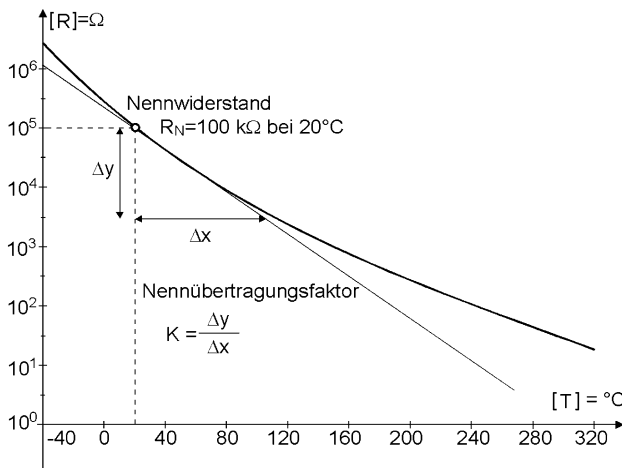


Bild 5-12: Statische Kennlinie eines Heißleiters

Das Ideal einer linearen statischen Kennlinie trifft aber meist nicht zu. So ist auch die Kennlinie des Temperaturempfängers PT 100 leicht gekrümmt und eine den wahren Funktionszusammenhang zwischen Widerstand und Temperatur besser beschreibende Funktion ist:

$$R = R_0 \cdot (1 + \alpha T + \beta T^2) \quad ,$$

d. h. die Funktion ist eine Parabel. Die beiden Koeffizienten haben folgende Werte:

$$\alpha = 3,9082 \cdot 10^{-3} 1/^{\circ}\text{C} \quad ,$$

$$\beta = -5,802 \cdot 10^{-7} 1/^{\circ}\text{C}^2 \quad .$$

Da der Koeffizient β sehr viel kleiner als α ist, ist die Abweichung vom linearen Verhalten nur gering und die Empfindlichkeit kann wie oben beschrieben bestimmt wer-

den. Liegt jedoch eine Kennlinie wie die eines Heißeiters vor (Bild 5-12), die stark vom linearen Verlauf abweicht und einen exponentiellen Abfall des Widerstandes aufweist, so ist die Empfindlichkeit von Punkt zu Punkt verschieden.

Hier wird man nur in kleineren Messbereichen die wahre Kennlinie durch eine gerade **Nennkennlinie** annähern können. Für den **Nennwiderstand** bei 20° C ist eine solche Nennkennlinie mit ihrem **Nennübertragungsfaktor** K eingezeichnet. In einem Messbereich von 0° C bis 100° C beschreibt diese Nennkennlinie mit konstanter Steigung die Empfindlichkeit des Heißeiters einigermaßen genau. Außerhalb dieses Bereichs treten allerdings große Abweichungen auf.

Wird der Messwert in einem Digitalrechner verarbeitet und sind Glieder der Messkette mit Nichtlinearitäten behaftet, so ist bei Kenntnis des funktionalen Zusammenhangs eine nachträgliche Kompensation der Nichtlinearität relativ einfach. Der nichtlineare Funktionszusammenhang kann beispielsweise in Form einer Wertetabelle im Speicher des Rechners abgelegt und die Abweichung vom linearen Verlauf anhand der Tabelle rechnerisch ausgeglichen werden.

Sind die Kennwerte eines Messgliedes nicht bekannt, so muss man dessen statische Kennlinie messtechnisch ermitteln. Diesen Vorgang bezeichnet man als **Kalibrierung**. Ist die Kennlinie linear, so genügt dazu die Erfassung zweier Ausgangswerte für zwei verschiedene Eingangswerte.

Anfangswert und Übertragungsfaktor eines Messgliedes müssen natürlich über längere Zeit gleich bleiben, damit immer der gleiche Zusammenhang zwischen Ein- und Ausgangsgröße erhalten bleibt. Es gibt jedoch Einflüsse wie Alterung von Werkstoffen oder Schwankungen von Temperatur oder Versorgungsspannung, die Einfluss auf die Kenngrößen haben. Durch diese Beeinflussung der Kennwerte werden Fehler hervorgerufen, die durch bestimmte Fehlerkenngrößen (Fehler des Nullpunktes, Fehler des Übertragungsfaktors) ausgewiesen werden.

5.1.3.2 Dynamische Kenngrößen

In der Regel sind die Messgrößen eines mechatronischen Systems dynamische Größen. Wie beim statischen Übertragungsverhalten ist natürlich auch beim dynamischen Übertragungsverhalten zu fordern, dass die Ausgangsgröße der Eingangsgröße entsprechend dem funktionalen Zusammenhang bei jeder Änderungsgeschwindigkeit genügend genau folgt. Reale Messglieder zeigen aber ein davon abweichendes dynamisches Übertragungsverhalten, das besonders bei hohen Änderungsgeschwindigkeiten der Eingangsgröße (hohen Frequenzen) stark vom idealen Übertragungsverhalten abweichen kann. Dies beruht darauf, dass reale Bauelemente energiespeichernde Komponenten enthalten. Dies sind mechanische Feder-/Masse-Systeme, elektrische Kapazitäten und Induktivitäten und thermische Wärmekapazitäten, die aus physikalischen Gründen nie ganz eliminiert werden können. Solche Energiespeicher bewirken eine Verzögerung des Signals zwischen Eingang und Ausgang eines Messgliedes. Diese Eigenschaft und verschiedene Beschreibungsmethoden und Kenngrößen werden im Abschnitt 7.3 über die Regelungstechnik noch ausführlicher behandelt werden, da dieses Verhalten bei allen Baugruppen eines Regelkreises auftritt. Auch im Bereich der mechanischen Komponenten wur-

de diese Problematik bereits behandelt.

Man beschreibt das dynamische Übertragungsverhalten von Messgliedern durch die Antwort des Systems auf bestimmte einfach zu reproduzierende Eingangstestfunktionen. Hierfür verwendet man im wesentlichen die sprungförmige Änderung der Eingangsgröße oder sinusförmige Testsignale veränderlicher Frequenz. Die Systemantwort auf eine sprungförmige Änderung der Eingangsgröße bezeichnet man als **Sprungantwort** oder **Übergangsfunktion**, die Darstellung der Ausgangsgröße in Abhängigkeit der Testfrequenz ω als **Frequenzgang**. Da nicht nur die Amplitude, sondern auch die Phasenlage zwischen Ein- und Ausgang frequenzabhängig sind, unterscheidet man noch zwischen **Amplitudengang** und **Phasengang**.

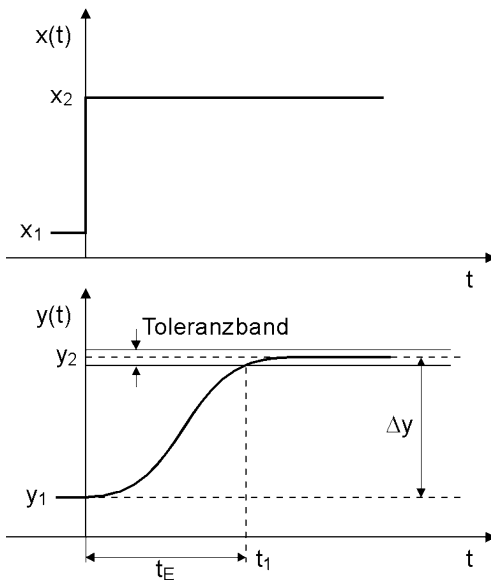


Bild 5-13: Sprungantwort eines Messglieds und Definition der Einstellzeit t_E durch ein Toleranzband

Im Bild 5-13 ist das typische zeitliche Übertragungsverhalten eines energiespeichernden Systems bei Anlegen eines sprungförmigen Testsignals an den Eingang (Sprung von x_1 auf x_2) dargestellt. Da der Energiespeicher erst aufgefüllt werden muss und dies gegen eine die Energieerhöhung behindernde Widerstandsgröße erfolgt (elektr. Widerstand, träge Masse, Wärmeleitwert), folgt die Ausgangsgröße der Eingangsgröße verzögert, entsprechend einer Exponentialfunktion (Übergang von y_1 auf y_2). Diese Sprungantwort besitzt einen für das Übertragungsverhalten charakteristischen Kennwert, die **Einstellzeit** t_E . Dies ist die Zeit zwischen dem Zeitpunkt des Sprungs und dem Zeitpunkt t_1 , zu dem die Ausgangsgröße nur noch um einen spezifischen prozentualen Betrag von ihrem stationären Endwert abweicht und danach dauerhaft in einem vorgegebenen Toleranzband verbleibt. Typische Toleranzbandbreiten sind hierbei $\pm 5\%$, $\pm 1\%$, und $\pm 0,1\%$. Je kürzer die Einstellzeit ist, um so schneller kann das Messglied auf schnelle Änderungen reagieren. Ist die Einstellzeit groß, so können unter Umständen bei schnellen Änderungen der Eingangsgröße gravierende Verfälschungen der Ausgangsgröße auftreten. Daher sollten alle Glieder einer Messkette gleiche Mindestanforderungen an die Einstellzeit erfüllen.

Gibt man ein sinusförmiges Signal auf ein Glied der Messkette, so erscheint in der Regel am Ausgang ebenfalls ein sinusförmiges Signal, das aber in seiner Amplitude verändert wird und dessen Nulldurchgang gegenüber dem Nulldurchgang des Eingangssignals phasenverschoben ist. Das Verhältnis von Ausgangsamplitude zu Eingangsamplitude bei unterschiedlichen Frequenzen des übertragenen Signals dient als Kennfunktion für das dynamische Übertragungsverhalten eines Messglieds. Dieses Übertragungsverhältnis in Abhängigkeit von der Kreisfrequenz ω wird

Amplitudenverhältnis oder **Amplitudenfunktion** $K(\omega) = |F(i\omega)|$ genannt. Dabei ist $F(i\omega)$ der sogenannte Frequenzgang des betrachteten Systems.

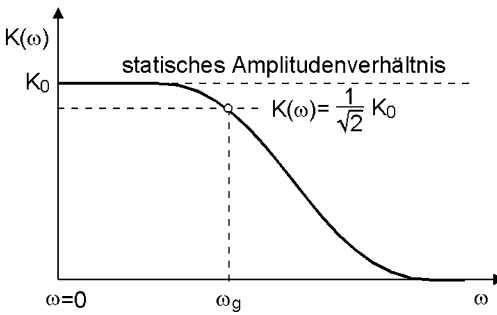


Bild 5-14: Amplitudenfunktion eines Verzögerungsgliedes mit Grenzfrequenz ω_g

Messglieder, die auf sprungförmige Testsignale verzögernd wirken, zeigen ein Amplitudenverhältnis, das mit steigender Kreisfrequenz abnimmt. Bild 5-14 zeigt das typische Übertragungsverhalten eines solchen Messgliedes. Im statischen Fall bei der Frequenz $\omega = 0$ ist das Amplitudenverhältnis K_0 entsprechend der statischen Verstärkung des Messgliedes. Bei höheren Frequenzen nimmt das Amplitudenverhältnis ab und kann bei hohen Frequenzen auf sehr kleine Werte sinken. Dies bedeutet, dass Signale hoher Frequenz

gar nicht mehr vom Messglied übertragen werden. Es gibt daher eine Grenzfrequenz, bis zu der das Übertragungsverhalten gegenüber dem statischen Verhalten kaum beeinträchtigt wird. Diese **Grenzfrequenz** ω_g ist ein wichtiger dynamischer Kennwert, der meist so festgelegt wird, dass dies diejenige Frequenz ist, bei der das Amplitudenverhältnis gegenüber dem statischen Fall auf einen Wert von $1/\sqrt{2} = 0,707$ abgesunken ist. Oberhalb dieser Grenzfrequenz kommt es zu starken Verfälschungen des Messwertes.

Um zu beurteilen, ob ein Messglied für ein dynamisches Messproblem geeignet ist, kann man sowohl die Einstellzeit als auch die Grenzfrequenz heranziehen. Bei nichtelektrischen Systemen ist zur Beurteilung des Übertragungsverhaltens in der Regel die Einstellzeit die wichtigere Kenngröße, da bei solchen Systemen das Erzeugen sinusförmiger Testsignale häufig nicht möglich ist. Elektrische Übertragungssysteme kann man günstig anhand ihrer Grenzfrequenz beurteilen.

5.1.3.3 Fehlerkenngrößen

Wie bereits mehrfach angemerkt, sind der Genauigkeit einer Messkette, aufgrund von Störgrößen und durch vom idealen Übertragungsverhalten abweichendes reales Übertragungsverhalten, Grenzen gesetzt. Die dadurch entstehenden Messfehler muss man beurteilen können, um eine Aussage über die Verlässlichkeit der Messung zu ermöglichen. Dabei ist es natürlich nicht immer erforderlich, so genau wie messtechnisch möglich zu messen, sondern es ist wichtig aus Kostengründen die Messung nur so genau wie erforderlich zu gestalten. Um dies auf einfache Weise beurteilen zu können, werden Messeinrichtungen in Fehlerklassen eingeteilt, wodurch ein weiterer wichtiger Kennwert gegeben ist.

Fehlerursachen wie Schwankungen von Temperatur, Luftdruck, Feuchtigkeit und Versorgungsspannung oder Störgrößen wie elektromagnetische Felder, aber auch zeitliche Änderungen der statischen und dynamischen Kenngrößen der Messglieder, führen dazu, dass der Messwert x vom tatsächlichen, dem **wahren Messwert**

x_w abweicht. Auch werden gleichartige Messgeräte aufgrund von Qualitätsschwankungen in der Fertigung bei gleichem wahren Messwert x_w nicht exakt den gleichen Messwert x anzeigen. Weitere Fehlermöglichkeiten bestehen, wenn die Messeinrichtung nicht genau auf das Messproblem abgestimmt ist und Rückwirkungen auf die Messgröße ausgeübt werden. Solche Rückwirkungen können den Messwert verfälschen.

Unabhängig von der Fehlerursache bezeichnet man die Abweichung zwischen x und x_w , also deren Differenz als **absoluten Fehler** Δ der Messeinrichtung:

$$\Delta = x - x_w \quad \text{Absoluter Messfehler}$$

Um die Größe dieses Fehlers zu bestimmen, kann man zwei Verfahren verwenden.

- Gibt es für die Messgröße ein genaues Messnormal, dessen Maßinformation man auf den Eingang der zu untersuchenden Messeinrichtung geben kann, so kann man das Ausgangssignal der Messeinrichtung mit einem möglichst genauen Messgerät messen, oder wenn die Messeinrichtung ein Messgerät ist, einfach von ihr ablesen. Von diesem Messwert x zieht man den wahren Eingangswert x_w des Messnormals ab und erhält dadurch den absoluten Fehler.
- Gibt es kein Messnormal, so benötigt man ein Vergleichsmessgerät, das in seiner Messgenauigkeit mindestens um eine Größenordnung genauer ist (Fehlergröße um eine Zehnerpotenz kleiner). Dann kann man einen typischen Messwert sowohl auf die zu untersuchende Messeinrichtung, als auch auf das Vergleichsmessgerät geben und wieder die Differenz aus beiden Messungen bestimmen. Dies ist der absolute Fehler.

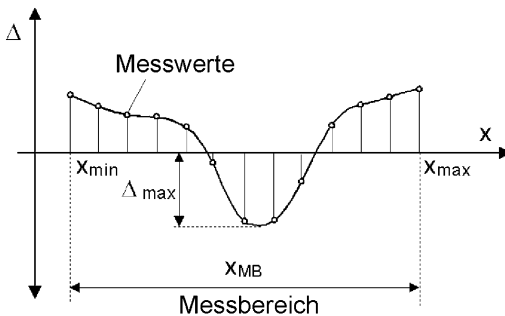


Bild 5-15: Fehlerkurve einer Messeinrichtung

Eine solche Messung des absoluten Fehlers muss man im gesamten Messbereich einer Messeinrichtung durchführen, da der absolute Fehler nicht überall gleich sein wird. Bild 5-15 zeigt einen typischen Beispiel für eine solche **Fehlerkurve**.

Normalerweise ist der absolute Messfehler einer Messeinrichtung nicht die entscheidende Größe, da bei gleichem absoluten Fehler die Genauigkeit einer Messeinrichtung vom absoluten Mess-

wert bzw. vom erforderlichen Messbereich abhängt. Will man die Brauchbarkeit verschiedener Messeinrichtungen vergleichen, so ist der **relative Messfehler** besser geeignet. Man kann den relativen Fehler so bilden, dass man den absoluten Fehler auf den Messwert bezieht. Bezieht man den absoluten Fehler auf den Messbereich, so spricht man auch vom reduzierten Fehler:

$$\delta = \frac{\Delta}{x} \quad \text{Relativer Fehler} \qquad \hat{\delta} = \frac{\Delta}{x_{MB}} \quad \text{reduzierter Fehler}$$

Die Hersteller von Messeinrichtungen leiten aus der Fehlerkurve die so genannte **Fehlerklasse** ab, ein Kennwert, der es dem Anwender erlaubt, schnell die Genauigkeit eines Fühlers oder Messwandlers beurteilen zu können. Der Hersteller garantiert mit der Fehlerklassenangabe, dass unter festgelegten Messbedingungen der relative Messfehler bezogen auf den Endwert des Messbereichs um nicht mehr als einen der Fehlerklasse entsprechenden prozentualen Wert nach oben und unten abweicht. Fehlerklasse 1 bedeutet beispielsweise bei einem Voltmeter, dass der Messfehler nicht größer als $\pm 1\%$ v. E. wird. Bei einem Messbereich von 10 Volt wäre daher der relative Messfehler $\leq \pm 0,1$ Volt. Für Betriebsmessgeräte sind Fehlerklassen 1; 1,5; 2,5; 5 üblich, Feinmessgeräte haben Fehlerklassen 0,1; 0,2; 0,5.

Die Möglichkeit Fehler abzuschätzen und gegebenenfalls zu korrigieren ist von der Art der Fehler abhängig. So gibt es Fehler, die aufgrund der Fehlerursache immer gleichen absoluten Betrag und gleiches Vorzeichen besitzen. Solche Fehler, die auch bei Wiederholung einer Messung reproduzierbar auftreten, heißen **systematische Fehler**. Ein solcher systematischer Fehler Δ_{sys} bei einem Temperaturfühler PT100 wäre beispielsweise die Annahme einer linearen Kennlinie. Da der Verlauf des durch diese Annahme auftretenden Fehlers bekannt ist, kann er bei jeder Messung berücksichtigt und das Messergebnis x auf den Wert x_{kor} korrigiert werden:

$$x_{\text{kor}} = x - \Delta_{\text{sys}}$$

Viel häufiger sind bei Messungen auftretende Fehler von zufälliger Natur, da ihre Ursachen in verschiedenartigen, nicht reproduzierbaren, regellosen Störungen liegen (Temperaturschwankungen, Reibvorgänge, magnetische Felder). Für sie ist charakteristisch, dass Vorzeichen und Betrag des Fehlers von Messung zu Messung verschieden sind.

Um Aussagen über solche Fehler machen zu können, muss man zu Methoden der **Statistik** greifen. Macht man eine Messung nur einmal, so ist eine einigermaßen zuverlässige Aussage über zufällige Fehler nicht möglich. Führt man jedoch eine Messung häufiger durch, so kommt man mit Hilfe entsprechender Rechenverfahren je nach Anzahl der Stichproben zu Aussagen über Fehlergrößen, die hohen Wahrscheinlichkeitswert besitzen. Zwei dieser statistischen Fehlerkennwerte sind der **arithmetische Mittelwert** $\bar{\Delta}$ und die **Standardabweichung** S , der quadratische Mittelwert aller Abweichungen der Einzelwerte vom arithmetischen Mittelwert:

$$\bar{\Delta} = \frac{1}{n} \sum_{i=1}^n \Delta_i \quad \text{mit } \Delta_i : \text{Stichprobenkennwert der } i\text{-ten Messung}$$

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\Delta_i - \bar{\Delta})^2}$$

Der arithmetische Mittelwert ist der wahrscheinlichste Wert für die wahre Größe des Fehlers Δ , die Standardabweichung ist ein Maß für die Genauigkeit dieses Wertes. Die Güte des arithmetischen Mittelwertes nimmt mit zunehmendem Stichprobenumfang zu, bleibt aber immer ein Schätzwert und wird erst für $n = \infty$ identisch mit dem tatsächlichen Fehler. Die Güte des Mittelwertes kann mit einem weiteren Kennwert, dem **Vertrauensbereich** $m_{\bar{\Delta}}$ abgeschätzt werden. Er ist definiert als:

$$m_{\Delta} = \pm \frac{t \cdot S}{\sqrt{n}}$$

Dabei ist t ein Korrekturfaktor der von n und der **Wahrscheinlichkeit** P für das Zutreffen der Fehleraussage abhängt. Die nachfolgende Tabelle zeigt Werte des Faktors t für verschiedenen Stichprobenumfang n und verschiedene übliche Wahrscheinlichkeiten P :

n	P = 68,3%	P = 95%	P = 99%	P = 99,73%
3	1,15	2,8	4,6	6,6
10	1,06	2,3	3,2	4,1
20	1,03	2,1	2,9	3,4
50	1,01	2,0	2,7	3,1
100	1,00	1,97	2,6	3,04
200	1,00	1,96	2,58	3,0

Der Fehler aus n Einzelmessungen ergibt sich mit dem Vertrauensbereich zu

$$\Delta = \bar{\Delta} \pm m_{\Delta}$$

für die statistische Wahrscheinlichkeit P . Eine Steigerung der Zahl n wirkt proportional zu $1/\sqrt{n}$ auf den Vertrauensbereich ein, d. h. mit der Steigerung von n auf große Werte (> 10) wird die Verbesserung des Vertrauensbereichs immer geringer. Daher ist ein Mittelwert aus $n > 10$ Messungen zu wählen.

Mit den Kennwerten $\bar{\Delta}$, S und m_{Δ} können zufällige Fehler analysiert und bewertet werden. Bei der Abschätzung des Vertrauensbereichs ist eine wichtige Bedingung zu prüfen: Die **absolute Häufigkeit** H der Einzelfehler muss annähernd normalverteilt sein. Rund 70% aller Fehlerwerte liegen dann innerhalb eines Intervalls $\pm S$ um den Fehlermittelwert. Die grafische Darstellung der Fehlerhäufigkeit ergibt bei einer **Normalverteilung** eine charakteristische **Glockenkurve** (Bild 5-16).

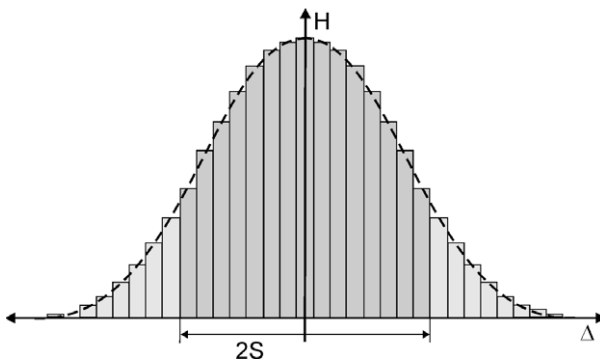


Bild 5-16: Häufigkeitsdiagramm normalverteilter zufälliger Fehler

Häufig wird eine Messanordnung aus mehreren Gliedern zusammengesetzt, die jeweils mit einem Messfehler behaftet sind. Dann ist es wichtig zu wissen, welchen Fehler die Einzelglieder der Messkette zum Gesamtfehler beitragen. Hierzu muss man die Gesetzmäßigkeiten der **Fehlerfortpflanzung** berücksichtigen.

Bei systematischen Fehlern Δx_1 , Δx_2 der unabhängigen Messgrö-

ßen x_1 und x_2 ergibt sich bei einem funktionalen Zusammenhang $y = f(x_1, x_2)$ der maximale absolute Fehler $\Delta y_{\max}$ entsprechend der nachfolgenden Tabelle:

Funktion	y	$\Delta y_{\max}$	$\Delta y_{\max}/y$
Summe, Differenz	$x_1 \pm x_2$	$ \Delta x_1 + \Delta x_2 $	$\frac{ \Delta x_1 + \Delta x_2 }{ x_1 \pm x_2 }$
Produkt	$x_1 \cdot x_2$	$ x_2 \Delta x_1 + x_1 \Delta x_2 $	$\left \frac{\Delta x_1}{x_1} \right + \left \frac{\Delta x_2}{x_2} \right $
Quotient	x_1 / x_2	$\frac{ x_2 \Delta x_1 + x_1 \Delta x_2 }{ x_2 ^2}$	$\left \frac{\Delta x_1}{x_1} \right + \left \frac{\Delta x_2}{x_2} \right $
Potenz	x_1^n	$n \cdot x_1^{n-1} \Delta x_1$	$n \cdot \left \frac{\Delta x_1}{x_1} \right $

Bei zufälligen Fehlern der Messgrößen x_1 und x_2 und zugehörigen Standardabweichungen S_1 und S_2 ergibt sich die Standardabweichung S_y des Messergebnisses für die Operationen Addition und Subtraktion zu

$$S_y = \sqrt{S_1^2 + S_2^2}$$

und für die Operationen Multiplikation oder Division zu

$$S_y = y \sqrt{\frac{S_1^2}{x_1^2} + \frac{S_2^2}{x_2^2}} \quad .$$

5.2 Messeffekte

Bei dem eigentlichen Messwertaufnehmer oder einfachen Sensor muss man zwischen dem im Aufnehmer verwendeten physikalischen Effekt und der zu messenden Größe unterscheiden, da der Wert vieler physikalischer Größen nur aus ihrem Einfluss auf bestimmte Messeffekte rückgeschlossen werden kann. Darüber hinaus muss unterschieden werden, ob durch den Messeffekt die Energie der Messgröße direkt in ein elektrisches Signal umgesetzt wird, oder ob die Energie der Messgröße nur zur Steuerung der Energie eines Signals aus einer anderen Quelle verwendet wird. Im ersten Fall ist der Sensor nur ein passiver Energiewandler, im zweiten Fall muss das Sensorsystem eine Energiequelle enthalten. Beispiele für solche Wandlertypen aus dem Bereich der Temperaturmessung sind das Thermoelement als passiver Wandler und das Widerstandsthermometer als aktiver Wandler. Wie man an diesem Beispiel leicht erkennen kann, benutzen unglücklicherweise passive Wandler aktive Bauelemente (Thermoelement liefert eine zur Temperatur proportionale Spannung) und aktive Wandler passive Bauelemente (temperaturproportionale Widerstandsänderung verursacht Stromänderung einer externen Spannungsquelle).

In [5.2] werden daher für den “aktiven Sensor” der Begriff **signalbearbeitender Sensor** (signal conditioning sensor) und für den “passiven Sensor” der Begriff **rezeptiver Sensor** (receptive sensor) vorgeschlagen. Der bearbeitende Sensor muss ein Signal erst noch konditionieren, während der rezeptive Sensor das Signal lediglich aufnimmt und umwandelt.

Die wichtigsten physikalischen Messeffekte sind:

- Thermisch
- Optisch
- Ohmscher Widerstandseffekt
- Magnetisch
- Kapazitiv
- Piezowiderstandseffekt
- Piezoelektrisch
- Pyroelektrisch
- Chemisch

Tabelle 5-3: Nutzung physikalischer Effekte für Sensoraufgaben

Physikalische Größe	Messeffekte								
	Thermisch	Optisch	Ohm'scher Widerstandseffekt	Magnetisch	Kapazitiv	Piezowider-	Piezoelektrisch	Pyroelektrisch	Chemisch
Ort		X	X	X	X	X	X		
Kraft			X			X	X		
Druck			X			X	X		
Temperatur	X	X						X	
Licht	X	X					X	X	
Gas		X							X
Magnetfeld				X					

Tabelle 5-3 enthält eine Übersicht, welche Messeffekte zur Messung welcher Art von physikalischen Größen verwendet werden. Tabelle 5-4 enthält eine Auflistung von interessierenden Messgrößen, für die käufliche Sensoren auf dem Markt angeboten werden.

Tabelle 5-4: Sensoren für physikalische Größen

Mechanische Größen an Festkörpern	Abstand, Beschleunigung, Dehnung, Dichte Dicke, Drehmoment, Drehzahl, Druck, Durchmesser, Form, Geschwindigkeit, Gewicht, Kraft, Länge, Höhe, Härte, Masse, Orientierung, Spannung, Weg, Winkel, usw.
Mechanische Größen an Flüssigkeiten und Gasen	Dichte, Druck, Durchfluss, Füllstand, Strömungsgeschwindigkeit, Viskosität, Volumen, usw.
Thermische Größen	Temperatur, Wärmeleitung, Wärmestrahlung, usw.
Optische Strahlung	Farbe, Intensität, Polarisation, Reflexion, Wellenlänge, usw.
Akustische Größen	Absorption, Intensität, Schalldruck, Schallfrequenz, Schallgeschwindigkeit, usw.
Kernstrahlung	Ionisationsgrad, Strahlungsenergie, Strahlungsfluss
Chemische Größen	Feuchtigkeit, Konzentration, Molekül- und Ionensorte, Partikelform und -größe, pH-Wert, Reaktionsgeschwindigkeit, usw.
Magnetische und elektrische Größen	Dielektrizitätskonstante, Frequenz, Induktivität, Kapazität, Leistung, Phase, Strom, Spannung, Widerstand, usw.
Sonstige Größen	Anzahl, Pulsdauer, Zeit, usw.

Um sich einen Überblick über Messverfahren in Bezug auf den Sensor zu verschaffen, kann man einerseits eine Unterteilung nach den zu messenden physikalischen Größen und andererseits nach den physikalischen Messeffekten vornehmen. Eine weitere Möglichkeit besteht in der Unterteilung nach den verschiedenartigen Technologien bei der Herstellung. Hier sind im wesentlichen die aus diskreten Bauelementen aufgebauten

- klassischen Messwertnehmer

und solche in miniaturisierter Bauweise mittels

- Si-Technologie (angelehnt an elektronische, integrierte Halbleiterschaltungen)
- Dünnschichttechnologie
- Dickschichttechnologie
- Faseroptische Sensoren

zu unterscheiden.

Da bei mechatronischen Systemen vor allem Sensoren zur Erfassung der Bewegungsgrößen (Lage, Geschwindigkeit, Beschleunigung) und der Bewegungsursachen (Kraft, Drehmoment, Druck) von Bedeutung sind, werden in den folgenden Abschnitten vor allem Sensoren (klassische Messfühler und integrierbare Sensoren) zur Messung dieser Größen behandelt. Die Gliederung erfolgt so, dass in den nachstehenden Unterabschnitten 5.2.1 - 5.2.3 zuerst Messeffekte ohne spezielle

Herausstellung der Anwendbarkeit für klassische Messfühler oder integrierbare Sensoren behandelt werden. Danach werden im Abschnitt 5.3 Sensoren für die physikalischen Messgrößen mechatronischer Systeme in konventioneller Bauweise und in **Mikrosystemtechnik** behandelt werden.

5.2.1 Widerstandseffekte

5.2.1.1 Ohmsche Widerstandseffekte

Ohm'sche Messaufnehmer sind dadurch gekennzeichnet, dass ihr ohm'scher Widerstand durch die jeweilige physikalische Messgröße verändert wird. Ohm'sche Widerstandseffekte werden bei **Potentiometern** zur Messung von Wegen oder Winkeln und bei **Dehnungsmessstreifen** zur Messung von Dehnung, Kraft, Druck, Weg, Winkel und Torsion verwendet. Darüber hinaus wird der Effekt in **Widerstandsthermometern** zur Temperaturmessung, in **Fotowiderständen** für Lichtgrößen und in **Feldplatten** für magnetische Größen verwendet.

Für den ohm'schen Widerstand R eines elektrischen Leiters gilt folgende Abhängigkeit

$$R = \rho \frac{l}{A} ,$$

mit dem **spezifischen Widerstand** ρ , der Leiterlänge l und dem Leiterquerschnitt A .

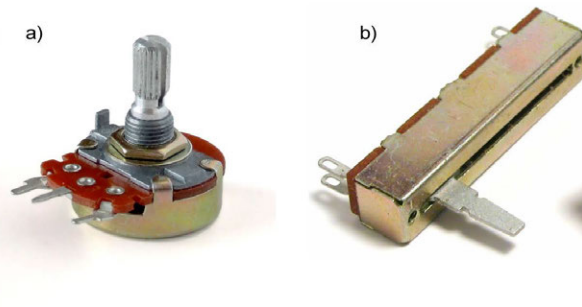


Bild 5-17: Potentiometer a) Dreh- b) Linear-

Bei einem Potentiometer (s. Kap. 5.1.2) wird eine Längenänderung als Messeffekt ausgenutzt, indem über einen bewegten Schleifkontakt die Widerstandsbahn (linear oder kreissegmentförmig) in der Länge l verändert wird. Bild 5-17 a) zeigt ein Dreh- und Bild 5-17 b) ein Schiebepotentiometer. Die Widerstandsänderung ist dem Verschiebeweg proportional.

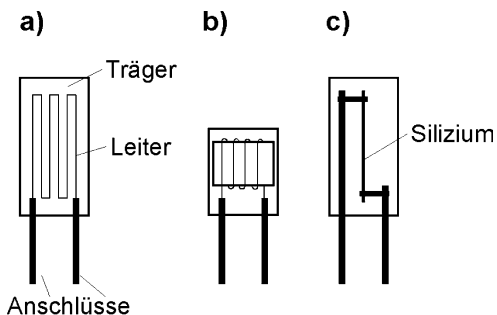


Bild 5-18: Bauarten von Dehnungsmessstreifen

Bei einem Dehnungsmessstreifen ist ein Draht- oder Halbleiterwiderstand entweder auf einen streifenförmigen Träger aus Papier, Kunststoff oder Aluminium (abhängig von der Einsatztemperatur) aufgeklebt, oder direkt auf das Siliziumsubstrat eines integrierten Schaltkreises aufdiffundiert.

Der klassische Dehnungsmess-

streifen (DMS) trägt eine mäanderförmige Wicklung aus Konstantandraht (Bild 5-18 a), um auf möglichst kleiner Fläche eine große Leiterlänge zu platzieren. Die Wicklung nach Bild 5-18 b) wird angewandt, wenn der DMS möglichst klein sein soll, um beispielsweise eine möglichst punktförmige Messung zu ermöglichen. Integrierte DMS-Wicklungen können von ihrer Ausdehnung her noch bedeutend kleiner sein. Zum Messen von mechanischen Spannungsfeldern und von Drehmomenten benutzt man Rosetten aus zwei oder drei DMS, die unter 90° oder 120° zueinander auf einem Träger angeordnet sind (Bild 5-19 b)). Ersetzt man das Konstantanwiderstandselement durch eine Schicht aus Silizium, so erhält man einen Halbleiter-DMS (Bild 5-18 c). Sein Vorteil ist eine höhere Empfindlichkeit, der Nachteil eine höhere Temperaturabhängigkeit. Die Widerstandsänderung des Siliziums beruht allerdings nicht auf dem normalen ohm'schen Widerstandseffekt, sondern auf dem noch zu behandelnden Piezo-Widerstandseffekt.

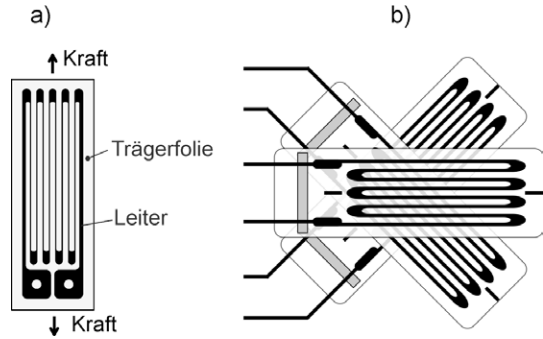


Bild 5-19: Dehnmessstreifen a) DMS für Messung in einer Richtung
b) Rosetten-DMS für mehrere Messrichtungen

Dehnt man nun einen DMS mit der Leiterlänge l in Längsrichtung, so nimmt dessen Widerstand R um ΔR zu, staucht man ihn, so nimmt sein Widerstand ab. Ist der Leiter senkrecht zur Dehnung mechanisch unbelastet, so beträgt die relative Änderung des Widerstandes

$$\frac{\Delta R}{R} = \frac{\Delta \rho}{\rho} + \frac{\Delta l}{l(1 + 2\nu)} .$$

Dabei ist ν die Querkontraktionszahl (Poisson-Zahl), die berücksichtigt, dass wegen der Volumenkonstanz mit der Längenänderung eine Querschnittsänderung einhergehen muss. Der zweite Term der Gleichung hängt nur von der Geometrie ab, der erste Term enthält den spezifischen Widerstand ρ , eine Stoffkenngröße, die sowohl von der Temperatur als auch von der Geometrie abhängt. Für den spezifischen Widerstand gilt:

$$\frac{1}{\rho} = q \cdot n \cdot \mu ,$$

wobei q die Elementarladung, n die auf das Volumen bezogene Ladungsträgerdichte und μ die Beweglichkeit der Ladungsträger sind. Im makroskopischen Bereich ändert sich die Anzahl der Ladungsträger nicht, so dass die Widerstandsänderung nur noch von der Beweglichkeit der Ladungsträger und der Länge des Leiters abhängt:

$$\frac{\Delta R}{R} = -\frac{\Delta \mu}{\mu} + \frac{2\Delta l}{l} .$$

Falls sich die Beweglichkeit nicht mit der Belastung ändert, findet man einen Wert K für das Verhältnis von relativer Widerstandsänderung und relativer Dehnung, der als K - oder **Gage - Faktor** bezeichnet wird:

$$K = \left(\frac{\Delta R}{R} \right) / \left(\frac{\Delta l}{l} \right) = 2 \quad .$$

Viele Metalle und Legierungen wie z. B. Konstantan verhalten sich nach dieser Gleichung. Es gibt jedoch auch Ausnahmen wie beispielsweise Platin-Iridium mit dem Wert 6,6. Für Silizium kann der Wert je nach Dotierung in einem weiten Bereich zwischen -150 und +200 liegen.

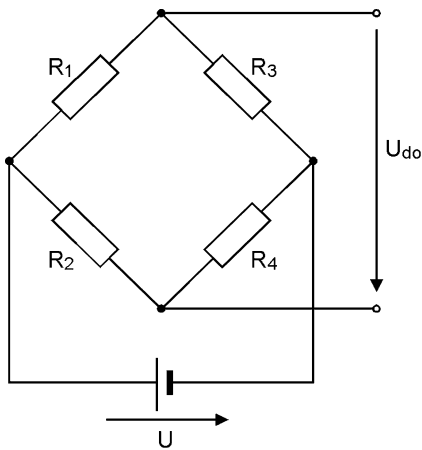


Bild 5-20: Wheatstone'sche Brücke

Normale relative Dehnungen liegen in der Größenordnung von 10^{-3} , so dass sich für typische Widerstandswerte von Draht-DMS zwischen 120 Ω und 600 Ω , Widerstandsänderungen von 0,12 Ω bis 0,6 Ω ergeben. Um so geringe Werte messen zu können, bedient man sich im allgemeinen einer Brückenschaltung, z. B. in Form der **Wheatstone'schen Brücke** (Bild 5-20), in der 1, 2 oder 4 der Widerstände DMS sein können. Sind alle vier Widerstandswerte der Brücke gleich, so ist die Brücke abgeglichen und die Brückenspannung $U_{D0} = 0$ (Diagonal-spannung). Dadurch wird der hohe Widerstandsgrundwert des DMS eliminiert, den man erhalten würde, wenn man die Widerstandsmessung durch eine Strommessung bei

bekannter Speisespannung U direkt über dem DMS-Widerstand vornehmen würde. In der Brückenschaltung kann man die Widerstandsmessung durch eine Spannungsmessung der Brückenspannung durchführen, die im unbelasteten Fall des DMS den Wert Null besitzt. Außerdem kann man je nach zu messendem Belastungsfall (Zug, Biegung) sogar zwei oder vier der Brücken-DMS mechanisch in Reihe schalten, wodurch sich der Messeffekt um den Faktor 2 oder 4 erhöht.

Die Brückenschaltung dient gleichzeitig der Temperaturkompensation, da der spezifische Widerstand eines Leiters temperaturabhängig ist. Dies ist ja auch beispielsweise der Messeffekt, den man in Widerstandsthermometern verwendet. Führt man wieder nur eine Strommessung an einem einzelnen DMS zur Widerstandsbestimmung aus, so kann man nicht unterscheiden, ob Widerstandsänderungen durch Dehnungen oder durch Temperaturveränderungen hervorgerufen werden. Ordnet man am Messort die beiden Widerstände einer Halbbrücke (z. B. R_1 und R_3) so an, dass ein DMS mechanisch belastet, der andere jedoch mechanisch unbelastet ist, so haben beide stets gleiche Temperatur und Widerstandsänderungen durch Temperaturänderungen kompensieren sich gerade. Die Kompensation von Temperaturgang und Grundwert des DMS gelingt natürlich um so besser, je genauer die Eigenschaften der unbelasteten DMS übereinstimmen. Diese Forderung lässt sich in idealer Weise mit Hilfe der Technik der integrierten Schaltungen verwirklichen.

5.2.1.2 Piezowiderstandseffekt

Die Eigenschaften eines Festkörpers hängen im allgemeinen vom Zustand seiner Dehnung ab. Wirkt eine mechanische Spannung auf einen Kristall, so verschieben sich die Atome gegeneinander. Die dabei auftretenden Änderungen der Gitterkonstanten bewirken eine Änderung der Struktur der Leitungs- und Valenzbänder. Fehlt den Elementarzellen der Kristalle ein Symmetriezentrum, treten also elektrische Dipole auf (polare Achsen), so ist bei geringer Leitfähigkeit des Stoffes Piezoelektrizität feststellbar (s. a. Kap. 1.4). Der piezoelektrische Effekt wirkt in der Regel nur dynamisch, weil äußere Ladungen immer rasch kompensiert werden. Sensoren, die ihn verwenden, sind rezeptive Sensoren.

Der **piezoresistive Effekt** wirkt dagegen auch statisch und kann bei bearbeitenden Sensoren ausgenützt werden. Beim piezoresistiven Effekt verändert sich, anders als beim normalen ohmschen Widerstandseffekt, der spezifische Widerstand der Materialien, solange sie einer Zug- oder Druckbelastung ausgesetzt sind. Er tritt auch in Materialien ohne polare Achsen auf und ist in Halbleitermaterialien wie z. B. Silizium gut ausgeprägt, bei Metallen jedoch sehr gering.

In vereinfachter skalarer Schreibweise ist der Zusammenhang zwischen relativer Widerstandsänderung und der anliegenden mechanischen Spannung σ wie folgt:

$$\frac{\Delta\rho}{\rho} = \pi \cdot \sigma \quad (5.1)$$

Dabei ist π der so genannte **piezoresistive Koeffizient**, der von der Kristallrichtung und den Messbedingungen abhängt. In Wirklichkeit ist die Gleichung (5.1) eine Tensorgleichung mit einem symmetrischen Tensor π , der im allgemeinen Fall 36 Komponenten besitzt. In einem kubischen Kristallgitter wie beispielsweise bei Silizium sind nur drei Komponenten unabhängig voneinander, nämlich π_{11} , π_{12} und π_{44} . Die Werte dieser Koeffizienten hängen vom Leitungstyp (p-dotiert, n-dotiert) und von der Höhe der Dotierung ab.

Beim Piezowiderstandseffekt lassen sich, aufgrund der möglichen mechanischen Belastungsfälle eines Körpers, drei verschiedene Effekte unterscheiden. Dies sind der longitudinale (Zugspannung in Richtung des betrachteten Stromflusses), transversale (Zugspannung quer zum betrachteten Stromfluss) und der Scher-Piezowiderstandseffekt (Scherbelastung quer zum betrachteten Stromfluss). Die entsprechenden Koeffizienten für die Zug-, Druck-Belastung π_L und π_T sind beide voneinander verschieden und lassen sich beispielsweise wie folgt berechnen:

$$\pi_L = -\frac{1}{3}(\pi_{11} + 2\pi_{12}) + \frac{2}{3}\pi_{14} \quad \pi_T = \pi_{12} + \frac{1}{6}(\pi_{11} - \pi_{12} - \pi_{44}) \quad .$$

Beide Koeffizienten sind somit stark von der Richtung der Belastung im Bezug auf die Orientierung des Kristalls abhängig. Die gezielte Orientierung der piezoresistiven Wandlerelemente in Abhängigkeit von der Kristallrichtung ist deshalb für die technische Anwendung von entscheidender Bedeutung.

In der Praxis wird die Piezoresistivität meistens für Elemente verwendet, die in Form von Widerstandsbahnen auf einem Verformungskörper angebracht werden. Als Verformungskörper kommen Biegebalken, insbesondere für Kraft- und Be-

schleunigungssensoren, und Biegeplatten in rechteckiger und runder Form zum Einsatz. Die Widerstände werden in Brückenschaltung in Bereichen maximaler mechanischer Spannungen aufgebracht. Dafür werden vier Widerstände mit möglichst gleichen Eigenschaften benötigt. Dies lässt sich in idealer Weise mit Hilfe mikroelektronischer Techniken realisieren, indem in einen Halbleiter bestimmter Dotierung vier Widerstände eindiffundiert werden.

5.2.2 Magnetische Effekte

5.2.2.1 Induktionsprinzip

Die Änderung der Induktivität einer Spule wird schon lange als Messeffekt für bearbeitende Sensoren zur Weg- und Winkelmessung eingesetzt. Bei den meisten wird eine Kombination aus einer Spule, einem Magneten oder Spulenkern und einem zu erfassenden Objekt verwendet. Das solchen Sensoren zugrundeliegende Wirkprinzip beruht auf dem **Induktionsgesetz**

$$U_{ind} = -N \frac{d\Phi}{dt} \quad , \quad (5.2)$$

mit U_{ind} als der in der Spule induzierten Spannung, $d\Phi/dt$ der magnetischen Flussänderung und N der Windungszahl der Spule. Als Messprinzip wird in der Regel die Änderung des magnetischen Flusses durch Dreh- oder Relativbewegungen benutzt. Um Möglichkeiten für die Flussänderung zu erkennen, kann man Gl. 5.2 wie folgt umformen:

$$U_{ind} = -N \frac{d(B \cdot A)}{dt} = -N \frac{d(\mu_r \cdot \mu_0 \cdot N \cdot I \cdot A/l)}{dt} \quad .$$

Darin sind B die magnetische Induktion, μ_r und μ_0 die Permeabilität des Spulenkerns und des Vakuums, I der Strom durch die Spulenwicklung, A die Querschnittsfläche der Spule und l die Spulenlänge. Eine Vielzahl der möglichen Einflussgrößen auf die Flussänderung ist auch in der Spuleninduktivität

$$L = \mu_r \cdot \mu_0 \cdot N^2 \cdot A/l$$

enthalten.

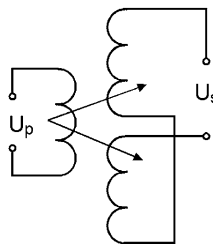
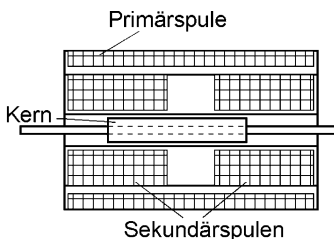


Bild 5-21: Induktiver Sensor in Form eines Differentialtransformators im Querschnitt und als Schaltbild

Von diesen lässt sich am einfachsten die relative Permeabilität μ ändern, indem man den Spulenkern in der Spule verschiebt. Ein nach diesem Prinzip arbeitender Sensor, mit dem man auch sehr kleine Verschiebungen erfassen kann, ist der in Bild 5-21 dargestellte **Differentialtransformator**. Beim Eindringen des Kerns in den Hohlraum

einer Spule wächst die Induktivität, jedoch ist die Wegabhängigkeit nicht linear. Durch eine Gegentaktnordnung wie in Bild 5-21 lässt sich die Kennlinie aber linearisieren. In dem rotationssymmetrischen Gehäuse ist eine Primärwicklung untergebracht, die von einem Oszillator mit einer Wechselspannung von einigen kHz gespeist wird. In den beiden gegentaktmäßig in Serie geschalteten Sekundärspulen wird, je nach Stellung des wegführenden Kerns, eine Spannung induziert. Dabei hängt die Kopplung von Primär- und Sekundärspule von der Stellung des Kerns ab. Nach Demodulation und Verstärkung liegt am Ausgang ein analoges Differenzsignal der beiden Sekundärspulen vor.

Das gleiche Prinzip kann als Sensor auch mit beweglichen Spulen genutzt werden, deren Kopplung durch Verschieben oder Verdrehen zueinander geändert werden kann. Die sich dadurch verändernde Amplitude der Sekundärspannung stellt den Messeffekt für eine Wegmessung dar.

Auch bei feststehendem Spulenkern kann durch Einbringen eines ferromagnetischen Objektes in das Spulenfeld das gleiche Prinzip für einen Annäherungssensor genutzt werden.

5.2.2.2 Galvanomagnetische Effekte

Galvanomagnetische Effekte sind solche, die es erlauben Magnetfelder verschiedener Stärke nachzuweisen. Durch Anwendung solcher Effekte in Sensoren, von denen der **Hall-Effekt** der bekannteste ist, kann man das Vorhandensein eines Magnetfeldes in ein elektrisches Signal transformieren. Weitere Effekte beruhen auf dem **Magnetowiderstand** und der **Magnetokonzentration** [5.1]. Als Maßeinheit für die Stärke eines Magnetfeldes wird die magnetische Induktion benutzt, deren Maßeinheit das Tesla ist ($1\text{T} = 1\text{ Vs/m}^2$). Der Messbereich von Magnetfeldern ist sehr groß. So liegen die Streufelder der magnetischen Domänen von Aufzeichnungsmedien im Bereich 10 mT bis 10 mT, während Permanentmagneten Felder von 5 mT bis 100 mT aufweisen.

Effekte wie der Hall-Effekt, entdeckt von dem amerikanischen Physiker E. Hall (1879), beruhen auf der Wirkung der Lorentzkraft (H.A. Lorentz, niederl. Physiker 1853-1928) auf bewegte Ladungen in einem Magnetfeld:

$$F_L = -e \cdot \vec{v} \cdot \vec{B} \quad .$$

Darin sind e die Elektronenladung, $\vec{v}$ die Ladungsträgersgeschwindigkeit und $\vec{B}$ die magnetische Induktion. Der Zusammenhang zwischen der magnetischen Induktion und der magnetischen Feldstärke $\vec{H}$ ist durch folgende Beziehung gegeben:

$$\vec{B} = \mu_r \cdot \mu_0 \cdot \vec{H}$$

Die Stärke der Lorentzkraft hängt demnach von der Permeabilität des verwendeten Materials ab. Bei ferromagnetischen Stoffen ist ihr Wert $\mu_r \gg 1$. Man benutzt dünne Metallfilme aus NiFe mit entsprechend hoher Empfindlichkeit. Dia- oder paramagnetische Stoffe wie etwa alle Halbleiter haben ein $\mu_r \approx 1$ und sind daher relativ unempfindlich. Diese Materialien (Si, GaAs, InSb) werden jedoch vielfach für Elemente verwendet, die den Hall-Effekt ausnutzen.

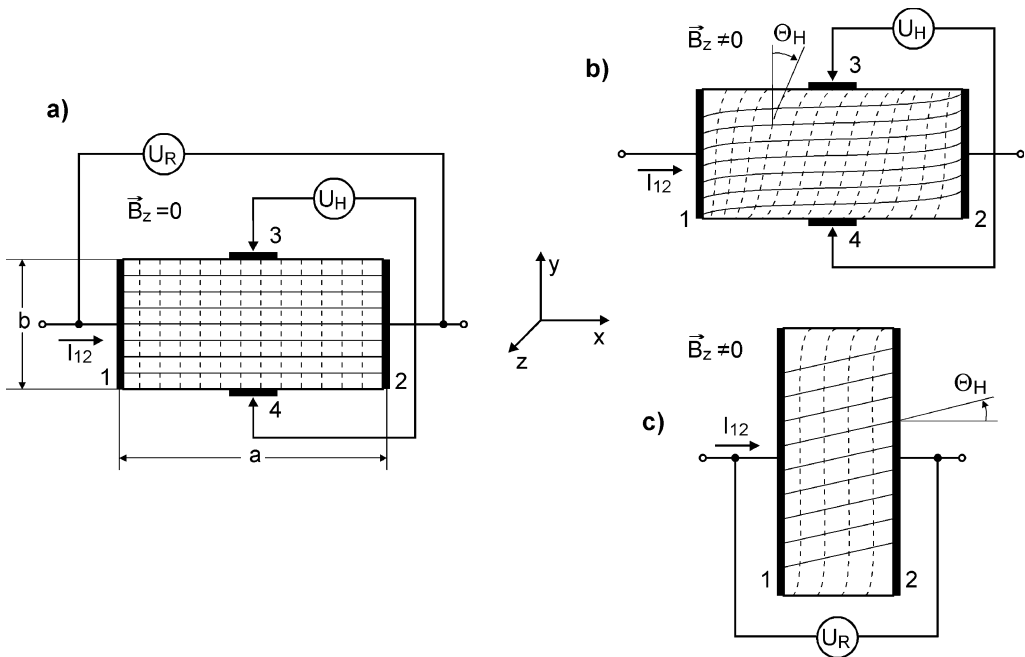


Bild 5-22: Einfluss eines Magnetfeldes auf die Verteilung von Strom- und Äquipotentiallinien in der rechteckigen Halbleiterschicht eines Hall-Generators

- a) ohne Magnetfeld
- b) Messung der Hall-Spannung am Hall-Generator
- c) Messung der Spannung am Gesamtwiderstand einer Feldplatte

Hall-Generatoren sind Sensorelemente, die den Hall-Effekt zur Messung der magnetischen Induktion ausnutzen. Sie bestehen im allgemeinen aus einem sehr dünnen Streifen eines Halbleitermaterials, an dessen gegenüberliegenden Seiten jeweils zwei Elektroden angebracht sind. Lässt man in einem feldfreien Raum in Längsrichtung des Halbleiters einen Strom I_{12} fließen (Bild 5-22 a)), so misst man über den Elektroden 1 und 2 die Spannung U_R , aus der sich der Gesamtwiderstand bestimmen lässt. Da in einem homogenen Halbleiter die Potentiallinien des elektrischen Feldes (gestrichelte Linien) parallel sind und senkrecht zu den Stromlinien (durchgezogene Linien) verlaufen, liegen die Elektroden 3 und 4 auf gleichem Potential und die Hall-Spannung U_H ist gleich Null. Wirkt nun ein Magnetfeld senkrecht zum Halbleitermaterial in Richtung $\vec{z}$, so werden aufgrund der Lorentzkräfte die Ladungsträger aus der ursprünglichen Stromflussrichtung in Richtung $\vec{y}$ abgelenkt (Bild 5-22 b)), wodurch sich nun die Elektroden 3 und 4 auf zwei verschiedenen **Äquipotentiallinien** befinden. Die Äquipotentiallinien werden um den Winkel Θ_H gedreht, im Bereich der Hall-Elektroden bleibt die ursprüngliche Stromrichtung erhalten. Dies bedeutet nichts anderes, als dass nun zwischen den Hall-Elektroden ein Potentialunterschied auftritt, d.h. die Hall-Spannung ist von Null verschieden. Um eine möglichst hohe Hall-Spannung zu bekommen, wählt man die Länge a gegenüber der Breite b möglichst groß.

In der Nähe der Steuerelektroden 1 und 2 werden die Stromlinien um den Hall-Winkel Θ_H gedreht, da diese metallisch leitenden Elektroden Äquipotentiallinien dar-

stellen. Dadurch wird eine Verlängerung der Strombahnen und somit eine Erhöhung des Widerstandes hervorgerufen. Aufgrund dieses magnetoresistiven Effektes steigt die Spannung U_R an.

Will man gerade diesen magnetischen Widerstandseffekt ausnutzen, so muss man die Länge a sehr viel kleiner als die Breite b wählen (Bild 5-22 c), weil dann die relative Widerstandsänderung besonders groß wird. Einen solchen Sensor bezeichnet man auch als **Feldplatte**. Um einen Grundwiderstand $R(0)$ solcher Feldplatten von einigen $100\ \Omega$ zu erhalten, werden in einer Feldplatte mehrere Streifen mit $a \ll b$ in Reihe geschaltet. Die Hall-Spannung hat folgenden Wert:

$$U_H = \frac{R_H}{d} I_{12} \cdot B \quad ,$$

und ist somit proportional zum Strom I_{12} , der magnetischen Induktion B und dem Quotienten aus Hall-Koeffizient R_H und der Dicke d . Für eine hohe Empfindlichkeit des Hall-Elementes sind daher dünne Schichten des Halbleitermaterials mit einem hohen spezifischen Widerstand erforderlich. Der Widerstand $R(B)$ einer Feldplatte zeigt für kleine magnetische Induktionen eine annähernd quadratische Abhängigkeit von B und ist unabhängig vom Vorzeichen von B . Die prinzipiellen Verläufe der Kennlinien eines Hall-Generators und einer Feldplatte sind in Bild 5-23 dargestellt.

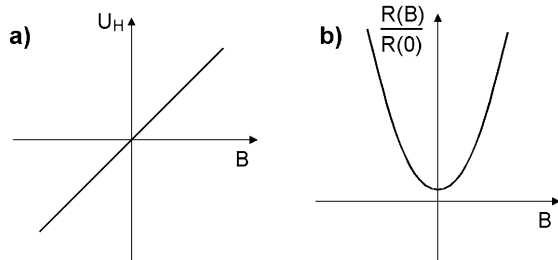


Bild 5-23: Kennlinie eines a) Hall-Generators und einer b) Feldplatte

5.2.2.3 Magnetoelastische Effekte

Magnetoelastische Materialien wie z. B. $Fe_{50}Co_{50}$ ändern ihre Permeabilität unter Druck oder Zug (Villari-Effekt) oder bei Torsion (Matteucci-Effekt) und dienen als magnetoelastische Sensoren für Kraft und Drehmoment. Umgekehrt führt die Magnetisierung gewisser ferromagnetischer Materialien z. B. $Ni_{81}Fe_{19}$ zu einer Änderung der Abmessungen (Magnetostriktion oder Joule-Effekt) und des Elastizitätsmoduls (ΔE -Effekt).

Bild 5-24 zeigt eine Positionsmesseinrichtung, die die **Magnetostriktion** in einem dünnen Rohr dieses Materials ausnutzt. In diesem Rohr befindet sich ein elektrischer Leiter. Der Strom eines Sendepulses $i(t)$ erzeugt ein zirkulares Magnetfeld der Stärke $B(i)$ in dem Rohr. An der Stelle der zu messenden Position befindet sich ein in Längsrichtung des Rohres verschiebbarer Permanentmagnet, dessen Magnetfeldlinien senkrecht zu dem zirkularen Magnetfeld des Sendepulses stehen. Die Überlagerung beider Magnetfelder regt den magnetostriktiven Vorgang an, und es entsteht ein mechanischer Echoimpuls im Ultraschallbereich, der sich mit etwa $v_s = 2850\text{ m/s}$ in beiden Richtungen entlang des Rohres ausbreitet. Zur Unterdrückung des Rückwand-Echoimpulses ist eine

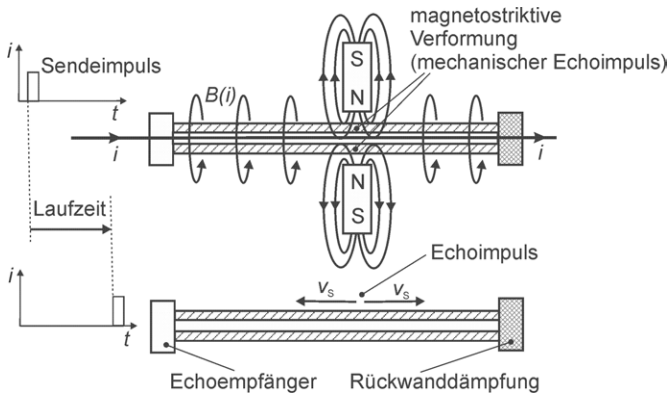


Bild 5-24: Magnetostriktive Positionsmessung

Dämpfung eingebaut. Am Ende des Rohres wird dieser Echoimpuls von einem Ultraschallsensor aufgenommen und auf Grund der Laufzeit zwischen Sendepuls und mechanischem Echoimpuls die Position des Permanentmagneten relativ zum Empfangssensor bestimmt.

5.2.3 Kapazitive Effekte

Kapazitive Effekte werden vor allem zur Erfassung von Abständen, Verschiebungen oder Verdrehungen verwendet. Die Kapazität C eines Kondensators verknüpft die Ladung Q und die Spannung U eines Kondensators wie folgt miteinander:

$$Q = C \cdot U \quad .$$

Die einfachste Form eines Kondensators ist der **Plattenkondensator** mit kreisförmigen Elektroden. Die Kapazität eines solchen Kondensators ist von der Fläche A der Elektroden, deren Abstand d , sowie den Dielektrizitätskonstanten ϵ_r des im Feld befindlichen Dielektrikums und ϵ_0 des Vakuums abhängig:

$$C = \epsilon_r \cdot \epsilon_0 \frac{A}{d} \quad .$$

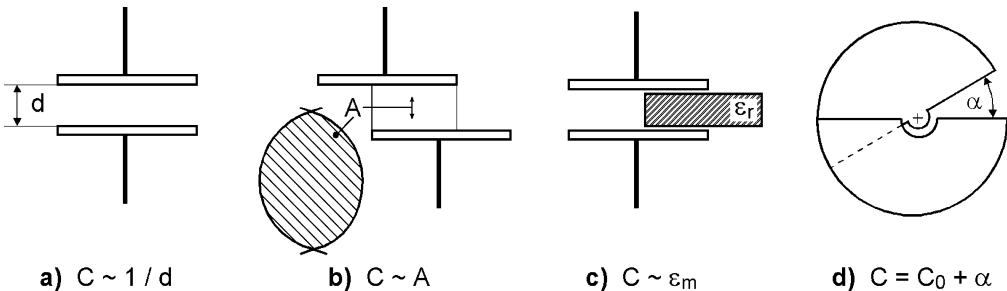


Bild 5-25: Messeffekte durch Kapazitätsänderung

Die in dieser Gleichung vorkommenden änderbaren Größen werden alle für die genannten Messaufgaben genutzt. Die erste Möglichkeit (Bild 5-25 a)) besteht in einer Abstandsänderung der Platten zueinander, die eine hyperbolische Änderung der Kapazität $C \sim 1/d$ in Abhängigkeit von d ergibt. Bei einer parallelen Verschiebung der Platten relativ zueinander (Bild 5-25 b)) ist eine direkte Proportionalität $C \sim A$ zur Flächenänderung gegeben, mit der Möglichkeit, gewünschte Kennlinienformen durch Wahl der Konturen zu beeinflussen. Drehkondensatoren erlauben die Umfor-

mung von Drehwinkeln in Kapazitätswerte (Bild 5-25 d)) mit einer durch die Kontur bedingten Kennlinie. Auch die Verschiebung eines festen Dielektrikums ist zur Messgrößenerfassung geeignet (Bild 5-25 c)). Aus dem Raumanteil ohne Dielektrikum und dem Teil mit der Dielektrizitätskonstanten ϵ_r ergibt sich eine mittlere Dielektrizitätskonstante ϵ_m . Diese hängt linear von der Verschiebung des Dielektrikums ab.

Die Hauptvorteile kapazitiver Sensoren liegen im einfachen Aufbau, der nahezu exakten Berechenbarkeit und der sehr guten Auflösung. Sie können bei erhöhten Temperaturen eingesetzt werden und liefern ein frequenzanaloges Ausgangssignal. Dem stehen Nachteile durch die kleinen Messsignale (Störempfindlichkeit) und durch den Einfluss von Kabelkapazitäten gegenüber.

5.2.4 Piezo- und Pyroelektrische Effekte

Piezelektrische Materialien zeichnen sich dadurch aus, dass unter mechanischer Druck- oder Zugbelastung an ihren Oberflächen elektrische Ladungen auftreten, die über Elektroden abgegriffen werden können. Die Ursache dieses Effektes wurde bereits in Kapitel 1 besprochen. Bei pyroelektrischen Materialien geschieht dies auch bei einer Temperaturänderung des Stoffes. Es gibt Einkristalle (α -Quarz, Lithiumniobat LiNbO_3), Keramik (Bariumtitanat BaTiO_3 , Bleizirkonattitanat $\text{PbZr}_{x-1}\text{Ti}_x\text{O}_3$, Bleimetaniobat PbNb_2O_6) und Polymere (Polyvinylidendifluorid PVDF) mit derartigen Eigenschaften. Halbleiter zeigen keine oder nur sehr geringe Neigung zur Piezoelektrizität.

Durch Verwendung geeigneter Materialien sind unter Nutzung des piezo- und pyroelektrischen Effektes rezeptive Sensoren mit einem weiten Linearitätsbereich herstellbar, deren Empfindlichkeit nur durch das mechanische und thermische Hintergrundrauschen begrenzt ist. Demgegenüber weisen piezoresistive Siliziumsensoren für ähnliche Aufgaben wie piezoelektrische Sensoren ein Rekombinationsrauschen auf, das um ein Mehrfaches über dem thermischen Rauschen liegt.

Besonders die Piezokeramiken und -polymere bieten wegen ihrer einfachen Formbarkeit eine große Gestaltungsfreiheit hinsichtlich Größe und Form des aktiven Sensorvolumens. Mechanische Anpassungen, wie sie im Bereich der Anwendung des piezoelektrischen Effektes für die Erzeugung von Ultraschall benötigt werden, sind möglich. Neben der Ultraschall-Prüf- und Messtechnik werden piezoelektrische Sensoren als Drucksensoren, Kraftsensoren, Beschleunigungsaufnehmer, Mikrofone und vieles andere mehr, eingesetzt. Da der piezoelektrische Effekt reversibel ist (inverser oder reziproker piezoelektrischer Effekt), werden Piezoelemente auch als Aktoren für kleine Wege und als Ultraschallsender eingesetzt. Bei der Umkehrung des Effektes führt das Material in einem elektrischen Feld Längenänderungen aus. Ist dies ein Wechselfeld, so wird das Material in Schwingung versetzt.

Da Piezo- und Pyroeffekt vielfach gemeinsam auftreten, muss man bei Piezosensoren im Bedarfsfall Kompensationsmaßnahmen gegen den störenden Pyroeffekt vorsehen und umgekehrt.

Ein Nachteil piezo- und pyroelektrischer Wandler ist, dass die von ihnen einmal er-

zeugte Ladung nicht beliebig lange erhalten bleibt. Sie fließt über den Isolationswiderstand der Sensor- und Verstärkerschaltung ab, auch wenn man diesen sorgfältig so hochohmig wie möglich auslegt. Statische Messungen hoher Genauigkeit über mehrere Minuten hinweg sind nur durch Piezomaterialien mit extrem hohem spezifischen Widerstand wie beispielsweise synthetischem α -Quarz mit $\rho \geq 10^{15} \Omega \cdot \text{m}$ möglich. Keramiken und Polymere liegen in der Zeitkonstante deutlich darunter ($\rho \geq 10^{12} \Omega \cdot \text{m}$). Um quasistatische Messungen durchführen zu können, muss man Integrations- oder Ladungsverstärker verwenden, mit denen nicht die Spannung des Piezoelementes, sondern die Ladung direkt gemessen wird. Wichtig dabei ist auch die Verwendung von speziellen Messkabeln mit hohem Isolationswiderstand und geringen Störspannungen.

Aus diesem Grund sind piezo- und pyroelektrische Sensoren am besten bei schnellen dynamischen oder transienten Vorgängen mit Frequenzen vom Hz- bis in den MHz-Bereich geeignet.

5.2.5 Optische Effekte

Sensoren auf Halbleiterbasis, die optische Effekte ausnutzen, haben in der Mess- und Automatisierungstechnik große Bedeutung. Dabei werden optische Sensoren nur in geringem Umfang zur Messung von Licht an sich eingesetzt, sondern in der Regel in Systemen zur Positions- oder Winkelmessung und in Form großer Arrays (Matrixanordnungen) in Kameras zur Erkennung und Lagebestimmung komplexer geometrischer Objekte. Aufgrund der Einfachheit und des niedrigen Preises werden vor allem optische Halbleitersensoren aus Silizium verwendet. Der Einsatz optischer Elemente in den oben genannten Sensorsystemen erfolgt, weil trotz alternativer Möglichkeiten folgende Vorteile gegenüber anderen Systemen bestehen:

- kontaktlose Messung
- Messung an schwer zugänglichen Stellen oder in aggressiver Umgebung, eventuell unter Verwendung von Lichtleitfasern
- hohe Genauigkeit beispielsweise bei interferometrischer Anordnung
- weitgehende Immunität optischer Signale gegenüber äußeren elektromagnetischen Feldern
- einfache Potentialtrennung
- Vermeidung elektrischer Funken (Explosionsschutz).

Optische Halbleitersensoren sprechen auf die Wechselwirkung elektromagnetischer Strahlung (sichtbares Licht, Infrarot, Ultraviolett) mit dem Halbleitermaterial an, in dessen Volumen die Strahlung durch Ionisation freie Ladungsträger erzeugt. Dazu ist eine bestimmte Energie notwendig, die im Halbleiterkristall absorbiert wird.

Trifft Licht auf einen Si-Halbleiterkristall, so werden durch die Energie der auftreffenden Photonen Elektronen aus dem Valenzband ins Leitungsband gehoben; es entstehen freie Elektronen und Löcher, also positive und negative Ladungsträger [5.5]. Fällt das Licht auf den PN-Übergang einer **Fotodiode** (Bild 5-26), so werden

die Ladungsträger durch eine Raumladungszone, die sog. Sperrschicht, voneinander getrennt. Dabei wandern die Elektronen in das N-Gebiet und die Löcher in das P-Gebiet ab. Es entsteht eine elektrische Potentialdifferenz, die annähernd logarithmisch mit der Bestrahlungsintensität ansteigt und sich schließlich dem Sättigungswert von etwa 0,6 Volt nähert. Die Spannung ist unabhängig von der Diodenfläche. Schließt man die Diode kurz, dann fließt ein Strom der proportional zur Bestrahlungsintensität und zur bestrahlten Fläche ist. Dieser Effekt lässt sich praktisch zur Messung von Lichtintensitäten und auch zur Erzeugung elektrischer Energie nutzen. Im ersten Fall spricht man von einem Fotoelement, im zweiten Fall von einer **Solarzelle**. Physikalisch gesehen ist beides das gleiche, nur die praktische Ausführung ist unterschiedlich und auf den jeweiligen Verwendungszweck hin optimiert.

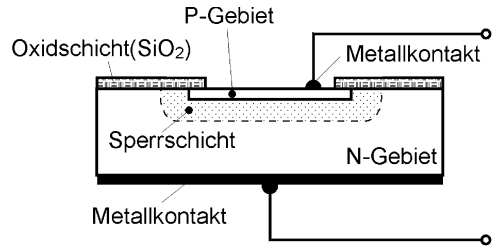


Bild 5-26: Aufbau einer einfachen PN-Fotodiode

Wird an die Diode eine Vorspannung in Sperrrichtung gelegt, dann fließt durch sie, bis auf einen minimalen Leckstrom, in der Dunkelheit kein Strom. Fällt Licht in die Sperrschicht, dann setzt ein Stromfluss (Sperrstrom) ein, der über 6 - 8 Größenordnungen hinweg proportional zur Beleuchtungsstärke ist. Die Diode wirkt als lichtabhängiger Widerstand. Im Fall einer derartigen Beschaltung (Bild 5-27 a) spricht man von einer Fotodiode. Im Normalbetrieb in Bild 5-27 b), in dem die Diode als lichtproportionale Stromquelle wirkt, bezeichnet man die Diode als **Fotoelement**. Diese beiden Betriebsarten sind immer möglich.

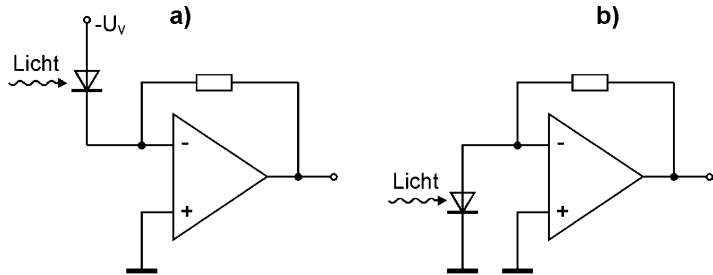


Bild 5-27: Verschiedene Betriebsarten einer lichtempfindlichen Diode
a) Betrieb als Fotodiode mit Vorspannung
b) Betrieb als Fotoelement

Im Fall einer derartigen Beschaltung (Bild 5-27 a) spricht man von einer Fotodiode. Im Normalbetrieb in Bild 5-27 b), in dem die Diode als lichtproportionale Stromquelle wirkt, bezeichnet man die Diode als **Fotoelement**. Diese beiden Betriebsarten sind immer möglich.

Beim Betrieb mit Vorspannung fließt durch die Diode immer ein geringer Leckstrom, der mit wachsender Spannung und wachsender Temperatur ansteigt. Dieser Leckstrom verursacht ein Rauschen und setzt eine untere Grenze für die erfassbare Lichtintensität. Beim Betrieb als Fotoelement gibt es keinen Leckstrom, weshalb das Rauschen niedriger und die Empfindlichkeit größer ist. Mit Vorspannung wird die Raumladungszone dicker, wodurch sich die Kapazität der Diode verringert. Da diese Kapazität zusammen mit den inneren Widerständen als Verzögerungsglied wirkt, nehmen dadurch die Anstiegs- und Abfallzeiten für sprunghafte Signaländerungen ab. Deshalb lassen sich in dieser Betriebsart höhere Frequenzen erfassen, während in der Betriebsart als Fotoelement die Reaktion auf dynamische Signale eher träge ist.

Auf einen Halbleiterkristall auffallendes Licht wird bereits nach kurzer Eindringtiefe

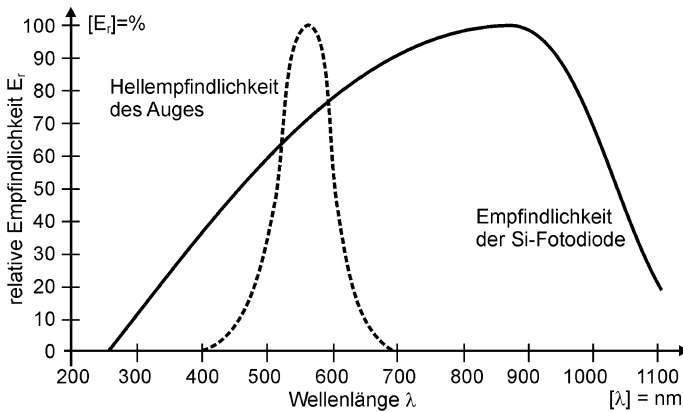


Bild 5-28: Relative Empfindlichkeit einer Si-Fotodiode und Hellempfindlichkeit des menschlichen Auges

absorbiert. Um bei Fotodioden eine möglichst hohe Empfindlichkeit zu erhalten, legt man die Sperrschicht so dicht wie möglich unter die Oberfläche. Der Chip besteht aus N-Silizium, in den auf der Oberseite eine dünne Schicht P-Silizium eindiffundiert wird.

Die Eindringtiefe elektromagnetischer Wellen in Halbleiterkristalle nimmt zu kürzeren Wellenlängen hin stark ab. Im na-

hen Infrarotbereich (Maximum bei 850 nm Wellenlänge), wo die Empfindlichkeit von Silizium-Fotodioden am größten ist (Bild 5-28), fällt die Intensität erst nach einigen 10 μm auf den e-ten Teil ab. Die Sperrschicht ist jedoch weit dünner, was zur Folge hat, dass der größte Teil der Lichtenergie erst tief im Kristall Ladungsträger erzeugt, die dann nicht genutzt werden können. Um die Empfindlichkeit zu erhöhen, legt man daher bei vielen Fotodioden zwischen P-Schicht und N-Schicht noch eine I-Schicht (engl. intrinsic = selbstleitend) aus hochreinem, undotiertem Silizium. Diese so genannten PIN-Dioden sind den normalen PN-Dioden an Empfindlichkeit weit überlegen und eignen sich wegen ihrer kleineren Sperrschichtkapazitäten auch für sehr hohe Signalfrequenzen (einige Gigahertz).

Die einfache Fotodiode lässt sich durch Integration einer zweiten Diode zum bipolaren **Fototransistor** erweitern. Dadurch ist eine Verstärkung des Fotostroms um den Faktor 100 bis 1000 möglich, d. h. die Empfindlichkeit ist deutlich verbessert. Allerdings sind aufgrund der großen Kollektor-Basis-Kapazität von Fototransistoren deren Anstiegs- und Abfallzeiten mit 5 bis 10 μs sehr viel höher als bei Fotodioden, wodurch ihre Grenzfrequenz bedeutend niedriger liegt.

Wie bei vielen anderen physikalischen Effekten lässt sich der bei der Fotodiode beschriebene Vorgang auch umkehren, so dass man Halbleiterdioden herstellen kann, die bei Anlegen einer in Durchlassrichtung gepolten Spannung Strahlung in einem engen Frequenzbereich aussenden. Liegt die Strahlung im sichtbaren Bereich, so spricht man von **Leuchtdioden** (LED: **L**ight **E**mitting **D**iode). Spezielle Halbleitermaterialien (AlGaAs/GaAs) und ein bestimmter geometrischer Aufbau (optischer Resonator) ermöglichen sogar die Erzeugung von Laser-Strahlung in der Regel im Infrarot-Bereich. Solche als Injektions-Laser bezeichneten Laser-Dioden werden zur optischen Nachrichtenübertragung, Abtastung von CD's, aber auch als Sendedioden für Entfernungsmessungen mit einem **Laserradar** (Lidar) benutzt.

In der Automatisierungstechnik werden häufig Kombinationen aus LED's als **Lichtschranke** oder als **Lichttaster** eingesetzt. Bild 5-29 zeigt eine solche Lichtschranke in Form einer „Gabellichtschranke“. In den beiden Armen sind gegenüberlie-

gend jeweils eine Leuchtdiode und eine Fotodiode eingebaut. Der von der LED ausgesendete Lichtstrahl wird beim Nichtvorhandensein eines Objektes zwischen den Gabelarmen von der Fotodiode empfangen und signalisiert damit den Zustand „Kein Objekt“. Tritt ein Objekt zwischen die Gabelarme, so wird der Lichtstrahl unterbrochen und die Fotodiode signalisiert aufgrund des fehlenden Lichtes „Objekt vorhanden“. Das ausgesendete Licht der LED liegt meist im nicht sichtbaren Infrarotbereich, damit die Empfangsdiode, die ebenfalls im Infrarotbereich empfindlich ist, nicht durch Streulicht der Umgebung gestört wird.

Bild 5-30 zeigt einen **Reflex-Lichttaster**. Hier sind Sender und Empfänger unter leichter Schrägstellung so zueinander angeordnet, dass bei Vorhandensein eines reflektierenden Objektes im Schaltabstand des Tasters der Lichtstrahl zum Empfänger reflektiert wird und dadurch das Objekt detektiert wird.

Fotowiderstände sind lichtempfindliche Halbleiterbauelemente aus Cadmium- oder Bleisulfid (CdS , PbS), die keine Sperrschicht enthalten. Bei Dunkelheit ist die Zahl der freien Ladungsträger sehr gering, so dass sich ein sehr hoher Widerstand im $\text{M}\Omega$ -Bereich ergibt. Licht setzt Ladungsträger frei, die einen Stromfluss ermöglichen und der Widerstand sinkt um 3-4 Größenordnungen. Da der Strom proportional zur angelegten Spannung ist, verhält sich der Fotowiderstand wie ein ohm'scher Widerstand. Es ist daher auch ein Betrieb mit Wechselspannung möglich.

Da der Fotowiderstand eine sehr niedrige Grenzfrequenz von einigen Hz besitzt und seine elektrischen Eigenschaften nicht zeitlich konstant sind, hat er heute gegenüber Fotodioden und -transistoren an Bedeutung verloren.

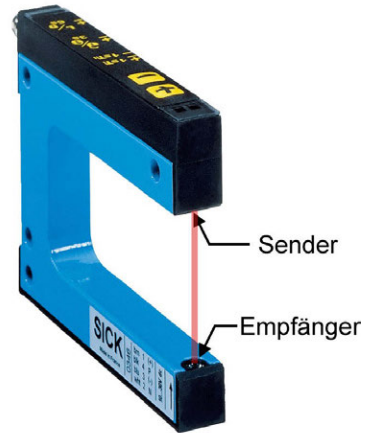


Bild 5-29: Gabellichtschranke (Sick)



Bild 5-30: Reflex-Lichttaster (Sick)

5.3 Sensoren für mechatronische Systeme

In mechatronischen Systemen, deren Aufgabe häufig in der gezielten Gestaltung von Bewegungsvorgängen besteht, sind Bewegungssensoren von besonderer Bedeutung. Daher werden in diesem Unterkapitel konkrete Sensorsysteme besprochen, die die im vorigen Unterkapitel beschriebenen Messeffekte verwenden.

5.3.1 Bewegungssensoren

Bewegungssensoren erfassen Bewegungsgrößen kinematischer Vorgänge. Diese Größen sind:

- Verschiebung (einschließlich Lage, Nähe, Größe)
- Geschwindigkeit bzw. Drehzahl
- Beschleunigung
- Ruck (Beschleunigungssprung)

Jeder der in der obigen Reihenfolge aufgeführten Größen ist die zeitliche Ableitung der vorherigen. Theoretisch wäre es daher zur Erfassung des Bewegungszustandes eines Systems möglich, nur eine dieser Variablen zu messen und die anderen durch Differentiation bzw. Integration daraus abzuleiten. In der Praxis lassen jedoch viele Faktoren dieses Vorgehen nicht zu. Dies kann seine Ursache in der Art des Signals (stationär, transient) oder seiner Frequenzanteile, der auftretenden Störungen und der aktuellen Verarbeitungsmöglichkeiten haben. So wird, wie in Kapitel 7.3 dargestellt, etwa bei einem Lageregelkreis für eine Bewegungsachse, sowohl die Position mit einem Wegmesssystem, als auch die Drehzahl mit einem Tachogenerator oder anderen Geschwindigkeitssensor und die Beschleunigung über die proportionale Größe "Ankerstrom" des Antriebsmotors bestimmt. Einige häufig verwendete Sensoren (Messwertaufnehmer) für Bewegungsgrößen werden in den nächsten Unterabschnitten ohne Anspruch auf Vollständigkeit behandelt.

5.3.1.1 Positionssensoren

Bei der Positionsmessung kann man im wesentlichen zwischen der **Abstands-** und **Lagemessung** unterscheiden. Die Abstandsmessung ist in der Regel eine Messung geringer Distanzen und die Lagemessung die Erfassung der Position eines bewegten Objektes innerhalb eines größeren Bewegungsbereiches. Da die Wegauflösung $\Delta l / l$ je nach Messsystem im Bereich 10^{-3} bis 10^{-5} liegt, erreicht man natürlich die höchsten Messgenauigkeiten bei der Messung kleiner Wege.

Bestimmte Messeffekte, wie die Beeinflussung der Induktivität einer Spule (s. 5.2.2.1), oder die Kapazität eines Kondensators (s. 5.2.3), reichen auch nur über relativ geringe Distanzen, liefern aber eine hohe örtliche Auflösung. Einige Möglichkeiten der Abstandsmessung zu ferromagnetischen Objekten mit induktiven Sensoren zeigt Bild 5-31.

Solche induktiven Näherungssensoren enthalten die in Bild 5-32 dargestellten Funktionselemente. Der Messeffekt beruht auf dem Induktionsgesetz. Der Sensor erzeugt mit Hilfe eines Oszillators, der einen LC-Schwingkreis enthält, ein elektromagnetisches hochfrequentes Feld. Die Spule des Schwingkreises befindet sich an der Frontfläche des Sensors, so dass das Feld aus der aktiven Fläche austritt. Dieses ruft in einem vorbeigeführten, elektrisch leitenden Material Wirbelströme hervor. Dadurch verändert sich die Amplitude des Schwingkreises. Das Signal wird demoduliert und einem Komparator zugeführt, der bei einer bestimmten Verände-

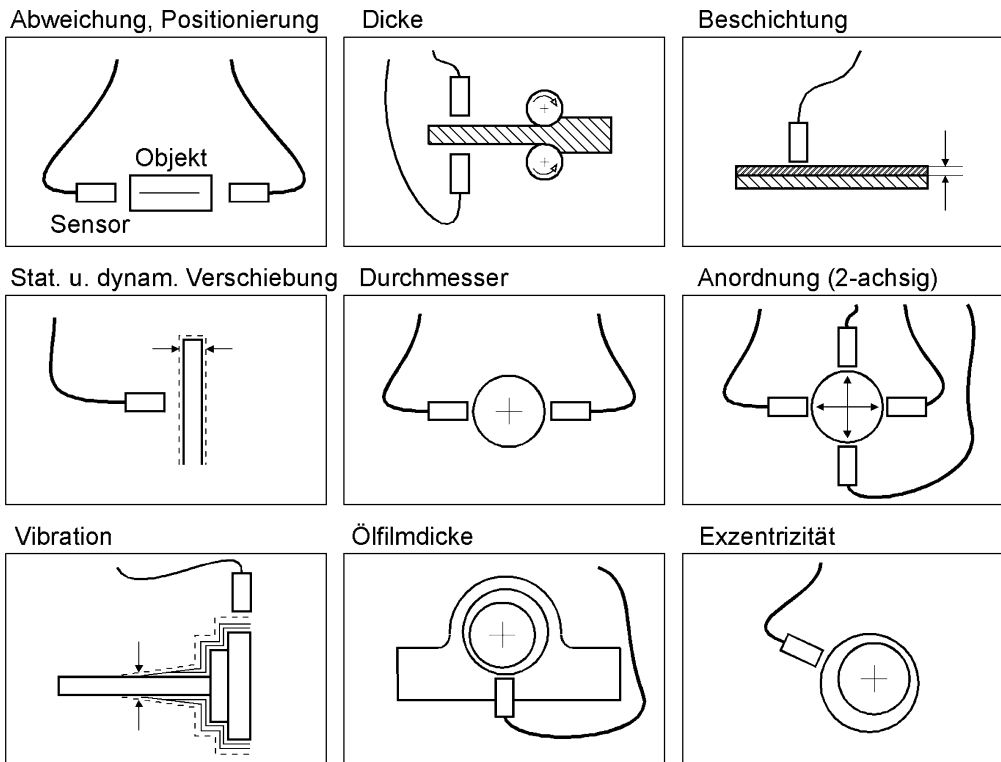


Bild 5-31: Anwendung induktiver Sensoren für unterschiedliche Messaufgaben

rung den Ausgangsverstärker durchgeschaltet. Wie am Beispiel des Schichtdickenmessgerätes in Bild 5-2 gezeigt, kann man die Beeinflussung des Oszillatorkreises auch für eine analoge Entfernungsmessung nutzen (Beschichtung in Bild 5-31).

Außer den schon im Abschnitt 5.2.3 dargestellten Möglichkeiten zur Erfassung von Verschiebungen durch Änderung der Kapazität eines Kondensators gibt es den so genannten **D-Feld-Sensor** (D steht für Dielektrikum) [5.1]. Wie in Bild

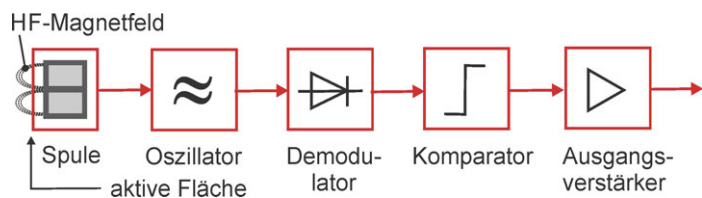


Bild 5-32: Aufbau der Messkette in einem induktiven Näherungsschalters

5-33 dargestellt, besteht der Sensor aus drei Elektroden. An den Elektroden 1 und 2 liegen gegenphasige Wechselspannungen an. Das daraus resultierende Potential wird am Ort der dritten Elektrode S aufgenommen und anschließend frequenz- und phasenselektiv weiter bearbeitet. Im abgeglichenen und ungestörten Zustand liegt die Elektrode S auf Nullpotential. Dringt ein leitendes oder nichtleitendes Objekt in das Messvolumen ein, ändern sich die Potentialverhältnisse. Je nach Lage der drei Elektroden kann man daraus eine einfache Abstandsmessung oder auch eine richtungsabhängige Abstandsmessung ableiten. Auch auf Änderungen der Sensorgeo-

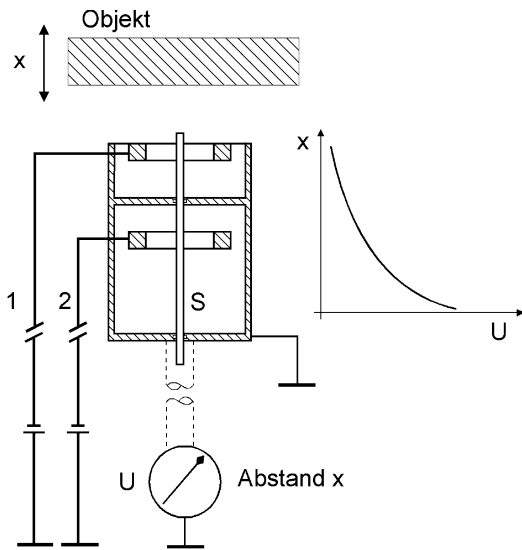


Bild 5-33: Prinzip eines D-Feld-Sensors

metrie reagiert das System sehr empfindlich. Die Kennlinie des Sensors ist nichtlinear (Bild 5-33), in kleinen Bereichen kann man jedoch von einem linearen Abstandsgesetz ausgehen, mit einer Messempfindlichkeit von 10^{-5} . Das Messsystem ist mit geeigneten keramischen Werkstoffen für die Haltematerialien bis zu Umgebungstemperaturen von 1000°C geeignet. Ein typischer Einsatzfall des Systems ist daher als Abstandssensor für das Schneiden von Werkstoffen mit Laserstrahlung. Im Kapitel 9 über Beispiele mechatronischer Systeme wird hiervon noch berichtet werden.

Um das Induktionsprinzip über größere Strecken als Messverfahren anwenden zu können, muss man den Sensor als

rotatorisches System ausführen, dessen Messsignale nach einer Umdrehungen sich zyklisch wiederholen. Ein typischer Vertreter eines solchen Systems ist der **Drehmelder** oder **Resolver**. In seinem Aufbau ähnelt der Resolver einem kleinen Elektromotor mit Präzisionswicklung. Der Stator trägt zwei Wicklungsgruppen, deren Wicklungsebenen senkrecht aufeinander stehen, also räumlich um 90° oder $\pi/2$ zueinander versetzt sind (Bild 5-34 a)). Der Rotor des Systems trägt eine Wicklung, die über Schleifkontakte mit Spannung versorgt werden kann, oder an der

man eine im Rotor induzierte Spannung messen kann. Je nach Anwendungsfall kann das Maschinenteil, dessen Verdrehwinkel φ gegenüber einer Bezugslinie mit dem Resolver bestimmt werden soll, mit dem Rotor oder dem Stator verbunden sein. Über einen Lineartrieb (z.B. Zahnstange-Ritzel-System) können auch Linearbewegungen erfasst werden.

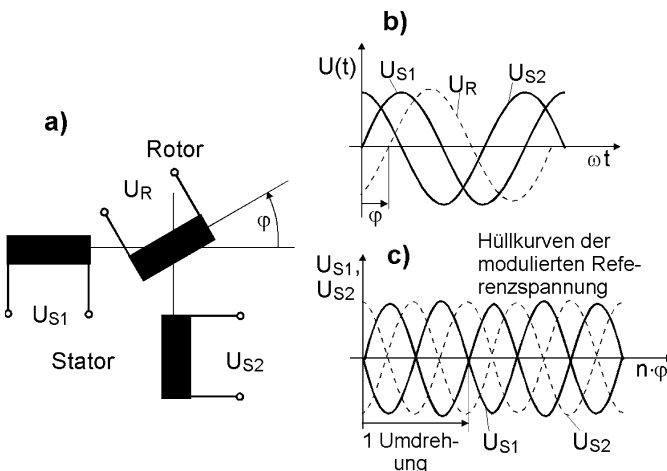


Bild 5-34: Resolver oder Drehmelder a) Funktionsprinzip
b) Phasenauswertung bei Speisung der Statorwicklungen mit Referenzwechselspannung
c) Amplitudenauswertung bei Speisung der Rotorwicklung

Im Prinzip kann man beide Betriebsarten verwenden. Entweder man speist die Statorwicklungen mit einer Wechselspannung

U_{ref} bei $\omega_{\text{ref}} \approx 10 \text{ KHz}$ und misst die im Rotor induzierte Spannung U_R , oder man speist den Rotor mit Wechselspannung und misst die in den Statorwicklungen induzierten Spannungen U_{S1} und U_{S2} .

Im ersten Fall besteht die Messung in der Ermittlung der Phasenverschiebung φ der Rotorspannung gegenüber der Referenzspannung. Diese Phasenverschiebung entspricht dem Verdrehwinkel der Rotorwicklung gegenüber der Statorwicklung. Die Spannung am Rotor ergibt sich aus den in der Rotorwicklung induzierten Spannungsanteilen aus beiden Statorwicklungen:

$$U_R(t) = U_{S1}(t) \cos \varphi + U_{S2}(t) \sin \varphi \quad .$$

Die einzelnen Statorspannungen haben folgende Werte:

$$U_{S1}(t) = U_{\text{Ref}} \sin \omega_S t$$

$$U_{S2}(t) = U_{\text{Ref}} \sin(\omega_S t - \frac{\pi}{2}) = U_{\text{Ref}} \cos \omega_S t \quad .$$

Daraus ergibt sich für die Rotorspannung

$$U_R(t) = U_{\text{Ref}} (\sin \omega_S t \cdot \cos \varphi + \cos \omega_S t \cdot \sin \varphi) = U_{\text{Ref}} (\sin \omega_S t + \varphi) \quad ,$$

d. h. aus der Phasenverschiebung zwischen $U_R(t)$ und $U_{S1}(t)$ lässt sich die Winkelinformation φ gewinnen (Bild 5-34 b)).

Speist man die Rotorwicklung mit der Referenzspannung U_{ref} und misst die beiden in den Statorwicklungen induzierten Spannungen, so haben diese folgenden Wert:

$$U_{S1}(t) = K \cdot U_{\text{Ref}} \cdot \sin \varphi \quad \text{bzw.} \quad U_{S2}(t) = K \cdot U_{\text{Ref}} \cdot \cos \varphi \quad .$$

Die Referenzspannung wird demnach durch den Verdrehwinkel φ in seinem Wert amplitudenmoduliert (Bild 5-34 c)). Aus der Messung der Spannung an einer Statorwicklung kann man Winkelwerte für φ in einem Quadranten ablesen, durch Vergleich beider Ausgangsspannungen wird es möglich, Winkelwerte zwischen 0° und 360° eindeutig zu messen. Da man jederzeit innerhalb einer Umdrehung aus der Messung zweier Spannungen eindeutig die Winkellage des Rotors bestimmen kann, spricht man von einer **absoluten** Lagemessung. Will man beispielsweise für die Lagebestimmung eines Linearantriebs die Winkellage über viele Umdrehungen des Resolvers bestimmen, so muss man die Winkelwiederholungen mitzählen, um den absoluten Lagemesswert zu bestimmen. Man spricht deshalb von einer **zyklisch absoluten** Messung.

Bei der oben erwähnten Frequenz der Referenzspannung kann man Winkelauflösungen von 10 Winkelminuten erreichen. Die kleine Bauweise und hohe Genauigkeit sind besondere Vorteile des Resolversystems, weshalb sie häufig zur Lagemessung in Antriebssystemen und Industrierobotern verwendet werden. Problematisch ist die mechanische Übertragung der Rotorspannung über Schleifringkontakte oder Bürsten. Um störende Kapazitäten und Übergangswiderstände so klein wie möglich zu halten, werden meist Edelmetallkontaktpaarungen verwendet.

Eine andere Möglichkeit magnetische Effekte in einem Wegmesssystem zu nutzen zeigt Bild 5-35. Es zeigt die Auswerteeinheit eines **inkrementalen** Wegmesssys-

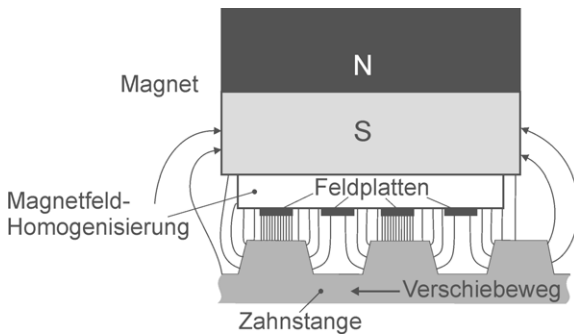


Bild 5-35: Inkrementales Wegmesssystem mit Feldplatten

tems für Linearbewegungen, das als Sensorsystem **Feldplatten** verwendet [5.11]. Die vier Feldplatten, die in einer Brückenschaltung zusammengeschaltet sind, befinden sich im Magnetfeld eines Permanentmagneten, dessen Feld durch ferromagnetische Plättchen homogenisiert wird. Jeweils zwei der Feldplatten sind so angeordnet, dass sie genau dem Raster der Zähne und Zahnücken der ferromagnetischen Zahnstange entsprechen. Durch Relativbewegung

der Zahnstange gegenüber dem Sensorkopf wird das Magnetfeld senkrecht zu den Feldplatten zyklisch verändert. Die dadurch auftretenden Widerstandsänderungen der Feldplatten ergeben die Zählimpulse für ein inkrementales Wegmesssystem.

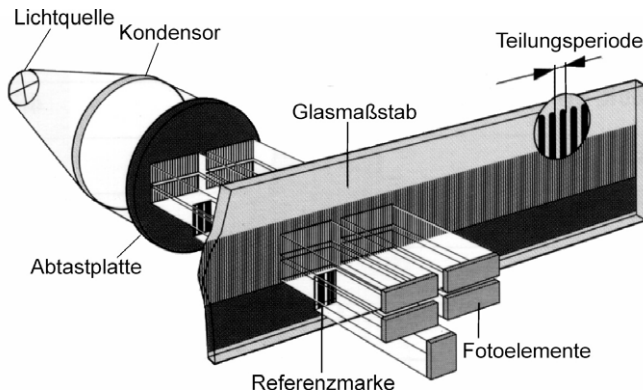


Bild 5-36: Fotoelektrische Abtastung eines Glasmaßstabs im Durchlichtverfahren (Heidenhain)

Um Winkellagen über viele Umdrehungen des Messsystems oder lineare Verschiebungen über größere Strecken messen zu können, hat sich jedoch eindeutig die optische Abtastung von Maßverkörperungen durchgesetzt. Solche **Maßverkörperungen** sind bei rotierenden Systemen Kreisscheiben und bei linearen Systemen Lineale. Diese tragen, je nach Verfahren, sehr feine Stricheinteilungen in der Größe der

Auflösung des Messsystems, oder darüber hinaus auch noch Markierungen mit geringerer Auflösung. Die Auswertung solcher Messsysteme liefert nicht wie beim Resolver analoge, sondern digitale Signale.

Man kann zwei wichtige unterschiedliche Messverfahren unterscheiden. Das **inkrementale**⁴ Messverfahren benutzt als Maßverkörperung einen Tragkörper aus Glas, auf den eine Gitterteilung aus lichtundurchlässigen Strichen und lichtdurchlässigen Lücken aufgebracht ist (Bild 5-36). Die Teilungsperiode beträgt bei sehr genauen Maßstäben 10 μm oder 20 μm . Zusätzlich sind auf einer zweiten Spur eine oder mehrere Referenzmarken aufgebracht. Die optische Abtasteinheit besteht aus einer Lichtquelle, einem Kondensor, der die Lichtstrahlen parallel ausrichtet, der Abtastplatte mit den Abtastgittern und den Fotoelementen. Bei einer Bewegung des Maß-

4 Es findet eine relative Messung kleiner Wegstücke (Inkrementale) statt, indem zu einem bereits bekannten Positionswert Inkrementale aufaddiert oder davon subtrahiert werden.

stabs relativ zur Abtasteinheit entstehen an den vier Fotoelementen der Inkremental-Teilung sich periodisch ändernde Signale. Um eine hohe Güte der Ausgangssignale zu erreichen, wird gleichzeitig eine große Anzahl von Strichen abgetastet. Als Ausgangssignale entstehen zwei sinusförmige Signale, die um 90° elektrisch zueinander phasenverschoben sind. Durch Zählung der Maxima der Signale kann man bei bekannter Teilungsperiode den zurückgelegten Verschiebeweg ermitteln. Durch gleichzeitige Auswertung der beiden phasenverschobenen Signale kann auch die Verschieberichtung bestimmt werden. Die Grundgenauigkeit in der Größenordnung der Teilung kann noch durch Mehrfachauswertung und Interpolation elektronisch auf beispielsweise $1\text{ }\mu\text{m}$, oder bis zum 100-fachen erhöht werden. Gleichartige Systeme können natürlich auch für rotatorische Bewegungen verwendet werden, bei denen die Maßverkörperung kreisförmig auf einer Scheibe angebracht ist.

Als Nachteil inkrementaler Längenmessung kann man anführen, dass durch fehlerhafte Abtastung auftretende Verfälschungen des Messergebnisses sich fortpflanzen und erst durch ein neuerliches Anfahren der Referenzmarke wieder korrigiert werden können. Das Gleiche ist erforderlich, wenn die aktuelle Positionsmessung durch Spannungsausfall verloren geht. Um dies zu vermeiden, wird meist eine Pufferbatterie in die Zählelektronik eingebaut. Positiv zu werten ist der einfache Aufbau der optischen Abtastung und der Maßverkörperung, der bei kodierten, absolut messenden Systemen sehr viel größer wird.

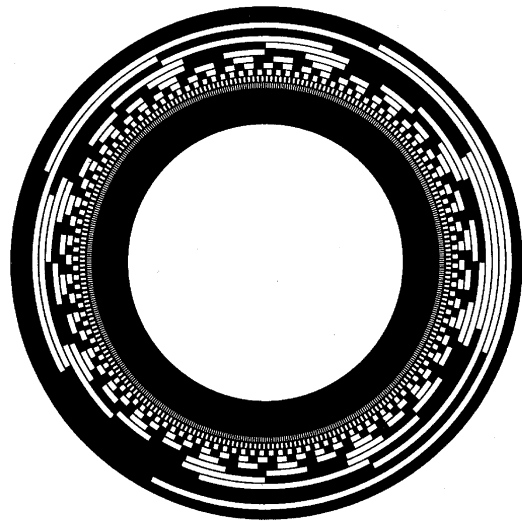


Bild 5-37: Kodierte Maßverkörperung eines Winkelmesssystems (Heidenhain)

Bild 5-37 zeigt die **kodierte**⁵ Maßverkörperung eines Winkelmesssystems. Auf der Glasscheibe sind, genauso wie beim inkremental geteilten Glasmaßstab, lichtdurchlässige und lichtundurchlässige Zonen aufgebracht. Die einzelnen Spuren (Anzahl 11) werden radial durch eine fotoelektrische Abtasteinheit in jeder Spur abgetastet. Die Teilung der am weitesten innenliegenden und am feinsten geteilten Spur bestimmt die Winkelauflösung des Messsystems. Entlang einer radialen Abtastlinie stellt dann die Hell-, Dunkel-Kodefolge einen binären Zahlenwert⁶ dar, der der zugehörigen Winkellage entspricht.

Je höher die Auflösung des Messsystems sein soll, um so mehr Spuren und damit auch Abtastelemente werden benötigt. Dadurch sind kodierte Messsysteme nicht

5 Der verwendete Kode kann ein einfacher Dualkode oder ein einschrittiger Kode (Gray-Code) sein.

6 Bei Durchlichtabtastung stellt beispielsweise ein dunkles Feld den logischen Wert 0 und ein helles Feld den Wert 1 dar. Bewertet man die Abtastwerte in den Spuren, von der innersten Spur ausgehend, mit der binären Wertigkeit 2^0 , 2^1 , 2^2 , ... usw., so kann man eine Dualzahl bilden.

so kostengünstig wie inkrementale, haben aber den Vorteil, dass zu jeder Zeit und auch bei zwischenzeitlichen Störungen, eine **absolute** Wegmessung erfolgt, d. h. die Messung bezieht sich immer auf den Nullpunkt des Messsystems. Das beschriebene rotatorische Prinzip kann natürlich als linear arbeitendes System an einem Linearantrieb für dessen Wegmessungen verwendet werden, bzw. es werden auch absolut kodierte lineare Wegmesssysteme hergestellt.

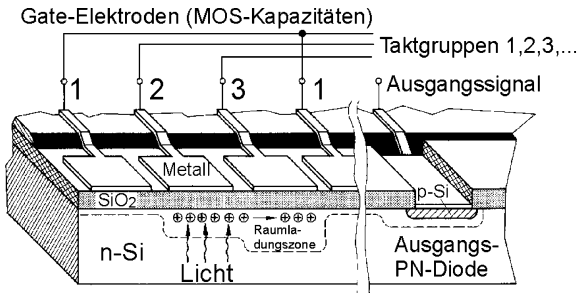


Bild 5-38: Aufbau eines CCD-Elementes

Zur Erfassung oder Vermessung komplexerer geometrischer Objekte werden fotooptische Bildsensoren in Form von CCD-Anordnungen verwendet. **CCD-Elemente** (Charge Coupled Device: Ladungsgekoppeltes Element) bestehen aus einer MOS-Struktur⁷. Eine Vielzahl von Gate-Elektroden aus Metall sind getrennt durch eine dünne Oxidschicht auf einen homogenen

Halbleiterkristall aus N-Silizium aufgebracht (Bild 5-38). Die Elektroden sind in drei Gruppen untereinander verbunden und ineinander geschachtelt angeordnet. Von der Rückseite der Halbleiterschicht auftretendes Licht erzeugt Minoritätsträger, die in einem Potentialtopf unter der Gateelektrode an der Si-SiO₂ Grenzschicht gesammelt und gespeichert werden. Dadurch entsteht ein Ladungsbild, dessen Intensität der Helligkeitsverteilung des lichtgebenden Objektes entspricht. Durch Anlegen geeigneter Taktspannungen an die Metallelektroden lassen sich die Potentialmulden im Halbleiter so verändern, dass die unter einer Zelle befindlichen Ladungen zur nächsten weitergeschoben werden. Ein CCD-Element ist somit ein Analog-Schieberegister, das aus einer Anzahl benachbarter MOS-Kapazitäten besteht. Der Ausgang besteht aus einem gesperrten PN-Übergang, der die durch den Verschiebeprozess ankommenden Ladungspakete in Stromimpulse wandelt. Diese werden bei zeilen- und matrixartigen Anordnungen solcher CCD-Elemente durch entsprechende Auslesetakte in Ausgaberegister verschoben und liefern die entsprechenden Helligkeitsinformationen. Dabei ist das Bild entsprechend der Anordnung und den Abmessungen der Gate-Elektroden gerastert.

Zeilen- und matrixartige CCD-Arrays (Bild 5-39) werden als lichtgebende Elemente in Kameras eingesetzt. Heute erhält man Zeilenarrays mit bis zu 4000 **Pixeln**⁸ und Matrixarrays mit bis zu 4000 x 2000 Pixeln Auflösung. Da die geometrische Auflösung eines Pixels unter Berücksichtigung der Abbildungseigenschaften der Optik genau bekannt ist, kann man solche Kameras sowohl zur Form- und Lageerkennung als auch zur Vermessung von Objekten (Winkel, Längen, Flächen) verwenden.

Ein weiteres optoelektronisches Element, mit dem die genaue Position eines Licht-

⁷ MOS: **M**etal **O**xide **S**emiconductor-Metalloxide-Halbleiter.

⁸ Pixel: picture element, kleinstes Bildelement eines bilderzeugenden (CCD-Array) oder bildgebenden Systems (Bildschirm, Display)

strahls bestimmt werden kann, ist der **PSD** (**P**osition **S**ensitive **D**etector) [5.6]. Als Positionssensor liefern diese so genannten **Lateraleffekt-Fotodioden** genaue Daten über die Lage eines über die Messfläche wandernden Lichtpunktes. Ein PSD ist wie eine PIN-Fotodiode mit relativ großen Abmessungen aufgebaut.

Die P-Si-Schicht (Bild 5-40) ist lichtempfindlich. Während die N-Si-Schicht durchgehend metallisiert ist, befinden sich nur an den beiden Enden der P-Si-Schicht die Metallelektroden A und B. Wird ein beliebiger Punkt des PSD von einem Lichtstrahl getroffen, so werden dort Ladungsträger freigesetzt, die einen Stromfluss zu den beiden Elektroden bewirken. Da die unbeleuchteten Teile der Schicht wie Widerstände wirken, ist das Verhältnis der Ströme I_A/I_B von der Lage des Lichtpunktes abhängig. Mittels einer Steuerelektronik kann aus dem Verhältnis der Ströme die genaue Entfernung ermittelt werden. Es gilt:

$$\frac{D - 2a}{D} = \frac{I_A - I_B}{I_A + I_B}.$$

Das Verhältnis der Ströme ist von der einfallenden Lichtmenge unabhängig, wodurch eine stabile Entfernungsmessung, ohne Beeinflussung durch das Reflexionsvermögen des Objektes, möglich ist.

Außer für lineare Positionsmessungen gibt es PSD's auch für zweidimensionale Lageerkennungen. PSD's zeichnen sich gegenüber anderen Sensoren durch hohe Auflösung der Positionsdaten, kurze Ansprechzeit und einfache Beschaltung aus.

In der mobilen Robotik und in der Verkehrstechnik werden häufig Sensoren zur Bestimmung von Entfernung, Geschwindigkeit oder Beschleunigung von Objekten im Umfeld des Gerätes benötigt. Hierzu eignen sich Laufzeitentfernungsmesser nach dem **Sonar-Prinzip** oder das **Laserradar** (Lidar). Beiden Messprinzipien ist gemeinsam, dass kurze Impulse von Schall- bzw. Lichtwellen ausgesendet werden, die an einem Objekt reflektiert und von einem Empfänger wieder registriert werden können. Auf diese Weise kann durch Bestimmung der Signallaufzeit bei bekannter Ausbreitungsgeschwindigkeit der Wellen die Entfernung zum Reflektor bestimmt werden. Das geht auch über größere Entfernungen und absolut berührungslos.

Bild 5-41 zeigt das Messprinzip der Entfernungsmessung mit Ultraschall. Dazu wird im Sende- und Empfangssensor eine dünne Scheibe aus piezoelektrischen Material durch einen elektrischen Sendeimpuls aufgrund des reziproken Piezoeffektes zu Schwingungen im Ultraschallbereich angeregt (Bild 5-41 a)). Durch den extrem kur-

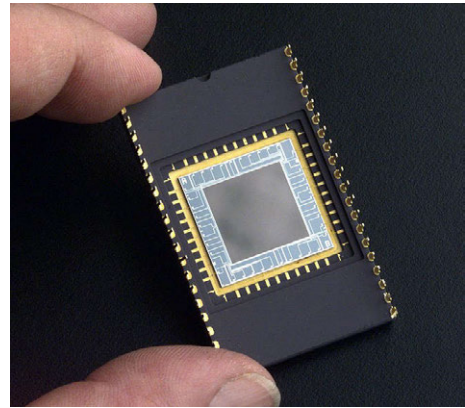


Bild 5-39: CCD-Sensor mit 2061×1887 Pixeln (Wikipedia)

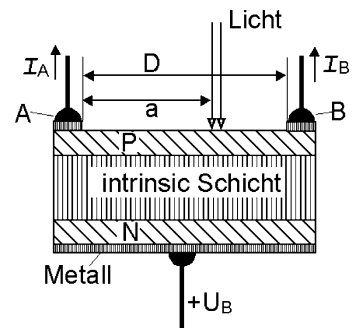


Bild 5-40: Aufbau eines PSD-Fotosensors

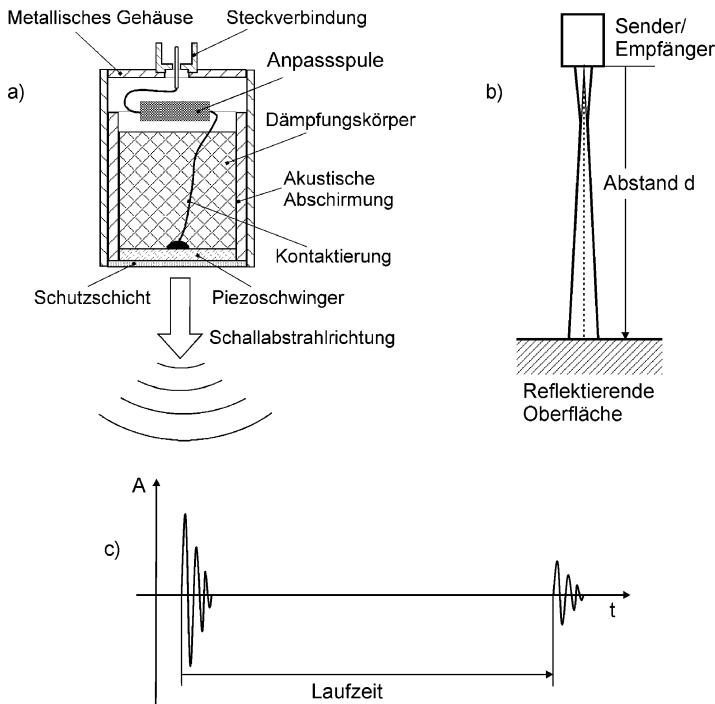


Bild 5-41: Entfernungsmessung mit Ultraschall a) Ultraschall-Sender/Empfänger b) Abstand d zwischen Sender und Reflektor c) Laufzeitmessung zwischen Send- und Empfangssignal

zen elektrischen Impuls ist der ausgesendete Ultraschallimpuls ebenfalls sehr kurz. Die Schallwellen breiten sich als Longitudinalwellen senkrecht zum Schwinger in der Luft aus (Bild 5-41 b)). Der ebenfalls zur Rückwand des Sensors ausgesendete Schallimpuls wird im Dämpfungskörper absorbiert. An der Grenzfläche eines Objektes, dessen Abstand d zum Sensor bestimmt werden soll, wird der Ultraschallimpuls reflektiert und gelangt auf dem gleichen Weg zum Sensor zurück. Beim Auftreffen auf den Piezoschwinger wird das Ultraschall-echo aufgrund des

Piezeffektes wieder in ein elektrisches Signal gewandelt. Durch eine Laufzeitmessung zwischen Send- und Echosignal kann dann auf die Entfernung zwischen Sensor und Reflektor zurück geschlossen werden (Bild 5-41 c)). Send- und Empfangssignal haben einen Amplitudenunterschied von mindestens 40 dB, so dass das Empfangssignal vor der Auswertung verstärkt werden muss. Bei der Abstandsbestimmung aus der Signallaufzeit ist im Empfänger noch die Schallgeschwindigkeit in Luft zu berücksichtigen, die zusätzlich noch temperaturabhängig ist. Auf diese Weise können in Luft Entfernungen bis zu 30 m und in Wasser (Echolot) sogar bis zu mehreren hundert Metern gemessen werden.

Beim Sonar werden von dem piezoelektrischen Schwinger Impulse von ca. 1mS im Ultraschallbereich zwischen 50 und 100 KHz in die Luft abgestrahlt. Bild 5-42 zeigt die Empfindlichkeit eines Sonarsensors, die nicht gleichmäßig über den überstrahlten Bereich verteilt ist, sondern aus keulenförmigen Mittel- und Seitenfeldern besteht [5.10]. Man sieht, dass innerhalb eines Kegels von 20° die Sensorempfindlichkeit entlang der Mittelachse der Ausstrahlung um 10 dB d.h. um das Zehnfache sinkt. Da die Schallgeschwindigkeit v_s in Luft bekannt ist, kann die Entfernung d des den Impuls reflektierenden Objektes aus dem Zeitintervall Δt zwischen Aussenden und Empfangen des Impulses berechnet werden:

$$d = \frac{1}{2} v_s \cdot \Delta t \quad . \quad (5.3)$$

Die kleinste Distanz $d_{\min}$, die mit dieser Methode noch gemessen werden kann, hängt von der Impulsdauer des Sendeimpulses δt ab:

$$d_{\min} = \frac{1}{2} v_s \cdot \delta t$$

Bei einer Impulsdauer von 1,2 mS liegt dieser Abstand bei etwa 20 cm. Die maximale Distanz $d_{\max}$, die noch gemessen werden kann, hängt vom Zeitintervall Δt_w

zwischen den einzelnen Sendeimpulsen ab. Aufgrund der Energieabnahme des Schallimpulses sind jedoch nur Reichweiten von 10-30 m möglich. Bei reiner Laufzeitmessung liegt die maximale Auflösung eines solchen Systems im Bereich der Wellenlänge λ der Ultraschallschwingung:

$$\lambda = \frac{v_s}{f} \quad \text{mit } f : \text{Frequenz der Ultraschallschwingung.}$$

Bei einer Sendefrequenz von 100 kHz beträgt dieser Wert ungefähr 34 mm. Wertet man die Phasenverschiebung zwischen Sende und Empfangsschwingung noch aus, so sind auch höhere Auflösungen möglich.

Laserentfernungsmesser arbeiten nach dem gleichen Prinzip wie die mit Sonarsensoren arbeitenden Systeme. Die Strahlungserzeugung erfolgt durch eine Laser-Diode im Infrarotbereich. Für GaAs-Halbleiter liegt die Wellenlänge im Bereich von 1 μm . Bei Entfernungsmessung nach dem Laufzeitprinzip gilt die Gleichung (5.3), in der die Schallgeschwindigkeit durch die Lichtgeschwindigkeit c ersetzt werden muss. Die Reichweite in Luft kann abhängig von der Reflexionseigenschaft der zu detektierenden Objekte bis zu mehreren hundert Metern betragen, die Auflösung ist aufgrund der kürzeren Wellenlänge im Bereich einiger Millimeter, oder bei Phasenauswertung noch besser.

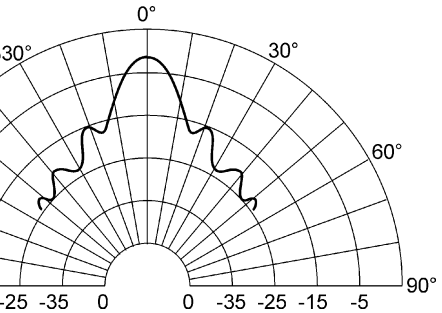


Bild 5-42: Richtungsempfindlichkeit eines Sonarsensors (Polaroidserie 600) bei 50 KHz



Bild 5-43: Laser-Scanner (Sick)

Aufgrund des Laser-Prinzips ist der Strahl stark gebündelt. Um daher z. B. zur Kollisionsmessung einen größeren Bereich überwachen zu können, wird der Strahl der Laserdiode durch einen Drehspiegel fächerförmig zyklisch abgelenkt (bis zu 180°), so dass der gesamte Bereich vor dem System (Bild 5-43) auf Objekte abgescannt werden kann. Man kann dadurch auch die gesamte Kontur eines Objektes abtasten und seine Geometrie und Entfernung ermitteln.

5.3.1.2 Geschwindigkeitssensoren

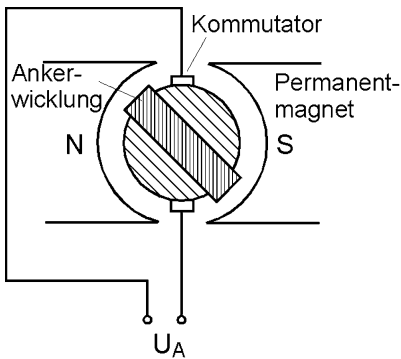


Bild 5-44: Prinzip eines Gleichstrom-Tachometers

Geschwindigkeiten, die in mechatronischen Systemen bestimmt werden müssen, liegen als Linear- oder als Winkelgeschwindigkeiten vor. Da Antriebe für lineare Bewegungen meist rotierende Systeme sind (z. B. Elektromotoren), kann man die Messung von Lineargeschwindigkeiten häufig auch auf Drehzahlmessungen an der Antriebsmaschine rückführen. Wie bereits gesagt, kann man prinzipiell die Geschwindigkeitsinformation durch Differentiation aus einer ohnehin stattfindenden Wegmessung ableiten. So ist das Messsignal eines Resolvers für diese Vorgehensweise gut geeignet. Häufiger wendet man aber die Messung der Drehzahl durch einen Tachogenerator (Gleichspannungsgenerator) an.

Bild 5-44 zeigt das Funktionsprinzip eines Dauermagnet-Gleichstromtachometers. Der Dauermagnet erzeugt ein konstantes und gleichmäßiges Magnetfeld. Die Bewegung des elektrischen Leiters (Spule) im Magnetfeld induziert in ihm eine der Geschwindigkeit proportionale Spannung. Der Läufer wird direkt mit dem rotierenden Objekt verbunden, dessen Drehzahl zu messen ist. Die in den Wicklungen induzierten Wechselspannungen werden durch den Kommutator mit den aufliegenden Bürsten mechanisch gleichgerichtet und zu einer Gleichspannung zusammengesetzt. Wegen der Induktion weist der Tachometer eine nicht zu vernachlässigende Zeitkonstante auf, so dass die induzierte Spannung bei schnellen Drehzahlwechseln nur verzögert der Drehzahl folgt. Kennzeichnend für Tachogeneratoren ist:

- großer Drehzahlbereich von $0,1 - 10.000 \text{ min}^{-1}$
- einfache Drehrichtungserkennung über Polaritätswechsel der Spannung
- geringe Oberwelligkeit der Tachospannung $< 2 \%$
- linearer Zusammenhang zwischen Drehzahl und Tachospannung mit Abweichungen $< 0,2 \%$
- mechanisch, elektrisch und thermisch sehr robust
- keine elektrische Hilfsenergie erforderlich.

Das mit den Bürsten und den damit verbundenen Verschleißerscheinungen entstehende Wartungsproblem kann man durch bürstenlose Tachogeneratoren umgehen. Bei diesen Generatoren induziert ein Permanentmagnetrotor in drei Wicklungssystemen drei zeitlich gegeneinander versetzte, sich überlappende Trapezspannungen (Bild 5-45).

Diese Signale steuern eine elektronische Gleichrichterschaltung so an, dass aus den positiven und negativen Trapezspannungen der drei Wicklungen die Teilbereiche 1-6 herausgeschnitten und zu einem Ausgangssignal zusammengesetzt werden. Das gleichgerichtete Signal wird in einen eingepprägten Strom umgesetzt, der

der Drehzahl in Betrag und Richtung proportional ist.

Bei digitalen Regelungen ist es wünschenswert, dass die gemessenen Istwerte schon als digitale Werte zur Verfügung stehen, so dass die sonst erforderliche A/D-Wandlung entfallen kann. Hierzu kann man bei Wegmessung mit einem optoelektrischen inkrementalen oder absoluten digitalen Messsystem, das höchstfrequente digitale Messsignal verwenden. Das ist bei inkrementaler Messung das Signal von einer Messspur und bei absoluter Messung die Spur mit dem höchstfrequenten Signal (LSB : Least

Significant Bit). Die Geschwindigkeitsmessung kann dann als Messung der Drehfrequenz erfolgen, wobei die Auflösung der Drehzahl natürlich um so besser wird, je mehr Signalwechsel das Wegmesssystem pro Umdrehung liefert.

Ähnlich arbeitende einfache Drehzahlmesser kann man auch mit Sensoren aufbauen, die mit induktiven Fühlern oder galvanomagnetischen Bauteilen arbeiten. Ordnet man auf dem sich drehenden Objekt beispielsweise ein Zahnrad aus ferromagnetischem Material an, an dessen äußeren Umfang ein induktiver Näherungsfühler radial vor den Zahnköpfen angebracht ist (Bild 5-46), so liefert der Sensor ebenfalls ein Signal, dessen Frequenz der Drehzahl proportional ist.

Will man für die gleiche Aufgabe das galvanomagnetische Bauteil "Feldplatte" verwenden, so muss man berücksichtigen, dass wie in Bild 5-24 b) dargestellt, die Feldplatte für kleine magnetische Flussdichten B eine nichtlineare Widerstandskennlinie $R(B)$ aufweist. Für größere Flussdichten oberhalb ca. 0,3 T verhält sie sich annähernd linear. Dieses Verhalten führt dazu, dass wie in Bild 5-47 [5.7] dargestellt, bei kleiner Flussdichte ein Eingangssignal ΔB zu einem völlig verfälschten Ausgangssignal $\Delta R(B)$ führt. Bringt man jedoch einen kleinen Permanentmagneten auf der Feldplatte an, so erhält sie durch dessen Magnetfeld eine Vormagnetisierung und dadurch einen vom Anfangswiderstand $R(0)$ abweichenden Wert R_{V0} . Wirkt nun das gleiche Eingangssignal ΔB auf die Feldplatte, so wird dieses einigermaßen proportional auf das Ausgangssignal abgebildet. In einer Anwendung mit Vormagnetisierung kann die Feldplatte dann einfach durch Annäherung eines ferromagnetischen Objektes (Zähne eines Zahnrades) angesteuert werden.

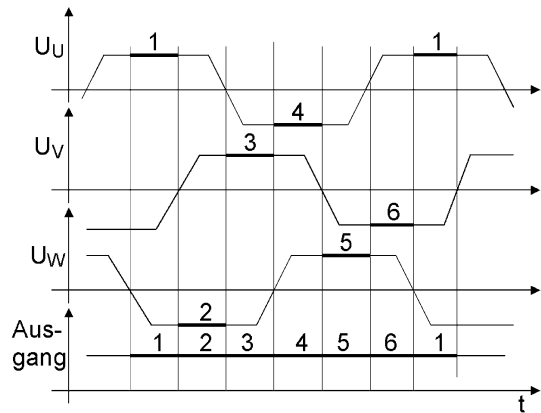


Bild 5-45: Signalverlauf des bürstenlosen Tachogenerators

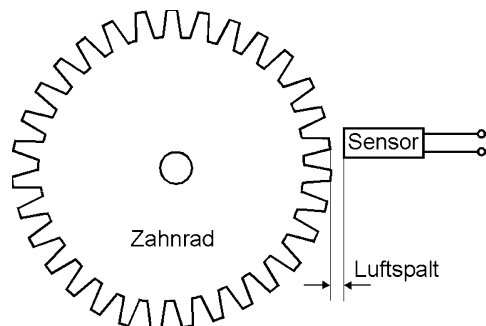


Bild 5-46: Prinzip der Abtastung eines Zahnades mit einem induktiven oder galvanomagnetischen Sensor

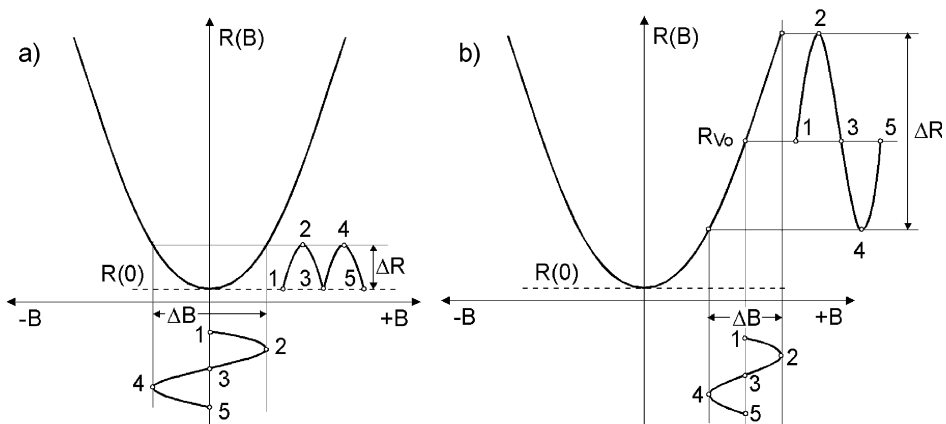


Bild 5-47: Ansteuerung einer Feldplatte mit einem sinusförmigen, magnetischen Signal
a) ohne Vormagnetisierung b) mit Vormagnetisierung

Feldplatten besitzen einen großen negativen Temperaturkoeffizienten, der je nach Arbeitspunkt - 0,3 bis - 0,4 % / K beträgt. Durch eine Differenzanordnung kann die thermische Nullpunktdrift und die thermische Empfindlichkeitsänderung weitgehend kompensiert werden. Dazu werden zwei Feldplatten in einem Sensor so angeordnet, dass sie gleichzeitig vom Messobjekt gegensinnig beeinflusst werden, wohingegen die Temperatur die beiden relativ nahe nebeneinander angeordneten Feldplatten im gleichen Sinne beeinflusst. Eingebaut in eine Wheatstonesche Brücke kann man bei einer solchen **Differentialfeldplatte** nicht nur die thermischen Einflüsse kompensieren, sondern auch die Übertragungskennlinie linearisieren und außerdem Messempfindlichkeit und Messbereich verdoppeln.

5.3.1.3 Beschleunigungssensoren

Beschleunigungssensoren sind in der Regel so aufgebaut, dass man eine kleine seismische Masse m in der Richtung, in der die Beschleunigung gemessen werden soll, federnd aufhängt. Wirkt eine Beschleunigung auf diese Masse, so wird diese in Richtung der Beschleunigung ausgelenkt (Weg, Verformung) und auf die Masse wirkt eine Kraft. Dadurch ergeben sich zwei Möglichkeiten der Auswertung, nämlich durch eine Kraftmessung (bei bekannter seismischer Masse) oder durch eine Messung der Auslenkung. Bei rein elastischer Verformung ist die Auslenkung der Kraft proportional.

Da die meisten Beschleunigungsvorgänge dynamisch sind, sind solche Sensoren auch zur Messung von Schwingungsvorgängen gut geeignet. Für Schwingungsmessungen mit Beschleunigungssensoren an mechanischen Systemen sollten diese daher einen großen Frequenzbereich erfassen können. Als untere Grenzfrequenz ist unter Umständen 2 Hz (für langsame Stöße, Flatterschwingungen) erforderlich, die obere Grenzfrequenz sollte mindestens 10 kHz (schnelle Stöße und akustische Schwingungen) betragen. Außerdem sollten solche Sensoren unempfindlich für Beschleunigungen sein, die nicht in Richtung der zu messenden Beschleunigung wirken und die Eigenschwingungen des Messobjektes möglichst ge-

ring beeinflussen (geringe Sensormasse). Eine geringe Sensormasse ist auch deshalb erforderlich, weil die Resonanzfrequenz des Sensors möglichst hoch über der Eigenfrequenz des Messobjektes liegen sollte.

Bild 5-48 zeigt schematisch eine solche Messanordnung, die auf einem Messobjekt befestigt ist, dessen Beschleunigung gemessen werden soll. Die bereits mehrfach für ein Feder-Masse-Dämpfer-System hergeleitete DGL lautet:

$$m\ddot{s}(t) + d\dot{s}(t) + ks(t) = m\ddot{x}(t) \quad \text{oder}$$

$$\frac{m}{k}\ddot{s}(t) + \frac{d}{k}\dot{s}(t) + s(t) = \frac{m}{k}\ddot{x}(t)$$

Misst man mit dem Sensor den Weg $s(t)$, der der zu messenden Beschleunigung $\ddot{x}(t)$ proportional ist, so sollte der Wert m/k möglichst groß sein, um eine hohe Empfindlichkeit zu erzielen. Die Resonanzfrequenz des Systems beträgt $\omega = \sqrt{k/m}$, d. h. für eine hohe Resonanzfrequenz sollte k/m möglichst groß, bzw. m/k möglichst klein sein. Da diese Forderungen sich gegenseitig widersprechen, wählt man für das Feder-Masse-System in der Regel eine steife Feder und eine kleine Masse, wodurch man hohe Eigenfrequenz bei geringer Empfindlichkeit erhält. Dies kann man durch einen direkt im Sensor integrierten Verstärker ausgleichen.

Für die Messung des Weges der seismischen Masse⁹ kann man unterschiedliche, bereits besprochene Messeffekte einsetzen. Bild 5-49 zeigt schematisch den Aufbau eines in Dickschichttechnik (Sandwich-Struktur) aufgebauten Beschleunigungssensors mit kapazitiver Erfassung des Weges. An einer Pendelmembran ist eine seismische Masse angebracht, die in horizontaler Richtung zwischen zwei metallisierten Keramikplatten frei schwingen kann. Die Kapazitätsänderungen zwischen der Membran und den beiden Platten, die bei Schwingungsvorgängen auftreten, bilden den Messeffekt.

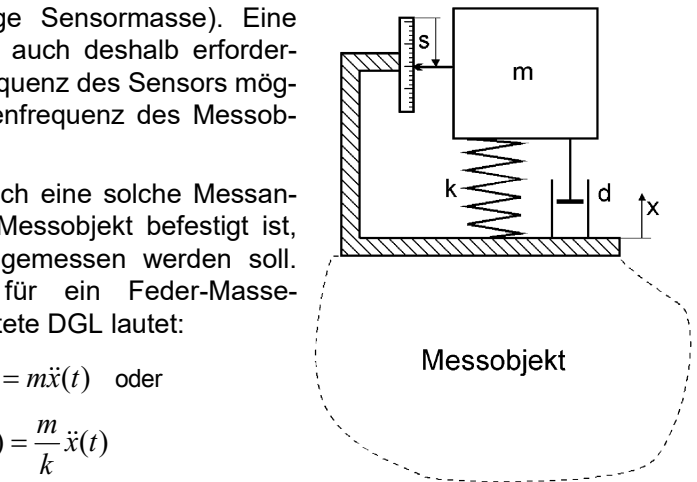


Bild 5-48: Prinzip der Beschleunigungsmessung durch Wegmessung an einem seismischen Schwinger

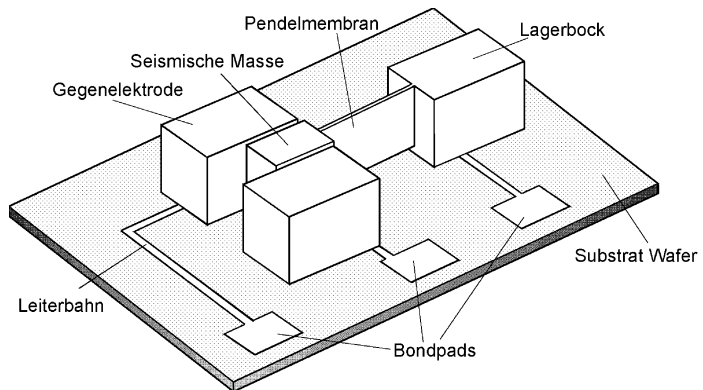


Bild 5-49: Prinzipskizze eines kapazitiven Beschleunigungssensors mit freischwingender seismischer Masse

⁹ elastisch gelagerte träge Masse.

Wie oben erwähnt, kann man eine Beschleunigungsmessung auch auf eine Messung derjenigen Kraft rückführen, die auf eine beschleunigte seismische Masse wirkt. Daher können für Beschleunigungssensoren alle bekannten Kraftmessverfahren angewendet werden.

Häufig wird der piezoelektrische und der piezoresistive Effekt für die Kraftmessung in Beschleunigungssensoren eingesetzt. Das piezoelektrische Messprinzip hat den

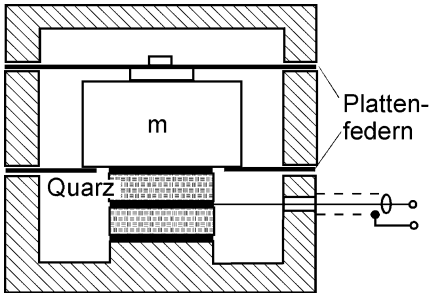


Bild 5-50: Prinzipskizze eines Beschleunigungssensors mit piezoelektrischer Kraftmessung

Vorteil, dass es ohne Hilfsenergie arbeitet und sehr empfindliche Sensoren ermöglicht, aber dafür sind keine statischen Messungen möglich und die hohe Ausgangsimpedanz macht aufwendige Verstärkertechniken erforderlich. Piezoresistive Aufnehmer ermöglichen auch statische Messungen und sind unempfindlicher gegen Temperaturschwankungen. Die niedrige Ausgangsimpedanz piezoresistiver Sensorelemente erleichtert außerdem die Anpassung an nachgeschaltete Geräte. Bild 5-50 zeigt schematisch den Aufbau eines Beschleunigungssensors mit mehreren hintereinander geschalteten Quarzscheiben als piezoelektrische Elemente.

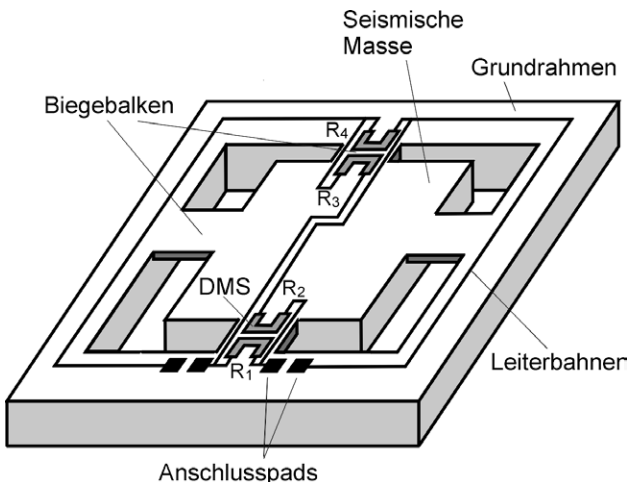


Bild 5-51: Prinzipskizze eines Beschleunigungssensors mit auf Biegebalken integrierten DMS in Brückenschaltung

Eine weitere Möglichkeit der Kraftmessung ist die mit Dehnungsmessstreifen, wobei die Widerstandsänderung der Verformung des DMS und damit der Massenkraft proportional ist. Bei bekannter seismischer Masse ist dann auch die Beschleunigung selber bestimmbar. In Bild 5-51 ist der Aufbau eines solchen Sensors dargestellt [5.2]. Die seismische Masse ist an mehreren Biegebalken aufgehängt, auf denen zu Wheatstonebrücken verschaltete DMS angebracht sind. Durch Wahl von Länge, Breite und Dicke der Balken und der Größe der seismi-

schen Masse, kann der Messbereich des Sensors eingestellt werden. Die Anordnung der Biegebalken ist wichtig, um die Empfindlichkeit des Sensors auf Beschleunigungen, die nicht in der Messrichtung liegen, zu unterdrücken. Die Querempfindlichkeit kann auch durch entsprechende Verschaltung der DMS korrigiert werden.

5.3.2 Kraft- und Momentsensoren

Schon im vorherigen Abschnitt wurde mehrfach die Kraftmessung erwähnt, die dort dazu verwendet wurde, die Massenkraft zu bestimmen, der eine seismische Masse unterliegt. Zur direkten Kraftmessung werden die gleichen Messprinzipien angewendet, wobei die zu messende Kraft hier einen im Kraftfluss liegenden Körper verformt. Dessen Verformung ist der eingeleiteten Kraft proportional. Die Verformung kann dann beispielsweise mit piezoelektrischen, piezoresistiven, oder mit DMS-Aufnehmern gemessen werden. Für große Kräfte zwischen $5 \cdot 10^2$ und $5 \cdot 10^7$ N werden massive Verformungskörper in Form von Kraftmessdosen verwendet, in denen ein Hohlzylinder gestaucht oder gedehnt wird. Für kleinere Kräfte werden Kraftaufnehmer mit Biegebalken verwendet, mit denen man Kräfte bis herunter zu 10^{-3} N messen kann.

Ein anderer Typ von Kraftaufnehmern arbeitet mit Elementen aus magnetoelastischem Material, das den magnetostriktiven Effekt zeigt. **Magnetostriktion** tritt bei gewissen ferromagnetischen Legierungen z. B. 80%igem Nickel-eisen auf und bewirkt, dass ein magnetisch isotropes Material unter mechanischen Druck- oder Spannungseinwirkungen anisotrop wird. Dadurch nimmt bei positiver Magnetostriktion und einer Druckbelastung des Materials die

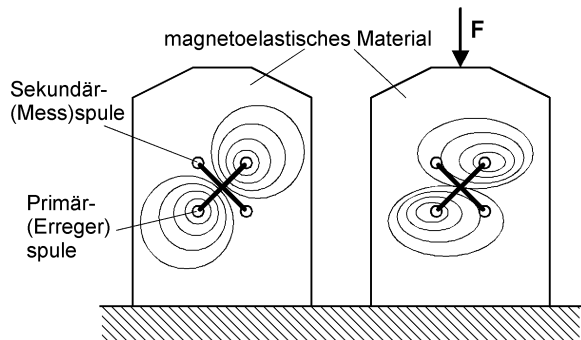


Bild 5-52: Kraftmessung mit magnetoelastischem Verformungskörper (Pressduktor)

zunächst homogene magnetische Permeabilität in Belastungsrichtung ab und senkrecht dazu steigt sie an. Die Permeabilitätsänderungen können durch eine Anordnung entsprechend Bild 5-52 [5.8] ausgewertet werden, die auch als **Pressduktor** bezeichnet wird. In entsprechendem magnetoelastischen Material sind zwei senkrecht zueinander stehende Spulen integriert. Die durch einen Wechselstrom gespeiste Primärspule erzeugt eine magnetische Induktion mit dem im linken Bildteil dargestellten symmetrischen Feldlinienverlauf. Der resultierende magnetische Flussanteil, der durch die von der Sekundärspule begrenzte Fläche durchtritt, hebt sich durch den symmetrischen Verlauf auf, so dass in der Sekundärspule keine Spannung induziert wird. Wie im rechten Bildteil zu sehen ist, wird der Feldlinienverlauf unter Einwirkung einer äußeren Druckkraft und der daraus resultierenden Anisotropie des Materials verzerrt. Dadurch verläuft ein Flussanteil durch die Sekundärspule und es wird eine Sekundärspannung induziert. Der Zusammenhang zwischen Druckkraft und resultierender induzierter Spannung weist eine gute Linearität auf.

Die Messung von Momenten ist natürlich auf die gleiche Art möglich wie die Kraftmessung. Verformungskörper werden den zu messenden Momenten unterworfen und die an ihnen auftretenden Verformungen werden beispielsweise mit DMS erfasst. Die Messung des Antriebsmomentes eines Gleichstrommotors lässt sich auch durch eine Messung des Ankerstroms vornehmen, da dieser dem Drehmo-

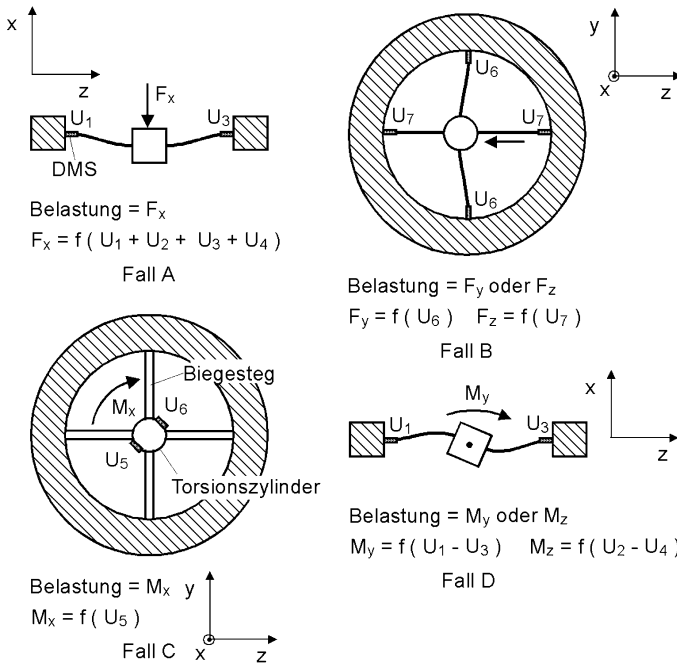


Bild 5-53: Kraft/Momenten-Sensor in Speichenradausführung

ment des Motors proportional ist.

Bei Handhabungs- und Bearbeitungsvorgängen mit Industrierobotern treten Reaktionskräfte und -momente auf, deren Kenntnis zur feinfühligsten Steuerung solcher Vorgänge benutzt werden kann. Dabei können Kräfte und Momente in bzw. um alle drei Richtungen eines kartesischen Koordinatensystems auftreten. Bild 5-53 zeigt einen Kraft/Momenten-Sensor in Speichenradausführung, mit dem man solche Belastungen erfassen kann [5.6]. Es erfolgt zunächst eine Umformung der angreifenden Kräfte

und Momente in kleine Längenänderungen der elastischen Stege, die durch DMS in elektrische Signale umgewandelt werden. Im Bild ist in vier Belastungsfällen gezeigt, welche Verformungsaufnehmer (DMS) jeweils Spannungsänderungen erzeugen und wie sich aus ihnen ein Messwert für die entsprechende Belastung ergibt. Andere Messeffekte zur Erfassung von Verformungen sind ebenfalls einsetzbar. Die Auflösung der Aufnehmersignale in die einzelnen Kraft/Momenten-Komponenten erfolgt mit Hilfe der so genannten Entkoppelungsmatrix:

$$\begin{pmatrix} F_x \\ F_y \\ F_z \\ M_x \\ M_y \\ M_z \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & k_{15} & k_{15} & k_{17} & k_{18} \\ 0 & k_{22} & 0 & k_{24} & 0 & 0 & 0 & 0 \\ k_{31} & 0 & k_{33} & 0 & 0 & 0 & 0 & 0 \\ k_{41} & k_{42} & k_{43} & k_{44} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & k_{55} & 0 & k_{57} & 0 \\ 0 & 0 & 0 & 0 & 0 & k_{66} & 0 & k_{68} \end{pmatrix} \times \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \\ U_5 \\ U_6 \end{pmatrix}$$

Mit Hilfe dieser Matrix kann das bauartabhängige, lineare Gleichungssystem aus den Kraft/Momenten-Komponenten gelöst werden. Dabei ist es günstig, wenn die Kräfte und Momente nur von wenigen Aufnehmersignalen abhängen.

Die gleichen Messprinzipien und Messaufbauten kann man natürlich auch für die **Druckmessung** verwenden, die in der Regel in flüssigen oder gasförmigen Medien stattfindet. Da ein Druck als eine auf eine Fläche bezogene Kraft definiert ist, lässt man den Druck auf eine im Sensor befindliche Membran bestimmter Fläche einwirken.

ken, die sich entsprechend verformt. Diese Verformung kann wieder mit DMS oder anderen Widerstandseffekten, oder durch piezoelektrische, kapazitive oder induktive Prinzipien ausgewertet werden. Bild 5-54 a) zeigt einen Querschnitt durch einen Drucksensor mit polykristallinen Piezowiderständen als Messaufnehmer [5.9] und im Bild 5-54 b) den gesamten, monolithisch integrierten Schaltkreis. Auf ihm sind zusätzlich noch bipolare Schaltkreise

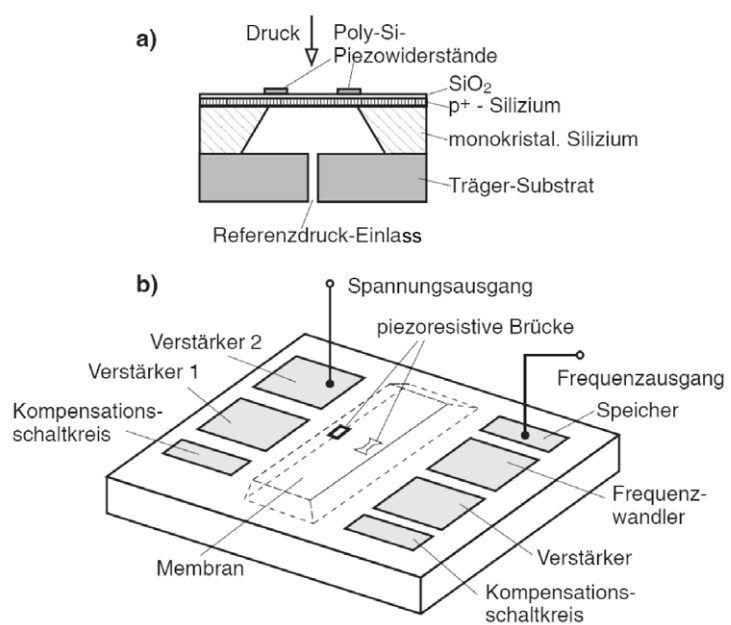


Bild 5-54: Silizium-Drucksensor mit piezoresistiver Signalumwandlung

a) Querschnitt b) schematische Ansicht

für eine spannungs- und frequenzanaloge Signalausgabe vorhanden. Das eigentliche Druckmesselement besteht aus einer allseitig eingespannten Membran, die sich verwölbt, wenn auf den beiden Seiten der Membranoberfläche eine Druckdifferenz auftritt. Die Membran besteht aus einer Siliziumfolie, die durch gezieltes Verdünnen des Siliziumwafers erzeugt wird. Auf der Membran ist eine Brücke aus piezoresistiven Widerständen aufgebracht, die bei Verformung der Membran ihren Widerstand ändern. Aufgrund der integrierten Elektronik kann der Sensor direkt ein digitales Ausgangssignal liefern.

6 Aktoren

Für die Erzeugung von Bewegungen, oder das Aufbringen von Kräften und Momenten, werden in mechatronischen Systemen Antriebe (Aktoren oder Aktuatoren) benötigt. Da sie häufig in lagegeregelten Antriebssystemen (Servosystemen) eingesetzt werden, kommen hier im wesentlichen drehzahl- oder geschwindigkeitsregelbare Antriebe in Frage. Diese Anforderung erfüllen die meisten elektrischen oder fluidischen Antriebe, auf denen das Hauptaugenmerk dieses Kapitels liegen wird. Der Begriff **Aktor** geht aber über den Begriff Antrieb (engl. actuator) hinaus und umfasst alle Arten von Ausgabeelementen für Bewegungen, Kräfte und Momente, die sowohl analog als auch binär wirken können. In Tabelle 6-1 ist eine Übersicht

Tabelle 6-1: Energiewandlungsprinzipien und darauf beruhende Aktoren

Energie	Elektrische Energie	Äußere elektrische und magnetische Feldkräfte	Elektromotor
			Elektromagnet
		Innere Molekularkräfte	Ferroelektrische Aktoren
			Magnetostriktive Aktoren
	Strömungsenergie	Pneumatisch erzeugte Druckkräfte	Überdruck-Stelleinrichtung
			Unterdruck-Stelleinrichtung
		Hydraulisch erzeugte Druckkräfte	Hydraulische Stelleinrichtung, Hydromotoren
	Thermische Energie	Wärmedehnung	Thermo-Bimetall
			Dehnstoff-Aktor
		Memoryeffekt	Einweg-Memory-Metall
			Zweiweg-Memory-Metall
	Chemische Energie	Elektrolysedruck	Elektrochemische Stelleinrichtung
		Explosionsdruck	Pyrotechnische Stelleinrichtung

über Energiewandlungsprinzipien und darauf beruhende Aktoren dargestellt. Eine umfassende Behandlung aller im Bild aufgeführten Aktorprinzipien geht über den Rahmen dieses Buches hinaus und kann hier nicht allzu sehr vertieft werden. Viele der aufgeführten Möglichkeiten, insbesondere beim Einsatz für mikromechanische Aktoren, sind inzwischen über das Laborstadium hinausgekommen bzw. in der Erprobungsphase. Diese werden daher wegen ihrer interessanten Einsatzmöglichkeiten in der Mechatronik behandelt, weit verbreitet und eingeführt sind jedoch im wesentlichen Aktoren, die mit elektrischer oder fluidischer Energie betrieben werden.

Wie in Bild 6-1 dargestellt, besteht ein Aktor immer aus zwei wesentlichen Komponenten, dem **Energiensteller**, der mit dem **Stellsignal** der Steuereinheit (Digitalrechner) des mechatronischen Systems beaufschlagt wird und dem **Energiewandler**, der die für die Bewegung oder Kraftausübung erforderliche Stellenergie liefert. Im Energiesteller wird dem Aktor **Hilfsenergie** zugeführt, die die eigentliche im Ak-

tor umgesetzte Energie liefert [6.2]. Energiesteller sind bei elektrischen Aktoren im einfachsten Fall ein Schalter (Relais) oder bei steuerbaren Aktoren mit Halbleitern bestückte Drehzahlsteller. Bei fluidischen Aktoren verwendet man analog einfache Schaltventile oder kontinuierlich im Durchfluss änderbare Servo- oder Proportionalventile. Diese Stelleinheiten werden jeweils zusammen mit den zugehörigen Wandlern (Motoren) im folgenden Abschnitt 6.1 behandelt.

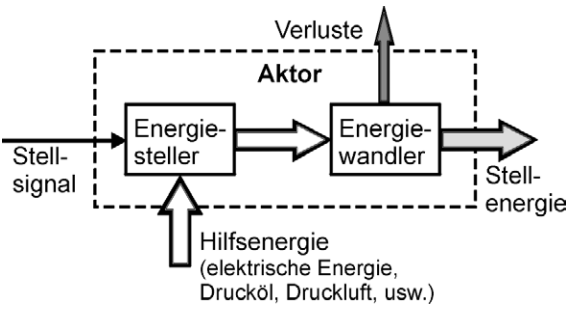


Bild 6-1: Strukturbild eines Aktors

Die Arbeit W , die zwischen den Zeitpunkten t_0 und t_1 vom Aktor geleistet wird, lässt sich folgendermaßen berechnen:

$$W = \int_{t_0}^{t_1} P(t) dt \quad .$$

Die von der Hilfsenergie eingebrachte Leistung P wird aber nie vollständig als Nutzarbeit abgegeben, sondern aufgrund von Verlusten, etwa durch Reibung bei mechanischen Systemen oder durch elektrische Verluste, in Wärme umgesetzt. Die aus dem Aktor abfließende Wärme Q , die bei hoher Leistungsdichte sogar durch aktive Kühlmaßnahmen entfernt werden muss, vermindert die zugeführte Arbeit:

$$W_{ab} = W_{zu} - Q \quad .$$

Dies gilt für den stationären Fall, in dem keine Energie im System gespeichert wird. Zur Charakterisierung der Verluste wird der Wirkungsgrad benutzt, der das Verhältnis zwischen abgegebener Nutzenergie W_{ab} und zugeführter Hilfsenergie W_{zu} angibt:

Tabelle 5-2: Einteilung drehzahl- und geschwindigkeitssteuerbarer Antriebe

Antriebe	rotierend	elektromagnetisch	Synchronmotor
			Asynchronmotor
			Gleichstrom
			Schrittmotor
	rotierend	hydraulisch	Flügelzellen-Motor
			Axialkolben-Motor
			Radialkolben-Motor
	linear	elektromagnetisch	Asynchron-Linearmotor
			Linear-Schrittmotor
		hydraulisch (pneumatisch)	Hydraulikzylinder

$$\eta = \frac{W_{ab}}{W_{zu}} .$$

Der Wirkungsgrad η ist eine wichtige Kenngröße für die Auswahl eines geeigneten Aktors. Dies ist vor allem für ortsbewegliche mechatronische Systeme von Bedeutung, da diese in der Regel aus Batterien gespeist werden.

Kontinuierlich arbeitende, drehzahlregelbare Antriebe, wie sie als Vorschubantriebe in Werkzeugmaschinen und Industrierobotern und auch in anderen Stellantrieben vorkommen, kann man wie in Tabelle 6-2 dargestellt klassifizieren:

Diese "klassischen Aktoren" und ihre Eigenschaften werden im nächsten Abschnitt behandelt werden. In einem weiteren Abschnitt werden dann einige neuartige Wirkprinzipien und darauf basierende "Neuartige Aktoren" behandelt werden.

6.1 Klassische Aktoren

6.1.1 Elektromotorische, rotierende Antriebe

In der Antriebstechnik werden zum größten Teil sich drehende elektrische Maschinen eingesetzt, die elektromagnetische Feldwirkungen als Antriebskraft nutzen. Andere Maschinentypen wie Linearmotoren, Piezomotoren oder elektrostatische Motoren werden relativ selten verwendet.

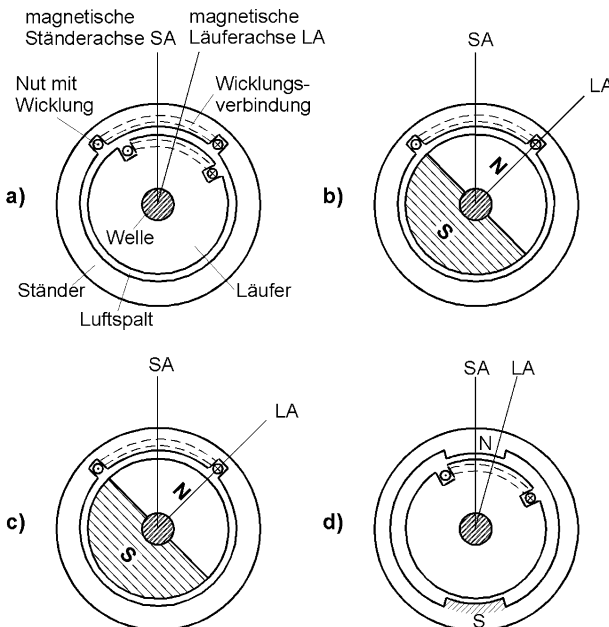


Bild 6-2: Prinzipieller Aufbau elektromechanischer umlaufender Energiewandler
a) Grundprinzip b) Synchronmaschine, Läufer als Permanentmagnet c) Gleichstrommaschine
d) Gleichstrommaschine mit Ständer als Permanentmagnet

Die Umsetzung elektrischer in mechanische Energie erfolgt in solchen Maschinen unter Ausnutzung elektromagnetischer Felder. Diese bewirken induzierte Spannungen oder Kräfte, die als Kräfte auf stromdurchflossene Leiter oder als Grenzflächenkräfte an den Trennflächen zwischen Gebieten verschiedener Permeabilität (z. B. Luft - Eisen) auftreten. Im ersten Fall erfahren dabei die Leiter einer Spule Kräfte im Feld eines Magneten oder einer anderen Spule, im zweiten Fall entstehen die Kräfte an einem ferromagnetischen Körper im Feld einer Spule.

Ein Energieaustausch tritt dann auf, wenn in Spulen Spannungen induziert werden, die zu Strömen führen. Dabei ist es gleichgültig, ob die Anordnung von Spule und Magnet im Stän-

der (feststehender Teil des Motors) und im Läufer (rotierender Teil des Motors) ist, oder auch umgekehrt. Bild 6-2 a) zeigt schematisch das Grundkonzept eines Elektromotors. Ständer und Läufer bestehen aus coaxialen ferromagnetischen Zylindern, in denen jeweils eine Spule untergebracht ist [6.1]. Speist man eine oder beide Wicklungen mit Wechselstrom, so hat man die Grundform der Drehfeldmaschinen, die als **Synchron-** oder **Asynchronmaschine** ausgeführt sein können. Bei der Synchronmaschine kann die Läuferwicklung auch durch einen Permanentmagneten ersetzt werden (Bild 6-2 b). Speist man Läufer- und Ständerwicklung mit Gleichstrom, so erhält man einen Gleichstrommotor. Da eine fortschreitende Drehbewegung des Motors aber nur mit Wechselstrom möglich ist, wird durch einen elektromechanischen Stromwender (Kommutator) oder durch elektronische Kommutierung (Stromwendung) die Stromrichtung fortschreitend umgekehrt. Die Läuferwicklung wird daher in Wirklichkeit auch mit Wechselspannung gespeist.

Wie in Bild 6-2 c) und d) dargestellt ist, kann man beim Gleichstrommotor entweder den Ständer oder den Läufer auch durch einen Permanentmagneten ersetzen. Im Fall c) muss dann die Ständerwicklung in Abhängigkeit von der Läuferstellung gespeist und im Fall d) die Läuferwicklung in Abhängigkeit von der Läuferstellung gespeist werden. Dies erfolgt jeweils so, dass durch die läuferstellungsabhängige Stromrichtung eine für fortschreitende Drehbewegung erforderliche Kraftwirkung zustande kommt.

Wird eine permanent erregte Synchronmaschine über einen **Frequenzumrichter** gespeist und taktet man die Steuerelektronik schrittweise weiter, erhält man eine Synchronmaschine mit Schrittschaltung, den sogenannten **Schrittmotor**.

Die wichtigsten mechanischen Kenngrößen solcher Motoren, das Drehmoment M und die Winkelgeschwindigkeit ω , haben den in Bild 6-3 gezeigten Zusammenhang. Synchron- und Schrittmotoren zeigen starres Verhalten, d. h. sie laufen bei vorgegebener Frequenz der Spannungsquelle immer mit der gleichen Winkelgeschwindigkeit. Arbeitet der Motor im **Nebenschluss** (Läufer- und Ständerwicklung werden unabhängig voneinander gespeist, oder einer der beiden ist als Permanentmagnet ausgeführt), so liegt ein steifes kaum nachgebendes Verhalten vor. Dies trifft auf Gleichstrommotore mit konstanter Erregung oder auf Asynchronmaschinen im Bereich zwischen Arbeitspunkt und Leerlauf zu. Es bedeutet, dass bei Vergrößerung des Lastmomentes und der dadurch erforderlichen Zunahme des Motormomentes, die Winkelgeschwindigkeit nur wenig abnimmt. Sind bei einem Gleichstrommotor beide Wicklungen hintereinander geschaltet, so liegt **Reihenschlussverhalten** vor. Die Maschine reagiert mit veränderlicher Drehzahlsteifigkeit und verhält sich daher sehr weich. Insbesondere bei hohen Drehzahlen führen Schwankungen des Lastmomentes zu starken Änderungen der Winkelgeschwindigkeit. Ein ähnliches Verhalten zeigt eine fremderregte Gleichstrommaschine im Feldstellbereich.

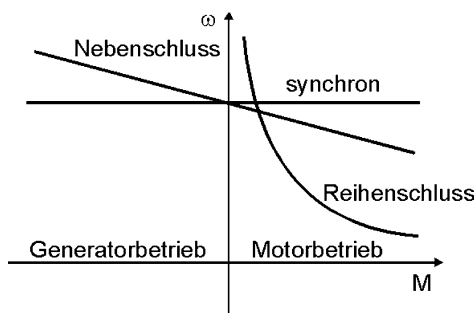


Bild 6-3: Grundtypen der Kennlinien von elektrischen Maschinen

Diese verschiedenartigen Maschinen und ihre Eigenschaften werden in den nächsten Abschnitten näher beschrieben.

6.1.1.1 Gleichstrommotoren

Gleichstrommotoren werden häufig als drehzahlregelbare Elektromotoren eingesetzt. Sie werden als Vorschubantriebe für Positionieraufgaben in Werkzeugmaschinen, Industrierobotern und anderen Vorrichtungseinrichtungen verwendet. Die für solche Aufgabenstellungen geforderte, stufenlose Drehzahlstellbarkeit über einen großen Bereich, hohe Drehzahlsteifigkeit (geringe Drehzahländerung unter Last im ungeregelten Fall), guter Gleichlauf und hohe Dynamik, werden von ihm in ausgezeichneter Form erfüllt.

Der allgemeine Fall des Gleichstrommotors liegt vor, wenn sich in Ständer und Läufer jeweils eine Spule befinden. Von den hier betrachteten Motoren sind für Servosysteme praktisch nur konstant erregte Maschinen (fremderregt durch separate Feldspannung an der Ständerwicklung oder Permanenterregung) von Bedeutung. Dies beruht auf deren hoher und konstanter Drehzahlsteifigkeit. Der prinzipielle Aufbau und das Schaltbild einer fremderregten Gleichstrommaschine zeigt Bild 6-4 [7.10].

Durch die Ständerwicklung fließt ein Strom I_F der ein magnetisches Feld mit dem magnetischen Fluss Φ erzeugt, das über den Luftspalt den Anker (Läufer) durch-

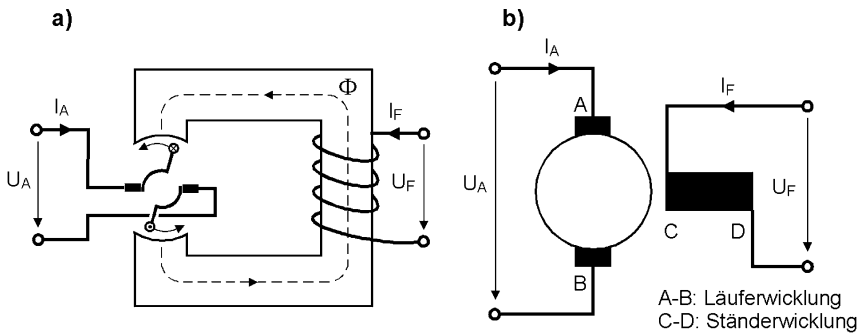


Bild 6-4: Fremderregter Gleichstrommotor a) prinzipieller Aufbau b) Schaltbild

dringt. Die im Bild 6-4 a) als einzelne Leiterschleife senkrecht zur Zeichenebene dargestellte Ankerwicklung wird gleichzeitig vom Strom I_A durchflossen. Im magnetischen Feld, das durch den Fluss Φ erzeugt wird, wirkt auf den stromdurchflossenen Leiter der Ankerwicklung eine **Lorentzkraft**, die am Anker ein Drehmoment in der angegebenen Richtung hervorruft. Die Lorentzkraft steht dabei senkrecht auf der Feld- und auf der Stromrichtung. Wenn die von der Leiterschleife der Ankerwicklung umschlossene Fläche in Richtung des Magnetfeldes liegt, wird die Lorentzkraft zu Null und der Anker würde in dieser Stellung verharren. Damit eine gleichbleibende Rotation zustande kommt, muss daher die Stromrichtung der Ankerwicklung nach einer halben Umdrehung des Läufers umgekehrt (kommutiert) werden. Dies geschieht durch den **Kommutator**, eine elektromechanische Schaltvorrichtung aus Schleifbahnen und Kohlebürsten, über die abwechselnd die Anker-

spannung umgepolt wird. Zur Erzeugung eines gleichmäßigen Drehmomentverlaufs sind bei realen Gleichstrommotoren mehrere Wicklungen (typabhängig bis zu 50) in Ständer und Läufer auf den Umfang verteilt.

Wichtig für das Verhalten des Motors ist noch, dass nach dem Induktionsgesetz in der im Magnetfeld bewegten Leiterschleife eine Spannung induziert wird, deren Richtung zur Ankerspannung entgegengesetzt ist. Der dadurch verursachte Stromfluss erzeugt die so genannte EMK (elektromotorische Kraft), die der Antriebskraft entgegengerichtet ist. Der Begriff der EMK ist heute veraltet und kann durch die induzierte Spannung

$$U_i = c_M \cdot \Phi \cdot \omega \quad (6.1)$$

ersetzt werden, wobei in Gl. 6.1 c_M eine Maschinenkonstante ist, die den Wert

$$c_M = \frac{z \cdot p}{2\pi \cdot i}$$

besitzt. Darin sind z die Ankerleitzahl, p die Polpaarzahl und i die Anzahl der parallelen Ankerzweigpaare.

Für den Ankerkreis kann man damit folgende Spannungsgleichung aufstellen:

$$U_A = c_M \cdot \Phi \cdot \omega + R_A \cdot I_A \quad , \quad (6.2)$$

wobei R_A der ohm'sche Widerstand der Ankerwicklung ist.

Um das vom Motor abgebbare Motormoment zu ermitteln, kann man eine Leistungsbilanz des Ankerkreises aufstellen:

$$U_A \cdot I_A = U_i \cdot I_A + R_A \cdot I_A^2$$

<i>zugeführte</i>	<i>abgegebene</i>	<i>aufretende</i>
<i>elektrische</i>	<i>mechanische</i>	<i>elektrische</i>
<i>Leistung P_{el}</i>	<i>Leistung P_{mech}</i>	<i>Verlustleistung P_v</i>

Die an die Welle des Motors abgegebene Leistung P_{mech} kann zur Berechnung des Motormomentes herangezogen werden:

$$M = \frac{P_{mech}}{\omega} = \frac{U_i \cdot I_A}{\omega} = \frac{c_M \cdot \Phi \cdot \omega \cdot I_A}{\omega} = c_M \cdot \Phi \cdot I_A \quad . \quad (6.3)$$

Beim **Anfahren** des Motors ist ω gleich Null, so dass nach Gl. (6.1) die Spannung auch Null ist. Da die induzierte Spannung der Ankerspannung entgegen gerichtet ist, muss nach Gl. (6.2) in diesem Zustand der höchste Ankerstrom I_{Amax} fließen:

$$\omega = 0: \quad I_{Amax} = \frac{U_A}{R_A} \quad . \quad (6.4)$$

Dementsprechend steht für das Anfahren aus dem Stillstand das größte Motormoment zur Verfügung, wobei für das Beschleunigen von Läufer und angekoppelter Last ja auch ein hohes Antriebsmoment erforderlich ist :

$$M_{\max} = c_M \cdot \Phi \cdot I_{A\max} = c_M \cdot \Phi \frac{I_A}{R_A} \quad (6.5)$$

Die induzierte Gegenspannung wächst mit steigendem ω proportional an, d. h. sie muss bei einer bestimmten Winkelgeschwindigkeit den Wert der angelegten Ankerspannung erreichen. Nach Gl. (6.2) muss dann der Ankerstrom den Wert Null annehmen, so dass das zur Verfügung stehende Motormoment ebenfalls Null sein muss. Daher beschleunigt der Motor nicht mehr, sondern verharrt bei dieser Winkelgeschwindigkeit, er ist im **Leerlauf**.

Im Betriebsbereich zwischen diesen beiden Extremfällen gilt für die Winkelgeschwindigkeit:

$$\omega = \frac{U_A}{c_M \cdot \Phi} - \frac{R_A}{(c_M \cdot \Phi)^2} M = \omega_0 - \Delta\omega \quad (6.6)$$

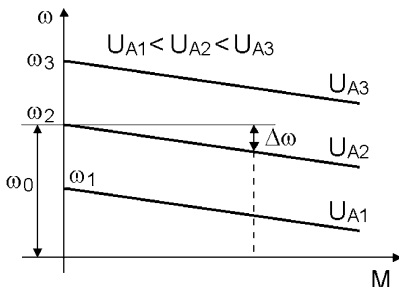


Bild 6-5: Kennlinien des Gleichstrommotors in Abhängigkeit von der Ankerspannung

Darin stellt der erste Summand die Leerlaufwinkelgeschwindigkeit ω_0 und der zweite die Winkelgeschwindigkeitsänderung bei Belastung $\Delta\omega$ dar. Dieser Zusammenhang ist in Bild 6-5 dargestellt, in dem für verschiedene Ankerspannungen und konstante Erregung ($\Phi = \text{const.}$) die Motorkennlinien (Zusammenhang zwischen Motormoment und Winkelgeschwindigkeit) eingetragen sind.

Wie man anhand der Gl. (6.6) erkennen kann, gibt es zwei Möglichkeiten die Winkelgeschwindigkeit oder Drehzahl des Gleichstrommotors zu verstellen. Eine Erhöhung der Ankerspannung U_A (**Ankerverstellung**) führt zu einer Erhöhung der Winkelgeschwindigkeit, eine Erhöhung des magnetischen Flusses Φ (**Feldverstellung**) führt zu einer Verminderung der Winkelgeschwindigkeit. Eine Änderung der Winkelgeschwindigkeit über einen veränderlichen Vorwiderstand im Ankerkreis scheidet wegen der Verluste weitgehend aus. In den beiden Verstellbereichen ergibt sich jeweils ein unterschiedliches Betriebsverhalten, das anhand der Veränderung verschiedener Kenngrößen in Abhängigkeit von der Winkelgeschwindigkeit ω in Bild 6-6 dargestellt ist. Beide Verstellmöglichkeiten werden in der Regel nicht gleichzeitig ausgenutzt, so dass bei Verstellung der einen die andere stets konstant bleibt. Im Bildteil 6-6 a) sind, entsprechend den Verstellbereichen, die beiden Einstellgrößen über der Winkelgeschwindigkeit aufgetragen. In Teilbild b) ist der Verlauf von Drehmoment M , Leistung P und zulässigem Ankerstrom I_A in Abhängigkeit von ω aufgetragen. Im einzelnen können dem Bild die Kennlinienschar $M(\omega)$ und die Grenzkennlinie für das maximale Motormoment $M_{\max}$, sowie die aufgenommene, maximale elektrische Leistung $P_{\max}$ und die an der Motorwelle zur Verfügung stehende maximale Leistung P_{mech} entnommen werden.

Der maximal zulässige Dauerankerstrom $I_{A\max}$ bei einer Einschaltdauer $ED=100\%$, der durch die thermische Belastbarkeit der Wicklung begrenzt ist, ist in beiden Verstellbereichen als konstant anzunehmen. Kurzzeitige Überlastungen sind je nach

Wärmekapazität des Läufers und der Umgebungstemperatur der Wicklung möglich.

Im **Ankerverstellbereich**, der bei kleinen und mittleren Winkelgeschwindigkeiten liegt, wird der maximale magnetische Fluss $\Phi_{\max}$ verwendet und konstant gehalten. Dadurch verlaufen die Momentenkennlinien $M(\omega)$ sehr steil, was gleichbedeutend mit hoher Drehzahlsteifigkeit ist. Die Steigung bleibt auch im gesamten Ankerverstellbereich konstant. Übliche Werte für den Bereich der Verstellung der Ankerspannung sind $U_{A\min}/U_{A\max} = 1/10$. Die Grenze für die maximale Ankerverspannung ist durch die Spannungsfestigkeit der Wicklungsisolation gegeben. Da magnetischer Fluss und maximaler Ankerstrom konstant sind, muss nach Gl. (6.5) das maximale Motormoment $M_{\max}$ ebenfalls konstant sein. Da für die maximal vom Motor aufgenommene Leistung gilt

$$P_{\max} = U_A \cdot I_{A\max} \quad ,$$

muss nach Gl. 6.2 diese Leistung unter Berücksichtigung der konstanten Größen linear mit ω zunehmen. Die an der Motorwelle zur Verfügung stehende mechanische Leistung P_{mech} , ist die um die elektrischen Verluste P_V verminderte elektrische Leistung. Die Verluste sind von ω unabhängig und werden in Wärme umgesetzt. Verluste durch Lager- und Luftspaltreibung, sowie Antriebsleistung für Lüfter zur Motor Kühlung, sind dabei nicht berücksichtigt.

Will man noch höhere Winkelgeschwindigkeiten erreichen, so besteht die Möglichkeit, bei maximaler Ankerverspannung das Feld zu verstellen. Um dies zu erreichen, muss der magnetische Fluss Φ im **Feldverstellbereich** geschwächt werden. Dies folgt unmittelbar aus Gl. (6.2), wenn man sie nach der Winkelgeschwindigkeit umstellt:

$$\omega = \frac{1}{c_M \cdot \Phi} (U_{A\max} - R_A \cdot I_{A\max}) \quad .$$

Man sieht, dass die Winkelgeschwindigkeit ω umgekehrt proportional zum magnetischen Fluss Φ ist und dass daher zur Vergrößerung der Winkelgeschwindigkeit der Fluss verkleinert werden muss. Übliche Werte für den Bereich der Feldverstellung sind $\Phi_{\min}/\Phi_{\max} = 1/3$. Die Folge dieser Schwächung des Erregerfeldes ist, dass die Steilheit der Kennlinien $M(\omega)$ abnimmt und die Drehzahlsteifigkeit des Motors nachlässt. Da nach Gl. (6.5) das maximale Motormoment dem Fluss proportional ist, nimmt dieses bei Feldschwächung entsprechend ab. Die vom Motor abgebbare mechanische Leistung bleibt konstant, da die Winkelgeschwindigkeit im gleichen

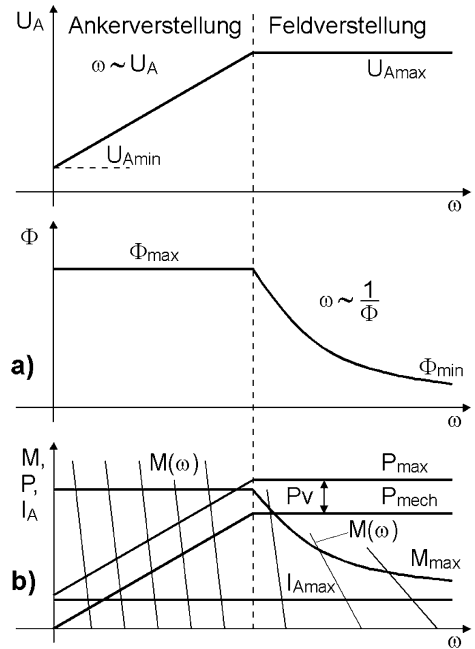


Bild 6-6: Anker- und Feldverstellbereich eines Gleichstrommotors
a) Bereichsdefinition b) Motor-kenngrößen in Abhängigkeit von ω

Maß zunimmt wie der Fluss abnimmt.

Da die beschriebenen Motoreigenschaften bei Feldschwächung (geringere und abnehmende Drehzahlsteifigkeit) für den Einsatz als dynamischer Vorschubantrieb eher ungünstig sind, wendet man sie in diesem Bereich heute nicht mehr an und geht eher auf eine konstante Felderzeugung durch Permanentmagnete aus Barium-Ferrit oder Samarium-Kobalt über. Dadurch vereinfacht sich der Motoraufbau auch erheblich. Die Verstellung der Winkelgeschwindigkeit eines solchen Motors, der in Bild 6-7 a) schematisch dargestellt ist, erfolgt ausschließlich durch Veränderung der Ankerspannung, die über Bürsten und Schleifkontakte den Ankerwicklungen zugeführt wird.

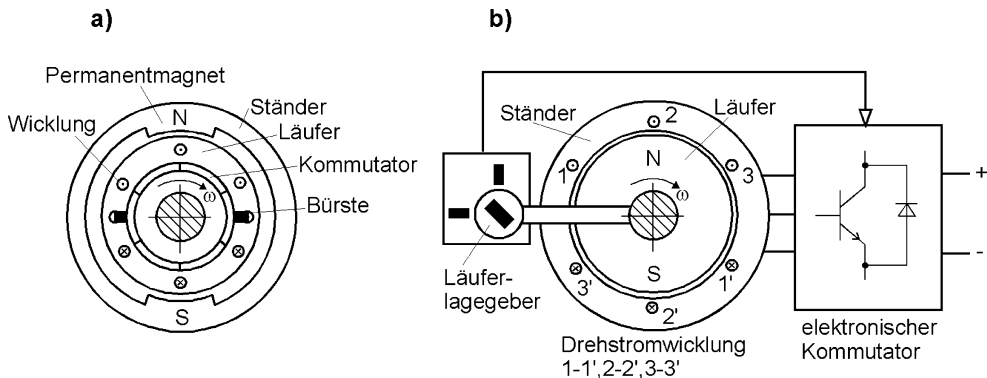


Bild 6-7: Gleichstrommotor mit Permanentmagneten a) permanente Felderregung und mechanische Kommutierung des Ankerstroms b) Permanentmagnetläufer und elektronisch kommutierte Ständerwicklung

Ein grundsätzlicher Nachteil dieses Motors ist die sog. **Kommutierungsgrenze**, bei der unzulässiges Kommutierungsfeuer an den Bürsten entsteht, das die Lebensdauer des Motors drastisch reduziert. Der Grund dafür liegt in der relativ großen Ankerinduktivität und hat eine Einschränkung der dynamischen Eigenschaften bei hohen Drehzahlen zur Folge. Die Kommutierungsgrenze wird meist durch eine Stromwendungsgrenzkurve vorgegeben, für die gilt:

$$\omega \cdot I_A = \text{const.} ,$$

d. h. mit zunehmender Drehzahl muss der maximal zulässige Ankerstrom zurückgenommen werden. In Bild 6-8 a) ist der Bereich dargestellt, in dem sich Moment und Drehzahl bewegen können und der einerseits durch das maximale Drehmoment und die maximale Drehzahl und andererseits durch die Stromwendungsgrenzkurve begrenzt ist. Diese muss von einer Höchstdrehzahl n_1 ab, bis zu der das maximale Drehmoment abgegeben werden kann, angesetzt werden. Der umfassendste Betriebsbereich ist nur für den dynamischen Betrieb anwendbar, in dem das maximale Drehmoment meist nur kurzzeitig gefordert wird. Will man ein bestimmtes Drehmoment dem Motor über längere Zeiten abverlangen, so ist der Betriebsbereich noch weiter eingeschränkt. Im Bild 6-8 a) sind ebenfalls Bereiche für den **Aussetzbetrieb** und den **Dauerbetrieb** eingezeichnet, die das maximale Drehmoment nach oben begrenzen. Die Kenngrößen S1 und S3 stehen dabei für die in VDE-Bestimmungen festgelegte Nennbetriebsarten, wobei neun verschiedene Betriebsarten S1

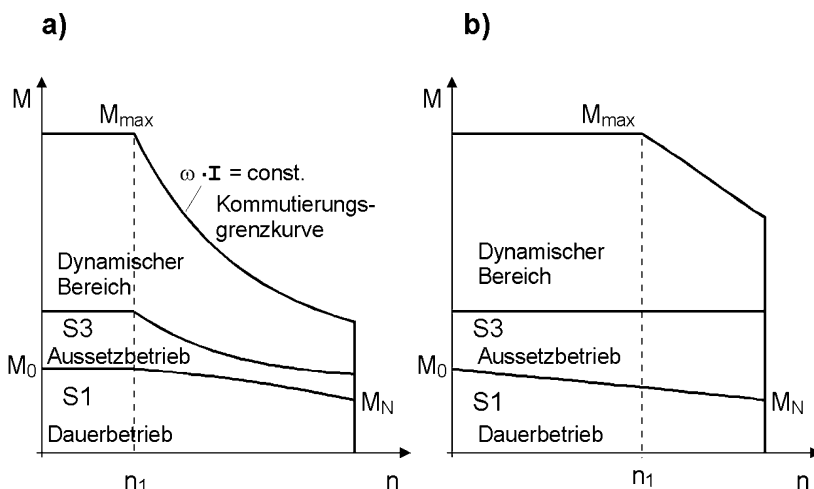


Bild 6-8: Betriebsbereiche von Gleichstrommotoren a) Betriebsbereich bei mechanischer Kommutierung b) Betriebsbereich eines *bürstenlosen*, elektronisch kommutierten Motors

- S9 unterschieden werden. Die im Bild genannten Betriebsarten sind folgendermaßen definiert:

- S1: **Dauerbetrieb** mit konstantem Belastungszustand, dessen Dauer ausreicht, den thermischen Beharrungszustand zu erreichen (Betriebsdauer t_B größer als die dreifache thermische Zeitkonstante T).
- S3: **Aussetzbetrieb ohne Einfluss des Anlaufvorgangs**, der sich aus einer Folge gleichartiger Lastspiele zusammensetzt, von denen jedes eine Zeit, mit konstanter Belastung und einer Pause umfasst, wobei der Anlaufstrom die Erwärmung kaum beeinflusst (Anlaufzeit t_A sehr klein gegen T).

Demnach hängt das maximale Drehmoment eines Gleichstrommotors auch stark von der Betriebsart ab. Lange Perioden einer gleichen Drehmomentanforderung bedingen ein geringeres maximales Drehmoment als im dynamischen Betrieb. Bei längerem Stillstand eines Motors mit mechanischer Stromwendung unter Last, besteht darüber hinaus die Gefahr der punktförmigen Überhitzung des Kollektors. Dies ist nur zu vermeiden, indem man den maximalen Ankerstrom und damit das Drehmoment reduziert.

Die Probleme, welche die mechanische Stromwendung aufwirft, kann man zum größten Teil durch Verwendung von bürstenlosen, elektronisch kommutierten Gleichstrommotoren vermeiden. Ein solcher Motor ist schematisch in Bild 6-7 b) dargestellt. Er arbeitet prinzipiell genauso wie der mechanisch kommutierte Gleichstrommotor, nur sind bei ihm stromdurchflossene Wicklung und Permanentmagnet miteinander vertauscht. Die Wicklungen sind im Ständer angeordnet und der Läufer besteht aus Permanentmagneten.

Um einen konstanten Kraftfluss zu erzeugen, wird die Lage des Läufers durch einen Positionsgeber festgestellt und dadurch die Drehung des Ankerfeldes mit der elektronischen Kommutierung synchronisiert. Die elektronische Schaltung kehrt die Stromflussrichtung durch eine der Ständerwicklungen, in Abhängigkeit von der La-

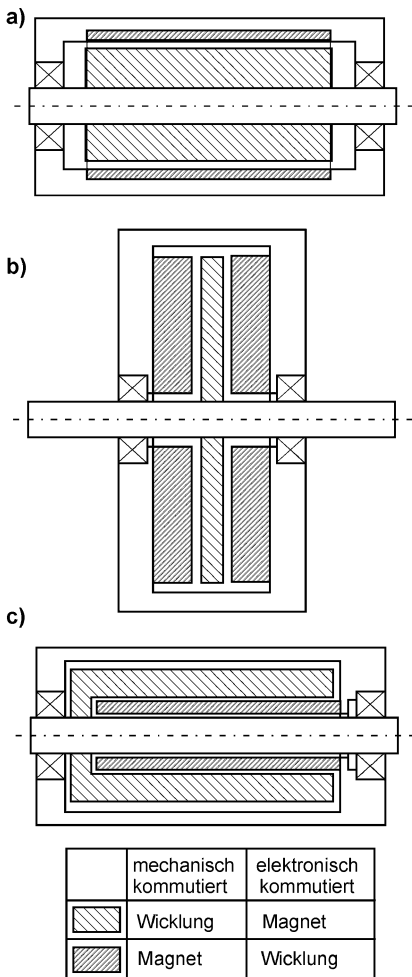


Bild 6-9: Verschiedene Läuferformen von Gleichstrommotoren

- a) Stabläufer
- b) Scheibenläufer
- c) Glockenläufer

ge des Läufers, im richtigen Moment um. Damit kann der mechanische Stromwender mit all seinen Nachteilen entfallen.

Da keine Begrenzung des Betriebsbereichs durch die Kommutierung besteht, können diese Motoren bis fast zur halben Nenndrehzahl mit dem maximalen Drehmoment belastet werden (Bild 6-8 b). Erst dann setzt eine Begrenzung durch die elektrischen Kenngrößen der Wicklung ein. Solche extrem robusten und wartungsfreien (kein Wechsel der Kohlebürsten) Motoren eignen sich hervorragend für den Einsatz in Industrierobotern, nicht zuletzt wegen ihrer guten Impulsfestigkeit.

Wichtig für das dynamische Verhalten des Motors ist, neben der Fähigkeit kurzzeitig starke Strombelastungen der Wicklung (verbunden mit hohem Drehmoment) gut vertragen zu können, das Trägheitsmoment J des Motors. Da nach Gl. 2-7 folgendes gilt:

$$\dot{\omega} = \frac{M_B}{J} ,$$

wobei M_B das Beschleunigungsmoment (stromabhängig) und J das Massenträgheitsmoment von Läufer und Welle sind, ist ebenfalls ein möglichst niedriges Trägheitsmoment für eine hohe Winkelbeschleunigung erforderlich. Da das Trägheitsmoment eines Zylinders (Motorläufer) sich wie folgt berechnet

$$J = \frac{m}{2} R^2 ,$$

kann man, sowohl durch Verkleinerung der Masse m , als auch durch Verkleinerung des Läuferradius R , das Trägheitsmoment verringern. Da der Läuferradius sogar in der zweiten Potenz in das Massenträgheitsmoment eingeht, ist ein hochdynamischer Servomotor beispielsweise wie in Bild 6-9 a) aufgebaut. Ein solcher **Stabankermotor** hat einen langgestreckten Rotor mit geringem Durchmesser und entsprechend geringem Massenträgheitsmoment. Man kann diese Bauform sowohl bei mechanisch als auch bei elektronisch kommutierten Motoren anwenden.

Um einen guten Rundlauf des Motors auch bei niedrigen Drehzahlen ($0,1 - 0,5 \text{ min}^{-1}$) zu erreichen, kann man beim mechanisch kommutierten Motor geschrägte Anker-nuten und einen fein unterteilten Kommutator verwenden. Diese guten Rundlaufeigenschaften erreicht man bei bürstenlosen Gleichstrommotoren meist nicht, da die Auflösung des Rotorlagegebers als Kompromiss zwischen Langsam- und Schnell-

lauf geringer ist. Bild 6-10 [3.5] zeigt im Vergleich zwei Servomotoren gleicher Leistung in Stabankerbauweise mit konventioneller und elektronischer Kommutierung. Das geringere Bauvolumen des bürstenlosen Gleichstrommotors rührt daher, dass Kupferverluste im Rotor entfallen und deshalb der Motor stark überlastbar (zwei bis dreimal mehr als der Bürstenmotor) ist. Das Massenträgheitsmoment des Magnetrotors ist außerdem erheblich geringer als der eines gewickelten Rotors.

Auch die übrigen in Bild 6-9 gezeigten Bauformen zeichnen sich durch eine sehr gute Dynamik aus und werden in den Versionen mit mecha-

nischer und elektronischer Kommutierung hergestellt. Bei der Bauform in Bild 6-9 b) ist der Läufer als dünne Scheibe ausgeführt. Außerdem ist diese Scheibe sehr leicht, da sie kein Eisen enthält, sondern aus Isoliermaterial besteht (glasfaserverstärktes Kunstharz), auf der sich bei der Bauform mit Permanentmagnetenerregung die Läuferwicklung aus massivem Kupferdraht befindet. Durch die geringe Läufermasse ist das Trägheitsmoment gering und durch den großen Scheibendurchmesser das Drehmoment aufgrund des langen Hebelarms groß. Der vielfach geteilte Kollektor, der oft als Flachkollektor mit axial angeordneten Bürsten ausgeführt wird, ermöglicht die Einleitung sehr hoher Kurzzeitströme zum Beschleunigen und Bremsen. Die Einleitung dieser dynamischen Stromspitzen ist auch bei Nenndrehzahl zulässig. Dies hat nahezu konstante Beschleunigung bis zur Nenndrehzahl zur Folge. Man erreicht ohne weiteres mechanische Zeitkonstanten von 5 msec.

Aufgrund der im Läufer angeordneten Wicklung, die auf eine Kunstharzscheibe aufgebracht ist, ist die Wärmeabfuhr natürlich nicht besonders gut. Das schon beim Stabläufer erläuterte Prinzip der elektronischen Kommutierung, bei dem die Wicklung im Ständer angeordnet ist, führt zu einer deutlich verbesserten Wärmeabfuhr über die Statorbleche an das Gehäuse. Dadurch sind bürstenlose **Scheibenläufermotoren** noch wesentlich robuster und höher überlastbar.

Der in Bild 6-9 c) dargestellte **Glockenläufermotor** ist genauso aufgebaut wie Scheibenläufermotoren, nur ist der Rotor zylindrisch ausgeführt, was zu noch geringeren Trägheitsmomenten und einer schlanken Bauart führt. Die elektrischen und dynamischen Eigenschaften, wie Überlastbarkeit, Drehzahlbereich, konstante Be-

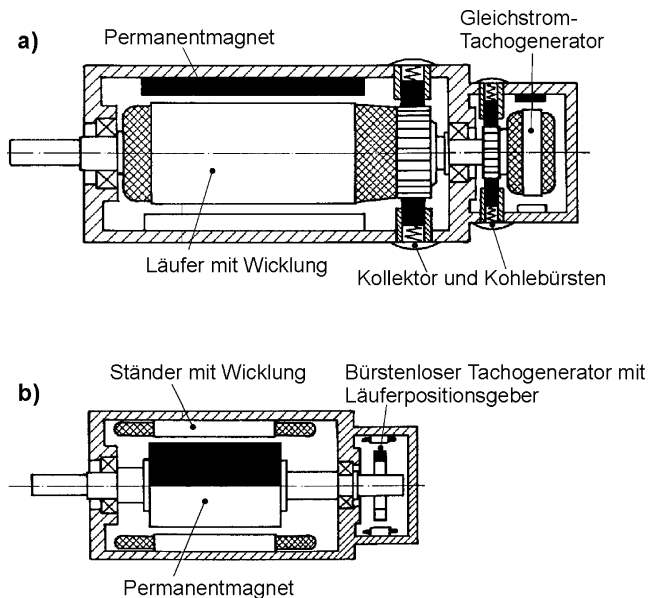


Bild 6-10: Vergleich zweier Stabläufer-Gleichstrommotoren mit mechanischer und elektrischer Kommutierung von gleicher Antriebsleistung a) mechanische b) elektrische Kommutierung

schleunigung über den ganzen Drehzahlbereich usw., sind mit denen von Scheibenläufern vergleichbar. Während jedoch Stab- und Scheibenläufermotoren bis in den Leistungsbereich von einigen Kilowatt gebaut werden, ist die Glockenläuferbauform nur mit Höchstleistungen bis ca. 100 W verfügbar.

Weitere Bauformen sind **Langsamläufer** und **Torque-Motoren**, die speziell für den Langsamlaufbereich ausgelegt sind und dort auch hohe Drehmomente abgeben können. Sie werden beispielsweise als **Direktantriebe** ohne Zwischengetriebe für Industrieroboter eingesetzt.

Zum Betrieb von kontinuierlich drehzahlverstellbaren Gleichstromantrieben benötigt man eine kontinuierlich verstellbare Gleichspannungsquelle. Da das überall verfügbare elektrische Netz Wechselspannung führt, muss die elektrische Energie erst durch eine entsprechende Leistungselektronik in die für den Betrieb des Motors verwendbare einstellbare Gleichspannung umgewandelt werden. Dafür sind im wesentlichen zwei verschiedene Schaltungsmöglichkeiten gebräuchlich. Dies sind:

- **netzgeführte Stromrichtergeräte** in verschiedenen Schaltungen (Bild 6-11 a) und
- **Steller** (Chopper) mit **Gleichspannungszwischenkreis** (Bild 6-11 b).

Ein Stromrichter besteht aus einer Anzahl von Stromventilen (Dioden, Thyristoren), die die Aufgabe haben, Strom mit nur einer Flussrichtung, also Gleichstrom, zu erzeugen. "Netzgeführt" bedeutet, dass die Kommutierung des Stromrichters, d. h. der Stromübergang von einem Ventil auf das nächste, von der Spannung des speisenden Netzes bestimmt wird.

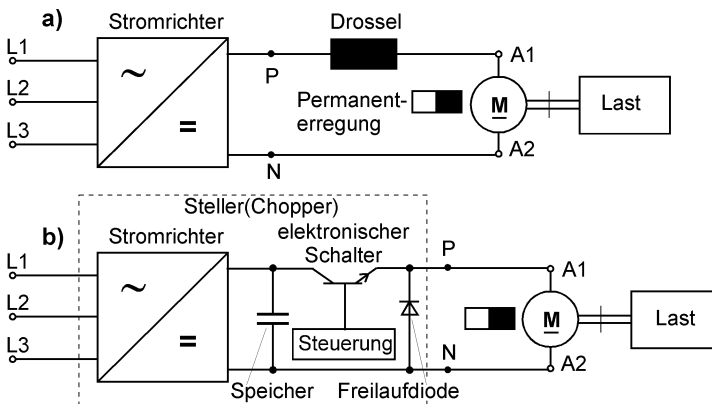


Bild 6-11: Prinzipschaltungen für den Betrieb von Servomotoren am Netz a) netzgeführter Stromrichter b) Steller (Chopper) mit Gleichspannungszwischenkreis

Man kann einen Stromrichter mit **Dioden** (nichtsteuerbare Ventile) oder mit **Thyristoren** (steuerbare Ventile) betreiben. Da für einen drehzahlregelbaren Servomotor die Gleichspannung in ihrer Höhe verstellbar sein muss, kommen für diesen Zweck nur steuerbare Ventile in Frage. Eine Diode ist eine Paarung von zwei unterschiedlich dotierten Halbleiter-

materialien, die nur dann einen Stromfluss durch die Diode ermöglichen, wenn die Anoden-Kathodenspannung U_{AK} positiv ist. Ein Thyristor hat prinzipiell ein Verhalten wie eine Diode, besitzt aber zusätzlich eine Steuerelektrode, über die der leitfähige Zustand mittels eines Zündimpulses erst eingeschaltet werden muss. Ist U_{AK} positiv, so führt der Thyristor erst vom Zeitpunkt des Zündimpulses ab Strom und der Stromfluss bleibt so lange erhalten, bis U_{AK} negativ wird. Ein netzgeführter

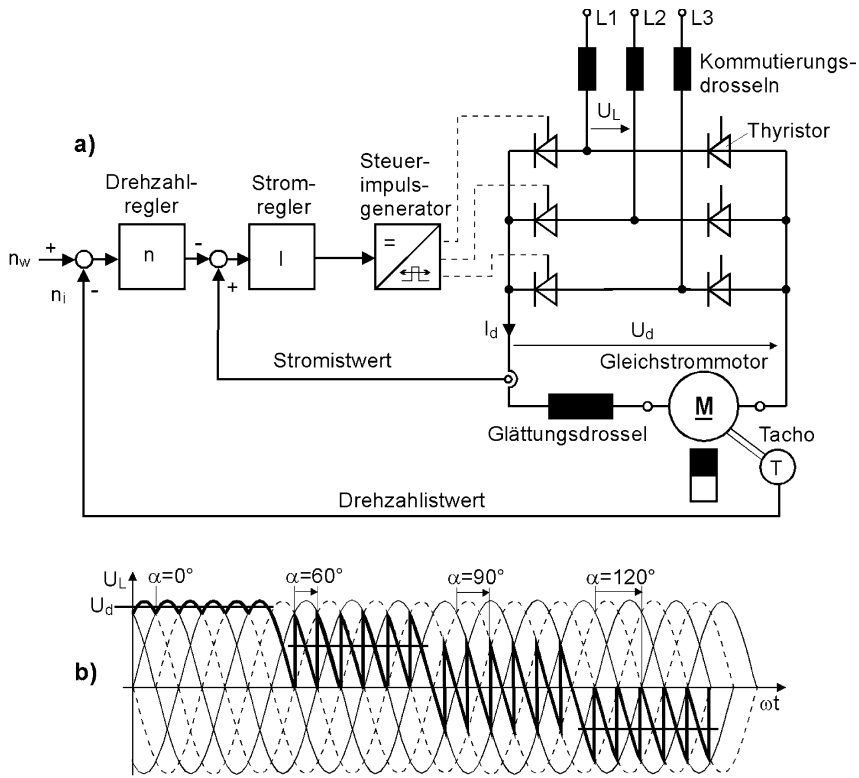


Bild 6-12: Drehzahlregelkreis für einen Gleichstrommotor mit vollgesteuerter Drehstrom-Thyristorbrücke a) Schaltbild b) Spannungsverlauf bei verschiedenen Phasenanschnittswinkeln

Stromrichter wird mit solchen in einer Brücke zusammen geschalteten Thyristoren am Wechsel- oder Drehstromnetz betrieben (Bild 6-12 a)). Durch die Brückenschaltung werden beide Halbwellen der Wechselspannung ausgenutzt, da bei jeder Halbwelle einer der beiden Stränge der Brückenschaltung stromführend ist. Ohne Zündimpulse sperren alle Thyristoren, so dass keine Spannung am Motor liegt. Der Zeitpunkt der Zündung der Thyristoren wird in Abhängigkeit vom Nulldurchgang (Netzführung) der Wechselspannung oder dem Schnittpunkt zweier Strangspannungen der Drehspannung gesteuert. Der Steuerimpulsgenerator erzeugt nach Durchlaufen des Phasenwinkels α , gerechnet vom Zeitpunkt der Netzsynchronisierung, den Zündimpuls, der den Stromfluss durch den Thyristor auslöst (Bild 6-12 b)). Diese Betriebsweise wird als **Phasenanschnittsteuerung** und der Steuerwinkel α als **Phasenanschnittwinkel** bezeichnet. Ist wie im Fall der Phasenanschnittsteuerung einer Drehstromversorgung $\alpha = 0^\circ$ (gerechnet vom Schnittpunkt der Strangspannungen), so liegt am Motor die maximale Spannung. Wird der Steuerwinkel auf beispielsweise $\alpha = 60^\circ$ erhöht, so werden die Thyristoren erst nach Auftreten des Spannungsmaximums der jeweiligen Strangspannung gezündet, wodurch der Thyristor die meiste Zeit keinen Strom führt. Dadurch wird die nutzbare Spannungs-Zeit-Fläche verringert und der Gleichspannungsmittelwert U_d herabgesetzt. Die dabei auftretenden großen Lücken in der Spannungsführung werden

durch die Ankerinduktivität und eine zusätzlich vor den Motor geschaltete Glättungsdrossel so überbrückt, dass eine mehr oder weniger wellige, mittlere Gleichspannung (Bild 6-12 b)) am Motoranker ansteht. Durch weiteres Erhöhen des Phasenanschnittwinkels kann dann die Motorspannung umgekehrt werden, d. h. die Gleichstrommaschine arbeitet dann im Generatorbetrieb (Nutzbremsung) und kann elektrische Leistung ins Netz zurückspeisen (2-Quadranten-Antrieb). Will man einen vollständigen 4-Quadranten-Antrieb realisieren (motorisch Beschleunigen und generatorisch Bremsen in beiden Drehrichtungen), so muss man zu jedem Thyristor der Brückenschaltung in Bild 6-12 a) einen weiteren Thyristor antiparallel schalten und entsprechend mit Zündimpulsen ansteuern. Wird der Gleichstrommotor in einem Drehzahlregelkreis eingesetzt (Bild 6-12 a)), so erhält der Steuerimpulsgenerator seine Stellsignale aus dem Regler des unterlagerten Stromregelkreises.

Aufgrund des Phasenanschnitt-Steuerungsprinzips arbeitet ein Drehzahlregelkreis mit netzgeführtem Stromrichter mit einer systembedingten nicht unerheblichen Totzeit. Die immer häufiger eingesetzten **Gleichstromsteller** (Chopper) arbeiten wie in Bild 6-11 b) dargestellt und haben eine viel geringere Totzeit, was die Dynamik des Antriebssystems deutlich verbessert. Sie werden auch in batteriebetriebenen Antrieben (Flurfördergeräte, autonome Roboter) eingesetzt, um die vom Ladezustand abhängenden Schwankungen der Batteriespannung auszugleichen oder die "gesparte" Verlustenergie, die sonst im Vorwiderstand umgesetzt würde, für den Antrieb zu nutzen.

Das gezeigte Prinzipschaltbild eines Stellers für den 1-Quadranten-Betrieb (Bild 6-11 b)) enthält drei wichtige Funktionsblöcke:

- den Eingangsstromrichter mit dem Anpasstransformator, der bei ortsveränderlichen Geräten durch eine Batterie ersetzt wird,
- den Energiespeicher (Zwischenkreiskondensator),
- den Maschinenstromrichter mit einem elektronischen Schalter und einer Freilaufdiode und der Steuer- und Überwachungseinheit.

Um die Ausgangsspannung zu vermindern, werden bei kleinen Leistungen auch rein analog arbeitende Gleichstromsteller eingesetzt. Der Längstransistor arbeitet dann als kontinuierlich verstellbarer Vorwiderstand. Die Verluste bei stetigem Steuerbetrieb werden jedoch sehr groß, wodurch der Wirkungsgrad des Antriebs entsprechend sinkt. Normalerweise wird daher die Ausgangsspannung über gepulst betriebene Transistorsteller erzeugt. Bei der Pulsung wird entweder die Pulsfrequenz f_p bei variabler Einschaltdauer t_e des Schalters, oder die Einschaltzeit bei variabler Periodendauer T konstant gehalten.

Die Drehzahl des Gleichstrommotors hängt vom Mittelwert der Speisegleichspannung ab. Über mehr oder weniger volle Durchsteuerung eines analog betriebenen Längstransistors, oder über seine Pulspausensteuerung (t_e/T) beim getakteten Steller, kann dieser Mittelwert und damit die Drehzahl kontinuierlich zwischen 0 und dem Maximalwert verstellt werden (Bild 6-13). Die Periodendauer $T = t_e + t_a$ wird in diesem Beispiel konstant gehalten (t_a : Ausschaltzeit) und durch Variation der Zeitanteile die Höhe des Spannungsmittelwertes U_{mittel} beeinflusst. Diese Art der Steuerung wird als **Pulsbreitensteuerung** bezeichnet. Der Wert der mittleren Ausgangs-

spannung hängt von der Speisegleichspannung U_z und einem Pulsbreitenverhältnis λ ab:

$$U_{\text{mittel}} = \lambda \cdot U_z \quad \text{mit} \quad \lambda = \frac{t_e}{t_e + t_a}$$

In der Durchschaltphase des Längstransistors in Bild 6-13 b) liegt die volle Gleichspannung $U_A = U_z$ des Zwischenkreises am Gleichstrommotor (Bild 6-14). Es baut sich ein Strom I_A durch den Anker auf, der dem Energiespeicher (Kondensator) des Zwischenspeichers entnommen wird (I_z). Schaltet der Transistor ab, fließt der Ankerstrom, getrieben durch die Induktivität der Ankerwicklung, als Freilaufstrom I_F im Kreis aus Motorwicklung und Freilaufdiode. Dabei ist die Spannung U_A am Motor Null, das Drehmoment behält jedoch seinen vollen Wert, da sich, wie in Bild 6-14 dargestellt, der Ankerstrom aus der Summe der Ströme I_z und I_F ergibt. Da die Schaltfrequenz im kHz-Bereich (1-20 kHz) liegt, ist die geringe Ankerinduktivität meist zur Stromglättung ausreichend. Dadurch ergeben sich sehr geringe Zeitkonstanten für den dynamischen Betrieb und durch die sehr geringe Totzeit sind die regelungstechnischen Eigenschaften in einem Drehzahlregelkreis sehr gut.

Fügt man den Motor in eine Brückenschaltung aus vier elektronischen Schaltern ein, so ist auch der Generatorbetrieb möglich und damit eine motorische Bremsung, sowie eine Drehrichtungsumkehr. Da bei Generatorbetrieb der Strom in den Zwischenkreisspeicher zurückgeladen wird, kann man diese Energie beispielsweise bei mehrachsigen Antrieben dazu nutzen, andere Achsen zu versorgen, oder gegebenenfalls über einen netzgeführten Stromrichter ins Netz zurück zu speisen.

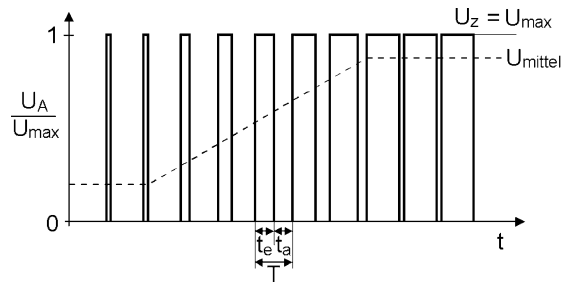


Bild 6-13: Taktung der konstanten Zwischenkreis-Spannung U_z mit Mittelwertbildung der Ausgangsspannung

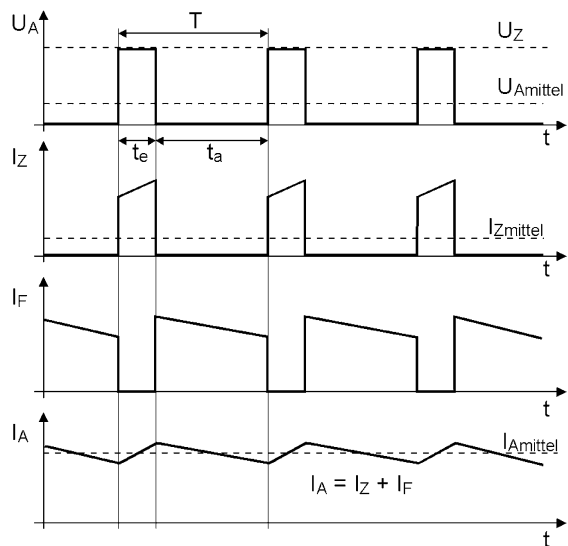


Bild 6-14: Verlauf von Spannung und Strömen bei einem 1-Quadranten Gleichstromsteller mit Pulsweitensteuerung

6.1.1.2 Drehfeldmotoren

Synchronmotoren sind ebenso wie **Asynchronmotoren** Drehfeldmaschinen, bei denen der Ständer eine oder mehrere Wechsel- oder Drehstromwicklungen trägt, die ein im Ständer umlaufendes Feld erzeugen. Im Bild 6-15 ist ein Prinzipbild eines

solchen Motors dargestellt. Da durch die Wechselspannungsspeisung des Ständerfeldes bereits ein sich drehendes Feld vorhanden ist, kann im Gegensatz zum Gleichstrommotor, der Läufer ein Magnet mit konstanter Feldrichtung sein. Ändert der magnetische Fluss im Ständer seine Richtung im Takt der Wechselspannung, so führt der Läufer durch abwechselnde Anziehung und Abstoßung bei jeder Netzperiode eine Umdrehung aus. Ist der Läufer eine gewickelte Spule, so kann die Speisespannung eine Gleichspannung sein, die wegen der Läuferdrehung jedoch über Schleifringe übertragen werden muss. Wie auch schon bei den Gleichstromantrieben kann man natürlich auch den Läufer durch einen Permanentmagneten ersetzen. Dadurch erhält man eine Maschine, die sehr einfach und robust aufgebaut ist. Betreibt man einen Synchronmotor am Wechselstromnetz, so ist seine Drehzahl durch folgenden Zusammenhang eindeutig festgelegt:

$$n = \frac{f}{p} \quad \text{mit } p: \text{Polpaarzahl und } f = 50 \text{ Hz}$$

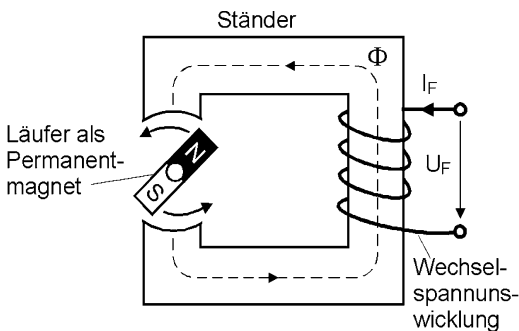


Bild 6-15: Prinzipbild eines Synchronmotors mit Permanentmagnet als Läufer

Bei einem solchen Betrieb mit einem Polpaar hat der Synchronmotor die feste Synchrondrehzahl $n_s = 3000 \text{ min}^{-1}$, oder einen sich durch mehrere Polpaare ergebenden kleineren Wert. Dies kann man dazu ausnutzen, um durch Umschalten der Polpaarzahl die Drehzahl in groben Sprüngen zu ändern.

Da eine solche Maschine im Netzbetrieb nur bei der Drehzahl, die durch die Polpaarzahl und Frequenz der Speisespannung vorgegeben ist, ein Drehmoment gleichbleibender Richtung entwickelt,

kann sie nicht selbstständig anlaufen. Sie muss beispielsweise mit einem Anwurfmotor oder mit einer Asynchronwicklung angelassen werden und läuft dann mit der Synchrondrehzahl bei hoher Drehzahlkonstanz. Überschreitet das Lastmoment ein sogenanntes maximales **Kippmoment**, so fällt der Motor außer Tritt und kommt zum Stillstand. Diese Betriebsart wird daher in mechatronischen Systemen selten vorkommen.

Hat man jedoch eine Einrichtung, mit der man die Frequenz des Ständerfeldes stufenlos einstellen kann, so erhält man einen stufenlos in der Drehzahl einstellbaren Servomotor, der praktisch wie ein bürstenloser Gleichstrommotor aufgebaut ist (Bild 6-7 b)). Dazu benötigt man einen **Frequenzumrichter** mit Gleichstromzwischenkreis, wie er als Blockschaltbild in Bild 6-19 dargestellt ist. Die dem Zwischenkreis entnommene Gleichspannung wird von einem entsprechenden steuerbaren Wechselrichter in eine stufenlos änderbare Frequenz umgesetzt, mit der die Ständerwicklung des Synchronmotors gespeist werden kann. Dadurch wird die Drehzahl des Motors stufenlos einstellbar. Ein ähnliches Funktionsprinzip hat auch der Schrittmotor, mit dem Unterschied, dass der Wechselrichter keine sinusförmigen, sondern digitale Stromwerte erzeugt. Das Prinzip des Frequenzumrichters wird im nächsten Abschnitt über die Asynchronmaschine ausführlicher behandelt.

6.1.1.3 Asynchronmotoren

Der Asynchronmotor ist sicher der am häufigsten eingesetzte Elektromotor. Der Ständer der **Asynchronmaschine** ist ähnlich dem der Synchronmaschine aufgebaut (Bild 6-15), wohingegen die Läuferwicklung nicht mit einer äußeren Spannungsquelle verbunden ist. An die Ständerwicklung wird eine Wechselspannung angelegt, wobei durch den Stromfluss ein im Ständer umlaufendes Magnetfeld hervorgerufen wird. Dieses Drehfeld induziert in den Leitern der Läuferwicklung eine Spannung, die einen Stromfluss und damit ein weiteres Magnetfeld hervorruft. Die Kräfte, die das Eigenmagnetfeld auf das umlaufende Ständermagnetfeld hervorruft, erzeugen ein Drehmoment, das den Läufer in Rotation versetzt. Der induzierte Strom ist um so geringer, je kleiner die Differenz zwischen der Drehzahl des Ständerdrehfeldes und der Läuferdrehzahl, der sogenannte **Schlupf**, ist. Die Winkelgeschwindigkeit ω des Läufers weicht im Betriebspunkt nur wenig von der Synchrondrehzahl ω_s des Netzes ab. Aus den beiden Drehzahlen lässt sich der Schlupf s berechnen:

$$s = \frac{\omega_s - \omega}{\omega_s} \quad (6.7)$$

Da dem Läufer keine Spannungen zugeführt werden müssen, ist dieser besonders einfach aufgebaut und benötigt keine störanfällige Übertragung mit Bürsten und Kommutatoren. Allerdings gibt es auch die Bauform des **Schleifringläufermotors**, bei dem die Läuferwicklungen über Schleifringe von außen variabel belastet werden können.

In der Bauform des Kurzschluss- oder Käfigläufermotors besteht die Wicklung nur noch aus massiven Stäben aus Kupfer oder Aluminium, die an den Enden durch elektrisch leitende Ringe kurzgeschlossen sind. Die Leiter liegen in den Nuten des Läuferkörpers, der aus geschichteten, isolierten Eisenblechen besteht, um Wirbelstrom- bzw. Ummagnetisierungsverluste gering zu halten. Ein solcher Motor zeichnet sich durch große Robustheit und Verschleißarmut und dadurch bedingte Servicefreundlichkeit aus. Die zugeführte Drehfeldleistung $P_d = P_{el}$ teilt sich in Abhängigkeit vom Schlupf wie folgt auf:

$$P_d = P_d \cdot (1-s) + P_d \cdot s$$

bzw.

$$P_{zu} = P_{mech} + P_v \quad (6.8)$$

Zugeführte	abgegebene	Läufer-
elektrische	mechanische	Kupfer-
Leistung	Leistung	Verlustleistung

Aus den Gleichungen (6.7) und (6.8) lässt sich die folgende Gleichung für die Winkelgeschwindigkeit des Läufers ableiten:

$$\omega = 2\pi \cdot \frac{f}{p} \left(1 - \frac{P_v}{P_{zu}} \right) \quad (6.9)$$

Demnach ergeben sich drei Möglichkeiten, die Winkelgeschwindigkeit der Asynchronmaschine kontinuierlich zu verändern:

- Änderung der Speisefrequenz f (über Frequenzumrichter)
- Änderung der Verlustleistung P_v im Läuferkreis (Schleifringläufer mit externer Beschaltung)
- Änderung der zugeführten Leistung P_{zu} (über Steller)

Die erste Möglichkeit wird weiter unten ausgeführt, zuerst soll jedoch der am häufigsten angewendete 50 Hz-Betrieb am Wechselstromnetz kurz behandelt werden.

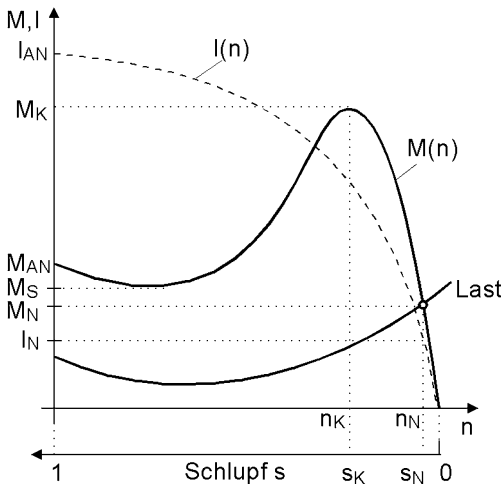


Bild 6-16: Drehzahl-Drehmoment-Kennlinie und Ständerstrom bei einer Asynchronmaschine

Im Bild 6-16 ist die Drehmoment-Drehzahl-Kennlinie eines Asynchronmotors eingezeichnet. Ebenfalls im Diagramm eingezeichnet ist die Kennlinie einer Last. Dort wo sich beide Kennlinien schneiden ist der Nennbetriebspunkt, bei dem der Motor mit Nenndrehzahl n_N läuft und dabei ein Nenndrehmoment M_N liefert. Der im Betriebspunkt vorliegende Nennschlupf s_N liegt bei 3 - 5 %. Der Anlauf der Maschine ist etwas problematisch. Bei Anlegen der Netzspannung an die Ständerwicklung fließt ein Anlaufstrom I_{AN} , der das 5-8fache des Nennstroms I_N im Arbeitspunkt beträgt. Trotzdem kann das Anlaufmoment M_{AN} , je nach genauem Verlauf der Kennlinien des Motors oder der Last, kleiner als das Nennmoment sein. Danach steigt das

Drehmoment zunächst über das Nennmoment bis auf das Kippmoment M_K an, wobei der Schlupf und der Strom durch die Ständerwicklung sich ständig verringern. Nach Erreichen des Kippmomentes fällt das Drehmoment bis auf das Nennmoment ab. Beim Hochlauf kann sich je nach Ausbildung der Läuferwicklung auch ein Sattelmoment M_S ergeben.

Das Drehmoment der Asynchronmaschine ist dem Quadrat der Speisespannung U direkt und dem Quadrat der Speisefrequenz f umgekehrt proportional:

$$M \sim \frac{U^2}{f^2} \quad .$$

Der Ständerstrom I hängt linear von der Spannung ab:

$$I \sim U \quad .$$

Um die Stromaufnahme größerer Motoren aus dem Drehstromnetz beim Anlauf zu begrenzen, wird daher für die Anlaufphase häufig eine **Stern-Dreieck-Umschaltung** der Ständerwicklung vorgenommen, bei der der Anlauf in Sternschaltung der Läuferwicklungen erfolgt und nach Überschreiten des Kippmomentes

auf Dreiecksschaltung umgeschaltet wird. Dadurch liegen in der Anlaufphase nur 220 V an den Wicklungen mit entsprechend geringeren Anlaufströmen, während beim Nennbetrieb 380 V anliegen, wodurch ein höheres Drehmoment geliefert werden kann.

Im Betriebspunkt hat die Drehzahl-Drehmoment-Kennlinie einen ähnlich steilen Verlauf wie die des Gleichstrommotors, wodurch sich hier eine gute Drehzahlsteifigkeit ergibt. Von den oben bereits angesprochenen Möglichkeiten die Winkelgeschwindigkeit eines Asynchronmotors stufenlos zu ändern, soll nun die am häufigsten verwendete Form des Betriebs mit variabler Speisefrequenz behandelt werden. Um ein möglichst hohes Drehmoment erzeugen zu können, wird der Motor mit konstantem Fluss Φ im Ständer betrieben. Der Fluss in der Ständerwicklung mit der Induktivität L_F , der durch die Feldspannung U_F hervorgerufen wird, ist proportional zum Strom I_F durch die Wicklung :

$$\Phi \sim I_F = \frac{U_F}{\omega \cdot L_F} = \frac{U_F}{2\pi \cdot f \cdot L_F} \quad .$$

Da der Fluss konstant sein soll und proportional zu U_F/f ist, muss dieser Quotient konstant gehalten werden. Daher muss bei einer Änderung der Speisefrequenz zur Verstellung der Winkelgeschwindigkeit auch die Speisespannung frequenzproportional verändert werden. Unter Vernachlässigung des Wirkwiderstandes der Läuferwicklung ergeben sich dann die in Bild 6-17 in Abhängigkeit der Winkelgeschwindigkeit gezeigten Kennlinien.

Beim Betrieb mit variabler Frequenz treten zwei Betriebsbereiche auf. In Anlehnung an die Gleichstrommaschine spricht man auch vom **Ankerverstellbereich**, der den Frequenzbereich bis zur Nenndrehzahl im Arbeitspunkt umfasst ($f < f_N$) und vom **Feldverstellbereich** oberhalb der Nenndrehzahl ($f > f_N$). Bei konstantem Feldstrom und konstantem Fluss ergibt sich im Ankerverstellbereich (U_F und f werden proportional gesteigert) mit steigender Spannung U_F eine linear zunehmende Leistung P bei konstantem Drehmoment M . Im Bereich niedriger Frequenzen wird zur Kompensation des Spannungsabfalls am ohmschen Widerstand der Wicklung ($I \cdot R$ -Kompensation) die Spannung angehoben, damit auch dort konstanter Fluss herrscht. Die Feldspannung erreicht im Arbeitspunkt bei der Winkelgeschwindigkeit ω_N ihren höchsten Wert. Die Momentenkennlinien $M(\omega)$ verlaufen im Ankerverstellbereich steil und sind untereinander parallel, der Schlupf s steigt bei sinkender Speisefrequenz mit $1/f$ an.

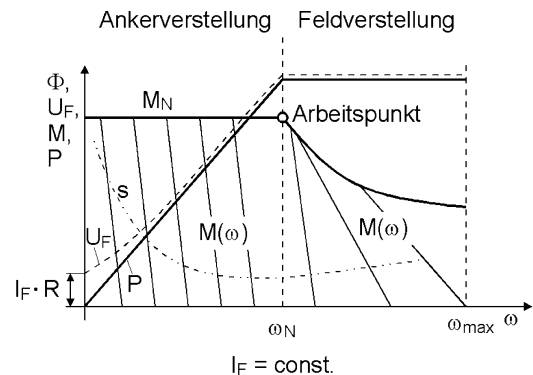


Bild 6-17: Kennlinien des Asynchronmotors bei Betrieb mit variabler Speisefrequenz

Bei weiterer Frequenzerhöhung steigt die Winkelgeschwindigkeit weiter an. Gleichzeitig sinkt jedoch der Fluss im Ständer, weil die Spannung U_F nicht mehr weiter

erhöht werden kann. Der Motor arbeitet dann quasi im Feldstellbereich, der mit einer Feldschwächung verbunden ist. Da beim Asynchronmotor das Drehmoment quadratisch von der Spannung abhängt, geht bei Frequenzsteigerung und konstanter Spannung das Kippmoment quadratisch mit der Frequenz zurück. In diesem Bereich steigt der Schlupf an und die Steilheit der Momentenkennlinien nimmt ab. Für den Betrieb als Servomotor ist daher, ähnlich wie beim Gleichstrommotor, der Feldstellbereich nicht gut geeignet.

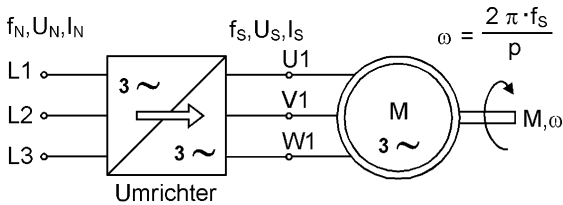


Bild 6-18: Frequenzumrichter zwischen Netz und Drehfeldmaschine

Um die oben beschriebene frequenzvariable Betriebsart einer Drehfeldmaschine durchführen zu können, wird der in Bild 6-18 schematisch dargestellte Frequenzumrichter benötigt. Das speisende 50 Hz-Drehstromnetz wird damit durch ein elektronisch erzeugtes, frequenzvariables Netz zur Speisung der Drehfeldmaschine ersetzt.

Eine übliche Bauart des Frequenzumrichters arbeitet, ähnlich wie beim Gleichstromsteller, mit einem Energiespeicher in einem Gleichstromzwischenkreis. Ein solcher Zwischenkreis-Stromrichter mit einem Kondensator als Energiespeicher besteht, wie in Bild 6-19 schematisch dargestellt, aus drei Baugruppen:

- einem ungesteuerten Eingangsstromrichter, der zur Energierückspeisung u. U. durch einen antiparallelen Rückspeisestromrichter ergänzt wird
- einem Kondensator als Energiespeicher im Zwischenkreis
- einem Maschinenstromrichter zur Speisung des Motors.

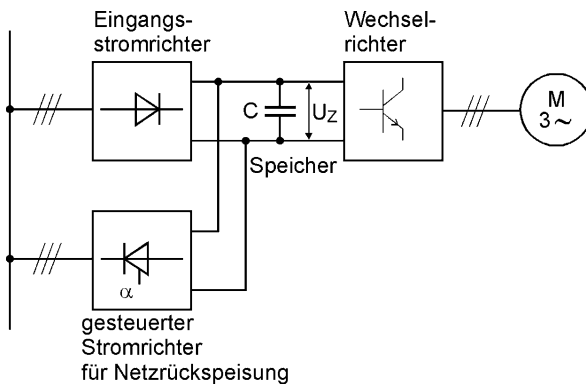


Bild 6-19: Prinzipieller Aufbau eines Zwischenkreis-Frequenzumrichters

Das Prinzip der elektronischen Drehfelderzeugung zeigt Bild 6-20 a). Das festfrequente Netz speist den Zwischenkreiskondensator über den ungesteuerten Eingangsstromrichter. Diese Anordnung stellt eine Konstantspannungsquelle mit der Zwischenkreisspannung U_Z dar. Die konstante Spannung des Zwischenkreises formt der Wechselrichter, der aus einer Drehstrom-Brückenschaltung mit 6 löschbaren Ventilen V1 bis V6 besteht, in ein Drehspannungsnetz um. Diese

Ventile können Transistoren oder GTO -Thyristoren (löschrare Thyristoren) sein. Im Gegensatz zu der sinusförmigen Spannung eines Drehstromgenerators sind die elektronisch erzeugten Spannungen nicht sinusförmig, sondern blockförmig, wodurch neben der Grundschwingung noch erhebliche Oberschwingungen vorhanden sind.

Die blockförmige, periodische, phasenverschobene Ansteuerung der Ventile (für U: 1 + 4, für V: 3 + 6, für W: 5 + 2) mit dem im oberen Teil des Bildes 6-20 b) dargestellten Impulsmuster ergibt drei um 120° phasenverschobene, blockförmige, getreppte Strangspannungen, die auf den Motor geschaltet werden. Diese Spannungen erzeugen in den Ständerwicklungen Ströme, die jeweils während eines Umlaufs in Sprüngen von 60° weitertakten und dadurch 4 verschiedene Spannungsniveaus erzeugen. Durch die induktive Last erhält man die im unteren Teil des Diagramms idealisiert dargestellten Stromverläufe.

Wie bereits oben erwähnt, muss man im Ankerstellbereich den Asynchronmotor bei

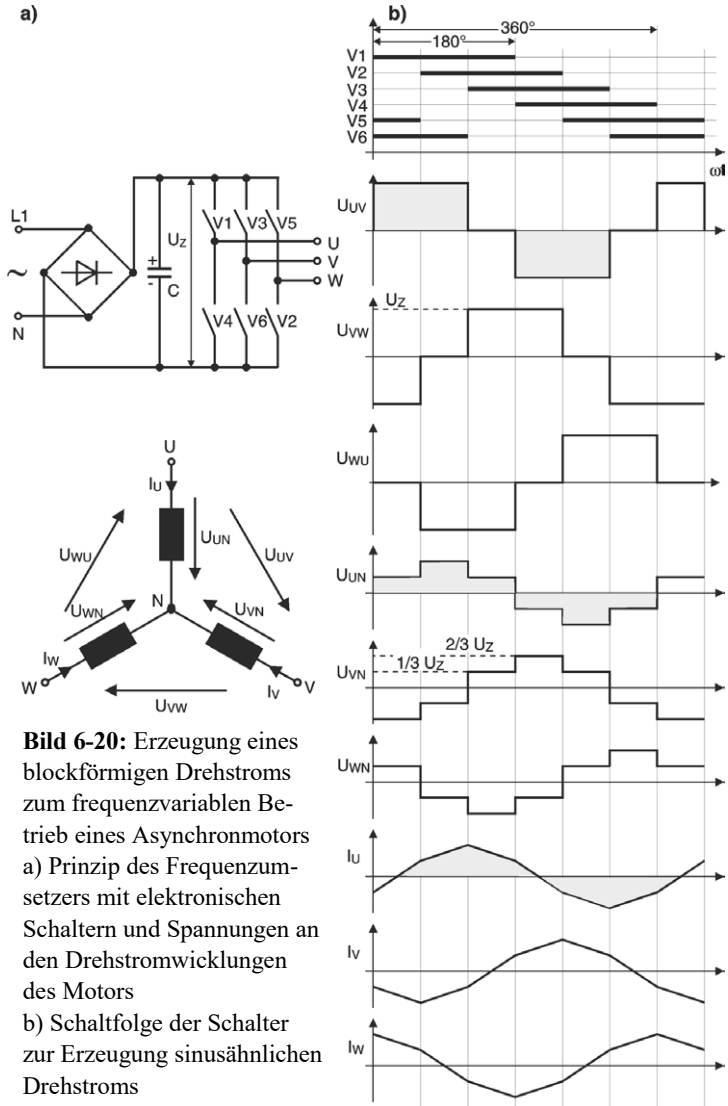


Bild 6-20: Erzeugung eines blockförmigen Drehstroms zum frequenzvariablen Betrieb eines Asynchronmotors
a) Prinzip des Frequenzumsetzers mit elektronischen Schaltern und Spannungen an den Drehstromwicklungen des Motors
b) Schaltfolge der Schalter zur Erzeugung sinusähnlichen Drehstroms

konstantem Fluss Φ im Ständer betreiben. Dazu muss man bei Änderung der Frequenz f auch proportional die Spannung U verstellen. Dies kann man beispielsweise dadurch erreichen, dass man den ungesteuerten Eingangsstromrichter in Bild 6-19 durch einen phasenanschnittgesteuerten Stromrichter mit variabler Ausgangsspannung U_z für den Zwischenkreis ausstattet. Eine andere Möglichkeit ist, den Eingangsgleichrichter ungesteuert zu belassen und die elektronischen Schalter sinusbewertet zu pulslen. Dadurch kann, wie im Blockschaltbild 6-21 a) dargestellt, sowohl die Ausgangsspannung als auch die Frequenz am Motor verändert werden. Im Bild 6-21 b) und c) sind zwei Beispiele dargestellt, wie jeweils durch die Änderung des Verhältnisses Pulsbreite zu Pulspause, sowohl die Ausgangsamplitude als auch die Ausgangsfrequenz verändert werden kann. Durch die hohe Pulsfrequenz werden an der Ankerinduktivität die Ströme geglättet, so dass der ebenfalls in Bild 6-21 b), c) eingetragene, sehr gut sinusförmige Verlauf auftritt. Mit Hilfe dieser Schal-

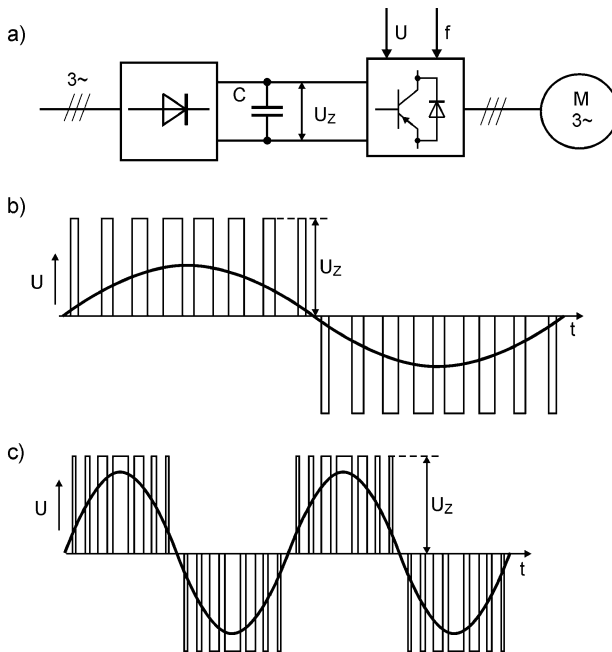


Bild 6-21: Erzeugung von Wechselspannungen variabler Höhe und Frequenz durch Pulsbetrieb des Wechselrichters
a) Blockschaubild b) sinusförmiger Stromverlauf durch Änderung der Pulsbreiten
c) Erhöhung von Ausgangsspannung und Frequenz mit Hilfe des Puls-Pausen-Verhältnisses

tungstechnik wird auch der starke Oberwellengehalt, den eine Schaltung nach Bild 6-20 erzeugt, deutlich reduziert, so dass der Motor viel ruhiger läuft.

Bis vor einigen Jahren wurden als Servoantriebe fast ausschließlich Gleichstrommotoren eingesetzt. Dies lag an der sehr einfachen und kostengünstigen Speise- und Regelungsmöglichkeit nur eines Stromes. Bei Drehstromantrieben müssen dagegen drei Strangströme in Abhängigkeit vom umlaufenden Rotorfeld geregelt werden, um das gewünschte Drehmoment einzustellen. Für diese komplexe Regelung ist außerdem die Messung der Ständerströme und der Rotorlage erforderlich. Seit mit modernen Leistungshalbleitern und Mikroelektronik die Realisierung komplexer Regelungs- und Umrichtermodule preiswert möglich ist, ist der Gleichstromservomotor aufgrund

der prinzipiellen Unterlegenheit seiner Antriebseigenschaften größtenteils durch Drehstromantriebe abgelöst worden [6.3]. Der entscheidende Nachteil des Gleichstrommotors ist der Kommutator. Er ist verschleißbehaftet und begrenzt die Motordynamik. Außerdem ist wegen der Einbrenngefahr der Bürsten im Stillstand nur eine begrenzte Drehmomentabgabe möglich. Die für Servoantriebe eingesetzten Drehstrommotoren sind wegen der fehlenden Schleifkontakte hingegen wartungsfrei. Vorteile von Gleichstrommotoren gegenüber Drehstrommotoren sind weiterhin die hohe Gleichlaufgüte und der große Drehzahlstellbereich.

6.1.1.4 Schrittmotoren

Der Schrittmotor ist im Prinzip eine permanenterregte Synchronmaschine mit einer hohen Anzahl von Polpaaren. Abweichend vom normalen Synchronmotor wird er über entsprechende Wechselrichter jedoch nicht mit periodischer, sinusförmiger Wechselspannung, sondern mit impulsförmigen, binären Stromsignalen betrieben. Diese können als Einzelimpulse oder als Impulsfolge auftreten. Darauf beruht die charakteristische Betriebsweise des Schrittmotors. Bei jedem Stromimpuls führt er

einen durch die Bauart genau definierten Winkelschritt α aus, so dass man durch Vorgabe einer bekannten Anzahl von Stromimpulsen und bei Einhaltung bestimmter Betriebsbedingungen eine Positionieraufgabe durchführen kann, ohne ein Lage- oder Geschwindigkeitsmesssystem zu verwenden. Damit ergibt sich ein einfacher und kostengünstiger technischer Aufbau von Positioniereinrichtungen. Jedoch ist zu beachten, dass es sich um eine reine Steuerung der Position und um keine echte Regelung handelt. Dies kann zu Problemen führen, wenn nämlich das äußere Lastmoment das Kippmoment der "Synchronmaschine" überschreitet. Dies hat zur Folge, dass ein vorgegebener Winkelschritt nicht ausgeführt wird, wodurch es zu dauerhaften Abweichungen zwischen angenommener und tatsächlicher Istposition kommt.

Im Bild 6-22 sind schematisch verschiedene Bauformen eines Permanentmagnet-Schrittmotors dargestellt. Teilbild a) zeigt einen Motor mit magnetisierten Polen im Läufer und einem Ständer mit mehreren Polschuhen und Erregungsspulen, Teilbild b) zeigt einen Mehrständermotor mit Klauenpolen und Einzelwicklungen. Der in Teilbild a) dargestellte Einständermotor hat vier Pole, von denen zwei vom gleichen Wicklungsstrang umschlungen sind und daher vom gleichen Strom durchflossen werden. Die beiden getrennten Wicklungen (Phasen) können getrennt angesteuert werden.

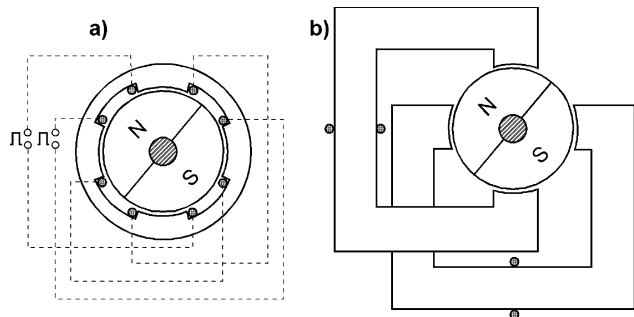


Bild 6-22: Schrittmotoren mit Permanentmagnet-Läufern und a) einem Ständer mit mehreren Wicklungen und b) mehrere Ständer mit jeweils einer Wicklung

Im Betrieb werden die Motorwicklungen über Wechselrichter mit variabler Steuerfrequenz und bestimmten Schaltsequenzen aus einem Gleichspannungszwischenkreis gespeist. Dies geschieht so, dass bei variabler Drehzahl genau die vorgegebene Schrittzahl ausgeführt wird. Die Größe des Schrittwinkels α wird durch die Strangzahl m (Anzahl der Phasen) und die Polpaarzahl p des Motors bestimmt:

$$\alpha = \frac{360^\circ}{2 \cdot m \cdot p} .$$

Da der Motor in Bild 6-22 a) zwei Stränge und ein Polpaar pro Strang besitzt, würde er pro Stromimpuls einen Winkelschritt von 90° ausführen. Reale niederpolige Motoren führen relativ große Winkelschritte von beispielsweise $\alpha = 7,5^\circ$ aus, während höherpolige Motoren Schritte von $\alpha = 0,72^\circ$ ermöglichen. Die erzielbare Winkelgeschwindigkeit hängt von der Frequenz f ab, mit der die Ansteuerimpulse auf die Wicklung gegeben werden:

$$\omega = \frac{2\pi \cdot f}{m \cdot p} .$$

Dabei werden Schrittfrequenzen bis zu 100 KHz verwendet.

Um sehr kleine Schrittwinkel zu erzielen, müssen die Pole in Längsrichtung gezahnt sein. Je nach Anzahl der Wicklungsstränge besteht der Rotor dann aus ebenso vielen axial nebeneinander angeordneten Polschuhen, die jeweils unterschiedlich magnetisiert sind. Bild 6-23 zeigt einen solchen 5-Phasen-Schrittmotor mit mehrfach gezahnten Statorpolen.

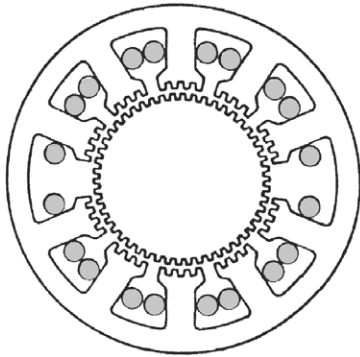


Bild 6-23: Querschnitt durch einen 5-Phasen-Schrittmotor

Solche Motoren können bis zu 500 Vollschrötte ausführen. Beim sogenannten Halbschrittbetrieb führt der Läufer jeweils einen Zwischenschritt aus, da die Bestromung der jeweiligen Wicklung vor dem Umpolen des Strangs auf Null geschaltet wird. Dadurch verdoppelt sich die Anzahl der möglichen Schritte. Auf elektronischem Wege kann man die Wicklungen auch nicht nur binär sondern gestuft bestromen. Durch diese sogenannte Mikroschrittansteuerung lassen sich die mechanisch möglichen Schritte noch weiter unterteilen und man kann Schrittzahlen von 10.000 Schritten pro Umdrehung erreichen.

Durch die oben genannten Eigenschaften zeichnet sich der Schrittmotor vor allem für Probleme kleiner Antriebsleistungen aus, weil man auf einfache Art sehr große Drehzahlbereiche und beim Positionieren hohe Genauigkeit erreichen kann. Da keine mechanische Kommutierung erforderlich ist und der Anker ein Permanentmagnet ist, ist die wartungsfreie Lebensdauer hoch und eine Überlastung praktisch nicht möglich.

Allerdings lässt die Drehzahlsteifigkeit zu wünschen übrig und das Leistungsvolumen ist ungünstiger als bei Gleichstrommotoren. Aufgrund der Ansteuerungsart entsteht ein hohes Laufgeräusch, das bei niedrigen Drehgeschwindigkeiten sich auch in Form von ruckhaften Bewegungen äußert.

6.1.2 Elektromotorische Linearantriebe

Für translatorische Bewegungen werden in der Regel rotierende elektromotorische Antriebe verwendet, deren Bewegung dann durch ein entsprechendes Getriebe (Ritzel-Zahnstange, Spindel-Mutter) in eine lineare Bewegung gewandelt wird. Man kann aber eine Linearbewegung auch direkt mit einem elektrischen **Linearmotor** erreichen, wodurch man Getriebe und die damit verbundenen Probleme und Kosten vermeidet.

Vom physikalischen Wirkprinzip her entspricht ein Linearmotor einem Drehstrom-Synchron- oder einem Asynchronmotor. Man kann sich den Linearmotor so entstanden denken, dass man den Ständer einschließlich der Ständerwicklung in eine Ebene ausbreitet (Bild 6-24). Anstelle des magnetischen Drehfeldes, das beim Asynchronmotor im Ständer umläuft, benötigt man hier ein lineares Wanderfeld, das sich mit der Synchrongeschwindigkeit v_s vom Anfang des ebenen Ständers zu dessen Ende hin bewegt. Oberhalb der ebenen Ständerwicklung, die beim Linearmotor als **Induktor** bezeichnet wird, befindet sich ein Reaktionsobjekt (Schiene,

Läufer) aus elektrisch leitendem Material. In dieser **Reaktionsschiene** werden, wie beim Kurzschlussläufermotor, durch das Wanderfeld aufgrund des Induktionsprinzips elektrische Ströme induziert, deren Magnetfelder in Wechselwirkung mit dem wandernden Magnetfeld im Induktor treten. Hierdurch wird eine Kraft erzeugt, die wie beim Asynchronmotor vom Schlupf abhängig ist. Sie bewegt die Reaktionsschiene mit einer Geschwindigkeit v im Wanderfeld, die bei kleinem Schlupf nur gering von der Wanderfeldgeschwindigkeit abweicht.

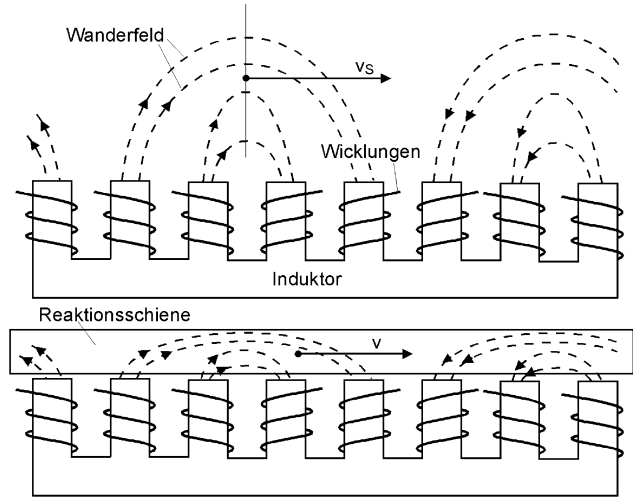


Bild 6-24: Aufbau und Wirkungsweise eines asynchronen Linearmotors

Da man wie beim Asynchronmotor eine solche Einrichtung auch mit einem Frequenzumrichter betreiben kann, besitzt der so ausgestattete Linearmotor auch die Eigenschaft, dass seine Lineargeschwindigkeit stufenlos geregelt werden kann.

Das gleiche Prinzip kann man natürlich auch auf die Arbeitsweise eines Schrittmotors übertragen und erhält damit einen **Linearschrittmotor**. Mit einem solchen Motor kann man dann sogar ohne weitere Wegerfassungen Positionierungen vornehmen. Dies ist sicher die einfachste Art einer linear arbeitenden Positionierung, da ja auch alle Getriebeelemente entfallen. Allerdings besitzt er ebenfalls die beim rotierenden Schrittmotor genannten Nachteile.

Neuere Weiterentwicklungen des Wirkprinzips mit einem ebenen Induktorkamm ermöglichen andererseits auch wieder den Antrieb eines rotierenden Systems. Wie in Bild 6-25 dargestellt [6.2], wird dazu ein kreissegmentförmiger Stator verwendet, der über einen Luftspalt direkt auf einen rotatorisch anzutreibenden Körper

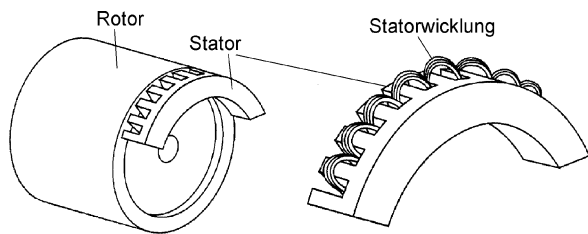


Bild 6-25: Linearmotor mit einem gekrümmten Induktorkamm für den rotatorischen Antrieb von Trommeln und Walzen

wie eine Trommel oder eine Walze wirkt, oder auf einen an der Stirnseite angebrachten Kurzschlussring. Entsprechend dem Induktionsprinzip werden wieder Ströme induziert, deren Magnetfeld vom Magnetfeld im Stator mitgenommen wird. Auf diese Weise kann die Antriebsenergie berührungslos und damit ohne Getriebe oder Übertragungselemente wie Antriebsriemen, auf das rotierende Bauteil übertragen werden. Dies führt einerseits zur Geräuschminderung und andererseits dazu,

dass bei ausgeschaltetem Linearmotor das rotierende Bauteil Freilaufverhalten besitzt.

6.1.3 Fluidische Aktoren

Unter fluidischen Aktoren versteht man solche Antriebe und Stelleinrichtungen, die zur Erzeugung von Kräften oder Bewegungen flüssige oder gasförmige Medien benutzen. Diese Medien besitzen eine über dem Umgebungsniveau liegende kinetische oder potentielle Energie, so dass sie in der Lage sind, in einem Motor oder Stellglied Arbeit zu verrichten. Vor allem für die Erzeugung von Linearbewegungen sind fluidische Aktoren sehr einfach aufgebaut und den meisten elektrischen Lösungen überlegen. Allerdings sind exakte Positionierungen problematisch und man benötigt eine Einrichtung zur Erzeugung des Druckmediums. In den letzten Jahren wurden hohe Anstrengungen unternommen, die Möglichkeiten der exakten Positionierung mit fluidischen Aktoren zu verbessern.

6.1.3.1 Pneumatische Aktoren

Es gibt zwar auch rotierende pneumatische Antriebe, die aber nur dann angewendet werden, wenn beispielsweise hohe Anforderungen an den Explosionsschutz gestellt werden. Aufgrund der Kompressibilität von Luft ist die Drehsteifigkeit sehr

gering, so dass sie in der Regel mit hoher Getriebeübersetzung ins Langsame verwendet werden. Ihr Einsatz in mechatronischen Systemen ist eher unwahrscheinlich.

Für sehr einfaches binäres Stellverhalten sind Pneumatikzylinder gut geeignet. Aufgrund der Kompressibilität von Luft ist ein exaktes Positionieren mit einfachen Zylindern nicht möglich und nur mit erheblichem zusätzlichen, regelungstechnischen Aufwand erreichbar. Arbeitet ein Zylinder jedoch zwischen zwei Schaltstellungen gegen feste Anschläge, so hat man ein sehr einfaches lineares Antriebssystem mit einstellbarer Geschwindigkeit. Beispielsweise in Handhabungsgeräten werden viele Greifgeräte oder Vorrichtungen mit Pneumatikzylindern betrieben. Bei den dort auftretenden einfachen Be-

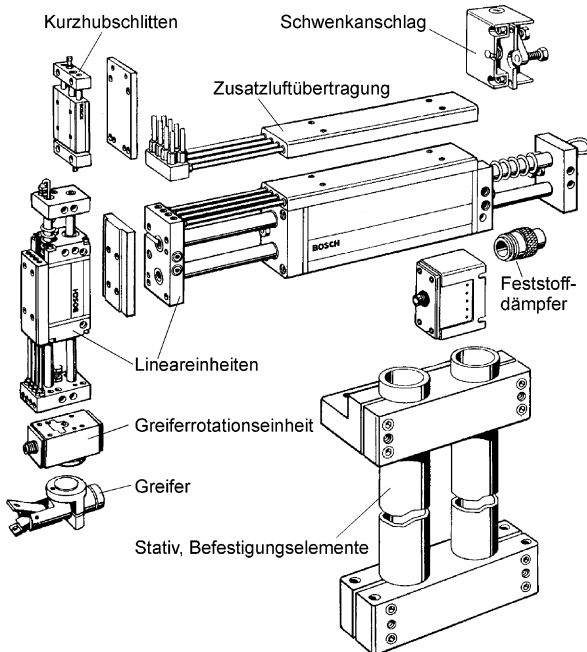


Bild 6-26: Modellübersicht eines Baukastensystems für Handhabungseinrichtungen mit pneumatischen Lineareinheiten (Bosch)

wegungsvorgängen ist der Pneumatikzylinder eine preiswerte Antriebsmöglichkeit, die sich auch durch geringes Gewicht auszeichnet. Bild 6-26 zeigt eine Modulübersicht eines Baukastensystems für einfache Handhabungsgeräte mit pneumatischen Linearantrieben. Um lineare Stellbewegungen auszuführen, ist ein Pneumatikzylinder, wie er in Bild 6-27 dargestellt ist, ein

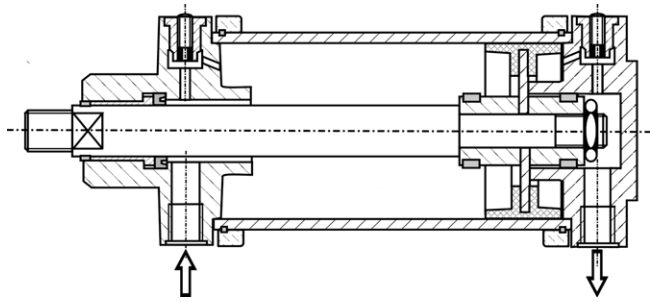


Bild 6-27: Schnitt durch einen doppelt wirkenden Pneumatikzylinder mit einstellbarer Dämpfung

sehr kostengünstiger einfacher Aktor, da er direkt ohne Getriebeelemente eingesetzt werden kann. Verglichen mit einem hydraulischen Antrieb besteht neben einigen Nachteilen auch noch der Vorteil, dass das Druckmedium nicht zum Erzeuger zurückgeführt werden muss, sondern nach seiner Entspannung in die Atmosphäre abströmen kann. Betrachtet man den Preis für elektrische Antriebsenergie, so ist der Preis für Druckluft jedoch immer noch hoch.

6.1.3.2 Hydraulische Aktoren

Ein Hauptvorteil hydraulischer Aktoren gegenüber elektrischen Antrieben und Stellvorrichtungen ist die weitaus höhere Kraftdichte und das damit verbundene geringere Bauvolumen und **Leistungsgewicht**. Es lassen sich daher mit relativ kleinen Baueinheiten große Kräfte und Momente erzeugen.

Nachteilig ist vor allem bei rotatorischen hydraulischen Antrieben ihr geringerer Wirkungsgrad, Schmutzempfindlichkeit, Temperaturabhängigkeit und der durch die hohe Kraftdichte in Verbindung mit der Reibung bewirkte höhere Verschleiß einzelner Bauelemente. Außerdem ist immer eine aufwendige Druckölerzeugung erforderlich, die mit einer hohen Geräuschentwicklung verbunden ist. Daher wird man hydraulische Antriebe heute vor allem einsetzen, wo das geringe Leistungsgewicht von Vorteil ist wie etwa bei ortsbeweglichen Baumaschinen.

Ähnlich wie bei pneumatischen Aktoren sind im Bereich der Hydraulik linear wirkende Zylinder als Stellvorrichtungen gut geeignet. Mit ihnen kann man auf sehr einfache Weise eine in der Geschwindigkeit durch Steuerung des Volumenstroms stufenlos einstellbare Linearbewegung erzeugen. Für Klemmeinrichtungen, die bei kurzem Hub hohe Kräfte aufbringen müssen, ist der Hydraulikzylinder unübertroffen.

Aber obwohl Hydrauliköl als Druckmedium verglichen mit Luft nicht kompressibel ist, lässt die Steifigkeit eines hydraulischen Linearantriebs, mit dem exakte Positionierungen unter Last vorgenommen werden sollen, zu wünschen übrig. Dies kann man leicht einsehen, wenn man die Steifigkeit einer Positioniereinrichtung mit einem hydraulischen Zylinder, mit einer solchen aus Elektromotor und beispielsweise einem Spindel-Mutter-System vergleicht. Als Maß für die Steifigkeit kann man die einer Ölsäule bestimmten Querschnitts und bestimmter Länge mit der einer Spindel

aus Stahl gleicher Baugröße vergleichen. Es zeigt sich, dass die Steifigkeiten der beiden Vergleichsobjekte sich wie die Elastizitätsmodule der Materialien verhalten:

$$\frac{k_{\dot{O}l}}{k_{\text{Spindel}}} = \frac{E_{\dot{O}l}}{E_{\text{Stahl}}} \quad (6.10)$$

Da die E-Module folgende Werte haben

$$E_{\dot{O}l} = (1,4 - 2) \cdot 10^3 \text{ N/mm}^2 \quad E_{\text{Stahl}} = 2,1 \cdot 10^5 \text{ N/mm}^2,$$

ergeben diese Werte in Gl. (6.10) eingesetzt ein Steifigkeitsverhältnis von 1 : 150 im ungünstigsten Fall. Das heißt nichts anderes, als dass eine Ölsäule unter einer bestimmten Lastkraft um den 150-fachen Betrag wie eine vergleichbare Stahlsäule komprimiert wird. Verbesserungen würden hier nur eine starke Erhöhung des Zylinderquerschnitts mit der damit verbundenen unerwünschten Erhöhung des Bauvolumens, oder eine Verkürzung des Zylinderhubs und damit verbundener Einschränkung des Arbeitsbereichs bringen.

Im Bereich der rotierenden hydraulischen Antriebe gibt es solche, die die kinetische Energie (**hydrodynamische Maschinen**) oder die potentielle Energie (**hydrostatische Maschinen**) des Ölmediums ausnutzen. Ein Beispiel für ein hydrodynamisches System ist der in PKW-Automatikgetrieben eingesetzte Föttinger Wandler. In ihm findet die hydrodynamische Leistungsübertragung zwischen einer Kreiselpumpe und einer Flüssigkeitsturbine statt. Mit einer solchen Einrichtung erreicht man eine stufenlose Drehzahlanpassung ohne Drehmomentwandlung.

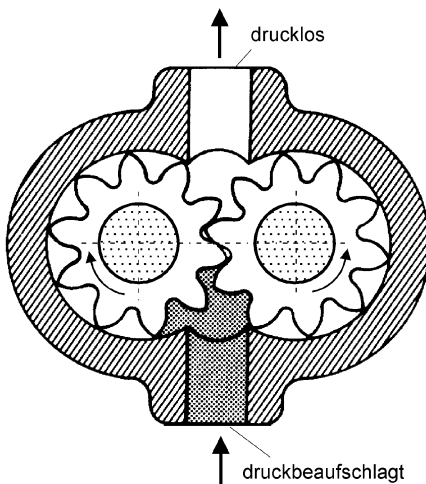


Bild 6-28: Zahnradmotor

Bei den hydrostatischen Antrieben, die die Druckenergie des Ölstroms zur Erzeugung einer Drehbewegung benutzen, unterscheidet man solche mit konstanter und solche mit stufenlos einstellbarer Drehzahl. Ein typisches Beispiel für einen Antrieb mit konstanter Drehzahl ist der **Zahnradmotor** (Bild 6-28). Bei ihm befinden sich in einem Gehäuse zwei Zahnräder, die in den Zahnücken mit dem Gehäuse sogenannte **Verdrängerräume** bilden. Die Trennung zwischen druckbeaufschlagtem und drucklosen Raum geschieht durch den Zahneingriff und zwischen Zahnkopf und Gehäuseinnenwand.

Wie bei allen hydrostatischen Antrieben beruht das Antriebsprinzip darauf, dass das einströmende Drucköl den Verdrängerraum aufweitet

und dabei entspannt wird. Durch den Aufweitungsvorgang wird auf einen, den Verdrängerraum begrenzenden Körper ein Drehmoment ausgeübt, so dass dieser in Rotation versetzt wird. Die Leistung, die ein solcher Motor abgeben kann, hängt bei idealisierten Verhältnissen (keine Leck- und Reibungsverluste) vom Volumenstrom Q und der Druckdifferenz Δp des Arbeitsmediums ab:

$$P_{\text{mech}} = Q \cdot \Delta p \quad , \quad \Delta p = p - p_0 \quad p_0 : \text{Druck auf der Abströmseite}$$

Der Volumenstrom Q wiederum ist durch das sog. **Schluckvolumen** V und die Winkelgeschwindigkeit ω des Motors bestimmt:

$$Q = \frac{\omega}{2\pi} V$$

Das Schluckvolumen bezeichnet das Ölvolumen, das pro Umdrehung durch die Verdrängerräume transportiert wird. Es beträgt für den Zahnradmotor:

$$V = 2\pi \cdot z \cdot b \cdot m^2 \cdot k_z$$

Dabei sind z die Zähnezahl, m der Zahnradmodul, b die Zahnbreite und k_z ein Korrekturfaktor, da die Zahnluken größer als die Zähne sind. Zahnradpumpen lassen Arbeitsdrücke bis zu $2 \cdot 10^7$ Pa zu.

Ähnlich wie bei den Elektromaschinen, die man im Motor- und im Generatorbetrieb verwenden kann, erlaubt die Umkehr des Motorbetriebs auch bei den hydraulischen Maschinen ihren Einsatz als Pumpe. Treibt man dazu den Rotor mit einem äußeren Moment an, so fördert er Öl in umgekehrter Richtung wie beim Motorbetrieb. In dieser Betriebsweise werden hydraulische Maschinen auch zur Erzeugung des Druckölstroms im Hydraulikaggregat verwendet (Bild 6-29). Sie werden dort in der Regel von einem Asynchronmotor angetrieben. In Fahrzeugen mit Verbrennungsmotor, wird dieser als Antrieb zum Betreiben der Pumpe benutzt.

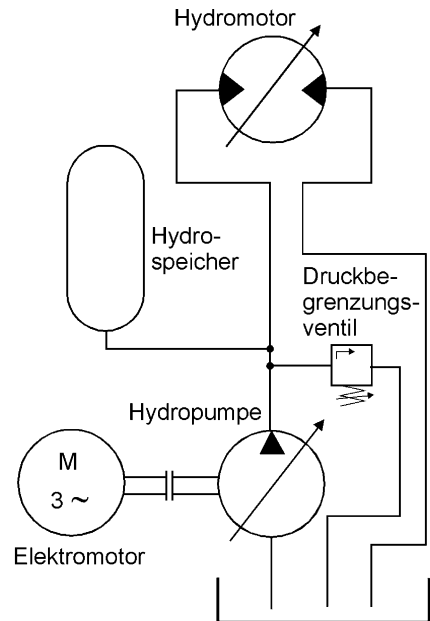


Bild 6-29: Druckölversorgung für einen Hydromotor

Im Bild 6-29 ist auch angedeutet, dass man das Fördervolumen einer Hydropumpe oder das Schluckvolumen eines Hydromotors verstellen kann. Durch beide Maßnahmen erreicht man eine Verstellbarkeit der Winkelgeschwindigkeit des Motors. Ein typischer Vertreter solcher Hydromaschinen ist der in Bild 6-30 a) dargestellte **Flügelzellenmotor**. In diesem System bilden Flügel, die radial in Schlitten gleiten, zusammen mit dem Rotor und dem Gehäuse die Verdrängerräume. Das Drucköl wird von außen in die Ver-

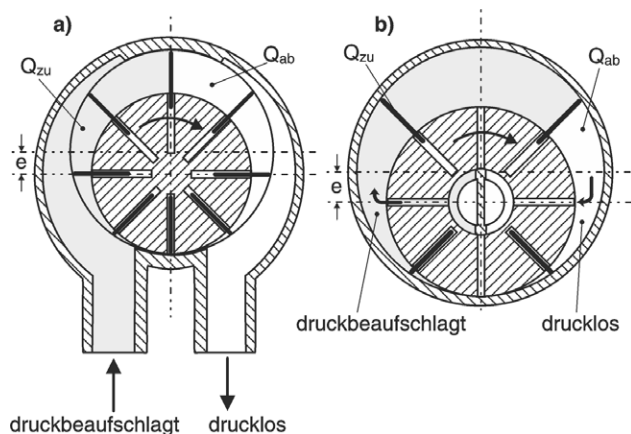


Bild 6-30: Flügelzellenmotor a) außen beaufschlagt
b) innen beaufschlagt

drängerräume eingeleitet und durch dessen Expansion wird ein Drehmoment auf den Rotor ausgeübt. Der Rotor selber ist exzentrisch gelagert, so dass sich das Volumen der Verdrängerräume bei Drehung des Rotors ändert. Hat aufgrund der Umdrehung das Volumen eines Verdrängerraums seinen größten Betrag erreicht, so wird er zwangsläufig mit der drucklosen Seite des Motors verbunden, so dass das Öl abströmen kann. Die in Bild 6-30 b) dargestellte andere Variante dieses Motorprinzips führt das Drucköl von innen durch die Rotorwelle über eine Art hydraulischen Kommutator, den sog. **Steuerspiegel**, in die Verdrängerräume. Dies ist eine von der Drehung mitgeführte mechanische Schiebereinrichtung, die die druckbeaufschlagte und die drucklose Seite des Systems trennt und die Leitungen mit den jeweiligen Verdrängerräumen verbindet.

Die Variation der Winkelgeschwindigkeit und des maximal möglichen Belastungsmoments wird durch Änderung der **Exzentrizität** e erreicht. Bei einem kreisförmigen Rotor kann man solche Motoren mit bis zu 10^7 Pa betreiben. Ein Problem bei diesen Motoren ist die Dichtung der Verdrängerräume zwischen der Seite mit hohem und mit niedrigem Druck. Ein erheblicher Teil der Verluste eines solchen Motors beruht dann auch auf **Leckverlusten** zwischen den beiden Motorseiten.

Bei Erhöhung der Exzentrizität vergrößert sich das Schluckvolumen des Motors und nimmt für die größte Exzentrizität

$$e_{\max} = \frac{D - d}{2}$$

ebenfalls sein Maximum an. Das Schluckvolumen lässt sich annähernd folgendermaßen berechnen:

$$V \approx b \left[\frac{\pi((d + 4e)^2 - d^2)}{4} - 2eaz \right],$$

worin d der Rotordurchmesser, b die Flügelbreite, a die Flügeldicke und z die Flügellanzahl sind.

Weitere verschiedene Hydromotoren arbeiten mit Zylinder-Kolben-Anordnungen, bei denen die Verdrängerräume, vergleichbar mit den Verbrennungsräumen des Verbrennungsmotors, durch die Zylinderbohrung im Gehäuse und den oszillierenden Kolben gebildet werden. Die Kolben sind entweder radial (Radialkolbenmotor) oder axial angeordnet (Axialkolbenmotor). Der **Radialkolbenmotor** kann durch Verstellung der Exzentrizität eines Rotors und der **Axialkolbenmotor** durch Verstellung der Schräglage einer Scheibe, auf der sich die axial im Rotor angebrachten Kolben abstützen, in seinem Schluckvolumen verändert werden. Ihr Betriebsverhalten und ihre Verstellbarkeit ist mit der des Flügelzellenmotors vergleichbar.

6.1.3.3 Geschwindigkeitsverstellung von hydraulischen Aktoren

Wie im vorherigen Abschnitt erläutert, können hydrostatische Motoren in ihrem Schluckvolumen und hydrostatische Pumpen in ihrem Fördervolumen verstellt werden. Verwendet man als Methode der Geschwindigkeitsstellung hydraulischer Aktoren die sog. **Pumpenverstellung**, so wird im Hydraulikaggregat entsprechend

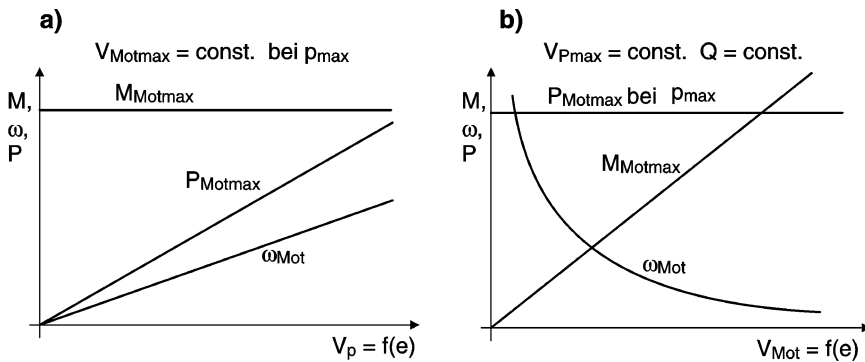


Bild 6-31: Kennlinien von Hydromotoren bei a) Pumpenverstellung b) Motorverstellung

Bild 6-31 eine variable Hydropumpe verwendet, bei der sich beispielsweise durch Veränderung der Exzentrizität der geförderte Volumenstrom ändert. Wird dieser Volumenstrom einem Hydrozylinder oder einem Hydromotor zugeführt, so ändert sich im gleichen Sinne dessen Geschwindigkeit bzw. Winkelgeschwindigkeit. Bild 6-31 a) zeigt verschiedene Kennlinien eines Hydromotors, wenn dieser konstantes, maximales Schluckvolumen $V_{Mot\ max}$ aufweist (auch Motoren mit konstantem Schluckvolumen, wie z. B. Zahnradmotoren). Die Kennlinien sind in Abhängigkeit des Fördervolumens V_p der Pumpe dargestellt, das wiederum eine Funktion der Exzentrizität ist. In dieser Betriebsart nimmt die Winkelgeschwindigkeit ω linear zu und das Motormoment ist im ganzen Drehzahlbereich konstant und maximal. Da die Motorleistung dem Volumenstrom Q_{zu} proportional ist, muss diese ebenfalls linear mit dem Fördervolumen zunehmen. Es ist unmittelbar einsichtig, dass bei einem Hydrozylinder die Lineargeschwindigkeit ebenfalls linear mit dem Fördervolumen zunimmt und dass die Kolbenkraft konstant ist.

Ist der Hydromotor ein Motor, der in seinem Schluckvolumen verstellbar ist, so kann bei ihm eine Verstellung der Winkelgeschwindigkeit bei konstantem Fördervolumen der Pumpe über die **Motorverstellung** erfolgen. Die entsprechenden Motorkennlinien bei dieser Betriebsart sind in Bild 6-31 b) dargestellt. Bei kleinem Motorschluckvolumen ist das Motormoment gering und nimmt linear mit der Vergrößerung des Schluckvolumens zu. Die Winkelgeschwindigkeit verhält sich umgekehrt proportional zum Schluckvolumen, wodurch sich der gezeigte hyperbolische Verlauf der Winkelgeschwindigkeit über dem Schluckvolumen ergibt. Die vom Motor abgegebene Leistung bleibt konstant, da ja die von der Pumpe gelieferte hydraulische Leistung konstant bleibt. Eine vergleichbare Betriebsart bei Hydrozylindern ist nicht möglich.

Vergleicht man die beiden Verstellmöglichkeiten eines Hydromotors mit der eines Gleichstrommotors, so entspricht die Pumpenverstellung der Verstellung der Ankerspannung und die Motorverstellung der Feldschwächung mit den damit verbundenen Eigenschaften.

Während die Drehzahlverstellung bei Elektromotoren mit sehr kurzen Zeitkonstanten (10-100 msec.) realisiert werden kann, ist die Verstellung von hydrostatischen Maschinen mit größeren Zeiten im Sekundenbereich verbunden. Dies beruht darauf, dass zur Verstellung nicht nur der Rotor beschleunigt, sondern auch noch sei-

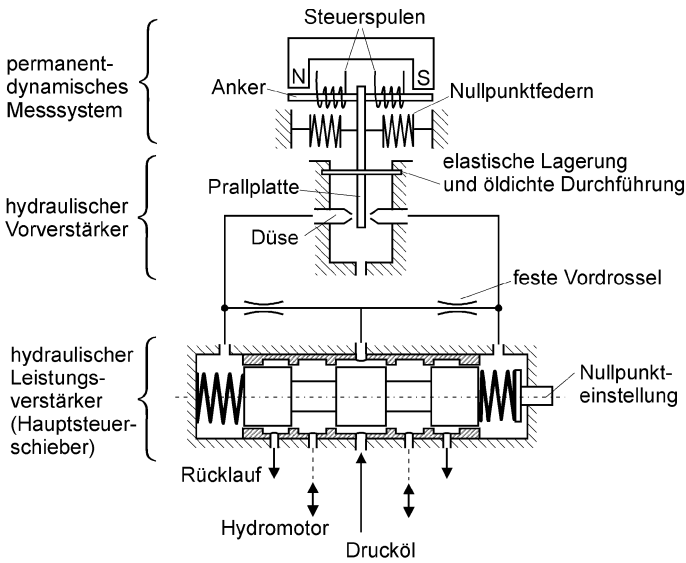


Bild 6-32: Elektrohydraulisches Servoventil

ne Exzentrizität verändert werden muss. Um diese Schwierigkeit für dynamisch anspruchsvollere Antriebsprobleme umgehen zu können, gibt es eine dritte Stellmöglichkeit. Da bei Pumpenverstellung ja als Stellgröße der geförderte Volumenstrom verändert wird, hat natürlich jede auch anders bewirkte Änderung des dem Motor zufließenden Volumenstroms die gleiche Auswirkung. Eine einfache Art den Volumenstrom des Öls bei konstantem Fördervolumen

der Pumpe zu verstellen ist, ihn mit Hilfe eines stufenlos im Durchlass verstellbaren Ventils zu drosseln. Neben den seit langem bekannten **Servo-** und **Proportionalventilen** wurden in letzter Zeit auch **Piezoventile** entwickelt, die als Stellglied ein Piezoelement benutzen [6.3].

Solche Stellventile haben gegenüber anderen Systemen den Vorteil, dass mit sehr kleinen Steuerleistungen große hydromechanische Leistungen in wenigen Millisekunden mit hoher Präzision gesteuert werden können.

Ein elektrohydraulisches Servoventil ist in Bild 6-32 als schematisches Funktionsbild dargestellt.

Das Servoventil besteht aus drei Baugruppen:

- elektromagnetische Eingangssteuerung mit permanentdynamischem Messsystem
- hydraulischer Vorverstärker mit Düse-Prallplatte als variable Öldrosselung
- hydraulischer Kraftschalter mit Folgekolben.

Die Ansteuerung des Ventils erfolgt mit elektrischen Signalen, die im Milliwattbereich liegen. Der eigentliche Hauptsteuerschieber, der den Ölstrom zum Hydromotor verstellt, kann Leistungen im Kilowattbereich steuern, so dass sich eine Leistungsverstärkung im Bereich $1:10^5$ ergibt. Die Eigenfrequenz eines Servoventils, die die maximale Änderungsfrequenz für dynamische Ansteuersignale begrenzt, liegt typisch über 200 Hz und kann bis zu 300 Hz betragen.

Fließt ein Strom durch die Steuerspulen, so wird auf den beweglichen Anker eine proportionale magnetische Kraft ausgeübt. Der Anker trägt an einem Ende eine Prallplatte, die sich zwischen zwei Düsen bewegen und je nach Abstand zur Düse, den Ölstrom durch die Düse drosseln kann. Die beiden durchflussveränderlichen

Düsen stellen veränderliche Drosseln dar und bilden zusammen mit den beiden festen Drosseln eine hydraulische Brückenschaltung (Bild 6-33). Fließt kein Strom durch die Steuerspulen, so ist die Brücke abgeglichen und der Hauptsteuerschieber, der hydraulisch in die Brücke eingefügt ist, verharrt in Mittelstellung, da auf beiden Kolbenseiten der gleiche Druck herrscht. Dadurch verschließen die Steuerkanten des Kolbens die Ölkanäle vom Druckölzufluss zu den Ausgängen, an denen sich die Zuleitungen zum Hydromotor befinden. Fließt ein Steuerstrom, so wird eine der variablen Drosseln verstellt. Dadurch steigt der Druck auf einer Seite des Steuerkolbens und es tritt eine Verschiebung nach rechts oder links auf, bis die zusätzliche Druckkraft im Gleichgewicht mit der auf der gegenüberliegenden Seite erhöhten Federkraft steht. Durch die Verschiebung kann über die jetzt geöffneten Steuerkanten Drucköl zum Hydromotor fließen und die drucklose Seite des Hydromotors wird mit dem Rücklauf verbunden. Je weiter der Steuerkolben verschoben wird, um so geringer wird die Drosselung des Ölstroms und die Winkelgeschwindigkeit des Motors steigt. Bei Auslenkung des Steuerkolbens in die andere Richtung werden die Zu- und Ableitungen umgeschaltet, so dass der Motor in der Drehrichtung umgekehrt wird. Ähnlich lässt sich natürlich auch die Linear- geschwindigkeit und Richtung der Bewegung eines Hydraulizylinders mit einem solchen Servoventil steuern.

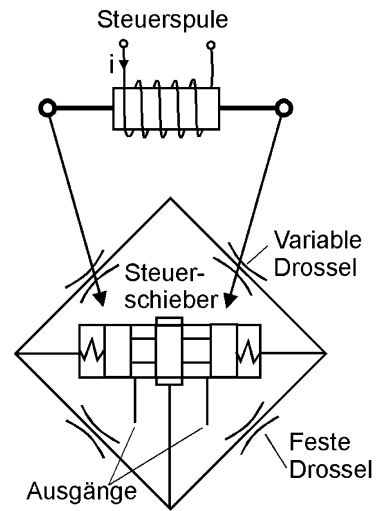


Bild 6-33: Hydraulische Brückenschaltung innerhalb eines Servoventils

Servoventile sind stark verschmutzungsempfindlich und stellen somit sehr hohe Anforderungen an die Filterung und Reinhaltung des hydraulischen Druckmittels. Andernfalls tritt erheblicher Verschleiß im Servoventil auf. Daher geht seit einigen Jahren der Trend weg von der empfindlichen Servohydraulik hin zu den Proportionalventilen, mit niedrigeren Eigenfrequenzen (bis ca. 100 Hz), aber wesentlich robusteren Betriebseigenschaften. Diese Systeme sind wesentlich weniger stör anfällig als Servoventile, da bei ihnen das

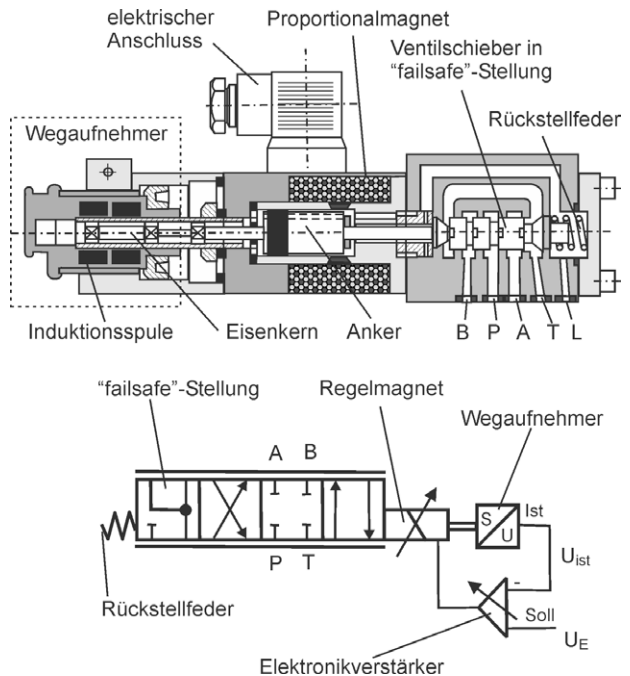


Bild 6-34: Aufbau und Prinzipbild eines Proportionalventils (Bosch)

Drucköl selbst nicht in den Stellvorgang eingebunden ist. Bild 6-34 zeigt den Aufbau und ein Prinzipbild eines Proportionalventils.

Im Gegensatz zum Servoventil (ca. 20 mW) benötigt das Proportionalventil zur Steuerung einer hydraulischen Ausgangsleistung von bis zu 100 kW eine wesentlich größere Eingangsleistung von ca. 40 W. Dies beruht auf der fehlenden hydraulischen Vorverstärkerstufe. Das in Bild 6-34 gezeigte Proportionalventil ist ein 4/4-Wegeventil mit vier Anschlüssen (Tank T, Pumpe P, Verbraucher A und B) und vier Schaltstellungen. Verstellt wird das Ventil durch einen Elektromagneten, der den Ventilschieber gegen die Rückstellfeder verschiebt. Dieser Schieber kann nicht so exakt positioniert werden, wie der Schieber des Servoventils. Zur genauen Positionierung des Schiebers und zur Vermeidung von Hysterese wird die Lage des Magnetankers durch ein Wegmesssystem erfasst und als Lage-Istwert einem Lageregelkreis zugeführt. Das bedeutet, dass die Lage des Ventilschiebers und damit der Öldurchfluss der elektrischen Eingangsspannung (U_{Soll}) direkt proportional ist.

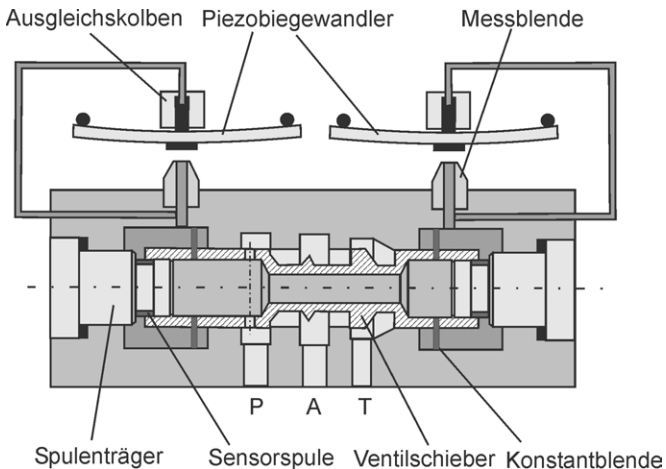


Bild 6-35: Schematischer Aufbau eines Piezoventils

systems erfolgt hierbei durch den hohlgebohrten Schieber, in dem sich auch die Durchlaufdüsen befinden. Im Deckel des Ventils sind die beiden Piezobiegewandler untergebracht, die den Widerstand der Düsen auf der Abflusseite gegensinnig verändern. Dies bedeutet, dass beim Öffnen der einen Düse die andere Düse im gleichen Maße verschlossen wird. Dadurch entsteht am Ventilschieber eine Druckdifferenz, welche in eine Beschleunigung des Schiebers umgesetzt wird. Je nach Richtung der Schieberbewegung wird Öl über die Steuerkanten vom Druck- zum Arbeitsanschluss oder vom Arbeitsanschluss in den Tank gefördert.

Je nach Amplitude kann man mit Piezoventilen obere Grenzfrequenzen von 550 Hz erreichen, was deutlich höher als bei Proportional- oder auch bei Servoventilen liegt. Ist daher eine hohe Dynamik der Ventilsteuerung gefordert, so bietet das Piezoventil die besten Voraussetzungen.

Bild 6-36 a) zeigt, wie ein als Schaltsymbol dargestelltes Servoventil und der Hydromotor zusammengeschaltet werden, um die Drehzahl des Motors zu verstellen. Das im folgenden Gesagte gilt genauso für Proportional- und Piezoventile. In Bild 6-36

Bei Stromausfall nimmt das Ventil die "failsave"-Stellung an, wodurch die Pumpe blockiert und der Öldruck im Ventil abgebaut wird. Der Steuerkolben wird hierbei durch die Feder in die Endposition geschoben.

Piezoventile sind eine neue Art von Servoventilen. Bild 6-35 zeigt den schematischen Aufbau eines Piezoventils [6.3]. Die Vorsteuerung des Ventils erfolgt mittels eines **Düse-Prallplatten-Systems**. Die Versorgung des Vorsteuer-

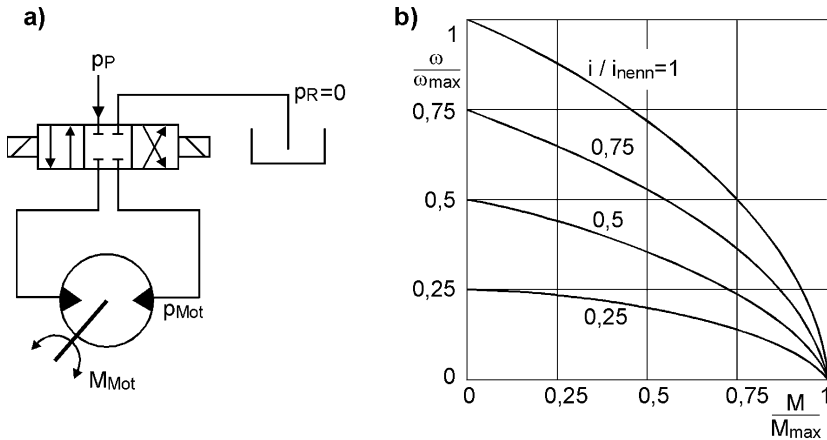


Bild 6-36: Steuerung der Winkelgeschwindigkeit eines Hydromotors mit einem Servoventil

a) Schaltschema b) Kennlinie

b) ist der Zusammenhang zwischen den Winkelgeschwindigkeiten und den Motormomenten für verschieden hohe Steuerströme dargestellt. Aufgrund der Durchflusscharakteristik eines Servoventils ist dieser Zusammenhang nichtlinear und stellt praktisch den Volumenstrom in Abhängigkeit der Druckdifferenz an einer Blende dar. Für den Durchfluss gilt:

$$Q \sim i \cdot \sqrt{\Delta p} = i \cdot \sqrt{p_P - p_{Mot}} \quad \text{mit} \quad p_{Mot} = \frac{2\pi}{V_{Mot}} M_{Mot}$$

Der Ventilsteuerstrom i bewirkt eine proportionale Auslenkung des Steuerkolbens und eine damit verbundene Veränderung des Durchströmquerschnitts. Da das Schluckvolumen des Motors konstant ist gilt für die Winkelgeschwindigkeit

$$\omega = \frac{2\pi Q}{V_{Mot}} = 2\pi C i \cdot \sqrt{p_P - \frac{2\pi}{V_{Mot}} M_{Mot}} \quad ; \quad \omega_{\max} = 2\pi C i_{nenn} \sqrt{p_P} \quad .$$

In dieser Gleichung ist C eine Motor / Ventilkonstante. Für die im Bild 6-36 b) dargestellte normierte Form gilt daher:

$$\frac{\omega}{\omega_{\max}} = \frac{i}{i_{nenn}} \sqrt{\frac{p_P - \frac{2\pi}{V_{Mot}} M_{Mot}}{p_P}} \quad \text{mit} \quad p_P = \frac{2\pi}{V_{Mot}} M_{\max} \quad ,$$

und schließlich

$$\frac{\omega}{\omega_{\max}} = \frac{i}{i_{nenn}} \sqrt{1 - \frac{M_{Mot}}{M_{\max}}} \quad .$$

Vergleicht man den Verlauf der Momentenkennlinien mit dem von Gleichstrommotoren im Ankerstellbereich oder der von Asynchronmotoren im Arbeitspunkt, so sieht man, dass die Drehsteifigkeit eines mit einem Servoventil geregelten Hydromotors von der Drehzahl und von der Ventilaussteuerung abhängig ist. Deshalb sind die

dynamischen Eigenschaften so geregelter Hydroantriebe schlechter als die Eigenschaften der genannten elektrischen Antriebe.

6.2 Neuartige Aktoren

Das Gebiet der Leistungsantriebe (einige Watt bis einige Kilowatt) ist heute im wesentlichen so weit entwickelt, dass in nächster Zeit keine revolutionären Neuerungen zu erwarten sind. Hier dominieren eindeutig die rotierenden, elektrischen Antriebe. Neuartige Aktoren wurden in den letzten Jahren vor allem für den Bereich kleiner Leistungen und für lineare Bewegungen entwickelt. Die Forschungs- und Entwicklungsaktivitäten auf diesem Gebiet liegen heute einerseits darin, bekannte Prinzipien zu erweitern und verbessern, aber es werden auch Wirkprinzipien, die schon länger bekannt, aber noch nicht praktisch in Aktoren eingesetzt wurden, zunehmend herangezogen. Eine besondere Rolle spielt hierbei auch die **Mikrosystemtechnik** [6.10], da vor allem neue Aktoren entwickelt werden, die zusammen mit Mikrosensoren und Mikrorechnern zu sehr kleinen Baueinheiten integriert werden können. Schaut man sich nochmals das Strukturbild eines mechatronischen Systems in Bild 1-24 an, so erkennt man, dass hier auch der größte Nachholbedarf bezüglich der Miniaturisierung liegt. Während die Miniaturisierung in der Elektrotechnik in den letzten 30 Jahren Verkleinerungen um mehrere Größenordnungen zur Folge hatte, ist dies im Bereich der mechanischen Stellglieder bei weitem nicht der Fall. So konnten Systeme aus dem Bereich Regelung und Steuerung sowie Sensorsysteme bedeutend verkleinert werden, der Bereich der Ausgabe hielt jedoch nicht Schritt. Dies beruht natürlich auch darauf, dass die hier umzusetzenden Leistungen in der Regel viel größer sind als in der Eingabe- und Datenverarbeitungsebene, aber auch wenn nur sehr kleine Stelleistungen erforderlich sind, sind mechanische Systeme häufig noch stark überdimensioniert. Die Entwicklungen der Mikromechanik der letzten Jahre haben dabei gute Fortschritte gemacht, ein mechatronisches System als Ganzes zu miniaturisieren. Manchmal haben solche Entwicklungen aber auch noch nicht die Praxisreife erreicht. Aber hier ist in den nächsten Jahren noch mit tiefgreifenden Verbesserungen zu rechnen. Anhand einiger repräsentativer Beispiele sollen nun solche neuartigen Aktoren vorgestellt werden.

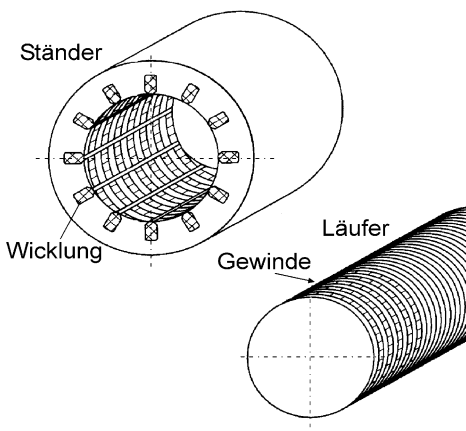


Bild 6-37: Schraubgewinde-Reluktanzmotor

In Positioniersystemen wird häufig die Funktion benötigt eine Linearbewegung auszuführen. Verwendet man die gebräuchlichen Elektromaschinen als Antriebssystem, so muss die Bewegung erst aus der Rotation durch ein schweres und teures Getriebe in eine Translation umgesetzt werden.

Dies umgeht der so genannte **Schraubgewinde-Reluktanzmotor** (Bild 6-37), der direkt eine Linearbewegung hoher Präzision erzeugt [6.5]. Reluktanzmotoren, die es auch als rotatorischen Antrieb gibt, sind prinzipiell Drehstrommotoren, deren Läufer weder Spulen noch Permanentmagnete tra-

gen, sondern die sich dadurch auszeichnen, dass sie am Umfang Pole und Lücken verschiedener magnetischer Leitfähigkeit tragen. Dies erreicht man beispielsweise durch Einlagen aus Aluminium (hoher magnetischer Widerstand) in einen Eisenläufer (niedriger magnetischer Widerstand), oder durch nicht rotationssymmetrische Ausführung des Läufers. Das im Ständer umlaufende Drehfeld übt auf den Läufer aufgrund der unterschiedlichen magnetischen Leitfähigkeit (Reluktanz) ein Drehmoment aus.

Beim Schraubgewinde-Reluktanzmotor ist in den Stator ein Gewinde eingeschnitten und der Läufer enthält ein gegenläufiges Gewinde. Beide Teile berühren sich im montierten Zustand nicht, sondern haben einen Luftspalt zwischen Statorbohrung und Läufer von 0,1 - 0,4 mm. Das Magnetfeld wirkt nach dem Reluktanzprinzip auf den durch das Schraubgewinde in axialer Richtung veränderlichen magnetischen Widerstand des Läufers und treibt dadurch den Läufer linear an. Die Bewegungsrichtung wird von der Drehrichtung des Magnetfeldes bestimmt. Die Lineargeschwindigkeit des Läufers ist proportional der Frequenz des Drehfeldes und der Gewindesteigung. Daher lässt sich die Lineargeschwindigkeit stufenlos mit Hilfe eines Stromrichters verstellen, wobei man Lineargeschwindigkeiten zwischen 1 mm/s und 800 mm/s erreicht. Im Stillstand ist eine Haltekraft vorhanden.

Da die Ständerwicklung nur Magnetisierungsstrom aufnimmt, beträgt die Erwärmung des Stators nur etwa 10-20 K, die des Läufers ist zu vernachlässigen. Die Schubkraft kann über den Strom eingestellt werden und ist proportional zum Läuferdurchmesser. Da die Motormasse aber quadratisch mit dem Läuferdurchmesser zunimmt, liegt eine technisch sinnvolle obere Grenze der Schubkraft solcher Motoren bei 500 N.

Auch auf die Miniaturisierung von Elektromotoren richten sich die Entwicklungsbemühungen. Bild 6-38 zeigt einen Permanentmagnet Motor mit integriertem Planetengetriebe, das einen Durchmesser von 1,9 mm besitzt.

Außer den bereits erwähnten elektromagnetischen oder elektrostatischen Kräften, sowie den statischen und dynamischen Kräften in einem strömenden Fluid, gibt es natürlich noch andere physikalische Prinzipien um Kräfte zu erzeugen. Der piezoelektrische Effekt wurde bereits mehrfach erwähnt, durch den in einem Kristall mit piezoelektrischen Eigenschaften elektrische Spannungen erzeugt werden, wenn äußere Kräfte Verformungen im Kristall hervorgerufen. Der inverse piezoelektrische Effekt ruft demzufolge bei Anlegen einer elektrischen Spannung an ein Piezoelement eine Verformung hervor. Ist dies eine Ausdehnung in der betrachteten Richtung, so werden auf anliegende Objekte Druckkräfte ausgeübt. Diese Verformungen liegen natürlich je nach Größe des Piezoelementes im Submillimeter- bis Submikrometerbereich. Sie können aber als Stellglieder für kleine Hübe mit Kräften bis zu einigen kN gebaut werden und zeichnen sich durch kurze Ansprechzeiten (100 μ s) aus.

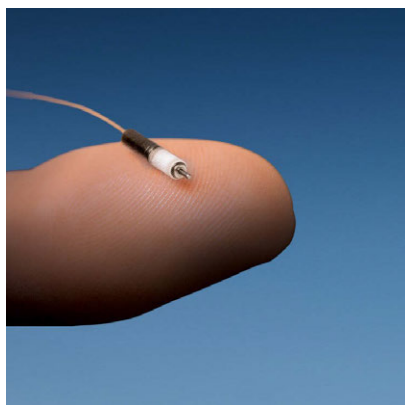


Bild 6-38: Miniaturmotor mit Planetengetriebe (Faulhaber)

- k_{ij} : Koppelkoeffizienten (dimensionslos): Die Koeffizienten stellen Energieverhältnisse dar, welche die Umwandlung von mechanischer in elektrische Energie oder umgekehrt beschreiben. Das Verhältnis von gespeicherter Energie (mechanisch oder elektrisch) zu aufgewendeter Energie (mechanisch oder elektrisch) wird durch k^2 ausgedrückt.

Um elektrische und mechanische Größen miteinander zu verbinden, werden doppelte Indizes (z.B.) eingeführt. Der erste Index gibt die Richtung der Erregung, der zweite die Richtung der Reaktion des Systems an.

Beispiel: Der Modul d_{33} ist entscheidend, wenn ein elektrisches Feld in Richtung der Polarisierung (Achse 3) anliegt und die Auslenkung in der gleichen Richtung betrachtet wird. Der Modul d_{31} kommt zum Tragen, wenn ein elektrisches Feld in der gleichen Richtung wie zuvor anliegt aber die Auslenkung in Richtung der Achse 1 (also orthogonal zur Polarisationsachse) betrachtet wird.

Piezokeramik für Aktoren wird in der Regel in Form dünner Scheiben eingesetzt. Die Längenänderung ΔL einer unbelasteten einzelnen Lage Piezokeramik kann mit der folgenden Gleichung abgeschätzt werden:

$$\Delta L \approx \pm E \cdot d_{ij} \cdot L_0 \quad ,$$

mit: $[L_0] = \text{m}$: Keramiklänge , $[E] = \text{V/m}$: elektrische Feldstärke, $[d_{ij}] = \text{m/V}$: Piezomodul.

Die Materialien, die in Piezoaktoren eingesetzt werden, weisen d_{33} -Werte von etwa 250 bis 550 pm/V und d_{31} -Werte von etwa -180 bis -210 pm/V auf und können mit einer maximalen Feldstärke von 1000 - 2000 V/mm betrieben werden. Daraus folgt das in Bild 6-40 a) dargestellte Verhalten einer zylindrischen Scheibe eines Piezoaktors. Legt

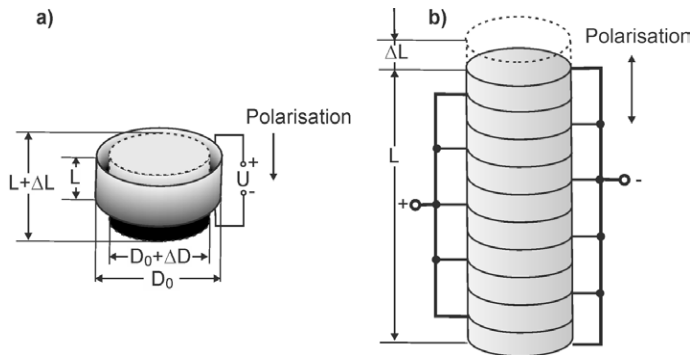


Bild 6-40: Scheibenförmiger Piezoaktor a) Einzelscheibe...
b) Stapeltranslator

man eine äußere Spannung in der Polarisationsrichtung (Achse 3) an, so dehnt sich die Ausgangslänge L_0 der Keramikscheibe um ΔL , gleichzeitig kontrahiert sich der Durchmesser D_0 (Achse 1) um ΔD . Dies beruht darauf, dass der Wert d_{31} negativ ist. Beispielsweise wäre die Dickenänderung einer Keramikscheibe der Dicke $L_0 = 0,5 \text{ mm}$ bei $d_{33} = 4 \cdot 10^{-10} \text{ m/V}$ und einer Feldstärke von 500 V/mm ungefähr $\Delta L = 0,1 \text{ }\mu\text{m}$.

Um größere lineare Stellwege zu ermöglichen werden so genannte **Stapeltranslatoren** aus einigen hundert Keramikscheiben von 25 - 100 μm Scheibendicke zusammengeschaltet (Bild 6-40 b). Je nach Länge des Stapeltranslators erreicht man so Verschiebewege von ungefähr 10 - 100 μm .

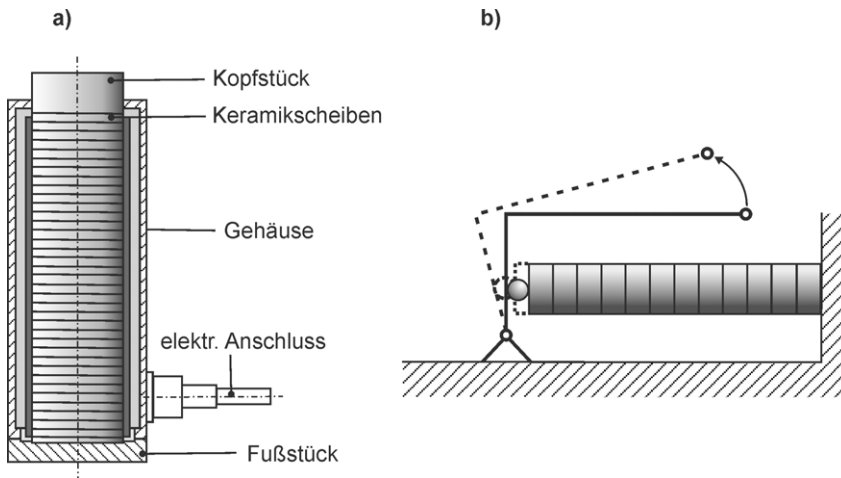


Bild 6-41: Anwendungsformen von Stapeltranslatoren a) linearer Piezoaktor b) Piezoaktor mit einfacher Hebelübersetzung

Der aktive Teil des Stapeltranslators in Bild 6-41 besteht aus einem Stapel von Keramikscheiben, die durch dünne Metallelektroden getrennt sind. Die maximale Betriebsspannung ist proportional zur Dicke der Scheiben. Die meisten Hochvoltaktoren sind aus Keramikschichten von 0,4 bis 1 mm Dicke aufgebaut. Bei Vielschichtaktoren werden die Keramiklagen und Elektroden zu einem Monolithen zusammengeintert und haben Dicken zwischen 25 und 100 μm .

Stapeltranslatoren können hohe Druckkräfte aufnehmen und bieten die höchste Steifigkeit aller Piezobauformen. Standardelemente mit bis zu 100 kN Druckbelastbarkeit sind verfügbar, und vorgespannte Aktoren können auch im Zugbetrieb arbeiten. Die Auslenkung von Piezoaktoren und -positioniersystemen lässt sich durch die Integration von Hebelmechanismen vervielfachen (Bild 6-41b). Das Hebel- und Führungssystem muss extrem steif, reibungs- und spielfrei sein, um hohe Dynamik und Auflösungen im Sub-Nanometerbereich erzielen zu können. Kugel- oder Rollenlager scheiden deshalb aus.

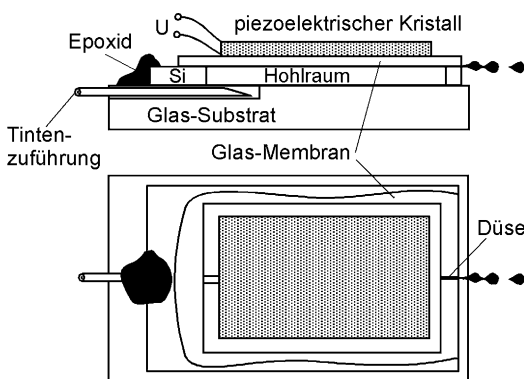


Bild 6-42: Tintentropfenerzeugung für einen Tintenstrahldrucker mit piezoelektrischem Aktor

Piezoelemente kann man auch als Antrieb in komplexeren Mechanismen verwenden. Bild 6-42 [5.9] zeigt eine Tintentropfenquelle für einen Tintenstrahldrucker, bei dem der Ausstoß eines Tintentropfens durch die Druckeinwirkung auf den Tintenhohlraum mit Hilfe eines piezoelektrischen Kristalls ausgeübt wird. Die Düsenelemente und das Tintenvorratsvolumen sind anisotrop aus einem (110) - orientierten Siliziumkristall herausgeätzt. Die Austrittsdüse ist durch isotropes Ausätzen eines 25 mm tiefen Kanals in das Silizium geformt worden.

Das Formänderungsverhalten von Piezomaterialien kann man auch für Motoranwendungen geringer Größe benutzen. Neben anderen Prinzipien gibt es das des **Langevin-Piezomotors**, dessen Antriebselement in Bild 6-43 a) dargestellt ist [6.6]. Das Antriebselement besteht aus einer Scheibe aus Piezomaterial, die auf beiden Seiten eine Schicht aus elastischem Material trägt. Das Piezoelement wird mit einer Wechsellspannung beaufschlagt, so dass das elastische Material zu Longitudinalschwingungen angeregt wird. Entspricht die halbe Wellenlänge der Anregeschwingung genau der Länge des Antriebselementes, so nehmen die Schwingamplituden und damit die Verschiebung der Enden des Schwingkörpers ihren maximalen Wert an.

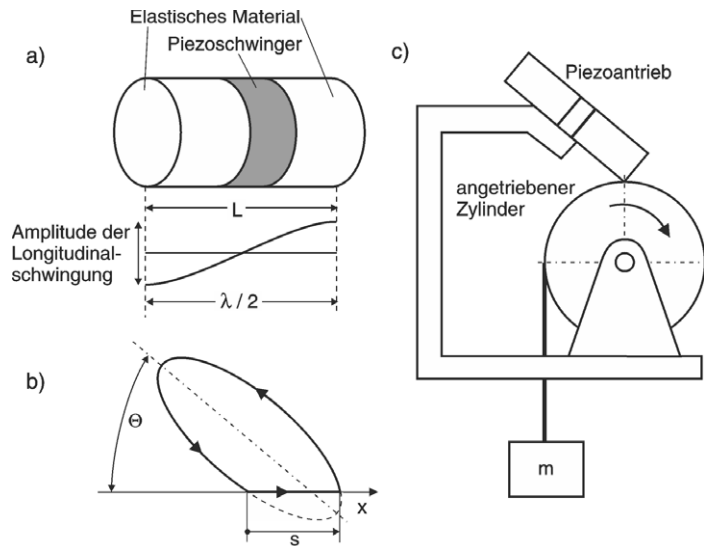


Bild 6-43: Piezoelektrischer Rotationsantrieb a) Piezoelektrisches Antriebselement b) Weg des Zylinderendes des Antriebselementes bei Berührung mit einem Objekt unter dem Winkel Θ c) Schematischer Aufbau eines Rotationsmotors

Drückt man nun ein solches Antriebselement, wie in Bild 6-43 b) dargestellt, unter dem Kontaktwinkel Θ gegen ein Objekt, so kann es Antriebsfunktionen ausüben. Das Endstück des Motors würde sich, wenn es frei liegt, auf einer elliptischen Bahn bewegen. Durch die Berührung mit dem Objekt wird im Kontaktfall diese Bahn gestört und entlang der Kontur des Objektes abgelenkt. Durch Reibschluss nimmt die Antriebseinheit das Objekt mit und verschiebt es dabei um den Weg s in Richtung x (Bild 6-43 b)). Wird die Antriebseinheit in einer Anordnung wie in Bild 6-43 c) gegen einen Zylinder gedrückt, so wird dieser bei Anlegen der Wechsellspannung in Rotation versetzt, so dass mit der Anordnung Arbeit geleistet werden kann. Wie im Bild dargestellt, kann durch Drehung des Zylinders beispielsweise eine Last gehoben werden. Ordnet man mehrere Schwinger um den Zylinder an, so kann das ausgeübte Drehmoment weiter erhöht werden.

6.2.2 Aktoren aus Formgedächtnislegierungen

Eine ganz neue Art von Aktoren wird aus sogenannten **Formgedächtnislegierungen (Memory-Metalle)**, kurz **FGL**, hergestellt. FGL's ändern bei Erreichen einer bestimmten Temperatur schlagartig ihren Zustand, ziehen sich zusammen, dehnen sich aus, oder werden **superelastisch** [6.7]. Die Eigenschaften sind unter den bekannten Stillelementen am ehesten mit einem **Bimetall** zu vergleichen. Auch dort

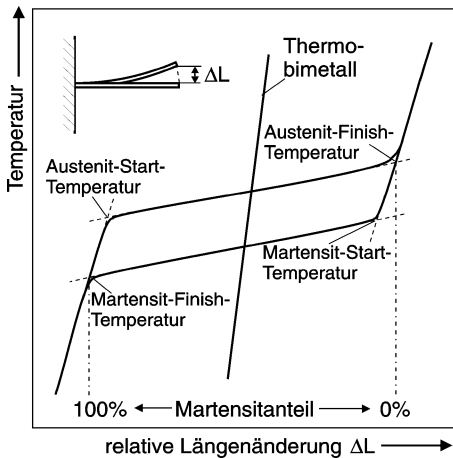


Bild 6-44: Gegenüberstellung des Temperatur-Dehnungsverhaltens eines Thermo-Bimetalls und einer Formgedächtnislegierung

findet man die Doppelfunktion, dass ein Element gleichzeitig Sensor und Aktor sein kann (Beispiel Thermostatschalter). Anders als beim Bimetall, das sich bei Erwärmung kontinuierlich verformt, reagiert die FGL jedoch spontan bei einer bestimmten Temperatur mit einer Formänderung. Ein Vergleich des Temperatur-Dehnungsverhaltens ist in Bild 6-44 dargestellt.

Das Phänomen der sich "erinnernden" Metalle wurde vor mehr als 40 Jahren zufällig an Kupferlegierungen entdeckt. Später fand man den Effekt auch bei Nickel-Titan-Legierungen, die sich bis heute für technische Anwendungen am leistungsfähigsten erwiesen haben. Die physikalische Erklärung liegt in zwei Kristallgitterstrukturen, zwischen denen diese Legierungen hin- und herpendeln. Die Kristallgitterstrukturen werden als

Austenit und **Martensit** bezeichnet (Bild 6-45). Je nach Legierungstyp und Legierungszusammensetzung unterscheidet man drei mögliche Effekte:

- beim **Einwegformgedächtnis** wird eine große bleibende Verformung durch Erwärmen um wenige Kelvin rückgängig gemacht (Bild 6-46).
- beim **Zweiwegeffekt** erinnert sich das Bauteil aus FGL an vorher eingeprägte, ganz unterschiedliche Formen und nimmt diese bei Erwärmen und Abkühlen abwechselnd an (Bild 6-46).

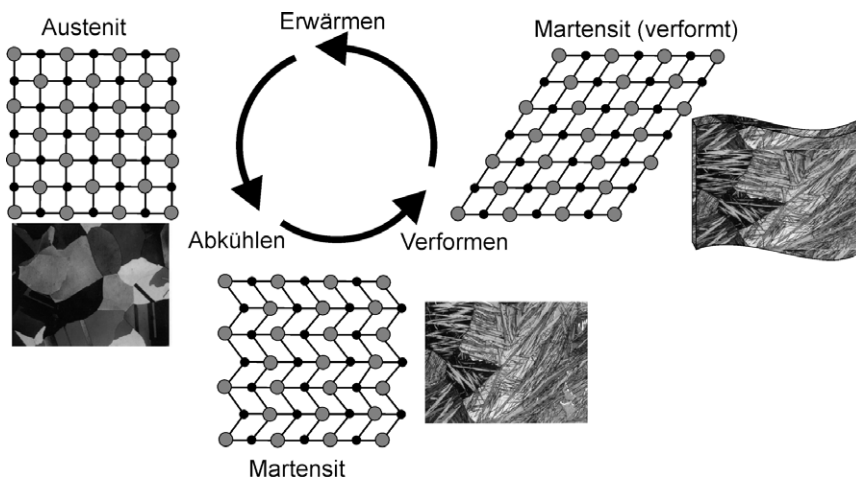


Bild 6-45: Formgedächtnislegierung: Beim Erwärmen und Abkühlen pendelt die Kristallgitterstruktur zwischen der Tieftemperaturphase Martensit und der Hochtemperaturphase Austenit hin und her.

- **Superelastizität** bewirkt ein gummiartiges Verhalten, wobei annähernd konstante Kräfte über große Verformungswege ausgeübt werden. Die Federkraft einer aus superelastischem Draht gewickelte Schraubenfeder ändert sich selbst bei Längenänderung der Feder um den Faktor 2 kaum.

Diese Vorgänge lassen sich 10-20 millionenfach wiederholen. Während der Effektfaltung wird eine hohe Arbeitsleistung pro Volumeneinheit erzeugt, so dass Bauteile aus FGL leistungsstarke Stellglieder abgeben können.

Bauteile aus FGL sind häufig Drähte. In Bild 6-46 sind der Ein- und Zweivegeeffekt schematisch am Beispiel eines Formteils aus Draht (Büroklammer) dargestellt. Das Bauteil wird in der Herstellungsphase in martensitischer Form erzeugt. Beim Fertigungsverfahren wird der Werkstoff unter Formzwang einer bestimmten Wärmebehandlung unterzogen. Liegt der Einvegeeffekt vor, so kann man das Bauteil verformen und diese Verformung wird bei Verformungstemperatur beibehalten. Erwärmt man nun den Werkstoff über eine für die jeweilige Legierungszusammensetzung charakteristische Temperatur ($70-80^{\circ}\text{C}$), so wandelt sich das Gefüge in Austenit um und nimmt dabei wieder die ursprüngliche Form an; das Bauteil hat sich an seine im Fertigungsprozess "erlernte" Form "erinnert". Bei erneuter Abkühlung wandelt sich das Gefüge wieder in Martensit um und es bleibt die dann vorhandene Form erhalten.

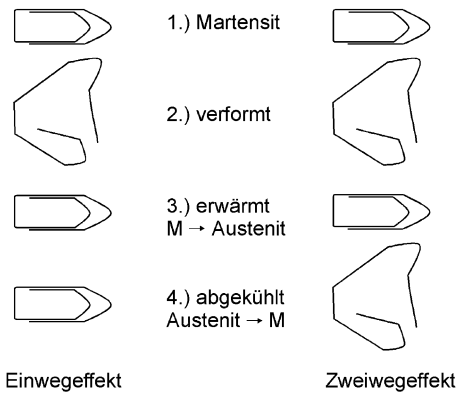


Bild 6-46: Schematische Darstellung des Ein- und Zweivegeeffektes von Formgedächtnislegierungen

Beim Zweivegeeffekt wird durch einen speziellen Fertigungsprozess die Legierung dazu gebracht ihre Form allein durch Temperaturwechsel zu ändern. Man kann ihr sozusagen zwei charakteristische Formen einprägen, die sich wechselseitig jeweils ineinander umwandeln, wenn eine charakteristische "Schalttemperatur" über- oder unterschritten wird.

Drähte, Bleche oder Federn können somit als Stellantriebe eingesetzt werden, die bei Erreichen der Schalttemperatur eine Stellbewegung durchführen. Die Wärmeübertragung auf das FGL-Bauelement kann dabei sowohl durch Strahlung als auch durch Wärmeleitung innerhalb eines Fluides erfolgen. Aber auch die Erwärmung mittels eines durch das Bauteil geleiteten Stroms kann verwendet werden, so dass das Element auf ein Stromsignal (Temperatur ist abhängig vom Stromfluss) mit einer Stellbewegung reagiert.

Bauteile aus Formgedächtnislegierungen werden als thermische Stellelemente in der Automobilindustrie und vielen weite-

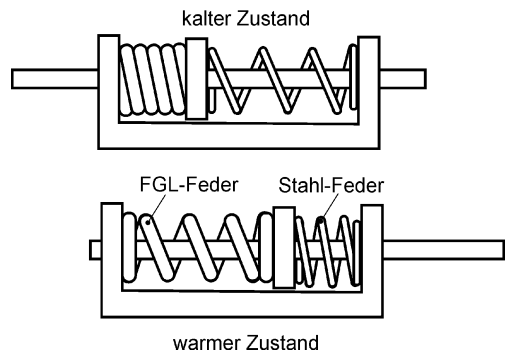


Bild 6-47: Stellantrieb mit FGL-Feder

ren Anwendungen in der Mess- und Regeltechnik, Hausgerätetechnik, Luft- und Raumfahrtindustrie sowie der Medizintechnik eingesetzt. Bild 6-47 zeigt schematisch einen Stellantrieb mit einer Antriebsfeder aus Einweg-FGL. Im kalten Zustand wird sie durch eine Stahlfeder zusammengedrückt. Bei Überschreiten der Schalttemperatur ändert die FGL-Feder ihre Form und drückt die Stahlfeder zusammen. Dadurch bewegt sich die Schaltachse in Stellrichtung.

6.2.3 Dehnstoffaktoren

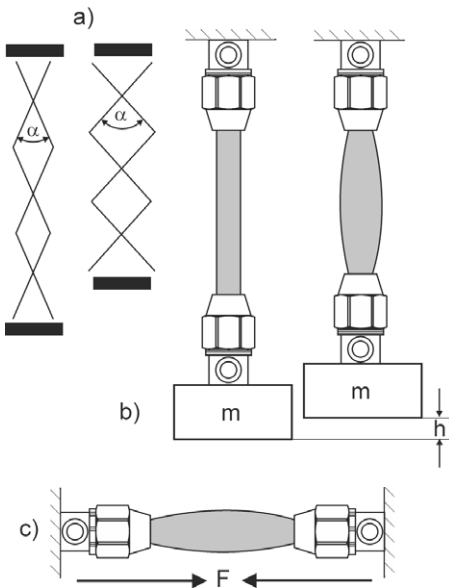


Bild 6-48: Fluidischer Muskel a) Funktionsprinzip b) Heben von Massen c) Kraftausübung

Um lineare Stellbewegungen auszuführen, ist ein Pneumatik-Zylinder, wie er in Bild 6-27 dargestellt ist, ein sehr kostengünstiger einfacher Aktor, da er direkt ohne Getriebeelemente eingesetzt werden kann. Der in Bild 6-48 dargestellte Aktor ist noch bedeutend einfacher im Aufbau und kann vergleichbare Bewegungen erzeugen. Dieses auch als "fluidischer Muskel" (Fluidic Muscle) bezeichnete Membran-Kontraktions-System ruft einer Längenänderung eines Kunststoffschlauches durch Zusammenziehung unter Druck hervor.

Dies beruht auf der Kombination von fluidisch dichtem, flexiblen Schlauch und Umspinnung mit festen Fasern in Rautenform. Im nicht verformten Zustand haben diese Fasern einen Kreuzungswinkel α (Bild 6-48 a)). Durch das einströmende Fluid wird die Gitterstruktur verformt (α wird größer), wodurch eine Zugkraft in Axialrichtung entsteht. Verkürzungen des Schlauches um bis zu 25% sind möglich. Dieser Aktor kann die 10-fache Kraft im Vergleich zu einem konventionellen Pneumatik-

Zylinder gleicher Baugröße aufbringen, bei einem auf 40% verringertem Energiebedarf. Abhängig vom Schlauchdurchmesser können bei 6 bar Luftdruck Kräfte von einigen Hundert bis zu einigen Tausend Newton aufgebracht werden.

Es sind Anwendungen zum Heben und Senken von Lasten (Bild 6-48 b)), Kraftübertragung (Bild 6-48 c)) oder als Luftfeder möglich. Wegen seiner hohen Anfangskraft und Beschleunigung ist der fluidische Muskel sehr gut für Aufgaben mit hoher Anfangsbeschleunigung geeignet, wobei außerdem kein Stick-Slip-Effekt auftritt. Da das System denkbar einfach aufgebaut und sehr leicht ist, ist es vor allem für mobile Anwendungen gut geeignet.

Ein weiterer Wirkmechanismus für Aktoren, ist die Verwendung **elektorrheologischer Flüssigkeiten** (ERF's). Daneben gibt es noch die artverwandten **magnetorheologischen Flüssigkeiten** (MRF's). Der in solchen Flüssigkeiten auftretende Effekt wurde bereits vor etwa 60 Jahren entdeckt. Er besteht darin, dass Suspensi-

onen von nichtmetallischen Teilchen in elektrisch nicht leitenden Ölen unter Einfluss eines elektrischen Feldes innerhalb von Sekundenbruchteilen ihre Viskosität ändern. Beim Einschalten des Feldes erstarrt die Flüssigkeit aufgrund von Polarisierungseffekten bei entsprechend hoher Feldstärke zu einem plastischen Körper. Nach Ausschalten des elektrischen Feldes wird die Suspension wieder dünnflüssig. Für diesen beliebig oft wiederholbaren Prozess ist nur eine geringe Leistung von wenigen Watt erforderlich. Anwenden kann man diesen Effekt beispielsweise für parametrisierbare Schwingungsdämpfer in Maschinen und Anlagen, die in ihrer Dämpfungskonstante verstellt werden können. Bild 6-49 zeigt als Beispiel Stoßdämpfer, bei denen die Eigenschaften der ERF entweder im Hauptzylinder oder in einem Bypass durch ein elektrisches Feld beeinflusst werden können [6.8]. Dadurch kann das Dämpfungsverhalten des Stoßdämpfers in Abhängigkeit des Straßenzustandes verstellt und der Fahrkomfort optimiert werden. Unter den Beispielen für mechatronische Systeme im nächsten Kapitel wird noch eine weitere Möglichkeit vorgestellt, die Dämpfung zwischen Rad und Karosserie eines Fahrzeugs zu verbessern.

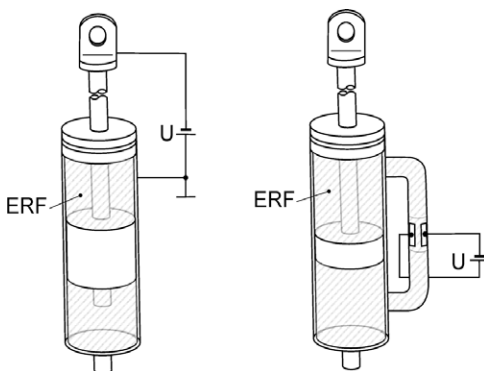


Bild 6-49: Stoßdämpfer mit einstellbarer Dämpfung durch Verwendung einer elektrorheologischen Flüssigkeit (ERF)

6.2.4 Mikrostrukturierte Aktoren

Große Entwicklungspotentiale liegen auch im Bereich der mikrostrukturierten Aktoren. Hierfür kommen vor allem Fertigungsverfahren wie die Silizium-Ätztechnik, die als Fertigungstechnik für Sensorelemente bereits angesprochen wurde, in Frage. Solche in der gleichen Mikrotechnologie aus Silizium hergestellten Mikrosysteme werden im internationalen Bereich auch als **MEMS** (**Micro-Electro-Mechanical System**) bezeichnet. Die bereits aus der Fertigung rein elektronischer Bauteile bekannten Verfahren zur Erzeugung von Mikrostrukturen bieten ausgezeichnete Möglichkeiten zur Erzeugung von MEMS aus Silizium. Das LIGA-Verfahren erweiterte in den 1980er-Jahren die der Mikrosystemtechnik zur Verfügung stehende Materialpalette um Kunststoffe, Metalle und Keramiken [6.9].

Bild 6-50 zeigt für das **LIGA-Verfahren** (Lithografie, Galvanoformung und Abformtechnik) in den Teilbildern a) bis e) die 5 Prozessstufen zur Herstellung von Mikrostrukturen. Zunächst wird der sogenannte Resist, eine bis zu 700 nm starke Schicht aus dem Kunststoff Polymethylmethacrylat (PMMA), auf ein metallisches Substrat aufgebracht. Mittels einer Maske und der parallel einfallenden Röntgenstrahlung einer Synchrotronquelle mit einer Wellenlänge von etwa 0,5 nm wird ein Strukturbild auf den Resist übertragen (a). Die energiereiche Strahlung verursacht in den belichteten Teilen des Kunststoffs Brüche an den langen Molekülketten. Mit einem geeigneten Lösungsmittel lassen sich diese Bereiche entfernen und es bleibt

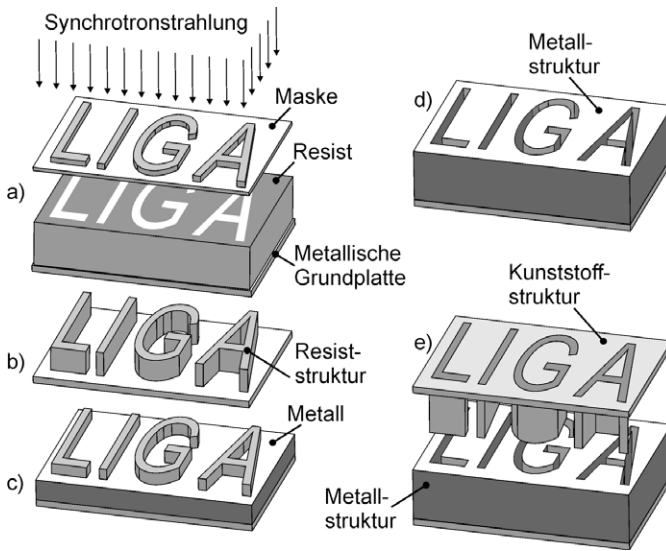


Bild 6-50: Prozessstufen des LIGA-Verfahrens

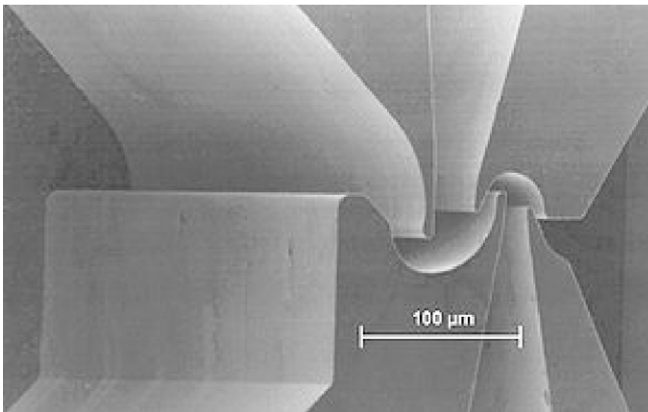


Bild 6-51: Mit dem LIGA-Verfahren gefertigte Trenndüse (KFZ Karlsruhe)

die Primärstruktur im PMMA stehen (b). Der dritte Schritt ist die Galvanoformung, in dem im galvanischen Bad auf dem Substrat eine Metallschicht aufwächst und den zuvor weggelösten Kunststoff ersetzt (c). Nach dem Herauslösen des unbestrahlten PMMA erhält man schließlich eine metallische Sekundärstruktur, die ein genaues Negativ der Primärstruktur darstellt (d). Sie lässt sich nun als Werkzeug in eine Spritzgieß- oder Prägemaschine einsetzen, um die Primärstruktur zu vervielfältigen (e). Auch diese abgeformten Kunststoffstrukturen lassen sich für die Massenfertigung durch erneute Galvanoformung in metallische Negativstrukturen überführen. Bild 6-51 zeigt ein mikroskopisches Bild einer mit dem LIGA-Verfahren hergestellten so genannten Trenndüse für die Trennung von Uranhexafluorid. Für diese Aufgabe wurde das Verfahren

ursprünglich am Kernforschungszentrum Karlsruhe entwickelt. Heute werden mit diesem Verfahren Mikrogetriebe und Bauteile für Mikro-Elektromotoren gefertigt.

Während im Makrobereich elektrische Aktoren meist mit Magnetkräften arbeiten (Elektromotor, Elektromagnet), kann man im Mikrobereich vor allem elektrostatische Kräfte nutzen. So wirkt in einem geladenen Plattenkondensator mit Luft als Dielektrikum eine Anziehungskraft von (Bedeutung der Größen siehe 5.2.3)

$$F = 0,5 \cdot \varepsilon_0 \cdot \varepsilon_r \cdot \frac{U^2}{d^2} \cdot A \quad .$$

In Makrosystemen sind diese Kräfte bei kleinen Spannungen U sehr klein. Ist der Plattenabstand d wie in Mikrosystemen (MEMS) aber sehr klein, so treten dort verwertbare Kräfte auf. Um ausreichend große Kräfte zu erzeugen, kann man zusätzlich auf kleinem Raum die Plattenoberfläche A durch ineinander geschachtelte

Strukturen wie beim **Kammantrieb** in Bild 6-52 vergrößern. Eine der Platten des Kamm-Antriebs ist feststehend zum Gehäuse, während die andere durch Spannungsänderung zwischen den Platten in Pfeilrichtung bewegt werden kann. Anders als bei einem elektromagnetischen Antrieb, lässt sich die zwischen den Antriebsstrukturen resultierende Kraftwirkung nicht umpolen. Für die Realisierung von Bauteilen mit positiver und negativer Wirkrichtung sind daher zwei entgegengesetzt wirkende Antriebe notwendig.

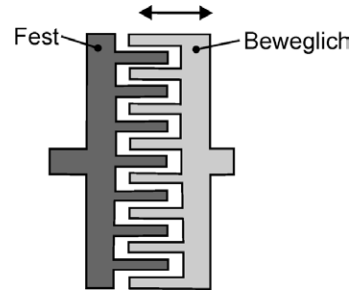


Bild 6-52: Prinzip des Kammantriebs

Elektrostatische Kammantriebe (Comb Drives) stellen eine der wichtigsten Komponenten der Mikromechanik dar. Bei Kammantrieben mit rechteckigen Fingern ändert sich die Kapazität linear mit der Auslenkung, was zu einer positionsunabhängigen Antriebskraft führt. Abweichungen von diesem Verhalten zeigen sich nur an den Enden des Auslenkungsbereichs. Bild 6-53 zeigt eine mikroskopische Aufnahme der kammartigen Plattenanordnungen innerhalb eines Antriebssystems, bei dem die Plattenfläche zur Erreichung größerer Kräfte durch mehrere Kammstrukturen vervielfacht wurde. Das gleiche Bewegungsprinzip kann natürlich auch für rotierende Antriebe genutzt werden. Kamm-Strukturen sind sowohl in Mikrosensoren als auch in Mikroaktoren anzutreffen, beispielsweise in Gyroskopen oder optischen Schaltern für die Telekommunikation.

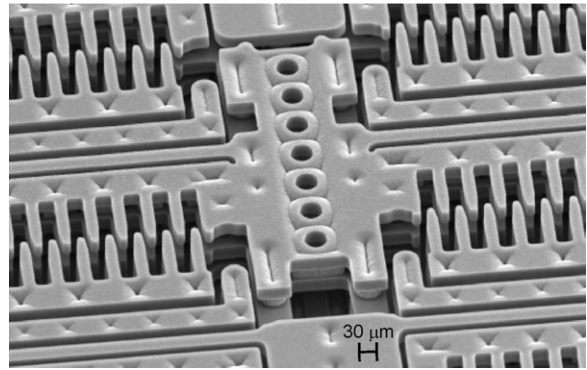


Bild 6-53: Elektrostatischer Kammantrieb

Weitere Methoden der Mikrostrukturierung sind durch abtragende Verfahren möglich. Bild 6-54 zeigt die rasterelektronenmikroskopische Aufnahme eines Mikrozahnrades mit einem Durchmesser von weniger als 100 µm, das mit Hilfe der Strahlung eines Eximer-Lasers (Strahlung im UV-Bereich) gefertigt wurde. Als Größenvergleich liegt das Zahnrad auf einem menschlichen Haar.

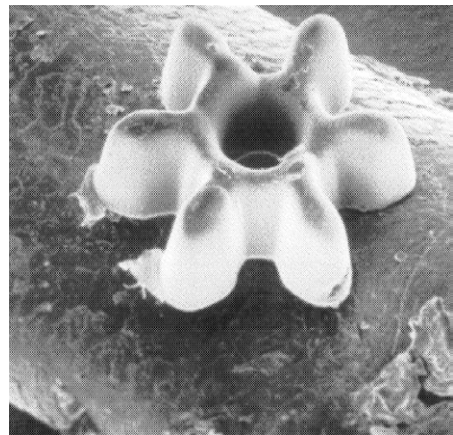


Bild 6-54: Mit Laserstrahlung erzeugtes Mikrozahnrad (FH Münster)

7 Automatisierungstechnik

Schaut man sich die bereits mehrfach angesprochene Grundstruktur eines mechatronischen Systems in Bild 7-1 an, so sieht man, dass die Beschreibung des Verhaltens eines Teils der Systemkomponenten bereits in den vorausgegangenen Kapiteln erfolgt ist. Die dunkelgrau unterlegten Systemteile sind mechanische Systeme, deren mathematische Behandlung in den Kapiteln 3 und 4 ausführlich vorgestellt wurde. Die hellgrau unterlegten Systemteile sind die Sensoren und Aktoren, welche ebenfalls in den Kapiteln 5 und 6 behandelt wurden.

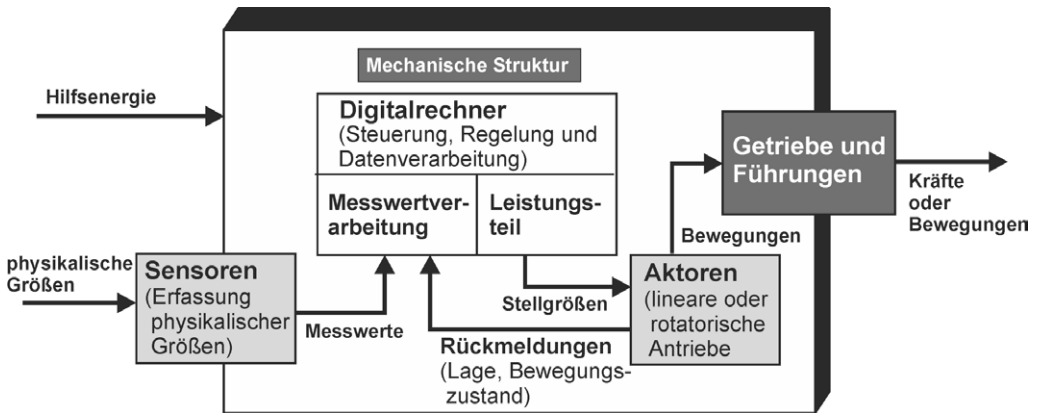


Bild 7-1: Grundstruktur eines mechatronischen Systems

Noch nicht behandelt wurde der zentrale Bereich eines mechatronischen Systems, die Verwertung von gemessenen Daten aus der Außenwelt und aus dem System selber, die Datenverarbeitung, sowie die Steuerung und Regelung der Stellsysteme.

Mit den jetzt bekannten Methoden kann man die Teile des Systems, die Kräfte oder Bewegungen erzeugen sollen, modellieren und mathematisch beschreiben. Es fehlen aber noch die Möglichkeiten, diesen Systemen ein dynamisches Sollverhalten vorzugeben und bei Abweichungen von diesem Verhalten entsprechend darauf zu reagieren. Die Methoden, Verfahren und Einrichtungen hierzu beschreibt die **Prozessdatenverarbeitung** und **Regelungstechnik** die man zusammenfassend auch als **Automatisierungstechnik** bezeichnet.

7.1 Automatisierungskonzepte

Die Automatisierungstechnik ist ein Gebiet, dessen erste Anfänge im Altertum¹ liegen, die aber vor allem mit dem Aufkommen der Dampfmaschine als Antrieb für Arbeitsmaschinen im frühen 18. Jahrhundert an Bedeutung gewann. Schon der englische Erfinder James Watt (1736-1819), der wesentliche Verbesserungen für die Dampfmaschine entwickelte, ließ sich 1788 die in Bild 7-2 dargestellte Drehzahlre-

1 So baute im 1. Jahrhundert n. Chr. Heron, ein Vertreter der berühmten Mechanikerschule von Alexandria, Automaten mit hydraulischer Steuerung

gelung mit Hilfe eines Fliehkraftpendels patentieren. Bei zunehmender Drehzahl, bewegen sich die Gewichte des Fliehkraftpendels durch die zunehmenden Fliehkräfte nach außen und heben über ein Gestänge einen Hebel an, der den Dampfeintritt in die Dampfmaschine über einen Schieber in der Dampfzuleitung verringert. Dadurch sinkt die Drehzahl wieder und der gegenteilige Effekt tritt ein, d. h. eine **Selbstregulierung** findet statt.

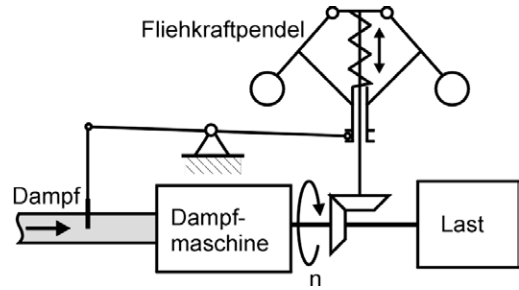


Bild 7-2: Drehzahlregelung einer Dampfmaschine mit Hilfe eines mechanischen Fliehkraftreglers

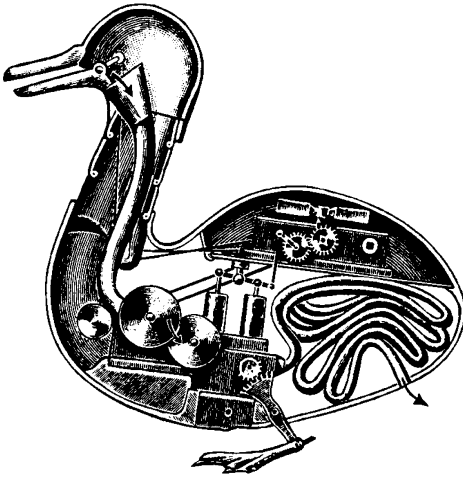


Bild 7-3: Mechanische Ente von J.D. Vaucanson um 1738

Das 18. Jahrhundert ist auch die Blütezeit der mechanischen Automaten, wie die von dem Franzosen Jaques de Vaucanson 1738 der Académie Royale de Sciences in Paris vorgeführte mechanische Ente (Bild 7-3), die mit den Flügeln schlagen, quaken, fressen und verdauen konnte. Ein anderes Beispiel für solche mechanischen Automaten, die ausschließlich zur Unterhaltung eines technisch noch stark beeindruckbaren Publikums dienten, ist der in Bild 7-4 dargestellte „automatische Schriftsteller“ von Pierre Jaquet-Droz und Jean Frédéric Leschot aus dem Jahre 1773. Er konnte Texte wie „Wir sind die Androiden Jaquez-Droz“ mit einem Stift auf Papier schreiben. Das Meyer'sche Lexikon definiert den Begriff 1888 so: „Automat - ein mechanisches Kunstwerk, welches mittels eines inneren Mechanismus die Tätigkeit lebender Wesen nachahmt und meist auch an Gestalt diesen nachgebildet ist“. Solche **intelligenten Maschinen** waren damals reine Automaten mit linearen Ablaufsteuerungen, deren **Programm** durch die Hardware festgelegt wurde.



Bild 7-4: Der „automatische Schriftsteller“ von P. Jaquet-Droz und J.F. Leschot von 1773

Moderne Automaten haben Automatisierungskonzepte, deren Arbeitsanweisungen nur noch durch Software-Programme bestimmt werden, aber häufig relativ starr sind. Zu den Automaten zählen aber auch die mechatronischen Systeme, bei denen man häufig versucht Eigenintelligenz, wie sie für Lebewesen charakteristisch ist, durch Software zu erzeugen, um solche Systeme sehr flexibel gegenüber Abweichungen vom Soll-Zustand zu machen.

7.1.1 Intelligente Maschinen

Maschinen und ihre Abläufe müssen gesteuert und geregelt werden. Bei einfachen Maschinen erfolgt dies durch einen Maschinenbediener. Automatisierte Maschinen enthalten heute in der Regel einen Computer auf dem ein Anwenderprogrammiersystem läuft, mit dessen Hilfe der Maschine Arbeitsabläufe vorgegeben werden können. Typische Beispiele für solche Maschinen sind numerisch gesteuerte Werkzeugmaschinen oder Industrieroboter. Das Automatisierungskonzept solcher Anlagen ist in Bild 7-5 dargestellt. Die auf diesen Maschinen programmierten Arbeitsabläufe ändern sich nur, wenn durch den Maschinenbediener oder durch den Fertigungsleitreechner das **Anwender-** oder **Teileprogramm** geändert wird, oder durch manuelle Eingaben des Maschinenbedieners Modifikationen des Anwenderprogramms vorgenommen werden.

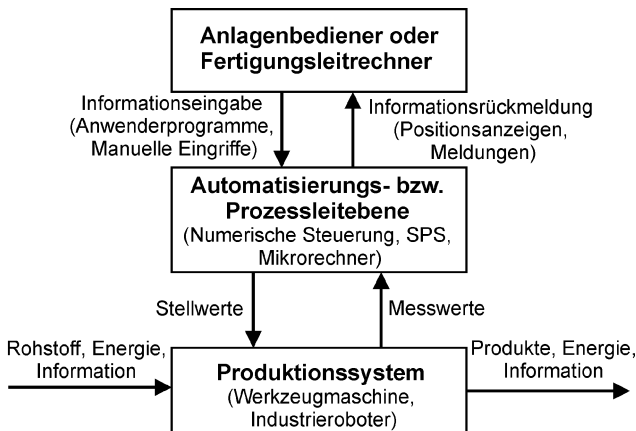


Bild 7-5: Automatisierungskonzept von Produktionssystemen

Es gibt jedoch auch Maschinen und Anlagen, die ein gewisses Maß an Selbstanpassung an bestimmte sich ändernde äußere oder innere Einflüsse besitzen. Beispiele hierfür sind Änderungen der Zerspanungswerte bei einer Werkzeugmaschine bei Ratterschwingungen oder Werkzeugverschleiß, oder Bahndatenbeeinflussung durch Sensordaten beim Schweißen mit einem Industrieroboter. Solche Maschinen kann man auch als **intelligente Maschinen**

bezeichnen [1.6]. Da schon die Definition des Begriffs Intelligenz für Menschen sehr problematisch und emotionsbeladen ist, soll hier für automatisierte Maschinen eine sehr einfache und pragmatische Definition gelten:

Intelligenz ist die Fähigkeit eines Systems ein Ziel, ein Verhalten oder einen bestimmten inneren Zustand bei Auftreten unbestimmter Zustände innerhalb oder außerhalb des Systems zu erreichen oder beizubehalten.

Dabei sind solche unbestimmten Zustände:

- Das Auftreten von nicht erwarteten Ereignissen, sowohl innerhalb des Systems, beispielsweise durch Bauteilversagen, als auch außerhalb des Systems, wie

nicht vorhersehbare Änderung im Umfeld des Systems, in dem es arbeitet.

- Unvollständige, unwahrscheinliche oder unzuverlässige Informationen, aufgrund derer das System seine nächsten Schritte planen soll.

Fest programmierte Maschinen, können nur eindeutigen Arbeitsanweisungen folgen. Im Gegensatz dazu bedeutet Intelligenz bei Maschinen demnach nicht mehr als die Fähigkeit, sich an solche Situationen eigenständig anpassen zu können, deren Auftreten in der Regel zwar hohe Wahrscheinlichkeit besitzt, aber nicht immer exakt vorhergeplant werden kann.

Auf der einfachsten Ebene der Intelligenz besitzt ein solches System Sensoren, mit denen es Informationen aus seinem Inneren und aus der Außenwelt aufnimmt und dadurch seine Handlungsweisen anpassen kann. Dies sind einfache mechatronische Systeme wie der in Bild 7-6 dargestellte Industrieroboter mit einem **Bildverarbeitungssystem** (Vision-System) [7.13]. Die Form der Teile (Kleinteile wie beispielsweise unterschiedliche Schrauben), die gehandhabt werden sollen und deren Lage, ist dem Industrieroboter unbekannt. Die Teile werden lose auf einem Förderband angeliefert, so dass die Greifpositionen und das Maß für das Öffnen des Greifers nicht vorprogrammiert werden können. Die fehlenden Informationen werden von einem Vision-System aus dem Bild einer CCD-Kamera berechnet. Dieses Vision-System bestimmt die Art der Teile, ihre Lage im Arbeitsbereich des Roboters, sowie die Information, ob das Teil frei liegt, so dass es ohne Kollisionsgefahr vom Greifer aufgenommen werden kann. Die Unbestimmtheit der Umwelt hält sich jedoch in gewissen Grenzen, da nur Teile erkannt werden können, deren Form dem System vorher durch Training beigebracht wurden.

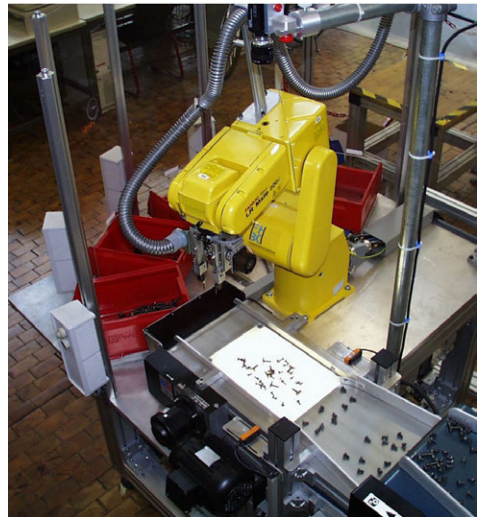


Bild 7-6: Industrieroboter mit Vision-System zum Sortieren von Kleinteilen

Es gibt natürlich weitaus intelligenteren Systeme, die in der Lage sind durch ihre Interaktion mit der Umgebung neue Fähigkeiten oder Verhaltensweisen zu lernen. Darüber hinaus können solche Systeme, um spezielle Ziele zu erreichen, sogar ihre Umwelt beeinflussen bzw. neue eigene Ziele formulieren und verfolgen. Im Bereich sich autonom bewegend und arbeitender Roboter sind solche Verhaltensweisen erforderlich.

Es besteht natürlich kein Zwang, dass automatisierte Maschinen Intelligenz besitzen müssen. Unter deterministischen d.h. vorhersagbaren Bedingungen ist es bedeutend einfacher und sicherer nach deterministischen Prinzipien ablaufende starre Programme zur Steuerung von Maschinen zu verwenden. Die Stärke von eindeutig programmiertem Verhalten ist Genauigkeit und Reproduzierbarkeit, seine Schwächen liegen in der Unfähigkeit auf unvorhergesehene Ereignisse zu reagieren.

7.1.2 Steuerung und Regelung

Die Drehzahlregelung mit einem Fliehkraftpendel (Bild 7-2) ist ein frühes Beispiel für ein selbstregulierendes System, das damit ein gewisses Maß von "Intelligenz" besitzt. Bei der Vorgabe von Sollwerten für eine Maschine, die der Maschine ein bestimmtes Systemverhalten aufprägen sollen, gibt es verschiedene Konzepte, die **Steuerung** und die **Regelung**.

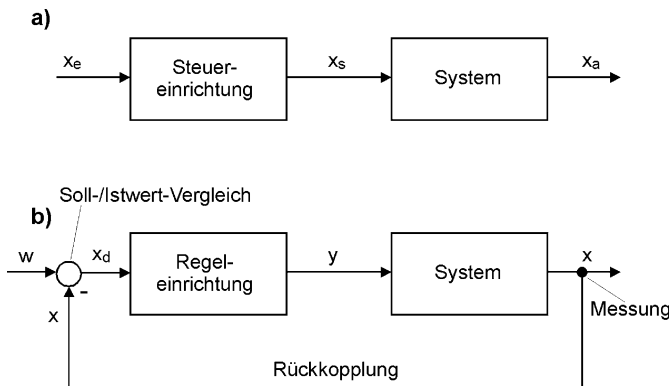


Bild 7-7: Wirkprinzipien von a) Steuerung b) Regelung

Im folgenden werden die zu betrachtenden Teilsysteme, wie in der Regelungstechnik üblich, als Blöcke behandelt. In diese Blöcke werden die Eingangsgrößen hineingeführt und aus ihnen kommen Ausgangsgrößen heraus. Innerhalb des Blocks befindet sich ein mathematisches Modell, das das Übertragungsverhalten des Systems vom Ein- zum

Ausgang beschreibt. Arbeiten mehrere solcher Blöcke (math. Modelle) zusammen, so sind sie durch Signale miteinander verbunden. Ein solches Schema aus Funktionsblöcken und den zugehörigen Signalaustausch bezeichnet man als **Signalflussplan**. Bild 7-7 zeigt den sehr einfachen **Signalflussplan** einer Steuerung und einer Regelung um diese begrifflich voneinander abzugrenzen. Beide Systeme verfolgen das gleiche Ziel, nämlich das dynamische Verhalten eines Systems nach bestimmten Vorgaben auf definierte Weise zu beeinflussen. Die Vorgehensweisen, um das jeweilige Ziel zu erreichen, sind jedoch bei einer Steuerung und einer Regelung grundsätzlich verschieden.

Wie man in Bild 7-7 a) sieht, besteht die Steuerung aus einer **offenen Wirkkette** (**Steuerkette**) des zu steuernden Systems und der Steuereinrichtung (**Steuerung**). Von links wird in die Steuerung ein Eingangssignal x_e hineingegeben, woraus diese ein **Stellsignal** x_s erzeugt, das das zu steuernde System zu definierten dynamischen Veränderungen der Ausgangsgröße x_a veranlasst. Natürlich wird mit der Steuerung beabsichtigt, durch die Vorgabe am Eingang ein genau vorhersehbares Ausgangssignal zu erzeugen. Voraussetzung dafür, dass am Ausgang der Steuerkette tatsächlich dieses Ausgangssignal auftritt, ist, dass das Systemverhalten durch das mathematische Modell exakt beschrieben wird und dass keine unbekannten Einflüsse (Störungen) von außen auf das System wirken. Enthält das mathematische Modell z. B. Parameter, die aufgrund von im Modell nicht erfassten Einflüssen schwanken können, oder wurden Störgrößen im Modell nicht erfasst, die aber wesentlichen Einfluss auf das System haben, so ist das Ausgangsverhalten der Steuerkette nicht vorhersehbar.

Bei rein **binär** arbeitenden Systemen und Steuereinrichtungen, die nur zwei deutlich unterschiedene stabile Zustände kennen, ist das Systemverhalten in der Regel

exakt bekannt und eindeutig mathematisch beschreibbar. Bei solchen Systemen ist die reine Steuerung der Regelfall. Bei **analogen** Systemen, deren Ein- und Ausgangsgrößen innerhalb eines bestimmten Wertebereiches kontinuierlich veränderbar sind, ist diese Eindeutigkeit der Beschreibung und des Verhaltens häufig nicht gegeben, d. h. das Modell ist nur unzureichend bekannt und die Größe und Art der möglichen Störungen ist nicht vorhersehbar. Hier wird daher das Wirkungsprinzip der **Regelung** eingesetzt. Wie man in Bild 7-7 b) sieht, sind wesentliche Komponenten wie schon bei der Steuerung die Regeleinrichtung (**Regler**) und das zu regelnde System, das auch häufig als **Regelstrecke** bezeichnet wird. Das gewünschte dynamische Verhalten der Ausgangsgröße x der Regelstrecke (**Regelgröße**) wird durch die Eingangsgröße w (**Führungsgröße**, **Sollwert**) vorgegeben. Wirken jetzt auf die Strecke unvorhergesehene Störungen ein, so wäre der gewünschte Zusammenhang zwischen Sollwert und Regelgröße (**Istwert**) mit einer offenen Wirkkette, wie bei der Steuerung, nicht mehr einzuhalten. Dieses Verhalten besitzt aber die Regelung, bei der der Istwert durch ein geeignetes Messverfahren bestimmt (gemessen) wird und das Ergebnis der Messung einer Vergleichsstelle mit dem Sollwert zugeführt wird (**Soll-/Istwertvergleich**). Der Vergleich erfolgt dadurch, dass vom Sollwert der Istwert abgezogen wird, was die sog. **Regeldifferenz** erzeugt. Im Idealfall, in dem die Regelgröße nicht vom Sollwert abweicht, ist die Regeldifferenz gleich Null und das Gesamtsystem aus Regler und Strecke ist in einem stabilen Zustand. Tritt durch äußere Störungen oder auch innere Schwankungen des Systems eine Abweichung der Regelgröße vom Sollzustand auf, so nimmt die Regeldifferenz einen von Null abweichenden Wert an. Da die Regelabweichung dem Regler zugeführt wird, "merkt" dieser, dass eine Abweichung aufgetreten ist und reagiert darauf in der Weise, dass er seine Ausgangsgröße, den **Stellwert** y , verändert. Dadurch wird die Regelgröße mehr oder weniger schnell in den alten Zustand zurückgeführt. Dies alles geschieht automatisch ohne das Eingreifen einer Person. Das Prinzip der Rückführung der gemessenen Ausgangsgröße an den Eingang zum Vergleich mit der Eingangsgröße bezeichnet man auch als **Rückkopplung**.

Das Prinzip der Selbstregulierung durch Rückkopplung ist in automatisierten Maschinen weit verbreitet. Wurden Steuerungen und Regelungen in der Vergangenheit meist durch digitale und analoge diskrete Bauteile aufgebaut, so werden heute die erforderlichen Algorithmen häufig auf einem Digitalrechner, meist einem Mikrorechner, als Programme implementiert. Solche Programme führen dazu folgende Schritte aus:

- 1.) Messe die Werte aller Variablen, die das Verhalten des Systems repräsentieren mit Hilfe von Sensoren.
- 2.) Vergleiche die Messwerte mit den Sollwerten.
- 3.) Ist eine Regelabweichung vorhanden so entscheide, welche Maßnahme die Abweichung minimieren kann.
- 4.) Führe die erforderliche Maßnahme durch.
- 5.) Springe zum Schritt 1.) und erzeuge dadurch die Rückkopplung.

Ein wichtiger Punkt, in dem sich die einfache Regelung in ihrem Grad von Intelli-

genz von anderen Prinzipien unterscheidet, ist der Entscheidungsmechanismus des Verfahrensschrittes 3.). Solche Prinzipien, die einen höheren Grad von Intelligenz umfassen, werden in den folgenden Abschnitten beschrieben.

7.1.3 Schlussfolgern und regelbasiertes Wissen

Hat man es mit einem teilweise unbekannten Umfeld zu tun, in dem eine Maschine arbeitet, so ist als intelligentes Verhalten mehr als nur einfache Selbstregulierung wie in Bild 7-7 b) erforderlich. Das kann daran liegen, dass bestimmte Eigenschaften des Umfeldes nicht messbar sind, oder dass sich mehrere Eigenschaften möglicherweise sogar in unerwarteter Art und Weise gleichzeitig ändern. Häufig ist dann auch nicht vorausberechenbar wie die Maschine auf alle möglichen Kombinationen von Veränderungen reagieren sollte. In einem solchen Fall sollte die Maschine in der Lage sein selbst **Schlussfolgerungen** aus der Situation zu ziehen. Damit das möglich ist, muss das System über einen Satz von Regeln verfügen, die sein Wissen über die Welt darstellt. Dazu gehören:

- Wissen über die eigenen Aufgaben und Ziele
- Wissen über die eigenen Fähigkeiten
- Wissen über das Umfeld, in dem die Maschine arbeitet

Die Maschine muss dann beim Auftreten einer Änderung ihr Regelwerk durchsuchen, um das Wissen aufzufinden, das ihr hilft, die Handlung zu schlussfolgern, die die Änderung ausgleichen oder sogar ausnutzen kann. Ist das Wissen der Maschine

über die Zukunft unvollständig oder unzuverlässig, so muss die Maschine ebenfalls aufgrund ihres gespeicherten Wissens die unter den gegebenen Umständen sinnvollste Handlungsweise schlussfolgern.

Um solches regelbasiertes Wissen zu erhalten, benutzen Entwickler solcher Systeme unterschiedliche Quellen. Dazu gehört auch, das Wissen von Maschinenbedienern, die bisher solche Maschinen erfolgreich gesteuert und geregelt haben, abzufragen und dieses in einer computerlesbaren **Wissensbasis** abzuspeichern. Eine solche Wissensbasis und der zugehörige Mechanismus zum Schlussfolgern (**Inferenzsystem**) werden auch als **Expertensystem** bezeichnet.

Eine typische Maschine, bei der derartige Mechanismen zur Steuerung erforderlich sind, ist ein sich **autonom** bewegendes Fahrzeug, das beispielsweise Güter ent-



Bild 7-8: Autonomer mobiler Roboter *Blue Kiss* (Hochschule Bochum)

lang eines Flurs transportiert oder an verschiedenen Stellen Arbeitsaufgaben verrichtet (Bild 7-8). Autonom bedeutet dabei, dass das Fahrzeug weder durch einen Bediener, noch durch Kommunikation mit einem anderen externen System gesteuert wird.

Ein solches Fahrzeug hat Sensoren wie Laser-Radar, Sonar und Video-Kamera um Informationen über die Umwelt zu sammeln. Außerdem benötigt es einen Lageplan seiner Umwelt, der Informationen über die möglichen Wege entlang der Flure wie Länge, Breite, mögliche Fahrgeschwindigkeiten und über alle Durchgänge oder möglichen Hindernisse enthält.

Während der Fahrt durch die Gänge nimmt das System über seine Sensoren ständig Informationen aus seinem Umfeld auf und vergleicht die Informationen mit seinem Lageplan. Dadurch **navigiert** das Fahrzeug entlang der Gänge und kann Hindernisse umfahren. Außer bekannten Objekten, die als feste Hindernisse dem Fahrzeug im Weg stehen, können aber auch neue feststehende oder bewegliche Objekte wie Personen, andere Fahrzeuge oder abgestellte Gegenstände plötzlich den Weg versperren. Um nun zu entscheiden, wie sich das Fahrzeug verhalten soll, muss es versuchen Art und Form des Objektes, die von den Sensoren aufgenommen wird, mit den Eigenschaften ihm bekannter Objekte zu vergleichen. Weiterhin muss festgestellt werden, ob das Objekt still steht oder wohin es sich mit welcher Geschwindigkeit bewegt.

Sind die zur Verfügung stehenden Informationen über das Objekt unvollständig, muss das Fahrzeug eine Entscheidung treffen, wie es sich verhalten soll. Mögliche Reaktionen wären ein Warnsignal abzugeben, langsamer zu werden oder die Fahrtrichtung zu ändern. Während des Entscheidungsprozesses durchsucht das System seine Wissensbasis nach allen möglichen Alternativen und wird wenn es die Situation erlaubt diejenige Option auswählen, die die eigentliche Arbeitsaufgabe am wenigsten stört oder verzögert. Nach Schlussfolgerung der angemessenen Aktion wird der entsprechende Stellbefehl an die Aktoren ausgegeben. Dies bewirkt eine Änderung im ursprünglich geplanten Ablauf, der dann erneut durch Verarbeitung der Sensorsignale und erneute Entscheidungsfindung überprüft werden muss. Man sieht also, dass es sich auch bei dieser komplexeren Verhaltensweise letztendlich um eine Regelschleife handelt, bei der der 3. Punkt der oben genannten 5 Schritte für eine Regelung besonders umfangreich ausfällt.

Sind die Informationen, die dem System für die Steuerung und Regelung zur Verfügung stehen unvollständig oder nicht eindeutig (unscharf), so verwendet man heute auch häufig Methoden der **Fuzzy Logik**, mit denen man trotz der vorhandenen Unbestimmtheit zu brauchbarem Verhalten kommt.

7.1.4 Autonome intelligente Agenten

Eine andere Vorgehensweise, um bei einer Maschine intelligentes Verhalten zu erzeugen, ist das Zusammenspiel **intelligenter autonomer Agenten**, die alle unabhängig voneinander arbeiten und jeweils ein eigenes Ziel verfolgen. Dabei ist ein Agent ein Computerprogramm, das vom Entwickler ohne genaue Kenntnis der Einzelheiten eines anderen Agenten geschrieben werden kann. Alle Agenten greifen

auf Daten in einem festgelegten Speicherbereich zu, so dass sie untereinander Informationen austauschen können. Bei der Entwicklung eines Agenten muss dem Programmentwickler nur bekannt sein, wie der Agent mit den gemeinsamen Daten umgehen muss. Das Zusammenspiel vieler solcher unabhängiger Agenten kann dann **emergente**² Eigenschaften besitzen, die ein einzelner Agent nicht besitzt, so dass durch das Zusammenspiel vieler Agenten eine höhere Stufe der Intelligenz erreicht wird.

Ein Beispiel für den Einsatz solcher voneinander unabhängiger Agenten ist eine Werkzeugmaschine, bei der ein Agent implementiert ist, der versucht die Maschine immer mit der für eine Arbeitsaufgabe höchstmöglichen Schnittgeschwindigkeit arbeiten zu lassen und einem anderen Agenten, der versucht die Maschine in einem optimalen Arbeitszustand zu halten. Normalerweise überwacht der erste Agent die Schnittgeschwindigkeit und wählt die für die aktuelle Bearbeitung höchstmögliche, während der zweite Agent den Werkzeugverschleiß überwacht. Erreicht die Verschleißmarkenbreite am Werkzeug einen festgelegten Grenzwert könnte nun der zweite Agent entscheiden, dass die Schnittgeschwindigkeit gesenkt werden muss, wenn nicht das Werkzeug noch während des laufenden Zerspanungsprozesses brechen soll. An dieser Stelle muss dann ein übergeordnetes Entscheidungssystem eingreifen, das den bestehenden Konflikt zwischen den beiden Agenten entscheidet. In diesem Fall würde man den geringfügigen Produktivitätsverlust durch Senkung der Schnittgeschwindigkeit für den Rest des aktuellen Zerspanungsvorgangs über den möglichen Schaden durch einen Werkzeugbruch stellen.

7.1.5 Lernen und Mustererkennung

Häufig sind bestimmte Informationen bei der Entwicklung von Agenten nicht verfügbar. Entweder besitzt man nur unvollständige Informationen über die Arbeitsaufgabe oder das Umfeld in dem die Maschine arbeitet, oder die Einsatzgebiete und die speziellen Arbeitsaufgaben liegen noch nicht fest. Außerdem kann es sein, dass die Umgebung, die Aufgabe oder die Maschine sich mit der Zeit in nicht vorhergesehener Weise ändern. Ein typisches Beispiel hierfür wäre eine mit optischen Sensoren (z. B. Video-Kamera) ausgestattete Maschine, die wichtige Arbeitsinformationen von diesen Sensoren erhält, in deren Umfeld sich aber ständig die Beleuchtungsverhältnisse in nicht vorhersehbarer Weise ändern. Hier kann die Fähigkeit des Lernens neuer Tatbestände eine Anpassung von Agenten an die oben beschriebenen Situationen erreichen. Lernen von Maschinen erfolgt häufig durch Rückkopplung (reinforcement learning), wodurch bestimmte Merkmale einer Arbeitsaufgabe durch ständige Wiederholung des Lernprozesses für die Entscheidungsfindung der Agenten verstärkt werden.

Die einfachste Form des Erlernens findet bei der Positionsdateneingabe von Robotern statt. Da die exakten Positionen für die Bewegungsaufgabe innerhalb einer komplexen Arbeitszelle nicht exakt bekannt sind oder vermessen werden können, werden die Positionen **geteacht**. Dabei wird im Handbetrieb der Roboter in die erforderliche Position bewegt und die Positionsdaten werden dann durch Abspeichern

2 sich herausbildend: Das Zusammenspiel verschiedener Funktionalitäten erzeugt zusätzliche, vorher nicht vorhandene Funktionalität; siehe auch Eigenschaften von Systemen

der Messwerte aller Lagemesssysteme "eingelernt". Dieser Vorgang muss eventuell durch erneutes Teachen verbessert werden, wenn sich im anschließenden automatischen Programmablauf herausstellt, dass die Position noch nicht der Arbeitsaufgabe genügt. Dieser Lernprozess erfolgt jedoch nicht automatisch sondern bedienergeführt.

Ein vollautomatisches Lernen kann man unter Verwendung **neuronaler Netzwerke** ermöglichen. Ein **künstliches Neuron** ist eine elektronische Schaltung, die die Eigenschaften von Gehirnzellen (Neuronen) nachbildet. Ein menschliches Gehirn hat etwa 10^{10} Neuronen, die jeweils mit einer Vielzahl benachbarter Neuronen verbunden sind. Anders als herkömmliche Computer, die die Befehle im allgemeinen seriell abarbeiten, verarbeiten biologische neuronale Netzwerke Informationen parallel, wobei die Signale elektrischer Natur sind und durch biochemische Vorgänge erzeugt werden. Auch die Speicherung von Informationen erfolgt anders als beim Computer. Ein Teil der Informationsspeicherung beruht auf der Änderung der Neuronenverschaltung untereinander bzw. durch Änderung der Intensität der Kopplung einzelner Neuronen. Die Speicherung erfolgt daher über das gesamte Neuronennetzwerk verteilt und nicht konzentriert auf bestimmte Speicherelemente. Die Eigenschaften der Parallelverarbeitung und der dezentralen Speicherung führen zu einer viel höheren Toleranzschwelle gegen Störungen und Ausfälle von Teilen des Netzwerkes.

Ebenso sind neuronale Netzwerke in der Lage, ganze Informationsmuster wie Bilder oder Klänge zu lernen und diese wiederzuerkennen, wenn sie dem Netzwerk in unvollständiger Weise erneut angeboten werde. Daher können künstliche neuronale Netzwerke auch dazu benutzt werden, intelligente Maschinen mit den Eigenschaften der **Mustererkennung** (pattern recognition) auszustatten oder um intelligentes Verhalten im Laufe des Lernprozesses zu entwickeln. Dies kann die Aufgabenstellung, eine möglichst vollständige Wissensbasis für das Bewältigen der Arbeitsaufgabe einer intelligenten Maschine zu erstellen, deutlich minimieren. Die prinzipielle Arbeitsweise neuronaler Netze wird später in diesem Kapitel behandelt werden.

7.1.6 Architektur intelligenter Maschinen

Mechatronische Systeme, deren Struktur in Bild 7-1 dargestellt ist, sind intelligente Maschinen. Wie schon mehrfach angemerkt, entsteht die Intelligenz solcher Systeme durch die Verbindung unterschiedlicher Teilsysteme, von denen kein einzelnes bereits die Eigenschaften des Gesamtsystems besitzt. Die Verbindung der Systeme untereinander entsteht durch Interfaces, die sowohl elektrisch als auch mechanisch sein können. Die Struktur und Topologie³ eines solchen Systems bezeichnet man auch als **Architektur**. Der Systemarchitektur widmet man heute große Aufmerksamkeit im Bereich der Kommunikation zwischen Teilsystemen. Sie besteht aus einer Anzahl von Regeln, die die Art des "Interfacing" zwischen dem System, seinen Teilsystemen und der Umwelt festlegen. Der Begriff stammt ja aus dem Bauwesen und die Architektur eines Gebäudes legt fest, wie Fenster, Treppenhäuser

3 Mathematisches Teilgebiet: Lehre von den Eigenschaften geometrischer Figuren im Raum; hier speziell, die Form der Verbindung von Kommunikationspartnern untereinander

und Dach sich in das Gebäude einfügen und miteinander verbunden sind und wie sich das Gebäude in das gesamte bauliche Umfeld einfügt. Ähnlich legt die Architektur eines Computers fest, wie die Hauptkomponenten der Hardware miteinander verbunden sind und wie die Hardware mit dem Ein- und Ausgabesystem verbunden ist.

Die Intelligenz einer Maschine hängt letztendlich auch davon ab, mit welcher Architektur ihre Elemente untereinander verbunden sind. Dabei werden drei Haupttypen von Architekturen unterschieden:

- Hierarchien
- Netzwerke
- Schichtarchitektur

In **modularen** Architekturen sind die Baugruppen, die das System bilden, so miteinander verbunden, dass möglichst jede Baugruppe weitestgehend autark arbeiten kann. Intelligente Maschinen sind oft modular aufgebaut, was den Vorteil besitzt, dass fehlerhafte Baugruppe schnell ausgetauscht und das System neu konfiguriert werden kann. Obwohl daher Modularität vom Standpunkt der Instandhaltung und für eventuelle Modifikationen ideal ist, enthält ein so aufgebautes System viele redundante Teile. Dies verursacht als Nachteil ein Mehr an Gewicht, Größe und Investitionskosten.

Im Gegensatz dazu steht die **integrale** Architektur deren Ziel es ist, die Bestandteile eines Systems so eng zusammenzubringen, dass Interface zur Verbindung der Baugruppe sich beinahe erübrigen. Man kann diese Bauweise beispielsweise dann günstig einsetzen, wenn die Notwendigkeit besteht zu verhindern, dass Baugruppen voneinander getrennt werden oder wenn das Hauptziel ist, das Bauvolumen des Systems so klein wie möglich zu halten.

7.1.6.1 Hierarchien

Bei einer **Hierarchie** handelt es sich um eine Architektur, die aus miteinander verbundenen Systemen besteht, die nicht alle gleichberechtigt sind, sondern eine definierte Rangfolge besitzen.

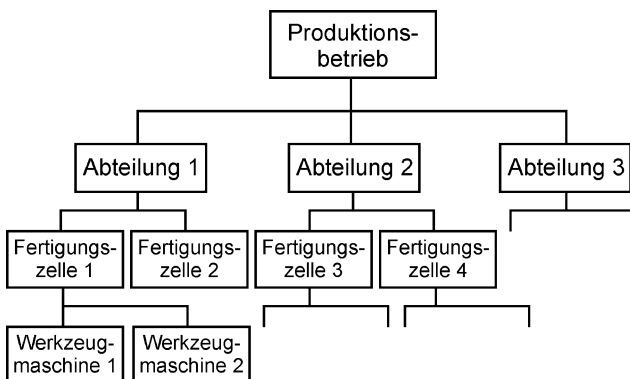


Bild 7-9: Hierarchische Steuerungsstruktur eines Produktionsbetriebs

Innerhalb einer solchen Hierarchie werden die im Rang höher stehenden Systeme auch als "Eltern" (parents) bezeichnet, die im Rang niedriger stehenden als "Kinder" (children). Jedes Eltern-teil kann mehrere Kinder haben, von Stufe zu Stufe entsteht so eine vielschichtige Hierarchie, in der jedes Kind wiederum Eltern-teil von weiteren Kindern sein kann. Eine

Hierarchie hat immer ein Element an der Spitze, das nur Elternteil und nicht gleichzeitig Kind ist. Am Ende der Hierarchie sind immer Elemente, die nur Kind sind aber nicht Elternteil. Als Beispiel für ein hierarchisches System ist in Bild 7-9 die Struktur eines Produktionsbetriebes dargestellt.

Der Grund für die hierarchische Strukturierung eines Systems liegt meist in der Komplexität, die eine Kommunikation aller mit allen anderen Partnern so unüberschaubar macht, dass das System nicht mehr handhabbar wäre. Dies trifft in der Regel auf große gesellschaftliche Strukturen wie Kommunen, Länder, Staaten oder Organisationen wie die EU oder UNO zu. In dem dargestellten Beispiel ist eine Abteilung des Produktionsbetriebs in Fertigungszellen aufgegliedert, die wiederum verschiedene Werkzeugmaschinen enthalten. Der Vorteil einer solchen Struktur ist nun, dass das ordnungsgemäße Arbeiten einer Werkzeugmaschine nicht mehr vom Zustand der übrigen Werkzeugmaschinen abhängt, da sie nicht direkt miteinander kommunizieren. Die Komplexität des Gesamtbetriebes ist an dieser Stelle auf ein überschaubares Maß reduziert. Je höher jedoch ein hierarchisches Element in der Struktur liegt, umso kritischer macht sich ein Ausfall auf das Gesamtsystem bemerkbar. Fällt beispielsweise der Rechner mit dem Produktions-Planungs- und Steuerungssystem (PPS) auf der höchsten Ebene des Produktionsbetriebes aus, so wird das gesamte System lahmgelegt.

7.1.6.2 Netzwerke

Im Gegensatz zu Hierarchien sind in **Netzwerken** die Teilsysteme nicht in Rangfolgen gegliedert, sondern gleichberechtigt. Alle Teilsysteme können jeweils mit jedem anderen Partner verbunden sein. Die Möglichkeit, dass jeder Knoten (Netzteilnehmer) des Netzes mit jedem anderen verbunden ist (Bild 7-10 a) ist die einfachste und naheliegendste Struktur, sie ist aber eher selten, da die bidirektionale Verbindung jedes Netzknotens mit jedem anderen sehr aufwendig ist. Die Anzahl der Verbindungen V steigt mit zunehmender Anzahl Netzknoten N entsprechend folgender Beziehung:

$$V = 0,5(N^2 - N) \quad .$$

Daher ist diese Struktur nur für kleine Netze geeignet.

In der Kommunikationstechnik werden Netzwerke in **Sternform** (Bild 7-10 b)) oder in **Ringform** (Bild 7-10 c)) häufig verwendet [7.14]. Man findet diese Strukturen vor allen in sogenannten **LAN's** (Local Area Network), mit denen verschiedene Rechner als Netzteilnehmer untereinander verbunden werden. In der Regel ist einer der Teilnehmer eine Zentrale, beim Sternnetz der Rechner, der mit allen anderen Teil-

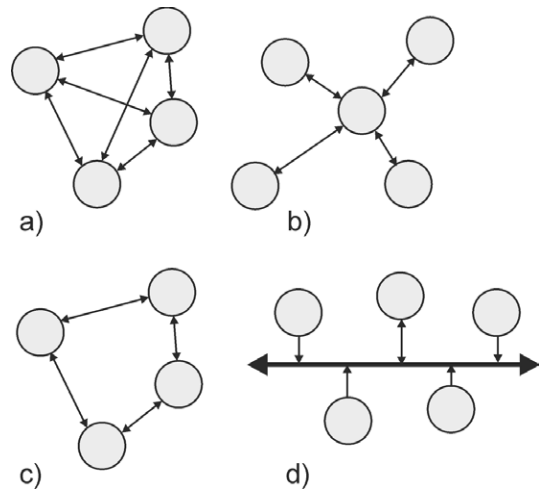


Bild 7-10: Topologien von Netzwerken

nehmern verbunden ist, beim Ringnetz ein beliebiger Rechner. Folgende weitere Eigenschaften besitzen die beiden Netztopologien Stern und Ring:

- Stern
 - Bei Verbindung zweier Kommunikationspartner ist immer die Zentrale beteiligt.
 - Es kann immer nur die Verbindung zwischen zwei Teilnehmern hergestellt werden, die Datenübertragung der Verbindung kann aber sehr schnell sein.
 - Der Teilnehmer hat einen bidirektionalen Anschluss, die Zentrale N solcher Anschlüsse, die Gesamtlänge der Verbindungen ist relativ hoch.
 - Der Ausfall einzelner Knoten ist für das Netz rückwirkungsfrei.
 - Veränderungen der Netztopologie sind ohne Unterbrechung des Betriebes möglich.
- Ring
 - Es wird eine gerichtete (unidirektionale) Kette aufgebaut, deren Enden geschlossen sind.
 - Die Knoten sind aktive Stationen, die die Informationen empfangen und regeneriert weiterleiten, wodurch hohe Teilnehmerzahlen und große geografische Ausdehnungen erreicht werden können, bei allerdings relativ langsamem Datenverkehr.
 - Störungen eines Teilnehmers oder des Busses wirken sich auf alle Teilnehmer aus.
 - Die Information muss die Adresse des Teilnehmers enthalten, damit dieser die für ihn bestimmte Nachricht aufnehmen kann.

Die in Bild 7-10 d) dargestellte Struktur ist der sehr häufig verwendete **Bus**. An einem Bus sind alle Teilnehmer parallel angeschlossen. Folgende Eigenschaften bestehen:

- Bus
 - Die Signalausbreitung erfolgt im Gegensatz zum Ring in beiden Richtungen des Busses. Daher müssen die Zugriffskanäle auf den Bus bidirektional ausgeführt werden. Die Übertragungsgeschwindigkeit ist mittelschnell.
 - Es ist immer ein Teilnehmer als Sender aktiv. Eine Bussteuerung sichert, dass es keine Zugriffskonflikte auf dem Bus gibt. Die Nachricht wird gleichzeitig an alle anderen Teilnehmer gesendet, wodurch sich eine Wegsuche (Routing) erübrigt. Daher muss in der Nachricht mitgeteilt werden, wer sie empfangen soll.
 - Busnetze sind sehr flexibel, da neue Teilnehmer einfach durch Parallelschalten eingefügt werden können.
 - Ist ein Teilnehmer gestört, so kann der Bus unbeeinflusst bleiben. Liegt jedoch

eine Busstörung vor, so sind alle Teilnehmer betroffen.

Ein vollständiges **Prozessleitsystem**, mit seinen verschiedenen Steuerungsebenen und den verschiedenen Steuerungssystemen ist in Bild 7-11 dargestellt. Das Prozessleitsystem ist hierarchisch gegliedert. Am Fuß der Pyramide befindet sich der Prozess, der zu beobachten, zu steuern und zu regeln ist.

Die Verbindung mit dem Prozessleitsystem erfolgt über Sensoren (S) und Ak-

toren (A), die es übernehmen, die physikalischen Größen in elektrische Signale umzuwandeln, die entweder direkt oder über ein Kommunikationssystem an die Automatisierungsgeräte (Steuerungen, Regeleinrichtungen, Computer) übertragen werden. Die Automatisierungsgeräte übernehmen die Verarbeitung der Informationen und beeinflussen über die Aktoren den Prozess.

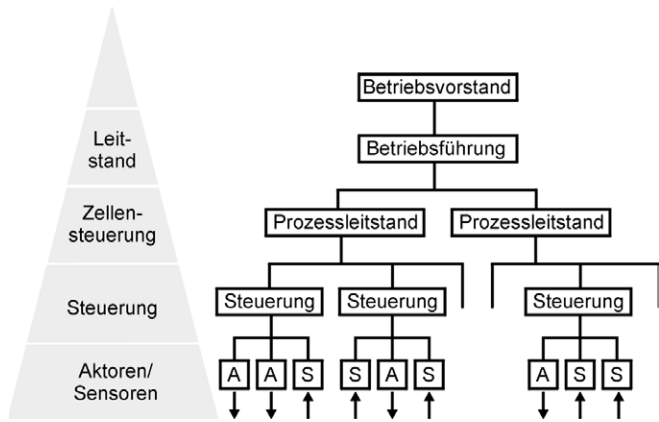


Bild 7-11: Steuerungsebenen mit Prozessleitsystem

Entsprechend der betrieblichen Hierarchie sind verschiedene Steuerungen in Bearbeitungszellen zusammengefasst. Diese Steuerungen sind zu koordinieren.

Schließlich ist ein Überblick über die Prozessabläufe erforderlich, der durch die Einrichtung größerer zusammenhängender Betriebsabschnitte mit Prozessleitständen realisiert wird.

Auf den verschiedenen Ebenen der Steuerungshierarchie werden verschiedenartige Netztopologien für die Prozesssteuerung benutzt. Bild 7-12 zeigt erneut die Kommunikationspyramide aus Bild 7-11. Im Bereich der verschiedenen Steuerungsebenen werden LAN, Prozessbusse und Sensor-Aktorbusse benutzt. Die letzten beiden Bussysteme werden auch als **Feldbusse** bezeichnet. Im Bild sind einige der gebräuchlichen Feldbussysteme genannt. Im Bereich der Feldbusse werden überwiegend **zeitkritische** Nachrichten zur zyklischen Abfrage oder zur Aktualisierung des Prozesses ausgetauscht.

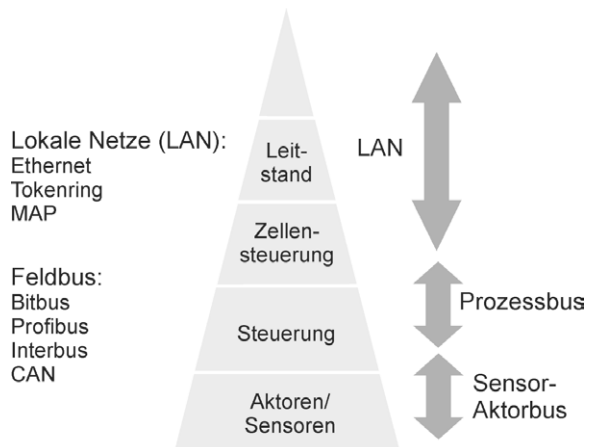


Bild 7-12: Verschiedene Netz- und Bussysteme innerhalb der Kommunikationspyramide

Auf den oberen Ebenen der Kommunikationspyramide werden in der Regel LAN

benutzt, von denen auch einige Netzsysteme im Bild 7-12 genannt werden. Die Nachrichten, die hier übertragen werden müssen, sind meist zeitunkritisch, dafür müssen in der Regel größere Datenpakete übertragen werden, wie NC- Programme, Prozessparameter und anderes.

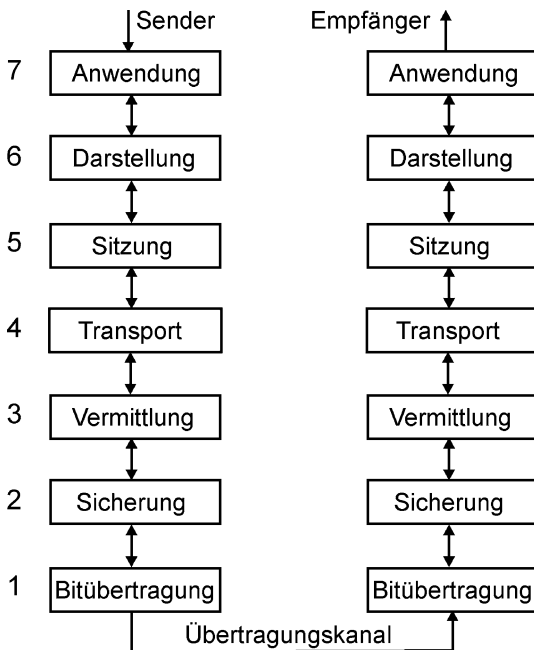


Bild 7-13: OSI 7 Schichten-Referenzmodell mit Übertragungsweg vom Sender zum Empfänger

Im Zusammenhang mit mechatronischen Systemen sind besonders Feldbussysteme für den Datenaustausch der Sensor-/Aktor-Ebene mit der übergeordneten Steuerungsebene von Bedeutung. Im Automobilbau, derjenigen Branche die sich heute am meisten mit mechatronischen Systemen befasst, kommt der Verbindung der verschiedensten Systeme mit den Sensoren und Aktoren über Bussysteme große Bedeutung zu.

Um direkte Kopplung unterschiedlicher Kommunikationssysteme mit unterschiedlichen Funktionalitäten gewährleisten zu können, ist durch die **ISO** (International Standards Organization) ein Schichtenmodell offener Systeme (OSI: Open systems interconnection) entwickelt und genormt worden. Die Offenlegung aller Einzelheiten der Kommunikation stellt sicher, dass beliebige Hersteller Hard- und Software für solche Kommunika-

tionssysteme entwickeln können. Das 7-Schichten **OSI-Referenzmodell** ist in Bild 7-13 dargestellt. Das OSI-Modell beinhaltet folgende Festlegungen:

- Der Kommunikationsweg vom Anwender bis zum Übertragungskanal durchläuft die bezeichneten 7 Schichten.
- Jede Schicht übernimmt eine abgegrenzte Kommunikationsfunktion.
- Zwischen den benachbarten Schichten sind Schnittstellen festgelegt. Die Schnittstelle definiert, welche einfachen Operationen und Dienste existieren.
- Jede Schicht bietet ihrer übergeordneten Schicht Dienste an, auf die diese zugreifen kann. Die übergeordnete Schicht sendet dazu ein Paket aus Dienstanforderungen und Daten.
- Ein Dienst sagt aus, welche Operationen eine Schicht der übergeordneten zur Verfügung stellt. Die Dienste werden nach einem Protokoll ausgeführt, das genau definiert, wie die Dienste ausgeführt werden.

Mit dieser Vorgehensweise kann die insgesamt komplizierte Kommunikationsaufgabe in kleine überschaubare Aufgaben zerlegt werden. Nicht jedes Kommunikations-

system muss zwingend alle 7 Schichten enthalten. Zur Übertragung einer Nachricht vom Sender zum Empfänger durchläuft diese zuerst die 7 Schichten von der "Anwendung" bis zur "Bitübertragung" (Bild 7-13), wird dann über den physikalischen Übertragungskanal geleitet und durchläuft beim Empfänger erneut die 7 Schichten in umgekehrter Reihenfolge. Die 7 Schichten haben dabei folgende Aufgaben mit zugehörigen Schwerpunkten:

- 1) Bitübertragung: Richtige Übertragung der Informationsbits über den Kanal.
 - Welches Übertragungsmedium?
 - Welcher Signalpegel für 0 und 1?
 - Welche Übertragungsgeschwindigkeit?
 - Wird in beiden Richtungen gleichzeitig übertragen?
 - Wie wird die Verbindung hergestellt und wie abgebrochen?
 - Wie viele Anschlüsse hat die Verbindung und wofür werden sie verwendet?
- 2) Sicherung: Datensicherung durch Coderedundanz und Quittungsmechanismen bei erkannten Fehlern.
 - Erzeugung von Datenrahmen (Frame).
 - Fehlererkennende und fehlerkorrigierende Codierung.
 - Wiederholung von Datenrahmen bei Fehlern.
 - Quittierung bei richtiger Übertragung.
- 3) Vermittlung:
 - Wegesuche.
 - Adressierung.
 - Koordinierung der zu übertragenden Datenpakete.
- 4) Transport: Auswahl der Wege vom Ursprungs- zum Bestimmungsort.
 - Zerlegung der Daten in kleinere Übertragungseinheiten.
 - Übertragung großer Datenmengen über mehrere Netzverbindungen.
 - Zusammenlegung mehrerer Transportverbindungen auf einer Netzverbindung.
- 5) Sitzung: Stellt gehobene Dienste für die Datenkommunikation bereit.
 - Verschaffung des Zugangs zu einem anderen Rechnersystem.
 - Steuerung der Richtung beider Kommunikationsteilnehmer.
 - Einbau von Synchronisations- und Wiederanlaufpunkten bei langen Übertragungen.
- 6) Darstellung: Umwandlung der Datenstruktur.
 - Konvertierung von Zeichensätzen und unterschiedlichen Zahlendarstellungen.

- Konvertierung von Steuerzeichenfolgen für Terminals.
- Datenkompression zur Reduzierung der zu übertragenden Datenmenge.
- Kryptographie für Vertraulichkeit und Authentizität.
- 7) Anwendung: Bereitstellung von Diensten für Endanwender.
 - Datenübertragung und Verzeichnisverwaltung (Löschen, Umbenennen, usw.).
 - Nachrichtenübertragungsdienste (z. B. Elektronische Post)
 - Auftragsüberwachung und -verwaltung.

Für Feldbussysteme sind vor allem die Schichten 1, 2, 3 und 7 wichtig. Die Schichten 4, 5 und 6 werden nicht benutzt. Im folgenden werden einige Systeme kurz erläutert.

Der **Bitbus** wurde 1983 durch die amerikanische Firma Intel entwickelt und in der IEEE 1118 genormt. Das Einsatzgebiet dieses Feldbussystems liegt in der Verbindung intelligenter Einheiten, wie sie in der Zellebene eines Prozessleitsystems vorhanden sind, d. h. Speicherprogrammierbare Steuerungen (SPS), Industrie-PC, intelligente Terminals, CNC-Maschinen. Der direkte Anschluss von nichtintelligenten Sensoren und Aktoren an den Feldbus ist nur über einen sogenannten Gateway möglich.

Grundlage für den **Profibus** ist eine herstellerunabhängige Festlegung, die in der seit 1990 bestehenden DIN 19245 genormt ist. Etwa 140 Firmen sind in einer Nutzerorganisation zusammengefasst. Der Profibus ist ein bitserialer Bus und nutzt nur die Schichten 1, 2 und 7 des OSI-Referenzmodells. Da er in unterschiedlichsten Steuerungsebenen eingesetzt werden kann, stellt der Profibus ein durchgängiges Konzept dar. Dies wird durch das Buszugriffsverfahren **Token-Passing**, einem so genannten Multi-Master-Konzept ermöglicht. Immer nur einer der Netzteilnehmer hat dabei das Zugriffsrecht als **Master** auf den Bus, während alle anderen als **Slaves** fungieren. Hat der aktuelle Master seine Kommunikation beendet, so gibt er das Zugriffsrecht (das Token⁴) an den nächsten Master weiter. Diese Weitergabe erfolgt ringförmig, während die Datenübertragung über einen Linienbus erfolgt. Die Datenübertragungsgeschwindigkeit zwischen Steuerungs- und Sensor-/Aktor-Ebene ist eher durchschnittlich.

Das ursprünglich von der Firma Bosch (1991) für Anwendungen im Automobil entwickelte serielle **CAN** (Controller area network) Feldbussystem wird zunehmend auch in der Industrieautomatisierung verwendet. Es ist international genormt in ISO/DIS 11898 und 11519-1. Es eignet sich vor allem für die Vernetzung intelligenter Ein-/Ausgabe-Einheiten im Automobilbau, Landwirtschaftsmaschinen, schiffstechnischen Anlagen, medizinischem Gerätebau sowie Textilmaschinen- und Sondermaschinenbau. Er nutzt wie der Profibus nur die Schichten 1,2 und 7 des OSI-Referenzmodells und verwendet ebenfalls ein Multi-Master-Konzept für den Buszugriff, gleichberechtigte Komponenten sind über einen Linienbus miteinander ver-

4 Token-Ring-Network (IBM): Lokales, ringförmiges Netzwerk, bei dem eine umlaufende (symbolische) Marke (Token) anzeigt, welcher angeschlossene Kommunikationsteilnehmer an der Reihe ist (Token-Passing).

bunden. Der Buszugriff erfolgt dadurch, dass die Teilnehmer am Bus abfragen, ob dieser frei ist. War der Bus frei und greifen zwei Teilnehmer gleichzeitig zu, dann setzt sich der Teilnehmer mit der höheren Priorität durch. Es existieren kostengünstige CAN-Protokollchips, mit denen angepasste Baugruppen entwickelt werden können, z. B. für PKW-Elektronik. Aufgrund seines ursprünglichen Einsatzes im Kraftfahrzeugbereich werden an dieses Feldbussystem besonders hohe Anforderungen an die Sicherheit der Datenübertragung gestellt. Deshalb verfügt das CAN-Protokoll über mehrere Maßnahmen zur Fehlererkennung.

Der **Interbus** wurde ab 1985 durch die Firma Phoenix entwickelt, um binäre und analoge Signale gleichzeitig über einen seriellen Feldbus mit der Steuerungsebene verbinden zu können. Der Interbus erlaubt daher sowohl den Anschluss einfacher Sensoren und Aktoren als auch den Anschluss intelligenter Sensor-/Aktor-Systeme. Er ermöglicht die für Steuerungs- und Regelungsaufgaben erforderlichen kurzen Zykluszeiten. Der Interbus ist in DIN 19258 genormt und wird von mehr als 400 Entwicklern und Anwendern in der Vereinigung des Interbus-Clubs vertreten.

Beim Interbus wird ein deterministisches Buszugriffsverfahren verwendet, so dass die Antwortzeiten für Steuerungen, die diesen Bus benutzen, immer berechenbar ist. Bei diesem Verfahren wird ein über die Busteilnehmer verteiltes Schieberegister verwendet, das gleichzeitig alle Busteilnehmer bedient. Jeder Busteilnehmer erhält Datenspeicherzellen, die über die Ringverbindung des seriellen Busses ein großes Schieberegister ergeben. In diesem aktiven Ring erfolgt eine Signalregenerierung in jedem Teilnehmer und die anschließende Weiterleitung zum nächsten Busteilnehmer. Dies ermöglicht auch große überbrückbare Übertragungswege.

7.1.6.3 Schichtarchitekturen

Solche Architekturen intelligenterer Maschinen bestehen aus einzelnen unabhängigen Baugruppen, die **Layer** genannt werden, die jeweils eigene Ein- und Ausgänge besitzen (Bild 7-14). Da das Verhalten, das durch die einzelnen Layer erzeugt wird, wegen der Unabhängigkeit voneinander im Konflikt zueinander stehen kann (s. Abschnitt 7.1.4), kann es notwendig sein, dass ein Entscheidungsmechanismus für die Entscheidungsfindung vorhanden ist, um ein fehlerfreies Gesamtverhalten des Systems zu gewährleisten.

Eine wichtige Eigenschaft einer solchen Schichtarchitektur ist, dass sie sehr gut dazu geeignet ist, die Entwicklung unterschiedlicher Baugruppen unabhängig voneinander zu betreiben. Da jeder Layer eigene unabhängige Ein- und Ausgänge besitzt, kann ein neuer Layer einem bestehenden System hinzugefügt werden, ohne die schon vorhandenen Layer stärker zu beeinflussen.

Welche der in den im Abschnitt 7.1.6 genannten Architekturen die optimale ist, lässt sich nicht losgelöst von der jeweiligen Aufgabenstellung beantworten. In der Regel wird bei größeren Systemen eine Mischung aus diesen Architekturen verwendet. Dabei ist zu beachten, dass diese Architekturen eher konzeptionelle Modelle als

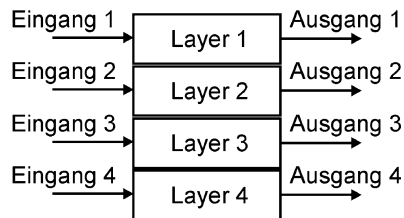


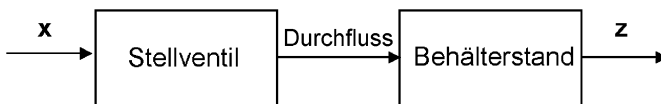
Bild 7-14: Vier Layer einer Schichtarchitektur

tatsächliche körperliche Realisierungen sind. Hat man beispielsweise einen Roboter, so kann es nützlich sein, zwischen den Aufgabenstellungen der Verarbeitung von Daten taktiler Sensoren und der Bewegungssteuerung des Robotergreifers zu unterscheiden. Dies bedeutet jedoch nicht, dass beide Programme auf unterschiedlichen Rechnern laufen müssen. Umgekehrt ist es beispielsweise nützlich einen taktilen Sensor räumlich am Greifer anzuordnen, was nicht gleichbedeutend damit ist, dass beide Funktionen konzeptionell gleich sind.

7.2 Steuerungstechnik

Wie bereits oben gesagt wurde, dient eine Steuerung dazu, automatisierten Einrichtungen eine Zielgröße vorzugeben. Dabei unterscheidet sich die **Steuerung** von einer **Regelung** dadurch, dass sie innerhalb einer offenen Wirkkette eingesetzt wird und dass der aktuelle Zustand der Steuergröße nicht überwacht wird. Daher werden bei unerwünschten Abweichungen vom Sollzustand keine automatischen Änderungen vorgenommen. Es war bereits erwähnt worden, dass aus diesem Grund Steuerungen vor allem dort eingesetzt werden können, wo das Modell der zu steuernden Einrichtung vollständig bekannt ist und der Einfluss von Störgrößen weitgehend vernachlässigt werden kann.

Kontinuierliche Größen: $x = \text{Zu}, \dots, 1/4\text{-Auf}, \dots, 3/4\text{-Auf}, \dots, \text{Auf}$
 $z = \text{Leer}, \dots, 1/4\text{-Voll}, \dots, 3/4\text{-Voll}, \dots, \text{Voll}$



Binäre Größe: $x = \text{Zu}, \text{Auf}$
 $z = \text{Füllstand nicht erreicht}, \text{Füllstand erreicht}$

Bild 7-15: Kontinuierliche und binäre Größen bei einer Behältersteuerung

Prinzipiell kann die zu steuernde Größe natürlich einen kontinuierlichen Wertebereich besitzen. Dann wird aber die Einhaltung der oben genannten Bedingungen sehr schwierig. Denkt man beispielsweise an die Füllung eines Behälters durch ein Stellventil (Bild 7-15), dessen Öffnung kontinuierlich vergrößert werden

kann, so ist es kaum vorstellbar genau vorherzusagen, wann bei 1/3 Öffnung des Ventils oder bei 2/3 Öffnung des Ventils der Behälter zu 3/4 gefüllt ist. Soll dies automatisch geschehen, so wird man in der Regel so vorgehen, dass man das Ventil ganz öffnet und den Behälterstand "3/4-Voll" mit einem Grenzwertschalter überwacht. Dieser registriert den gewünschten Zustand und veranlasst dann das vollständige Schließen des Ventils.

Diese Vorgehensweise schränkt den Wertevorrat der **Stellgröße** x "Ventilöffnung" auf die Werte "Zu" und "Auf" ein. Auch die überwachte Größe z "Füllstand" hat eigentlich nicht mehr die möglichen Werte 1/4-, 2/4-, 3/4-Voll, sondern es interessiert nur noch die Größe "Grenzwertschalter-Ein" oder "Grenzwertschalter-Aus". Da jetzt die Stellgröße und die Zielgröße nur noch jeweils zwei genau festgelegte Zustände kennen, bezeichnet man solche Größen als **binäre** Größen. Nimmt man diese Einschränkung vor, so wird das Modell der Strecke und der Steuerungseinrichtung sehr einfach und Störungen werden in der Regel keinen Einfluss auf die Steuerung mehr haben. Öffnet beispielsweise das Ventil bei Auftreten des "Öffnen"-Signals

nicht vollständig, sondern nur zu 90%, so wird das Erreichen des Füllstandes "3/4-Voll" zwar zeitlich etwas später eintreten, aber der gewünschte Zustand "zu 3/4 gefüllt" wird genau erreicht. Wäre der Volumenstrom durch die Behälterzuleitung (kontinuierliche Größe) eine wichtige einzuhaltende Kenngröße, so wäre eine solche Steuerung natürlich wieder ungeeignet.

Es gibt in der Automatisierungstechnik eine Menge von Aufgaben, die durch eine Steuerung gelöst werden können. Da der Geräteaufwand verglichen mit einer Regelung hier geringer und die Funktionssicherheit höher ist, ist eine große Anzahl von automatisierten Geräten mit Steuerungen ausgerüstet. Besonders durch die stürmische Entwicklung des Digitalrechners in den vergangenen Jahrzehnten hat die Verarbeitung binärer Signale stark zugenommen. Gerätetechnisch arbeiten auch Regelungsgeräte für kontinuierliche Größen heute meist mit Digitalrechnern in der Form von Mikrorechnern, auf denen die kontinuierlichen Wirkabläufe durch digital bearbeitete Algorithmen behandelt und realisiert werden.

Daneben haben sich Methoden für einen weiteren Bereich von Steuerungen und Regelungen entwickelt, die man einsetzt, wenn das Modell der Strecke völlig fehlt oder nur sehr unklar formuliert werden kann. Da hier weder die Eindeutigkeit von Binärgrößen weiterhilft, noch wegen des fehlenden Streckenmodells ein klarer funktionaler Zusammenhang zwischen Eingang und Ausgang vorhanden ist, werden Systemgrößen verwendet die unscharf formuliert sind und daher im engl. Sprachgebrauch als **fuzzy** bezeichnet werden.

Vielfach besteht heute das Erstellen einer Steuerung im Anfertigen eines Steuerprogramms für einen **Mikrorechner** oder eine **speicherprogrammierte Steuerung** (SPS). Die in Bild 7-16 dargestellte Einteilung von Steuerungen nach der Art, wie sie programmiert werden, zeigt neben den **speicherpro-**

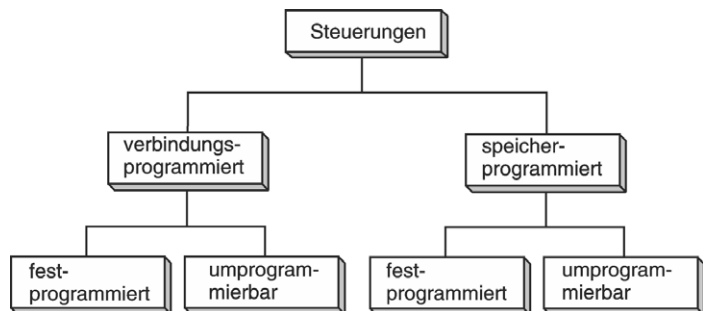


Bild 7-16: Einteilung von Steuerungen nach der Art der Programmierung

grammierten auch **verbindungsprogrammierte Steuerungen**. Typische Vertreter hiervon sind Relaissteuerungen und Steuerungen mit elektronischen Schaltelementen, die untereinander mit elektrischen Signalleitungen verbunden sind. Vor allem für solche Steuerungen sind die Methoden des digitalen Steuerungsentwurfs von Bedeutung. Diese werden hier nicht vertieft behandelt sondern auf entsprechende Standardwerke verwiesen [7.1], [7.11]. Verbindungsprogrammierte Steuerungen können außer elektrischer Energie als Signalmedium auch andere Informationsträger benutzen, wie Druckluft oder Drucköl. Diese sind aber nur bei gewissen Spezialanwendungen im Bereich der Stelltechnik von Bedeutung. Obwohl verbindungsprogrammierte Steuerungen immer mehr an Bedeutung verlieren, muss man zum Verständnis der Steuerungstechnik auch die elementaren Regeln und Gesetzmäßigkeiten der Steuerungstechnik beherrschen, zumal die Programmierung elektro-

nischer Steuerungen meist als Simulation einer Verbindungsprogrammierung abläuft (Kontaktplan, Funktionsplan). Informationen über die Gerätetechnik finden sich in [7.15].

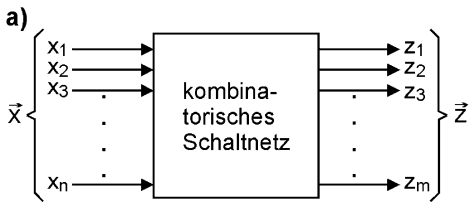
Eine weitere wichtige Unterscheidung bei binären Steuerungen, die sich nicht auf die Gerätetechnik, sondern auf das mathematische Modell des Systems bezieht, ist die Abhängigkeit der Ausgangsgrößen von den Eingangsgrößen. Da bei Steuerungen eine große Anzahl solcher Ein- und Ausgangsvariablen auftreten können, ist es häufig sinnvoll, diese als Vektor darzustellen:

$$\vec{x} = (x_1, x_2, \dots, x_n) \quad \text{Vektor der Eingangsvariablen}$$

$$\vec{z} = (z_1, z_2, \dots, z_n) \quad \text{Vektor der Ausgangsvariablen}$$

Ist der Ausgangsvektor zu jedem Zeitpunkt eindeutig durch den Eingangsvektor bestimmt (Bild 7-17 a), gilt also stets

$$\vec{z} = f(\vec{x}) \quad ,$$



so spricht man von einer **kombinatorischen Steuerung** oder von einem **Schaltnetz**.

Wird ein Teil der Ausgangsgrößen, der Zustandvektor $\vec{y}$, zum Eingang zurückgeführt und als Eingangsgröße dem Schaltnetz erneut zugeführt (Bild 7-17 b), so spricht man von einer **sequentiellen Steuerung** oder einem **Schaltwerk**. Für eine solche sequentielle Steuerung gilt:

$$\vec{z} = f(\vec{x}, \vec{y}) \quad .$$

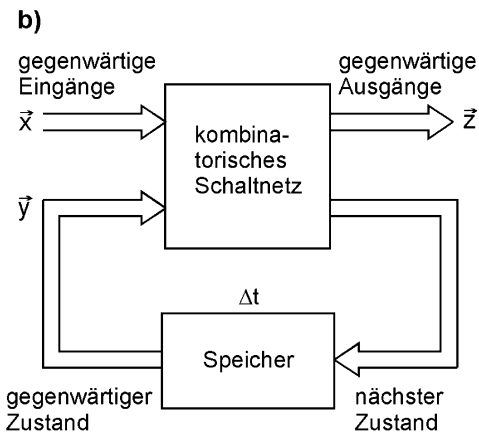


Bild 7-17: Verschiedene Arten von Steuerungen

- a) kombinatorische Steuerung
- b) sequentielle Steuerung

geschichte des aktuellen Zustands beinhalten. Der Ausgang der Speicher trägt immer die Information des gegenwärtigen Zustands, während an deren Eingängen die Information über den nächsten Zustand anliegt. Eingang und Ausgang der Speicherelemente unterscheiden sich daher um eine Zeitverzögerung Δt . Die Funktions-

5 Dieses Automatenmodell stammt von G.H. Mealy [7.2], dass das von E.F. Moore [7.3] entwickelte Automatenmodell verallgemeinert. Jeder Moore-Automat kann in einen Mealy-Automat überführt werden.

weise des mathematischen Modells lässt sich fehlerfrei beschreiben, weil die dabei verwendeten idealen Schaltelemente völlig vorhersehbar arbeiten. Reale binäre Schaltelemente zeigen davon abweichendes, nicht ideales Schaltverhalten, was zu unvorhersehbaren Problemen (Hazards) bei der Realisierung des modellierten Automaten führen kann. Diese Probleme umgeht man dadurch, dass man Änderungen der Belegungen aller beteiligten Eingangs- und Zustandsvariablen nicht zu jedem beliebigen Zeitpunkt zulässt, sondern den Zeitpunkt für Änderungen durch ein Taktsignal synchronisiert. Man spricht dann von **synchronen Automaten**, ein Arbeitsprinzip, mit dem alle heutigen Digitalrechner arbeiten. Da diese Arbeitsweise die Arbeitsgeschwindigkeit des Automaten herabsetzt, werden **asynchrone Automaten** nur bei Realisierung mit ohnehin langsam arbeitenden Schaltelementen wie beispielsweise Relais verwendet.

7.2.1 Boole'sche Algebra

Die Ein- und Ausgangsvektoren eines kombinatorischen Schaltnetzes sind binäre Größen, also Größen, die nur zwei verschiedene Zustände (Ventil geschlossen, Ventil geöffnet) annehmen können. Man könnte außer binären Größen und kontinuierlichen Größen natürlich auch mehrwertige Größen verwenden, die beispielsweise drei verschiedene Zustände kennen. Man hat sich in der Vergangenheit aber bevorzugt mit zweiwertigen Größen beschäftigt, weil ihre mathematische Behandlung denkbar einfach ist, und weil eine technische Realisierung mit elektrischen Komponenten so naheliegend ist. Jeder elektrische Schalter kennt nur die beiden Zustände offen oder geschlossen und kann dadurch sehr einfach durch eine zweiwertige Variable modelliert werden. Dazu kommt, dass der mathematische Formalismus zur Behandlung zweiwertiger Variablen schon im 19. Jahrhundert durch G. Boole (engl. Mathematiker 1815-1864) entwickelt worden ist. Dieser Formalismus wird als **Boole'sche Algebra** bezeichnet. Zur Zeit der Entwicklung dieser Algebra gab es für sie keine technischen Anwendungen, sondern sie wurde vor allem im Bereich der Aussagenlogik verwendet. Der amerikanische Mathematiker C.E. Shannon übertrug in seiner "information theory" 1948 die Gesetzmäßigkeiten der Boole'schen Algebra auf die Anwendung in der elektrischen Schalttechnik, weshalb diese auch als **Schaltalgebra** bezeichnet wird.

Es hat sich in der Vergangenheit eingebürgert, die beiden möglichen Belegungen einer zweiwertigen Variablen (auch als **logische Variable** bezeichnet) mit den beiden mathematischen Ziffernsymbolen für die Dezimalzahlen 0 und 1 darzustellen. Bei der Anwendung auf elektrische Schaltsignale ist es natürlich sinnvoll, den Zustand "Schalter offen" oder "kein Stromfluss" mit dem Symbol "0" und den Zustand "Schalter geschlossen" oder "Stromdurchgang" mit dem Symbol "1" darzustellen. Auch die umgekehrte Darstellung wäre denkbar, aber sicher nicht nützlich für die Anschaulichkeit.

Das erste Axiom der Boole'schen Algebra definiert den Charakter der Zweiwertigkeit einer logischen Variablen x :

$$x = 0 \text{ , wenn } x \neq 1 \text{ ; } x = 1 \text{ , wenn } x \neq 0 \text{ .} \quad (7.1)$$

Das zweite Axiom beschreibt die Komplementbildung (Negation), das besagt, dass

der Zustand "0" das Gegenstück zum Zustand "1" ist. Die Komplementbildung wird normalerweise als Querstrich über dem binären Zahlenwert oder der binären Variablen dargestellt:

$$\overline{0} = 1, \quad \overline{1} = 0 \quad . \quad (7.2)$$

Um bei mehr als einer Variablen die möglichen Kombinationen der Belegungen des Eingangsvektors und die daraus resultierende Belegung des Ausgangsvektors darzustellen, benutzt man schon in der Aussagenlogik so genannte **Wahrheitstabilen**. Wie Bild 7-18 für zwei Variable x_1 und x_2 zeigt, werden alle denkbaren Kombinationen in der Wahrheitstabelle untereinander geschrieben. Die Anzahl der möglichen Kombinationen beträgt bei n Variablen 2^n , d. h. bei zwei Variablen gibt es vier Kombinationsmöglichkeiten.

x_1	x_2	z	x_1	x_2	z
0	0	0	0	0	0
0	1	1	0	1	0
1	0	1	1	0	0
1	1	1	1	1	1

ODER-Funktion

UND-Funktion

Bild 7-18: Wahrheitstabellen für die Oder- und für die UND-Funktion

Die Wahrheitstabelle definiert, welchen Wert die Ausgangsvariable z bei welcher Kombination der Belegungen der Eingangsvariablen besitzt. Aus der Aussagenlogik sind einige Grundfunktionen bekannt, mit deren Hilfe man zweiwertige Aussagen untereinander verknüpfen kann. Die wichtigsten sind die Funktionen **ODER (Disjunktion)** und **UND (Konjunktion)**.

Diese beiden Grundfunktionen sind im Bild 7-18 in Form von Wahrheitstabellen dargestellt. Die Namensgebung der Funktionen kann man leicht anhand der Wahrheitstabelle erläutern. Bei der Verknüpfung der Eingangsvariablen durch die

ODER-Funktion wird die Ausgangsvariable z dann 1, wenn entweder x_1 oder x_2 die Belegung 1 besitzen. Dies gilt auch, wenn beide Variable den Wert 1 haben. Bei der Verknüpfung der Eingangsvariablen durch die UND-Funktion wird die Ausgangsvariable z nur dann 1, wenn x_1 und x_2 die Belegung 1 besitzen. Als Verknüpfungszeichen sind für die beiden Grundfunktionen folgende Symbole üblich:

$$\vee \text{ oder } + : \text{Disjunktion} \quad \wedge \text{ oder } \cdot : \text{Konjunktion} \quad .$$

Bei der Konjunktion ist es wie in der Schreibweise der "normalen" Algebra oft üblich, das Verknüpfungszeichen fortzulassen. Bei Verwendung der Zeichen "+" und "." ist stets zu beachten, dass es sich nicht um die Operationszeichen der normalen Algebra handelt.

Wendet man die oben genannten Grundfunktionen auf die boole'schen Werte 0 und 1 an, so kann man weitere Axiome der Boole'schen Algebra formulieren:

$$0 + 0 = 0, \quad 1 + 1 = 1, \quad 0 + 1 = 1 + 0 = 1 : \text{Disjunktion} \quad (7.3a)$$

$$0 \cdot 0 = 0, \quad 1 \cdot 1 = 1, \quad 0 \cdot 1 = 1 \cdot 0 = 0 : \text{Konjunktion} \quad (7.3b)$$

Den gleichen Sachverhalt, der in den Wahrheitstabellen dargestellt ist, kann man mit Hilfe der Boole'schen Algebra als Funktion darstellen:

$$z = x_1 + x_2 \quad \text{ODER-Funktion} \quad z = x_1 \cdot x_2 \quad \text{UND-Funktion} \quad .$$

Weitere Gesetzmäßigkeiten und Regeln der Boole'schen Algebra, wie sie teilweise

auch aus der normalen Algebra bekannt sind, sind in der folgenden Auflistung zusammengestellt:

$$x+1=1, x+0=x, x \cdot 0=0, x \cdot 1=x \quad (7.4)$$

$$x+x=x, x \cdot x=x \quad (7.5)$$

$$\overline{\overline{x}} = x, \overline{\overline{\overline{x}}} = \overline{x} \quad (7.6)$$

$$x+y=y+x, x \cdot y=y \cdot x \quad \text{Kommutative Gesetze} \quad (7.7)$$

$$x+\bar{x}=1, x \cdot \bar{x}=0 \quad (7.8)$$

$$x+x \cdot y=x, x \cdot (x+y)=x, (x+\bar{y}) \cdot y=x \cdot y, x \cdot \bar{y}+y=x+y \quad (7.9)$$

$$x+y+z=(x+y)+z=x+(y+z) \quad \text{Assoziative} \quad (7.10a)$$

$$x \cdot y \cdot z=(x \cdot y) \cdot z=x \cdot (y \cdot z) \quad \text{Gesetze} \quad (7.10b)$$

$$x \cdot y+x \cdot z=x(y+z), (x+y)(x+z)=x+yz \quad \text{Distributive Gesetze} \quad (7.11)$$

$$\overline{(x+y+z)}=\bar{x} \cdot \bar{y} \cdot \bar{z}, \overline{(x \cdot y \cdot z)}=\bar{x}+\bar{y}+\bar{z} \quad (7.12)$$

Die beiden Gleichungen (7.12) sind auch als **De Morgan'sches Theorem** bekannt, mit dessen Hilfe sich invertierte Ausdrücke umwandeln lassen.

Am einfachsten kann man die Schaltfunktion, die das Übertragungsverhalten eines Schaltnetzes vom Eingang zum Ausgang beschreibt, durch eine Wahrheitstabelle bestimmen. Soll man beispielsweise das Übertragungsverhalten einer kombinatorischen Schaltung bestimmen, so müsste man an die Eingänge der Schaltung hintereinander alle möglichen Kombinationen der Eingangsvariablen anlegen (linke Seite der Wahrheitstabelle) und die zugehörigen logischen Werte der Ausgangsvariablen (rechte Seite der Wahrheitstabelle) ermitteln. In Bild 7-19 ist eine solche Wahrheitstabelle für drei Eingangsvariablen x_0, x_1, x_2 und eine Ausgangsvariable z dargestellt. Will man nun die Schaltfunktion ermitteln, so kann man wie folgt vorgehen. Jede Zeile der Wahrheitstabelle lässt sich durch einen konjunktiv verknüpften Term aller Eingangsvariablen beschreiben. Dieser so genannte **Minterm** nimmt genau dann den logischen Wert 1 an, wenn die für die zugehörige Zeile entsprechende Belegung vorliegt. Beispielsweise hat die siebte Zeile der Wahrheitstabelle in Bild 7-19 die Belegung des Eingangsvektors $\vec{x}=(0,1,1)$. Dieser Kombination entspricht der Minterm m_6 , weil sein logischer Wert 1 wird, wenn die Variablen entsprechend belegt sind. Setzt

Minterm	x_0	x_1	x_2	z
$m_0 = \bar{x}_0 \bar{x}_1 \bar{x}_2$	0	0	0	1
$m_1 = x_0 \bar{x}_1 \bar{x}_2$	1	0	0	1
$m_2 = \bar{x}_0 x_1 \bar{x}_2$	0	1	0	1
$m_3 = x_0 x_1 \bar{x}_2$	1	1	0	1
$m_4 = \bar{x}_0 \bar{x}_1 x_2$	0	0	1	0
$m_5 = x_0 \bar{x}_1 x_2$	1	0	1	0
$m_6 = \bar{x}_0 x_1 x_2$	0	1	1	1
$m_7 = x_0 x_1 x_2$	1	1	1	0

Bild 7-19: Wahrheitstabelle einer Schaltfunktion

man nämlich in den Minterm die Belegungen des Eingangsvektors ein, so erhält man den konjunktiv verknüpften Ausdruck

$$\bar{0} \cdot 1 \cdot 1 = 1 \cdot 1 \cdot 1 = 1 \quad .$$

In der Wahrheitstabelle gibt es mehrere Belegungen des Eingangsvektors, für die die Ausgangsvariable den Wert 1 annimmt. Da entweder für die Belegung des Eingangsvektors der ersten, oder der zweiten, oder der dritten, oder der vierten, oder der siebten Zeile der Wahrheitstabelle die Ausgangsgröße den Wert 1 annimmt, erhält man die Schaltfunktion als Disjunktion der entsprechenden Minterme. Diese Form der Schaltfunktion bezeichnet man als **disjunktive Normalform**. Für die im Beispiel dargestellte Funktion lautet diese Normalform:

$$z = \bar{x}_0 \bar{x}_1 \bar{x}_2 + x_0 \bar{x}_1 \bar{x}_2 + \bar{x}_0 x_1 \bar{x}_2 + x_0 x_1 \bar{x}_2 + \bar{x}_0 x_1 x_2 \quad (7.13)$$

Diese Normalform wird meist zur Ermittlung der Schaltfunktion herangezogen, weil häufig in kombinatorischen Schaltungen die Anzahl der Belegungen des Eingangsvektors, für die der Ausgang den Wert 1 annimmt, kleiner ist als die Hälfte der möglichen Kombinationen der Belegungen der Eingangsgrößen. In diesem Fall ist der Ausdruck für die Schaltfunktion, den man als disjunktive Normalform bekommt günstiger (minimaler) als der in Form der ebenfalls möglichen **konjunktiven Normalform**. Da im angeführten Beispiel die Anzahl der Einsen für z größer ist als die Anzahl der Nullen, müsste die konjunktive Normalform die günstigere sein. Zur Bildung dieser Normalform verwendet man die so genannten **Maxterme**. Die Maxterme sind eine disjunktive Verknüpfung aller Eingangsvariablen und sie ergeben den logischen Wert 0, wenn man die Belegung des Eingangsvektors der zugehörigen Zeilen, für die die Ausgangsgröße 0 wird, in sie einsetzt. Setzt man in den zur sechsten Zeile der Wahrheitstabelle aus Bild 7-19 gehörenden Maxterm $M_5 = \bar{x}_0 + x_1 + \bar{x}_2$ die zugehörige Belegung des Eingangsvektors $\bar{x} = (1,0,1)$ ein, so ergibt sich $\bar{1} + 0 + \bar{1} = 0 + 0 + 0 = 0$. Die der Gl. (7.13) entsprechende konjunktive Normalform ergibt sich dann als Konjunktion der Maxterme, für die der Wert der Ausgangsgröße 0 ist. Da dies im vorliegenden Beispiel nur drei Maxterme sind, ergibt sich folgende einfachere konjunktive Normalform der Schaltfunktion:

$$z = (x_0 + x_1 + \bar{x}_2)(\bar{x}_0 + x_1 + \bar{x}_2)(\bar{x}_0 + \bar{x}_1 + \bar{x}_2) \quad . \quad (7.14)$$

Mit Hilfe der Gesetzmäßigkeiten der Boole'schen Algebra kann man die Form der Gl. (7.14) nun noch weiter vereinfachen. Vor den einzelnen im folgenden dargestellten Schritten ist jeweils die Gleichung notiert, nach der die jeweilige Vereinfachung erfolgt:

Bildung der Konjunktion der ersten beiden Klammern nach Gl. (7.11) :

$$z = (x_0 \bar{x}_0 + x_0 x_1 + x_0 \bar{x}_2 + \bar{x}_0 x_1 + x_1 + x_1 \bar{x}_2 + \bar{x}_0 \bar{x}_2 + x_1 \bar{x}_2 + \bar{x}_2)(\bar{x}_0 + \bar{x}_1 + \bar{x}_2) \quad .$$

Ersetzen der ersten Konjunktion in der ersten Klammer durch 0 nach Gl. (7.8) und "Ausklammern" von x_1 und $\bar{x}_2$ innerhalb der ersten Klammer nach Gl. (7.11) :

$$z = (0 + x_1(x_0 + \bar{x}_0 + 1 + \bar{x}_2) + \bar{x}_2(x_0 + \bar{x}_0 + 1))(\bar{x}_0 + \bar{x}_1 + \bar{x}_2) \quad .$$

Ersetzen der inneren Klammern durch 1 und Weglassen von 0 nach Gl. (7.4) :

$$z = (x_1 \cdot 1 + \bar{x}_2 \cdot 1)(\bar{x}_0 + \bar{x}_1 + \bar{x}_2) = (x_1 + \bar{x}_2)(\bar{x}_0 + \bar{x}_1 + \bar{x}_2) \quad .$$

Erneutes Vorgehen wie im ersten Schritt nach Gl. (7.11) :

$$z = \bar{x}_0 x_1 + x_1 \bar{x}_1 + x_1 \bar{x}_2 + \bar{x}_0 \bar{x}_2 + \bar{x}_0 \bar{x}_2 + \bar{x}_1 \bar{x}_2 + \bar{x}_2 \quad .$$

Ersetzen der zweiten Konjunktion durch 0 nach Gl. (7.8) und "Ausklammern" von nach Gl. (7.11) :

$$z = \bar{x}_0 x_1 + 0 + \bar{x}_2 (x_1 + \bar{x}_0 + \bar{x}_1 + 1) \quad .$$

Ersetzen der Klammer durch 1 und Weglassen von 0 nach Gl. (7.4) :

$$z = \bar{x}_0 x_1 + \bar{x}_2 \cdot 1 = \bar{x}_0 x_1 + \bar{x}_2 \quad .$$

Dies ist nun endgültig die **Minimalform** der Schaltfunktion, die sich nicht (7.15) mehr weiter vereinfachen lässt.

7.2.1.1 Kombinatorische Steuerungen

Das logische Verhalten solcher elementarer Schaltfunktionen kann man mit technischen Mitteln nachbilden, oder umgekehrt technische Steuerungseinrichtungen kann man mit Hilfe solcher Schaltfunktionen mathematisch modellieren. Die einfachste Form einer elektrischen Realisierung von Schaltfunktionen ist die Verwendung elektromechanischer Schalter oder **Relais**. Ein solches Relais mit mehreren Schaltkontakten ist in Bild 7-20 a) dargestellt. Es besteht aus einem Elektromagneten, dessen Spule den Eingang des Bauelementes darstellt. Fließt Strom durch die Spule (Eingangsgröße nimmt den Wert 1 an), so wird ein beweglicher Anker betätigt, der Schaltkontakte je nach Bauart öffnen oder schließen kann. Wird der Stromfluss unterbrochen (Eingangsgröße nimmt den Wert 0 an), so wird der Anker, meist durch Federbelastung, in seine Ruhelage zurückgestellt, wodurch die mit dem Anker verbundenen Schaltkontakte wiederum betätigt werden. Es gibt drei verschiedene Arten

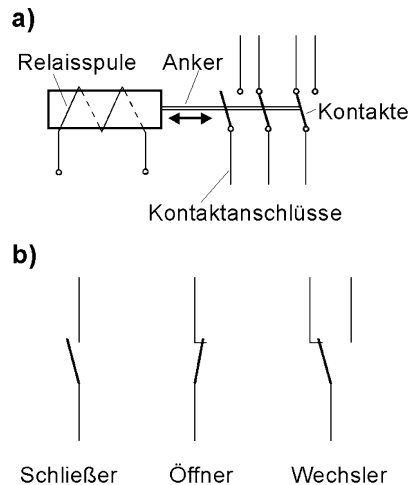


Bild 7-20: Elektromechanisches Relais mit Schaltkontakten
a) Schema b) Kontaktarten in Stromlaufplandarstellung

von Schaltkontakten (Bild 7-20 b), die in grafischen Darstellungen stark vereinfacht gezeichnet werden. Für die Definition ihrer Funktion ist die nicht betätigte Stellung der Kontakte entscheidend. Ist ein Kontakt im **nicht betätigten Zustand** des Relais geöffnet, so spricht man von einem **Schließkontakt**, **Arbeitskontakt** oder **Schließer** (wird bei Betätigung geschlossen). Ist ein Kontakt im **nicht betätigten Zustand** des Relais geschlossen, so spricht man von einem **Ruhekontakt** oder **Öffner** (wird bei Betätigung geöffnet). Enthält eine Kontaktanordnung zwei Kontakte, von denen je nach Erregungszustand des Relais jeweils einer geöffnet und einer geschlossen ist, so spricht man von einem **Wechselkontakt** oder **Wechsler**. Man kann nun mehrere solcher Kontakte untereinander verschalten um komplexere Schaltfunktio-

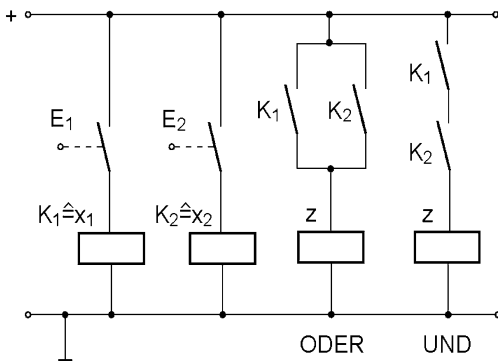


Bild 7-21: Stromlaufplan mit Realisierung der Schaltfunktionen ODER und UND

nen zu realisieren. Die Darstellung einer solchen Anordnung ist ein **Stromlaufplan**. In ihm werden nicht nur die für die Funktion der Schaltung wichtigen Schaltkontakte, sondern auch die zur Betätigung erforderlichen Relaispulen als rechteckige Kästen dargestellt. Um die Zuordnung zu verdeutlichen, gibt es eine entsprechende Abhängigkeitsnotation. So trägt die Spule, die einen bestimmten Kontakt betätigt, die gleiche Bezeichnung wie der zugehörige Kontakt. Im Bild 7-21 ist in einem Stromlaufplan dargestellt, wie man mit Relais die

beiden logischen Grundfunktionen UND und ODER der beiden Eingangsgrößen x_1 und x_2 realisiert. Betätigung der Endschalter E_1 oder E_2 erregt die zugehörigen Relaispulen K_1 oder K_2 , die den Eingangsgrößen x_1 und x_2 entsprechen. Diese Relaispulen betätigen die Schaltkontakte mit der gleichen Bezeichnung. Der Schaltfunktion ODER entspricht eine Parallelschaltung der Kontakte. Beim Schließen eines oder beider Kontakte wird die Relaispule z erregt, die die Ausgangsgröße der Schaltfunktion darstellt. Eine Reihenschaltung der Kontakte entspricht der Schaltfunktion UND, da nur dann die Relaispule z erregt wird, wenn beide Kontakte geschlossen sind.

Während der Stromlaufplan auch dazu dient, die körperliche Anordnung der Bauelemente darzustellen, gibt es auch noch die Darstellungsart, des **Kontaktplans**, in dem nur die für die logische Funktion entscheidenden Kontakte und ihre Verbindungen dargestellt sind. Diese Darstellungsart wird vor allem im Zusammenhang mit speicherprogrammierbaren Steuerungen benutzt.

Der Stromlaufplan und der Kontaktplan sind eng an die physikalische Realisierung einer Schaltfunktion, nämlich die mit elektromechanischen Schaltkontakten angelehnt. Moderne elektronische Bauelemente lassen auch die Realisierung komplexerer Schaltfunktionen direkt in einem Baustein zu. Die ersten dieser höher integrierten Bauelemente waren elektronische Schaltgatter, die in ihrem Inneren aus einer Vielzahl von Halbleiterbauelementen wie Dioden und Transistoren aufgebaut sind. Deren genaue Anordnung und Dimensionierung sind für die logische Funktion und damit für den Anwender aber unerheblich. Man betrachtet das Innenleben eines solchen Schaltgatters als "black box" und interessiert sich nur für das logische Schaltverhalten. Hieran orientiert sich eine andere grafische Darstellungsart, der sogenannte **Funktionsplan** oder **Logikplan**.

Bild 7-22 zeigt die Darstellung der Funktionsplansymbole für die beiden bereits bekannten Schaltfunktionen UND und ODER, sowie die beiden einfachsten Schaltfunktionen in einer Variablen, die **Identität** und die **Negation**. Die Identitätsfunktion hat keinen Einfluss auf den logischen Wert eines Signals und wird in der Schaltungstechnik beispielsweise zur Signalanpassung benutzt. Die Negation invertiert entsprechend der Gl. (7.2) jeweils den Wert des Eingangssignals. Diese Funktion wird häufig in technischen Schaltungen benötigt. Im Bild 7-22 sind auch die ande-

ren Darstellungsarten des gleichen Sachverhaltes in Form der Schaltfunktion und der Wahrheitstabelle dargestellt. Sie sind alle miteinander äquivalent und dienen nur zur klareren Darstellung bestimmter Aspekte verschiedenartiger Beschreibungen.

Außer den beiden bis jetzt vorgestellten Grundfunktionen von zwei Variablen gibt es natürlich noch weitere. Das Bild 7-23 zeigt in Form einer Wahrheitstabelle, dass 16 verschiedene Belegungen der Ausgangsvariablen möglich sind. Die Anzahl der Kombinationen von n Eingangsvariablen ergibt sich, wie schon erwähnt, als $2^n = 4$, während die Anzahl möglicher Ausgangsbelegungen $2^{2^n} = 16$ beträgt. Im Bild sind die beiden bereits bekannten Schaltfunktionen $z_2 \triangleq \text{UND}$ und $z_8 \triangleq \text{ODER}$ markiert. Ebenfalls markiert sind drei weitere wichtige Grundfunktionen $z_7 \triangleq \text{Antivalenz}$, $z_9 \triangleq \text{NOR}$, $z_{15} \triangleq \text{NAND}$. Die Funktion **Antivalenz** oder auch **Exklusiv-Oder** hat die Eigenschaft, dass die Ausgangsgröße nur dann den Wert 1 annimmt, wenn die beiden Eingangsvariablen unterschiedliche Belegung haben. Die Bezeichnung Exklusiv-Oder weist darauf hin, dass diese Funktion der ODER-Funktion ähnlich ist, lediglich für den Fall, dass beide Eingangsgrößen gleich sind, ergibt sich ein Unterschied. Die Funktion **NOR** ist die Invertierung der ODER-Funktion, woraus sich auch ihr Name herleitet, der eine Zusammenziehung der englischen Worte **NOT** und **OR** (Nicht Oder) ist. Die Funktion **NAND** ist die Invertierung der UND-Funktion, der Name bedeutet **NOT AND** (Nicht Und).

Grafisches Symbol

Wahrheitstabelle

Schaltfunktion

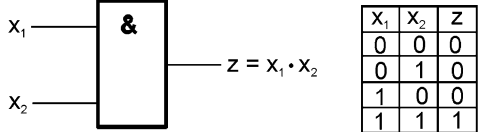
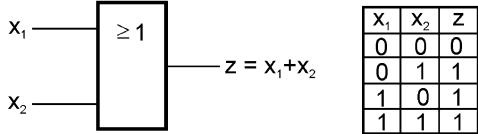
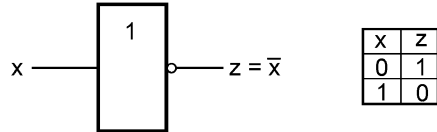
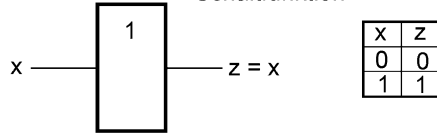


Bild 7-22: Grafisches Symbol, Schaltfunktion und Wahrheitstabelle einfacher Logikfunktionen

		UND				ODER											
x_1	x_2	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
		Antivalenz				NOR				NAND							

Bild 7-23: Mögliche Belegungen der Ausgangsvariablen für unterschiedliche Boole'sche Funktionen zweier Eingangsvariablen

7.2.1.2 Sequentielle Steuerungen

Wie schon zu Beginn des Abschnitts 7.2 (Bild 7-17) dargestellt, gibt es Steuerungen, deren Ausgänge sich zu jedem Zeitpunkt durch eine Schaltfunktion beschreiben lassen, die nur von den Eingangsvariablen abhängt. Solche Steuerungen heißen kombinatorisch. Bei sequentiellen Steuerungen (Schaltwerken) benötigt man zur Aufstellung der Schaltfunktion noch einen Vektor von Zustandsvariablen, die den vom Schaltwerk jeweils erreichten Zustand charakterisieren. Der Zustand wird in Speicherelementen, deren Ausgänge die Zustandsvariablen sind, bis zur nächsten Zustandsänderung gespeichert. Bei elektromechanischen Relais ist die ein-

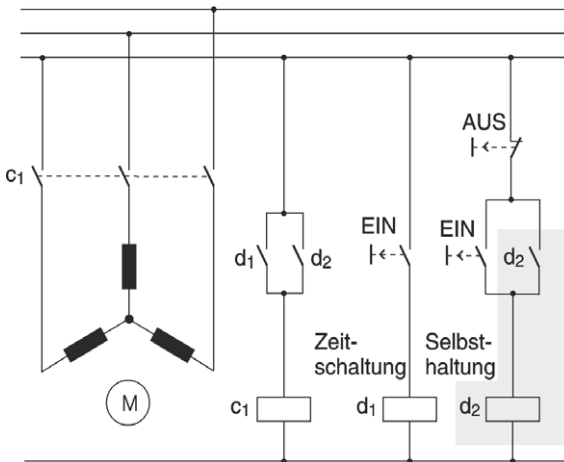


Bild 7-24: Motoransteuerung durch kombinatorische Zeitschaltung und sequentielle Selbsthaltung

fachste Form eines Speicherelementes die sogen. **Selbsthaltung**. Im Bild 7-24 ist eine solche Selbsthaltung für die steuerungstechnische Funktion "Motor EIN-/AUS" dargestellt. Im gleichen Bild wird eine solche Funktion für das Ein- und Ausschalten eines Motors, beispielsweise eines Vorschubmotors einer Werkzeugmaschine, ebenfalls durch eine rein kombinatorische Steuerung dargestellt. Verwendet man eine kombinatorische Schaltung in Form eines Einschalttasters mit Federrückstellung zum Betätigen des Motorschützes⁶, so läuft der Motor, solange der Bediener die Taste gedrückt hält. Nach dem Loslassen

des Tasters bleibt der Motor durch selbsttätiges Öffnen des Kontaktes und Abfallen des Motorschützes stehen. Aufgrund dieses Schaltverhaltens spricht man auch von **Zeitschaltung**, wie man sie für den Handbetrieb der Vorschubachse benötigen würde.

Fügt man parallel zum Einschaltkontakt einen Schließkontakt hinzu, der vom Motorschütz selber betätigt wird, so schließt dieser Parallelkontakt und hält über den parallelen Zweig den Stromfluss aufrecht, auch wenn der Einschaltkontakt durch Loslassen des Handtasters wieder geöffnet wird. Der Motor läuft dann ununterbrochen weiter, die Einschaltinformation wurde gespeichert, man sagt auch der Speicher wurde **gesetzt**. Da das Schütz sich über den zusätzlichen Parallelkontakt selbst in erregtem Zustand hält, spricht man von **Selbsthaltung**. Um nun den Motor wieder ausschalten zu können, wird ein weiterer Kontakt benötigt. Er befindet sich in Form eines Handtasters mit Öffnerfunktion in der Stromzuleitung des Selbsthaltekreises. Das Drücken dieses Tasters führt zur Unterbrechung des Stromflusses durch das Motorschütz, die gespeicherte Information wird gelöscht, der Speicher wurde **rückgesetzt**. Fügt man dieser Schaltung in Reihe noch einen weiteren durch einen

6 Elektromagnetischer Wechselstrom-Leistungsschalter.

Endschalter⁷ betätigten Öffnerkontakt hinzu, so kann man mit dieser Einrichtung den automatischen Positioniervorgang einer Vorschubachse steuern.

Eine solche Schaltung ist nicht mehr kombinatorisch sondern sequentiell und arbeitet nach dem Prinzip des Automatenmodells in Bild 7-17 b). Die Rückkopplung der Zustandsvariablen ist hier elektromechanisch; der vom Selbsthaltekontakt durch die Spule geleitete Strom baut ein Magnetfeld auf, das den Anker und damit den Kontakt in erregter Stellung hält. Will man dieses logische Verhalten mit den bekannten Darstellungsmethoden beschreiben, so kommt man beispielsweise auf die in Bild 7-25 gezeigte Wahrheitstabelle. In den beiden Fällen, in denen nur das **Setzsignal** (**EIN**) oder nur das **Rücksetzsignal** (**AUS**) den Wert 1 haben, ist das logische Verhalten eindeutig. Das Setzsignal speichert die Information ein und das Rücksetzsignal löscht sie wieder. Haben beide Signale den Wert 1, dann dominiert das Rücksetzsignal über das Setzsignal und der Speicher wird gelöscht. Man bezeichnet daher einen solchen **Speicher** als **dominierend rücksetzend**. Im vierten Fall, wenn beide Signale den Wert 0 besitzen, ist für die Ausgangsgröße d_2 der Wert d_2^* eingetragen. Dies bedeutet, dass in diesem Fall der vorher erreichte Zustand des Speichers erhalten bleibt, ein gesetzter Speicher ($d_2 = 1$) bleibt gesetzt und ein rückgesetzter Speicher ($d_2 = 0$) bleibt rückgesetzt.

EIN	AUS	d_2
0	0	d_2^*
0	1	0
1	0	1
1	1	0

Bild 7-25: Dominierend rücksetzende Selbsthalteschaltung

Bei elektronischen Realisierungen eines einfachen Speichers, verwendet man die in Bild 7-26 a) gezeigte symmetrische Struktur eines so genannten **Flip-Flops** oder **Triggers**, das man wie gezeigt aus zwei NOR-Gattern aufbauen kann. Das Flip-Flop hat zwei Eingänge die mit **R** (Rücksetzeingang) und **S** (Setzeingang) bezeichnet sind und zwei Ausgänge Q und $\bar{Q}$. Diese beiden

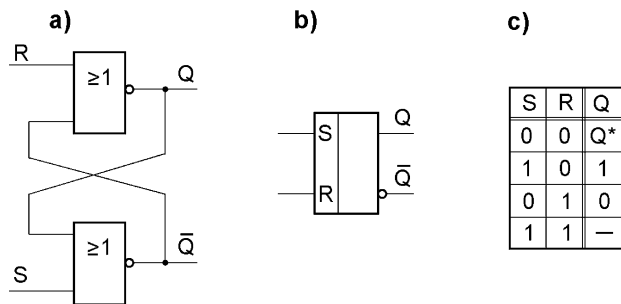
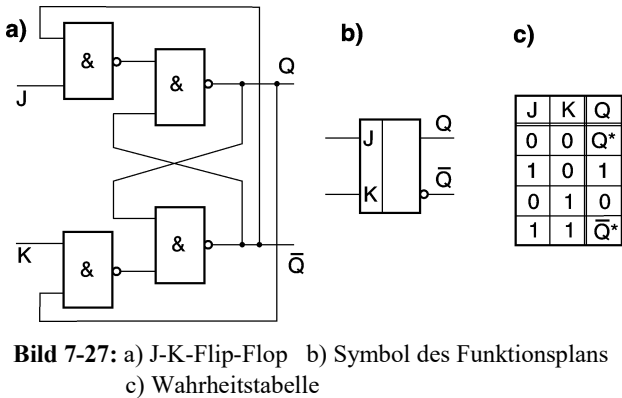


Bild 7-26: a) RS-Flip-Flop aus zwei NOR-Gattern, b) Symbol des Funktionsplans und c) Wahrheitstabelle

Ausgänge führen stets den entgegengesetzten logischen Wert. Für ein solches Flip-Flop gibt es ein vereinfachtes Symbol für die Verwendung im Funktionsplan (Bild 7-26 b). In Bild 7-26 c) ist die Wahrheitstabelle des RS-Flip-Flops dargestellt. Hier fällt gegenüber der Wahrheitstabelle des Selbsthaltekreises auf, dass der Fall $S = R = 1$ zu einem unbestimmten Ausgangswert führt. Dieser ist natürlich nicht wirklich unbestimmt. Er hängt davon ab, welche Bauelemente (Gatter) zur Realisierung des RS-Flip-Flops verwendet wurden. Benutzt man NOR-Gatter wie in Bild 7-26, so ist im genannten Fall $Q = 0$, $\bar{Q} = 0$, was nicht zur Definition der beiden Ausgänge passt. Man kann anstelle der NOR- auch NAND-Gatter verwenden, bei denen bei gleichzeitig anliegendem Setz- und Rücksetzsignal an den Eingängen

7 Elektromechanischer Schalter, dessen Position im Verbahrbereich der Vorschubachse manuell verstellt und selbsttätig durch ein bewegtes Maschinenteil gedrückt werden kann.

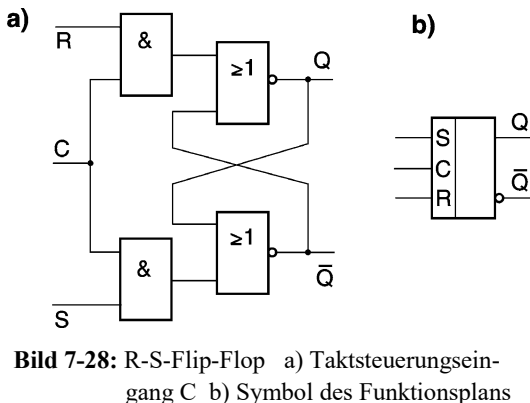
und $\bar{Q}$ die Ausgänge das beschriebene Verhalten zeigen. Der Eintrag Q^* für den $\bar{S}$ Ausgangswert Q bei einer Eingangsbelegung $R = S = 0$ bedeutet wieder, dass der alte Wert gespeichert bleibt.



Die Problematik des einfachen RS-Flip-Flops, dass es Eingangsbelegungen gibt, für die die Ausgangsgrößen unbestimmt sind, kann man mit dem sog. **JK-Flip-Flop** umgehen (Bild 7-27 a, b). Wie man der zugehörigen Wahrheitstabelle in Bild 7-27 c) entnehmen kann, führt hier die Belegung $J = K = 1$ ($J \triangleq S$, $K \triangleq R$) dazu, dass der alte gespeicherte Wert invertiert wird (Q^*), wo-

durch ein eindeutiger Wert vorliegt.

Alle Realisierungen von kombinatorischen und sequentiellen Schaltungen zeigen Abweichungen von dem Modellverhalten, das durch Schaltfunktionen beschrieben wird (s. Abschnitt 7.2.2). Schaltfunktionen enthalten **Hazards** (engl. hazard = Gefahr) die unter gewissen Umständen in realen Schaltungen zu Signalveränderungen führen und in sequentiellen Schaltungen durch Speicherung auch zu dauerhaftem Fehlverhalten führen können. Diese insbesondere bei gleichzeitigem Wechsel mehrerer Eingangsvariablen auftretende Problematik kann durch Synchronisierung der Signalwechsel in einzelnen Stufen der Schaltung vermieden werden. Diese Synchronisierung, mit der Digitalrechner grundsätzlich arbeiten, kann durch ein zusätzliches Taktsignal bewirkt werden, das allen Stufen zugeführt wird. Die **Syn-**



chronisation erfolgt bei Flip-Flops mit Hilfe eines zusätzlichen Takteingangs C (engl. clock), der Änderungen der Setz- und Rücksetzsignale nur zulässt, wenn das **Taktsignal** einen bestimmten Wert aufweist. Man spricht von Zustandssteuerung, wenn wie in Bild 7-28 a) gezeigt $C = 1$ sein muss, um den Wert am Dateneingang D zu übernehmen, oder von Flankensteuerung, wenn am Takteingang ein bestimmter Signalwechsel auftreten muss.

Ein weiterer Typ von Flip-Flop, das nur als taktgesteuertes Flip Flop sinnvoll ist, ist das **D-Flip-Flop** (Daten-Flip-Flop). Bei diesem in Bild 7-29 dargestellten Flip-Flop gibt es nur einen Informationseingang D , aus dem durch Invertierung die beiden stets mit entgegengesetztem logischen Wert erscheinenden R- und S-Signale erzeugt werden. Die kritischen Fälle bei anderen Flip-Flops, wenn Setz- und Rücksetzsignale gleich sind, treten daher bei diesem

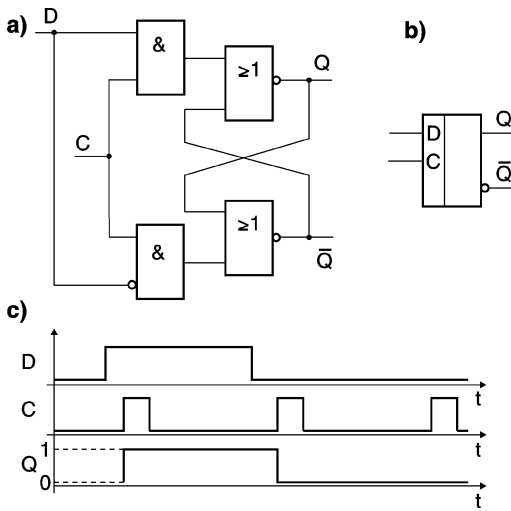


Bild 7-29: a) D-Flip-Flop b) Symbol des Funktionsplans c) Schaltfolgediagramm

Flip-Flop gar nicht auf. Wird das Takt-signal 1, so wird der Wert am Dateneingang ins Flip-Flop eingespeichert (bei $D = 0$ wird es rückgesetzt und bei $D = 1$ wird es gesetzt). Um die zeitliche Abfolge in sequentiellen Schaltungen darzustellen benutzt man so genannte **Schaltfolgediagramme**, in denen die zeitliche Abhängigkeit zwischen den Ein- und Ausgängen und den Zustandsgrößen dargestellt wird. In Bild 7-29 c) ist ein Schaltfolgediagramm für das Einspeichern und Löschen von Daten in einem D-Flip-Flop dargestellt.

7.2.2 Probleme der Modellbildung digitaler Systeme

Die boole'schen Gleichungen, die das logische Verhalten einer realen kombinatorischen oder sequentiellen Steuerung beschreiben, gehen von einem idealisierten Schaltverhalten der verwendeten Bauelemente aus. Wie im Teil a) des Schaltfolgediagramms in Bild 7-30 gezeigt, bedeutet **ideal** dabei, dass die Übergänge vom logischen Wert 0 auf den Wert 1 und umgekehrt keine Zeit beanspruchen, also unendlich schnell erfolgen. Unabhängig davon, ob man nun elektronische, elektromagnetische oder pneumatische Schaltelemente benutzt, ist mit dem Arbeiten immer ein verzögerndes dynamisches Verhalten verbunden, da alle Systeme energiespeichernde Elemente enthalten. Das reale Schaltverhalten sieht daher eher wie im

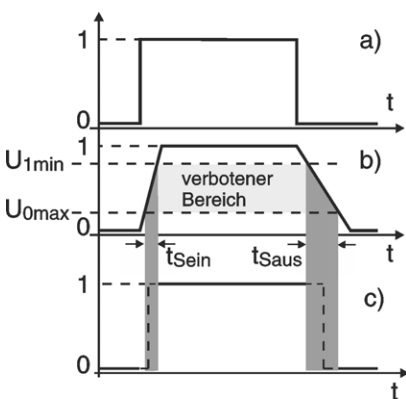


Bild 7-30: Schaltverhalten von elektronischen Bauelementen
a) ideal b) real
c) Schalten nachfolgender Elemente

Teil b) des Schaltfolgediagramms dargestellt aus. Die Übergänge zwischen den Signalpegeln, die den logischen Werten 0 und 1 entsprechen, benötigen endliche Zeit und erfolgen, je nach Richtung des Signalübergangs, auch mit unterschiedlicher Geschwindigkeit. Das im Bild dargestellte elektrische Signal darf aber höchstens eine Spannung von U_{0max} haben, um noch als Signalwert 0 interpretiert werden zu können und es muss für den Signalwert 1 mindestens U_{1min} betragen. Alle dazwischen liegenden Spannungen bezeichnet man als im "verbotenen Bereich", weil hier technisch keine eindeutige Zuordnung zu den binären Signalwerten mehr möglich ist. Innerhalb der Einschaltzeit t_{Sein} und der Ausschaltzeit t_{Saus} kann nicht exakt vorausgesagt werden, wann ein nachfolgendes Schaltelement (Bildteil c)) den Signalpegel wechseln wird.

Außerdem benötigen Signale, die über Leitungen zwischen den Schaltelementen ausgetauscht werden, endliche Signallaufzeiten, wobei die Ausbreitungsgeschwindigkeit von der jeweiligen Technologie abhängt⁸. Da mehrere sich gleichzeitig ändernde Signale sich auf verschiedenen Signallaufwegen mit verschiedenen Signalverzögerungszeiten ausbreiten, kann es beim Wiederausammenführen dieser Signale zu zeitlichen Flankenverschiebungen kommen. Da alle Schaltfunktionen so genannte Funktionshazards enthalten [7.4], können dann fehlerhafte Signalwerte in einer Schaltung beim gleichzeitigen Wechsel von Variablen auftreten.

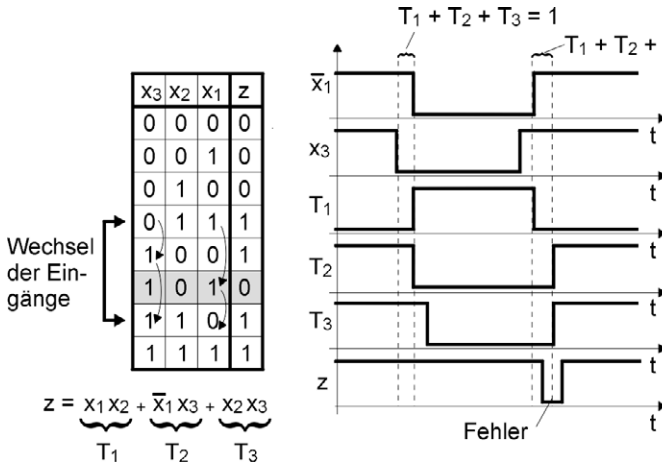


Bild 7-31: Schaltfunktion mit Funktionshazard und Schaltfolgediagramm für doppelten Signalwechsel von x_1 und x_3

In Bild 7-31 ist als Beispiel für die beschriebene Problematik eine Schaltfunktion und ihre Wahrheitstabelle abgebildet. Betrachtet wird der Übergang vom Feld mit dem Minterm $x_1 x_2 \bar{x}_3$ zu dem Feld mit dem Minterm $\bar{x}_1 x_2 x_3$, oder umgekehrt. Der Funktionswert soll zu Anfang und zu Ende des Übergangs jeweils 1 sein und daher seinen Wert nicht ändern. Um diesen Übergang zu

bewirken, müssen gleichzeitig die Variablen x_1 und x_3 ihren Wert ändern. Dieser Übergang ist in der Wahrheitstabelle durch einen starken Doppelpfeil angedeutet. Da man davon ausgehen kann, dass der Signalwechsel der beiden Variablen nicht absolut gleichzeitig stattfindet, wird der Übergang schrittweise, entsprechend den dünnen Pfeilen ablaufen. Der Übergangsbereich für diesen Übergang, besteht aus den vier Zeilen, die durch den kräftigen Doppelpfeil eingeklammert werden.

Innerhalb des **Übergangsbereiches** ist die Funktion nicht konstant gleich 1, da der Minterm $\bar{x}_1 x_2 \bar{x}_3$ (grau unterlegte Zeile) den Funktionswert 0 zur Folge hat. Ob ein Fehler in Form des Einbruchs des 1-Signals am Ausgang einer Schaltungsrealisierung auftritt hängt daher davon ab, in welcher Reihenfolge der Wechsel von x_1 und x_3 erfolgt. Ist die Reihenfolge der Änderungen so, dass der Übergang über das Feld mit dem Funktionswert 0 erfolgt, so zeigt das Ausgangssignal kurzzeitig für die Dauer der Verschiebung der Signalfanken einen Einbruch des 1-Signals auf den Wert 0. Bei anderer Reihenfolge der Änderungen verläuft der Übergang über ein Feld, das auch den Funktionswert 1 enthält und es tritt kein Fehler auf.

8 Bei elektrischen und elektronischen Schaltungen ist dies die Lichtgeschwindigkeit, bei pneumatischen Bauelementen ein Wert kleiner als die Schallgeschwindigkeit in Luft. Dies erscheint bei elektrischen Signalen zwar sehr schnell, wenn man aber bedenkt, dass eine durchschnittliche Signalanstiegszeit bei elektronischen Schaltgattern 10^{-8} Sekunden beträgt, so legt ein elektrisches Signal in dieser Zeit nur einen Weg von 3 Metern zurück.

Diese dynamischen Vorgänge sind im Schaltfolgediagramm des Bildes 7-31 dargestellt, das die Änderungen der Eingangsvariablen bzw. ihrer Negation zeigt. Die Verschiebung der Flanken von $\bar{x}_1$ und x_3 tritt selbst bei absolut gleichzeitiger Änderung von x_1 und x_3 auf, da der für die Invertierung benötigte Negationsbaustein eine endliche Signalverzögerung hervorruft. Ebenfalls im Schaltfolgediagramm sind die Ausgangsgrößen der Schaltelemente dargestellt, die den drei Mintermen der disjunktiven Normalform der Schaltfunktion entsprechen. Man sieht, dass der besagte Signalübergang in der einen Richtung zu keinen Änderungen des Ausgangssignals z führt, dass aber in der anderen Richtung aufgrund der vorhandenen Signalverzögerungen ein kurzzeitiger Signaleinbruch auf den Wert 0 auftritt.

Bei rein kombinatorischen Schaltungen tritt dieser Fehler nur ganz kurzzeitig auf und könnte vielfach toleriert werden. Sind die Schaltungen jedoch sequentiell, so können solche kurzen Impulse in den Speicherelementen gespeichert werden und dadurch ein dauerhaftes Fehlverhalten des Schaltwerks hervorrufen. Insbesondere bei komplexen Schaltwerken mit vielen hintereinander geschalteten Stufen kann dieses Schaltverhalten nicht mehr beherrscht werden. Daher verwendet man die in Bild 7-32 dargestellte Methode der Synchronisation aller Ausgangssignale einer Stufe. Ändern sich mehrere Eingangssignale des Eingangsvektors $\bar{x}$ in diese Stufe gleichzeitig, so werden an den Ausgängen der Verknüpfungsbaugruppen kurzzeitige Fehlimpulse aufgrund verschiedener Signallaufzeiten auftreten. Diese werden jedoch nicht weitergeleitet, da die Flip-Flops die Ausgangswerte erst beim Auftreten des Taktsignals übernehmen. Dieses erscheint erst wenn die fehlerhaften Impulse abgeklungen sind und übernimmt dann die gültigen Verknüpfungsergebnisse am Ende der Übergangsvorgänge und gibt den Ausgangsvektor $\bar{z}$ aus. Auf diese Weise wird in allen getakteten, digitalen Steuerungen wie auch bei normalen Digitalrechnern ein Fehlverhalten aufgrund unvermeidlicher Funktionshazards vermieden.

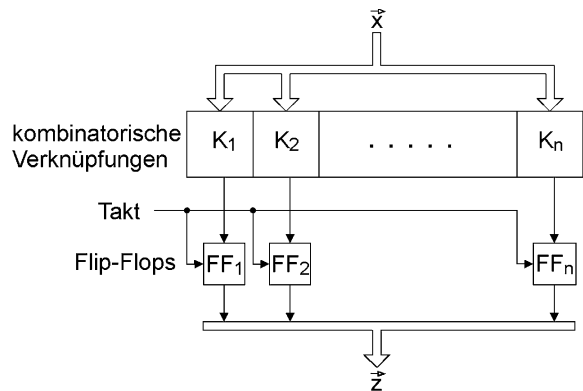


Bild 7-32: Synchronisation einer Verknüpfungsstufe mit Hilfe getakteter Ausgangsspeicher

Der Preis für diesen Vorteil ist die Herabsetzung der Arbeitsgeschwindigkeit solcher Schaltungen, da jetzt Signaländerungen nicht mehr mit der maximalen Signalgeschwindigkeit erfolgen können. Dem trägt man beispielsweise in modernen Digitalrechnern dahingehend Rechnung, dass man die Taktfrequenz immer weiter steigert. Betrug bei einem PC zu Ende der 80'er Jahre die Taktfrequenz noch 16 MHz, so ist sie heute bereits aufgrund technologischer Verbesserungen der Hardware auf das hundertfache und mehr gesteigert worden.

7.2.3 Mehrwertige und unscharfe Logik (Fuzzy Logic)

Im Kapitel über die Modellbildung wurde bereits angesprochen, dass mathematische Modelle technischer Systeme zwangsläufig die Realität vereinfachen müssen und bei ihrer Anwendung die Randbedingungen und der Gültigkeitsbereich des Modells zu beachten sind. Aufgrund der Einfachheit des mathematischen Modells für digitale Systeme (nur zwei Funktionswerte 0 und 1 sind zugelassen), das man mit Hilfe der Boole'schen Algebra beschreiben kann, kann es wie in Abschnitt 7.2.2 ausgeführt, in gewissen Situationen zu Abweichungen zwischen Modellverhalten und realem Verhalten technischer digitaler Systeme kommen.

Gerade die Einfachheit des Modells besticht aber und liegt auf der Entwicklungslinie unseres physikalischen Denkens. In der Physik ist es seit langem üblich, sich mit einfachen, möglichst ungestörten Systemen zu beschäftigen. Dadurch lassen sich einfache mathematische Modelle bilden, die auch mit den konventionellen Möglichkeiten der Mathematik behandelt werden können. Nur durch dieses Denken in scharfen, exakten Begriffen wurde die Aufstellung der Gesetze der klassischen Mechanik möglich, wie sie beispielsweise Newton in Form seiner Bewegungssaxiome entwickelt hat. So ist die Feststellung, dass sich ein einmal beschleunigter Körper mit konstanter Geschwindigkeit fortbewegt, eine reine Fiktion, denn niemand kann dies in seiner normalen Umwelt nachprüfen oder beobachten. Der Augenschein zeigt jedem das Gegenteil der obigen Aussage an, da ein durch Kräfte beschleunigtes Fahrzeug bei Wegfall des Antriebs in kürzerer Zeit stehen bleibt. Jedem ist unmittelbar klar, warum das so ist. Das Modell, in dem obiges Gesetz gilt, tritt in dieser Form nicht auf, da in realen Systemen immer Reibkräfte auf einen bewegten Körper wirken. Trotzdem haben solche Vereinfachungen erst zu dem Erkenntnisstand in der Physik und Technik geführt, den wir heute besitzen.

Die Denkweise, alle Modelle und Aussagen über die Physik von Systemen möglichst einfach zu gestalten, hat uns in der Vergangenheit aber auch teilweise blind für die Wirklichkeit gemacht. Der Drang, die Unbestimmtheit in einem Modell nicht zuzulassen, hat uns Lösungen für technische Probleme, in denen ein hoher Grad von Unbestimmtheit oder Unschärfe enthalten ist, lange Zeit unmöglich gemacht. Es ist daher verständlich, dass eine mathematische Modellbildung und die zugehörige **Theorie der Unschärfe (Fuzzy-Set-Theorie)**, wie sie bereits 1965 von dem amerikanischen Mathematiker L. A. Zadeh entworfen und veröffentlicht wurde, lange Zeit in den exakten Wissenschaften keine Anerkennung fand [7.5], [7.6]. Mit Hilfe der von ihm entwickelten Theorie lassen sich aber bereits heute Steuerungen und Regelungen mit teilweise recht einfachen Mitteln realisieren, wo früher exakte Algorithmen und Gesetzmäßigkeiten keine hinreichende Problemlösung ermöglichten. Eine Darstellung derzeitiger technischer Anwendungen findet sich in [7.7]. Im folgenden verwendete Beispiele sind [7.5] und [7.6] entnommen.

Die Gesetzmäßigkeiten der Boole'schen Algebra lassen sich auf die allgemeine Mengentheorie zurückführen, wie sie bereits Ende des 19. Jahrhunderts durch den deutschen Mathematiker Cantor (1845-1918) entwickelt wurde. Die Mengen, mit denen die Boole'sche Algebra umgeht, enthalten jedoch nur die zwei Elemente 0 und 1. Zustände innerhalb eines binären Systems, können nur den einen oder den anderen Wert besitzen. Im alltäglichen Leben ist diese "Schwarz-Weiß Eindeutig-

keit" eher unwahrscheinlich, und häufig lassen sich solche eindeutigen Zuordnungen nicht treffen. Meist sind Eigenschaften oder physikalische Größen mehrwertig oder gar nicht exakt festgelegt, sie unterliegen gewissen **Unschärfen**.

Dabei muss man drei verschiedene Arten der Unschärfe unterscheiden:

- **Stochastische Unschärfe:** Sie wird durch das Maß der Wahrscheinlichkeit für das Eintreffen eines Ereignisses definiert und kann beliebige Werte zwischen den Werten 0 und 1 annehmen. Beispiel: Die Wahrscheinlichkeit eines Würfelergebnisses ist $1/6$.
- **Sprachliche Unschärfe:** Häufig bleibt bei der sprachlichen Zuweisung von Eigenschaften der Objekte eine gewisse Unklarheit über die Eigenschaft, die häufig nur aus dem Kontext geschlossen werden kann. Beispiele: großes Gebäude, schnelle Maschine, hohe Temperatur.
- **Unschärfe der Information:** Sie beruht auf mangelndem Wissen über ein System oder fehlende Informationen, die sich häufig auch nicht beschaffen lassen. Beispiel: Umformbarkeit von Blech, Schweißbarkeit von Stahl.

Nun fällt es leicht, Objekten anstatt scharfer Eigenschaften unscharfe Eigenschaften zuzuordnen. Um aber mit unscharfen Informationen in technischen Prozessen arbeiten zu können, müssen diese in eine durch einen Digitalrechner verarbeitbare Form gebracht werden. Dazu muss die eindeutige Mengenzugehörigkeit von Elementen (Element gehört zur Menge oder Element gehört nicht zur Menge) durch den Begriff des Zugehörigkeitsgrades zu einer Menge ersetzt werden.

Betrachtet man beispielsweise die Menge M der reellen Zahlen von 2 bis 3

$$M := \{x \mid x \in \text{Re}, 2 \leq x \leq 3\} \quad ,$$

dann gehören Werte für x von 1; 3,3 oder 5 nicht zur Menge M , während Werte für x von 2; 2,8 oder 3 dazugehören. Diese Zugehörigkeit zur Menge M kann man auch grafisch darstellen (Bild 7-33). In dieser grafischen Darstellung erscheint die **Zugehörigkeitsfunktion** $\mu(x)$, die die Zugehörigkeit eines Elementes x zur Menge M angibt.

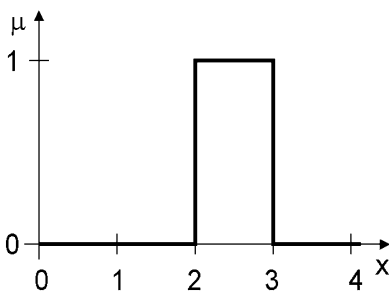


Bild 7-33: Zugehörigkeitsfunktion zur Menge M der reellen Zahlen x zwischen 2 und 3

In der Cantor'schen Mengenlehre kann die Zugehörigkeitsfunktion $\mu(x)$ nur die Werte 0 und 1 (gehört zur Menge, gehört nicht zur Menge) annehmen. Verwendet man nun Begriffe für Zahlen wie "viel größer als", so wird diese Definition der Zugehörigkeit zu einer Menge problematisch. Betrachtet man beispielsweise die Menge der reellen Zahlen, die sehr viel größer als 1 sind

$$M := \{x \mid x \in \text{Re}, x \gg 1\} \quad ,$$

so ist zuerst einmal unklar, welche Zahlen dieser Menge zuzuordnen sind. Dies liegt daran, dass der Begriff "viel größer als" mathematisch nicht eindeutig definiert ist, so dass ohne genauere Festlegungen die Zugehörigkeit einer Zahl zur Menge unklar (unscharf) bleibt. Die duale Modellierung im cantor'schen Sinne muss daher

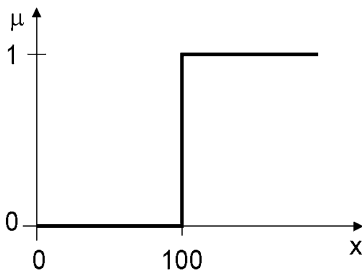


Bild 7-34: Zugehörigkeitsfunktion zur Menge der reellen Zahlen *sehr viel größer als 1*

durch eine andersartige Modellierung erweitert werden. Man könnte jetzt den unklaren Begriff so schärfen, dass man alle Zahlen die größer oder gleich 100 sind als "sehr viel größer als 1" definiert. Die Zugehörigkeitsfunktion $\mu(x)$ von Zahlen x zur Menge M würde dann wie in Bild 7-34 gezeigt aussehen. Dadurch ergäbe sich die Festlegung, dass 100 zur Menge gehört, die reelle Zahl 99,99 jedoch nicht. Dieses unstetige Verhalten der Zugehörigkeit zur Menge der Zahlen, die sehr viel größer als 1 sind, entspricht nun überhaupt nicht dem sprachlichen Empfinden. Die Modellierung mit Hilfe der Cantor'schen Mengenlehre ist deshalb für solche unscharfen Probleme ungeeignet.

Die eigentliche Unzulänglichkeit der oben aufgeführten Definition scheint im unstetigen Verlauf der Zugehörigkeitsfunktion zu liegen. Im normalen Verständnis kann bei einer kontinuierlichen Größe wie den reellen Zahlen nicht eine Zahl nicht zur Menge gehören und eine nur unwesentlich größere Zahl doch dazu gehören. Es wäre naheliegender, die Entfernung von der fiktiven Unstetigkeitsstelle der Zugehörigkeitsdefinition als abgestufte Zugehörigkeit festzulegen. So wäre beispielsweise völlig sicher, dass der Bereich um die Zahl 1 nicht zur Menge gehört

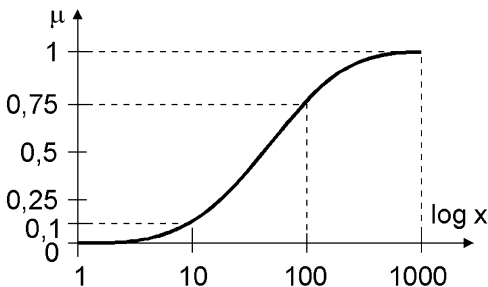


Bild 7-35: Zugehörigkeitsfunktion bei unscharfer Modellierung

(Zugehörigkeitsgrad $\mu = 0$) und auch dass der Bereich um die Zahl 1000 vollständig zur Menge gehört (Zugehörigkeitsgrad $\mu = 1$). Dazwischen könnte man jetzt einen kontinuierlichen Übergang der Zugehörigkeit modellieren (Bild 7-35), so dass man beispielsweise für die Zahl 10 der Zugehörigkeitsgrad $\mu = 0,1$ und für die Zahl 100 den Wert $\mu = 0,75$ annimmt. Die Wertepaare, also den Wert x des Elementes einer Menge und der Zugehörigkeitsgrad $\mu(x)$ zur Menge bezeichnet man als **Singleton**.

Unschärfe Mengen (Fuzzy-Mengen) zeichnen sich dadurch aus, dass Elemente dieser Mengen auch Zugehörigkeitsgrade zwischen 0 und 1 besitzen können. Die Cantor'schen Mengen sind daher Spezialfälle von Fuzzy-Mengen, bei denen nur die Zugehörigkeitsgrade 0 und 1 möglich sind. Fuzzy-Mengen bestehen aus Singletons oder aus durch die Zugehörigkeitsfunktion zur Menge definierten Wertepaaren.

Häufig kommt im Zusammenhang mit Fuzzy-Mengen die Frage auf, wo der Unterschied zur Wahrscheinlichkeit liegt. Diese Begriffe dürfen nicht verwechselt oder synonym benutzt werden. Die von dem französischen Mathematiker Blaise Pascal (1623 - 1663) entwickelte **Wahrscheinlichkeitsrechnung** befasst sich mit Situationen, wo Vorhersagen über zukünftige Ereignisse gemacht werden sollen, deren Ausgang zufällig erfolgt, die aber beliebig häufig wiederholt werden können.

Ein typisches Beispiel dafür ist das Fallen eines Würfels. Wenn man einen Würfel sehr oft wirft, wird er ziemlich genau in einem Sechstel aller Fälle den Wert 1 zeigen. Die Wahrscheinlichkeit des Ereignisses lautet:

$$\text{Wahrscheinlichkeit} = \frac{\text{Anzahl der günstigen Ereignisse}}{\text{Anzahl aller möglichen Ereignisse}}$$

Die Wahrscheinlichkeit, eine Zahl kleiner 6 zu werfen, ist beispielsweise 5/6.

Von Wahrscheinlichkeiten redet man meist im Hinblick auf zukünftige Ereignisse. Sie sind daher Ausdruck für einen Mangel an Information. Es ist ja nicht wirklich so, dass der Würfel zufällig eine 6 anzeigt, sondern im konkreten Fall ist der Grund für das Auftreten des Zahlenwertes die aktuelle Wurfbahn mit all ihren Randbedingungen. Der Zufall ist nur der Ausdruck über unsere Unkenntnis der genauen physikalischen Bedingungen beim Wurf. Liegt die konkrete Information vor, so verschwindet die Wahrscheinlichkeit. Hieraus ergibt sich der Unterschied zur Zugehörigkeit zu einer Fuzzy-Menge.

Angenommen, man möchte Mineralwasser kaufen, dessen Nitratgehalt unterhalb des Grenzwertes liegt, so besteht bei einer Wahrscheinlichkeit von 95%, dass dies zutrifft die Möglichkeit, dass man ein Mineralwasser bekommt, dessen Nitratgehalt deutlich oberhalb des Grenzwertes liegt. Gehören die in Frage kommenden Mineralwässer aber zu einem Grad von 0,95 zur Menge der Mineralwässer mit niedrigerem Grenzwert, so wird der Nitratgehalt auf jeden Fall unterhalb des Grenzwertes liegen.

Um nun mit Fuzzy-Mengen technische Sachverhalte beschreiben und später auch steuern oder regeln zu können, muss man zuerst die in Frage kommenden Größen **fuzzifizieren**. Betrachtet man beispielsweise die Temperatur von Wasser, so kann man durch Messung natürlich ihren Wert exakt bestimmen. Zu einer Fuzzy-Größe wird die Temperatur erst, wenn man sie Begriffen wie niedrig, mittel oder hoch zuordnen will. Die Temperatur bezeichnet man dann als **linguistische Variable** (sprachliche Variable) und die genannten Eigenschaften als **linguistischen Terme**.

Bei der Modellierung der Zugehörigkeitsfunktion einer bestimmten Temperatur zur Menge der niedrigen Temperaturen, kann man natürlich verschiedenartige Kennlinien benutzen. Wie in Bild 7-36 gezeigt, kann die Kennlinie beispielsweise glockenförmig oder dreiecksförmig sein. Die

Annahme einer komplizierten Kennlinienform bringt in der Praxis aber kaum Vorteile, so dass man häufig dreiecks- oder trapezförmige Kennlinien annimmt, da dann das Rechnen mit den Zugehörigkeitsgraden vereinfacht wird. Die Modellierung der dreiecksförmigen Kennlinie erfolgt hier so, dass Temperaturen unter 15°C und über 45°C einen Zugehörigkeitsgrad von $\mu = 0$ und die Temperatur von 30°C einen Zugehörigkeitsgrad von $\mu = 1$ erhält.

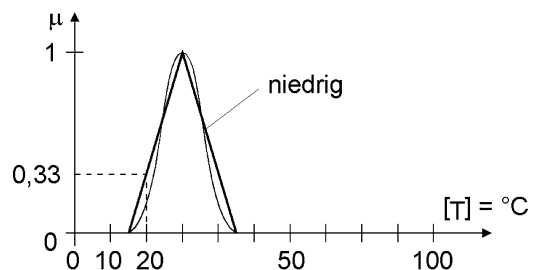


Bild 3-36: Zugehörigkeitsfunktion für *niedrige* Temperaturen

hörigkeitsgrad von $\mu = 1$ zur Menge der niedrigen Temperaturen besitzen. Die Temperatur von beispielsweise 20°C besitzt demnach einen Zugehörigkeitsgrad von $\mu = 0,33$.

Um den gesamten Temperaturbereich, in dem Wasser unter Normaldruck flüssig ist, erfassen zu können, benötigt man weitere linguistische Terme, wie etwa sehr niedrig, niedrig, mittel, hoch und sehr hoch.

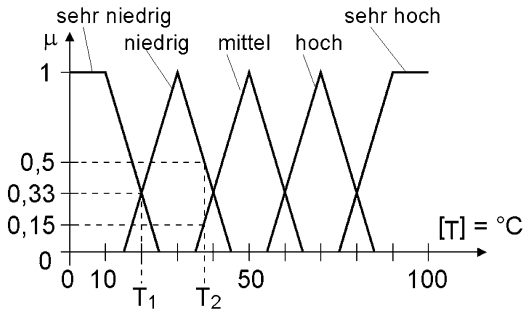


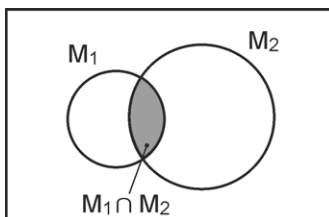
Bild 7-37: Zugehörigkeitsfunktion zur Menge der reellen Zahlen sehr viel größer als 1

Wie viele Terme man verwendet, hängt vom Anwendungsfall ab, die Anzahl liegt normalerweise im Bereich 2 bis 7. Als Kennlinien für die Zugehörigkeitsfunktion am rechten und linken Rand des Definitionsbereiches verwendet man häufig trapezförmige Verläufe. Eine Aufteilung der verschiedenen Fuzzy-Mengen auf den gesamten Temperaturbereich zeigt Bild 7-37. Im Bild sind die beiden exakten Temperaturen von $T_1 = 20^\circ\text{C}$ und $T_2 = 37,5^\circ\text{C}$ eingezeichnet. Wie man ablesen kann, ist die

Temperatur T_1 jeweils zu einem Grad von 0,33 **sehr niedrig** oder auch **niedrig**. Die Temperatur T_2 ist zu einem Grad von 0,5 **niedrig** und zu einem Grad von 0,15 **mittel**. Würde man die Temperatur T_2 rein linguistisch definieren, so könnte man sagen, sie ist niedrig bis mittel, eher niedrig.

Aus dem scharfen Temperaturwert T wird damit ein Fuzzy-Temperaturwert T^* , der durch ein n -Tupel von Zugehörigkeitsgraden beschrieben werden kann, wobei n die Anzahl der bei der Fuzzifizierung verwendeten linguistischen Terme ist:

$$T^* = (\mu_{\text{sehr niedrig}}(T), \mu_{\text{niedrig}}(T), \mu_{\text{mittel}}(T), \mu_{\text{hoch}}(T), \mu_{\text{sehr hoch}}(T))$$



Die Fuzzifizierung liefert demnach für die beiden Temperaturen T_1 und T_2 :

$$T_1^* = (0,33; 0,33; 0; 0; 0) \quad T_2^* = (0; 0,5; 0,15; 0; 0)$$

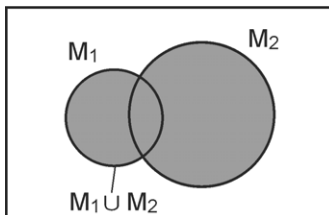


Bild 7-38: Venn-Diagramme von Durchschnitt und Vereinigung

In der Regel wird ein Problem, das mit der Fuzzy-Technologie gelöst werden soll, durch mehrere linguistische Variablen beschrieben, die man dann durch mathematische Operationen miteinander verknüpfen möchte. Aus der Cantor'schen Mengenlehre sind Verknüpfungsoperationen wie **Durchschnitt** ($\cap$) und **Vereinigung** ($\cup$) bekannt. Die Bildung der Durchschnittsmenge und der Vereinigungsmenge zweier Mengen M_1 und M_2 ist in Bild 7-38 in Form so genannter **Venn-Diagramme** dargestellt. In der Boole'schen Algebra entsprechen diese Operationen den Funktionen UND bzw. ODER.

Diese Mengenoperationen kann man direkt auf Fuzzy-

Mengen übertragen und so die in Frage kommenden Eingangsgrößen aus linguistischen Variablen fuzzifizieren. Mit Hilfe ähnlicher Verknüpfungsmethoden und Bildungsvorschriften für Ausgangsgrößen kann man dann einen mehrwertigen Steuerungs- oder Regelungsalgorithmus erstellen, mit dem man ansonsten nur schwer fassbare technische Problemstellungen lösen kann. Am Ende des Modellierungsprozesses steht die so genannte **Fuzzy-Inferenz**, mit der die Fuzzy-Größen defuzzifiziert werden, um wieder exakte Ausgangsgrößen zu erzeugen. Genauer zur Vorgehensweise findet sich in [7.5] - [7.7].

7.2.4 Neuronale Netzwerke

Neuronale Netzwerke werden verwendet, wenn man die Eigenschaften des menschlichen Gehirns nutzen will, in denen sich dieses von einem "Elektronengehirn", d. h. einem Computer, unterscheidet. Wie bereits in Abschnitt 7.1.5 erwähnt, besteht ein Gehirn aus ca. 10^{10} Neuronen genannten Zellen, die in einer komplizierten Vernetzung miteinander verbunden sind. Die Signalübertragung zwischen den Zellen erfolgt elektrisch, wobei die dazu erforderlichen elektrischen Potentiale auf biochemischem Wege erzeugt werden. Die Gesamtstruktur ist hochgradig parallel ausgelegt, im Gegensatz zur seriellen Verarbeitung von Informationen durch Digitalrechner. Dadurch besitzen neuronale Netzwerke eine hohe Redundanz und Fehlertoleranz. Aus dieser komplexen Struktur erwächst "Intelligenz", so dass es naheliegend ist, zu versuchen, intelligentes Verhalten von Systemen durch so strukturierte Netzwerke zu erreichen [7.18].

Eine besonders wichtige Eigenschaft des Gehirns ist aber seine Fähigkeit zu lernen. Dieses **Lernen** erfolgt durch das Wachsen neuer Verbindungen zwischen den Neuronen, was bei konventionellen Rechnern einer Erstellung neuer Programme entsprechen würde.

Im Bereich solcher Aufgabenstellungen wie Mustererkennung oder Bewegungssteuerung in unbekanntem Umfeld sind Methoden des Erlernens von Informationen sehr viel effektiver und robuster als der Versuch, alle denkbaren Möglichkeiten beispielsweise in Form eines Expertensystems im Voraus abzuspeichern.

Aufgrund des biologischen Vorbilds baut man neuronale Netzwerke ähnlich wie das menschliche Gehirn auf, indem man versucht, die Eigenschaft eines biologischen Neurons in einer elektronischen Realisierung zu modellieren und eine Vielzahl solcher künstlicher Neuronen in Schichten hintereinander anzuordnen und zu vernetzen.

Dabei werden verschiedene Modelle verwendet, die eine, zwei oder mehrere Schichten verwenden. Hat ein Netzwerk nur eine Schicht, so beinhaltet jedes Neuron die Funktionen Ein-, Ausgabe und Verarbeitung der Informationen. Da es hier keine Verbindungsstruktur gibt, wird der Anpassungsprozess (Lernen) des Netzwerks durch veränderliche Gewichtung der Eingänge der Neuronen erreicht. Ein typisches Beispiel für solche Art von Neuronen ist das in Bild 7-55 gezeigte

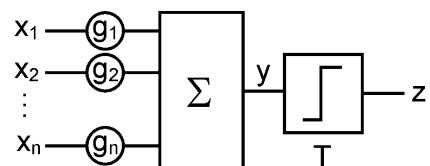


Bild 7-39: McCulloch-Pitts-Neuron

McCulloch-Pitts-Neuron [5.10], benannt nach zwei amerikanischen Biophysikern (1943).

Ein vereinfachtes biologisches Neuron, das das Modell für ein künstliche Neuron liefert, besteht aus dem **Soma** genannten Zellkörper, der die eigentliche Berechnung durchführt, den **Dendriten**, die die Eingangsinformationen liefern, und einem oder mehreren **Axons**, die die Verbindung zu den übrigen Neuronen herstellen. Die ausgetauschten Signale sind elektrische Impulsfolgen, deren Frequenz die Signalstärke beinhaltet. Das vereinfachte biologische Neuronenmodell geht davon aus, dass die Frequenz des Ausgangssignals proportional zur Aktivität des Neurons ist. Bei künstlichen Neuronen wird je nach Anwendungsgebiet das Ausgangssignal analog belassen, oder es wird durch Vergleich mit einem Schwellwert in ein binäres Signal gewandelt. Beim McCulloch-Pitts-Neuron werden die Eingangssignale der biologischen Neuronen durch ein kontinuierliches Eingangssignal (keine Impulsfolge) ersetzt. Die "biochemische Kodierung" der Neurotransmitter in den Synapsen⁹ werden durch eine multiplikative Gewichtung der Stärke g_n (Bild 7-39) ersetzt. Die **Schwellwertfunktion** (Threshold) des biologischen Neurons wird durch einen **Komparator** realisiert und die Impulsfolge des Ausgangssignals durch einen Binärwert ersetzt.

Das McCulloch-Pitts-Neuron berechnet zunächst die gewichtete Summe y aller n Eingangssignale x_n :

$$y = \sum_{i=1}^n x_i \cdot g_i \quad .$$

Die mit den festen Gewichten g_i gewichtete Summe y wird dann mit einem festen Schwellwert T verglichen um die Ausgangsgröße z zu ermitteln. Ist y größer als T , so gilt $z = 1$ (Neuron aktiv), andernfalls ist $z = 0$ (Neuron inaktiv). Bei korrekt gewählten Eingangsgewichtungen g_i kann man mit Hilfe einer Anordnung solcher Neuronen einen Universalrechner bauen, der jede beliebige Funktion berechnen kann. Das Problem besteht aber in der "korrekten" Wahl aller entsprechender Gewichtungen.

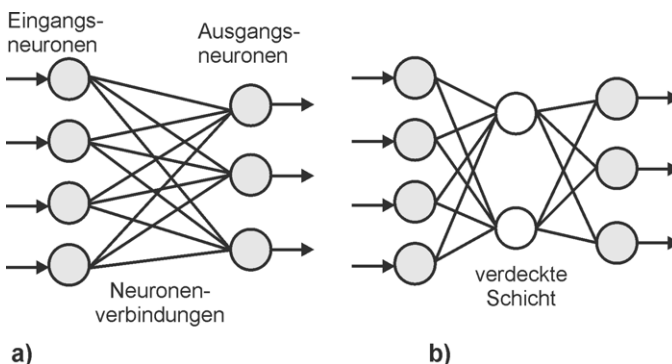


Bild 7-40: Schematische Darstellung mehrschichtiger neuronaler Netzwerke a) Perceptron b) Backpropagation-Modell

Bei zweischichtigen Netzwerken gibt es jeweils eine Ein- und eine Ausgabeschicht (Input-, Output-Layer). Ein typischer Vertreter einer solchen Struktur ist das **Perceptron**, ein von Rosenblatt 1962 vorgeschlagenes Netzwerk (Bild 7-40 a)). Häufig werden auch Strukturen mit einer oder mehreren Schichten zwischen

⁹ Koppelspalt zwischen dem Axon des sendenden und dem Dendrit des empfangenden Neurons. Der dort vorhandene Neurotransmitter kann bewirken, dass eine Synapse erregend oder hemmend ist.

den Ein-, Ausgabeschichten verwendet. Diese so genannten **verdeckten Schichten** (hidden layer), die keine Verbindung zur Außenwelt besitzen, sind nur für die interne Verarbeitung zuständig. Die Struktur eines solchen **Backpropagation-Modells** zeigt Bild 7-40b.

Das Lernverfahren, mit dem Eigenschaften erlernt werden, ist abhängig von der Anzahl der Schichten. Je mehr Schichten, umso komplexer gestaltet sich der erforderliche **Lernalgorithmus**. Bei mehr als zwei Schichten vereinheitlicht sich in den meisten Fällen der Lernalgorithmus, da es prinzipiell unbedeutend ist, ob der Algorithmus die Änderungen in der Stärke der Neuronenverbindungen für eine oder für mehrere Schichten berechnen muss.

7.2.4.1 McCulloch-Pitts-Neuronen

Die Steuerung eines einfachen Systems mit Hilfe künstlicher Neuronen kann man gut an einer Art von einfachen Bewegungssystemen, den von seinem Erfinder so bezeichneten Braitenberg-Vehikeln [7.19] darstellen. Dies sind ganz einfache Fahrzeuge (Bild 7-57) mit zwei durch Einzelmotoren angetriebenen Hinterrädern und zwei Sensoren am Vorderende, die rechts und links vom Fahrzeug angeordnet sind. Solchen einfachen Fahrzeugen kann man durch Abfrage der Sensoren und entsprechende Ansteuerung der Motoren mit z. T. sehr einfachen Steuerungen quasi "intelligentes" Bewegungsverhalten geben.

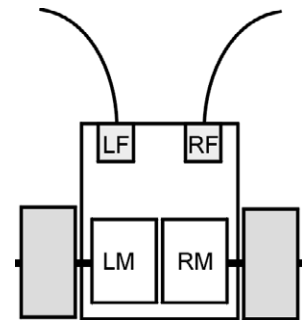


Bild 7-41: Braitenberg-Vehikel mit Fühlersensoren

In dem hier zu zeigenden Fall sollen zwei Neuronen so eingerichtet werden, dass bei Verwendung von Berührungsfühlern als Sensoren das Fahrzeug bei seiner Bewegung Hindernissen ausweichen kann, bei ansonsten geradeaus gerichtetem Fahrverhalten. Die mit LF und RF bezeichneten Fühler an der linken und rechten Vorderseite des Fahrzeugs sollen 1-Signal zeigen, wenn sie berührt werden, ansonsten 0-Signal. Die mit LM und RM bezeichneten Motoren links und rechts am Fahrzeug können sich vorwärts bewegen, wenn sie 1-Signal (positive Spannung) erhalten und bewegen sich rückwärts, wenn sie einen Signalwert "-1" (negative Spannung) erhalten.

Um das erwünschte Verhalten zu erzeugen, fertigt man am besten eine Funktionstabelle an (Bild 7-42). Die beiden zu verwendenden McCulloch-Pitts-Neuronen müssen als Ausgangssignale entweder "1" oder "-1" liefern. Zur Bestimmung der notwendigen Gewichtungen g_{RF} und g_{LF} für den linken Motor *LM* muss zuerst der Schwellwert T festgelegt werden. Für diese Wahl kann man die erste Zeile der Funktionstabelle in Bild 7-40 heranziehen, die besagt, dass beide Motoren laufen müssen, wenn beide Fühler keinen Kontakt melden:

LF	RF	LM	RM
0	0	1	1
0	1	-1	1
1	0	1	-1
1	1	beliebig	beliebig

Bild 7-42: Funktionstabelle für Hindernisausweichen

$$g_{LF} \cdot LF + g_{RF} \cdot RF > T$$

Ist der Schwellwert kleiner "0", so gilt das automatisch, da LF und RF gleich "0" sind. Daher wird $T = -0,01$ gewählt. Der zweiten Zeile der Funktionstabelle kann man entnehmen, dass $g_{RF} < T$ sein muss, damit bei $RF = 1$ der linke Motor das Ausgangssignal $LM = -1$ erhält. Daher wird $g_{RF} = -0,3$ gewählt. Entsprechend sieht man aus Zeile drei der Tabelle, dass $g_{LF} > T$ sein muss, damit für $LF = 1$ der Motor $LM = 1$ erhält. Also wird beispielsweise $g_{LF} = 0,3$ gewählt. Da die erforderlichen Ausgangssignale für den rechten Motor genau entgegengesetzt zu denen für den

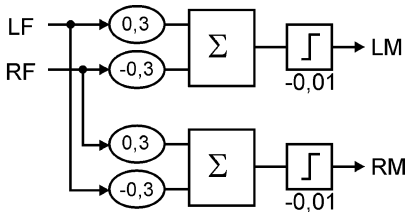


Bild 7-43: Neuronales Netzwerk für „Hindernisausweichen“

linken Motor sein müssen, braucht man hier die Gewichte ebenfalls nur entgegengesetzt zum linken Motorneuron zu wählen. Das komplette neuronale Netzwerk für das geforderte Verhalten „Hindernisausweichen“ ist in Bild 7-43 dargestellt. Da für die vierte Zeile der Funktionstabelle, bei der beide Fühler gleichzeitig gedrückt sind, kein Motorverhalten spezifiziert war, ergibt sich natürlich ein unvorhergesehenes Verhalten für diesen Fall. Da die gewichteten Eingangssignale durch die Wahl eines Absolutwertes von 0,3 sich

gerade gegenseitig aufheben, würde nun das gleiche Verhalten wie für die erste Zeile erzeugt, d. h. das Fahrzeug fährt geradeaus. Sinnvoll wäre als Ausweichmanöver aber ein Rückwärtsfahren, d. h. $LM = RM = -1$. Damit dies erreicht wird muss $g_{RF} + g_{LF} < T$ sein. Für eine Wahl von $g_{LF} = 0,3$ und $g_{RF} = -0,5$ für das Motorneuron, das LM liefert, werden alle vier Zeilen der modifizierten Funktionstabelle erfüllt. Entsprechend müssen die Gewichte des anderen Motorneurons $g_{LF} = -0,5$ und $g_{RF} = 0,3$ sein.

Die Gewichte konnten in diesem Beispiel intuitiv gewählt werden, da das Problem einfach und leicht überschaubar war. Für kompliziertere Funktionen ist dieses Vorgehen aber schwierig. Besser wäre ein Lernmechanismus, der die benötigten Gewichtungswerte selbst bestimmen könnte. Dadurch würde ein neuronales Netzwerk die geforderte Eigenschaft der Lernfähigkeit bekommen. Dies erreicht man mit dem Perceptron, einem zweischichtigen Netzwerk von McCulloch-Pitts-Neuronen.

7.2.4.2 Perceptron

Das Perceptron ist ein „einschichtiges“ neuronales Netzwerk, das aus zwei Schichten von McCulloch-Pitts-Neuronen aufgebaut ist (Bild 7-44). Da die Eingangsschicht die Signale nur weitergibt und auf die zweite Schicht verteilt, die die eigentlichen Berechnungen ausführt, spricht man von einem einschichtigen Netzwerk. Dieses Netzwerk ist einfach zu implementieren, benötigt auf dem Modellrechner nur geringe Rechnerkapazität und zeigt hohe Lerngeschwindigkeit.

Das Ausgangssignal der Einheit z_j wird folgendermaßen bestimmt:

$$z_j = f\left(\sum_{k=1}^n g_{jk} \cdot x_k\right) = f(\vec{g}_j \cdot \vec{x})$$

Hierbei ist $\bar{g}_j$ der individuelle Gewichtungsfaktor der Ausgangseinheit j , n die Zahl der Eingangseinheiten und f die sog. **Transferfunktion**. Diese Funktion des Perceptrons ist wegen des Schwellwertkomparators am Ausgang als Schrittfunktion (Signum-Funktion) definiert:

$$f(x) = \begin{cases} 1 & \forall x > T \\ 0 \text{ (oder } -1) & \text{andernfalls} \end{cases}$$

Es gibt auch entsprechende Netzwerke ohne Komparator, die ein kontinuierliches Ausgangssignal liefern und **Assoziativspeicher** genannt werden.

Wie schon in 7.2.4.1 gesagt, hat das Perceptron die Eigenschaft, dass es die Wahl der richtigen Gewichtungen für das Erzeugen einer bestimmten Funktionalität durch Lernen selbst bestimmt. Das Lernen erfolgt iterativ in diskreten Schritten nach einer bestimmten Regel unter Anleitung durch einen Lehrer. Das Inkrement, um das ein Gewicht beim Lernen jeweils verändert werden soll, wird folgendermaßen berechnet:

$$\Delta \bar{g}_k(t) = \eta(t)(\tau_k - z_k) \bar{x} \quad .$$

Dabei ist der erwünschte Ausgangswert für die Einheit k und z_k der tatsächlich erhaltene Ausgangswert. Die Lerngeschwindigkeit wird von der Lernrate $\eta(t)$ bestimmt. Der Maximalwert für $\eta(t)$ beträgt 1, ein hoher Wert von beispielsweise $\eta(t) = 0,8$ würde ein Netzwerk ergeben, das sich sehr schnell auf Veränderungen einstellen kann. Als nachteilig wäre dabei anzusehen, dass ein solches Perceptron "neurotisch" reagiert, d. h. sobald ein paar Zufallssignale eintreffen, vergisst es alles Gelernte und erlernt Neues. Umgekehrt ergibt ein niedriger Wert von beispielsweise $\eta(t) = 0,1$ ein "lethargisches" Netzwerk, das lange braucht, um eine Funktion zu erlernen. Nach jeder Bestimmung der Gewichtsänderung um Δg_k während des Lernprozesses, müssen dann iterativ die Gewichte nach folgender Gleichung neu berechnet werden:

$$\bar{g}_k(t+1) = \bar{g}_k(t) + \Delta \bar{g}_k \quad .$$

Um die Vorgänge beim Lernen zu verdeutlichen, soll nochmals das Beispiel aus Bild 7-41 behandelt werden, in dem ein Braitenberg-Vehikel mit Berührungsfühlern einem Hindernis ausweichen soll. Die Struktur des zu verwendenden Perceptrons ist in Bild 7-44 dargestellt. Die Gewichte sind durch Indizierung mit dem jeweiligen Fühler- und Motornamen eindeutig gekennzeichnet.

Als Anfangswahl werden die Schwellwerte beider Komparatoren wieder auf $T = -0,01$

festgelegt und eine mittlere Lernrate von $\eta(t) = 0,3$ festgelegt. Der Lernprozess kann beispielsweise damit beginnen, dass die Gewichte mit

$$g_{LFLM} = g_{RFLM} = g_{LFRM} = g_{RFRM} = 0$$

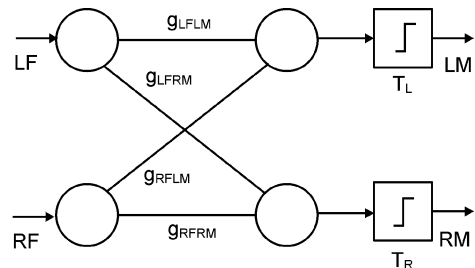


Bild 7-44: Struktur eines Perceptrons für das „Hindernisausweichen“

starten. Da dadurch die Eingangswerte in die beiden Komparatoren Null sind, die Schwellwerte aber $-0,01$ betragen, geben beide Motorknoten jeweils den Wert $z_{LM} = z_{RM} = 1$ aus. Entsprechend der Gleichung für den Lernprozess ergibt sich dann für die erste Zeile der Funktionstabelle in Bild 7-42, die $\tau_{LM} = \tau_{RM} = 1$ erfordert:

$$g_{LFLM} = 0 + 0,3(1-1)0 = 0 \quad g_{RFLM} = 0 + 0,3(1-1)0 = 0$$

$$g_{LFRM} = 0 + 0,3(1-1)0 = 0 \quad g_{RFRM} = 0 + 0,3(1-1)0 = 0$$

Für die zweite Zeile der Funktionstabelle, in der $LF = 0$, $RF = 1$ gilt, muss $\tau_{LM} = -1$, $\tau_{RM} = 1$ sein, während die Ausgänge tatsächlich zur Zeit $z_{LM} = z_{RM} = 1$ betragen. Dadurch errechnet sich im nächsten Lernschritt:

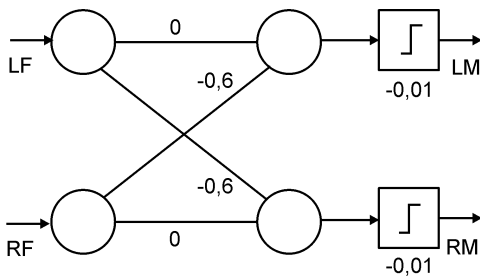
$$g_{LFLM} = 0 + 0,3(-1-1)0 = 0 \quad g_{RFLM} = 0 + 0,3(-1-1)1 = -0,6$$

$$g_{LFRM} = 0 + 0,3(1-1)0 = 0 \quad g_{RFRM} = 0 + 0,3(1-1)1 = 0$$

Entsprechend ergibt sich für Zeile drei der Funktionstabelle, in der für $LF = 1$, $RF = 0$ die Ausgangswerte $\tau_{LM} = 1$, $\tau_{RM} = -1$ erreicht werden sollen:

$$g_{LFLM} = 0 + 0,3(1-1)1 = 0 \quad g_{RFLM} = -0,6 + 0,3(-1-1)0 = -0,6$$

$$g_{LFRM} = 0 + 0,3(-1-1)1 = -0,6 \quad g_{RFRM} = 0 + 0,3(-1-1)0 = 0$$



Überprüft man das nun mit Wertigkeiten belegte Perceptron (Bild 7-45) daraufhin, ob es bereits nach diesen drei Schritten in der Lage ist, das Problem des „Hindernisausweichens“ zu lösen, so stellt man fest, dass dies zutrifft. Auch die Eigenschaft, bei gleichzeitig gedrückten Fühlern durch eine Rückwärtsbewegung auszuweichen, ist erfüllt.

Bild 7-45: Perceptron mit gelernten Gewichten für das „Hindernisausweichen“

Das Perceptron ist ein **feed forward-Netzwerk**, da die Neuronenschichten nur Verbindungen nach vorne besitzen und

auch keine Querverbindungen zwischen den Neuronen einer Schicht existieren. Der Lernalgorithmus ist auf einen **Lehrer** angewiesen. Für das Lernen mit Lehrer muss zu jedem zu erlernenden Muster ein Ziel-Ausgangs-Muster vorhanden sein. Der Lehrer berechnet dann die für den nächsten Schritt des Lernprozesses erforderliche Änderung der Verbindungsstärken, in die die vorhandene Abweichung des ungelerten Netzes zwischen tatsächlichem Ausgangsmuster und dem Soll-Ausgangsmuster eingeht. Nur über diese durch den Lehrer generierten Informationen kann das Netz bestimmen, wie viel und in welche Richtung (hemmend, erregend) die internen Neuronenverbindungen geändert werden müssen.

Unter 7.2.4 war festgestellt worden, dass man aus McCulloch-Pitts-Neuronen einen Universalrechner bauen kann, der in der Lage ist, jede beliebige Funktion zu berechnen. Beim Perceptron findet man leicht, dass es Fälle von Funktionen gibt, die sich nicht mit seiner Hilfe realisieren lassen. Ein Beispiel für diese Funktion ist das Exklusiv-Oder, deren Ausgang immer dann den Wert 1 annimmt, wenn die Belegung der beiden Eingangsgrößen verschieden ist. Sollte ein Ausgang des Percept-

rons wie in Bild 7-44 diese Funktion realisieren, so müssten folgende Ungleichungen erfüllt sein:

$$g_{LF} > T, g_{RF} > T, g_{LF} + g_{RF} < T.$$

Die ersten beiden Ungleichungen ergeben zusammen $g_{LF} + g_{RF} > 2T$, was im Gegensatz zur dritten Ungleichung steht. Das bedeutet, dass es keine Werte für die Gewichte eines Perceptrons gibt, die die Funktion Exklusiv-Oder der Eingangsgrößen realisiert. Allgemein gilt, dass Perceptrons nicht in der Lage sind Funktionen zu erlernen, die nicht **linear trennbar** sind.

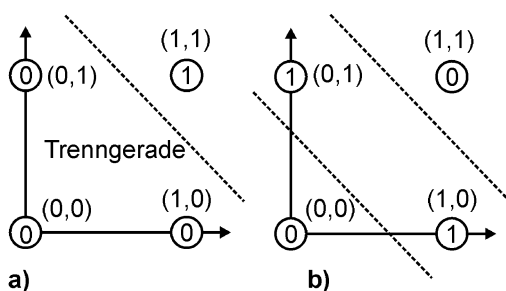


Bild 7-46: Lineare Trennbarkeit von Funktionen
a) UND-Funktion b) Exklusiv-Oder-Funktion

Was dies bedeutet, ist in Bild 7-46 für die UND- und die Exklusiv-Oder-Funktion dargestellt. Im Bildteil a) und b) ist jeweils in der Ebene ein Achsenkreuz eingezeichnet, wobei entlang jeder Achse eine der Eingangsgrößen aufgetragen wird. Diese hat jeweils am Anfang der Achse den Wert 0 und am Ende den Wert 1. Die Kombination der Belegungen der Eingangsgrößen (0,0), (0,1), (1,0) und (1,1) ergibt 4 Punkte in der Ebene, an denen die jeweilige Funktion einen bestimmten Wert hat, der in einen

Kreis an dieser Stelle eingetragen ist. Bildteil a) zeigt die UND-Funktion. Innerhalb der Ebene der Funktion lässt sich eine Gerade einzeichnen, die die Bereiche, für die die Funktion den Wert 1 bzw. 0 besitzt, voneinander trennt. Das Gleiche ist für die in Teilbild b) dargestellte Exklusiv-Oder-Funktion nur durch zwei Geraden zu erreichen, die Funktion ist nicht linear trennbar. Da es viele solcher Funktionen gibt, die nicht linear trennbar sind, ist das einfache neuronale Netzwerk "Perceptron", obwohl es eine hohe Lerngeschwindigkeit besitzt, nicht immer einsetzbar. Eine Implementierung beliebiger Funktionen ist jedoch stets mit einem mehrschichtigen Netzwerk, wie in Bild 7-40 b) dargestellt, möglich.

7.2.4.3 Backpropagation-Netzwerk

Dieser Typ von neuronalen Netzwerken wurde erstmals von Rumelhart, Hinton und Williams im Jahr 1986 exakt beschrieben. Es hat eine andere Topologie wie das Perceptron und eine andere Transferfunktion. Es hat einen prinzipiellen Aufbau wie in Bild 7-40 b), der zusätzlich zur Eingabe- und Ausgabe-Schicht des Perceptrons eine weitere verdeckte Schicht besitzt, die keinen direkten Kontakt zur Außenwelt hat. Durch diese zusätzliche Schicht ist das **Backpropagation-Netzwerk** in der Lage, eine interne Repräsentation des angelegten Eingangsmusters aufzubauen. Eine weitere die Leistungsfähigkeit steigernde Änderung fand bei der Transferfunktion statt, mit der in einem neuronalen Netz bestimmt wird, bei welcher Summe der Eingangssignale das Neuron aktiviert wird. Statt einer linearen Schwellwertfunktion, wie sie meist beim Perceptron verwendet wird, benutzt man die nichtlineare **Sigmoiden-Funktion**. Zur Bestimmung der Gewichte zwischen den einzelnen neuronalen Knoten wird der so genannte Backpropagation-Algorithmus verwendet. Bei

diesem Netz kann die Änderung der Gewichtung der Neuronenverbindungen zwischen der Eingangsschicht und der verdeckten Schicht nicht aufgrund der Soll-Ausgangswerte vorgenommen werden, da diese für die verdeckte Schicht unbekannt sind. Anstelle dessen muss der Fehler, den das Netz bei der Berechnung des Ausgangs-Vektors (Ausgänge der Ausgangsschicht) macht, durch das Netz "zurückgereicht" werden, da auch die Gewichte der Neuronenverbindungen der verdeckten Schichten zu diesem Fehler beigetragen haben. Aufgrund dieser Art der Änderung der Neuronenverbindungen nennt man dieses Modell *Fehlerrückführungsnetz*, **error back propagation** oder kurz **Backpropagation-Netz**.

Der Lernalgorithmus eines solchen Netzes beginnt damit, dass alle Gewichte des Netzes auf Zufallswerte gesetzt werden. Um Konvergenzprobleme des Lernalgorithmus zu vermeiden werden keine festen Schwellwerte T verwendet, sondern deren Bestimmung wird Bestandteil des Algorithmus. Es werden anfangs alle Schwellwerte auf Null gesetzt und jedem Neuron ein zusätzlicher fester Eingangswert von +1 aufgeschaltet (bias: engl. Tendenz). Wenn das Netz initialisiert ist, wird es mit vorgegebenen Eingangs-/Ausgangspaaren, die die erwünschte Funktion repräsentiert, trainiert. Diese Trainingsmuster werden dann zur Aktualisierung der Gewichtungswerte verwendet. Der Ausgangswert z_j jeder Einheit j des Netzwerkes wird nun folgendermaßen berechnet:

$$\bar{z}_j = f(\bar{g}_j \cdot \bar{i}) \quad .$$

Dabei ist $\bar{g}$ der Gewichtungsvektor der betrachteten Einheit j und $\bar{i}$ der Eingangsvektor der Einheit. Die Transferfunktion f ist anders als die einfache binäre Schwellwertfunktion "Signum-Funktion" beim Perceptron nun die differenzierbare Sigmoid-Funktion (Bild 7-47):

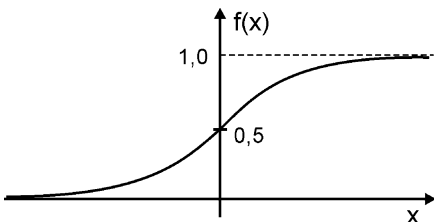


Bild 7-47: Sigmoid-Funktion

$$f(x) = \frac{1}{1 + e^{-kx}} \quad ,$$

wobei k eine positive Konstante ist, die die Steigung der Sigmoiden bestimmt. Für $k \rightarrow \infty$ geht die Sigmoid-Funktion in die Signum-Funktion über. Sobald die Ausgangswerte aller Neuronen, sowohl in der Ausgangsschicht als auch in der verdeckten Schicht, berechnet sind, wird das Netzwerk trainiert. Alle Gewichte g_{ij} , die von Neuron i nach Neuron j führen, werden nach folgender Gleichung iterativ verändert:

$$g_{ij}(t+1) = g_{ij}(t) + \eta \cdot \delta_{pj} \cdot i_{pj} \quad .$$

Dabei ist η die Lernrate, δ_{pj} das Fehlersignal für die Einheit j und i_{pj} das Eingangssignal für Neuron j , kommend von Neuron p . Zuerst werden die Fehlersignale für die Ausgangsneuronen und dann für die verdeckten Neuronen bestimmt. Für jedes Neuron j der Ausgangsschicht bestimmt sich das Fehlersignal zu:

$$\delta_{pj}^{out} = (t_{pj} - z_{pj}) z_{pj} (1 - z_{pj}) \quad .$$

Dabei ist t_{pj} das gewünschte Ausgangssignal für das betreffende Ausgangsneuron und z_{pj} das tatsächlich dort erhaltene Ausgangssignal. Sobald die Fehlersignale für

die Ausgangsneuronen bestimmt sind, wird die verdeckte Schicht des Netzwerkes durch Rückpropagierung des Ausgangsfehlers aktualisiert:

$$\delta_{pj}^{hid} = z_{pj}(1 - z_{pj}) \sum_k \delta_{pk} \cdot g_{kj} \quad .$$

Dabei ist nun z_{pj} das tatsächliche Ausgangssignal des betreffenden Neurons der verdeckten Schicht, δ_{pk} der Fehler des Neurons k in der nächsten Schicht des Netzwerkes und g_{kj} die Gewichtung zwischen dem verdeckten Neuron j und dem Neuron k in der nächsthöheren Schicht. Der Trainingsvorgang wird solange wiederholt, bis der Ausgangsfehler des Netzwerkes unterhalb einer vom Benutzer definierten Toleranzschwelle liegt.

Das mehrschichtige Perceptron (Backpropagation-Netzwerk) hat gegenüber dem einfachen Perceptron den Vorteil, dass sich auch nicht linear trennbare Funktionen erlernen lassen. Allerdings ist der Lernvorgang hier wesentlich langsamer. Wo das einfache Perceptron nur einige wenige Lernschritte benötigt, braucht ein mehrschichtiges Perceptron meist mehrere hundert Durchgänge, um die erwünschte Eingangs-/Ausgangszuordnung zu erlernen. Dies wäre beispielsweise für ein sich in Echtzeit bewegendes Fahrzeug nicht akzeptabel, da es zum Training Hunderte von Fehlversuchen für bestimmte Bewegungen durchführen müsste. Für zeitunkritische Anwendungen wie die Muster- oder Objekterkennung zur Qualitätskontrolle, zur Unterschriftenerkennung u.ä. sind solche Netze jedoch heute im Einsatz.

7.3 Regelungstechnik

Im Kapitel 2 waren bereits die Begriffe des Systems und der Modellbildung behandelt worden. Mit diesen Begriffen und Methoden arbeitet die Regelungstechnik schon seit langem und hat mathematische Methoden zur Beschreibung des Systemverhaltens und für die Auslegung entwickelt. Deshalb werden diese Methoden, die auch schon im Kapitel 2 hätten behandelt werden können, aber an dieser Stelle abgehandelt, da man sie auch in den meisten Standardwerken dort findet.

Die Aufgabe einer Regelung wird in DIN 19226 wie folgt formuliert:

Die Regelung ist ein Vorgang, bei dem der vorgegebene Wert einer Größe fortlaufend durch Eingriff aufgrund von Messungen dieser Größe hergestellt und aufrechterhalten wird. Hierdurch entsteht ein Wirkungsablauf, der sich in einem geschlossenen Kreis (Regelkreis) vollzieht, denn der Vorgang läuft ab aufgrund von Messungen einer Größe, die durch den Vorgang selbst wieder beeinflusst wird. Dieser Wirkungskreis wird Regelkreis genannt. Eine selbsttätige Regelung (im folgenden kurz "Regelung" genannt) liegt vor, wenn dieser Vorgang ohne menschliches Zutun abläuft.

Wie bereits zu Anfang des Kapitels 7 erläutert, wird die **Regelung** eines dynamischen Systems immer dann benutzt, wenn das Systemmodell analoge Größen enthält und die Gefahr besteht, dass das System Störungen oder Parameterschwankungen unterliegt. Nur bei der Verwendung sehr einfacher binärer Komponenten ist das Systemmodell so eindeutig und wenig störanfällig, dass man das Prinzip der **Steuerung** anwenden kann. Der Unterschied zwischen Steuerung und Regelung war bereits in Bild 7-7 grafisch dargestellt worden.

Wie die Modellbildung von Systemen erfolgt, haben wir in den vorherigen Kapiteln bereits ausführlich behandelt.

Im Zusammenhang mit der Regelungstechnik kommt der Modellbildung eine große Bedeutung zu. Häufig tritt der Fall auf, dass das mathematische Modell der Strecke unbekannt ist, was eine sinnvolle Auslegung der Regeleinrichtung (Regler) unmöglich macht. Hier müssen Methoden zur Streckenidentifikation bereitgestellt werden. Dieses Teilgebiet bezeichnet man als **Systemanalyse**.

Ist das Modell der Strecke bekannt oder identifiziert, so ist der am besten geeignete Regler mit einer entsprechenden optimalen Reglereinstellung auszuwählen. Hier stellt die Regelungstechnik eine Anzahl von **Syntheseverfahren** zur Verfügung. Im folgenden sollen daher die Analyse und Synthese von Regelungen und die gebräuchlichsten dort verwendeten Verfahren und Beschreibungsmethoden behandelt werden.

7.3.1 Beschreibung und Analyse regelungstechnischer Systeme

Schon beim Thema Modellbildung dynamischer Systeme und bei der Behandlung von Schwingungsvorgängen haben wir die Beschreibung im **Zeitbereich** durch Differentialgleichungen kennen gelernt. Wir haben gesehen, dass man einfache technische Systeme wie Feder-Masse-Dämpfer-Systeme oder elektrische Schwingkreise durch lineare Differentialgleichungen mit konstanten Koeffizienten der Form

$$\overset{(n)}{b_n} x_a(t) + \dots + b_1 \dot{x}_a(t) + b_0 x_a(t) = a_0 x_e(t) + \dots + \overset{(m)}{a_m} x_e(t) \quad (7.16)$$

beschreiben kann.

Häufig wird versucht, die Anfangsbedingungen oder den Gültigkeitsbereich des Modells eines Systems so zu vereinfachen, dass die Beschreibung durch derartige DGL's möglich ist. Dies tut man deshalb, weil diese Art der Differentialgleichung der einfachen expliziten Lösung zugänglich ist und weil für lineare Systeme das **Verstärkungsprinzip** und das **Superpositionsprinzip** (s. Kap. 2.3) gelten. Reale Systeme sind in Wirklichkeit in den seltensten Fällen durch ein lineares Übertragungsverhalten gekennzeichnet und viele Systeme sind sogar ausgesprochen nichtlinear.

Regelungen sollen eine Strecke so beeinflussen, dass sie entweder einen vorgegebenen Sollwert oder Arbeitspunkt möglichst genau einhält, oder dass die Strecke einem sich ständig ändernden Sollwert mit möglichst geringen Abweichungen folgt. Soll eine nichtlineare Übertragungsstrecke durch eine Regelung in einem Arbeitspunkt gehalten werden, so ist vor allem das Übertragungsverhalten der Strecke im Bereich dieses Arbeitspunktes von Bedeutung. Betrachtet man aber nur einen kleinen Ausschnitt aus einer nichtlinearen, statischen Kennlinie eines Systems, so kann man häufig eine **Linearisierung** der Kennlinie in diesem Arbeitspunkt vornehmen.

Typische nichtlineare Kennlinien ergeben sich für den Durchfluss durch ein Ventil. In Bild 7-48 ist eine solche Kennlinie dargestellt, wobei die Eingangsgröße x_e dieses Systems der Ventilhub y und die Ausgangsgröße x_a der Durchfluss Q durch das Ventil ist. Die Linearisierung im Arbeitspunkt (x_{e0} , x_{a0}) wird häufig so vorgenommen,

dass man die gekrümmte Kennlinie im Arbeitspunkt durch ihre Tangente annähert. Es ist unmittelbar einleuchtend, dass diese Linearisierung dann auch nur für kleine Schwankungen des Systems um diesen Arbeitspunkt ausreichende Genauigkeit liefert.

Für die Linearisierung kann man die Funktion $x_a = f(x_e)$ im Arbeitspunkt x_{e0} in einer **Taylorreihe** entwickeln und die Reihe nach dem ersten Glied abbrechen:

$$f(X_e) = f(X_{e0}) + \frac{X_e - X_{e0}}{1!} \cdot \frac{d}{dX_e} f(X_e) + \frac{(X_e - X_{e0})^2}{2!} \cdot \frac{d^2}{dX_e^2} f(X_e) + \dots$$

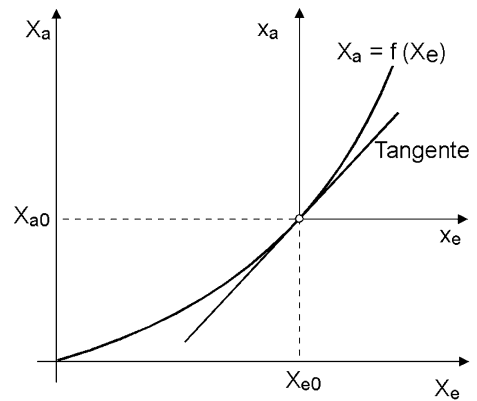


Bild 7-48: Linearisierung einer statischen Ventilkennlinie durch eine Tangente im Arbeitspunkt

Der Abbruch der Reihe nach dem ersten Glied liefert:

$$X_a = X_{a0} + (X_e - X_{e0}) \cdot \left. \frac{dX_a}{dX_e} \right|_{X_{e0}}$$

wobei die Ableitung der Ausgangsgröße nach der Eingangsgröße an der Stelle X_{e0} genommen werden muss.

Stellt man diese Gleichung nochmals wie folgt um

$$(X_a - X_{a0}) = \left. \frac{dX_a}{dX_e} \right|_{X_{e0}} \cdot (X_e - X_{e0}) \quad ,$$

so erhält man die endgültige Form der linearisierten Kennlinie:

$$\Delta X_a = x_a = K(X_{e0}) \cdot \Delta X_e = K(X_{e0}) \cdot x_e \quad . \quad (7.17)$$

Dabei stehen durch Kleinbuchstaben gekennzeichnete Größen für die Abweichungssignale. Die **Übertragungskonstante** K hat eine Maßeinheit von der gilt:

$$[K] = \frac{[x_a]}{[x_e]} \quad .$$

Um diese Vorgehensweise nochmals zu verdeutlichen, möge folgendes Beispiel der Durchflussmessung mit einer Blende dienen [7.9]. Für den Durchfluss durch eine Messblende gilt:

$$Q = c \cdot A \cdot \sqrt{\frac{2\Delta p}{\rho}} = k \cdot \sqrt{\Delta p} \quad .$$

Dies ist ein typischer nichtlinearer Zusammenhang zwischen dem Durchfluss Q und der Druckdifferenz an der Blende Δp (Bild 7-49). Für den Arbeitspunkt gilt:

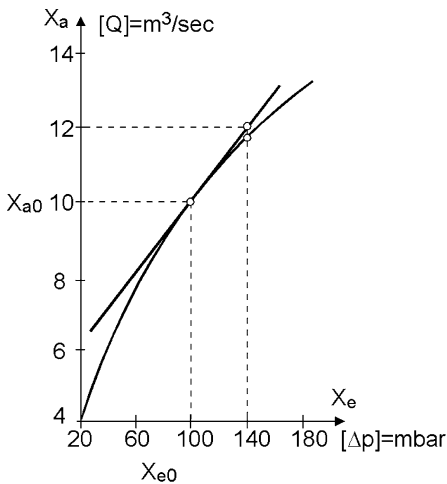


Bild 7-49: Statische Kennlinie für eine Durchflussmessblende

$$X_{a0} = Q_0 = 10 \text{ m}^3/\text{sec} \quad ,$$

$$X_{e0} = \Delta p_0 = 100 \text{ mbar} \quad .$$

Für die Übertragungskonstante K gilt:

$$\left. \frac{dQ}{d\Delta p} \right|_0 = \frac{12 - 10}{140 - 100} = 0,05 \frac{\text{m}^3}{\text{sec} \cdot \text{bar}} \quad .$$

Damit lautet die Gleichung der linearisierten Kennlinie:

$$Q = Q_0 + 0,05(\Delta p - \Delta p_0) \quad .$$

Für einen vom Arbeitspunkt abweichenden Wert von beispielsweise $\Delta p = 140 \text{ mbar}$ ergibt sich dann $Q = 12 \text{ m}^3/\text{sec}$, was vom genauen Wert, der $Q = 11,85 \text{ m}^3/\text{sec}$ beträgt, nur geringfügig abweicht.

7.3.1.1 Laplace-Transformation

Besitzt das zu untersuchende System als mathematisches Modell eine lineare Differentialgleichung vom Typ der Gl. 7.16, so ist sie geschlossen lösbar, und es gibt darauf aufbauend in der Regelungstechnik, aber natürlich auch in anderen Disziplinen, verschiedenartige Möglichkeiten, das Verhalten des Systems zu beschreiben oder darzustellen. Ein Teil dieser Darstellungsweisen wie der **Frequenzgang** oder die **Ortskurve** wurden im Kapitel 4 über Schwingungen bereits andeutungsweise behandelt.

Die in Kapitel 4 bereits behandelte klassische Lösung einer linearen DGL mit konstanten Koeffizienten erfolgt dadurch, dass zuerst die homogene DGL mit Hilfe des Exponentialansatzes $x(t) = C \cdot e^{st}$ gelöst und dadurch **algebraisiert** wird. Danach muss eine partikuläre Lösung der inhomogenen DGL gefunden und abschließend müssen durch Einsetzen der Anfangsbedingungen die Integrationskonstanten bestimmt werden. Dazu sind n lineare Gleichungen mit n Unbekannten zu lösen, wodurch die klassische Methode unter Umständen sehr aufwendig werden kann.

Ebenfalls in Kapitel 4 wurde bereits die Methode der Verwendung einer Funktionaltransformation am Beispiel der Fourier-Transformation behandelt. Bei solchen Transformationen werden Funktionen, die im **Originalbereich** nur schwer zu handhaben sind, in einen **Bildbereich** transformiert. Im Bildbereich sind dann komplexe Rechenoperationen auf einfachere Operationen rückführbar, so dass die Probleme hier viel leichter behandelt oder gelöst werden können. Danach wird das Problem wieder in den Originalbereich zurücktransformiert, wodurch man die Problemlösung im Originalbereich erhält. Als einfaches Beispiel für eine solche Funktionaltransformation war das Logarithmenrechnen bereits erwähnt worden. Der Vorgang des Logarithmenrechnens sei am folgenden Beispiel einer Potenzfunktion dargestellt:

$$y = x^n \xrightarrow[\text{in Bildbereich}]{\text{Transformation}} \log y = n \cdot \log x \xrightarrow[\text{in Originalbereich}]{\text{Rücktransformation}} y = \text{INV}(n \cdot \log x)$$

Man logarithmiert zuerst die Gleichung und transformiert dadurch die Funktion in den Bildbereich. Hier wird die Funktion "einfacher", da nun die Exponierung durch eine Multiplikation ersetzt wird. Nach Durchführung der Multiplikation im Bildbereich erfolgt die Rücktransformation in den Originalbereich durch Delogarithmierung des Produktes, wodurch man den Wert von y erhält.

Mit Hilfe der auf Laplace (Franz. Mathematiker 1749-1827) zurückgehenden **Laplace-Transformation** gelingt es, ohne den oben beschriebenen umständlichen Weg über die allgemeine Lösung mit unbestimmten Konstanten, direkt die Lösung einer DGL zu den gegebenen Anfangsbedingungen zu finden. Da bei den erwähnten linearen DGL's der Originalbereich der **Zeitbereich** ist, benötigt man Funktionaltransformationen zwischen dem Zeitbereich und einem für die Lösung günstigen Bildbereich. Die Laplace-Transformation gehört ebenso wie die Fourier-Transformation zu den Integraltransformationen, für die die allgemeine Transformationsgleichung wie folgt lautet:

$$F(s) = \int_{t_1}^{t_2} f(t) \cdot K(s, t) dt \quad .$$

Wird in diese Gleichung eine Zeitfunktion $f(t)$ eingesetzt, so ergibt sich nach Bestimmung des Integrals eine Zahl, die noch von der Größe s abhängig ist, also eine Funktion $F(s)$. Der Ausdruck $K(s, t)$ heißt Kern der Transformation, durch dessen Beschaffenheit sich die Integraltransformationen voneinander unterscheiden. So wie eine einfache Funktion einer bestimmten Zahl genau eine andere Zahl zuordnet, ordnet die Integraltransformation einer Funktion der Variablen t eine neue Funktion der Variablen s zu. Bei der Laplace-Transformation gilt für den Kern der Transformation

$$K(s, t) = e^{-st} \quad \text{und} \quad s = \delta + i\omega \quad .$$

d. h. s ist eine komplexe Variable. In der Transformierten wird die Zeit t als unabhängige Variable der reellen Funktion $f(t)$ eliminiert und durch die komplexe Variable s ersetzt.

Da es sich bei den mit der Laplace-Transformation zu behandelnden Problemen um physikalische Vorgänge handelt, die man von einem willkürlichen Zeitpunkt $t = 0$ ab betrachten will, kann man die untere Integrationsgrenze grundsätzlich als $t_1 = 0$ annehmen, als obere Grenze wählt man $t_2 = \infty$. Damit lautet das Laplace-Integral dann:

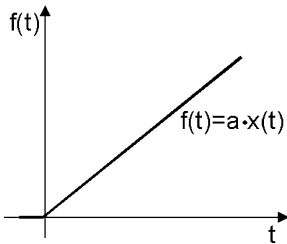
$$F(s) = \int_0^{\infty} f(t) \cdot e^{-st} dt \quad . \quad (7.18)$$

Voraussetzung für die Existenz dieses Integrals einer Zeitfunktion $f(t)$ ist, wie bei der Fourier-Transformation, dass das Integral konvergiert, d. h. der Wert des Integrals ist $< \infty$. Eine abkürzende Schreibweise für das Laplace-Integral ist:

$$\mathcal{L}[f(t)] = F(s) \quad .$$

Für diese Zuordnung (Korrespondenz) ist auch die folgende Schreibweise üblich:

$$f(t) \circ \longrightarrow F(s) \quad .$$



Die Bestimmung der Laplace-Transformierten einer Zeitfunktion "Rampe" (Bild 7-50) soll nun beispielhaft durchgeführt werden. Die Rampenfunktion hat folgende Definition:

$$f(t) = a \cdot x(t) = a \cdot t$$

$$x(t) = t \text{ für } t \geq 0 ; x(t) = 0 \text{ für } t < 0 \quad .$$

Die Laplace-Transformierte wird nach Gl. (7.18) gebildet:

Bild 7-50: Rampenfunktion

$$\mathcal{L}[f(t)] = \int_0^{\infty} t \cdot e^{-st} dt \quad .$$

Durch partielle Integration erhält man:

$$\begin{aligned} \mathcal{L}[f(t)] &= a \cdot \underbrace{t \left(-\frac{1}{s} \right) e^{-st}}_{=0} \Big|_0^{\infty} - a \cdot \int_0^{\infty} \left(-\frac{1}{s} \right) \cdot e^{-st} dt \\ &= \frac{a}{s} \left(-\frac{1}{s} e^{-st} \right) \Big|_0^{\infty} = \frac{a}{s^2} \quad . \end{aligned}$$

Dies kann man abgekürzt schreiben als

$$a \cdot t \circ \longrightarrow \frac{a}{s^2} \quad .$$

Mit der so gefundenen Bildfunktion kann man nun im Bildbereich die erforderlichen Rechenoperationen vornehmen und muss dann die erhaltene Ergebnisfunktion in den Originalbereich (Zeitbereich) rücktransformieren. Die Operation der Rücktransformation wird wie folgt dargestellt:

$$\mathcal{L}^{-1}[F(s)] = f(t) \quad ; \quad F(s) \bullet \longrightarrow f(t) \quad .$$

Dieser Gesamtvorgang für die Lösung linearer DGL's ist in Bild 7-51 nochmals dargestellt.

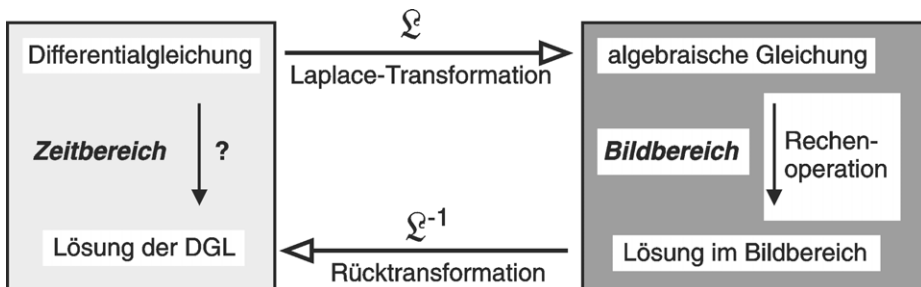


Bild 7-51: Lösung einer linearen Differentialgleichung mit Hilfe der Laplace-Transformation

Häufig brauchen die Vorgänge der Transformation und Rücktransformation nicht rechnerisch durchgeführt werden, weil es in der Literatur umfangreiche Tabellen mit Korrespondenzen von Original- und Bildfunktionen gibt. Die Durchführung von Transformation und Rücktransformation reduziert sich dann auf das Aufsuchen entsprechender Korrespondenzen in den Tabellen, was die Vorgehensweise stark vereinfacht.

Für die Anwendung der Laplace-Transformation zur Lösung von Differentialgleichungen sind noch einige Rechenregeln erforderlich, die im folgenden kurz behandelt werden sollen.

Wendet man die Laplace-Transformation auf die einfachste lineare DGL, die inhomogene DGL 1. Ordnung an, so geht das folgendermaßen vor sich:

$$\mathcal{L}[T \cdot \dot{x}_a + x_a] = \mathcal{L}[a_0 \cdot x_e] \quad (7.19)$$

Aufgrund des für lineare DGL's geltenden Superpositions- und Verstärkungsprinzips lässt sich dies auch wie folgt schreiben:

$$T \cdot \mathcal{L}[\dot{x}_a] + \mathcal{L}[x_a] = a_0 \cdot \mathcal{L}[x_e] \quad ,$$

bzw. durch Einsetzen des Laplace-Integrals

$$T \cdot \int_0^{\infty} \dot{x}_a \cdot e^{-st} dt + \int_0^{\infty} x_a \cdot e^{-st} dt = a_0 \cdot \int_0^{\infty} x_e \cdot e^{-st} dt \quad (7.20)$$

Das Laplace-Integral über die zeitliche Ableitung von $x_a(t)$ findet man wieder durch partielle Integration:

$$\begin{aligned} \int_0^{\infty} \dot{x}_a(t) \cdot e^{-st} dt &= x_a(t) \cdot e^{-st} \Big|_0^{\infty} + s \int_0^{\infty} x_a(t) \cdot e^{-st} dt \\ &= \underbrace{\lim_{t \rightarrow \infty} x_a(t) \cdot e^{-st} - x_a(0)}_{\rightarrow 0 \text{ (Vorausgesetzt das Integral konvergiert)}} + s \cdot \underbrace{\int_0^{\infty} x_a(t) \cdot e^{-st} dt}_{s \cdot \mathcal{L}[x_a(t)]} = \mathcal{L}[\dot{x}_a(t)] \quad . \end{aligned}$$

Demnach ist

$$\mathcal{L}[\dot{x}_a(t)] = s \cdot \mathcal{L}[x_a(t)] - x_a(0) = s \cdot X(s) - x_a(0) = s \cdot F(s) - f(0) \quad .$$

Da laut Voraussetzung nur Zeiten größer oder gleich Null betrachtet werden sollen, handelt es sich bei dem Wert von $x_a(0)$ um den sog. **rechtsseitigen Grenzwert**

$$\lim_{\tau \rightarrow 0^+} x_a(0 + \tau) = x_a(0+) \quad .$$

Dies ist von Bedeutung, wenn die Zeitfunktion, wie beispielsweise bei der Sprungfunktion, zum Zeitpunkt $t = 0$ eine Unstetigkeitsstelle besitzt.

Wie oben hergeleitet, erhält man auch die Transformierten für die höheren Ableitungen:

$$\left. \begin{aligned}
 x(t) &\circ \longrightarrow X(s) \\
 \dot{x}(t) &\circ \longrightarrow s \cdot X(s) - x(0+) \\
 \ddot{x}(t) &\circ \longrightarrow s^2 \cdot X(s) - s \cdot x(0+) - \dot{x}(0+) \\
 &\vdots \\
 x^{(n)}(t) &\circ \longrightarrow s^n \cdot X(s) - s^{n-1} \cdot x(0+) - \dots - x^{(n-1)}(0+)
 \end{aligned} \right\} \quad (7.21)$$

Diese Gleichungen stellen den **Differentiationssatz** der Laplace-Transformation dar. In ihm kommen der Anfangswert $x(0+)$ und seine Ableitungen vor, die normalerweise nicht bekannt sind. Für die zu betrachtenden physikalischen Probleme kann man jedoch annehmen, dass der rechtsseitige Grenzwert gleich dem linksseitigen Grenzwert $x(0-)$ ist, und dass dies auch für seine Ableitungen gilt. Da man bei dynamischen Untersuchungen eines physikalischen Systems immer von dem Anfangszustand $x_a(0-) = x_e(0-) = 0$, sowie vom gleichen Anfangszustand für die Ableitungen ausgehen kann (man spricht von (**verschwindende Anfangsbedingungen**)) kann man für die Anwendung des Differentiationssatzes in der Systemdynamik einige Vereinfachungen vornehmen.

Wendet man ihn auf die DGL (7.16) an, so erhält man.

$$b_n s^n X_a(s) + \dots + b_1 s^1 X_a(s) + b_0 X_a(s) = a_0 X_e(s) + \dots + a_m s^m X_e(s) \quad (7.22)$$

Für die lineare Differentialgleichung 2. Ordnung erhält man daher im Bildbereich:

$$b_2 \ddot{x}_a + b_1 \dot{x}_a + b_0 x_a = a_0 x_e + a_1 \dot{x}_e \circ \longrightarrow (b_2 s^2 + b_1 s + b_0) X_a(s) = (a_0 + a_1 s) X_e(s) \quad .$$

Schreibt man dies in folgender Form auf

$$X_a(s) = \frac{Q(s)}{R(s)} X_e(s) = G(s) \cdot X_e(s) = \frac{a_0 + a_1 s}{b_0 + b_1 s + b_2 s^2} \cdot X_e(s) \quad ,$$

so erhält man die Schreibweise der **Übertragungsfunktion**, einer der wichtigsten Begriffe der linearen Regelungstechnik. Die Übertragungsfunktion ist das Verhältnis der Bildfunktionen der linken und rechten Seite der linearen DGL für verschwindende Anfangsbedingungen:

$$G(s) = \frac{X_a(s)}{X_e(s)} \quad (7.23)$$



Bild 7-52: Typische Aufgabenstellung der Regelungstechnik

Die in Bild 7-52 dargestellte Aufgabe der Ermittlung des Ausgangssignals $x_a(t)$ eines dynamischen Systems mit bekannter Differentialgleichung in Abhängigkeit einer beliebigen Eingangssignals $x_e(t)$ bei verschwindenden Anfangsbedingungen lässt sich mit Hilfe der Übertragungsfunktion folgendermaßen behandeln:

- 1.) Ermittlung der Übertragungsfunktion **G(s)** direkt aus der Differentialgleichung
- 2.) Laplace-Transformation der Eingangs-Zeitfunktion $x_e(t)$, z. B. mit Korrespon-

denztabelle $x_e(t) \circ \longrightarrow X_e(s)$

3.) Bildung von $G(s) \cdot X_e(s) = X_a(s)$

4.) Rücktransformation der Bildfunktion $X_a(s)$ in den Zeitbereich, z. B. mit Korrespondenztabelle $X_a(s) \bullet \longrightarrow x_a(t)$

Für die Operationen der Hin- und Rücktransformation benötigt man so genannte **Korrespondenztabelle** (Tabelle in Bild 7-53) und weitere Rechenregeln (Tabelle in Bild 7-54) der Laplace-Transformation.

$F(s)$	$f(t)$	$F(s)$	$f(t)$
1	$\delta(t)$ (Dirac-Impuls)	$\frac{1}{(1 \pm a \cdot s)^2}$	$\pm \frac{t}{a^2} \cdot e^{\mp t/a}$
$\frac{1}{s}$	$\sigma(t)$ (Sprungfunktion)	$\frac{2}{(s \pm a)^3}$	$t^2 \cdot e^{\mp at}$
$\frac{1}{s^2}$	$\rho(t)$ (Rampenfunktion)	$\frac{n!}{(s \pm a)^{n+1}}$	$t^n \cdot e^{\mp at}, n = 1, 2, 3, \dots$
$\frac{n!}{s^{n+1}}$	$t^n ; n = 1, 2, 3, \dots$	$\frac{1}{(s-a)(s-b)}$	$\frac{e^{at} - e^{bt}}{a-b}$
$\frac{1}{s \pm a}$	$e^{\mp t/a}$	$\frac{1}{(1+a \cdot s)(1+b \cdot s)}$	$\frac{e^{-t/a} - e^{-t/b}}{a-b}$
$\frac{1}{1 \pm a \cdot s}$	$\pm \frac{1}{a} \cdot e^{\mp t/a}$	$\frac{\omega}{s^2 + \omega^2}$	$\sin(\omega \cdot t)$
$\frac{1}{s(s \pm a)}$	$\pm \frac{1}{a}(1 - e^{\mp at})$	$\frac{s}{s^2 + \omega^2}$	$\cos(\omega \cdot t)$
$\frac{1}{s(1 + a \cdot s)}$	$1 - e^{-t/a}$	$\frac{\omega}{(s-a)^2 + \omega^2}$	$e^{at} \cdot \sin(\omega \cdot t)$
$\frac{1}{s^2(1 + a \cdot s)}$	$t - a + a \cdot e^{-t/a}$	$\frac{s-a}{(s-a)^2 + \omega^2}$	$e^{at} \cdot \cos(\omega \cdot t)$
$\frac{1}{(s \pm a)^2}$	$t \cdot e^{\mp at}$	$\frac{\omega}{s^2 - \omega^2}$	$\sinh(\omega \cdot t)$

Bild 7-53: Korrespondenzen für die Laplace-Transformation

Unter diesen Regeln findet man den **Verschiebungssatz**, der für die Behandlung von Vorgängen mit **Totzeit** recht nützlich ist. Darunter versteht man Vorgänge, bei denen erst eine gewisse Zeit t verstreichen muss, bis die Ausgangsgröße auf eine Änderung der Eingangsgröße reagiert. Ein Beispiel hierfür ist das Zahnflankenspiel bei einem Zahnradgetriebe; erst nachdem das antreibende Rad das Spiel durchlaufen hat, beginnt sich das angetriebene Rad zu drehen.

Liegt im Bildbereich ein Produkt bekannter Bildfunktionen $F_1(s) \cdot F_2(s)$ vor, so entspricht dies im Originalbereich der Operation **Faltung** der Zeitfunktionen $f_1(s) * f_2(s)$.

Operation	Originalbereich $f(t) = \mathcal{L}^{-1}[F(s)]$	Bildbereich $\mathcal{L}[f(t)] = F(s)$
Addition im Original- und Bildbereich	$f_1(t) + f_2(t) + \dots$	$F_1(s) + F_2(s) + \dots$
Ähnlichkeit	$f(a \cdot t)$	$\frac{1}{a} \cdot F\left(\frac{s}{a}\right)$
Verschiebung im Originalbereich	$f(t - \tau)$	$e^{-s\tau} \cdot F(s)$
Verschiebung im Bildbereich	$f(t) \cdot e^{\mp at}$	$F(s + a)$
Multiplikation im Bildbereich, entspricht Faltung im Originalbereich	$\int_0^\infty f_1(\tau) \cdot f_2(t - \tau) d\tau =$ $f_1(t) * f_2(t)$ Faltung	$F_1(s) \cdot F_2(s)$
Integration im Originalbereich	$\int_0^\infty f(\tau) d\tau$	$\frac{F(s)}{s}$
Differentiation im Originalbereich	$\frac{d^n f(t)}{dt^n}$	$s^n \cdot F(s)$
Endwertsatz (nur für stabile Vorgänge)	$\lim_{t \rightarrow \infty} x(t)$	$\lim_{s \rightarrow 0} s \cdot X(s) = s \cdot X(s) \Big _{s=0}$

Bild 7-54: Rechenregeln der Laplace-Transformation

Diese Operation ist durch das in der Tabelle angegebene Integral definiert. Mit Hilfe dieser Integralbildung kann man somit solche Produkte von Bildfunktionen, die man nicht in der Korrespondenztabelle hat, rücktransformieren.

Der Bildung der n -ten Ableitung einer Zeitfunktion entspricht, wie oben behandelt, die Multiplikation der zugehörigen Bildfunktion mit s^n . Die Integration einer Zeitfunktion entspricht der Division der Bildfunktion durch s .

Es soll nun noch an einem Beispiel dargestellt werden, wie die Aufgabenstellung aus Bild 7-52 konkret gelöst werden kann. Dazu sollen auf ein System mit der DGL

$$3\dot{x}_a(t) + 1,5x_a(t) = x_e(t)$$

drei verschiedene Eingangs-Zeitfunktionen zum Zeitpunkt $t = 0$ aufgegeben werden. Diese drei einfachen Eingangsfunktionen, die in Bild 7-55 a) dargestellt sind, sind die Sprungfunktion $\sigma(t)$, der Dirac-Impuls $\delta(t)$ und die Rampenfunktion $\rho(t)$. Der erste Schritt, der oben beschriebenen Vorgehensweise, besteht in der Bestimmung der Übertragungsfunktion des Systems, derjenigen Funktion im Bildbereich, die das Übertragungsverhalten vom Ein- zum Ausgang charakterisiert:

$$3\dot{x}_a(t) + 1,5x_a(t) = x_e(t) \longrightarrow 3s \cdot X_a(s) + 1,5X_a(s) = X_e(s)$$

Daraus folgt:

$$(3s + 1,5)X_a(s) = X_e(s) \quad .$$

Nach Gl. (7.23) lautet dann die Übertragungsfunktion:

$$G(s) = \frac{X_a(s)}{X_e(s)} = \frac{1}{3s + 1,5} = \frac{0,6\bar{6}}{2s + 1} \quad .$$

Als Lösung der gestellten Aufgabe wird die Ausgangs-Zeitfunktion $x_a(t)$ in Abhängigkeit von der Eingangs-Zeitfunktion $x_e(t)$ gesucht, die hier die Sprungfunktion sein soll. Entsprechend muss die obige Gleichung umgestellt, und die Korrespondenzfunktion $\mathcal{L}[\sigma(t)] = 1/s$ eingesetzt werden:

$$X_a(s) = G(s) \cdot X_e(s) = \frac{0,6\bar{6}}{2s + 1} \cdot \frac{1}{s} = \frac{0,6\bar{6}}{s(2s + 1)} \quad .$$

Nun kann man durch Rücktransformation von $X_a(s)$ unter Verwendung der Korrespondenztabelle in Bild 7-53 direkt die gesuchte Lösung finden:

$$X_a(s) \bullet \longrightarrow x_a(t) = 0,6\bar{6}(1 - e^{-t/2}) = K_p(1 - e^{-t/T}) \quad .$$

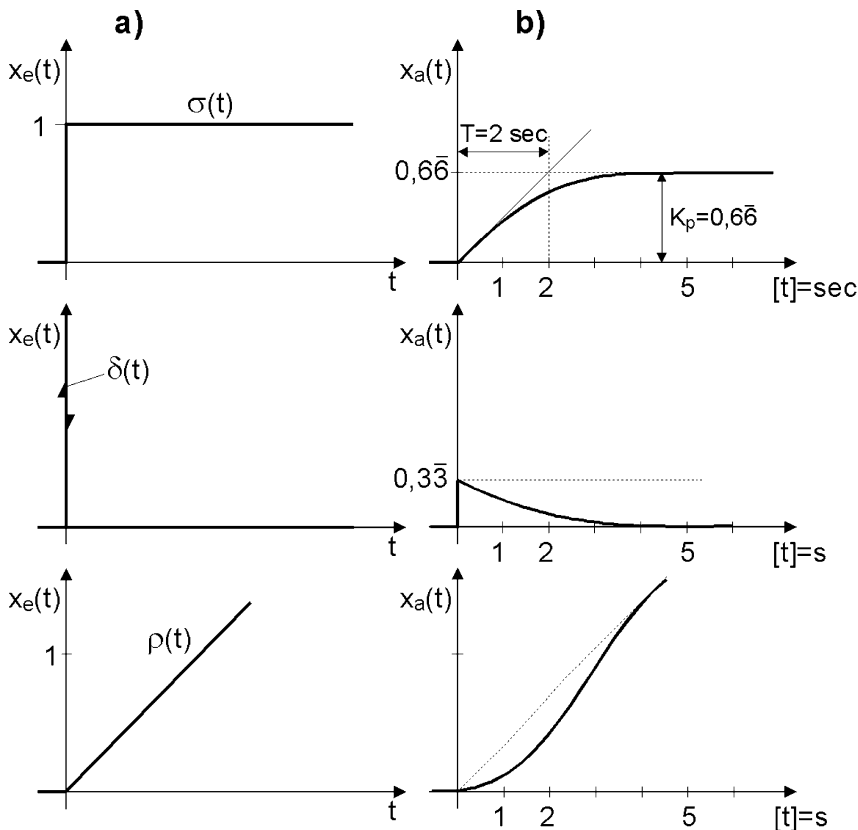


Bild 7-55: Antworten eines einfachen dynamischen Systems auf verschiedene Eingangsfunktionen
a) Eingangsgrößen b) Ausgangsgrößen oder Übergangsfunktion $h(t)$

Diese Ausgangs-Zeitfunktion, die man auch als **Sprungantwort** oder **Übergangsfunktion** $h(t)$ des Systems bezeichnet, ist in Bild 7-55 b) dargestellt. Der Wert T ist dabei die **Zeitkonstante** des Systems und K_p seine **Verstärkung**.

Ist die Eingangs-Zeitfunktion der Delta-Dirac-Impuls $\delta(t)$, dessen Korrespondenzfunktion $\mathcal{L}[\delta(t)] = 1$ beträgt, so erhält man:

$$X_a(s) = G(s) \cdot X_e(s) = \frac{0,6\bar{6}}{2s+1} \cdot 1 = \frac{0,6\bar{6}}{2s+1} \quad .$$

Daraus erhält man durch Rücktransformation:

$$X_a(s) \bullet \longrightarrow x_a(t) = 0,6\bar{6}(0,5 \cdot e^{-t/2}) = 0,3\bar{3} \cdot e^{-t/2} \quad .$$

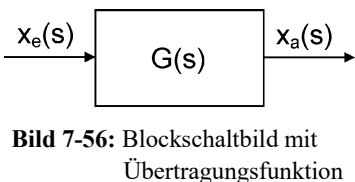
Die Ausgangs-Zeitfunktion, die als Antwort auf einen Eingangs-Dirac-Impuls am Systemausgang auftritt, wird auch als **Gewichtsfunktion** $g(t)$ des Systems bezeichnet. Ihr Verlauf kann für das Beispiel ebenfalls dem Bild 7-55 b) entnommen werden.

Als letztes soll noch die Antwort auf die Rampenfunktion $\rho(t)$ untersucht werden, deren Korrespondenzfunktion $\mathcal{L}[\rho(t)] = 1/s^2$ ist:

$$X_a(s) = G(s) \cdot X_e(s) = \frac{0,6\bar{6}}{2s+1} \cdot \frac{1}{s^2} = \frac{0,6\bar{6}}{s^2(2s+1)} \quad .$$

Die Rücktransformation ergibt:

$$X_a(s) \bullet \longrightarrow x_a(t) = 0,6\bar{6}(t - 2 + 2e^{-t/2}) \quad .$$



Mit Hilfe dieser Methodik ergibt sich nun eine Darstellung eines dynamischen Systems als Blockschaltbild wie in Bild 7-56 dargestellt. Der Block enthält als mathematisches Modell die Übertragungsfunktion, die eine Eingangsgröße im Bildbereich in die Ausgangsgröße des Bildbereichs überführt.

7.3.1.2 Blockschaltbilder

In der Regel sind reale dynamische Systeme komplexerer Natur, so dass sich ihr Blockschaltbild aus einer Anzahl miteinander verbundener Blöcke ergibt. Solche Blöcke können in einigen charakteristischen Verschaltungsarten miteinander verbunden sein. Im folgenden sind solche Zusammenschaltungen dargestellt, und es wird erläutert, wie man die für die Zusammenschaltung resultierende Übertragungsfunktion erhält.

Die beiden einfachsten Zusammenschaltungen sind die **Serien-** oder **Reihenschaltung** und die **Parallelschaltung** zweier Blöcke. Die Serienschaltung zweier Blöcke mit den Einzelübertragungsfunktionen $G_1(s)$ und $G_2(s)$ ist in Bild 7-57 dargestellt. Diese Übertragungsfunktionen lauten:

$$G_1 = \frac{X_{a1}}{X_{e1}} \quad , \quad G_2 = \frac{X_{a2}}{X_{e2}} \quad .$$

Die Gesamtübertragungsfunktion der Serienschaltung lässt sich folgendermaßen herleiten:

$$\begin{aligned} X_a(s) &= X_{a2} = G_2 \cdot X_{e2} = G_2 \cdot X_{a1} \\ X_{a1} &= G_1 \cdot X_{e1} = G_1 \cdot X_e(s) \quad (7.24) \\ X_a(s) &= G_1 \cdot G_2 \cdot X_e(s) \Rightarrow G = G_1 \cdot G_2 \end{aligned}$$

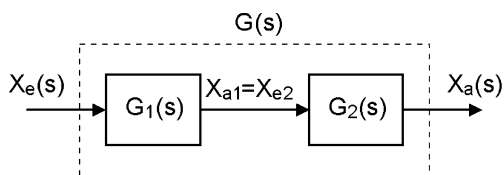


Bild 7-57: Blockschaltbild der Serienschaltung

Verallgemeinert man dies auf eine beliebige Anzahl von in Reihe geschalteter Übertragungsblöcke, so ist die Gesamtübertragungsfunktion das Produkt aller Einzelübertragungsfunktionen:

$$G = \prod_i G_i \quad (7.25)$$

In Bild 7-58 ist die Parallelschaltung zweier Übertragungsblöcke mit den Einzelübertragungsfunktionen $G_1(s)$ und $G_2(s)$ dargestellt. Der in diesem Blockschaltbild verwendete Kreis stellt eine so genannte **Summationsstelle** dar. Diese summiert im Gegensatz zur einfachen Verzweigungsstelle (alle Punkte vor und nach einer Verzweigungsstelle führen das gleiche Signal) die Eingangssignale dem Vorzeichen entsprechend zum Ausgangssignal auf. Dem Bild kann man folgendes entnehmen:

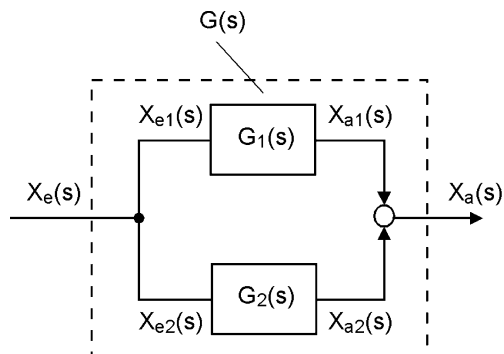


Bild 7-58: Blockschaltbild der Parallelschaltung

$$X_e = X_{e1} = X_{e2} \quad , \quad X_a = X_{a1} + X_{a2} \quad .$$

Entsprechend der Definition der Übertragungsfunktion nach Gl. 7.23 gilt dann:

$$X_a = X_e \cdot G_1 + X_e \cdot G_2 = (G_1 + G_2)X_e \Rightarrow G = \frac{X_a}{X_e} = G_1 + G_2 \quad (7.26)$$

Bei mehreren parallel geschalteten Übertragungsblöcken gilt demnach, dass die Gesamtübertragungsfunktion die Summe aller Einzelübertragungsfunktionen ist:

$$G = \sum_i G_i \quad (7.27)$$

Eine andere wichtige Kombination zweier Übertragungsblöcke ist die so genannte **Rückkopplung** oder **Kreisschaltung**. In dieser Schaltung sind die Blöcke quasi parallel und gleichzeitig in Reihe angeordnet. Das entspricht dem Funktionsprinzip einer Regelung (Bild 7-59). Je nach Vorzeichen an der Summationsstelle, spricht man von **Gegenkopplung** oder **Mitkopplung**. Die Gegenkopplung, die bei Differenzbildung von Eingangs- und Rückführungssignal vorliegt, ist das am häufigsten in der Regelungstechnik angewendete Wirkprinzip.

Für die Gesamtübertragungsfunktion gelten folgende Zusammenhänge:

$$X_{e1} = X_e \mp X_{a2} \quad , \quad X_e = X_{e1} \pm X_{a2} \quad , \quad X_a = X_{a1} = X_{e2} \quad .$$

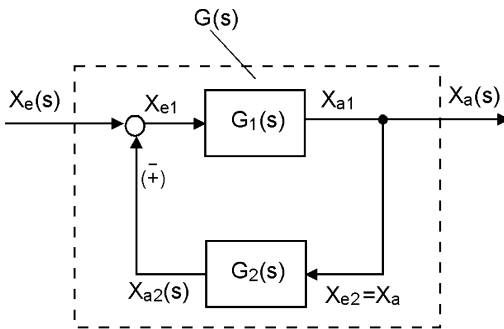


Bild 7-59: Blockschaltbild der Rückkopplung

Entsprechend der Definition der Übertragungsfunktion gilt weiterhin:

$$X_a = X_{e1} \cdot G_1 \quad , \quad X_{a2} = X_a \cdot G_2$$

$$\Rightarrow X_e = \left(\frac{1}{G_1} \pm G_2 \right) X_a$$

$$\Rightarrow G = \frac{G_1}{1 \pm G_1 \cdot G_2} \quad . \quad (7.28)$$

Die Rückkopplung ist nicht nur eines der wichtigsten Wirkprinzipien einer Regelung, sondern sie wird auch in der

Verstärkertechnik benutzt, um gezielt die Verstärkung und auch das dynamische Übertragungsverhalten eines Verstärkers einzustellen. Als Beispiel soll der in der analogen Schaltungstechnik vielfach verwendete **Operationverstärker** dienen, wie er in Bild 5-8 und 5-9 im Kapitel 5 dargestellt ist. Wird dieser ohne Beschaltung als Differenzverstärker betrieben, so ist der lineare Verstärkungsbereich wegen der hohen Spannungsverstärkung $K_V = 10^3 - 10^7$ eng begrenzt und der Verstärker gerät schon bei sehr kleinen Eingangssignalen in die Übersteuerung.

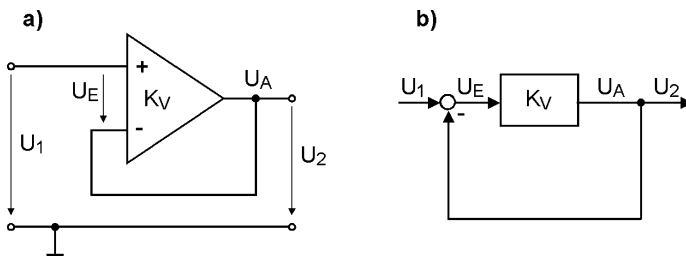


Bild 7-60: Rückgekoppelter Operationsverstärker (Spannungsfolger) a) Stromlaufplan b) Blockschaltbild

Fügt man nun wie in Bild 7-60 a) dargestellt eine Rückkopplung vom Verstärkerausgang auf den invertierenden Eingang ein, so ergibt sich für diese Anordnung ein Blockschaltbild entsprechend Bild 7-60 b). Anhand des Blockschaltbildes kann man unter Verwendung

von Gl. (7.28) die Gesamtübertragungsfunktion herleiten. Dabei ist die Übertragungsfunktion der Rückführung $G_2 = 1$:

$$G = \frac{K_V}{1 + K_V \cdot 1} = \frac{1}{1 + \frac{1}{K_V}} \quad .$$

Da für den Operationsverstärker definitionsgemäß $K_V \gg 1$ ist, kann man den Term $1/K_V$ gleich Null setzen, so dass gilt:

$$G = 1 \quad .$$

Durch den Einfluss der Rückkopplung hat der Verstärker nun die Schleifenverstärkung $K_S = 1$, d. h. die Ausgangsspannung U_2 folgt unter gewissen dynamischen Einschränkungen exakt dem Verlauf der Eingangsspannung U_1 . Aus diesem Grund wird die Schaltung auch als **Spannungsfolger** bezeichnet. Die Schaltung nimmt also keinen Einfluss auf die Signalspannung, sondern dient vor allem der Widerstands- oder Leistungsanpassung, da in der Regel ihr Eingangswiderstand groß

und ihr Ausgangswiderstand klein ist.

Will man den Verstärker für eine von "1" abweichende Schleifenverstärkung nutzen, so kann man die in Bild 7-61 a) dargestellte Schaltung verwenden. Hier wird nicht die volle Ausgangsspannung, sondern nur die am Spannungsteiler aus den Widerständen R_1 und R_2 sich einstellende Spannung der Größe

$$U_2 \frac{R_2}{R_1 + R_2} = U_2 \cdot K_R$$

rückgekoppelt. Der im Teilbild 7-61 b) in der Rückführung eingezeichnete Block stellt praktisch eine Abschwächung der Rückführungsgröße um den Faktor K_R dar. Als Gesamtübertragungsfunktion ergibt sich damit:

$$G = \frac{K_V}{1 + K_V \cdot K_R} = \frac{1}{\frac{1}{K_V} + K_R} = \frac{1}{K_R} \quad \text{für } K_V \gg 1.$$

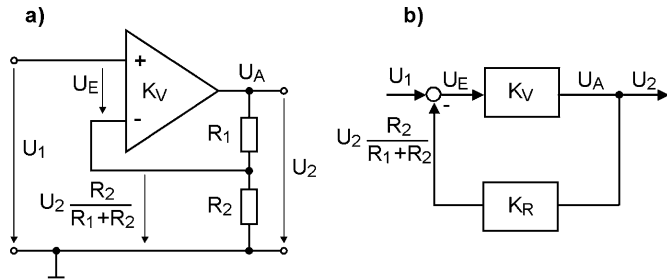


Bild 7-61: Rückgekoppelter Operationsverstärker mit einstellbarer Spannungsverstärkung

Da der Wert von K_R kleiner als 1 ist, bekommt man so praktisch einen Verstärker mit einer Schleifenverstärkung K_S , die größer als 1 ist. Verwendet man dann in der Rückführung noch Schaltelemente, die nicht nur eine lineare Verstärkungsbeeinflussung, sondern auch eine Beeinflussung des dynamischen Verhaltens bewirken (Kondensator oder Spule mit frequenzabhängigem Widerstand), so kann man mit solchen Schaltungen beliebiges dynamisches Verhalten erzeugen. Seine Eigenschaften hängen dann im wesentlichen nur von der Art der Rückführung und nicht vom Verstärker im Vorwärtszweig ab.

Hat man nun ein reales System, dessen Übertragungsfunktion und Differentialgleichung noch nicht bekannt sind, so kann man durch Aufstellung eines Blockschaltbildes und durch schrittweise Vereinfachung dieses Blockschaltbildes ebenfalls die Gesamtübertragungsfunktion ableiten. Dies soll im folgenden am Beispiel des bereits bekannten Einmassenschwingers gezeigt werden. Zwar kennen wir seine Differentialgleichung und könnten daher die Übertragungsfunktion sofort anschreiben, aber hier soll bewusst zur Verdeutlichung der Zusammenhänge dieser Weg beschritten werden.

Dazu wird nochmals das Schema des Einmassenschwingers aus Bild 4-2 herangezogen. Durch eine äußere Kraft $F(t)$ wird die Masse m um den Weg x ausgelenkt. Dieser Beschleunigungskraft wirken die Federkraft F_k und die Dämpferkraft F_d entgegen, so dass zur Beschleunigung der Masse eine Beschleunigungskraft

$$F_b = F(t) - F_k - F_d = m \cdot \ddot{x}$$

verbleibt. Dies ist in Bild 7-62 a) als Summationsstelle dargestellt. Die Beschleunigungskraft beschleunigt die Masse m mit der Beschleunigung $\ddot{x}$ auf die Geschwin-

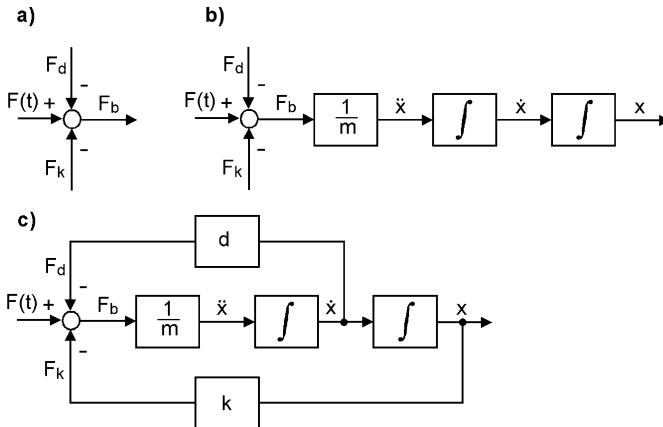


Bild 7-62: Entwicklung des Blockschaltbildes für den Einmassenschwinger im Zeitbereich a) Kräftegleichgewicht b) Ermittlung des Weges aus der Beschleunigung durch Integration c) Gesamtschaltbild

digkeit $\dot{x}$. Die Größen x , $\dot{x}$ und $\ddot{x}$ sind wichtige Zustandsgrößen des Systems. Die Kräftesummi-
on aus Bild 7-62 a) liefert den Wert $m \cdot \ddot{x}$, so dass man, wie in Bild 7-62 b) dargestellt, nach einem Divisionsblock mit der Übertragungsfunktion $1/m$ die Beschleunigung $\ddot{x}$ erhält. Aus dieser ergibt sich wiederum durch einen Block, der $\ddot{x}$ integriert, die Geschwindigkeit $\dot{x}$ und durch einen weiteren Integrationsblock schließlich der Weg x . Für den Ein-

massenschwinger sind die beiden Kräfte F_k und F_d bereits bekannt:

$$F_k = k \cdot x \quad , \quad F_d = d \cdot \dot{x} \quad .$$

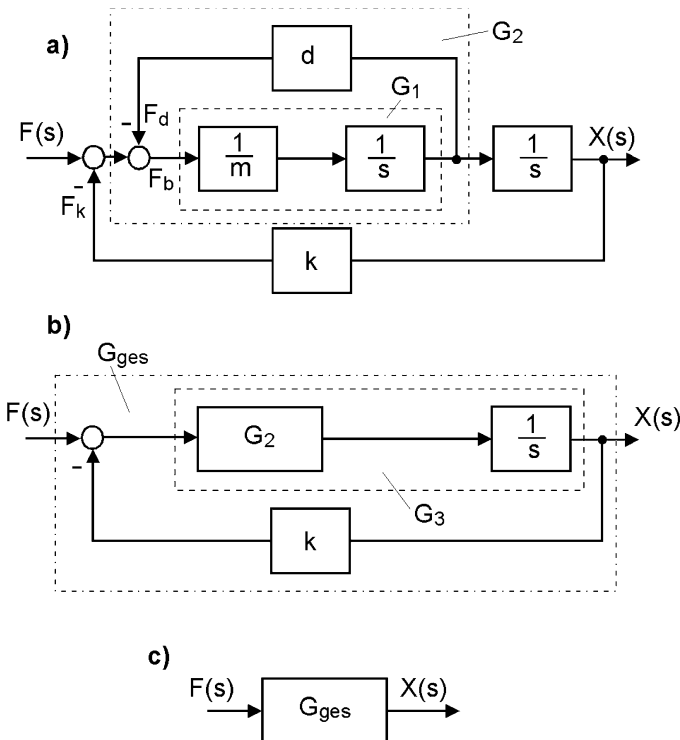


Bild 7-63: Entwicklung des Blockschaltbildes für den Einmassenschwinger im Bildbereich a) Gesamtschaltbild b) Vereinfachungen c) Gesamtübertragungsfunktion

Daher kann man wie in Bild 7-62 c) dargestellt, diese Kräfte durch je einen Multiplikationsblock aus dem Weg bzw. der Geschwindigkeit erhalten. Damit ist der gesamte Wirkablauf im Zeitbereich als Blockschaltbild dargestellt. Dieses würde man, wie wir noch im Kapitel 8 sehen werden, auch als Grundlage für die Simulation des Einmassenschwingers auf einem Digitalrechner benutzen.

Im Moment soll aber die Gesamtübertragungsfunktion anhand des Blockschaltbildes ermittelt werden. Dazu muss in alle Blöcke die Einzelübertragungsfunktionen im Bildbereich eingetragen werden. Multiplikation und Division ändern sich beim

Übergang vom Zeit- in den Bildbereich der Laplace-Transformation nicht. Die Integration wird im Bildbereich durch die Bildfunktion $1/s$ ersetzt, die Summationsstelle wird in zwei Summationsstellen aufgespalten. Dadurch ergibt sich ein Blockschaltbild entsprechend Bild 7-63 a). In diesem Teilbild sind zwei Stufen der Vereinfachung eingezeichnet. Die beiden Blöcke im gestrichelten Rahmen sind in Reihe geschaltet, so dass ihre Gesamtübertragungsfunktion lautet:

$$G_1 = \frac{1}{m \cdot s}$$

Der Block mit der Übertragungsfunktion G_1 und der Block mit dem Faktor d sind als Rückkopplung geschaltet, deren Gesamtübertragungsfunktion folgendermaßen lautet:

$$G_2 = \frac{G_1}{1 + d \cdot G_1} = \frac{1}{m \cdot s + d} \quad .$$

Nach diesen Vereinfachungen hat man ein Blockschaltbild entsprechend Teilbild 7-63 b). Hier kann man wiederum die Serienschaltung von G_2 und dem Block mit der Übertragungsfunktion $1/s$ als Produkt zu G_3 zusammenfassen:

$$G_3 = G_2 \cdot \frac{1}{s} = \frac{1}{m \cdot s + d} \cdot \frac{1}{s} = \frac{1}{m \cdot s^2 + ds} \quad .$$

Danach liegt nur noch eine Rückkoppelschaltung aus G_3 und dem Block mit dem Faktor k vor, deren Gesamtübertragungsfunktion lautet:

$$G_{ges} = \frac{1}{\frac{1}{G_3} + k} = \frac{1}{m \cdot s^2 + ds + k} = \frac{X(s)}{F(s)} = \frac{X_a(s)}{X_e(s)} \quad .$$

Da dies die Gesamtübertragungsfunktion des Systems "Einmassenschwinger" ist, lautet die Laplace-Transformierte der Differentialgleichung:

$$(m \cdot s^2 + ds + k)X_a(s) = X_e(s) \quad .$$

Hieraus kann man sofort die DGL ablesen:

$$m\ddot{x}_a + d\dot{x}_a + kx_a = x_e \quad \text{oder} \quad m\ddot{x}(t) + d\dot{x}(t) + kx(t) = F(t) \quad .$$

Dies stimmt wie zu erwarten mit der bereits für das System aus einer Kräftebilanz ermittelten DGL überein.

Die Darstellung als Blockschaltbild ist besonders gut für die Darstellung der Gesamtstruktur eines Regelkreises aus Strecke und Regler geeignet. In Bild 7-64 a) ist ein solches Blockschaltbild mit den beiden Eingangsgrößen **Sollwert** W und **Störgröße** Z dargestellt. Da für die Beschreibung des Einflusses des Sollwertes auf die Ausgangsgröße meist eine andere Differentialgleichung (mathematisches Modell) als für die Beschreibung des Einflusses der Störgröße gilt, wird die letztere dem Regelkreis über einen eigenen Übertragungsblock mit der **Stör-Übertragungsfunktion** G_{Sz} "zugeführt" (die Störung wirkt auf die Strecke ein). Die Störübertragungsfunktion beschreibt das **Störverhalten** der Strecke, d. h. die Ein-

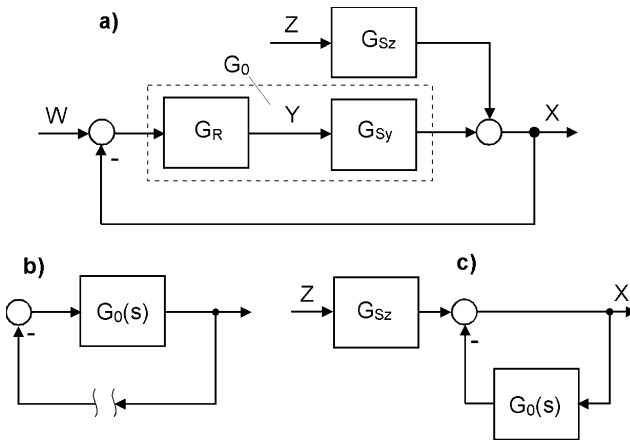


Bild 7-64: Blockschaltbild des Regelkreises

- a) Stör- und Führungsübertragungsfunktion
- b) Übertragungsfunktion des offenen Regelkreises
- c) Stör-Übertragungsfunktion

flüsse von Störungen auf die Strecke. Das **Stellverhalten** der Strecke, d. h. die Einflüsse von Änderungen der Führungs- oder Stellgröße, wird durch die **Führungs-Übertragungsfunktion** G_{Sy} beschrieben.

Das Produkt der in Serie geschalteten Übertragungsfunktionen von Regler und Stellverhalten der Strecke wird folgendermaßen zusammengefasst:

$$G_0(s) = G_R(s) \cdot G_{Sy}(s) \quad .$$

Die Übertragungsfunktion $G_0(s)$ wird als Übertragungsfunktion des **offenen** oder **aufgeschnittenen Regelkreises** bezeichnet. Sie beschreibt das Übertragungsverhalten des Systems aus Regler und Strecke für $W = Z = 0$, wenn man sich die Rückkopplung aufgeschnitten denkt (Bild 7-64 b).

Aus Bild 7-64 a) kann man nun die Führungs- und Stör-Übertragungsfunktion ableiten, für die jeweils die andere Einflussgröße gleich Null gesetzt wird, d. h. Führungsverhalten $Z = 0$ und Störverhalten $W = 0$. Unter Verwendung der Übertragungsfunktion des offenen Regelkreises lautet die Führungs-Übertragungsfunktion:

$$G_W(s) = \frac{X(s)}{W(s)} = \frac{G_0(s)}{G_0(s) + 1} = G(s) \quad . \quad (7.29)$$

Da die Führungs-Übertragungsfunktion nur vom Übertragungsverhalten von Regler und Strecke abhängt, wird diese oft auch einfach als Übertragungsfunktion $G(s)$ des Regelkreises bezeichnet. Die Stör-Übertragungsfunktion lautet:

$$G_Z(s) = \frac{X(s)}{Z(s)} = \frac{G_{Sz}(s)}{G_0(s) + 1} \quad . \quad (7.30)$$

Die letzte Gleichung erhält man leicht, wenn man das in Bild 7-64 c) dargestellte Blockschaltbild verwendet, das man aus dem Blockschaltbild 7-64 a) durch Umzeichnen für den Fall $W = 0$ erhält.

7.3.1.3 Frequenzgang und Ortskurve

Im Kapitel 4 haben wir zur Lösung von linearen Differentialgleichungen mit konstanten Koeffizienten den klassischen Lösungsansatz

$$x(t) = A \cdot e^{st} + B \cdot e^{-st}$$

mit $s = i \cdot \omega$ verwendet. Setzt man nun in der Übertragungsfunktion $G(s)$ den Wert

des Laplace-Operators $s = i \cdot \omega$ ein, so erhält man den Frequenzgang $F(i\omega)$ des Systems. Es gilt also :

$$F(i\omega) = G(s) \Big|_{s=i\omega} \quad (7.31)$$

Dies ist wie die Übertragungsfunktion eine komplexe Funktion, für die gilt

$$F(i\omega) = \Re\{F\} + i \cdot \Im\{F\} = A(\omega) \cdot e^{i\alpha(\omega)} \quad , \quad (7.32)$$

wobei $A(\omega)$ die frequenzabhängige Amplitude und $\alpha(\omega)$ die frequenzabhängige Phasenverschiebung ist. Es ist leicht erkennbar, dass diese Art der Darstellung sich vor allem für Aussagen über ein System im Zeitbereich bei Eigenschwingvorgängen oder bei erzwungenen Schwingungen eignet. Die Darstellung im Bildbereich mit Hilfe der Übertragungsfunktion hingegen ist für beliebige Anregungssignalformen oder Eigenvorgänge geeignet.

Es bietet sich nun an, die komplexe Funktion $F(i\omega)$ in der komplexen Zahlenebene darzustellen. Dazu muss man den Parameter ω von Null bis zu hohen Werten variieren und dann entweder Real- und Imaginärteil, oder Amplitude und Phasenwinkel als Koordinaten in der komplexen Ebene benutzen (Bild 7-65 b).

Diese Vorgehensweise soll anhand des Beispiels eines elektrischen Schwingkreises, der schon in Kapitel 1 behandelt wurde, dargestellt werden. Der Schwingkreis, der in Bild 7-65 a) dargestellt ist, hat die Differentialgleichung

$$LC\ddot{u}_a(t) + RC\dot{u}_a(t) + u_a(t) = u_e(t)$$

Daraus kann man sofort die Übertragungsfunktion ableiten:

$$G(s) = \frac{1}{LCs^2 + RCs + 1} \quad .$$

Ersetzt man nun s durch $i \cdot \omega$ so erhält man den Frequenzgang:

$$\begin{aligned} F(i\omega) &= \frac{1}{LC(i\omega)^2 + RC(i\omega) + 1} \\ &= \frac{1}{(1 - LC\omega^2) + RCi\omega} \quad . \end{aligned}$$

Um den Frequenzgang in der komplexen Ebene darzustellen, muss man nun Real- und Imaginärteil ermitteln und deren Zahlenwert für verschiedene Werte des Parameters ω bestimmen. Zur Ermittlung von Real- und Imaginärteil wird der Bruch zuerst mit dem konjugiert komplexen Wert des Nenners erweitert:

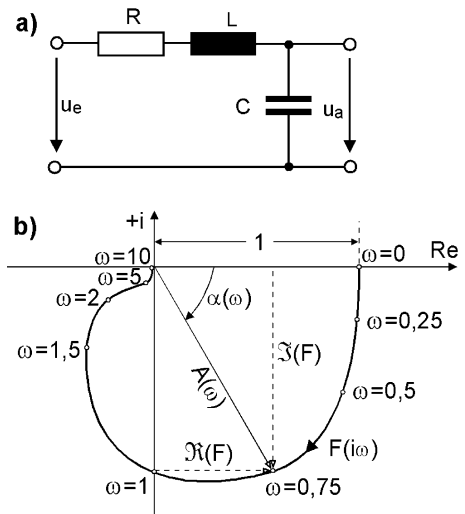


Bild 7-65: Ortskurve eines Schwingkreises
a) Schaltbild b) Ortskurve

$$F(i\omega) = \frac{(1 - LC\omega^2) - RCi\omega}{(1 - LC\omega^2)^2 + (RC\omega)^2} \cdot$$

Damit haben Real- und Imaginärteil des Frequenzgangs folgende Werte:

$$\begin{aligned} \Re\{F(i\omega)\} &= \frac{(1 - LC\omega^2)}{(1 - LC\omega^2)^2 + (RC\omega)^2} \\ &= \frac{1 - LC\omega^2}{L^2C^2\omega^4 + (R^2C^2 - 2LC)\omega^2 + 1} \cdot \\ \Im\{F(i\omega)\} &= \frac{-RC\omega}{L^2C^2\omega^4 + (R^2C^2 - 2LC)\omega^2 + 1} \cdot \end{aligned}$$

Um die Beträge von Real- und Imaginärteil beispielhaft ausrechnen zu können, müssen die Werte von L , C und R festgelegt werden:

$$L = 10^2 \text{ H}, \quad C = 10^{-2} \text{ F}, \quad R = 10^2 \Omega$$

Mit diesen Werten betragen Real- und Imaginärteil:

$$\begin{aligned} \Re\{F(i\omega)\} &= \frac{1 - \omega^2}{\omega^4 + (1 - 2)\omega^2 + 1} = \frac{1 - \omega^2}{\omega^4 - \omega^2 + 1} \cdot \\ \Im\{F(i\omega)\} &= \frac{-\omega}{\omega^4 - \omega^2 + 1} \cdot \end{aligned}$$

Die folgende Tabelle zeigt die für verschiedene Werte von ω ermittelten Real- und Imaginärteile :

ω	0	0,25	0,5	0,75	1	1,5	2	5	10
$\Re\{F\}$	1	0,996	0,923	0,58	0	-0,326	-0,230	-0,04	-0,01
$\Im\{F\}$	0	-0,266	-0,615	-0,995	-1	-0,393	-0,154	-0,08	-0,001

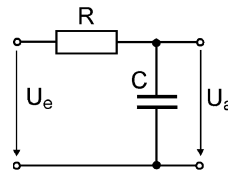
Mit Hilfe dieser Funktionswerte kann man nun einen Funktionsverlauf der komplexen Funktion **Frequenzgang** in der komplexen Zahlenebene zeichnen (Bild 7-65 b)). Der Funktionsverlauf stellt praktisch die Orte in der komplexen Ebene dar, die man erhalten würde, wenn man den Eingang eines Systems mit Sinusschwingungen konstanter Amplitude und veränderlicher Frequenz beaufschlagen würde und das Ausgangssignal dann in der komplexen Ebene aufträgt. Diese Ausgangssignale sind Ortsvektoren, die vom Nullpunkt zu einer Stelle dieser als **Ortskurve** bezeichneten Funktion zeigen und durch den Betrag des Frequenzgangs $A(\omega)$ (Ausgangsamplitude) und den Phasenwinkel $\alpha(\omega)$ (Phasenverschiebung zwischen Ein- und Ausgangsschwingung) gekennzeichnet sind (Bild 7-65 b)). Das gleiche Ergebnis erhält man natürlich, wie im obigen Beispiel gezeigt, wenn man die entsprechenden Punkte durch ihren Real- und Imaginärteil bestimmt. Da man mit Hilfe der Ortskurve die Übertragungseigenschaften eines Systems sehr gut beschreiben kann, sollen im nächsten Abschnitt verschiedenartige Typen von Übertragungssystemen anhand der Ortskurve charakterisiert werden.

7.3.1.4 Verschiedenartige Übertragungssysteme

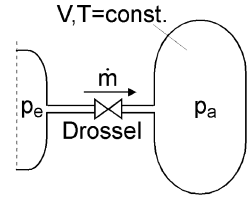
Die einfachsten dynamischen Systeme werden durch Differentialgleichungen 1. Ordnung beschrieben. Eine solche DGL lautet beispielsweise:

$$b_1 \dot{x}_a + b_0 x_a = a_0 x_e \quad (7.33)$$

Typische Systeme, die sich mit einer solchen DGL modellieren lassen, sind der Ladevorgang eines Widerstands - Kondensator - Systems mit einer Eingangsspannung U_e auf eine Ausgangsspannung U_a , oder als vergleichbares mechanisches System das Aufpumpen eines Volumens mit Innendruck p_a auf einen höheren Außendruck p_e . In Bild 7-66 sind solche Systeme mit ihren Differentialgleichungen dargestellt. Diese Differentialgleichungen findet man, wie im folgenden kurz aufgezeigt, aus bekannten physikalischen Gesetzen.



$$RC \dot{U}_a + U_a = U_e$$



$$\frac{V}{RTk} \dot{p}_a + p_a = p_e$$

Bild 7-66: Dynamische Systeme mit DGL 1. Ordnung

Ladevorgang Kondensator:

Ladestrom: $I = C \cdot \dot{U}_a$ Ohmsches Gesetz für Ladewiderstand: $I = \frac{U_e - U_a}{R}$

Resultierende DGL: $RC \dot{U}_a + U_a = U_e$

Aufpumpen Volumen:

Massenstrom durch Drosselstelle $\dot{m} = k(p_e - p_a)$

k : Drosselwiderstand

Gasgleichung für Volumen: $p_a \cdot V = mRT$ abgeleitet $\dot{p}_a \cdot V = \dot{m}RT$

R : Gaskonstante T : Temperatur V : Volumen

Resultierende DGL: $V \cdot \dot{p}_a = RTk(p_e - p_a)$ bzw. $\frac{V}{RTk} \cdot \dot{p}_a + p_a = p_e$

Die Differentialgleichungen beider Systeme haben die gleiche Struktur, die allgemeingültig wie folgt aufgeschrieben werden kann:

$$T \cdot \dot{x}_a + x_a = x_e \quad (7.34)$$

Der Koeffizient T wird als **Zeitkonstante** bezeichnet, da durch sie wichtige Informationen über das zeitliche Verhalten des Systems gegeben sind. Die Übertragungsfunktion und der Frequenzgang eines solchen **Übertragungsgliedes 1. Ordnung** lauten :

$$G(s) = \frac{X_a(s)}{X_e(s)} = \frac{1}{1 + T \cdot s}$$

$$F(i\omega) = \frac{1}{1 + T \cdot i\omega} = \Re\{F\} + i \cdot \Im\{F\} = \frac{1}{1 + (T\omega)^2} - i \frac{T\omega}{1 + (T\omega)^2}.$$

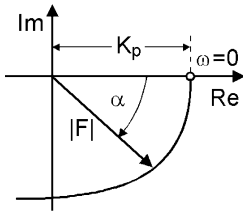


Bild 7-67: Ortskurve von Verzögerungsgliedern 1. Ordnung

Am Frequenzgang derartiger Übertragungsglieder erkennt man, dass sich bei kleinen Frequenzen immer positive Realteile und negative Imaginärteile ergeben. Der Funktionswert liegt dann im 4. Quadranten der komplexen Zahlenebene (Bild 7-67). Daher ist der Phasenwinkel α negativ und die Ausgangsgröße eilt der Eingangsgröße hinterher. Man spricht daher auch von Verzögerungsgliedern und hier im speziellen Fall von **Verzögerungsgliedern 1. Ordnung**.

Für den Fall in Gl. (7.34) ist der Koeffizient b_0 des allgemeinen Falls nach Gl. (7.33) mit $b_0 = 1$ angenommen worden.

Meist wird dieser Koeffizient von 1 verschieden sein, so dass für den Frequenzgang bei $\omega = 0$ (statischer Fall) gilt:

$$F(i\omega) = \frac{1}{b_0} = K_p = \text{const.}$$

	P-Glied	PT ₁ -Glied	PT ₂ -Glied
Übertragungsfunktion G(s)	K_p	$\frac{K_p}{1 + T s}$	$\frac{K_p}{1 + 2DT s + T^2 s^2}$
Pole von G(s)			
Frequenzgang F(iω)	K_p	$\frac{K_p}{1 + T i \omega}$	$\frac{K_p}{1 + 2DT i \omega + T^2 (i\omega)^2}$
Ortskurve von F(iω)			
Übergangsfunktion h(t)			

Bild 7-68: Übertragungsfunktion, Pole der Übertragungsfunktion, Frequenzgang, Ortskurve und Übergangsfunktion von Proportionalgliedern ohne und mit Verzögerung

Der Faktor K_p wird als **Verstärkungsfaktor** oder **Verstärkung** bezeichnet. Da der Zusammenhang zwischen Ausgangs- und Eingangsgröße dann durch folgende Gleichung beschrieben wird,

$$x_a(t) = K_p \cdot x_e(t) \quad ,$$

bezeichnet man dieses Übertragungsverhalten als **proportionales** Verhalten oder kurz **P-Verhalten**.

Verzögerungsglieder mit P-Verhalten haben die Eigenschaft, dass die Ausgangsgröße nach einer Änderung der Eingangsgröße mit einem bestimmten Übergangsverhalten einem neuen Beharrungswert zustreben. Sie werden daher als Systeme oder **Strecken mit Ausgleich** bezeichnet. Dies sieht man deutlich an den Übergangsfunktionen $h(t)$ (Sprungantworten) der proportional wirkenden Systeme in Bild 7-68. Bei diesen Systemen handelt es sich um ein reines P-Glied, sowie um proportionale Verzögerungsglieder 1. und 2. Ordnung. Wie wir bereits früher festgestellt haben, kann es reine Proportionalglieder als physikalische Systeme nicht geben, da ein gewisses Verzögerungsverhalten durch innere Energiespeicher immer gegeben ist. Die Ortskurven proportional wirkender Verzögerungsglieder beginnen immer für $\omega = 0$ bei $\Re\{F\} = K_p$ auf der reellen Achse und führen für zunehmendes ω durch so viele Quadranten, wie es der Ordnung des Verzögerungsgliedes entspricht. Ebenfalls in Bild 7-68 sind die Polstellen der Übertragungsfunktion dargestellt, also Werte von s , für die die Übertragungsfunktion gegen ∞ strebt. Über die Anordnung der Pole in der komplexen Ebene kann man auch wichtige Aussagen über ein System gewinnen.

Beim Verzögerungsglied 1. Ordnung oder kurz **PT₁-Glied** ist die Bedeutung der Zeitkonstante T in der Darstellung der Übergangsfunktion eingetragen. Dies ist die Zeit, nach der die Tangente an die Sprungantwort zum Zeitpunkt $t = 0$ den Amplitudenendwert für $t = \infty$ schneidet. Zu diesem Zeitpunkt beträgt der Amplitudenwert 63,2% des Endwertes. Dies war schon im Beispiel in Bild 7-55 dargestellt worden und beruht auf der Lösung der DGL für ein sprungförmiges Eingangssignal:

$$x_a(t) = K_p(1 - e^{-t/T}) \quad x_a(t = T) = K_p(1 - e^{-1}) = K_p(1 - 0,368) = K_p \cdot 0,632 \quad .$$

Verzögerungsglieder 2. Ordnung haben wir bereits als gedämpfte Feder-/Masse-Systeme oder als elektrische RLC-Systeme in Kapitel 4 kennen gelernt. Dort wurde festgestellt, dass bei Auslenkungen der Masse durch eine Erregerkraft, je nach vorhandenem Lehr'schen Dämpfungsmaß, das System eine gedämpfte Schwingung ausführt. Dies sieht man sowohl an der Übergangsfunktion, in der das Schwingverhalten nach einem Eingangssprung angedeutet ist, wie auch an der Polstellenkonfiguration. Bei Verzögerungsgliedern 2. Ordnung gibt es einen konjugiert komplexen Pol und somit eine konjugiert komplexe Lösung des charakteristischen Polynoms der DGL. Dies hatten wir bereits in Kapitel 4 als Voraussetzung für schwingfähiges Verhalten festgestellt.

Übertragungsglieder können aber nicht nur proportional wirken. Im Bild 7-69 ist beispielsweise das Blockschaltbild einer **Lageregelung** dargestellt. Die Strecke besteht dabei aus der Hintereinanderschaltung eines Gleichstrommotors und eines Linearantriebs mit einem Spindel-/ Mutter- System. Der Gleichstrommotor hat das

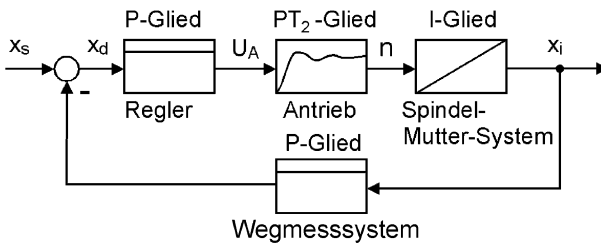


Bild 7-69: Blockschaltbild eines Lageregelkreises

umsetzt. Dieser Istweg wird über ein Wegmesssystem mit P-Verhalten gemessen, um zur Bildung der Eingangsgröße für den Regler dem Soll-Istwertvergleich zugeführt (rückgekoppelt) zu werden. Sind Soll- und Istweg gleich, so wird die Eingangsgröße x_d in den Regler Null, und der Motor bleibt stehen.

Hier soll nun besonders das Übertragungsverhalten des Linearantriebs betrachtet werden. Seine Wirkung wird durch folgende DGL beschrieben:

$$x_i = K_I \int_0^t n dt \quad \text{bzw.} \quad x_a = K_I \int_0^t x_e dt \quad \text{oder} \quad \dot{x}_a = K_I \cdot x_e \quad (7.35)$$

Dies beschreibt ein Zeitverhalten, bei dem die Änderungsgeschwindigkeit der Ausgangsgröße der Änderung der Eingangsgröße proportional ist. Formal bedeutet das für die allgemeine DGL Gl. (7.33), dass der Koeffizient $b_0 = 0$ und $a_0 = 1$ ist. Der Frequenzgang des Systems ist demnach:

$$F(i\omega) = \frac{1}{b_1 \cdot i\omega}.$$

Im Beharrungszustand (Motordrehzahl $n = \text{const.}$, bzw. Drehzahländerungsfrequenz $\omega = 0$) gilt daher:

$$F(i\omega) \rightarrow \infty,$$

d. h. dass eine noch so kleine Eingangsgröße x_e eine unendlich große Abweichung x_a hervorruft, oder dass ein Ruhezustand nur für $x_e = 0$ eintreten kann. Bei technischen Systemen wie einem Spindel-/Mutter-System kann der Ausgangswert (Weg) natürlich nicht unendlich groß werden, sondern wird durch eine nichtlineare Vorrichtung, wie ihn beispielsweise ein Anschlag darstellt, begrenzt. Ansonsten gilt jedoch das gerade Gesagte. Aus diesem Grund werden solche Übertragungsglieder, die die Ausgangsgröße nach einer Änderung der Eingangsgröße nicht auf einen neuen Beharrungszustand ausgleichen, auch als **Strecken ohne Ausgleich** bezeichnet. Da wie oben angeführt, die Ausgangsgröße zu jedem Zeitpunkt dem Integral über der Eingangsgröße proportional ist, bezeichnet man das Übertragungsverhalten auch als **integrierendes Verhalten** oder kurz I - Verhalten. Andere technische Vorgänge mit I - Verhalten sind beispielsweise das Laden eines Kondensators aus einer Konstantstromquelle oder die Kraftbeaufschlagung eines viskosen Dämpfers.

Reines integrierendes Verhalten tritt in technischen Systemen in der Regel nicht auf, sondern ist mit Verzögerungseigenschaften aufgrund von anfänglich aufzufül-

Übertragungsverhalten eines PT₂-Gliedes und bekommt als Stellsignal vom P-Regler (reine P-Verstärkung) eine Spannung U_A , die proportional der Differenz aus Sollweg x_s und Istweg x_i ist. Die Ausgangsgröße des Motors ist die Drehzahl n , welche das Übertragungsglied **Linearantrieb** in den Istweg x_i

lenden oder zu leerenden Energiespeichern verbunden. Dies wird auch klar, wenn man den allgemeinen Frequenzgang eines integrierenden Systems ($b_0 = 0$) wie folgt umformt:

$$F(i\omega) = \frac{1}{b_1(i\omega) + b_2(i\omega)^2 + \dots} = \frac{1}{b_1(i\omega)} \cdot \frac{1}{1 + \frac{b_2}{b_1}(i\omega) + \dots}$$

$$= \underbrace{\frac{K_I}{(i\omega)}}_{\text{Ideales I-Glied}} \cdot \underbrace{\frac{1}{1 + T_1(i\omega) + \dots}}_{\text{Verzögerung}}$$

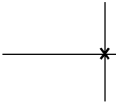
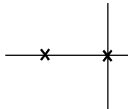
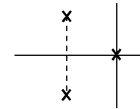
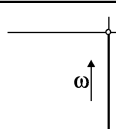
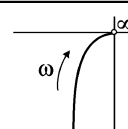
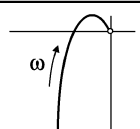
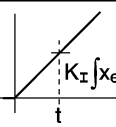
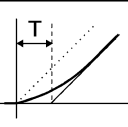
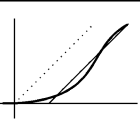
	I-Glied	I / PT ₁ -Glied	I / PT ₂ -Glied
Übertragungsfunktion G(s)	$K_I \frac{1}{s}$	$\frac{K_I}{s(1 + T s)}$	$\frac{K_I}{s(1 + 2DT s + T^2 s^2)}$
Pole von G(s)			
Frequenzgang F(iω)	$K_I \frac{1}{i\omega}$	$\frac{K_I}{i\omega (1 + T i\omega)}$	$\frac{K_I}{i\omega (1 + 2DT i\omega + T^2(i\omega)^2)}$
Ortskurve von F(iω)			
Übergangsfunktion h(t)			

Bild 7-70: Übertragungsfunktion, Pole der Übertragungsfunktion, Frequenzgang, Ortskurve und Übergangsfunktion von integrierenden Übertragungsgliedern ohne und mit Verzögerung

Im Bild 7-70 sind wieder verschiedene Möglichkeiten von idealem und realem integrierendem Verhalten dargestellt. Die Ortskurven der integrierend wirkenden Verzögerungsglieder kommen für $\omega \rightarrow \infty$ aus dem negativ imaginär Unendlichen und durchlaufen wieder so viele Quadranten, wie es der Ordnung der Verzögerungsglieder entspricht.

Ein dritter wichtiger Typ von Systemen mit speziellen Übertragungseigenschaften ist in dem Beispiel in Bild 7-71 a) enthalten. Gibt man auf die abgebildete Kondensator / Widerstandsschaltung einen Spannungssprung als Eingangsgröße, so erfolgt am Ausgang ein Spannungsverlauf (Bild 7-71 b)), der zuerst schlagartig auf einen Höchstwert springt und dann einer Exponentialfunktion folgend auf Null zurückgeht. Diesem Verhalten liegt wiederum ein idealisiertes Verhalten zugrunde, das in Bild 7-71 c) dargestellt ist. Als Antwort auf einen Eingangssprung erfolgt bei

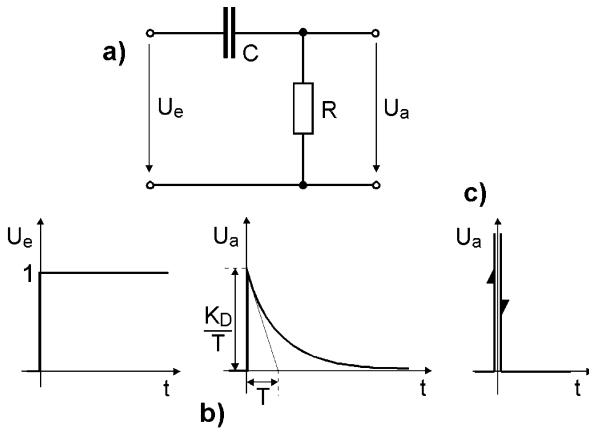


Bild 7-71: D-Verhalten eines D-Gliedes a) Schaltbild
b) Sprungantwort reales D-Glied
c) Sprungantwort ideales D-Glied

dem idealisierten Verhalten innerhalb eines infinitesimalen Zeitraums ein Sprung der Ausgangsgröße ins Unendliche und wieder zurück. Es ist unmittelbar einsichtig, dass dieses dem **Delta-Dirac-Impuls** entsprechende Verhalten in technischen Systemen nicht möglich ist. Das Verhalten wird durch folgende DGL beschrieben:

$$\dot{x}_a = K_D \cdot \dot{x}_e \quad (7.36)$$

Da also die Ausgangsgröße bei solchen Systemen der Änderungsgeschwindigkeit (Differentialquotient) der Eingangsgröße proportional ist, bezeichnet man

dieses Übertragungsverhalten als **differentiales** oder kurz **D-Verhalten**. Formal bedeutet dies, bezogen auf die allgemeine DGL nach Gl. (7.33), dass die Koeffizienten $b_0 = 1$ und $b_1 = 0$, sowie $a_0 \neq 0$ sind. Dies ergibt folgenden Frequenzgang:

$$F(i\omega) = K_D \cdot i\omega$$

Enthält die DGL auf der rechten Seite auch Ableitungen der Eingangsgröße, so ist ebenfalls differentiales Verhalten höherer Ordnung möglich. Für die 1. Ordnung lautet eine solche DGL allgemein:

	D-Glied	PD ₁ -Glied	PD ₂ -Glied	D/PT ₁ -Glied
Übertragungsfunktion G(s)	$K_D s$	$(1 + T s)$	$(1 + 2DT s + T^2 s^2)$	$\frac{K_D s}{(1 + T s)}$
Pole und Nullstellen von G(s)				
Frequenzgang F(iω)	$K_D i\omega$	$(1 + T i \omega)$	$(1 + 2DT i \omega + T^2 (i\omega)^2)$	$\frac{K_D i\omega}{(1 + T i\omega)}$
Ortskurve von F(iω)				

Bild 7-72: Übertragungsfunktion, Pole und Nullstellen der Übertragungsfunktion, Frequenzgang und Ortskurve von D-Gliedern ohne und mit Verzögerung

$$b_0 x_a = a_0 x_e + a_1 \dot{x}_e \quad .$$

Der zugehörige Frequenzgang für $a_0 = b_0 = 1$ und $a_1 = T$ lautet dann

$$F(i\omega) = 1 + T \cdot i\omega \quad .$$

Ein Übertragungsglied mit diesem Verhalten wird als **PD₁-Glied** bezeichnet. In Bild 7-72 sind wiederum einige Darstellungsformen verschiedener Übertragungsglieder mit D-Verhalten zusammengestellt. Da die Übertragungsfunktion von D-Gliedern als Extremstellen keine Pole sondern Nullstellen besitzt, sind diese in der komplexen Ebene eingetragen.

Wie bereits oben erwähnt kann es ein Übertragungsglied mit idealem D-Verhalten nicht geben, sondern das Verhalten realer technischer Systeme ist, wie bei den meisten realen Systemen mit I-Verhalten, mit zusätzlichem Verzögerungsverhalten verbunden. Der Frequenzgang eines entsprechenden realen D-Gliedes lautet demnach beispielsweise:

$$F(i\omega) = \frac{K_D \cdot i\omega}{1 + T \cdot i\omega} \quad .$$

Die Ortskurve dieses Übertragungssystems ist ebenfalls in Bild 7-72 dargestellt. Die zugehörige Übergangsfunktion wurde bereits in Bild 7-71 b) gezeigt. Übergangsfunktionen der idealen D-Glieder in Bild 7-72 sind nicht darstellbar, sondern ebenfalls nur im Zusammenhang mit zusätzlichem Verzögerungsverhalten sinnvoll.

Als weitere Typen von Übertragungsgliedern sind natürlich noch beliebige Kombinationen der Grundtypen möglich, deren Verhalten aber aus dem der Grundtypen abgeleitet werden kann. Außerdem tritt häufig im Zeitverhalten eines Übertragungsgliedes eine sog. **Totzeit** auf, eine Zeit, in der beispielsweise nach einem Eingangssprung, das System noch überhaupt nicht mit irgend einer Änderung reagiert. Ein typisches Beispiel für ein solches Verhalten ist bei einem mechanischen System ein vorhandenes Spiel, das bei einem Bewegungsvorgang vom antreibenden Element zuerst durchlaufen werden muss, bis das angetriebene Element überhaupt reagiert.

7.3.1.5 Frequenzkennlinien

Eine Möglichkeit der Darstellung der komplexen Funktion **Frequenzgang**

$$F(i\omega) = A(\omega) \cdot e^{i\alpha(\omega)} \quad ,$$

ist die bereits behandelte Ortskurve in der komplexen Zahlenebene. In dieser Darstellung sind die beiden Informationen über Phasenlage und Amplitude der Ausgangsschwingung bei Anliegen einer harmonischen Eingangsschwingung der Frequenz ω in einem Bild zusammengefasst. Eine weitere übliche Darstellungsart ist die als **Frequenzkennlinien**, in der die Amplituden- und die Phaseninformation in Abhängigkeit von ω als Einzeldiagramme verwendet werden.

Besonders günstig ist die Verwendung der Frequenzkennlinien in logarithmischer Darstellung, die auch als **Bode-Diagramm** bekannt ist. Logarithmiert man die obige

Gleichung des Frequenzgangs, so erhält man:

$$\ln F(i\omega) = \ln A(\omega) + i\alpha(\omega) \quad \text{oder}$$

$$\log F(i\omega) = \log A(\omega) + i\alpha(\omega) \cdot \log e \quad (7.37)$$

Diese Art der Darstellung der logarithmierten Amplitudenwerte über der Frequenz (**Amplitudengang**) und der Phasenwerte über der Frequenz (**Phasengang**) hat den Vorteil, dass aus dem Produkt zweier Frequenzgänge (Reihenschaltung zweier Übertragungsglieder)

$$F(i\omega) = F_1(i\omega) \cdot F_2(i\omega)$$

durch Logarithmierung folgender Ausdruck entsteht:

$$\ln F(i\omega) = \ln F_1(i\omega) + \ln F_2(i\omega) = \ln A_1(i\omega) + \ln A_2(i\omega) + i[\alpha_1(\omega) + \alpha_2(\omega)]$$

Der Gesamtfrequenzgang einer Reihenschaltung ergibt sich demnach durch einfache Addition der einzelnen Kennlinien. Diese Tatsache lässt sich einerseits günstig

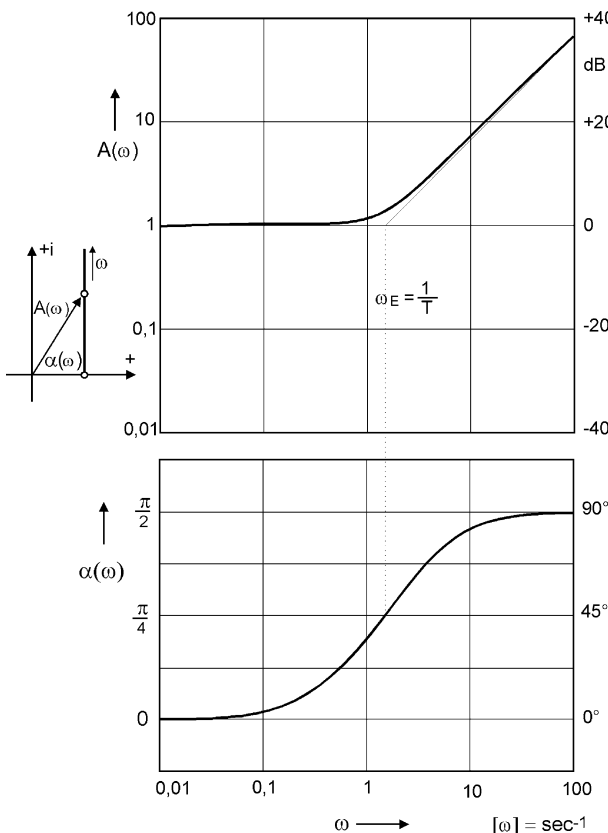


Bild 7-73: Amplituden- und Phasenkennlinie eines PD₁-Gliedes

zeigt die Frequenzkennliniendarstellung eines PD₁-Gliedes.

Die Frequenzkennlinien für einfache P, I und D-Glieder sind Geraden. Ihren Verlauf

für die grafische Konstruktion der Frequenzkennlinien einer komplexen Reihenschaltung anwenden. Hat man andererseits einen Übertragungsblock mit einer komplizierten Übertragungsfunktion, so lässt sich für bestimmte Fälle (nur negativ reelle Wurzeln von Zähler- und Nennerpolynom) zeigen, dass diese stets als Produkt (Reihenschaltung) einfacher Grundformen (P-, I-, D, PT₁-, PD₁-Verhalten) von Übertragungsfunktionen dargestellt werden können. Zeichnet man dann die Einzelfrequenzkennlinien, so erhält man die Gesamtkennlinie durch grafische Addition der Einzelkennlinien. Führt man noch für die Darstellung des Amplitudengangs die logarithmische **Dezibel-Skala** ein

$$\frac{A}{\text{dB}} = 20 \cdot \log A$$

so kann man sowohl den Amplitudengang als auch den Phasengang in einem linearen Maßstab auftragen. Bild 7-73

kann man leicht aus den zugehörigen Frequenzgängen ermitteln.

P-Verhalten:

$$F(i\omega) = A(\omega) \cdot e^{i\alpha(\omega)} = K_p \Rightarrow |F| = A = K_p \quad \alpha = 0$$

Der Amplitudengang ist demnach eine Parallele zur ω -Achse im Abstand K_p , die Phasenkennlinie verläuft in der ω -Achse.

D-Verhalten:

$$F(i\omega) = K_D \cdot i\omega = i \cdot T \cdot \omega = T \cdot \omega \cdot e^{i\pi/2} \Rightarrow A = T \cdot \omega \quad \alpha = \frac{\pi}{2}.$$

Der Amplitudengang ist eine Gerade mit der Steigung T , da $\log A = T \cdot \log \omega$ gilt. Der Phasengang verläuft konstant bei einem Wert von $\pi/2$.

I-Verhalten:

$$F(i\omega) = \frac{K_I}{i \cdot \omega} = \frac{1}{i \cdot T \cdot \omega} = -i \frac{1}{T \cdot \omega} = \frac{1}{T \cdot \omega} \cdot e^{-i\pi/2} \Rightarrow A = \frac{1}{T \cdot \omega} \quad \alpha = -\frac{\pi}{2}.$$

Der Amplitudengang ist eine Gerade mit der Steigung $-1/T$, da

$$\log A = \log \left(\frac{1}{T} \cdot \frac{1}{\omega} \right) = \frac{1}{T} (\log 1 - \log \omega) = -\frac{1}{T} \cdot \log \omega$$

ist. Der Phasengang verläuft konstant bei einem Wert von $-\pi/2$.

Wie in Bild 7-73 gezeigt, verlaufen die Frequenzkennlinien eines PD₁-Gliedes und analog auch die eines PT₁-Gliedes etwas komplizierter.

PD₁-Verhalten:

$$F(i\omega) = 1 + i \cdot T \cdot \omega \Rightarrow A = \sqrt{1 + (T \cdot \omega)^2} \quad \alpha = \arctan(T \cdot \omega).$$

Für den Amplitudengang gilt:

$$\log A(\omega) = \frac{1}{2} \log(1 + (T\omega)^2)$$

$$\omega \rightarrow 0 \Rightarrow \log A \rightarrow \log 1 = 0 \quad \omega \rightarrow \infty \Rightarrow \log A \rightarrow \log(T\omega) = \log T + \log \omega.$$

Die Geraden $\log A = 0$ und $\log A = \log T + \log \omega$ sind nach der obigen Grenzwertbetrachtung Asymptoten an den Verlauf der Amplitudenkennlinie. Die Asymptote für größere Werte von ω hat die Steigung 1 und schneidet die mit der ω -Achse zusammenfallende Asymptote bei der so genannten **Eckfrequenz** $\omega_E = 1/T$ (Bild 7-73). Der wahre Verlauf der Amplitudenkennlinie weicht für Werte oberhalb und unterhalb von der Eckfrequenz nur wenig vom Verlauf der Asymptoten ab, so dass diese für das Zeichnen der Amplitudenkennlinie eine gute Näherung darstellen. Im Bereich der Eckfrequenz können Abweichungen von maximal 3 dB = 41% auftreten. Die daher für diesen Bereich erforderliche Amplitudenkorrektur ist in Bild 7-74 zusammen mit den Frequenzkennlinien abgebildet. Da wie schon bei den D- und I-Gliedern sich die jeweils andere Kennlinie durch Spiegelung an der ω -Achse ergab, kann

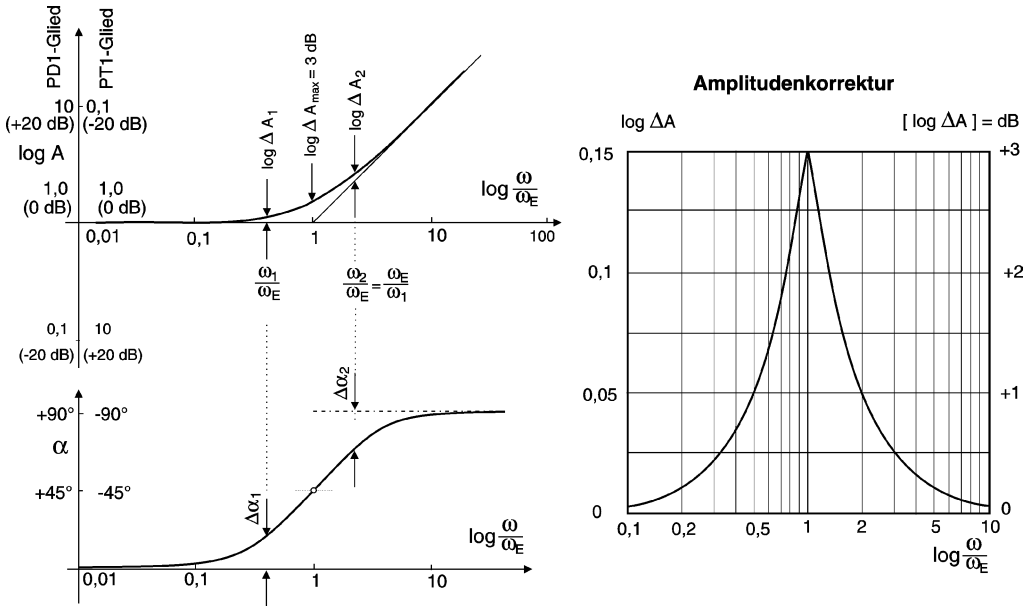


Bild 7-74: Frequenzkennlinien mit Asymptoten von PD₁- und PT₁-Gliedern und Amplitudenkorrektur

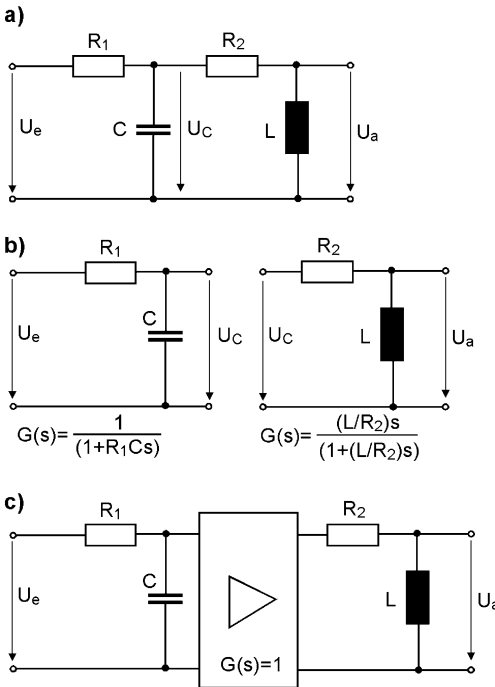


Bild 7-75: RLC-Netzwerk a) Zusammenschaltung b) getrennte RC- und RL-Glieder c) RC- und RL-Glieder verbunden durch rückwirkungs-freien Verstärker

man die Kennlinien für ein PT₁-Glied auf gleiche Weise aus denen des PD₁-Gliedes ableiten. Sie sind daher ebenfalls im Bild 7-74 dargestellt. Die Kennliniendarstellung in Abhängigkeit von der Kreisfrequenz ω erfolgt bezogen auf die Eckfrequenz ω_E .

Als Näherung für die Phasenkennlinie kann man drei Geraden verwenden. Für kleine Werte von ω eine Gerade bei $\alpha = 0^\circ$ und bei großen Werten eine Gerade bei $\alpha = 90^\circ$. Im Bereich der Eckfrequenz wechselt die Phasenkennlinie von 0° auf 90° . Dies kann man durch eine Gerade annähern, die durch $\alpha = 45^\circ$ verläuft und eine Steigung von 45° pro Dekade aufweist.

Im folgenden soll nun gezeigt werden, wie man zeichnerisch die Frequenzkennlinien eines elektrischen RLC-Netzwerks, das in Bild 7-75 a) dargestellt ist, ermittelt. Man könnte nun versucht sein, das Netzwerk wie in Bild 7-75 b) gezeigt in zwei einfache Teilnetzwerke mit bereits

bekannten Übertragungsfunktionen aufzuspalten und die Gesamtübertragungsfunktion dann als Reihenschaltung dieser beiden Teilnetzwerke zu ermitteln. Dies geht aber nur, wenn wie bei der Behandlung von Reihenschaltungen im Kapitel 7.3.1.2 stillschweigend vorausgesetzt wurde, keine Rückwirkung des zweiten auf den ersten Block stattfindet. Die hier verwendeten RC- (Verzögerungsglied 1. Ordnung) und RL-Glieder (D-Glied mit Verzögerung 1. Ordnung) erfüllen diese Bedingung jedoch nicht, da durch die Verbindung beider Glieder die ursprüngliche Kondensatorspannung wesentlich beeinflusst wird. Die geforderte Rückwirkungsfreiheit ließe sich, wie in Bild 7-75 c) gezeigt, durch die Verbindung über einen Trennverstärker mit der Übertragungsfunktion $G(s) = 1$ erreichen, wobei dann die Gesamtübertragungsfunktion des Netzwerkes mit Trennverstärker

$$\begin{aligned}
 G(s) &= G_1(s) \cdot G_2(s) \cdot G_3(s) = \frac{1}{1 + R_1 C s} \cdot 1 \cdot \frac{L / R_2 s}{1 + L / R_2 s} = \frac{T_2 \cdot s}{(1 + T_1 s)(1 + T_2 s)} \\
 &= \frac{\frac{s}{T_1}}{s^2 + \left(\frac{T_1 + T_2}{T_1 \cdot T_2} \right) s + \frac{1}{T_1 \cdot T_2}} \quad (7.38)
 \end{aligned}$$

betragen würde. Ohne den Trennverstärker ist das Netzwerk jedoch nicht rückwirkungsfrei, so dass das RC-Glied und das RL-Glied nicht unabhängig voneinander behandelt werden können. Man muss daher die Übertragungsfunktion nach den Regeln von Spannungsteilern unter Berücksichtigung der Blindwiderstände

$$R_C = \frac{1}{C \cdot s} \quad \text{und} \quad R_L = L \cdot s$$

ermitteln. Man kann nun zuerst einmal das Verhältnis der Eingangsspannung U_a zur Spannung U_C am Kondensator ermitteln, wenn man wie in Bild 7-76 a) gezeigt, die Reihenschaltung aus R_1 mit der Parallelschaltung von C und der Reihenschaltung aus R_2 und L bildet.

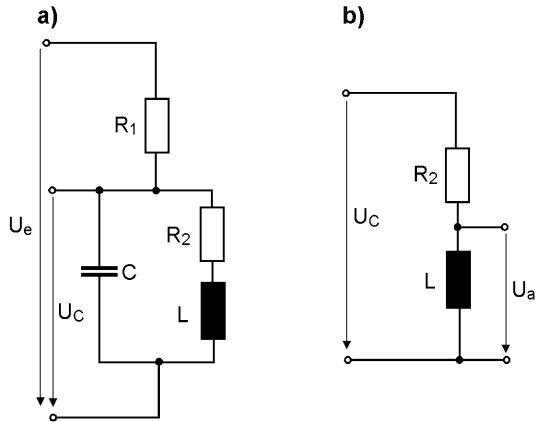


Bild 7-76: Herleitung der Übertragungsfunktion eines nicht rückwirkungsfreien RLC-Netzwerkes

$$\begin{aligned}
 \frac{U_C}{U_e} &= \frac{\frac{(R_2 + L \cdot s) \cdot \frac{1}{C \cdot s}}{(R_2 + L \cdot s) + \frac{1}{C \cdot s}}}{R_1 + \frac{\frac{(R_2 + L \cdot s) \cdot \frac{1}{C \cdot s}}{(R_2 + L \cdot s) + \frac{1}{C \cdot s}}}{1}} = \frac{(R_2 + L \cdot s) \cdot \frac{1}{C \cdot s}}{R_1 R_2 + L R_1 \cdot s + \frac{R_1 + R_2}{C \cdot s} + \frac{L}{C}}
 \end{aligned}$$

Anschließend stellt man das Spannungsverhältnis für den Spannungsteiler aus R_2 und L auf (Bild 7-76 b)) :

$$\frac{U_a}{U_c} = \frac{L \cdot s}{R_2 + L \cdot s}$$

Kombiniert man die beiden letzten Gleichungen, so erhält man die Gesamtübertragungsfunktion:

$$G(s) = \frac{U_a}{U_e} = \frac{\frac{L}{C} \cdot s}{LR_1 \cdot s^2 + \left(R_1 R_2 + \frac{L}{C}\right) \cdot s + \frac{R_1 + R_2}{C}} \quad .$$

Schreibt man dies wieder mit den schon in Gl. (7.38) verwendeten Zeitkonstanten

$T_1 = R_1 \cdot C$ und $T_2 = L / R_2$, so erhält man:

$$G(s) = \frac{s/T_1}{s^2 + \left(\frac{T_1 + T_2}{T_1 \cdot T_2}\right) \cdot s + \frac{1 + R_1/R_2}{T_1 \cdot T_2}} \quad (7.39)$$

Vergleicht man die beiden Gleichungen (7.38) und (7.39), so sieht man einen Unterschied im PT_2 -Glied des Nenners der Übertragungsfunktionen, der die Rückwirkung des LR-Gliedes auf die Kondensatorspannung des RC-Gliedes beinhaltet. Der Unterschied im Übertragungsverhalten hängt von den Werten von R_1 und R_2 ab. Nur wenn $R_2 \gg R_1$ ist, sind die beiden Übertragungsfunktionen annähernd gleich; es findet dann keine nennenswerte Rückwirkung mehr statt.

Zum Zeichnen der Frequenzkennlinien der Übertragungsfunktion (7.39) werden nun konkrete Werte für R , L und C gewählt:

$$R_1 = 10^2 \, \Omega, \quad R_2 = 10^2 \, \Omega, \quad C = 86 \cdot 10^{-3} \, \text{F}, \quad L = 10^2 \, \text{H} \quad .$$

Damit errechnen sich die Zeitkonstanten zu

$$T_1 = 8,6 \, \text{Sek.} \quad , \quad T_2 = 1,4 \, \text{Sek.}$$

und die Übertragungsfunktion Gl. (7.39) bekommt folgenden Wert:

$$G(s) = \frac{0,116 \cdot s}{s^2 + \frac{5}{6} \cdot s + \frac{1}{6}} = \frac{0,697 \cdot s}{6s^2 + 5s + 1}$$

Diese Übertragungsfunktion kann man nun in ein Produkt einfacher Grundfunktionen aufspalten, von denen der Verlauf der Frequenzkennlinien bekannt ist:

$$G(s) = 0,697 \cdot s \cdot \frac{1}{(3s+1)} \cdot \frac{1}{(2s+1)}$$

P
D
(PT₁)₁
(PT₁)₂

Man kann also die Gesamtübertragungsfunktion als Produkt von vier Einzelübertragungsfunktionen schreiben, nämlich einem P-Glied, einem D-Glied und zwei PT₁-Gliedern. Die Eckfrequenzen für die beiden PT₁-Glieder und die Verstärkung haben folgende Werte:

$$T_1 = 3 \text{ sec} \quad \omega_{E1} = \frac{1}{T_1} = 0,33 \frac{\text{rad}}{\text{sec}} \quad T_2 = 2 \text{ sec} \quad \omega_{E2} = \frac{1}{T_2} = 0,5 \frac{\text{rad}}{\text{sec}} \quad K_p = 0,697 = -3,5 \text{ dB}$$

Wie oben erläutert, kann man nun die sehr einfachen Kennlinien dieser Grundglieder in ein Diagramm mit logarithmischem Maßstab eintragen und zur Ermittlung der Gesamtkennlinien die der Einzelglieder grafisch addieren. Dies ist in Bild 7-77 dargestellt. Die Amplitudenkennlinie des P-Gliedes ist eine Parallele zur ω -Achse im Abstand -3,48 dB (= 0,697), die Kennlinie des D-Gliedes geht durch den Wert $\omega = 1$ und steigt mit 20 dB/Dekade (Frequenzdekade). Die beiden PT₁-Glieder werden jeweils in Form zweier Asymptoten gezeichnet, die sich bei der zugehörigen Eckfrequenz schneiden. Die eine Asymptote (für kleine ω -Werte) verläuft auf der ω -Achse, die andere fällt mit 20 dB/Dekade. Ebenfalls im Amplitudendiagramm enthalten ist die Summenkurve der Einzelverläufe und der wirkliche Verlauf mit Ausrundungen der Kurve im Bereich der Eckfrequenzen. Man sieht am Amplitudengang des RLC - Gliedes, dass es sich um einen so genannten Bandpass handelt, der nur einen bestimmten Frequenzbereich optimal durchlässt. Diese Durchlassfrequenz liegt hier bei ca. $\omega = 0,4 \text{ rad/sec}$.

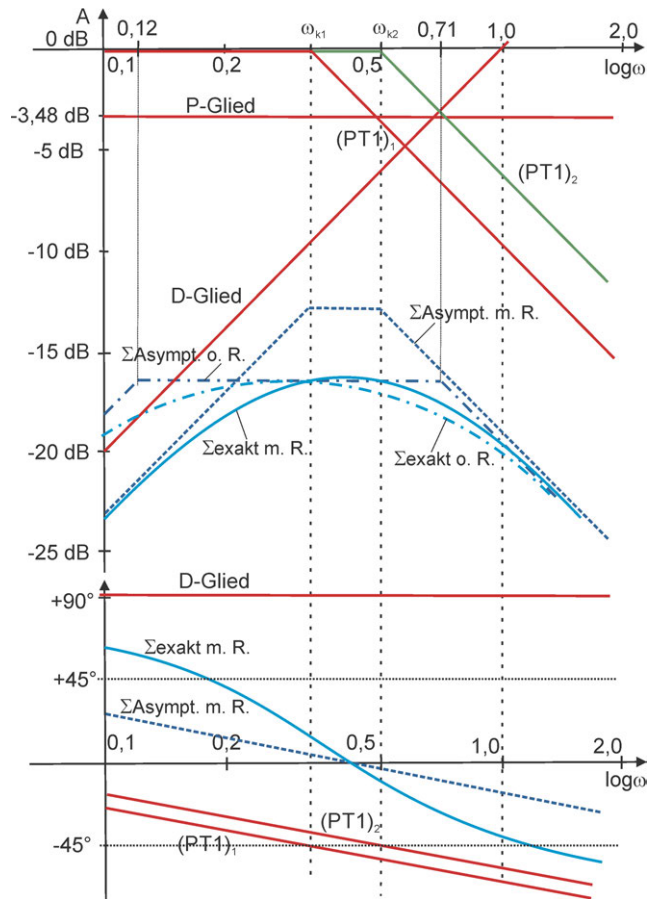


Bild 7-77: Frequenzkennlinien eines RLC-Netzwerks

Das P-Glied hat keinen Einfluss auf die Phasenkennlinie, das D-Glied verursacht eine konstante Phasenverschiebung von 90°. Die beiden PT₁-Glieder verursachen jeweils eine Phasenverschiebung von -90° für höhere Frequenzen mit Übergängen im Bereich der Eckfrequenzen. Die Phasenkennlinie beginnt daher bei 90° für sehr kleine Frequenzen ($\omega < 0,033$) und verläuft bei -90° bei hohen Frequenzen ($\omega > 5$).

Hätte man die RC- und die RL-Kombination getrennt durch einen rückwirkungsfreien Verstärker behandelt und dementsprechend die Gl. (7.38) als Übertragungsfunktion benutzt, so hätte diese für das Zahlenbeispiel gelautet:

$$G(s) = \frac{1,394 \cdot s}{(8,6s + 1) \cdot (1,4s + 1)}$$

Daraus ergeben sich folgende Werte für die Verstärkung und die Eckfrequenzen:

$$\omega_{E1} = 0,12 \frac{\text{rad}}{\text{sec}} \quad \omega_{E2} = 0,71 \frac{\text{rad}}{\text{sec}} \quad K_p = 1,394 = 2,88 \text{ dB}$$

Der für diese Übertragungsfunktion gültige Amplitudengang ist ebenfalls im Amplitudendiagramm in Bild 7-77 enthalten. Man sieht, dass dann die Durchlassfrequenz weniger ausgeprägt, die Durchlassdämpfung höher und die Mittenfrequenz niedriger liegt.

Die Methodik der zeichnerischen Ermittlung der Frequenzkennlinien in Form des Bode-Diagramms stammt aus einer Zeit, bevor leistungsfähige Digitalrechner mit grafischer Ausgabe zur Verfügung standen. Heute lassen sich die Frequenzkennlinien relativ einfach unter Verwendung eines Simulationssystems ausplotten. Bild 7-78 zeigt entsprechende Plots des Systems 20-sim.

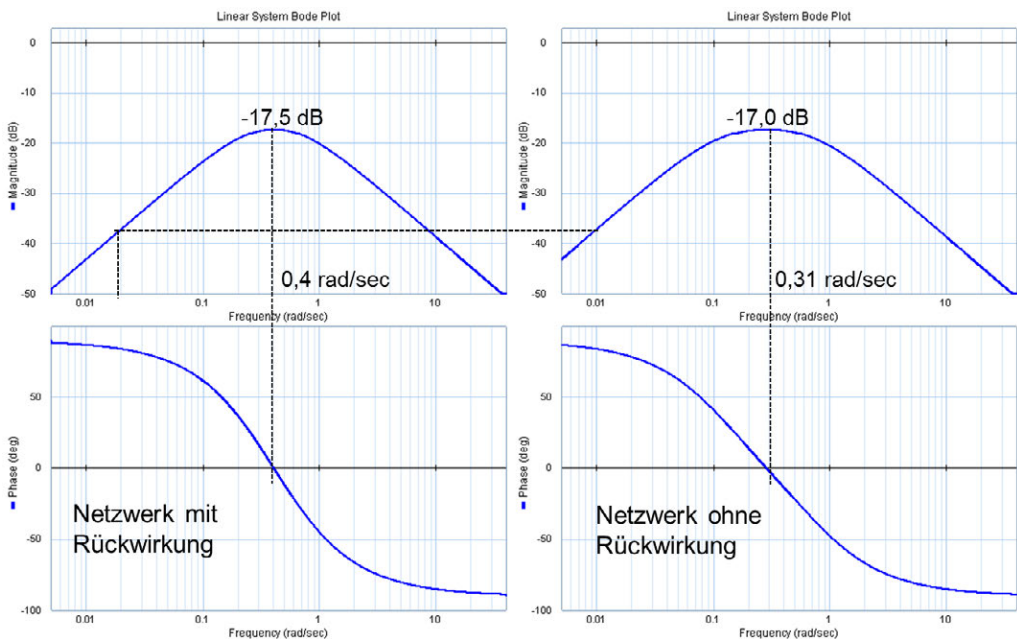


Bild 7-78: Frequenzkennlinienplot des RLC-Netzwerks mit und ohne Rückwirkung

7.3.1.6 Zustandsraumdarstellung

Die bisher betrachteten Beschreibungsformen dynamischer Systeme haben alle gemeinsam, dass sie lediglich das Übertragungsverhalten vom Eingang zum Aus-

gang des Systems beschreiben. Man kann mit den zugehörigen Modellgleichungen zwar den Verlauf der Ausgangsgröße bei Vorhandensein einer bekannten Eingangsgröße ermitteln, erhält aber keinerlei Informationen darüber, was sich im Inneren der durch die Modellgleichung beschriebenen Einheit abspielt. Diese Darstellungsweisen haben sich entwickelt, weil ein Lösen der Modellgleichungen ohne Rechneinsatz so relativ einfach ist.

Heutzutage stehen leistungsfähige Digitalrechner zur Verfügung und Methoden zur Lösung von Gleichungssystemen, die nicht mehr auf das klassische Lösen von Differentialgleichungen angewiesen sind. Ende des 19. Jahrhunderts entwickelte der russische Mathematiker A. M. Ljapunow das Konzept der Systemanalyse mit Hilfe so genannter **Zustandsgrößen**, das aber erst in den 60-iger Jahren des 20. Jahrhunderts sinnvoll eingesetzt werden konnte, nachdem leistungsfähige Rechner zur Verfügung standen [7.9].

Das Zustandsraummodell steht, wie in Bild 7-79 gezeigt, natürlich immer noch als Übertragungsblock zwischen Ein- und Ausgangsgrößen, (hier in Anlehnung an die in der Literatur übliche Schreibweise mit u und y bezeichnet) beinhaltet aber implizit einen Vektor $\vec{x}$, der als Komponenten die Zustandsgrößen enthält. Dies sind Hilfsgrößen, die den dynamischen

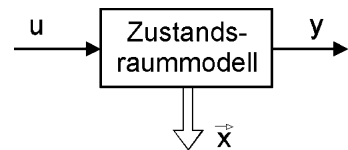


Bild 7-79: Zustandsraummodell

Zustand des Übertragungsgliedes beschreiben. Der Raum, der von den Zustandsgrößen aufgespannt wird, wird auch als **Zustandsraum** bezeichnet. Die Darstellung in einem zweidimensionalen Zustandsraum haben wir in Kapitel 4 in Form des Phasendiagramms eines Pendels kennen gelernt, das den Verlauf der Winkelgeschwindigkeit in Abhängigkeit vom Winkel darstellt. Winkel und Winkelgeschwindigkeit sind zwei mögliche Zustandsgrößen des Systems.

Mit Hilfe dieser voneinander unabhängigen Zustandsgrößen kann man das klassische Modell für ein dynamisches System, das sich durch eine Differentialgleichung n -ter Ordnung beschreiben lässt, durch ein System von n Differentialgleichungen erster Ordnung ersetzen. Dies ist für die Lösung mit numerischen Verfahren auf einem Digitalrechner viel angenehmer, da hier zur Lösung nur eine einfache Integration notwendig ist.

Ein weiterer Vorteil der Zustandsraummethode ist, dass bei Kenntnis des aktuellen Zustandes eines Systems die Voraussage über künftige Zustände allein aus den Informationen über die Eingangsgrößen gewonnen werden kann. Kenntnisse aus der Vorgeschichte des Systems sind nicht erforderlich.

Die Einführung von Zustandsvariablen für ein bestimmtes System ist nicht eindeutig. Man kann für das gleiche System verschiedene Zustandsgrößen wählen, um eventuell bestimmte Beschreibungsmöglichkeiten zu vereinfachen. Dies soll an einem einfachen System gezeigt werden, das durch eine DGL 2. Ordnung beschrieben werden kann.

Die DGL möge lauten:

$$\ddot{y} + 3\dot{y} + 2y = u \quad .$$

Als Zustandsgrößen werden nun willkürlich folgende Hilfsgrößen eingeführt:

$$x_1 = y \quad x_2 = \dot{y} \quad .$$

Aufgrund dieser Festlegung ergibt sich

$$\dot{x}_1 = x_2 \quad .$$

Setzt man dies in die DGL ein und stellt nach $\dot{x}_2$ um, so erhält man:

$$\dot{x}_2 = -2x_1 - 3x_2 + u \quad .$$

Die beiden letzten Gleichungen, in denen die Zustandsgrößen nur als direkte zeitabhängige Größen und ihre ersten Ableitungen nach der Zeit vorkommen, werden **Zustandsgleichungen** genannt. Das System, das vorher durch eine DGL zweiter Ordnung beschrieben wurde, kann nun mit Hilfe der Zustandsgrößen durch zwei DGL's erster Ordnung beschrieben werden. Da in der Regel der Wert der Ausgangsgröße interessiert, benötigt man noch eine **Ausgangsgleichung**, die im vorliegenden Fall besonders einfach ist:

$$y = x_1 \quad .$$

Hätte man dagegen die Zustandsgrößen wie folgt gewählt

$$x_1 = 2y + \dot{y} \quad x_2 = -y - \dot{y} \quad ,$$

so würden sich andere Zustands- und Ausgangsgleichungen ergeben. Aus den letzten beiden Gleichungen kann man ableiten

$$y = x_1 + x_2 \quad \text{und} \quad \dot{y} = -x_1 - 2x_2 \quad ,$$

woraus wiederum durch Kombination

$$\dot{x}_1 + \dot{x}_2 = -x_1 - 2x_2$$

folgt. Dies ist eine erste Zustandsgleichung. Eine weitere erhält man direkt aus der DGL nach Umstellung auf die Form

$$\ddot{y} = -3\dot{y} - 2y + u$$

und wegen

$$\ddot{y} = \frac{d}{dt}(-x_1 - 2x_2) = -\frac{dx_1}{dt} - 2\frac{dx_2}{dt} = -\dot{x}_1 - 2\dot{x}_2$$

ergibt sich

$$-\dot{x}_1 - 2\dot{x}_2 = -3(-x_1 - 2x_2) - 2(x_1 + x_2) + u$$

Diese beiden Zustandsgleichungen lassen sich auf folgende Form vereinfachen:

$$\begin{aligned} \dot{x}_1 &= -x_1 & + u \\ \dot{x}_2 &= & - 2x_2 - u \quad . \end{aligned}$$

Die Ausgangsgleichung war bereits durch die Wahl der Zustandsgrößen mit

$$y = x_1 + x_2$$

festgelegt worden. Der Vorteil der zweiten Auswahl der Zustandsgrößen ist, dass

die Zustandsgleichungen nicht mehr wie im ersten Fall miteinander gekoppelt sind, da nun in einer Zustandsgleichung außer der einen Zustandsgröße und ihrer ersten Ableitung nur noch die Ausgangsgröße y vorkommt.

Für den allgemeinen Fall, dass ein System mit einer linearen Differentialgleichungen n -ter Ordnung beschrieben werden kann und das r Eingangsgrößen u erhält, lauten die n Zustandsgleichungen:

$$\begin{aligned}\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n + b_{11}u_1 + \dots + b_{1r}u_r \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n + b_{21}u_1 + \dots + b_{2r}u_r \\ &\dots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \dots a_{nn}x_n + b_{n1}u_1 + \dots + b_{nr}u_r \quad .\end{aligned}$$

Dies kann man durch Einführung von Vektoren und Matrizen auch wie folgt schreiben:

$$\vec{\dot{x}} = \mathbf{A} \cdot \vec{x} + \mathbf{B} \cdot \vec{u} \quad (7.40)$$

Dabei ist

$$\begin{aligned}\vec{x} &= (x_1, x_2, \dots, x_n)^T && \text{Vektor der } \mathbf{Zustandsgrößen} \\ \vec{\dot{x}} &= (\dot{x}_1, \dot{x}_2, \dots, \dot{x}_n)^T && \text{Vektor der 1. Ableitung der Zustandsgrößen} \\ \vec{u} &= (u_1, u_2, \dots, u_n)^T && \text{Vektor der } \mathbf{Eingangs-} \text{ oder } \mathbf{Steuergrößen} \\ \mathbf{A} &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} && \mathbf{Systemmatrix} \text{ (Format } n \times n \text{)} \\ \mathbf{B} &= \begin{pmatrix} b_{11} & \dots & b_{1r} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nr} \end{pmatrix} && \mathbf{Steuermatrix} \text{ (Format } n \times r \text{)}\end{aligned}$$

Zwischen den Zustandsgrößen $\vec{x}$ und den Ausgangsgrößen $\vec{y}$ besteht eine lineare Beziehung

$$\vec{y} = \mathbf{C} \cdot \vec{x} + \mathbf{D} \cdot \vec{u} \quad (7.41)$$

Mit $\vec{y} = (y_1, y_2, \dots, y_p)^T$ Vektor der **Ausgangsgrößen**

$$\begin{aligned}\mathbf{C} &= \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \dots & \dots & \dots \\ c_{p1} & \dots & c_{pn} \end{pmatrix} && \mathbf{Ausgangs-} \text{ oder } \mathbf{Beobachtungsmatrix} \\ &&& \text{(Format } p \times n \text{)} \\ \mathbf{D} &= \begin{pmatrix} d_{11} & \dots & d_{1r} \\ \dots & \dots & \dots \\ d_{p1} & \dots & d_{pr} \end{pmatrix} && \mathbf{Durchgangsmatrix} \text{ (Format } p \times r \text{)}\end{aligned}$$

Wenn es nur eine Ausgangsgröße gibt, dann besteht die Matrix **C** aus einer einzigen Zeile. In der Regel gibt es keine direkte Verbindung zwischen den Eingangs- und Ausgangsgrößen, so dass dann die Matrix **D** gleich Null ist.

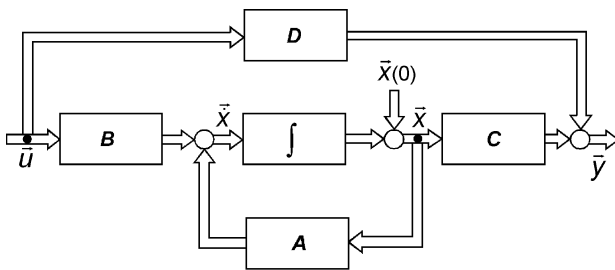


Bild 7-80: Blockschaltbild des Zustandsraummodells

Man kann die in den beiden Gleichungen (7.40) und (7.41) dargestellten Zusammenhänge auch als Blockschaltbild darstellen (Bild 7-80), wodurch der Zusammenhang und das Modell noch klarer werden.

An dem folgenden Beispiel soll noch einmal das Aufstellen der Zustandsgleichungen verdeutlicht werden.

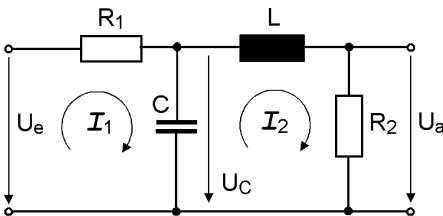


Bild 7-81: RLC-Netzwerk

Die sich bei Beaufschlagung der Schaltung mit einer dynamischen Eingangsgröße (Sprung, harmonische Schwingung, usw.) einstellenden dynamischen Vorgänge innerhalb der Schaltung sind durch die Speicherung von Energie in Kondensator und Induktivität bestimmt. Für die im Kondensator und in der Induktivität gespeicherte Leistung gilt:

$$P_C = \frac{1}{2} C \cdot U_C^2 \quad P_L = \frac{1}{2} L \cdot I_L^2 \quad .$$

Demnach ist zur Beschreibung des Kondensatorzustands die Kondensatorspannung U_C und zur Beschreibung des Zustands der Induktivität der Strom durch die Spule I_L geeignet. Daher werden als die zwei erforderlichen Zustandsgrößen $x_1 = U_C$ und $x_2 = I_L = I_2$ gewählt.

Um die beiden Zustandsgleichungen zu bestimmen, kann man die beiden folgenden Maschengleichungen aufstellen:

$$R_1 I_1 + \frac{1}{C} \int I_1 dt - \frac{1}{C} \int I_2 dt = U_e \quad (7.42)$$

$$-\frac{1}{C} \int I_1 dt + \frac{1}{C} \int I_2 dt + R_2 I_2 + L \dot{I}_2 = 0 \quad (7.43)$$

Durch Eliminierung der Ströme aus den beiden Maschengleichungen kann man die DGL des RLC-Netzwerks herleiten:

$$LCR_1 \ddot{U}_a + (L + R_1 R_2 C) \dot{U}_a + (R_1 + R_2) U_a = R_2 U_e \quad .$$

In den beiden Maschengleichungen kommen als Variablen die Ströme I_1 und I_2 vor, die nun in Beziehung zu den gewählten Zustandsgrößen gesetzt werden müssen:

$$x_1 = U_c = \frac{1}{C} \int (I_1 - I_2) dt \quad \text{Kondensatorspannung}$$

$$x_2 = I_L = I_2 \quad \text{laut Wahl der Zustandsvariablen.}$$

Daraus folgt für die Maschenströme:

$$I_2 = x_2 \quad I_1 = I_2 + C \cdot \dot{x}_1 = x_2 + C \cdot \dot{x}_1 \quad .$$

Setzt man dies in die Maschengleichungen ein,

$$R_1 x_2 + R_1 C \dot{x}_1 + x_1 = U_e \quad \text{aus Gl. (7.42)}$$

$$-x_1 + R_2 x_2 + L \dot{x}_2 = 0 \quad \text{aus Gl. (7.43)}$$

so erhält man nach Umordnen die beiden Zustandsgleichungen:

$$\dot{x}_1 = -\frac{1}{R_1 C} x_1 - \frac{1}{C} x_2 + \frac{1}{R_1 C} U_e$$

$$\dot{x}_2 = \frac{1}{L} x_1 - \frac{R_2}{L} x_2 \quad .$$

Die Ausgangsgleichung findet man sofort aus Bild 7-81 zu:

$$U_a = R_2 \cdot I_2 = R_2 \cdot x_2 \quad .$$

In Matrizenschreibweise lauten die Zustandsgleichungen und die Ausgangsgleichung schließlich:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{R_1 C} & -\frac{1}{C} \\ \frac{1}{L} & -\frac{R_2}{L} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \frac{1}{R_1 C} \\ 0 \end{pmatrix} \cdot U_e$$

$$U_a = (0, R_2) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad .$$

7.3.1.7 Regler

Bei Reglern oder Regeleinrichtungen, wie sie als Block innerhalb eines vollständigen Regelkreises (Bild 7-7) enthalten sind, handelt es sich prinzipiell nicht um andere Typen von Übertragungsgliedern, als wir sie bereits als Strecken kennen gelernt haben. Die dort üblichen Grundformen des Übertragungsverhaltens, nämlich P-, I- und D-Verhalten treten auch hier auf.

Verwendet man einen einfachen Verstärker mit P-Verhalten und einem Verstärkungsfaktor vom Wert K_p zur Regelung einer Strecke mit Ausgleich (beispielsweise PT_1 -Verhalten), so stellt man fest, dass es in jedem Arbeitspunkt der Strecke und bei beliebiger Verstärkung immer zu einer bleibenden Regelabweichung nach einer Eingangsänderung oder beim Auftreten von Störungen kommt.

Das Blockschaltbild in Bild 7-64 a) zeigt den Gesamtregelkreis mit den Eingangs-

größten Sollwert W und Störgröße Z . Für dieses Blockschaltbild lautet die Gleichung der Laplace-Transformierten für die Ausgangsgröße:

$$X(s) = \frac{G_0(s)}{G_0(s)+1} W(s) + \frac{G_{sz}(s)}{G_0(s)+1} Z(s) \quad .$$

Betrachtet man zuerst das Führungsverhalten ($Z = 0$) für beispielsweise eine sprungförmige Änderung der Führungsgröße $W(s) = 1/s$, so erhält man für die Regeldifferenz:

$$X_d(s) = W(s) - X(s) = \left(\frac{G_0}{G_0+1} \right) \cdot W(s) + \frac{1}{G_0+1} \cdot \frac{1}{s} \quad .$$

Nach dem Endwertsatz der Laplace-Transformation gilt (Bild 7-54) :

$$\lim_{t \rightarrow \infty} x_d(t) = \lim_{s \rightarrow 0} s X_d(s) = \frac{1}{G_0(s)+1} \Big|_{s=0} \quad .$$

Für einen P-Regler und eine PT₁-Strecke hat die Übertragungsfunktion des offenen Regelkreises folgenden Wert:

$$G_0(s) = K_p \cdot \frac{K_S}{1+T \cdot s} \quad .$$

Setzt man dies in den Endwertsatz ein, so erhält man:

$$\lim_{t \rightarrow \infty} x_d(t) = \frac{1}{K_p K_S + 1} = \frac{1}{K_0 + 1} = R_S \quad .$$

Der Wert R_S wird **statischer Regelfaktor** genannt, K_0 ist die sog. **Kreisverstärkung**.

Bei gleicher Betrachtung des Systems für Störverhalten ($W = 0$), erhält man nach Einsetzen der Übertragungsfunktionen von Regler und Strecke einen ähnlichen Ausdruck:

$$\lim_{t \rightarrow \infty} x_d(t) = \frac{K_{sz}}{K_p K_S + 1} = \frac{K_{sz}}{K_0 + 1} = K_{sz} \cdot R_S \quad .$$

Man sieht, dass sowohl für das Führungsverhalten wie auch für das Störverhalten die Größe der bleibenden Regelabweichung nach einer sprungförmigen Änderung mit zunehmender Kreisverstärkung K_0 immer kleiner wird, aber nie verschwindet. Die Kreisverstärkung kann am Regler durch Vergrößerung von K_p erhöht werden, wobei aber beachtet werden muss, dass der Regelkreis dadurch nicht instabil wird. Dies wird im nächsten Abschnitt behandelt werden.

Während ein P-Regler gerätetechnisch nichts anderes als ein linear arbeitender Verstärker ist, benötigt man als Regler mit I-Verhalten ein System, bei dem die Ausgangsgröße dem Integral über der Eingangsgröße proportional ist. Im Beispiel in Bild 7-61 war gezeigt worden, dass man einen P-Verstärker dadurch erhält, dass man einen Teil der Ausgangsspannung eines Operationsverstärkers, der eine sehr hohe Verstärkung besitzt, auf den Eingang mit negativem Vorzeichen rückkoppelt (gegengekoppelt). Die Übertragungsfunktion der Rückkopplung war eine einfache P-

Verstärkung, so dass das dynamische Übertragungsverhalten der Gesamtschaltung nicht beeinflusst wurde. Eine weitere Möglichkeit, einen solchen Verstärker mit P-Verhalten aufzubauen ist die in Bild 7-82 a) gezeigte Verschaltung. Diese als **invertierender Verstärker** bezeichnete Schaltung führt sowohl die Eingangsspannung als auch den rückgekoppelten Anteil der Ausgangsspannung dem invertierenden Eingang zu, während der nichtinvertierende Eingang nahezu auf Bezugspotential liegt. Aus der Tatsache, dass wegen der hohen Verstärkung des Operationsverstärkers der Eingangsstrom nahezu Null sein muss, kann man sein Übertragungsverhalten zu

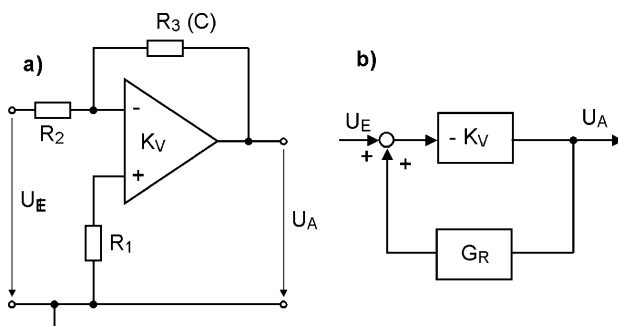


Bild 7-82: Rückgekoppelter, invertierender Verstärker
a) Stromlaufplan b) Blockschaltbild

bestimmen. Stellt man diese Schaltung als Blockschaltbild dar (Bild 7-82 b)), so ist die Übertragungsfunktion der Rückkopplung:

$$\frac{U_A}{U_E} = -\frac{R_3}{R_2} \quad (7.44)$$

und unter Verwendung der Regeln zur Vereinfachung von Blockschaltbildern und der Tatsache, dass $K_V \gg 1$ ist, erhält man damit die Übertragungsfunktion Gl. (7.44):

$$G_R = \frac{R_2}{R_3} \quad ,$$

Benutzt man nun in dieser Schaltung als Rückkopplung Bauelemente, deren Übertragungsfunktion Einfluss auf das dynamische Verhalten haben (Kondensatoren, Induktivitäten), so bekommt man ein anderes Gesamtübertragungsverhalten als das reine P-Verhalten. Ersetzt man den Widerstand R_3 in der Rückkopplung aus Bild 7-82 a) durch einen Kondensator mit der Kapazität C , so erhält man unter Verwendung der Übertragungsfunktion eines Kondensators als Gesamtübertragungsfunktion:

$$G(s) = \frac{U_A(s)}{U_E(s)} = \frac{-K_V}{1 + K_V \cdot \frac{R_2}{R_3}} = \frac{-1}{\frac{1}{K_V} + 1 \cdot \frac{R_2}{R_3}} = -\frac{R_3}{R_2} \quad .$$

Benutzt man nun in dieser Schaltung als Rückkopplung Bauelemente, deren Übertragungsfunktion Einfluss auf das dynamische Verhalten haben (Kondensatoren, Induktivitäten), so bekommt man ein anderes Gesamtübertragungsverhalten als das reine P-Verhalten. Ersetzt man den Widerstand R_3 in der Rückkopplung aus Bild 7-82 a) durch einen Kondensator mit der Kapazität C , so erhält man unter Verwendung der Übertragungsfunktion eines Kondensators als Gesamtübertragungsfunktion:

$$G(s) = \frac{U_A(s)}{U_E(s)} = \frac{-1}{\frac{1}{K_V} + 1 \cdot \frac{R_2}{R_3}} = -\frac{1}{R_2 \cdot C \cdot s} = -\frac{1}{T \cdot s} = -\frac{K_I}{s} \quad .$$

Dies ist die Übertragungsfunktion eines I - Gliedes und somit stellt das Schaltbild 7-82 a), mit einem Kondensator in der Rückkopplung, einen I - Regler dar. Wie schon

beim P-Regler gezeigt, kann man wieder mit Hilfe des Endwertsatzes der Laplace-Transformation die Regelabweichung bei Verwendung eines I-Reglers und einer Strecke mit PT_1 -Verhalten nach Aufgeben einer sprunghörmigen Eingangsgröße berechnen:

$$\lim_{t \rightarrow \infty} x_d(t) = \lim_{s \rightarrow 0} s \cdot X_d(s) = \frac{1}{G_0(s) + 1} \Big|_{s=0} = 0 \quad \text{mit} \quad G_0(s) = \frac{K_I}{s} \cdot \frac{1}{1 + T \cdot s}.$$

Dies bedeutet, dass nach einer sprunghörmigen Änderung der Führungsgröße oder auch nach Auftreten einer sprunghörmigen Störung keine bleibende Regelabweichung auftritt.

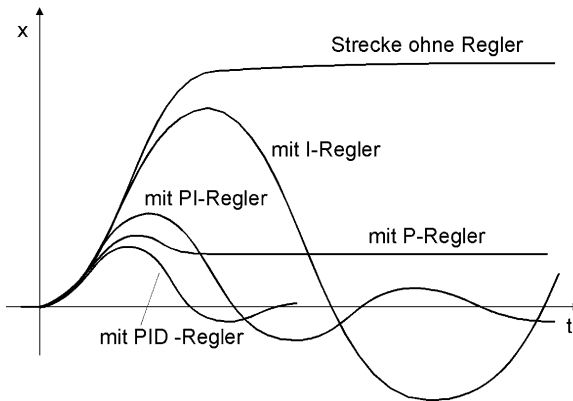


Bild 7-83: Vergleich des Regelverhaltens verschiedener Reglertypen

Vergleicht man nun insgesamt die dynamischen Eigenschaften von P- und I-Regler (Bild 7-83), so sieht man am Beispiel der Regelung einer Verzögerungsstrecke, dass der P-Regler verhältnismäßig gute dynamische Eigenschaften besitzt, da er auf eine sprunghörmige Störung schnell reagiert. Er verursacht aber starkes Überschwingen mit relativ schwacher Dämpfung und führt zu einer bleibenden Regelabweichung. Der I-Regler regelt zwar nach einer Störung die Regelabweichung wieder vollständig aus, verursacht aber noch

stärkeres Überschwingen und neigt zur Instabilität. Es liegt daher der Gedanke nahe, beide Reglerarten miteinander zu kombinieren, um die Vorteile von beiden zu nutzen. Wie man in Bild 7-83 sieht, ist das Überschwingen eines solchen kombinierten PI-Reglers nicht mehr so groß wie beim reinen I-Regler, die Dämpfung ist verbessert und es tritt keine bleibende Regelabweichung auf.

Erinnert man sich, dass ein Übertragungsglied mit D-Verhalten, beispielsweise auf eine sprunghörmige Eingangsgröße augenblicklich einen hohen Ausgangswert liefert (Delta-Dirac-Impuls), so ist leicht einsehbar, dass das Hinzufügen eines D-Anteils zu einem PI-Regler dessen dynamisches Verhalten noch weiter verbessern müsste. Das in Bild 7-83 dargestellte Regelverhalten des PID-Reglers bestätigt dies auch, der nun mit noch geringerem Überschwingen und guter Dämpfung bei verschwindender Regelabweichung reagiert. Da alle Übertragungstypen gleichzeitig wirksam sein müssen, müssen der P-, I- und D-Anteil im Blockschaltbild parallelgeschaltet werden, bzw. in der Übertragungsfunktion müssen die Anteile addiert werden. Unter Vernachlässigung der unvermeidlichen, verzögernden Anteile reiner P-, I- und D-Glieder lautet die DGL des PID-Reglers dann:

$$\underbrace{\dots + T_1 \cdot \dot{y}(t) + y(t)}_{\text{vernachlässigen}} = K_P \cdot x_d(t) + K_I \int_0^t x_d(t) dt + K_D \cdot \dot{x}_d(t)$$

bzw.

$$y(t) = K_P \left(x_d(t) + \frac{K_I}{K_P} \int_0^t x_d(t) dt + \frac{K_D}{K_P} \cdot \dot{x}_d(t) \right) .$$

Die in der DGL auftauchenden Konstanten werden nach DIN 19226 wie folgt definiert:

$$\frac{K_P}{K_I} = T_n \quad : \text{Nachstellzeit} \qquad \frac{K_D}{K_P} = T_v \quad : \text{Vorhaltezeit} .$$

Mit diesen Konstanten lautet die Übertragungsfunktion des idealen PID-Reglers:

$$G_R(s) = K_P \left(1 + \frac{1}{T_n} \cdot \frac{1}{s} + T_v \cdot s \right) . \quad (7.45)$$

P - I - D - Anteil

Da aber zumindest ein D-Glied ohne Verzögerung technisch nicht zu realisieren ist, lautet die Übertragungsfunktion eines realen PID-Reglers

$$G_R(s) = K_P \left(1 + \frac{1}{T_n} \cdot \frac{1}{s} + \frac{T_v \cdot s}{1 + T_1 \cdot s} \right) . \quad (7.46)$$

Zur Realisierung eines PID-Reglers kann man wieder den invertierenden Verstärker verwenden, bei dem die beiden Widerstände und durch Kombinationen von Widerständen und Kondensatoren, parallel und in Reihe geschaltet, ersetzt werden (Bild 7-84). Da der Eingangsstrom in den Operationsverstärker ungefähr Null sein muss, gilt $I_E = -I_R$. Für die Ströme gelten folgende Beziehungen:

$$I_E = \frac{U_E}{R_2} + C_2 \cdot \dot{U}_E$$

$$I_R = \frac{I_R \cdot R_3}{R_3} = C_3 \cdot \frac{d}{dt} (U_A - I_R \cdot R_3) .$$

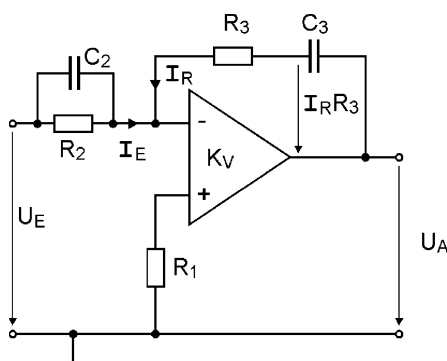


Bild 7-84: PID-Regler mit Operationsverstärker

Durch Anwendung der Laplace-Transformation erhält man daraus:

$$I_E(s) = \left(\frac{1}{R_2} + C_2 \cdot s \right) \cdot U_E(s)$$

$$\begin{aligned} I_R(s) &= C_3 \cdot s \cdot U_A(s) - R_3 \cdot C_3 \cdot s \cdot I_R(s) \\ &= \frac{C_3 \cdot s \cdot U_A(s)}{1 + R_3 \cdot C_3 \cdot s} . \end{aligned}$$

Da der Eingangsstrom und der Rückkoppelstrom wie oben erwähnt betragsmäßig gleich sein müssen, folgt daraus:

$$(1 + R_2 \cdot C_2 \cdot s) U_E(s) = \frac{R_2 \cdot C_3 \cdot s}{1 + R_3 \cdot C_3 \cdot s} U_A(s) \quad .$$

womit man die Übertragungsfunktion erhält:

$$G(s) = \frac{U_A(s)}{U_E(s)} = - \frac{(1 + R_2 \cdot C_2 \cdot s)(1 + R_3 \cdot C_3 \cdot s)}{R_2 \cdot C_3 \cdot s}$$

$$= - \frac{R_2 C_2 + R_3 C_3}{R_2 C_3} \left(1 + \frac{1}{R_2 C_2 + R_3 C_3} \cdot \frac{1}{s} + \frac{R_2 R_3 C_2 C_3}{R_2 C_2 + R_3 C_3} \cdot s \right) \quad .$$

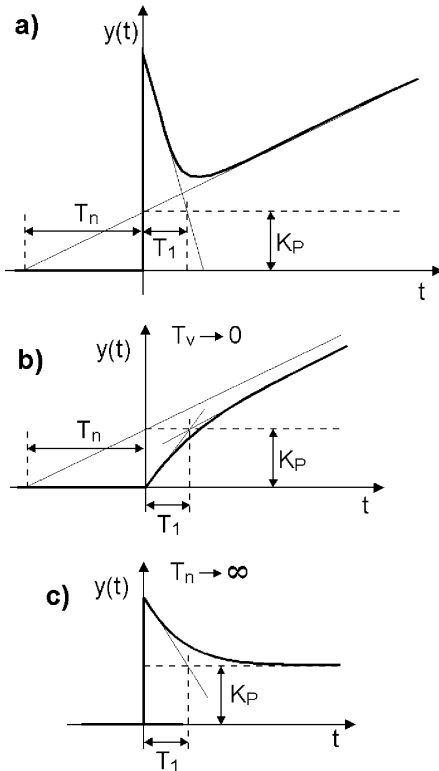


Bild 7-85: Übergangsfunktion eines PID-Reglers bei verschiedenen Einstellungen

Dies entspricht der Übertragungsfunktion des PID-Reglers aus Gl. (7.45). Durch Wahl der Widerstände und Kondensatoren kann man in dieser Schaltung den P-, I- und D-Anteil des Reglers festlegen.

Die Übergangsfunktion eines PID-Reglers, also die Antwort auf eine sprunghafte Änderung der Eingangsgröße, ist in Bild 7-85 a) dargestellt. Da wie gesagt ein D-Anteil ohne Verzögerung technisch nicht realisiert werden kann, ist die Höhe der Sprungantwort zum Zeitpunkt $t = 0$ auf einen bestimmten Wert $y(t)$ begrenzt und es tritt die Zeitkonstante T_1 des Verzögerungsgliedes auf.

Stellt man die Vorhaltezeit T_v eines solchen Reglers auf Null, dann verschwindet der D-Anteil des Reglers und es ergibt sich das Verhalten eines PI-Reglers (Bild 7-85 b). Ebenso führt ein Wert der Nachstellzeit $T_n = \text{Max.}$ ($T_n = \infty$) dazu, dass der I-Anteil des Reglers verschwindet und ein PD-Regler entsteht.

Moderne Regler werden nicht mehr in Analogtechnik mit beschalteten Operationsverstärkern, wie oben dargestellt, aufgebaut, sondern enthalten in der Regel Digitalrechner in Form von Mikrorechnern. Auf diesen werden die oben beschriebenen Regelalgorithmen digital simuliert und es können daher alle Regelparameter in weiten Grenzen frei gewählt werden.

Im Unterschied zu analog arbeitenden Systemen, auf die ein **zeitkontinuierliches** Eingangssignal wirkt, benötigt man für eine digitale Regelung ein **zeitdiskretes** Eingangssignal $x_e(t)$, das normalerweise in äquidistanten Intervallen T_0 abgefragt wird. Diese Diskretisierung wird als **Abtastung** bezeichnet, bei der eine Folge von Signalwerten $x_e(k \cdot T_0)$ entsteht. Benötigt das nachgeschaltete Übertragungsglied wieder zeitkontinuierliche Signale, so erfolgt eine Rückwandlung mittels eines Halteglie-

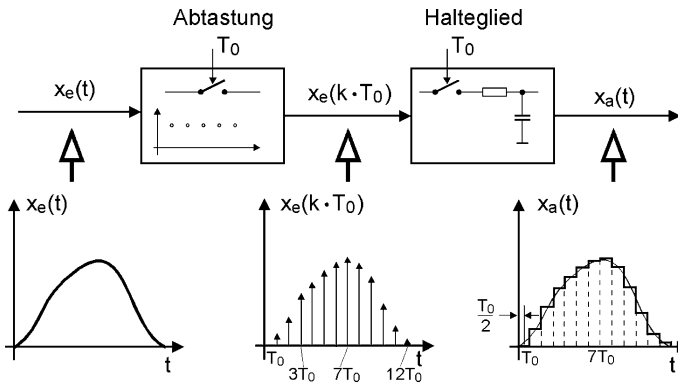


Bild 7-86: Der Abtast- und Haltevorgang

des. Dieser Abtast- und Haltevorgang (**sample and hold**) ist in Bild 7-86 dargestellt [7.10]. Das Halteglied erzeugt aus einem Abtastimpuls $x_e(k \cdot T_0)$ durch Speicherung des Abtastwertes in einem Kondensator eine betragskonstante Ausgangsgröße $x_a(t)$ und hält diese bis zum Eintreffen des nächsten Abtastimpulses $x_e[(k+1) \cdot T_0]$ kon-

stant. Es entsteht ein treppenförmiger Verlauf der Ausgangsgröße, deren äquivalenter kontinuierlicher Verlauf um $T_0/2$ zur Eingangsfunktion zeitlich versetzt ist.

Aufgrund dieser Tatsache muss man die regelungstechnische Behandlung zeitdiskreter Systeme im Hinblick auf das dynamische Übertragungsverhalten und die Stabilitätsanalyse unterscheiden. Ist bei Systemen mit Ausgleich die Ausgleichszeit mehr als zehnfach so groß wie die Abtastzeit $T_0 (T_g > 10 T_0)$, so hat der Abtastprozess keinen wesentlichen Einfluss auf die Gesamtdynamik des Systems. Diese so genannten quasikontinuierlichen Abtastsysteme können wie kontinuierliche Systeme behandelt werden. Das Übertragungsverhalten von Abtast- und Halteglied geht als in Reihe geschaltetes Totzeitglied mit $T_t = T_0/2$ ein.

Gilt für die Verhältnisse zwischen Abtastzeit und Ausgleichzeit $T_g < 10 T_0$, so hat der Abtastvorgang wesentlichen Einfluss auf die Dynamik des Gesamtsystems. Zur Behandlung müssen andere Methoden wie die aus der Laplace-Transformation hergeleitete Übertragungsfunktion eingesetzt werden. Die für die Behandlung solcher Probleme erforderliche **z-Transformation** ermöglicht wie die Fourier-Transformation und die Frequenzgangdarstellung stetiger Systeme eine vereinfachte mathematische Beschreibung diskret arbeitender Systeme. Auf diese Transformation kann hier nicht weiter eingegangen werden [7.11], [7.12].

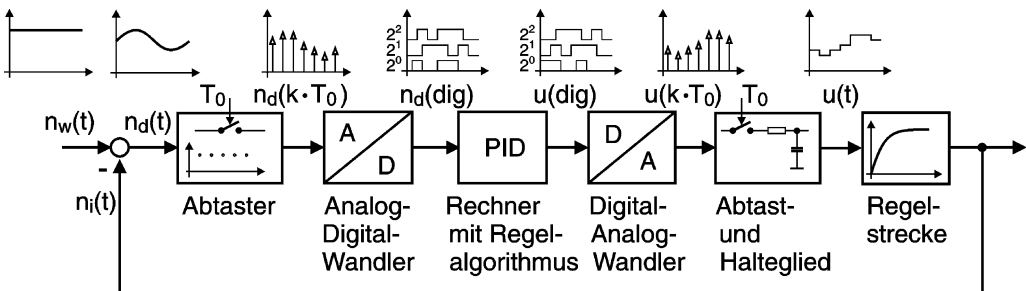


Bild 7-87: Drehzahlregelung eines Gleichstrommotors mit Abtastregelung

Als Beispiel für eine Abtastregelung ist in Bild 7-87 die Drehzahlregelung eines Gleichstrommotors dargestellt. Der dem Drehzahlregelkreis vorgegebene Drehzahl-

sollwert $n_w(t)$ wird mit dem Drehzahlwert $n_i(t)$ als analoge Größe verglichen und so die Regelabweichung $n_d(t)$ erzeugt. Diese Größe wird in gleichmäßigen Intervallen T_0 abgetastet, wodurch diskrete Signalwerte $n_d(k \cdot T_0)$ entstehen. Für die Durchführung des PID-Algorithmus' auf dem Mikrorechner benötigt dieser den jeweiligen Abtastwert als digitales Signal. Daher wird der Abtastwert durch einen Analog-Digital-Wandler in eine Dualzahl $n_d(dig)$ gewandelt. Der im Mikrorechner daraus ermittelte Stellwert $u(dig)$ ist wieder ein digitaler Wert, wohingegen der Gleichstrommotor als Stellsignal eine analoge Spannung benötigt. Daher wird $u(dig)$ durch einen Digital-Analog-Wandler in eine analoge Spannung umgesetzt. Da der Stellwert für das nächste Zeitintervall der digitalen Regelung vom Mikrorechner nur kurzzeitig ausgegeben wird, ist hinter dem D-/A-Wandler noch ein Abtast- und Halteglied geschaltet, das einen treppenförmigen Spannungsverlauf an den Gleichstrommotor bzw. an das hier nicht gezeigte Leistungsstellglied des Motors ausgibt.

Darüber hinaus gibt es andere Reglertypen, die nicht das Ein-/Ausgangsverhalten des Regelsystems in Bezug auf eine Regelgröße beeinflussen, sondern vom Zustandsraum-Modell ausgehen. Dies sind Regler mit andersartiger Struktur, die die sog. Zustandsrückführung verwenden [7.12]. Weitere Reglertypen müssen der Fachliteratur entnommen werden.

7.3.1.8 Stabilität von Regelkreisen

In Kapitel 4 ist bereits das Schwingungsverhalten von Systemen behandelt worden. Auch Regelsysteme aus Strecke und Regler können je nach Systemverhalten stabil arbeiten oder instabil werden. Bei stabilem Systemverhalten klingen Systemerregungen durch Störungen oder Änderungen der Eingangsgröße entweder aperiodisch auf einen stabilen Endwert ab, oder das System kann in mehr oder weniger stark gedämpften Schwingungen sich einem Endwert nähern. Bei den in Kapitel 4 beispielhaft behandelten mechanischen Schwingungen tritt selten instabiles Schwingverhalten auf, das ist das Anfachen einer Ausgangsschwingung mit zunehmender Amplitude. Dies setzt nämlich voraus, dass dem schwingenden System aus einer Energiequelle ständig neue Energie zugeführt werden kann. Allerdings kann ein solcher Vorgang auch in mechanischen Systemen in Form einer sog. **Galopping-Schwingung** auftreten. Ein Beispiel dafür wäre das Auftreten einer aufklingenden Schwingung bei durch Windkräfte angeregten Freileitungen, die zur Zerstörung führen können. In Regelsystemen steht als Energiequelle, die eine erregte Schwingung anfachen kann, der Regler selber zur Verfügung, so dass je nach Systemdämpfung und Lage der komplexen Eigenwerte der Systemdifferentialgleichung, ein instabiles aufschwingendes Verhalten auftreten kann. Ein solches Verhalten muss natürlich vermieden werden und in den meisten Fällen ist es wünschenswert, dass bei einer Störung oder Änderung der Führungsgröße, sich die Ausgangsgröße $x_a(t)$ möglichst ohne Schwingvorgang schnell aperiodisch einem neuen Endwert nähert. Je nach Art und Eigenschaften des Reglers kann dieser prinzipiell eine stabile Strecke im geschlossenen Regelkreis zur Instabilität führen oder andererseits auch eine instabile Strecke stabilisieren.

Kann man das System mit einer linearen DGL beschreiben, so setzt sich die Lösung der DGL für den untersuchten Erregungsfall aus der Lösung der homogenen

DGL und einer partikulären Lösung zusammen:

$$x_a(t) = x_{ah}(t) + x_{ap}(t) \quad .$$

Die Lösung der homogenen DGL beschreibt dabei den Eigenvorgang des Systems, der für stabiles Verhalten nach endlicher Zeit abklingen muss. Diese Lösung der allgemeinen linearen homogenen DGL mit konstanten Koeffizienten

$$b_n x_{ah}^{(n)}(t) + \dots + b_1 \dot{x}_{ah}(t) + b_0 x_{ah}(t) = 0$$

mit der charakteristischen Gleichung

$$b_n s^n + \dots + b_1 s^1 + b_0 = 0$$

lautet wie in Kapitel 4 gezeigt in allgemeiner Form:

$$x_{ah}(t) = \underbrace{\sum_j C_j \cdot e^{\delta_j t}}_{\text{reelle Eigenwerte}} + \underbrace{\sum_k e^{\delta_k t} (A_k \cdot e^{i\omega_k t} + B_k \cdot e^{-i\omega_k t})}_{\text{konjugiert komplexe Eigenwerte}} \quad .$$

Dabei haben die Indizes folgende Werte:

$$j = 1, 2, \dots, u \quad k = 1, 2, \dots, v \quad n = u + 2v \quad .$$

Schaut man sich diese Gleichung an, so sieht man, dass sowohl die reellen als auch die komplexen Eigenwerte jeweils einen Dämpfungsfaktor $e^{\delta t}$ enthalten. Eine Schwingung klingt daher ab ($x_a(t)$ wird in endlicher Zeit zu Null), wenn sowohl alle δ_j der reellen Eigenwerte, als auch alle δ_k der komplexen Eigenwerte negativ sind. Eine Stabilitätsuntersuchung wird daher darin bestehen können festzustellen, ob alle Eigenwerte in der linken Hälfte der komplexen Zahlenebene liegen. Um dies zu tun, müssen die Wurzeln der charakteristischen Gleichung bestimmt werden, die die Eigenwerte der homogenen Differentialgleichung darstellen.

Bei einem Eigenvorgang eines Regelkreises gilt: $w = z = 0$. Dies bedeutet, dass das Verhalten des Systems ohne Stör- oder Führungsgrößen betrachtet wird. Sieht man sich das Blockschaltbild in Bild 7-64 a) an, für das in Gl. (7.29) die Führungsübertragungsfunktion ($z = 0$) abgeleitet wurde und setzt in dieser Gleichung ebenfalls $w = 0$, so folgt für den Eigenvorgang eines geschlossenen Regelkreises:

$$G_0(s) + 1 = 0 \quad \text{mit} \quad G_0(s) = \frac{Q_0(s)}{R_0(s)} \quad \text{Übertragungsfunktion des } \underline{\text{offenen}} \text{ Regelkreises.}$$

Daraus kann man die charakteristische Gleichung des geschlossenen Regelkreises mit Hilfe des Zähler- und Nennerpolynoms der Übertragungsfunktion des offenen Regelkreises formulieren:

$$Q_0(s) + R_0(s) = 0 \quad . \quad (7.47)$$

Im folgenden soll ein Beispiel für eine Stabilitätsuntersuchung durchgeführt werden. Es sei eine Regelstrecke mit Verzögerungsverhalten 3. Ordnung (PT₃-Verhalten) und der Übertragungsfunktion

$$G_0(s) = \frac{1}{(s+1)(s+2)(s+3)} = \frac{1}{s^3 + 6s^2 + 11s + 6} \quad ,$$

sowie ein P-Regler mit der Übertragungsfunktion

$$G_R(s) = K_P = 80$$

gegeben. Die Übertragungsfunktion des offenen Regelkreises ergibt sich als Produkt der Übertragungsfunktionen von Strecke und Regler zu

$$G_0(s) = \frac{80}{s^3 + 6s^2 + 11s + 6} \quad .$$

Die charakteristische Gleichung des offenen Regelkreises hat damit folgendes Aussehen:

$$s^3 + 6s^2 + 11s + 6 = 0$$

und besitzt die Eigenwerte (Nullstellen)

$$s_1 = \delta_1 = -1 \quad s_2 = \delta_2 = -2 \quad s_3 = \delta_3 = -3 \quad .$$

Dies sind drei reelle negative Eigenwerte, so dass das offene System aus Strecke und Regler stabil sein muss. Schließt man die Rückkopplung, so lautet die charakteristische Gleichung entsprechend Gl. (7.47):

$$s^3 + 6s^2 + 11s + 86 = 0 \quad .$$

Die Eigenwerte dieser charakteristischen Gleichung dritten Grades lauten:

$$s_1 = \delta_1 = -6,39 \quad s_{23} = \delta_{23} \pm i\omega_{23} = 0,193 \pm i \cdot 3,66 \quad .$$

Da nun der Realteil der beiden konjugiert komplexen Eigenwerte positiv ist, muss das rückgekoppelte System aus Strecke und Regler instabil geworden sein, d. h. das System schwingt ohne äußere Einflüsse auf und führt eine Dauerschwingung aus.

Reduziert man nun die Verstärkung des P-Reglers auf den Wert 10, so ergeben sich folgende Eigenwerte der charakteristischen Gleichung des geschlossenen Regelkreises:

$$s_1 = \delta_1 = -4,313 \quad s_{23} = \delta_{23} \pm i\omega_{23} = -0,843 \pm i \cdot 1,73 \quad .$$

Wie man sieht, sind nun wieder alle Realteile der Eigenwerte negativ, weshalb dieses System bei reduzierter Verstärkung stabil ist. Es gibt also eine kritische Verstärkung zwischen **10** und **80**, bei der das System instabil wird.

Wendet man die oben aufgezeigte Methode zur Stabilitätsuntersuchung an, so muss man die charakteristische Gleichung lösen und deren Nullstellen bestimmen. Bei einer Gleichung zweiten Grades ist das leicht möglich, allgemeine Lösungen für Gleichungen dritten und vierten Grades sind bekannt, ihre Berechnung ohne Rechner aber schon sehr umständlich. Für Gleichungen höheren Grades gibt es nur numerische Näherungsverfahren, die ebenso wie Nullstellen-Suchverfahren sinnvoll nur mit einem Rechner durchgeführt werden können. Es sind daher in der Vergangenheit eine größere Anzahl von alternativen Methoden und Kriterien entwickelt worden, mit denen man Systeme auf Stabilität untersuchen kann ohne die charakteristische Gleichung zu lösen. Dabei unterscheidet man Verfahren, die die Koeffi-

zienten der charakteristischen Gleichung untersuchen und **algebraische Stabilitätskriterien** genannt werden, sowie Verfahren, die die Stabilitätsuntersuchung anhand der Ortskurve ermöglichen und **Ortskurvenkriterien** heißen.

Ein häufig genanntes und bereits 1895 von dem Mathematiker A. Hurwitz entwickeltes algebraisches Kriterium zur Stabilitätsuntersuchung ist das **Hurwitz-Kriterium**. Für charakteristische Polynome niedriger Ordnung (2., 3., 4., 5. Ordnung) liefert es mit vertretbarem Aufwand Antworten auf die Stabilitätsfrage. Allerdings ist es für Polynome höherer Ordnung schlecht geeignet, da dann eine Vielzahl von Determinanten gelöst werden müssen. Wie oben ausgeführt, liegt Stabilität vor, wenn die Realteile aller Eigenwerte negativ sind. Man kann zeigen, dass in einem solchen Fall alle Koeffizienten des Polynoms vorhanden sind und gleiches Vorzeichen haben müssen. Diese Bedingung ist zwar notwendig aber nicht hinreichend für Stabilität. Es kann Polynome geben, bei denen alle Koeffizienten vorhanden und positiv sind, deren zugehöriger geschlossener Regelkreis aber instabil ist.

Das Hurwitz-Kriterium geht nun von der notwendigen Bedingung für Stabilität als Voraussetzung aus und formuliert Koeffizientendeterminanten, die existieren und größer als Null sein müssen.

Für die DGL 3. Ordnung

$$b_3 s^3 + b_2 s^2 + b_1 s + b_0 = 0$$

lautet die Stabilitätsbedingung nach Hurwitz

$$\begin{vmatrix} b_1 & b_3 \\ b_0 & b_2 \end{vmatrix} = b_1 \cdot b_2 - b_0 \cdot b_3 > 0 \quad .$$

Die Stabilitätsgrenze ist gerade für den Fall

$$b_1 \cdot b_2 - b_0 \cdot b_3 = 0$$

erreicht, bei der das System eine Dauerschwingung mit der kritischen Kreisfrequenz

$$\omega_{krit} = \sqrt{\frac{b_1}{b_3}}$$

ausführt.

Bei Polynomen höherer Ordnung müssen dann mehrere Koeffizientendeterminanten als Stabilitätsbedingung vorhanden und größer Null sein. So gilt für ein charakteristisches Polynom 4. Ordnung:

$$\begin{vmatrix} b_1 & b_3 \\ b_0 & b_2 \end{vmatrix} = b_1 \cdot b_2 - b_0 \cdot b_3 > 0 \quad \text{und}$$

$$\begin{vmatrix} b_1 & b_3 & b_5 \\ b_0 & b_2 & b_4 \\ 0 & b_1 & b_3 \end{vmatrix} = \begin{vmatrix} b_1 & b_3 & 0 \\ b_0 & b_2 & b_4 \\ 0 & b_1 & b_3 \end{vmatrix} = (b_1 \cdot b_2 - b_0 \cdot b_3) b_3 - b_1^2 b_4 > 0 \quad .$$

Bei Überschreiten der Stabilitätsgrenze entsteht wiederum eine Dauerschwingung mit der gleichen kritischen Kreisfrequenz wie beim Polynom 3. Ordnung. Dieses Kriterium kann dann entsprechend auf ein Polynom n -ter Ordnung erweitert werden, wobei $n - 1$ Determinanten größer Null sein müssen.

Als Beispiel soll nun nochmals das geschlossene System, das bereits untersucht wurde, für eine Reglerverstärkung $K_P = 40$ untersucht werden. Sein charakteristisches Polynom lautet dann:

$$s^3 + 6s^2 + 11s + 46 = 0 \quad .$$

Die notwendige aber nicht hinreichende Bedingung für Stabilität, dass alle Koeffizienten vorhanden sind und gleiches Vorzeichen haben, ist erfüllt. Die für die Stabilitätsuntersuchung zu lösende Hurwitz-Determinante lautet:

$$\begin{vmatrix} 11 & 1 \\ 46 & 6 \end{vmatrix} = 11 \cdot 6 - 46 \cdot 1 = 20 > 0 \quad .$$

Für die Verstärkung von $K_P = 40$ ist also der geschlossene Regelkreis stabil. Interessant ist natürlich auch, wo die Stabilitätsgrenze liegt. Für diesen Fall lautet das charakteristische Polynom

$$s^3 + 6s^2 + 11s + 6 + K_{P_{krit}} = 0 \quad .$$

In die Hurwitz-Determinante eingesetzt ergibt sich daraus:

$$\begin{vmatrix} 11 & 1 \\ 6 + K_{P_{krit}} & 6 \end{vmatrix} = 11 \cdot 6 - (6 + K_{P_{krit}}) \cdot 1 = 0 \Rightarrow K_{P_{krit}} = 60 \quad .$$

Ab einer Reglerverstärkung von $K_P = 60$ entsteht im Regelkreis demnach ohne äußere Anregung eine Dauerschwingung mit der Kreisfrequenz von

$$\omega_{krit} = \sqrt{\frac{11}{1}} = 3,32 \frac{\text{rad}}{\text{sec}} \quad .$$

Weiterhin ist häufig die Aussage wichtig, wie groß die Stabilitätsgüte oder **Stabilitätsreserve** ist, d. h. wie weit man von der Stabilitätsgrenze entfernt ist und wie hoch daher die Dämpfung δ des Regelkreises ist. Liegt Stabilität vor, so haben ja alle Eigenwerte einen Abstand von $-k$ von der imaginären Achse. Setzt man eine Dämpfung $\delta = -k$ voraus, so kann man für diesen Fall die Stabilität prüfen, indem man die Stabilitätsbedingung für das Polynom mit $s^* = s + k$ prüft. Weitere Stabilitätskriterien, die sich teilweise auch besser in rechnergestützten Verfahren anwenden lassen, finden sich in [7.11].

Variiert man einen Stabilitätsparameter wie die Verstärkung K_P , so wandern die Eigenwerte oder Wurzeln des charakteristischen Polynoms des geschlossenen Regelkreises auf bestimmten Wegen durch die komplexe Ebene. Dabei ergeben sich Kurven, auf denen sich die Wurzelorte bewegen, die daher auch **Wurzelortskurven** genannt werden. Kennt man die Lage und den Verlauf der Wurzelortskurve eines Systems, so kann man am Durchstoßpunkt der Wurzelortskurve den Stabilitätsrand beurteilen und allgemein Aussagen über die Stabilität machen.

Bei dem schon mehrfach untersuchten Beispiel kennen wir bereits die Lage der Wurzeln (Pole der Übertragungsfunktion des geschlossenen Systems) für die Fälle $K_p = 0, 10, 80$. Für die ebenfalls behandelten Fälle liegen sie bei:

$$K_p = 40 : s_1 = -5,517 \quad s_{2,3} = -0,242 \pm i \cdot 2,875$$

$$K_p = 60 : s_1 = -6 \quad s_{2,3} = \pm i \cdot 3,317$$

Die letzten Wurzeln bestätigen nochmals die Stabilitätsgrenze für $K_p = 60$, da hier die Realteile der konjugiert komplexen Wurzel verschwinden. Mit diesen Werten kann man nun die in Bild 7-88 dargestellte Wurzelortskurve zeichnen.

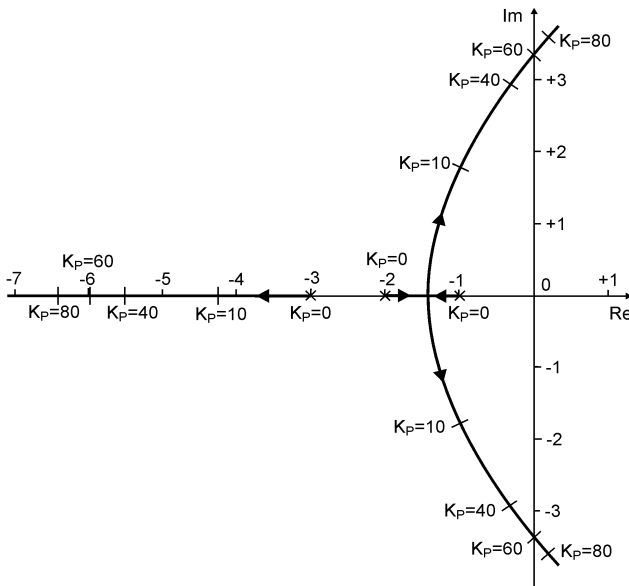


Bild 7-88: Wurzelortskurve für das charakteristische Polynom $s^3 + 6s^2 + 11s + 6 + K_p = 0$

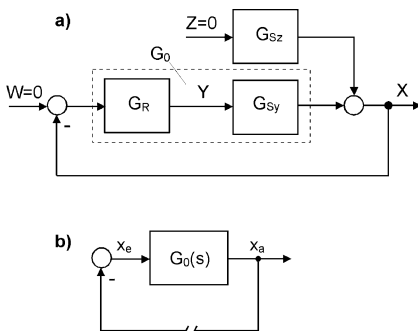


Bild 7-89: a) geschlossener und b) aufgeschnittener Regelkreis für $w = z = 0$

Ein ebenfalls häufig benutztes **grafisches Verfahren**, das die Stabilität anhand der Ortskurve ermittelt, ist das 1932 von H. Nyquist beschriebene **Nyquist-Kriterium**. Die einfachste Fassung dieses Kriteriums wird im folgenden behandelt. Diese einfache Form behandelt nur Regelsysteme, deren Übertragungsfunktion des offenen Kreises $G_0(s)$ nur Pole mit negativem Realteil besitzt und höchstens einen reellen Pol im Ursprung ($s = 0$), d. h. der offene Regelkreis ist stabil.

Folgende anschauliche Betrachtung, die man mit Hilfe der Funktionentheorie auch exakt beweisen kann, soll die

Bedeutung des Nyquist-Kriteriums erläutern. Betrachtet man den geschlossenen Regelkreis mit der Übertragungsfunktion $G_0(s)$ von Regler und Strecke in Bild 7-89 a) für den Fall $w = z = 0$ und trennt die Rückführung, wie in Bild 7-89 b) gezeigt auf, so wird für die Veranschaulichung der Fall betrachtet, dass das Eingangssignal in $G_0(s)$

$$x_e = A \cdot \sin \omega t \quad (7.48)$$

beträgt. Dieser Fall ist interessant für die Betrachtung, da es an der Grenze zwischen stabilen Eigenvorgängen mit abnehmender Amplitude und instabilen Eigenvorgängen mit wachsender Amplitude liegt. Für die Ausgangsgröße x_a aus dem Übertragungsblock mit der Übertragungsfunktion $G_0(s)$ und dem zugehörigen Frequenzgang $F_0(i\omega)$ gilt dann:

$$x_a = |F(i\omega_0)| \cdot A \cdot \sin(\omega t + \varphi) \quad .$$

Wäre jetzt x_a gerade gleich x_e , so braucht man die Eingangsschwingung gar nicht mehr aufzugeben, da sich diese in der Rückkopplenschleife selbsttätig aufrechterhalten würde. Der Fall

$$x_e = -x_a$$

bedeutet für eine Sinusschwingung aber gerade eine Phasenverschiebung um den Winkel π , so dass gelten muss

$$x_a = |F(i\omega_0)| \cdot A \cdot \sin(\omega t + \varphi + \pi) \quad . \quad (7.49)$$

Da, wie oben gesagt, die ursprüngliche Schwingung nach Gl. (7.48) mit der rückgeführten nach Gl. (7.49) identisch sein muss, um ohne Veränderung die Rückkopplung schließen zu können, kann dies nur für den Fall

$$\varphi + \pi = 0 \Rightarrow \varphi = -\pi$$

und zusätzlich

$$|F(i\omega_0)| \cdot A = A \Rightarrow |F(i\omega_0)| = 1$$

zutreffen.

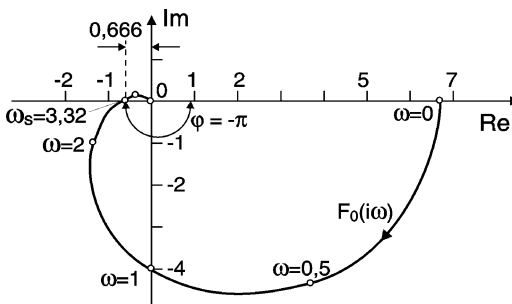


Bild 7-90: Ortskurve für das Beispiel bei $K_P = 40$

Das Schließen der Rückkopplung ist daher nur bei dieser speziellen **Durchtrittsfrequenz** möglich. Da die Stabilitätsgrenze gerade für

$$|F(i\omega_0)| = 1$$

erreicht wird, muss für stabile Systeme (abklingende Schwingung)

$$|F(i\omega_0)| < 1$$

gelten. Schaut man sich die Ortskurve in Bild 7-90 an, von deren zugehörigem geschlossenen Regelkreis wir bereits die Stabilität nachgewiesen haben, so sieht man, dass beim Schnitt der Ortskurve mit der reellen Achse ($\varphi = -\pi$) der Betrag der Übertragungsfunktion des offenen Regelkreises kleiner als 1 ist. Dies bedeutet nach dem oben gesagten Stabilität, da eine auftretende Eigenschwingung dann abklingt. Um die Ortskurve zu erhalten, muss man Real- und Imaginärteil des aus der Übertragungsfunktion abgeleiteten Frequenzgangs bestimmen:

Betrachtet man dazu beispielweise die Ortskurve des bereits oben behandelten Verzögerungssystems dritter Ordnung bei einer Verstärkung von $K_P = 40$ (Bild 7-90), so bedeutet ein Phasenwinkel von $\varphi = -\pi$ den Punkt, an dem die Ortskurve die negative, reelle Achse schneidet. Der an der Ortskurve stehende Parameter ω hat hier den Wert $\omega = \omega_s$. Das Schließen der Rückkopplung ohne Änderung der Dauerschwingung

$$\begin{aligned}
 F_0(i\omega) &= \frac{40}{(1+i\omega)(2+i\omega)(3+i\omega)} = \frac{40}{(6-6\omega^2)+i(11\omega-\omega^3)} \\
 &= \frac{40[(6-6\omega^2)-i(11\omega-\omega^3)]}{(6-6\omega^2)^2+(11\omega-\omega^3)^2} \\
 \Re\{F_0(i\omega)\} &= \frac{40(6-6\omega^2)}{(6-6\omega^2)^2+(11\omega-\omega^3)^2} \\
 \Im\{F_0(i\omega)\} &= \frac{-40(11\omega-\omega^3)}{(6-6\omega^2)^2+(11\omega-\omega^3)^2} .
 \end{aligned}$$

Für die Durchtrittsfrequenz muss gelten:

$$\Im\{F_0(i\omega)\} = 0 \Rightarrow \omega_s(11-\omega_s^2) = 0 \Rightarrow \omega_s^2 = 11 \Rightarrow \omega_s = 3,317 .$$

Mit dieser Frequenz kann man die Amplitude für $\varphi = -\pi$ berechnen:

$$A(\omega_s) = |\Re\{F_0(i\omega)\}| = 0,666 .$$

Da dies kleiner als 1 ist, muss das System, wie bereits vorher gezeigt, stabil sein. Unter den oben genannten Voraussetzungen und für Systeme, deren Ortskurve für unendlich hohe Frequenzen im Nullpunkt endet (gilt für die meisten einschleifigen Regelkreise), kann man das Nyquist-Kriterium wie folgt formulieren:

Ein geschlossener Regelkreis verhält sich dann stabil, wenn beim Durchlaufen der Ortskurve $F_0(i\omega)$ im Sinne zunehmender Frequenz der kritische Punkt der komplexen Ebene $(-1, \pm i0)$ stets zur Linken liegt.

Da die Frequenzkennlinien nur eine andere Darstellungsart der Ortskurve sind, kann das Nyquist-Kriterium auch entsprechend dort angewendet werden. Für weitere Details sei auf die Literatur verwiesen.

Im folgenden soll die Analyse und die Reglerauslegung eines mechatronischen Systems an einem Beispiel aus [7.8] nochmals ausführlich dargestellt werden. Die Aufgabe besteht darin, die in Bild 7-91 a) abgebildete magnetische Wägeeinrichtung für ein Gewicht der Masse m auf Stabilität zu untersuchen und einen Regler auszulegen, der die gestellte Regelungsaufgabe ermöglicht. Der Wägevorgang beruht darauf, dass unter der Voraussetzung, dass der Regler die Spaltbreite H zwischen der zu wiegenden Masse und der Magnetspule

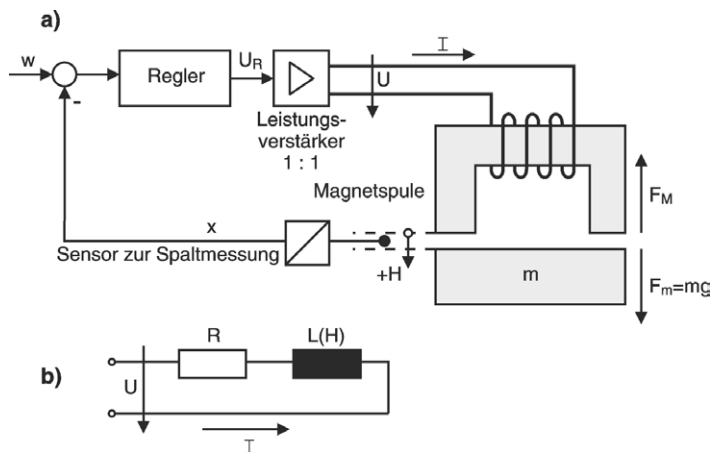


Bild 7-91: Magnetische Wägeeinrichtung a) Blockschaltbild
b) Ersatzschaltbild der Magnetspule

konstant hält, man aus einer Messung des Stromes I durch die Spule auf die Masse zurückschließen kann. Die Magnetkraft F_M ist dabei dem Strom durch die Magnetspule proportional. Für eine fehlerfreie Wägung ist daher eine stabile Regelung der Spaltbreite H erforderlich.

1.) Berechnung des mathematischen Modells der Regelstrecke

Die Induktivität der Spule ist umgekehrt proportional zur Spalthöhe H , wobei die Anfangsinduktivität bei Spalthöhe 1 mm durch die Konstante C_1 gegeben ist:

$$L(H) = \frac{C_1}{H + C_2} \quad C_1 = 5,52 \cdot 10^{-1} \frac{\text{V} \cdot \text{s} \cdot \text{mm}}{\text{A}} \quad C_2 = 1 \text{ mm} \quad (7.50)$$

Die Magnetkraft F_M der Spule lässt sich aus der Gleichung für die Energie eines Magnetfeldes

$$W = \frac{1}{2} L \cdot I^2 = F \cdot s$$

ermitteln:

$$F_M \cdot s = \frac{1}{2} L(H) \cdot I^2 \Rightarrow F_M = \frac{1}{2} L(H) \cdot \frac{I^2}{H + C_2} = 0,5 \frac{I^2 C_1}{(H + C_2)^2} \quad \text{mit } s = H + C_2 \quad (7.51)$$

Der Weggeber für die Spalthöhe arbeitet proportional:

$$x = C_3 \cdot H \quad C_3 = 2,5 \text{ V/mm} \quad (7.52)$$

Als weitere Angaben sind bekannt:

$$m = 50 \text{ kg}; \quad R = 15 \Omega; \quad \text{angestrebter Arbeitspunkt: } H_0 = 2 \text{ mm}.$$

Zur Aufstellung der DGL für den elektromagnetischen Anteil der Strecke kann man die Kirchhoffsche Maschenregel verwenden. Für die Spannung über der Induktivität ist zu beachten, dass sie nicht wie im Normalfall mit

$$U = L \cdot I \quad \text{mit } L = \text{const.}$$

angenommen werden kann, da sie mit der Spalthöhe H variiert:

$$U = R \cdot I + \frac{d}{dt}(L(H) \cdot I) \quad (7.53)$$

Die beiden Variablen in dem abzuleitenden Klammerausdruck sind beide von der Zeit abhängig, zusätzlich ist die Induktivität von der Spalthöhe abhängig:

$$L(H) = f(t, H).$$

Es muss für diese Variable daher die totale Ableitung gebildet werden:

$$\frac{dL(H)}{dt} = \frac{\partial L(H)}{\partial t} + \frac{\partial L(H)}{\partial H} \cdot \frac{dH}{dt}.$$

In diesem Ausdruck kann man die partielle Ableitung der Spuleninduktivität nach der Zeit gleich Null setzen, da die Abhängigkeit der Spuleninduktivität von der Spalthöhe nicht von der Zeit abhängt (Gl. (7.50)).

Unter Anwendung der Produktregel für die Differentiation des Klammerausdrucks in Gl. (7.53) erhält man dann:

$$U = R \cdot I + \frac{\partial L(H)}{\partial H} \cdot \frac{dH}{dt} \cdot I + L(H) \frac{dI}{dt} . \quad (7.54)$$

Dies ist eine nichtlineare DGL, die nun zur weiteren Behandlung im angestrebten Arbeitspunkt H_0 zu linearisieren ist. Dazu wird die Taylorreihenentwicklung mit Abbruch nach dem ersten Glied, wie sie zu Anfang des Abschnitts 7.3.1 vorgestellt wurde, verwendet:

$$X_a = X_{a0} + (X_e - X_{e0}) \cdot \left. \frac{dX_a}{dX_e} \right|_{X_{e0}}$$

Dabei ist zu beachten, dass wegen der funktionalen Abhängigkeit in Gl. (7.54)

$$R \cdot I + \frac{\partial L(H)}{\partial H} \cdot \frac{dH}{dt} \cdot I + L(H) \frac{dI}{dt} - U = 0 = f(I, \dot{I}, H, \dot{H}, U)$$

auch die partiellen Ableitungen nach diesen Variablen gebildet werden müssen:

$$f = f_0 + \left. \frac{\partial L(H)}{\partial H} \right|_0 \cdot \Delta I + \left. \frac{\partial f}{\partial \dot{I}} \right|_0 \cdot \Delta \dot{I} + \left. \frac{\partial f}{\partial H} \right|_0 \cdot \Delta H + \left. \frac{\partial f}{\partial \dot{H}} \right|_0 \cdot \Delta \dot{H} + \left. \frac{\partial f}{\partial U} \right|_0 \cdot \Delta U = 0 .$$

Da aus dieser Gleichung folgt, dass $f_0 = 0$ sein muss, und im Arbeitspunkt gilt

$$\dot{H}_0 = \dot{I}_0 = 0 ,$$

ergibt sich folgende DGL

$$f = R \cdot \Delta I + L(H)_0 \cdot \Delta \dot{I} + \left. \frac{\partial L(H)}{\partial H} \right|_0 \cdot \dot{I}_0 \cdot \Delta H + \left. \frac{\partial L(H)}{\partial H} \right|_0 \cdot I_0 \cdot \Delta \dot{H} - \Delta U = 0 .$$

bzw.

$$\Delta U = R \cdot \Delta I + L(H_0) \cdot \frac{d\Delta I}{dt} + \left. \frac{\partial L(H)}{\partial H} \right|_0 \cdot I_0 \cdot \frac{d\Delta H}{dt} .$$

Mit den Beziehungen

$$u = \frac{U - U_0}{U_0} = \frac{\Delta U}{U_0} ; h = \frac{\Delta H}{H_0} ; i = \frac{\Delta I}{I_0} ; \left. \frac{\partial L(H)}{\partial H} \right|_0 = - \frac{L(H_0)}{H_0 + C_2}$$

erhält man eine erste linearisierte DGL für den elektromagnetischen Teil der Strecke:

$$u = \frac{R \cdot I_0}{U_0} \cdot i + L(H_0) \cdot \frac{I_0}{U_0} \frac{di}{dt} - \frac{L(H_0)H_0I_0}{(H_0 + C_2)U_0} \cdot \frac{dh}{dt} . \quad (7.55)$$

Der mechanische Teil wird durch eine Bewegungsgleichung nach dem d'Alembert-schen Prinzip modelliert:

$$F_M - F_m - m \cdot \frac{d^2}{dt^2}(-H) = 0 . \quad (7.56)$$

Einsetzen bekannter Kräfte ergibt:

$$m \cdot \ddot{H} = -F_M + m \cdot g = -\frac{C_1 \cdot I^2}{2(H + C_2)^2} + m \cdot g \quad .$$

Auch dies ist wieder eine nichtlineare DGL, die im Arbeitspunkt linearisiert werden muss. Für den Beharrungszustand gilt:

$$m \cdot \ddot{H} = 0 = -F_{M0} + m \cdot g \Rightarrow F_{M0} = m \cdot g \Rightarrow m \cdot \ddot{H} = -F_M + F_{M0} \quad .$$

Die Magnetkraft ist eine Funktion der Spalthöhe und des Spulenstroms, so dass ihre Linearisierung lauten muss:

$$F_M = f(H, I) = F_{M0} + \left. \frac{\partial F_M}{\partial H} \right|_0 \cdot \Delta H + \left. \frac{\partial F_M}{\partial I} \right|_0 \cdot \Delta I \quad .$$

Mit den oben genannten Beziehungen im Arbeitspunkt wird daraus:

$$\begin{aligned} -m \cdot \ddot{H} &= F_{M0} + \left. \frac{\partial F_M}{\partial H} \right|_0 \cdot \Delta H + \left. \frac{\partial F_M}{\partial I} \right|_0 \cdot \Delta I - F_{M0} \\ &= -\frac{C_1 \cdot I_0^2}{(H_0 + C_2)^3} \cdot \Delta H + \frac{C_1 \cdot I_0}{(H_0 + C_2)^2} \cdot \Delta I \quad . \end{aligned}$$

Mit den Beziehungen

$$h = \frac{\Delta H}{H_0} ; i = \frac{\Delta I}{I_0}$$

erhält man die linearisierte DGL für den mechanischen Teil des Systems:

$$\frac{m(H_0 + C_2)^3}{C_1 \cdot I_0^2} \cdot \ddot{h} - h = \frac{(H_0 + C_2)}{H_0} \cdot i \quad . \quad (7.57)$$

Um das System auf Stabilität zu untersuchen wird die Übertragungsfunktion der Strecke benötigt. Dazu werden zuerst die Koeffizienten der beiden DGL's Gl. (7.55) und Gl. (7.57) im Arbeitspunkt ($H = 2$) bestimmt:

$$\begin{aligned} m \cdot g &= F_{M0} = -\frac{C_1 \cdot I_0^2}{2(H_0 + C_2)^2} \Rightarrow I_0^2 = \frac{2(H_0 + C_2)^2 \cdot m \cdot g}{C_1} \\ I_0^2 &= \frac{2 \cdot 9 \cdot 50 \cdot 9,81}{5,53 \cdot 10^{-1} \cdot 10^3} \cdot \frac{\text{mm}^2 (\text{kg} \cdot \text{m} \cdot \text{s}^{-2}) \text{m}}{\text{V} \cdot \text{s} \cdot \text{mm} \cdot \text{A}^{-1} \cdot \text{mm}} = 16 \text{A}^2 \Rightarrow I_0 = 4 \text{A} \\ L(H_0) &= \frac{C_1}{H_0 + C_2} = \frac{5,52 \cdot 10^{-1}}{3} \cdot \frac{\text{V} \cdot \text{s} \cdot \text{mm}}{\text{A} \cdot \text{m}} = 0,184 \frac{\text{V} \cdot \text{s}}{\text{A}} \\ U_0 &= R \cdot I_0 = 1,5 \cdot 4 \Omega \cdot \text{A} = 6 \text{V} \quad . \end{aligned}$$

Mit diesen Zahlenwerten ergibt sich das linearisierte, mathematische Modell der Regelstrecke aus elektromagnetischem und mechanischem Teil zu:

$$\begin{aligned} u &= i - \frac{0,184 \cdot 2 \cdot 4}{3 \cdot 6} \cdot \frac{\text{V} \cdot \text{s} \cdot \text{mm} \cdot \text{A}}{\text{A} \cdot \text{mm} \cdot \text{V}} \cdot \dot{h} + \frac{0,184 \cdot 4}{6} \cdot \frac{\text{V} \cdot \text{s} \cdot \text{A}}{\text{A} \cdot \text{V}} \cdot \dot{i} \quad \begin{array}{l} \text{elektromag.} \\ \text{Teil} \end{array} \quad (7.58a) \\ &= i - 0,08177 \cdot \dot{h} + 0,12266 \cdot \dot{i} \end{aligned}$$

$$\frac{50 \cdot 27}{0,552 \cdot 16} \cdot \frac{\text{kg} \cdot \text{mm}^3 \cdot \text{A}}{\text{V} \cdot \text{s} \cdot \text{mm} \cdot \text{A}^2} \cdot \ddot{h} - h = -\frac{3}{2} \cdot \frac{\text{mm}}{\text{mm}} \cdot i \quad \begin{array}{l} \text{mechanischer} \\ \text{Teil} \end{array} \quad (7.58b)$$

$$0,1528 \cdot 10^{-3} \cdot \ddot{h} - h = -1,5i$$

Zur Ermittlung der Übertragungsfunktion der Strecke müssen die Laplacetransformierten der Gl. (7.58 a)) und Gl. (7.58 b)) im Bildbereich berechnet werden:

$$u(s) = i(s) - 0,08177 \cdot s \cdot h(s) + 0,12266 \cdot s \cdot i(s) \quad (7.59a)$$

$$0,1528 \cdot 10^{-3} \cdot s^2 \cdot h(s) - h(s) = -1,5 i(s) \quad (7.59b)$$

Da die Spalthöhe H über die Messgröße X ermittelt wird, müssen diese beiden Größen noch in Zusammenhang gebracht werden:

$$x = \frac{\Delta X}{X_0} ; h = \frac{\Delta H}{H_0} ; \Delta X = C_3 \cdot \Delta H_0 \Rightarrow x = \frac{C_3 \cdot H_0}{X_0} \cdot h = \frac{2,5 \cdot 2}{5} \cdot h = h .$$

Mit diesem Zusammenhang und durch Einsetzen von Gl. (7.59 b)) in Gl. (7.59 a)) kann man $i(s)$ eliminieren und erhält die Übertragungsfunktion der Strecke zu:

$$G_s(s) = \frac{x(s)}{u_R(s)} = \frac{1,5}{1 - 0,1528 \cdot 10^{-3} s^2 - 0,0187 \cdot 10^{-3} s^3} . \quad (7.60)$$

Zur Untersuchung der Stabilität der Strecke müssen die Eigenwerte der charakteristischen Gleichung bestimmt werden. Diese Gleichung lautet:

$$-0,0187 \cdot 10^{-3} s^3 - 0,1528 \cdot 10^{-3} s^2 + 1 = 0 \quad (7.61)$$

bzw.

$$s^3 - 8,171 s^2 + 5,347 \cdot 10^4 = 0 \Rightarrow s_1 \approx 35 .$$

Da die charakteristische Gleichung der Strecke einen positiven reellen Eigenwert besitzt, muss diese instabil sein. Es muss daher im folgenden untersucht werden, ob sie durch einen geeigneten Regler stabilisiert werden kann. Mit der Übertragungsfunktion $G_R(s)$ des Reglers lautet die charakteristische Gleichung des geschlossenen Regelkreises:

$$0,0187 \cdot 10^{-3} s^3 + 0,1528 \cdot 10^{-3} s^2 + (-1 - 1,5 \cdot G_R) = 0 . \quad (7.62)$$

An Gl. (7.62) kann man erkennen, dass der geschlossene Regelkreis nur mit einer negativen Verstärkung des Reglers zu stabilisieren ist, da nur dann die Nullstelle s_1 einen negativen Wert bekommt (Bedingung für Stabilität). Dies bedeutet, dass zur Stabilisierung des Regelkreises die Rückkopplung als Mitkopplung und nicht als Gegenkopplung realisiert werden muss. Diese Eigenschaft der Regelstrecke ist leicht einsehbar: Eine Vergrößerung der Spalthöhe H hätte bei Gegenkopplung und positiver Verstärkung eine Verkleinerung der Stellgröße u_R und damit eine Verkleinerung der Magnetkraft zur Folge. Daher würde die Magnetkraft mit der Vergrößerung von H abnehmen, wodurch das zu wiegende Gewicht schon bei der kleinsten Störung des Spaltes abfallen würde. Einer Vergrößerung von H muss als eine Vergrößerung der Stellgröße u_R bewirken, d. h. die rückgekoppelte Messgröße x muss mit positivem Vorzeichen dem Regler zugeführt werden, oder der Regler muss eine negative Verstärkung besitzen.

Da eine notwendige Bedingung für Stabilität eines geschlossenen Regelkreises war, dass alle Koeffizienten des charakteristischen Polynoms vorhanden und von gleichem Vorzeichen sind, kann man an Gl. (7.62) sofort sehen, dass ein P- Regler das Regelsystem nicht stabilisieren kann, da im Polynom die Potenz s^1 fehlt. Der Regler muss noch zusätzlich ein bestimmtes Zeitverhalten aufweisen.

Verwendet man einen **PI- Regler**, dessen Übertragungsfunktion

$$G_R(s) = -K_p \left(1 + \frac{1}{T_n \cdot s} \right)$$

beträgt, so lautet die charakteristische Gleichung des geschlossenen Regelkreises:

$$b_3 s^4 + b_2 s^3 + (1,5 \cdot K_p - 1)s + \frac{1,5 \cdot K_p}{T_n} = 0 \quad .$$

In diesem Polynom fehlt die Potenz s^2 ($b_1 = 0$), so dass der Regelkreis auch mit einem PI - Regler nicht zu stabilisieren ist.

Verwendet man einen **PD - Regler** mit der Übertragungsfunktion

$$G_R(s) = -K_p (1 + T_v \cdot s)$$

so lautet die charakteristische Gleichung:

$$b_3 s^3 + b_2 s^2 + 1,5 \cdot K_p T_v s + (1,5 K_p - 1) = 0 \quad .$$

Unter bestimmten Bedingungen können nun alle Koeffizienten der charakteristischen Gleichung positiv und von Null verschieden sein, so dass es Werte von K_p geben kann, für die alle Eigenwerte des Systems einen negativen Realteil besitzen. Für Stabilität muss nach dem Hurwitz-Kriterium gelten:

$$\begin{vmatrix} b_1 & b_3 & 0 \\ b_0 & b_2 & 0 \\ 0 & b_1 & b_3 \end{vmatrix} = \begin{vmatrix} 1,5 K_p T_v & 0,0187 \cdot 10^{-3} & 0 \\ 1,5 K_p - 1 & 0,1528 \cdot 10^{-3} & 0 \\ 0 & 1,5 K_p T_v & 0,0187 \cdot 10^{-3} \end{vmatrix} > 0 \quad .$$

Dies setzt folgende Relationen voraus:

$$0,1528 \cdot 10^{-3} \cdot 1,5 K_p T_v > 0,0187 \cdot 10^{-3} (1,5 K_p - 1) \quad .$$

Daher muss für Stabilität gelten:

$$T_v > 0,1223 - \frac{0,0816}{K_p} \Rightarrow K_p > \frac{2}{3} \quad .$$

Für diese Werte von K_p und T_v ist das System zwar immer stabilisierbar, aber der PD - Regler verursacht eine bleibende Regelabweichung. Sinn der Einrichtung ist eine Wägung zu ermöglichen, was nur dann aus der Messung des Stroms durch den Magneten möglich ist, wenn die Spalthöhe konstant bleibt. Der für das Rechenbeispiel eingesetzte Wert der Masse von 50 kg ist nicht bekannt und deshalb ist eine bleibende Regeldifferenz bei der Sollwertvorgabe für H bei der Messung nicht zu berücksichtigen. Um die Regelabweichung verschwinden zu lassen, benötigt der Regler einen I - Anteil, so dass ein **PID-Regler** sinnvoll ist. Die Übertragungsfunkti-

on eines solchen Reglers lautet:

$$G_R(s) = -K_P \left(1 + \frac{1}{T_n} \cdot \frac{1}{s} + \frac{T_v \cdot s}{1 + T_1 \cdot s} \right) .$$

Die zugehörige charakteristische Gleichung des geschlossenen Regelkreises lautet damit:

$$b_3 s^4 + b_2 s^3 + 1,5 \cdot K_P T_v s^2 + (1,5 K_P - 1)s + \frac{1,5 K_P}{T_n} = 0 .$$

Nach dem Hurwitz-Kriterium muss für Stabilität wiederum gelten:

$$\begin{vmatrix} 1,5 K_P - 1 & 0,1528 \cdot 10^{-3} & 0 & 0 \\ 1,5 K_P T_n^{-1} & 1,5 K_P T_v & 0,0187 \cdot 10^{-3} & 0 \\ 0 & 1,5 K_P - 1 & 0,1528 \cdot 10^{-3} & 0 \\ 0 & 1,5 K_P T_n^{-1} & 1,5 K_P T_v & 0,0187 \cdot 10^{-3} \end{vmatrix} > 0 .$$

Entsprechend dem Kriterium, müssen für Stabilität die eingerahmte Unterdeterminante und der Koeffizient $(1,5 K_P - 1)$ größer als Null sein. Außerdem muss gelten:

$$1,5 K_P T_v [0,0187 \cdot 0,1528 \cdot 10^{-6} - 0,0187^2 \cdot 10^{-6} (1,5 K_P - 1)] (1,5 K_P - 1) > 1,5 K_P T_n^{-1} \cdot 0,1528^2 \cdot 10^{-6} 0,0187 \cdot 10^{-3}$$

und weiterhin

$$(1,5 K_P - 1) > 0 \Rightarrow K_{Pkrit} = 0,6\bar{6} .$$

Wenn $K_P > K_{Pkrit}$ ist und für T_n und T_v sinnvolle Werte wie beispielsweise

$$T_n = T_v = 1 \text{ sec.}$$

verwendet werden, sind diese Stabilitätsbedingungen immer erfüllt, so dass die Regelung mit einem PID-Regler stabil ist und bei Störungen oder Sollwertsprüngen keine bleibende Regelabweichung auftritt. Es ist unmittelbar einleuchtend, dass die genaue Wahl der Parameter jedoch nicht beliebig sein kann, sondern dass es sicher Werte gibt, bei denen die Regelung optimal funktioniert. Dieses Thema wird noch in einem weiteren Unterkapitel zu behandeln sein.

7.3.1.9 Systemidentifikation

Bis jetzt wurde ausschließlich die Modellbildung aufgrund einer theoretischen Systemanalyse durchgeführt. Sie verlangt umfangreiche Informationen über das System und die physikalischen Vorgänge, die in ihm ablaufen. Das Ergebnis einer solchen Systemanalyse ist ein mathematisches Modell in Form einer Differentialgleichung, aus der man die Übertragungsfunktion gewinnen kann. Liegt das reale System, das zu modellieren ist, bereits vor und ist die mathematische Modellierung wegen der physikalischen Zusammenhänge sehr unübersichtlich, so besteht auch die

Möglichkeit, das Modell experimentell durch ein **Identifikationsverfahren** zu ermitteln. Die experimentelle Analyse ermöglicht die Bildung von mathematischen Modellen durch Messung der Ein- und Ausgangsgrößen bei beliebig aufgebauten Systemen. Ein großer Vorteil dieser Vorgehensweise besteht darin, dass die selben experimentellen Analysemethoden auf die verschiedenartigsten, beliebig komplizierten Systeme anwendbar sind. Solche Methoden beruhen auf der Ermittlung des Ausgangsverhaltens als Antwort auf genau definierte und bekannte Eingangsgrößen. Im Prinzip ist die Vorgehensweise schon erwähnt worden, beispielsweise die Ermittlung der Sprungantwort, oder die Messung des Frequenzgangs.

Theoretische Modellbildung	Experimentelle Identifikation
Modellstruktur folgt aus den Naturgesetzen.	Modellstruktur muss angenommen werden.
Beschreibung des Verhaltens der inneren Zustandsgrößen und des Ein-/Ausgangsverhaltens.	Es wird nur das Ein-/Ausgangsverhalten identifiziert.
Modellparameter werden als Funktion von Systemgrößen angegeben.	Modellparameter sind reine Zahlenwerte, die i. a. keinen Zusammenhang mit den physikalischen Systemgrößen erkennen lassen.
Die wesentlichen internen Vorgänge des Systems müssen bekannt und mathematisch beschreibbar sein.	Innere Vorgänge des Systems müssen nicht bekannt sein.
Meist großer Zeitaufwand erforderlich.	Meist relativ kleiner Zeitaufwand erforderlich

Bild 7-92: Gegenüberstellung der Eigenschaften von theoretischer Modellbildung und Identifikation

Durch die Messung von Ein- und Ausgangssignalen kann man jedoch nur Modelle für das Ein- und Ausgangsverhalten der Systeme ermitteln, also keine Modelle, die die wirkliche innere Struktur beschreiben. Solche Ein-/Ausgangsmodelle reichen jedoch für viele Anwendungszwecke aus. Falls das System die Messung innerer Zustandsgrößen zulässt, kann man natürlich auch Informationen über die innere Struktur erhalten. In der Tabelle in Bild 7-92 sind einige Eigenschaften der theoretischen Modellbildung und der experimentellen Identifikation einander gegenübergestellt [7.12].

Um den Anwendungsbereich der jeweiligen Vorgehensweise bei der Modellbildung zu bewerten, muss man die erreichbare Modellgenauigkeit und den erforderlichen Aufwand an Zeit und Kosten miteinander vergleichen. Dazu kann man beispielsweise die erreichbare Modellgüte (bewertet nach einem im Folgekapitel noch zu behandelnden Kriterium) und die dafür erforderliche Kenntnis der inneren Zusammenhänge des Systems zueinander in Relation setzen (Bild 7-93 a)). Man sieht, dass für entsprechende Erhöhungen der Modellgüte bei der theoretischen Modellbildung ein entsprechender Aufwand für die Kenntnis der inneren Systemzusammenhänge zu leisten ist. Bei der experimentellen Identifikation benötigt man zumindest bei linearisierbaren Systemen bedeutend weniger Kenntnisse über die inneren Vorgänge des Systems, um Modelle gleicher oder besserer Güte zu erreichen.

Bezieht man nun aber noch die Verschiedenartigkeit von Prozessen in die Betrachtung ein, wie beispielsweise verfahrenstechnische und energietechnische Prozesse und Anlagen, sowie die bei mechatronischen Systemen am häufigsten vertretenen mechanischen und elektrischen Systeme, so ändert sich die Aussage für bestimmte Systeme. Wie Bild 7-93 b) zeigt, ist es für die Verfahrenstechnik häufig günstiger, das Modell experimentell als durch theoretische Modellbildung zu ermitteln. Bei mechanischen und elektrischen Systemen, mit denen sich

die Mechatronik am häufigsten befasst, verschwindet der Vorteil der Modellbildung durch Identifikationsverfahren fast. Dies liegt daran, dass die Methoden der theoretischen Modellbildung für diese Systeme seit langem gut bekannt sind, so dass man das theoretische Modell mit geringem Aufwand ermitteln kann. Weniger einfach wird jedoch der Sachverhalt, wenn dissipative und nichtlineare Prozesse wie Reibung, Spiel oder elektrische Verluste eine wichtige Rolle bei der Modellbildung spielen.

Man kann auch beide Verfahren miteinander verbinden, indem man sich ein relativ einfaches theoretisches Modell des Systems bildet und dieses mit einem entsprechenden rechnergestützten System (früher wurden hierfür Analogrechner verwendet) simuliert. Gleichzeitig führt man Identifikationsverfahren an dem realen System durch und vergleicht die Ergebnisse mit denen der Rechnersimulation. Abweichungen zwischen dem realen und dem simulierten Ein-/Ausgangsverhalten können dann Hinweise auf Verbesserungen des theoretischen Modells liefern, wobei die Verbesserung durch erneute Simulation überprüft werden kann. Solche rechnergestützten **Simulationssysteme** werden im Kapitel 8 behandelt.

Die einfachsten Identifikationsverfahren verwenden die Messung der Ausgangsgröße eines Systems bei genau bekannten Testsignalen als Eingangsgröße. Verwendet man als Testsignal einen Sprung, so bekommt man als Ausgangsgröße die **Sprungantwort** oder **Übergangsfunktion $h(t)$** . Alternativ kann man auch kurze Impulse als Testsignal verwenden und erhält dadurch am Ausgang die **Stoßübergangsfunktion** oder **Gewichtsfunktion $g(t)$** . Die dritte häufig verwendete Möglichkeit ist die Verwendung harmonischer Testsignale, die man in der Frequenz variiert. Zeichnet man die Ausgangsamplituden für viele verschiedene Frequenzen auf und

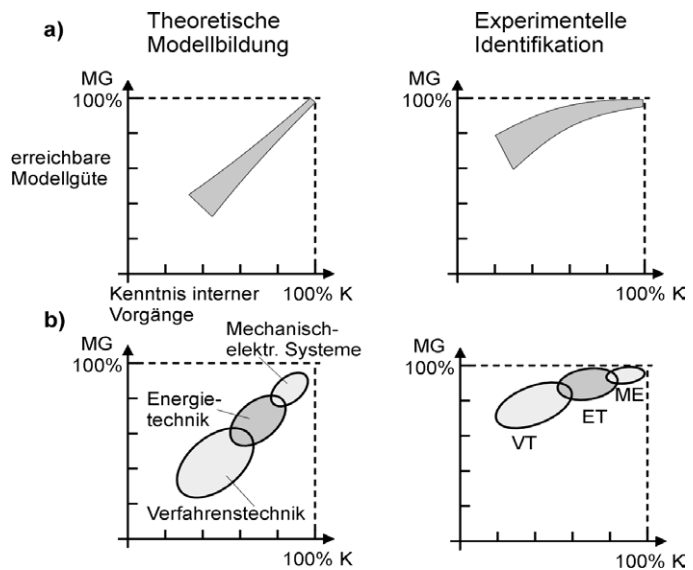


Bild 7-93: Qualitative Zusammenhänge zwischen Modellgüte und Prozesskenntnis bei theoretischer Modellbildung und experimenteller Identifikation a) allgemein
b) prozessabhängig

setzt sie in Beziehung zur Amplitude der Eingangssignale und misst gleichzeitig die Phasenverschiebung zwischen Ein- und Ausgangssignal, so erhält man praktisch den **Frequenzgang**.

Um nochmals an einem Beispiel aufzuzeigen, welche Informationen man aus einer solchen Identifikation erhalten kann, soll die Übergangsfunktion von Verzögerungsgliedern betrachtet werden. Liegt zum Beispiel ein Verzögerungssystem 2. Ordnung vor, so lautet seine DGL

$$T^2 \ddot{x}_a + 2DT \dot{x}_a + x_a = K_P x_e$$

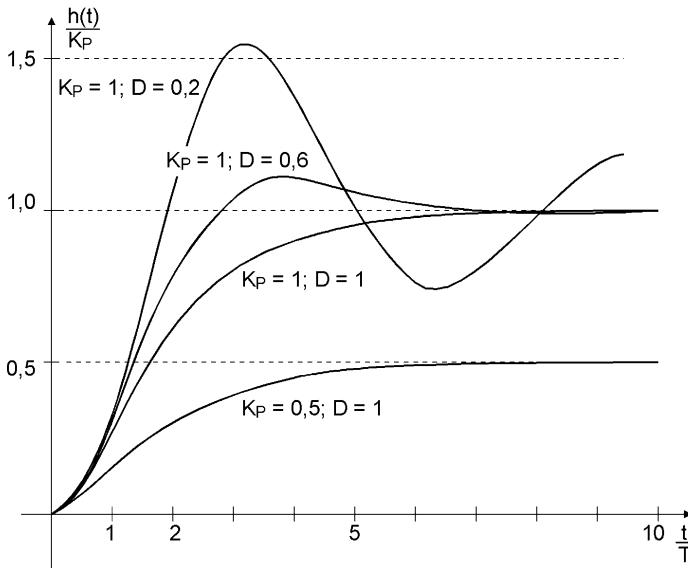


Bild 7-94: Übergangsfunktion $h(t)$ eines PT_2 -Gliedes für unterschiedliche Werte von K_P und D

Im Bild 7-94 sind verschiedene Verläufe der Sprungantwort dargestellt. Die Unterschiede im zeitlichen Verlauf der Signale und in ihrer Endamplitude beruhen auf Änderungen der Systemparameter K_P und D . Ausgehend von dem Fall $K_P = D = 1$ bewirkt eine Verkleinerung der Verstärkung auf $K_P = 0,5$ eine Halbierung des asymptotischen Endwerts der Amplitude für $t \rightarrow \infty$. Umgekehrt kann man bei der Identifikation, bei einer Verringerung der Ausgangsamplitude gegenüber der

Amplitude des Eingangssprungs auf die Hälfte, auf eine Streckenverstärkung $K_P = 0,5$ schließen.

Verkleinert man die Dämpfung D der Strecke, so sieht man, dass der Endwert der Amplitude nicht mehr durch asymptotische Annäherung, sondern mit einem anfänglichen **Überschwingen** über den Endwert und mehr oder weniger stark gedämpftem Ausschwingen erreicht wird. Während bei einer Dämpfung von $D = 0,6$ im wesentlichen nur ein einfaches Überschwingen auftritt, erfolgt die Annäherung an den Endwert bei $D = 0,2$ schon mit einem längeren Ausschwingvorgang. Umgekehrt kann man aus diesem Verlauf bei der Identifikation auf die Größe der Dämpfung D schließen.

Welche dynamischen Eigenschaften bei einem technischen PT_2 -Glied wünschenswert sind, kann man nicht ohne die Anwendung beurteilen. Einerseits gibt es Fälle wo ein Überschwingen über den Endwert absolut unzulässig ist (z. B. bei einem Lageregelkreis), weshalb die Dämpfung D immer ausreichend groß sein muss. Andererseits sieht man, dass bei kleiner Dämpfung das erstmalige Erreichen des Endwerts viel schneller erfolgt, dafür aber Überschwingen auftritt. Ein guter Kompro-

miss zwischen beiden ist die Dämpfung $D = 0,7$, da hier die Überschwingweite nur 5 % von K_P beträgt.

Hat man nun bei der Identifikation eine Sprungantwort wie in Bild 7-95 gemessen, so kann man aus ihr die Streckenparameter K_P , D und T bestimmen. Für die Verstärkung gilt:

$$K_P = \lim_{t \rightarrow \infty} h(t) \quad .$$

Die Lösung der DGL eines PT2-Gliedes bei einer Anregung mit einem Sprung lautet:

$$h(t) = K_P \left[1 - \frac{1}{\sqrt{1-D^2}} e^{\delta t} \cos(\omega t + \varphi) \right]$$

mit

$$\delta = -\frac{D}{T} ; \quad \omega = \frac{1}{T} \sqrt{1-D^2}$$

$$\tan \varphi = \frac{\delta}{\omega} ; \quad \sin \varphi = -D \quad .$$

Die Abnahme des Überschwingens wird durch die Dämpfung D bestimmt. Diese Abnahme kann man aus dem Verhältnis aufeinanderfolgender Überschwingweiten gewinnen. Um aus dem Bild 7-95 die Überschwingweiten $A_0, A_1, \dots, A_n$ zu ermitteln, benötigt man den Wert von $h(t)$ an den entsprechenden Minima und Maxima. Hierfür gilt:

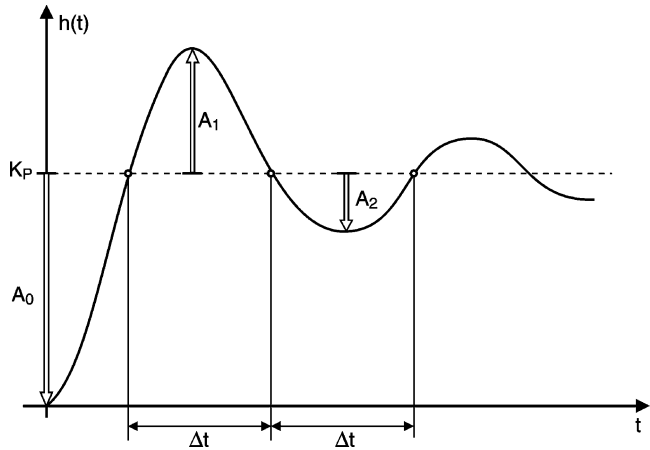


Bild 7-95: Sprungantwort eines PT2-Gliedes bei geringer Dämpfung

$$\frac{dh(t)}{dt} = K_P \frac{1}{T \sqrt{1-D^2}} e^{\delta t} \sin \omega t = 0 \quad \text{für} \quad t = t_{\dot{u}b} \quad .$$

Somit muss gelten:

$$\sin \omega t_{\dot{u}b} = 0 \Rightarrow t_{\dot{u}b} = \frac{k\pi}{\omega} \quad .$$

Damit kann man die jeweilige Überschwingweite berechnen:

$$|A_n| = |h(t_{\dot{u}b}) - K_P| = K_P \cdot e^{\delta n(\pi/\omega)} \quad .$$

Bestimmt man nun das Verhältnis zweier aufeinander folgender Überschwingweiten, so erhält man:

$$v = \left| \frac{A_{n+1}}{A_n} \right| = e^{\delta(\pi/\omega)} = \exp \left(-\frac{\pi D}{\sqrt{1-D^2}} \right) \quad .$$

Da das Verhältnis v nur von der Dämpfung D abhängt, lässt sich diese aus der Messung zweier aufeinander folgender Überschwingweiten bestimmen:

$$D = \sqrt{\frac{(\ln v)^2}{\pi^2 - (\ln v)^2}} \quad .$$

Der letzte Parameter, die Schwingungsdauer T , lässt sich aus der halben Periode Δt der Schwingung berechnen. Diese Zeitdauer Δt wird zwischen den Zeitpunkten gemessen, an denen für die Amplitude der Übergangsfunktion gilt:

$$h(t_n) = K_p \quad .$$

Aus Gl. (7.63) folgt dann:

$$\frac{1}{\sqrt{1-D^2}} e^{\delta t_n} \cos(\omega t_n + \varphi) = 0$$

bzw.

$$\omega \cdot t_n + \varphi = \left(n + \frac{1}{2}\right)\pi \Rightarrow t_n = \frac{1}{\omega} \left(n\pi + \frac{\pi}{2} - \varphi\right) \quad .$$

Daraus lässt sich die Dauer der halben Periode der Schwingung ermitteln:

$$\frac{T_0}{2} = \Delta t = t_{n+1} - t_n = \frac{\pi}{\omega} = \frac{T\pi}{\sqrt{1-D^2}} \quad .$$

Da D bereits bekannt ist, lässt sich hieraus der Parameter T berechnen:

$$T = \frac{\Delta t \cdot \sqrt{1-D^2}}{\pi} \quad .$$

Damit kann das mathematische Modell dieses Verzögerungsgliedes 2. Ordnung vollständig aus gemessenen Werten der Sprungantwort bestimmt werden.

Darüber hinaus kann man an der gemessenen Sprungantwort von Verzögerungsgliedern höherer Ordnung mit bestimmten Verfahren auch die Ordnung eines Verzögerungsgliedes bestimmen. Weitere Methoden zur Identifikation von Strecken finden sich in [7.12].

7.3.2 Synthese von Regelkreisen

Im Kapitel 7.3.1 wurden vor allem Methoden dargestellt, mit denen man folgende Aspekte eines Regelsystems behandeln kann:

- Mathematische Modellbildung von Regler und Strecke
- Untersuchung wesentlicher Eigenschaften dieser Systeme, wie beispielsweise der Stabilität einer Regelung
- Bestimmung der Ausgangsgröße eines Systems bei bekannter Eingangsgröße.

Alle diese Aufgaben zählt man zur **Analyse** von Regelkreisen. Unter **Synthese** versteht man dagegen einen Vorgang, der sich nicht mit der Beschreibung vorhandener Eigenschaften befasst, sondern bei dem versucht wird, durch Änderung von Parametern oder sogar durch Änderung der Struktur einem Regelkreis ein Verhal-

ten zu geben, das einem durch eine **Spezifikation** vorgegebenen Ziel entspricht. Dies kann beispielsweise eine Vorgabe sein, wie das Regelsystem auf ein am Eingang vorgegebenes Testsignal am Ausgang reagieren soll.

Da die Regelstrecke meist vorgegeben und nicht veränderlich ist, richten sich die Verfahren der Synthese vor allem auf die Ermittlung optimaler Reglerparameter oder auf Optimierungen der Reglerstruktur.

7.3.2.1 Spezifikationen

Die Spezifikationen einer Regelung können sowohl das statische, als auch das dynamische Verhalten eines Regelsystems betreffen. Eine typische statische Spezifikation, die bleibende Regelabweichung, wurde bereits im Abschnitt 7.3.1.7 im Zusammenhang mit den Eigenschaften von verschiedenen Reglern behandelt. Sie wird auch als **Positionsfehler** bezeichnet. Diese bei P-Strecken (Strecken mit Ausgleich) und Einsatz von Reglern ohne integrierenden Anteil typischerweise auftretende Eigenschaft, wurde dort für den P-Regler durch den statischen Regelfaktor R_S charakterisiert. Die Definition der Positionsabweichung eines Regelkreises bezieht sich auf das Verhalten für $t \rightarrow \infty$ nach Aufgeben einer sprungförmigen Eingangsgröße in Form eines Sollwertsprungs, oder nach Auftreten einer sprungförmigen Störgröße. Eben solche statischen Spezifikationen kann man natürlich für nicht sprungförmige, typische Eingangsgrößenänderungen definieren. So wird als **Geschwindigkeitsfehler** die Regeldifferenz im Beharrungszustand bei einer rampenförmigen Änderung der Führungsgröße bezeichnet. Er ist bei Systemen mit Ausgleich unendlich groß und bei Systemen ohne Ausgleich bei einfachem integralem Verhalten der Geschwindigkeitszeitkonstanten K_I umgekehrt proportional. Bei mehrfacher Integration im System verschwindet der Geschwindigkeitsfehler sogar ganz.

Neben diesen das Langzeitverhalten nach Eingangsänderungen beschreibenden statischen Spezifikationen gibt es auch dynamische Spezifikationen, die das kurzzeitige Übergangsverhalten qualifizieren. Dabei wirken Veränderungen zur Verbesserung einer statischen Spezifikation auch auf dynamische Spezifikationen ein. In Bild 7-96 ist das Verhalten eines Regelkreises bei einer sprungförmigen Eingangsstörung dargestellt. Die Strecke

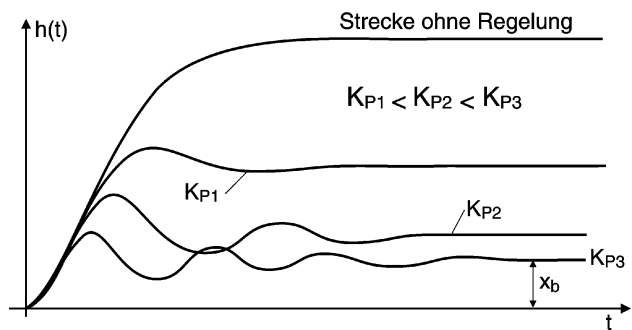


Bild 7-96: Sprungantwort einer Regelstrecke mit Verzögerungsverhalten bei offenem und geschlossenem Regelkreis und unterschiedlicher Reglerverstärkung

zeigt Verzögerungsverhalten, das ohne Regelung zu einem Positionsfehler in Größe des Störsprungs führt. Die Regeleinrichtung ist ein P-Regler, so dass auch der geschlossene Regelkreis mit einer bleibenden Regelabweichung x_b reagiert. Diese nimmt mit zunehmender Reglerverstärkung K_P ab, gleichzeitig sieht man jedoch, dass die Kreisdämpfung ebenfalls abnimmt. Dadurch kommt es zu einem

mehr oder weniger starken Überschwingen, das durch eine dynamische Spezifikation charakterisiert werden muss.

Die Regeleinrichtung hat die Aufgabe, die Regelgröße möglichst genau auf einem vorgegebenen, zeitlich konstanten oder veränderlichen Wert zu halten und den Einfluss von Störgrößen auf ein Mindestmaß zu begrenzen. Weiterhin ist zu fordern, dass das Ausregeln von Störungen oder Sollwertänderungen in möglichst kurzer Zeit erfolgt, wobei vorübergehende Abweichungen der Regelgröße (Überschwingen) sich innerhalb bestimmter Grenzen bewegen müssen. Die im Abschnitt 7.3.1.8 behandelte Stabilität des Regelkreises ist dabei eine absolute Minimalforderung. In dem in diesem Abschnitt als Beispiel behandelten Regelkreis (Magnetische Waage) herrschte immer Stabilität, wenn man die entsprechenden Reglerparameter ober-

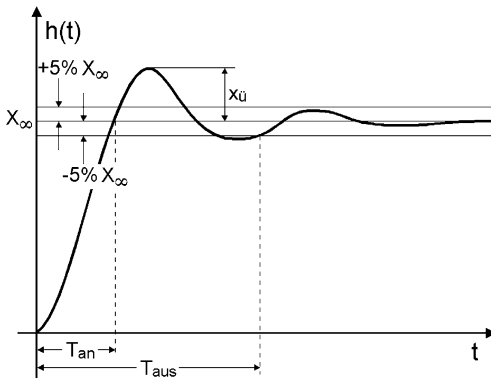


Bild 7-97: Verlauf der Regelgröße nach einer sprunghaftigen Sollwertänderung auf einen neuen Sollwert x_∞ mit dynamischen Spezifikationsgrößen

halb bestimmter Grenzen wählte. Keine Antwort konnte dort darauf gegeben werden, wie diese Parameter nun für optimales Regelverhalten des Systems gewählt werden müssten. Um zu definieren, wann eine Regelung optimal arbeitet, muss man daher erst einmal **Gütekriterien** definieren.

In Bild 7-97 ist der Verlauf der Regelgröße eines geschlossenen Regelkreises nach einer sprunghaftigen Sollwertänderung dargestellt. In diesem Bild sind einige Kenngrößen, die den dynamischen Übergang auf den neuen Sollwert charakterisieren, bezeichnet. Als dynamische Spezifikationen sind hier vor allem folgende Größen von Bedeutung:

- Die Anregelzeit T_{an} als die Zeit, nach der die Regelgröße erstmals den neuen Endwert x_∞ erreicht.
- Die Ausregelzeit T_{aus} als die Zeit, nach der die Regelgröße innerhalb eines Toleranzbandes (typischerweise $\pm 5\%$) um den neuen Endwert bleibt.
- Die Überschwingweite x_u als die maximale Amplitudenüberhöhung über den neuen Sollwert x_∞ , die in Prozent vom Sollwert ein Maß für die Dämpfung des Regelkreises darstellt.

Wie bereits erläutert wurde, kann man durch Änderung der Reglerparameter nicht alle Spezifikationswerte gleichzeitig optimieren. Man muss daher für verschiedene Belange des Regelergebnisses verschiedenartige Optimierungskriterien definieren.

Nimmt die Sprungantwort einer Regelung einen monotonen Verlauf wie in Bild 7-97, so ist das Regelergebnis um so günstiger, je kleiner Überschwingweite und Ausregelzeit sind. Dies ist der Fall, wenn die Fläche zwischen der aktuellen Regelgröße $x(t)$ und dem Endwert x_∞ ein Minimum annimmt, d. h. das Integral über der Regeldifferenz muss ein Minimum werden:

$$I_1 = \int_0^{\infty} (x(t) - x_{\infty}) dt = \int_0^{\infty} x_d(t) dt \rightarrow \text{Min.} \quad (7.64)$$

In Bild 7-98 b) ist diese sog. lineare Regelfläche des monoton abklingenden Regelvorgangs in Teilbild a) dargestellt. Man sieht, dass diese positive und negative Anteile besitzt und der Wert des Integrals falsche Informationen liefern kann, wenn nicht gleichzeitig die Dämpfung mit berücksichtigt wird. Anstelle eines **linearen Intergralkriteriums** zur Optimierung eines Regelvorgangs verwendet man daher besser ein **betragslineares Kriterium** wie in Bild 7-98 c) dargestellt. Bei diesem Kriterium werden positive und negative Flächenanteile (Betragslineare Regelfläche) gleichermaßen gewertet:

$$I_2 = \int_0^{\infty} |x_d(t)| dt \rightarrow \text{Min.} \quad (7.65)$$

Ist eine kleine Überschwingweite und eine kurze Anregelzeit von Bedeutung, so liefert ein **quadratisches Kriterium**, bei dem durch Quadrieren große Amplituden in der quadratischen Regelfläche besonders stark gewertet werden (Bild 7-98 d), noch schärfere Bedingungen:

$$I_3 = \int_0^{\infty} x_d^2(t) dt \rightarrow \text{Min.} \quad (7.66)$$

Dieses letztgenannte Kriterium ist das im Bereich der Regleroptimierung am meisten verwendete Kriterium. Da die Regelabweichung x_d eine Funktion verschiedener Parameter des Reglers ist, ist die Berechnung der entsprechenden Integrale sehr schwierig und unter Umständen nur auf einem Digitalrechner mit entsprechenden Näherungsverfahren möglich.

Für einfache Regleroptimierungen sind daher in der Vergangenheit verschiedene **Einstellregeln** entwickelt worden.

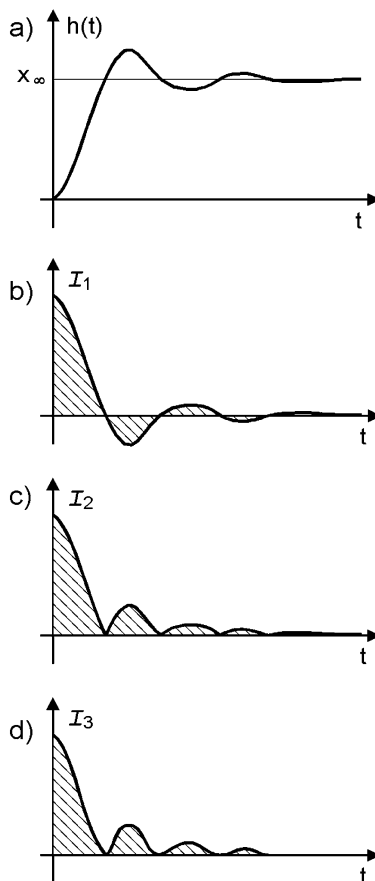


Bild 7-98: Verschiedene Regelflächen: Verlauf
a) der Regelgröße
b) der Regelabweichung
c) des Betrages der Regelabweichung
d) des Quadrates der Regelabweichung

7.3.2.2 Einstellregeln

Um eine Parameteroptimierung vornehmen zu können, muss das Modell der Regelstrecke bekannt sein, oder durch Identifikation ermittelt werden. Dazu kann man im einfachsten Fall die Übergangsfunktion messen, oder gewisse Grenzsituationen herbeiführen, aus denen man auf die statischen und dynamischen Eigenschaften

der Regelstrecke schließen kann. Schließt man den Regelkreis mit Hilfe eines Reglers, so kann man durch Veränderung der Reglerverstärkung die Stabilitätsgrenze als eine Grenzsituation ermitteln. Ein bereits im Jahre 1942 veröffentlichtes Verfahren zur optimalen Reglereinstellung durch Ermittlung der Stabilitätsgrenze ist das von **Ziegler/Nichols**. Dazu muss der Regler auf reines P-Verhalten geschaltet werden, d. h. bei Verwendung eines PID-Reglers muss T_v auf den kleinstmöglichen und T_n auf den größtmöglichen Wert verstellt werden. Danach wird die Verstärkung bis zum Erreichen der Stabilitätsgrenze erhöht. Durch Beobachtung des Auftretens einer Dauerschwingung des Regelkreises wird die kritische Verstärkung $K_{P_{krit}}$ und die zugehörige Schwingungsdauer T_{krit} ermittelt. Aus diesen beiden Werten bekommt man dann Werte für die optimale Einstellung der Reglerparameter:

- P-Regler $K_P = 0,5 \cdot K_{P_{krit}}$
- PI-Regler $K_P = 0,45 \cdot K_{P_{krit}}$, $T_n = 0,85 \cdot T_{krit}$
- PID-Regler $K_P = 0,6 \cdot K_{P_{krit}}$, $T_n = 0,5 \cdot T_{krit}$, $T_v = 0,12 \cdot T_{krit}$

Diese Reglereinstellung ist vor allem für ein gutes Störverhalten geeignet und liefert bei einem Dämpfungsgrad von ungefähr $D = 0,25$ ein maximales Überschwingen von 40 %.

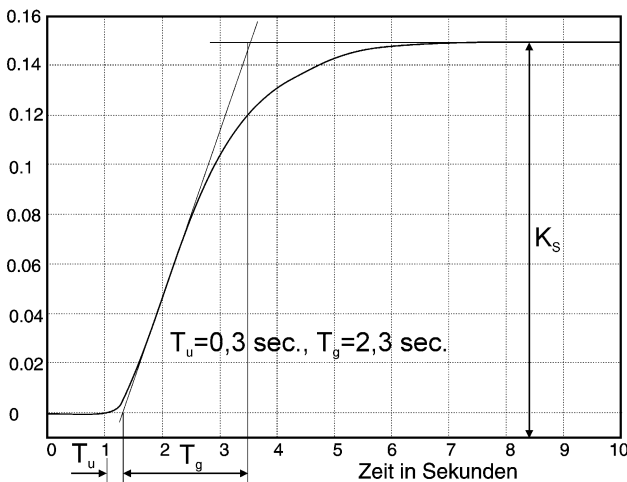


Bild 7-99: Übergangsfunktion einer Regelstrecke mit Ausgleich und charakteristischen Größen für die Einstellregeln nach Chien/Hrones und Reswick

Ein großer Nachteil dieser einfachen Einstellregel ist die Tatsache, dass die Regelung für die Dauer mehrerer Schwingungen an der Stabilitätsgrenze betrieben werden muss. Dies lässt sich nicht für jede Regelstrecke realisieren.

Eine weitere Einstellregel nach **Chien/Hrones/Reswick** vermeidet das Aufsuchen der Stabilitätsgrenze und erfordert auch keine Streckenidentifikation. Es geht von der Übergangsfunktion der Strecke aus (Bild 7-99), aus der die charakteristischen Größen **Übertragungskonstante K_S** , **Verzugszeit T_u** und die **Aus-**

gleichszeit T_g ermittelt werden können. Daher ist das Verfahren so nur auf Strecken mit Ausgleich anwendbar. Die Tabelle in Bild 7-100 enthält für verschiedene Regler, verschiedenartiges Regelungsverhalten und verschiedene Regelungsziele eine Zusammenstellung der Einstellregeln nach dem genannten Verfahren.

In Bild 7-101 sind die Führung- und die Störübergangsfunktion des geschlossenen Regelkreises mit PID-Regler und einer Strecke nach Bild 7-99 dargestellt. Die Reglerparameter wurden jeweils nach den oben genannten Einstellregeln nach Chien/Hrones/Reswick gewählt und der Verlauf der Übergangsfunktionen mit einem Simu-

	Aperiodischer Regelvorgang		20% Überspringen	
	Führung	Störung	Führung	Störung
P-Regler	$K_P = \frac{0,3}{K_S} \cdot \frac{T_g}{T_u}$	$K_P = \frac{0,3}{K_S} \cdot \frac{T_g}{T_u}$	$K_P = \frac{0,7}{K_S} \cdot \frac{T_g}{T_u}$	$K_P = \frac{0,7}{K_S} \cdot \frac{T_g}{T_u}$
PI-Regler	$K_P = \frac{0,35}{K_S} \cdot \frac{T_g}{T_u}$ $T_n = 1,2T_g$	$K_P = \frac{0,6}{K_S} \cdot \frac{T_g}{T_u}$ $T_n = 4T_u$	$K_P = \frac{0,6}{T_u}$ $T_n = T_g$	$K_P = \frac{0,7}{K_S} \cdot \frac{T_g}{T_u}$ $T_n = 2,3T_u$
PID-Regler	$K_P = \frac{0,6}{K_S} \cdot \frac{T_g}{T_u}$ $T_n = T_g, T_v = 0,5T_u$	$K_P = \frac{0,95}{K_S} \cdot \frac{T_g}{T_u}$ $T_n = 2,4T_u$ $T_v = 0,42T_u$	$K_P = \frac{0,95}{K_S} \cdot \frac{T_g}{T_u}$ $T_n = 1,35T_g$ $T_v = 0,47T_u$	$K_P = \frac{1,2}{K_S} \cdot \frac{T_g}{T_u}$ $T_n = 2T_u$ $T_v = 0,42T_u$

Bild 7-100: Einstellregeln für unterschiedliche Reglertypen und unterschiedliches Regelverhalten nach Chien/Hrones/Reswick

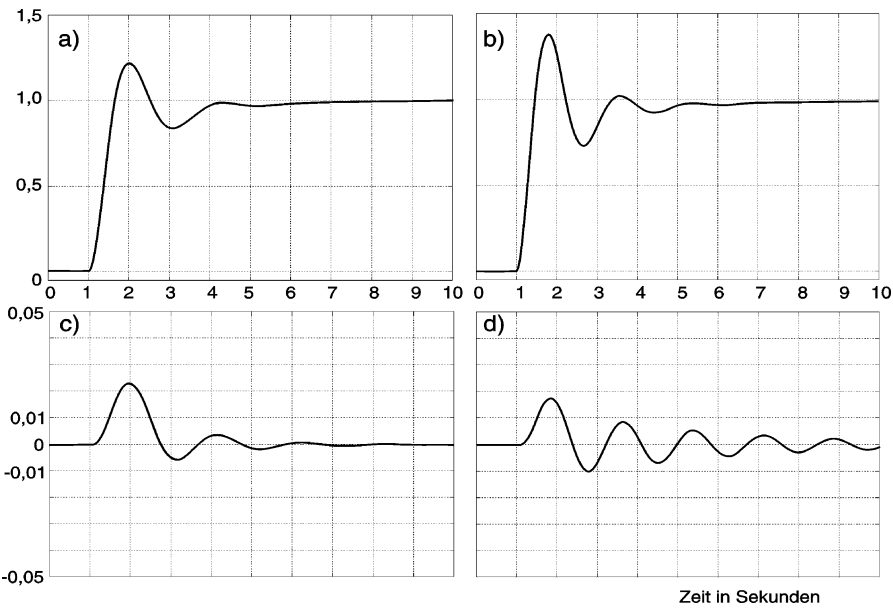


Bild 7-101: Übergangsfunktionen einer Regelung aus einer Strecke mit Ausgleich und einem PID-Regler; Parameteroptimierung nach den Einstellregeln von Chien/Hrones/Reswick

- a) Führungsübertragungsfunktion (FÜF), aperiodische Regelung
- b) FÜF, 20% Überspringen
- c) Störungsübertragungsfunktion (SÜF), aper. Regelung
- d) SÜF, 20% Überspringen

lationssystem ermittelt. Hier zeigt sich, dass solche Methoden ihre Grenzen in der Genauigkeit des Ablesens der Kenngrößen der Regelstrecke finden. Wie man in

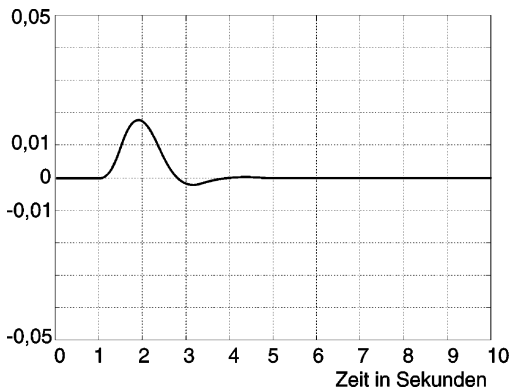


Bild 7-102: SÜF der Strecke mit Ausgleich und PID-Regler. Veränderte Vorhaltezeit $T_{v\text{neu}} = 2 T_v$

den Teilbildern 7-101 a) bis d) sieht, zeigen die Übergangsfunktionen, die mit den optimierten Parametereinstellungen ermittelt wurden, keinen aperiodischen Verlauf. Erhöht man aber beispielweise die Vorhaltezeit T_v des PID-Reglers auf den doppelten Wert, so ergibt sich wie in Bild 7-102 gezeigt, tatsächlich ein aperiodischer Regelverlauf. Mit einem Simulationssystem, wie sie im nächsten Kapitel behandelt werden, kann man solche Untersuchungen sehr einfach durchführen. Darüber hinaus gibt es sowohl für Betrachtungen im Zeitbereich als auch für solche im Frequenzbereich weitere Einstellregeln und Optimierungsmethoden.

Dafür sei auf die entsprechende, in diesem Kapitel genannte Literatur verwiesen.

7.3.2.3 Mehrschleifige Regelkreise

Alle bisherigen Betrachtungen bezogen sich auf Regelsysteme, die nur eine Regelgröße (Eingrößensysteme) verwendeten und mit einer Rückführschleife (einschleifige Regelsysteme) auskamen. Der Regelverlauf für Eingrößensysteme kann jedoch wesentlich verbessert werden, wenn nicht nur Informationen über die Regelgröße selbst zur Verfügung stehen, sondern auch Informationen über deren Ableitungen. Die Differentialgleichungen vieler mechanischer Systeme, bei denen der Weg die Regelgröße ist (**Lageregelung**), enthalten die erste Ableitung (Geschwindigkeit) und die zweite Ableitung (Beschleunigung). Werden diese Größen in einem Lageregelkreis ebenfalls erfasst, so kann man ihren Verlauf in der Re-

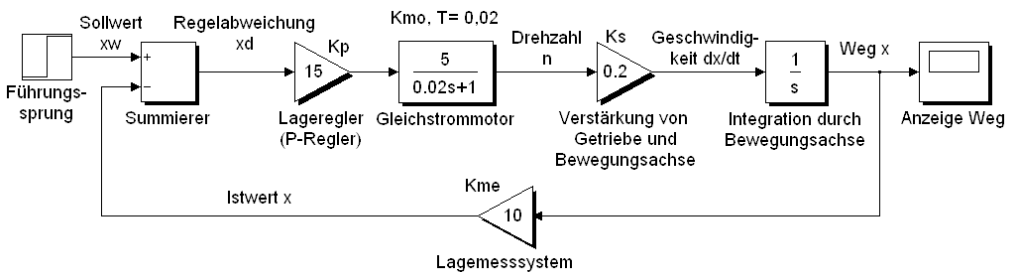


Bild 7-103: Blockschaltbild eines einfachen Lageregelkreises (Simulationssystem Simulink)

gelung berücksichtigen. Dadurch entsteht ein mehrschleifiger Regelkreis, der auch als **Kaskadenregelung** bezeichnet wird. Solche Lageregelungen sind Bestandteil der meisten Positioniersysteme in Arbeitsmaschinen wie Werkzeugmaschinen oder Industrierobotern. In Bild 7-103 ist das Blockschaltbild einer einfachen Lageregelung dargestellt, wie man es aus der Modellierung mit einem Simulationssystem

(Simulink) erhält. Der Lageregelkreis besteht aus einem P-Regler, einem Gleichstrommotor und einer Bewegungsachse, deren Position (Lageistwert) durch ein Lagemesssystem bestimmt wird. Der Motor und die trägen Massen sind durch ein Verzögerungsglied erster Ordnung mit der Zeitkonstante $T = 0,02$ Sek. modelliert, die Verfahrenseinheit besteht aus einem Getriebe und der Bewegungsachse. Diese Einheit ist als Reihenschaltung von P-Glied und I-Glied modelliert, da sie die Drehzahl n des Motors zum Weg x aufintegriert. Das Lagemesssystem hat proportionales Übertragungsverhalten.

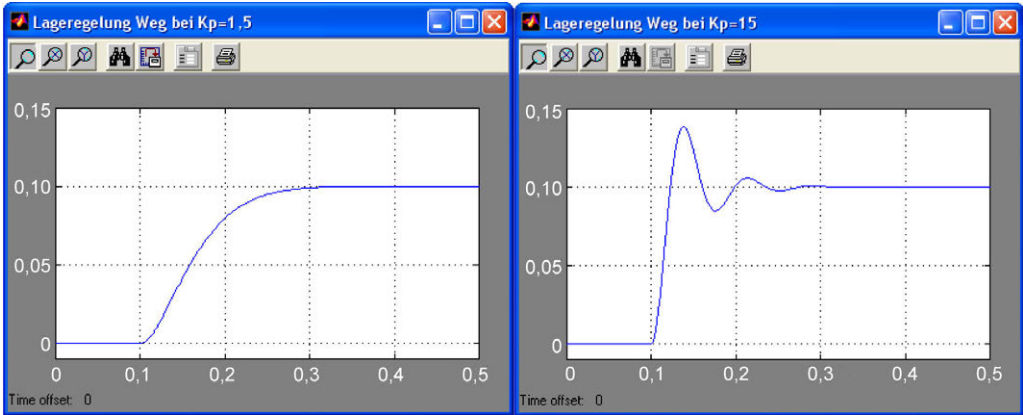


Bild 7-104: Wegverlauf (von Weg 0 –0,1 mm) bei einschleifiger Lageregelung und den unterschiedlichen Verstärkungen $K_P = 1,5$ bzw. 15; Maßstab auf der Zeitachse in Sekunden

Wie man an den Simulationsergebnissen in Bild 7-104 sieht, erhält man mit dem verwendeten P-Regler bei Anlegen eines Führungsgrößensprungs nur für geringe Verstärkung ($K_P = 1,5$) einen aperiodischen Regelvorgang. Die Zeit bis zum Erreichen der neuen Sollposition (0,1) beträgt annähernd 0,2 Sekunden. Erhöht man die Reglerverstärkung zur Verbesserung der Dynamik des Regelkreises auf $K_P = 15$, so wird die Sollposition zwar schon nach ca. 0,02 Sek. erreicht, es kommt aber zu starkem Überschwingen. Dies ist insbesondere für eine Lageregelung nicht wünschenswert. Der Lageregelkreis besitzt offenbar eine zu geringe Dämpfung.

Diese Ergebnisse kann man nun verbessern, wenn man einen zweischleifigen Re-

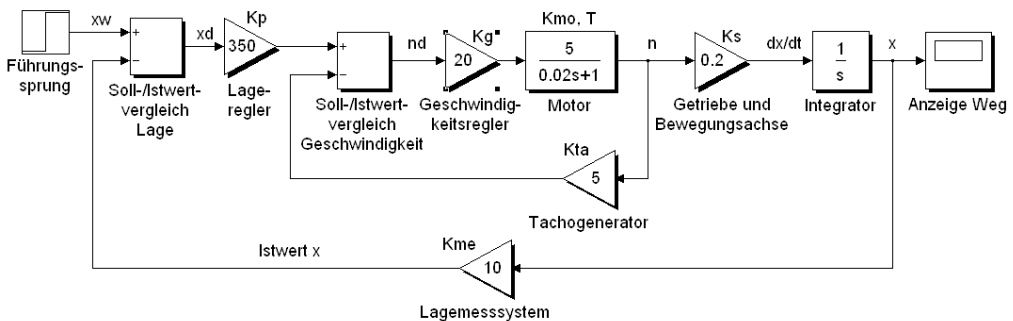


Bild 7-105: Blockschaltbild eines Lageregelkreises mit unterlagierter Geschwindigkeitsregelung

gelkreis einführt. Hierzu wird, wie in Bild 7-105 gezeigt, in den Lageregelkreis ein unterlagertes Geschwindigkeitsregelkreis (Drehzahlregelkreis) eingefügt. Die Ist-drehzahl wird mit einem Tachogenerator gemessen und mit der Solldrehzahl verglichen. Die Regeldifferenz wird dem Geschwindigkeitsregler (Drehzahlregler) zugeführt, der den Motor ansteuert. Die Solldrehzahl ist die Ausgangsgröße des Lagereglers, die durch die Regeldifferenz zwischen Soll- und Istposition bestimmt wird.

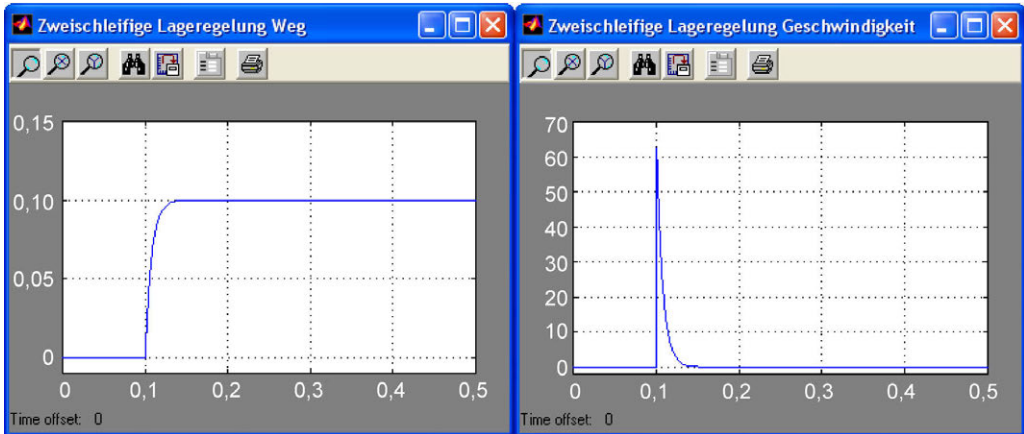


Bild 7-106: Weg- und Geschwindigkeitsverlauf bei einer Lageregelung mit unterlagertem Geschwindigkeitsregelkreis

Gibt man nun wieder eine sprungförmige Führungsgrößenänderung vor, so lässt sich die Verstärkung K_P des Lagereglers bei einer Verstärkung $K_g = 20$ des Geschwindigkeitsreglers und einer Verstärkung $K_{ta} = 5$ des proportionalwirkenden Tachos auf sehr hohe Werte einstellen, ohne das Überschwingen auftritt. Bild 7-106 zeigt Weg- und Geschwindigkeitsverlauf des Lageregelkreises für $K_P = 100$. Wie man sieht, erreicht der Weg dabei bereits nach ca. 0,04 Sekunden seinen Endwert und kann daher das Beschleunigungsvermögen des mechanischen Systems mit einer Zeitkonstante $T = 0,02$ Sek. voll ausnutzen. Die Gesamtdämpfung des Systems ist offensichtlich stark angestiegen.

Dies wird klar, wenn man sich die Übertragungsfunktionen der beiden Regelkreise mit und ohne Geschwindigkeitsregelung anschaut. Ohne Geschwindigkeitsregelung gilt für die Lageregelung:

$$G(s) = \frac{\frac{K_P K_{mo} K_s}{s(1+Ts)}}{1 + \frac{K_{me} K_P K_{mo} K_s}{s(1+Ts)}} = \frac{\frac{1}{K_{me}}}{1 + \frac{1}{K_{me} K_P K_{mo} K_s} s + \frac{1}{K_{me} K_P K_{mo} K_s} s^2}.$$

Diese Übertragungsfunktion kann man abgekürzt folgendermaßen schreiben und erhält dadurch die Normalform eines Verzögerungsgliedes 2. Ordnung:

$$G(s) = \frac{K}{1 + T_1 s + T_2^2 s^2}.$$

Die Dämpfung eines Verzögerungsgliedes 2.Ordnung bestimmt sich daraus zu:

$$D = \frac{1}{2} \frac{T_1}{T_2} = \frac{1}{2} \frac{1}{\sqrt{K_p K_{mo} K_s K_{me} T}} .$$

Setzt man die Werte für die in Bild 7-104 dargestellten Fälle ein, so erhält man folgende Dämpfungsgrade D :

$$K_p = 1,5: D = 0,91 \quad K_p = 15: D = 0,3 .$$

Diese Werte für die Dämpfung des Lageregelsystems erklären das in Bild 7-104 gezeigte Verhalten auf einen Sprung der Führungsgröße.

Will man die gleiche Betrachtung für den Lageregelkreis mit unterlagertem Geschwindigkeitsregelkreis durchführen, so muss man zuerst die Übertragungsfunktion des Geschwindigkeitsregelkreises bestimmen:

$$G_g(s) = \frac{\frac{K_g K_{mo}}{1 + Ts}}{1 + \frac{K_{ta} K_g K_{mo}}{1 + Ts}} = \frac{1}{\left(K_{ta} + \frac{1}{K_g K_{mo}} \right) + \frac{T}{K_g K_{mo}} s} .$$

Die Übertragungsfunktion des gesamten Regelkreises lautet damit:

$$\bar{G}_g(s) = \frac{\frac{K_p G_g(s) K_s}{s}}{1 + \frac{K_{me} K_p G_g(s) K_s}{s}} = \frac{\frac{1}{K_{me}}}{1 + \frac{1 + K_g K_{mo} K_{ta}}{K_{me} K_p K_g K_{mo} K_s} s + \frac{T}{K_{me} K_p K_g K_{mo} K_s} s^2} .$$

Auch dies ist die Übertragungsfunktion eines Verzögerungsgliedes 2. Ordnung, aus der man wieder die Dämpfung berechnen kann:

$$\bar{D} = \frac{1}{2} \frac{T_1}{T_2} = \frac{1}{2} \frac{1 + K_g K_{mo} K_{ta}}{\sqrt{K_p K_{mo} K_s K_{me} K_g T}} .$$

Die Nachrechnung der jetzt vorhandenen Dämpfung ergibt für den in Bild 7-106 gezeigten Fall:

$$K_p = 100, K_g = 20: D = 3,775 .$$

Diese trotz hoher Verstärkung immer noch außerordentlich gute Dämpfung erklärt das dynamische Verhalten des Lageregelkreises mit unterlagertem Geschwindigkeitsregelkreis.

Die oben aufgezeigte Verbesserung des Regelverhaltens wurde dadurch erreicht, dass die Information über die Zustandsgröße "Geschwindigkeit" gemessen und in die Regelung mit einbezogen wurde. Daher ist es naheliegend, eine weitere Zustandsgröße wie die "Beschleunigung" ebenfalls in die Regelung mit einzubeziehen, um das Regelverhalten noch weiter zu verbessern. Da die direkte Messung der Beschleunigung schwierig ist, verwendet man meist eine zur Beschleunigung proportionale Größe nämlich den Strom durch den Anker des Motors. Durch einen weiteren dem Geschwindigkeitsregelkreis unterlagerten Stromregelkreis erhält man dann ei-

an den Prozess weitergibt [7.16]. Der Realzeitrechner wird anders als andere Rechner durch externe Ereignisse unterbrochen, die Priorität der Unterbrechung wird interpretiert, und bei zeitkritischen Aufgaben wird sofort korrigierend in den Prozess eingegriffen. Durch modular aufgebaute Hardwarestrukturen (Mikroprozessoren), genormte Interfaces, standardisierte Bussysteme und strukturierte Programme lassen sich effiziente und fehlerarme Realzeitsysteme auf eine Anwendung hin zuschneiden. Der Rechner muss aber auch seine internen Ereignisse, also sich selbst überwachen können, um sicher zu sein, dass er funktionstüchtig ist und den Prozess unter Kontrolle hat.

Externe Ereignisse können dabei folgenden Ursprungs sein:

- Eintreffen eines Zeittaktes (Uhr)
- Unterbrechung durch ein Sensorsignal (Prozessstatus)
- Anforderung eines Peripheriegerätes (interaktiver Betrieb)

Die beiden ersten Ereignisse treten in der Regel bei der Steuerung und Regelung eines Prozesses auf, das letzte Ereignis im Zusammenhang mit der Kommunikation mit Benutzern oder zu Dokumentationszwecken.

7.4.1 Mikrorechner

Der Mikrorechner ist die konsequente Weiterentwicklung von anfänglichen Miniaturisierungsstufen der Elektronik, die mit der Erfindung des Transistors (1948, Bell-Telephone-Laboratories) begann.

Mikrocomputer wurden durch zwei technische Neuerungen in der Mikroelektronik ermöglicht: dem integrierten Schaltkreis (IC: Integrated Circuit; 1959 entwickelt von J. Kilby und R. Noyce) und dem **Mikroprozessor** (erstmalig 1971, Intel 4004, 4-Bit Mikroprozessor). Der IC gestattete die Miniaturisierung von Computerspeicherschaltungen und Peripheriebaugruppen. Der Mikroprozessor verringerte die Größe der Computer-CPU auf die Größe eines einzelnen Siliziumchips.

Weil eine CPU Berechnungen und logische Operationen durchführt, Betriebsanweisungen enthält und den Datenfluss verwaltet, existierte das Potential, ein spezielles System zu entwickeln, das wie ein vollständiger Mikrocomputer funktioniert. Das erste dieser Systeme, wurde speziell für den persönlichen Gebrauch entworfen und erschien 1974. Es wurde von der Firma Micro Instrumentation Telemetry Systems (MITS) angeboten. Die Besitzer des Systems wurden vom Herausgeber einer in den USA populären Technologiezeitschrift aufgefordert, einen Computerbausatz zu entwerfen und als Versandgeschäft über die Zeitschrift anzubieten. Der Computer mit dem Namen Altair wurde für etwas weniger als 400 \$ angeboten.

1981 führte IBM ihr eigenes Mikrocomputermodell ein, den IBM-PC, ausgestattet mit einem 16-Bit Mikroprozessor Intel 8088). Obwohl er nicht die neueste Computertechnologie einsetzte, war der PC ein Meilenstein in diesem aufblühenden Marktsegment. Er zeigte, dass die Mikrocomputerindustrie mehr als eine aktuelle Modeerscheinung war und dass der Mikrocomputer tatsächlich ein notwendiges Werkzeug für Wirtschaft und Industrie darstellte. Die Verwendung des 16-Bit-

Mikroprozessors im PC regte die Entwicklung schnellerer und leistungsfähigerer Mikrorechner an.

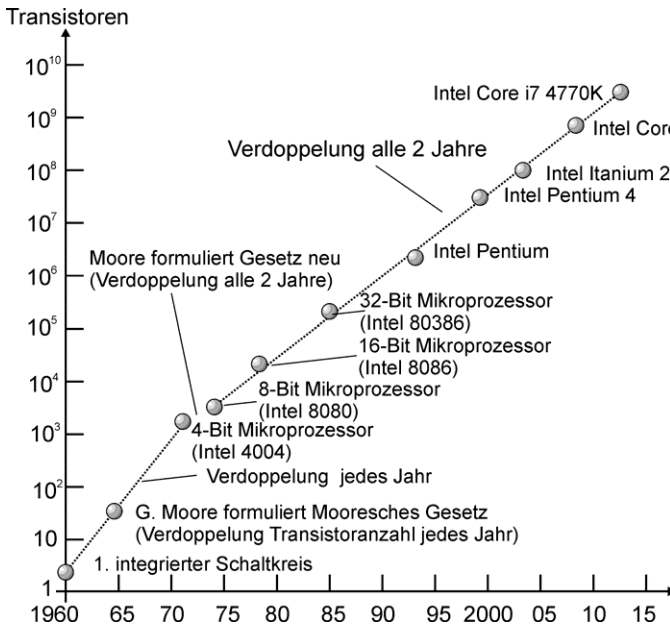


Bild 7-108: Entwicklung der Integrationsdichte von Mikroprozessoren nach dem Moore'schen Gesetz

Die Integrationsdichte schreitet weiter rasant fort, wobei das von G. Moore im Jahr 1965 erstmals formulierte sog. **Moore'sche Gesetz** besagt, dass sich die Zahl der Transistoren auf einem Siliziumchip alle 11 Monate verdoppelt. Diese keinem wirklichen Naturgesetz entsprechende Erfahrungsaussage wurde von ihm im Jahre 1975 daraufhin modifiziert, dass alle 24 Monate (18 Monate) sich die Anzahl der Bauelemente verdoppelt und die Packungsdichte sich alle drei Jahre vervierfacht. Bild 7-108 zeigt diese Entwicklung beginnend mit geringen

Integrationsgraden SSI (Small Scale Integration), in denen erste IC's gefertigt wurden, bis zu VLSI (Very Large Scale Integration), die im Jahr 2004 den Intel Mikroprozessor Itanium 2 mit 9 MB Speicher und 592 Millionen Transistoren ermöglichte. Im Jahr 2013 enthielt der Intel Prozessor Core i7 4770K bereits 2,27 Milliarden Transistoren. Diese Entwicklung wird sich weiter fortsetzen, hat aber von der Integrationsdichte noch lange nicht die Komplexität erreicht, die im menschlichen Gehirn vorliegt. So besitzen heutige leistungsfähige PC's etwa eine Rechenkapazität, wie sie das Nervensystem einer Spinne besitzt. Seine Begrenzung wird die Siliziumtechnologie einerseits durch die im Chip entstehende Verlustwärme und durch den bei immer kleiner werdenden Strukturen unvermeidlichen quantenmechanischen Tunnelstrom finden, der bei Strukturgrößen von 2-3 nm zu erwarten ist.

7.4.1.1 Aufbau von Mikrorechnern

Ein Mikrorechner ist eine Zusammenschaltung eines Mikroprozessors mit Speicherelementen und Ein-/Ausgabebaugruppen. Dies können separate integrierte Bausteine sein, die auf einer oder mehreren Platinen zusammen geschaltet sind, oder auch ein Einchip-Mikrorechner, der als **Mikrocontroller** bezeichnet wird. Bild 7-109 zeigt die typischen Baugruppen, die in der Regel zu einem Mikrorechner zusammen geschaltet werden [1.2], [7.17]. Er besteht aus den Baugruppen:

- **CPU** (central prosessor unit): Mikroprozessor (Zentraleinheit) zur logischen und arithmetischen Verarbeitung von Daten, die durch die Programmierung festgelegt wird.
- **Counter/Timer** : Taktgeber. Frequenz bestimmt Arbeitsgeschwindigkeit.
- **RAM** (random access memory): Arbeitsspeicher aus Speicherbausteinen mit Schreib-/Lesezugriff.
- **ROM** (read only memory): Festwertspeicher mit Nur-Lesezugriff, der bei der Herstellung maskenprogrammiert wird. Enthält beispielsweise den unveränderlichen Teil des Betriebssystems.
- **PROM** (programmable ROM), **EPROM** (erasable PROM), **EEPROM** (electrically erasable PROM): Festwertspeicher, der durch den Anwender programmiert werden kann und bei EPROM/EEPROM auch wieder löscherbar ist.

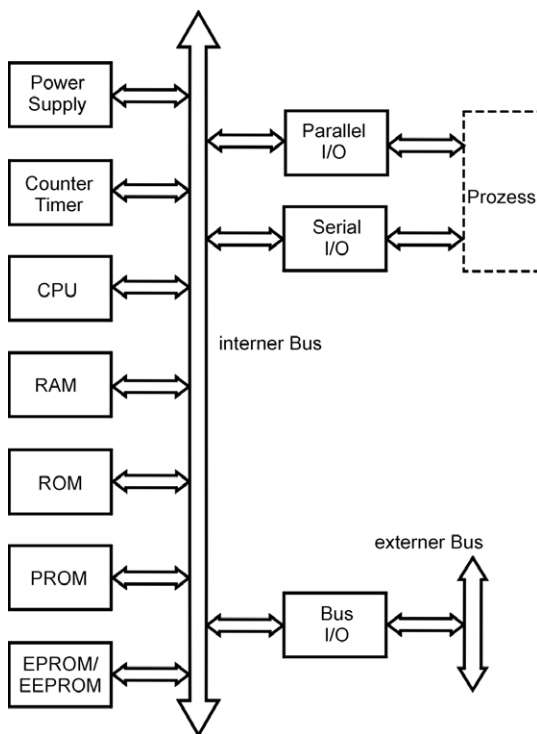


Bild 7-109: Baugruppen eines Mikrorechners

- **Parallele I/O** : Ein- und Ausgabebaugruppen, die mehrere Bits parallel bearbeiten. Verwendung für Peripheriegeräte und Sensoren/Aktoren, da schnell.
- **Serielle I/O** : Ein- und Ausgabebaugruppen, die Datenbits seriell bearbeiten. Verwendung meist für Peripheriegeräte, da langsam.
- **Bus I/O** : Kommunikation mit externem Bus.
- **Power Supply** : Versorgung aller Baugruppen mit den erforderlichen stabilisierten Spannungen.

Für Digitalrechner gibt es unterschiedliche Architekturen. Unter der Architektur versteht man das Operationsprinzip und die Hardwarestruktur eines Digitalrechners. Das verbreitetste Prinzip ist die **von Neumann-Maschine**¹⁰ [7.16].

Kennzeichnend für eine von Neumann-Maschine sind die streng serielle Bearbeitung¹¹ von Programmbefehlen und die Tatsache, dass kein prinzipieller Unterschied

10 John von Neumann (1903-1957). Berühmter amerikanischer Mathematiker. Er wurde auch durch seine schnellen elektronischen Computer bekannt; er entwarf 1952 den MANIAC 1, den ersten Computer mit einem flexiblen Speicherprogramm.

11 Serielle Verarbeitung ist für die Vermeidung von Hazard-Problemen (Kap. 7.2.2) erforderlich und schränkt dadurch die Bearbeitungsgeschwindigkeit ein, was zu immer höheren Taktfrequenzen geführt hat (z.Z. max. 3,5 GHz)

zwischen Daten und Programmbefehlen besteht, die deshalb auch im gleichen Speicherbereich stehen können. Das von **Neumann-Organisationsprinzip** kennt nur die folgenden Maschinen-Datentypen:

- Befehl (Instruction)
- Datum
- Adresse

Alle Operationen, die zu diesen drei Maschinen-Datentypen gehören, manipulieren die sogen. **von Neumann-Variable**. Diese Variable ist durch zwei Komponenten definiert:

von Neumann-Variable = (Identifikator, Wert)

Die erste Komponente dieser Variablen (Identifikator) bezeichnet eine logische oder physikalische **Adresse** (Kennzeichnung eines Speicherplatzes). Die zweite Komponente (Wert) ist der Inhalt der Adresse, die durch den Identifikator angegeben wird. Dieser Inhalt ist eine strukturlose Bitkette. Die Zuweisung einer Programmvariablen an die von Neumann-Variable erfolgt durch eine Wertzuweisung im Programm. Diesem Wert wird eine feste Adresse (Identifikator) zugewiesen.

Der Vorteil dieses Ansatzes liegt darin, dass der Speicherplatz im Arbeitsspeicher (RAM) sehr ökonomisch verwendet wird, da die Programmvariable stets an der gleichen Adresse abgelegt wird, obwohl sich ihr Wert im Laufe einer Programmausführung häufig ändern kann¹².

Alle drei oben genannten Maschinen-Datentypen manipulieren die von Neumann-Variable. Zu dem Maschinen-Datentyp „Befehl“ gehören im wesentlichen (bis auf Befehlszähler erhöhen) die beiden folgenden Operationen:

- 1.) Programmbefehl holen.
- 2.) Programmbefehl dekodieren und die Adresse des zugehörigen Datums berechnen.

Die Operationen, die mit den Maschinen-Datentypen „Datum“ und „Adresse“ verknüpft sind, können in drei Phasen eingeteilt werden:

- 1.) Datum (Adresse) aus dem Speicher holen.
- 2.) Datum (Adresse) bearbeiten.
- 3.) Datum (Adresse) in den Speicher zurückschreiben.

Die Interpretation, ob der Wert einer von Neumann-Variablen einen Befehl, ein Datum oder eine Adresse beschreibt, ist nur mit Hilfe von Maschinenzuständen möglich. Die **Kontrollstruktur** einer von Neumann-Maschine kennt daher drei entsprechende Maschinenzustände. Im Befehlszustand wird der Speicherinhalt als Befehl interpretiert und im Regelfall die Adresse des zugehörigen Datums berechnet. Die Information, ob eine Informationsstruktur als Datum oder als Adresse zu interpretie-

12 Dies war in der Anfangsphase der Digitalrechner-technik wichtig, da Arbeitsspeicher sehr teuer und knapp war, und in der Größenordnung von einigen Kilobit zur Verfügung stand. Heute haben Mikrorechner durchaus Arbeitsspeicher in der Größenordnung von einigen Gigabit.

ren ist, wird dem Befehlsformat hinzugefügt. Im Datumszustand wird das Datum gehalten, bearbeitet und wieder zurückgespeichert (Befehl ausführen). Entsprechendes gilt für den „Adresszustand“.

Jede Programmanweisung wird in einem starren Zwei-Takt-Zyklus aus Befehlszustand und Datumszustand abgearbeitet, der mit dem Befehlszustand beginnt. Der Kontrollfluss eines Programms ist streng sequentiell, wodurch in der Regel Programmanweisung nach Programmanweisung linear abgearbeitet wird. Abweichungen von dieser Regel sind nur möglich, wenn im Programm Verzweigungen, Wiederholungen, rekursive Prozeduraufrufe und Unterprogrammaufrufe stehen. Nach diesen **Programmsprüngen** ist der Kontrollfluss wieder streng sequentiell. Die Gesamtheit der Programmanweisungen

wird daher durch den von Neumann-Kontrollfluss in partiell geordnete Programmabschnitte zerlegt, die aber alle selbst vollständig geordnet sind. Zur physikalischen Steuerung dieser sequentiellen Programmabarbeitung genügt ein einziger **Programmzähler**. Er wird üblicherweise nach jedem abgearbeiteten Befehl um 1 erhöht (inkrementiert).

Zur Steigerung der Verarbeitungsgeschwindigkeit sind auch **parallele Rechnerarchitekturen** entwickelt worden. Bei einer parallelen Abarbeitung von Programmanweisungen muss ein ganzer Befehlsspeicher angelegt werden, der für jede gleichzeitig ausführbare Anweisung einen eigenen Befehlsspeicher enthält.

Für eine von Neumann-Maschine gibt es zwei unterschiedliche Strukturen, die **Princeton-Struktur** und die **Harvard-Struktur** (Bild 7-110). Der Princeton-Rechner hat einen Arbeitsspeicher, der sowohl die Daten als auch das Programm aufnimmt, und einen einzigen Bus, der auch **Universalbus** genannt wird. Der Harvard-Rechner enthält getrennte Speicher für die Daten und das Programm. Daher enthält er auch zwei interne Busse, den Befehls- und den Datenbus. Durch diese Teil-Parallelisierung wird die

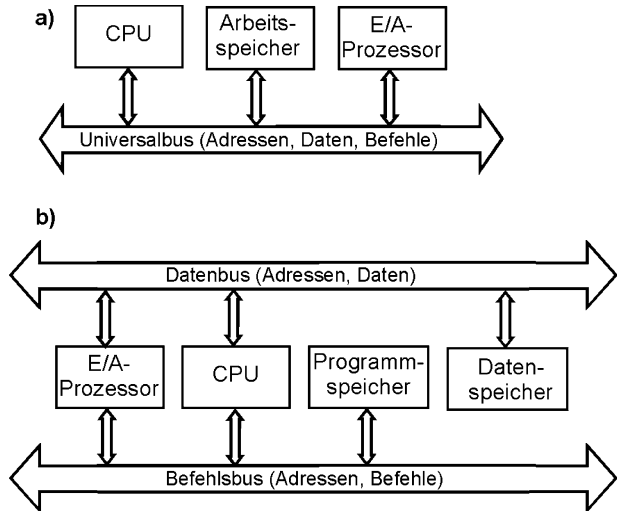


Bild 7-110: Unterschiedliche Architekturen der von Neumann-Maschine a) Princeton-Struktur
b) Harvard-Struktur

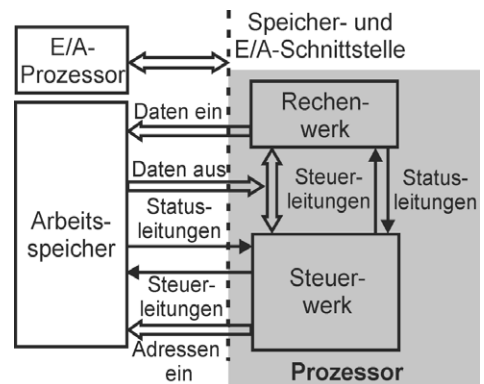


Bild 7-111: Verschaltung von Arbeitsspeicher und E/A-Prozessor mit dem Prozessor über die selbe Schnittstelle

Verarbeitungsgeschwindigkeit gegenüber der Princeton-Struktur gesteigert, da nun nicht mehr Befehle und Daten über den gleichen Bus transportiert werden müssen. Gleichzeitig kann die Breite des jeweiligen Busses der Harvard-Architektur an die unterschiedlichen Wortlängen im Programmspeicher (z. B. 32 bit) und Datenspeicher (z. B. 16 bit) angepasst und optimiert werden.

Bild 7-111 zeigt die Struktur eines Rechners mit Princeton-Architektur etwas detaillierter. Die CPU (Mikroprozessor) ist nochmals in **Rechenwerk** und **Steuerwerk** aufgeteilt. Die Arbeitsabläufe finden darin wie folgt statt:

- Das Steuerwerk ermittelt die Arbeitsspeicheradresse des nächsten Befehls. Diese Adresse wird über Adressleitungen in den Speicher übertragen.
- Über die Datenausgabe wird der Befehl aus dem Speicher in das Steuerwerk geholt. Dort wird er interpretiert und weitere Funktionen werden ausgelöst.
- Die Adresse des Operanden im Speicher wird ermittelt. Diese Adresse wird in den Arbeitsspeicher gelegt und der Operand steht nach dem Speicherzugriff auf den Datenausgabeleitungen bereit.
- Das Rechenwerk führt unter Kontrolle des Steuerwerks die Operation durch. Das Ergebnis wird im Rechenwerksregister abgelegt.
- Das Steuerwerk geht zur Bearbeitung des nächsten Befehls über.

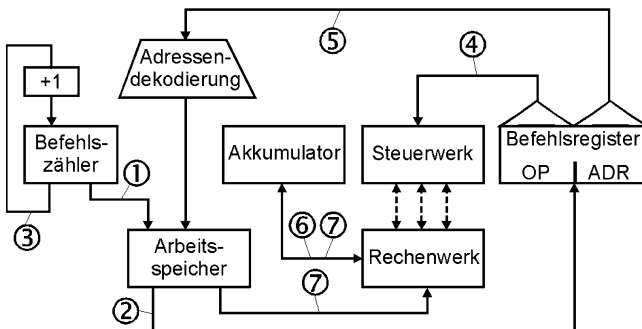


Bild 7-112: Grundorganisation eines Additionsbefehls

Das Steuerwerk erhält den Operationsteil eines Befehls aus dem Befehlsregister, das auch Instruktionsregister genannt wird. Zur Ermittlung der Befehlsadresse dient der Befehlszähler, dessen Inhalt die Adresse des nächsten Befehls ist. Bei einem linearen Programmablauf wird der Befehlszähler nach jedem Befehl um 1 erhöht. Um diesen Ablauf zu verdeutlichen,

soll unter Zuhilfenahme von Bild 7-112, das die Einzelelemente des Rechen- und Steuerwerks als separate Blöcke enthält, beispielhaft eine Addition schrittweise durchgeführt werden. Die einzelnen Vorgänge sind:

- 1.) Inhalt des Befehlszählers an den Arbeitsspeicher übergeben.
- 2.) Der unter 1.) adressierte Speicherinhalt wird gelesen, als Befehl interpretiert und ins Befehlsregister eingetragen.
- 3.) Der Befehlszähler wird um 1 erhöht.
- 4.) Der Operationsteil aus dem Befehlsregister wird im Steuerwerk dekodiert.
- 5.) Der Adressteil des Befehlsregisters wird als Adresse an den Arbeitsspeicher gegeben.

- 6.) Der Inhalt des Akkumulators wird ins Rechenwerk gegeben.
- 7.) Der Inhalt der in 5.) adressierten Speicherstelle wird ebenfalls ins Rechenwerk gegeben, zum Akkumulatorinhalt addiert und in den Akkumulator zurückgegeben.

Arbeitsspeicher oder **Akkumulatoren** sind lokale Speicher, die Operanden oder Verknüpfungsergebnisse aufnehmen. Realzeitrechner enthalten meistens nur ein einziges Universalrechenwerk. Es dient zur Verknüpfung von arithmetischen und logischen Operationen und für weitere Operationen wie etwa der Berechnung von Operandenadressen. Wesentlicher Bestandteil des Rechenwerks ist neben den Arbeitsregistern die Verknüpfungslogik. Beide zusammen werden ALU (Arithmetic Logic Unit) genannt und haben die Aufgabe, Verknüpfungen zwischen den Operanden herzustellen.

Die Architektur der Harvard-Maschine findet sich vor allem bei **Signalprozessoren**. Diese Prozessoren haben neben der genannten Architektur noch zwei weitere Elemente, die die Anzahl der je Sekunde möglichen Operationen gegenüber konventionellen Prozessoren erhöhen. Dies sind zusätzliche spezielle Hardware-Komponenten wie Addierer, Multiplizierer, Coprozessoren und das Fließband (pipeline)-Konzept. Hiermit sind z. B. Multiplikationszeiten bei Festkommazahlen bis zu 15 ns zu erreichen. Ein konventioneller Mikroprozessor ist ungefähr 50 mal langsamer.

Analoge Signalbearbeitung und -auswertung ist, besonders wenn höhere Genauigkeiten gefordert werden, teuer und störanfällig. Deshalb versucht man die Informationsbearbeitung in mechatronischen Systemen weitestgehend auf digitaler Ebene, also nach erfolgter Analog-/Digitalwandlung vorzunehmen. Signale mit hohen Grenzfrequenzen müssen dann aber mit sehr hohen Abtastraten digitalisiert werden. Die dabei entstehenden hohen Datenmengen müssen anschließend mit hoher

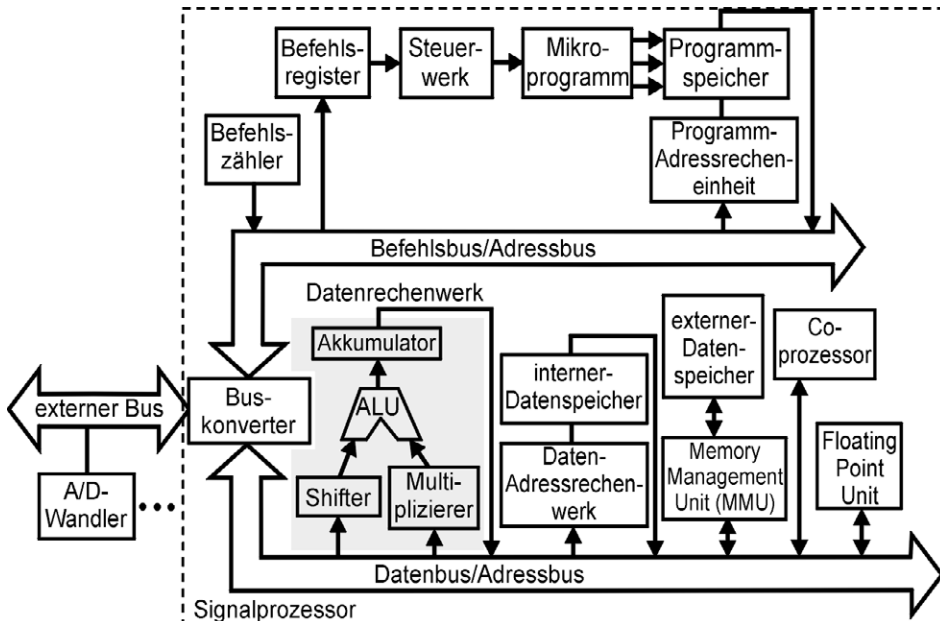


Bild 7-113: Prinzipieller Aufbau eines Signalprozessors

Geschwindigkeit weiterverarbeitet werden. Mit Signalprozessoren kann man Signale verarbeiten, die mit Grenzfrequenzen bis etwa 20 MHz abgetastet wurden. Somit gibt es viele Anwendungen für Signalprozessoren wie digitale Filterung, Telekommunikation, Bewegungssteuerung von Robotern, Spracherkennung, Bildverarbeitung und vieles mehr.

Bild 7-113 zeigt den prinzipiellen Aufbau eines Signalprozessors. Wie bereits erwähnt, ist die Struktur dieser Harvard-Maschine dadurch gekennzeichnet, dass der interne Bus in einen Befehlsbus und einen Datenbus aufgeteilt ist. In diese beiden Busse sind auch jeweils die Adressbusse integriert. Der ebenfalls noch vorhandene Steuerbus ist übersichtshalber weggelassen.

Der Buskonverter wird benutzt, um einzelne Signale zwischen Befehls- und Datenbus auszutauschen. Das Datenrechenwerk entnimmt seine Eingabeparameter dem Datenbus und verarbeitet diese Daten. Hierfür stehen ein Schieberegister (Shifter), ein schneller Multiplizierer für Festkommazahlen, eine arithmetisch, logische Einheit (ALU) und ein aufsummierendes Akkumulatorregister zur Verfügung.

Der Zugriff auf einzelne Worte im Programm- und Datenspeicher wird jeweils durch spezifische Adressrecheneinheiten beschleunigt. Die Arbeit des Zentralprozessors kann durch spezielle, leistungsfähige Coprozessoren unterstützt werden. Beispiele für solche Prozessoren sind Gleitkommaprozessoren (Floating Point Unit), Prozessoren für die Verwaltung des Arbeitsspeichers (Memory Management Unit) oder spezielle Bildverarbeitungsprozessoren (z. B. für die Fast Fourier Transformation). Bei Verwendung solcher Coprozessoren geht die Aktivität immer von der Zentraleinheit (Datenrechenwerk) aus, man spricht vom **Slave Modus** des Coprozessors.

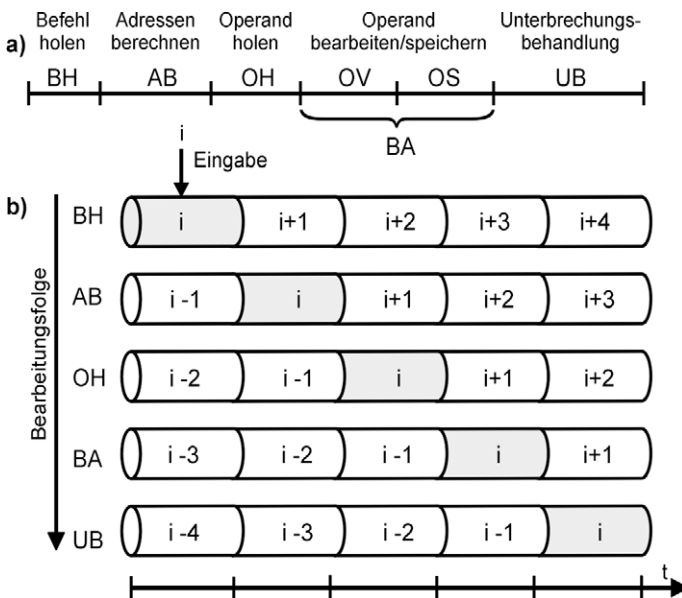


Bild 7-114: Anwendung des Fließbandprinzips auf den Befehls-Bearbeitungszyklus einer von Neumann Maschine
a) Sequentieller Befehls-Verarbeitungszyklus
b) 5-fach Pipeline

Die Zentraleinheit holt den Befehl, führt notwendige Adressberechnungen durch und sendet danach den Befehl mit den dazugehörigen Daten dem Coprozessor zur Ausführung. In einem Identifikationsfeld des Befehls wird spezifiziert, welcher Coprozessor den Befehl ausführen soll.

Eine weitere Leistungssteigerung erreicht man mit dem **Fließbandprinzip (pipelining)**, das sowohl bei Signalprozessoren aber auch bei konventionellen Prozessoren benutzt wird. Hierdurch kann die Anzahl der Operationen (Instruktionen), die ein Rechner pro Zeit-

einheit¹³ durchführen kann, deutlich erhöht werden. Ein Fließband von Verarbeitungseinheiten entsteht, wenn diese linear wie ein Fließband angeordnet sind. Mit Hilfe dieser Einheiten können gleichartige Befehle nacheinander und verschiedenartige Befehle gleichzeitig auf einem Informationsstrom (meist Datenstrom) ausgeführt werden. Typische Anwendungen des Prinzips erfolgen für den Befehls- Bearbeitungszyklus einer Multiplikation oder Addition zweier ganzer (Festkomma-) oder reeller (Gleitkomma-) Zahlen.

In Bild 7-114 a) sind nochmals die einzelnen Schritte des Befehls-Bearbeitungszyklus einer von Neumann-Maschine dargestellt, wobei die beiden Teilschritte OV und OS häufig als zusammengesetzter Schritt Befehl ausführen, BA, beschrieben werden. Diese Befehlsschritte werden normalerweise sequentiell hintereinander abgewickelt. In Bild 7-114 b) ist eine 5-stufige Befehlspipeline dargestellt, deren Befehlsnummern durch i , $i+1$ usw. beschrieben werden. Der Befehl mit der Nr. i wird nacheinander abgearbeitet. Die fünf Hardware-Module, die die unterschiedlichen Schritte des Befehls-Verarbeitungszyklus repräsentieren, bearbeiten parallel fünf verschiedene Befehle. Gleichzeitig wird der Befehl i geholt, die Adresse für Befehl $i-1$ berechnet, der Operand für Befehl $i-2$ geholt, Befehl $i-3$ ausgeführt und für den Befehl $i-4$ wird die Unterbrechungsbehandlung durchgeführt. Durch die gleichzeitige Bearbeitung von fünf Befehlen wird die Rechenleistung eines Mikroprozessors deutlich gesteigert.

Beim Fließbandkonzept, das wie beim Signalprozessor intern auf die Befehlsausführung und auf arithmetische Operationen angewendet wird, spricht man von **Mikro-Pipelining**. Rechner mit Mikro-Pipelining sind immer noch von Neumann-Rechner, da weder die Struktur noch das Operationsprinzip verändert werden. Erst wenn vollständige Rechenwerke oder komplette Prozessoren zu einer Makro-Pipeline zusammengefasst werden, liegt eine von der von Neumann-Struktur abweichende Rechnerstruktur vor, die aber immer noch das von Neumann-Operationsprinzip benutzt. Wird auch das Operationsprinzip eines Rechners verändert, so hat man Rechner die vollständig anders arbeiten wie beispielsweise die **Feld-** oder **Vektor-Rechner**. Wichtiger Vertreter dieser sehr aufwendigen Konzepte ist z. B. der Feldrechner CRAY-2 und Nachfolger.

Geht man von der Princeton-Architektur des von Neumann-Rechners aus und möchte diesen prinzipiellen Rechneraufbau nicht verändern, so gibt es für die Steigerung der Effizienz neben dem Mikro-Pipelining nur noch den Ansatz des **RISC-Rechners** (Reduced Instruction Set Computer). Wie der Name bereits sagt, wird hier der Befehlssatz des Rechners auf 30 -100 Befehle verkleinert. Die Befehlsausführung wird dadurch deutlich schneller, da nur einfache Befehlssätze mit einheitlichem Befehlsformat abgearbeitet werden. Konventionelle Rechner nutzen einen Satz von mehr als 100 komplexen Befehlen (CISC-Maschinen: Complex Instruction Set Computer), die verschiedenartige Formate haben und in der Regel **mikroprogrammiert** sind. Die Befehlsbearbeitung erfolgt dabei durch Interpretation eines Befehls mit einem Mikrocode, wodurch mehrere Taktzyklen benötigt werden und eine komplexere Steuerung im Zentralprozessor erforderlich ist.

13 Dieser Wert wird in MIPS (Million Instructions Per Second) gemessen. Der Mikroprozessor Intel Core i 7 5960x besitzt beispielsweise eine Rechenleistung von 336.000 MIPS bei 3,5GHz.

Integriert man neben der Zentraleinheit auch den Arbeitsspeicher und Peripheriebausteine auf einem Chip, so spricht man von einem **Mikrocontroller**. Solche Systeme, die sich durch ihren geringen Platzbedarf und Stromverbrauch, sowie durch hohe Zuverlässigkeit auszeichnen, sind auch für kleine Mess-, Steuerungs- und Regelungsaufgaben gut geeignet. Die Anwendungen erstrecken sich von Haushaltsgeräten (Waschmaschine, Herdsteuerung, usw.) über Geräte der Unterhaltungselektronik (TV, Video, HiFi, CD, usw.) bis hin zu professionellen Anwendungen in den Bereichen KFZ, Telekommunikation und Industrie. Gerade im Automobilbereich gibt es viele Anwendungen in Form von ABS, ASR, ESP, Airbag, Motorsteuerung und Navigationssystem, die ohne Mikrocontroller gar nicht möglich gewesen wären, und die typische Beispiele für mechatronische Systeme sind.

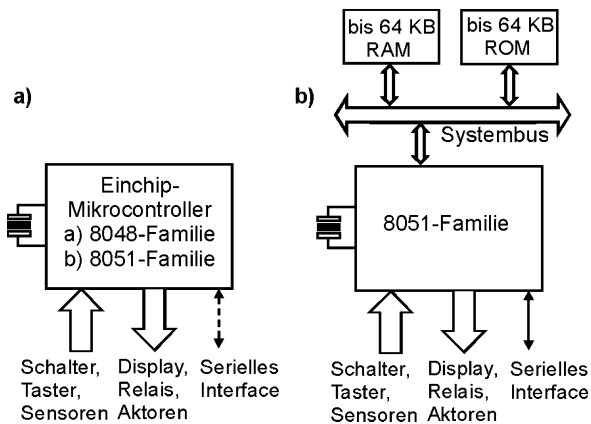


Bild 7-115: Einsatz von Mikrocontrollern in Steuerungssystemen a) für einfache Anwendung mit maskenprogrammiertem ROM
b) mit externem Speicher für komplexe Anwendungen

Je nach Architektur stehen 4 Bit-, 8 Bit-, 16 Bit- und 32 Bit-Mikrocontroller zur Verfügung, wobei die Applikation die erforderliche Komplexität vorgibt. Typische Vertreter solcher Mikrocontroller sind die 8048- und 8051-Serien, die von Intel entwickelt wurden und inzwischen von anderen Herstellern wie Siemens und Philips in Lizenz gebaut werden. In Bild 7-115 sind die Einsatzmöglichkeiten solcher Mikrocontroller-Steuerungen dargestellt. Der Mikrocontrollerchip enthält in der Regel ein herstellenseitig maskenprogrammiertes ROM, das beim 8051-Prozessor durch einen externen Programm- und Datenspeicher

von max. 64 KByte ergänzt werden kann. Der in CMOS-Technologie hergestellte Controller 80C51 verfügt neben dem 8051-Rechnerkern mit einem 8-Bit Datenbus über 4 K ROM, 128 Byte RAM, 2 x 16 Bit Timer und eine durch das 40-Pin-Gehäuse begrenzte Anzahl von E/A-Ports. Andere Derivate wie die 80C16x Mikrocontroller mit 16-Bit Datenbus verfügen über zusätzliche Funktionen wie A/D-Schnittstellen, DMA-Kanäle, PCMCIA-Interface, LCD-Controller, CAN-Controller, Fuzzy-Hardware und vieles mehr.

Der Einchip-Mikrocontroller PCB 83C552 der Firma Phillips (Bild 7-116) ist ein in CMOS-Technologie gefertigter Mikrocontroller, der zur 80C51-Familie gehört. Er wird vor allem zur Lösung von Echtzeitaufgaben verwendet, z. B. im Bereich industrieller Steuerungen, in der Medizintechnik, bei Instrumentierungsaufgaben sowie im Kraftfahrzeugsektor, wo er insbesondere bei der Motor- und Getriebesteuerung verwendet wird. Außer den Standardfunktionen des 80C51 weist der Mikrocontroller PCB 83C552 noch weitere Hardwarefunktionen auf, die ihn für die genannten Anwendungen besonders geeignet machen.

Beispielsweise ist auf dem Chip ein 10 Bit A/D-Wandler integriert, der bis zu 8 Ein-

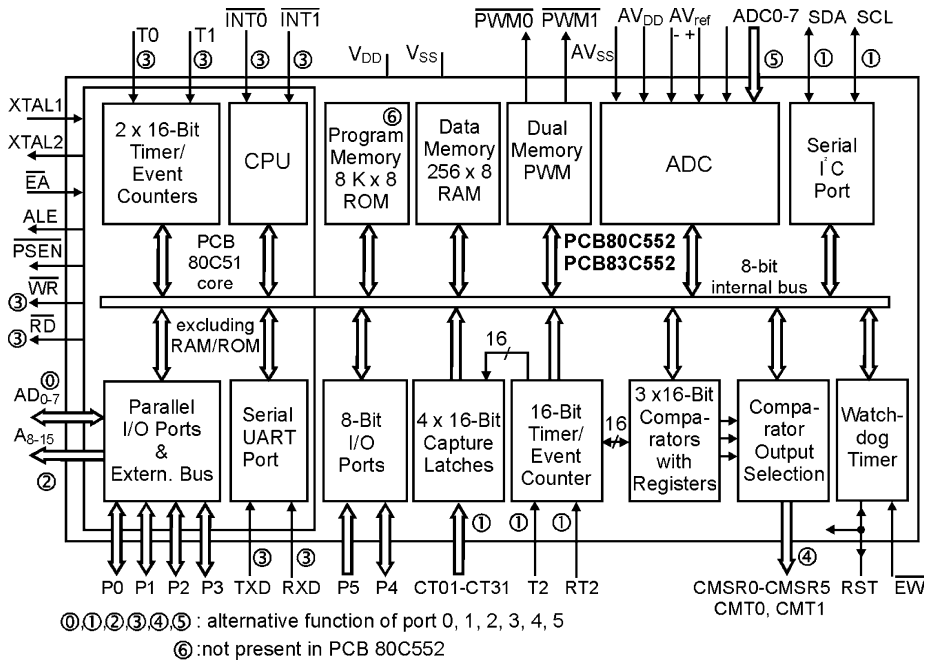


Bild 7-116: Blockschaltbild des Einchip-Mikrocontrollers PCB 83C552

gangsgrößen im Multiplexverfahren digitalisieren kann. Zusätzlich zu den beiden 16-Bit-Standard-Timern des 80C51 ist ein dritter Timer eingebaut, der besonders für Echtzeitanwendungen günstig ist. Mit diesem Timer sind vier Capture-Register verknüpft, die unabhängig voneinander mit dem Zählerinhalt geladen werden, wenn ein externes Signal sie triggert und dann einen Interrupt erzeugt. Drei Compare-Register, die ebenfalls mit diesem Timer verknüpft sind, werden bei jedem Maschinenzyklus mit ihm verglichen; bei Identität wird ebenfalls ein Interrupt erzeugt. Ein weiterer Interrupt kann bei einem 8-Bit oder 16-Bit-Überlauf erzeugt werden.

Mit dem Watch-Dog Timer lässt sich die ordnungsgemäße Abarbeitung des Programms in der CPU überwachen. Bei fehlerhafter Abarbeitung (z. B. Endlosschleife) wird ein Watchdog-Timer Reset ausgelöst, der die On-Chip-Peripherie zurücksetzt und den Instruction Pointer mit dem Wert Null lädt. Er hat einen 11-Bit-Vorteiler und einen 8-Bit-Zähler, der nach seiner Aktivierung den Timer Reset auslöst, wenn er nicht alle 2-500 ms neu beschrieben wird.

Zwei 8-Bit-PWM-Ausgänge (pulsbreitenmodulierte Ausgänge) können analoge Signale zur Verfügung stellen. Sie erzeugen Ausgangssignale, deren Länge und Intervalle programmierbar sind. Die Wiederholfrequenz kann zwischen 92 Hz und 23,5 kHz, das Pulsweitenverhältnis zwischen 0 und 255/255 mit Inkrementen von 1/255 variieren.

Über die seriellen Schnittstellen (UART, I²C-Bus)¹⁴, kann das System modular er-

14 UART: Universal Asynchronous Receiver Transmitter (Serielle Datenschnittstelle)

I²C-Bus-Schnittstelle: serieller E/A-Port mit byteorientierten Master- und Slave Funktionen.

weitert werden. Von der hier beschriebenen Schaltung existieren drei Versionen:

- PCB 83C552 : 8 Kbytes maskenprogrammierbares ROM, 256 Bytes RAM
- PCB 80C552 : ROM-lose Version des PCB 83C552
- PCB 87C552 : EPROM-Version

Der Befehlssatz des PCB 83C522 ist der gleiche wie beim 80C51. Er besteht aus über 100 Befehlen: 49 Ein-Byte-, 45 Zwei-Byte- und 17 Drei-Byte-Befehle. Zum Beispiel werden mit einem 12 MHz-Quarz (Max. 16 Mhz) 58% aller Befehle in 1 μ s und 40% in 2 μ s ausgeführt. Die Befehle Multiplizieren und Dividieren erfordern 4 μ s.

Der PCB 83C522 lässt sich auch als arithmetischer Prozessor betreiben, und zwar kann er Operationen sowohl in binärer als auch in BCD-Arithmetik ausführen. Er kann wahlweise mittels Software auf zwei Betriebsarten mit verringerter Aktivität (Idle und Power Down) geschaltet werden.

7.4.1.2 Software für Mikroprozessoren

Die Entwicklung der Programme für Mikroprozessoren erfolgt in der Regel mit Hilfe von **Entwicklungssystemen**. Dies kann direkt auf dem Zielrechner oder auf einem anderen **Host-Rechner** erfolgen. Auf einem Entwicklungssystem werden in einer Hochsprache wie C, C++, Pascal oder Python oder in einer **Assemblersprache** die Quellprogramme geschrieben und anschließend durch ein Compiler-Programm in die Objektprogramme übersetzt, die den Maschinencode des jeweiligen Prozessors enthalten. Weiterhin kann man mit Hilfe von **Debugger**-Programmen Fehler im Anwenderprogramm suchen und die lauffähigen Programme in PROMs oder EPROMs eingeben. Die Entwicklungssysteme enthalten eine umfangreiche System-Software und eine geeignete Peripherie. Hiermit kann eine Hardwareverbindung zwischen Mikroprozessor und Entwicklungssystem hergestellt werden.

Die Softwareerstellung mittels Host-Rechner (z.B. PC) erfolgt mit einem speziellen **Cross-Compiler**. Dieser übersetzt das Quellprogramm in den Maschinencode des Zielsystems, ein Monitorprogramm erlaubt das Herunterladen des Maschinencodes auf den Mikrorechner.

Bei Mikrocontrollern sind Cross-Compiler für eine ganze Mikrocontroller-Familie ausgelegt. Es gibt komfortable Bedienoberflächen unter Windows und integrierte Entwicklungssysteme (IDE: integrated development environment) mit einer Kombination aus Editor, Compiler und Debugger. Bei großen Stückzahlen, wo die Leistung und die Speichergröße der Mikrocontroller aus Kostengründen gering gehalten werden muss, wird in der Regel in der Assembler-Sprache oder direkt in Maschinencode programmiert.

Für das Debugging solcher Programme kann man in einfachen Fällen Simulatoren verwenden, die den Mikrocontroller auf dem Entwicklungsrechner abbilden. Will man unter Berücksichtigung des Zeitverhaltens das Zusammenspiel von Hard- und Software testen, so verwendet man Entwicklungssysteme mit **In-Circuit-Emulatoren**, bei denen der Mikrocontroller direkt in der Schaltung durch einen

Emulator ersetzt wird. Bei Emulatoren unterscheidet man zwischen echtzeitfähigen und nicht echtzeitfähigen Varianten.

7.4.2 Anwendungsspezifische Prozessoren und Bauelemente

Die heutigen VLSI-Techniken erlauben es auch hochintegrierte anwendungsspezifische Hardware die sogenannten **ASIC's** (application specific integrated circuit) zu entwickeln. Dabei unterscheidet man zwischen zwei Möglichkeiten: der vollkunden-spezifische (full-custom) Entwurf und der halbkunden-spezifische (semi-custom) Entwurf. Beim full-custom Entwurf wird die VLSI-Schaltung individuell für das Schaltungsproblem bis zur untersten Ebene optimiert, so dass kleinste Chipflächen, höchste Schaltungsgeschwindigkeiten und geringste Verlustleistungen erreicht werden. Der Entwurf solcher Schaltungen ist aufwendig und teuer und wird daher in der Regel nur für Standardschaltungen in großen Stückzahlen beschränkt. Ein semi-custom Entwurf nutzt vorhandene Grundstrukturen von Bauelementen und Schaltungen nach vorgegebenen Bibliotheken. Die Entwurfs- und Fertigungskosten sind deshalb geringer, auf Kosten optimaler Eigenschaften. Bild 7-117 zeigt einen Überblick über programmierbare integrierte Schaltungen, die sowohl als full- als auch als semi-custom Schaltungen gefertigt werden. Die im Bild abgekürzten Technologien haben dabei folgende Bedeutung:

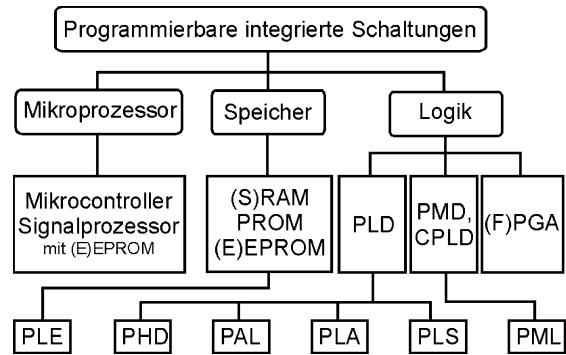


Bild 7-117: Programmierbare integrierte Schaltungen

- (S)RAM: (Static) Random Access Memory
Statischer flüchtiger Speicher mit wahlfreiem Zugriff. Wird bei einigen programmierbaren Gate-Arrays (CMOS-PGA) zur dynamisch veränderbaren Verbindungsprogrammierung benutzt.
- PROM: Programmable Read-Only Memory
Einmalig programmierter Festwert/Nur-Lese-Speicher. Gleiche Technologie wie für bipolare PLD's mit trennbaren Verbindungen (Fuse).
- (E)EPROM: (Electrically) Erasable PROM
Reprogrammierbarer, nicht flüchtiger CMOS-Festwertspeicher. Die Programmierung erfolgt elektrisch, das Löschen entweder auch elektrisch oder durch UV-Licht, das über ein Quarzglasfenster auf den Chip geleitet wird.
- PLD: Programmable Logic Device (Programmierbare Logik)
Sammelbegriff für alle anwenderprogrammierbaren Logikbauelemente mit niedriger Komplexität.
- PMD: Programmable Multilevel/Macro Device
Architekturen mit hoher Funktionsdichte, welche weder den PGA, noch den

CPLD zuzuordnen sind.

- CPLD: Complex- oder Clustered PLD
Sammelbegriff für komplexe PLD's mit hoher Funktionsdichte auf einem Kristall. Kennzeichnend ist eine definierte Verzögerungszeit, wie sie auch bei einfachen PLD's aufgrund der fest vorgegebenen Verdrahtungswege vorliegt.
- (F)PGA: (Field) Programmable Gate-Array
Anwenderprogrammierbares Gate-Array mit einer sehr großen Anzahl von frei verbindbaren Basiselementen (z.B. Gattern), sowie Verdrahtungskanälen, die teilweise in mehreren Ebenen angeordnet sind. Neben flüchtigen, reprogrammierbaren Versionen auf RAM- Basis gibt es nichtflüchtige Versionen, die über Fuses programmiert werden.
- PLE: Programmable Logic Element
Architekturbezeichnung für Speicherbauelemente, z.B. PROM's
- PHD: Programmable High-Speed Decoder
Herstellerbezeichnung für (schnelle) programmierbare Adressdekoder.
- PLA: Programmable Logic Array
Architekturbezeichnung für PLD's mit frei programmierbarem UND- und ODER-Array und wählbarer Ausgangspolarität.
- PAL: Programmable (AND-) Array- Logic
Architekturbezeichnung für PLD's mit festem Oder-Array
- PLS: Programmable Logic Sequencer
Architekturbezeichnung für universell programmierbare Zustandsmaschinen- (State Machine) Bauelemente mit Array-Aufbau wie beim PLA.
- PML: Programmable Makro Logic
Herstellerbezeichnung für eine PMD- Architektur mit hoher Funktionsdichte, basierend auf rückgekoppelten NAND- Gattern mit bis zu 256 Eingängen.

Viele dieser unterschiedlichen programmierbaren Bausteintypen haben eine vorgegebene Hardwarestruktur, die einmalig auf bestimmte Logikfunktionen hin programmiert werden kann. Die in Bipolar- oder Mischtechnologie (BiCMOS) hergestellten PLD's enthalten in der Regel als Programmiererelement so genannte **Fuses** (Schmelzsicherung) aus Nickel-Chrom- oder Titan-Wolfram-Material, die gezielt durch einen Stromstoß zerstört werden können, wodurch eine potentielle Verbindung unterbrochen wird. Es gibt auch reprogrammierbare Zellen, die adressierbare Verbindungszellen darstellen.

In Bild 7-118 sind einige Grundarchitekturen verschiedener PLD's dargestellt. Es handelt sich um eine vereinfachte Darstellung zur Verdeutlichung des Prinzips, bei der die Vielzahl der Eingänge in ein Gatter nur durch jeweils eine Leitung dargestellt ist. Die Programmiermatrizen sind als Kreuzfeldsymbol dargestellt, bei dem die Programmierung durch "Durchbrennen" der Sicherung erfolgt, die zwei sich kreuzende Leitungen miteinander verbindet.

Die im Bild 7-118 oben abgebildete PLE-Struktur wird typischerweise für PROM/ EPROM verwendet. Diese einfachste Struktur, bei der lediglich die ODER-Matrix

programmierbar ausgeführt ist, findet man bei PROM's in bipolar-, MOS- oder CMOS-Technologie. Die Eingänge sind über eine festverdrahtete Dekoder-Funktion mit UND-Gattern verbunden, deren Ausgänge über programmierbare Verbindungen einer Reihe von ODER-Gattern zugeführt werden. Diese rein kombinatorischen Bausteine werden in Anwendungen benutzt, bei denen einer Eingangsinformation in Form einer Adresse ein bestimmtes Bitmuster an den Ausgängen zugeordnet werden soll. Für allgemeine Logikanwendungen ist diese Architektur jedoch zu starr, da über die festverdrahtete Dekodereinheit alle möglichen Belegungen am Eingang dekodiert werden und folglich auch mit einem Bitmuster belegt werden müssen.

Für nicht speicherorientierte Anwendungen eignet sich daher die in der Mitte von Bild 7-118 dargestellte PAL-Architektur. Bei kombinatorischen PAL's folgt auf eine programmierbare UND-Matrix eine starr in Gruppen verdrahtete ODER-Matrix. Hierdurch kann bestimmten Eingangsbelegungen ein relativ starres Ausgangsbitmuster zugeordnet werden. Vorzugsweise werden solche Bausteine in Adressdekoder-Anwendungen in Mikrorechnersystemen verwendet.

Im unteren Teil von Bild 7-118 ist die PLA-Architektur dargestellt. Durch ihre vielfältigen Programmiermöglichkeiten bietet sie höchste Flexibilität. Mit dieser Struktur können praktisch beliebige kombinatorische Logikfunktionen realisiert werden.

Werden der PLA-Architektur noch zusätzlich programmierbare Flip-Flops hinzugefügt (Bild 7-119), so erhält man ein PLS. Die Bezeichnung Sequencer deutet darauf hin, dass es

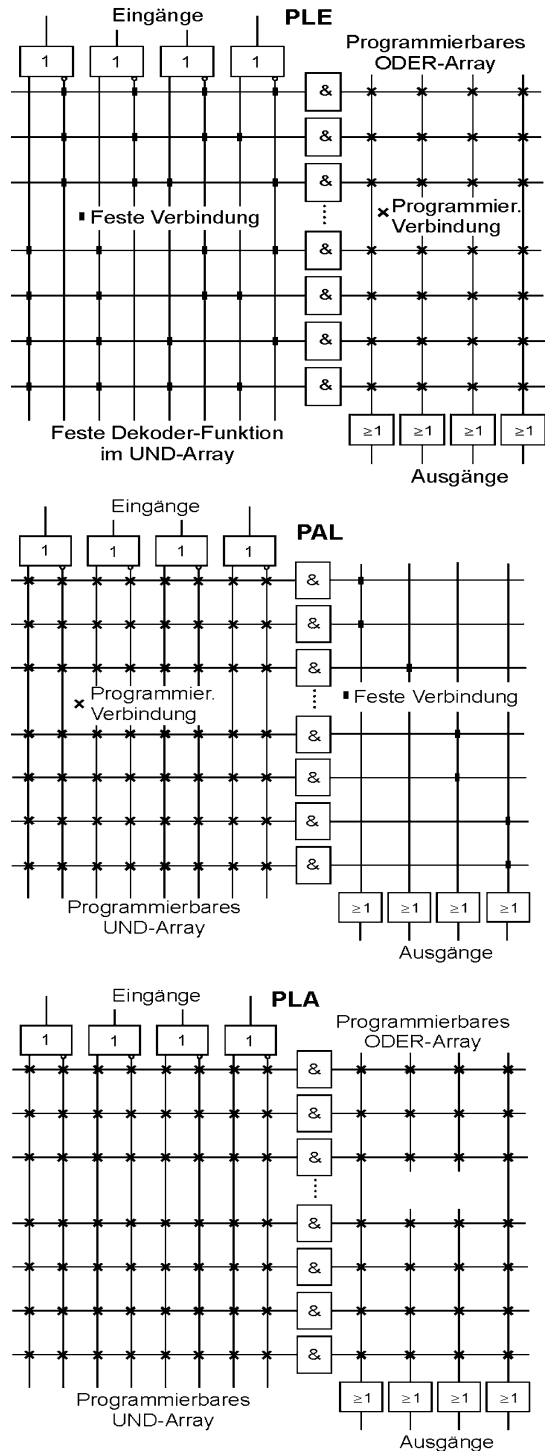


Bild 7-118: Grundarchitekturen programmierbarer Logikbausteine

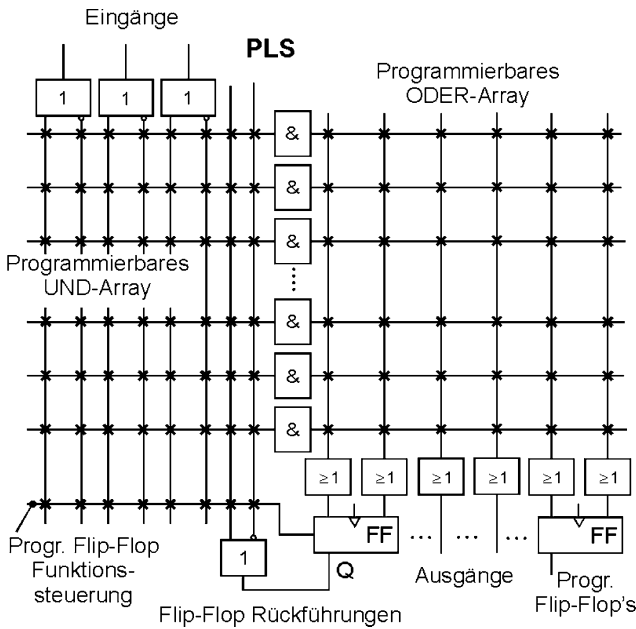


Bild 7-119: PLS-Grundarchitektur

sich hierbei um Bauelemente für sequentielle Anwendungen handelt. Mehrere universelle Flip-Flop-Einheiten, die wahlweise für die Eingangs- oder Ausgangssignalspeicherung verwendet werden können, gestatten dem Entwickler die Programmierung von „Automaten“ (Kap. 7.2), bei denen die Ausgangswerte nicht nur von der Eingangsbelegung, sondern auch von der „Historie“ abhängig sind.

8 Simulation

In der Entwicklungs- oder Planungsphase von komplexen Bewegungssystemen mit unübersichtlichen Bewegungsvorgängen ist es heute vielfach üblich, solche Systeme nicht an körperlich vorhandenen Prototypeneinrichtungen zu erproben und zu optimieren, sondern sie auf einem Digitalrechner zu simulieren. Die dazu erforderliche Software wird als **Simulationssystem** bezeichnet. Solche Simulationssysteme gibt es zur Simulation der Kinematik (Beispiel Robotersimulationssystem), zur Simulation dynamischer Vorgänge (Beispiel regelungstechnisches Simulationssystem) oder auch zur Simulation statischer Belastungen (Beispiel Finite-Element-System). Weitere Simulationssysteme befassen sich mit der Simulation des elektrischen Verhaltens von Schaltungen und Bewegungssystemen.

In den vorherigen Kapiteln ist schon häufiger die Simulation dynamischer Vorgänge auf Digitalrechnern angesprochen worden und es wurden mit einem Simulationssystem ermittelte Ergebnisse solcher Vorgänge gezeigt. Die Technik der numerischen Simulation bezieht sich natürlich auf die mathematischen Modelle realer Systeme, die man im Modellbildungsverfahren (s. Kap. 2) ermittelt hat. Hat man als mathematisches Modell eine lineare Differentialgleichung gefunden, so ist die geschlossene Lösung mit konventionellen Methoden ohne Rechneinsatz möglich, aber sehr zeitaufwendig. Insbesondere wenn man verschiedene Fälle ausrechnen will, oder die Auswirkungen von Parameteränderungen studieren möchte, kann der Einsatz eines Simulationssystems auf einem Digitalrechner mit grafischer Ausgabe viel Zeit und Mühe sparen und die Visualisierung der Ergebnisse sehr gut unterstützen. Ist das Modell sogar nichtlinear, so ist eine Lösung der Differentialgleichung in der Regel nur noch durch numerische Näherungsverfahren möglich, die man sinnvollerweise rechnergestützt durchführen sollte.

Als einfache lineare DGL mit konstanten Koeffizienten war beispielsweise die eines gedämpften Einmassenschwingers behandelt worden:

$$m\ddot{x} + d\dot{x} + kx = 0 \quad . \quad (8.1)$$

In Kapitel 4 wurde gezeigt, wie diese Gleichung mit Hilfe des klassischen Ansatzes gelöst werden kann und in Kapitel 7 waren anhand dieser Darstellungsform andere Beschreibungsmethoden für ein dynamisches System, wie etwa die Übertragungsfunktion, abgeleitet worden. Ebenfalls in Kapitel 7 war auch eine andere Darstellungsform, nämlich die Zustandsraumdarstellung vorgestellt worden. Anstatt wie in Gl. (8.1) das System durch eine DGL 2. Ordnung zu beschreiben, wird in der Zustandsraumdarstellung das System durch 2 DGL's 1. Ordnung beschrieben:

$$\dot{x}_1 = x_2 = v \quad \dot{x}_2 = -\left(\frac{k}{m}\right)x_1 - \left(\frac{d}{m}\right)x_2 = \dot{v} = a \quad . \quad (8.2)$$

Von dieser Darstellungsform gehen viele Verfahren zur Simulation dynamischer Systeme aus, weil sich hieran die Vorgehensweise zur Lösung der Differentialgleichungen auf die einfache Integration der Zustandsgrößen zurückführen lässt:

$$x_1 = \int_0^t x_2 dt \quad x_2 = -\left(\frac{k}{m}\right) \int_0^t x_1 dt - \left(\frac{d}{m}\right) \int_0^t x_2 dt \quad . \quad (8.3)$$

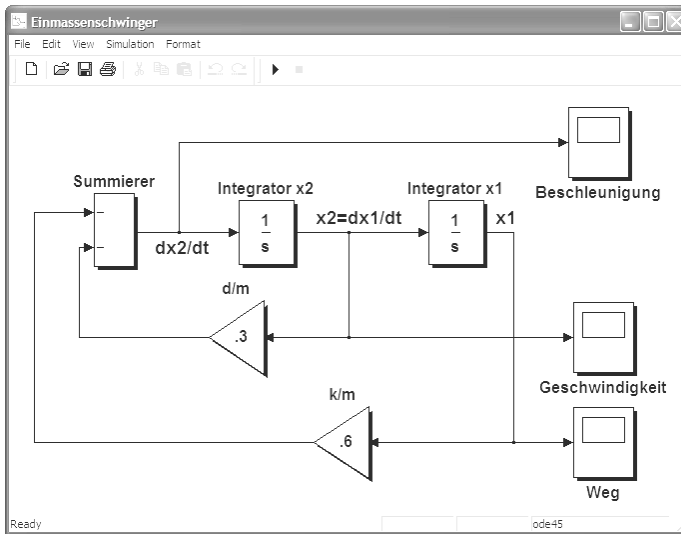


Bild 8-1: Simulationsdiagramm für die DGL eines Einmassenschwingers mit Dämpfung

Viele Simulationssysteme benutzen für die Darstellung des zu simulierenden dynamischen Vorgangs eine Grafikform in der Art der Blockdiagramme. In Bild 8-1 ist ein solches Simulationsdiagramm dargestellt, das die gleichen Zusammenhänge wie die Gl. (8.1) und (8.2) beschreibt. Die Darstellung erfolgt mit dem Simulationssystem **Simulink**¹. Der Zusammenhang mit der Zustandsraumdarstellung in Gl. (8.2) ist sofort erkennbar. Beide

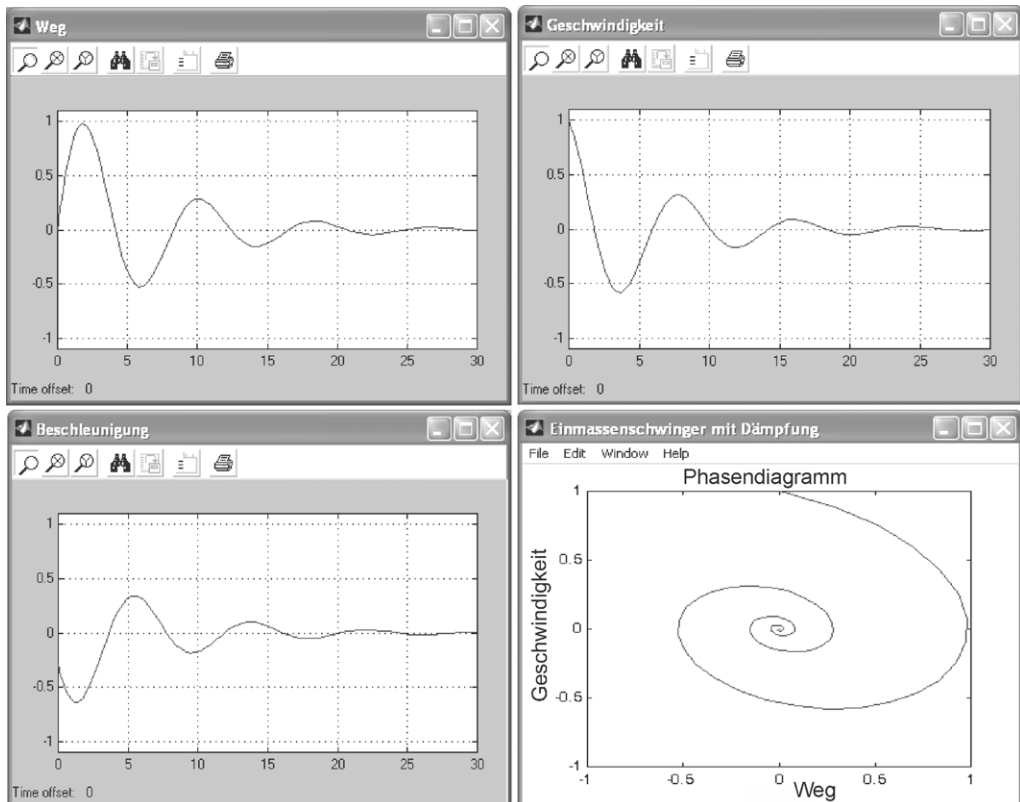


Bild 8-2: Simulationsergebnisse für die Simulation eines Einmassenschwingers

¹ Simulink ist ein Produkt der Fa. Scientific Computers GmbH.

Gleichungen werden simultan von zwei Integratoren, einem Addierer und zwei Koeffizientenbildnern gelöst. Durch Variation der Koeffizienten kann man dann mit Hilfe des Simulationssystems sich schnell die Änderungen des dynamischen Verhaltens eines Systems darstellen lassen. Bild 8-2 zeigt die Ergebnisse für den Weg, die Geschwindigkeit und die Beschleunigung des Simulationsdiagramms in Bild 8-1. Die Simulation wurde dabei für einen Anfangswert der Geschwindigkeit von 1 m/s gestartet, der Zeitmaßstab ist Sekunden. Auch andere Darstellungen der Simulationsergebnisse wie das Phasendiagramm (XY-Plot von Weg und Geschwindigkeit) sind mit solchen Systemen möglich (Bild 8-2 rechts unten).

Um numerische Simulationen durchführen zu können, ist als wesentliche mathematische Aufgabe das Problem der numerischen Integration zu lösen. Verschiedene Verfahren werden im folgenden Unterkapitel beschrieben und können in der Literatur, die sich mit der numerischen Mathematik befasst, vertieft werden.

8.1 Numerische Integration

Numerische Integrationsverfahren für Zeitfunktionen gehen davon aus, dass man eine zu integrierende Funktion in Stützstellen im Abstand eines Zeitintervalls Δt diskretisiert (Bild 8-3) und schrittweise iterativ den jeweiligen neuen Wert des Integrals bestimmt. Von der Vielzahl der Verfahren zur numerischen Integration sollen hier im wesentlichen die **expliziten** Verfahren in Einschritt- und Mehrschrittforn behandelt werden. Diese Verfahren ermitteln ausgehend von einem Funktionswert x_i den Funktionswert x_{i+1} , wobei die Bestimmungsgleichung nach x_{i+1} aufgelöst ist (explizite Darstellung).

Bei den hier untersuchten Weg-Zeitfunktionen ist der Wert x_i der zurückgelegte Weg des betrachteten Objektes zum Zeitpunkt t_i . Um nun von einer Stelle i der Zeitfunktion zu einer Stelle $i+1$ zu kommen, kann man eine **Taylor-Reihenentwicklung** vornehmen:

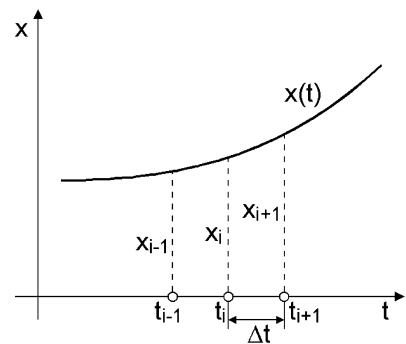


Bild 8-3: Zeitdiskretisierung der Funktion $x(t)$

$$x_{i+1} = x_i + \dot{x}_i \cdot \Delta t + \frac{\ddot{x}_i \cdot \Delta t^2}{2!} + \frac{\dddot{x}_i \cdot \Delta t^3}{3!} + \dots \quad (8.4)$$

Bricht man die Taylor-Reihe nach dem ersten Glied ab, so erhält man daraus

$$x_{i+1} \approx x_i + \dot{x}_i \cdot \Delta t \quad (8.5)$$

Dabei entsteht zwar ein Abbruchfehler, dieser wird aber bei genügend kleinem Δt ebenfalls klein werden. Mit einer derartigen Näherung kann man die Differentialgleichung integrieren. Dabei geht man von den in der Regel bekannten Anfangswerten x_0 und v_0 aus und setzt diese in das Differentialgleichungssystem 8.2 ein, das man damit auch wie folgt schreiben kann:

$$\dot{x}_0 = v_0$$

$$\dot{v}_0 = -\left(\frac{k}{m}\right)x_0 - \left(\frac{d}{m}\right)v_0 \quad .$$

Hieraus kann man $\dot{x}_0$ und $\dot{v}_0$ gewinnen und in die obige Taylor-Reihe einsetzen, um wiederum x_1 und v_1 zu bestimmen:

$$\dot{x}_1 = x_0 + \dot{x}_0 \cdot \Delta t$$

$$\dot{v}_1 = v_0 + \dot{v}_0 \cdot \Delta t \quad .$$

Dies kann man wieder in die DGL einsetzen und gewinnt dadurch für den nächsten Iterationsschritt die Werte $\dot{x}_1$ und $\dot{v}_1$.

$$\dot{x}_1 = v_1$$

$$\dot{v}_1 = -\left(\frac{k}{m}\right)x_1 - \left(\frac{d}{m}\right)v_1 \quad .$$

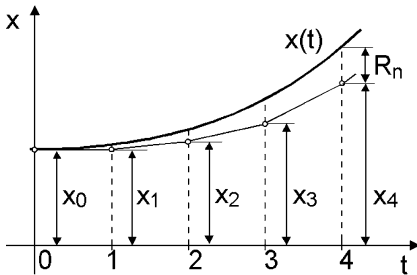


Bild 8-4: Annäherung der Funktion nach dem Euler-Cauchy Verfahren durch einen Polygonzug

Dies Verfahren kann man dann beliebig fortsetzen. Es ist unter dem Namen **Euler-Cauchy-Regel** bekannt und nähert die zu integrierende Funktion durch einen Polygonzug an. Wie in Bild 8-4 gezeigt, folgt die Annäherung der ursprünglichen Funktion mehr oder weniger gut.

Der auftretende **Abbruchfehler** kann durch das nächste Glied der Taylor-Reihe nach dem Abbruch abgeschätzt werden:

$$R_n \approx \dot{x}_i \frac{\Delta t^2}{2!} \quad .$$

Der Fehler nimmt also quadratisch mit kleiner werdender Schrittweite Δt ab. Ein weiterer Fehler bei der Berechnung auf einem Digitalrechner ist der **Rundungsfehler**. Der Einfluss der Rundungsfehler wächst mit der Anzahl der Rechenoperationen, also mit kleiner werdender Schrittweite, während der Abbruchfehler abnimmt. Es ist daher eine optimale Schrittweite in Abhängigkeit der Rechengenauigkeit zu erwarten.

Um die Genauigkeit der numerischen Integration zu verbessern, kann man verschiedene Maßnahmen treffen. Eine Möglichkeit besteht darin, mehr Stützstellen in die Rechnung mit einzubeziehen. Ein Beispiel dafür ist die Integrationsformel nach **Adams-Bashfort**:

$$x_{i+1} = x_i + (b_0 \dot{x}_i + b_1 \dot{x}_{i-1} + b_2 \dot{x}_{i-2}) \Delta t \quad . \quad (8.6)$$

Sie unterscheidet sich von der Integrationsformel nach Euler-Cauchy in Gl. (8.5) durch den Klammerausdruck, der den Zuwachs für den nächsten Iterationsschritt enthält. Da in ihn nicht nur Ausdrücke aus dem i-ten Schritt, sondern auch solche aus einem, bzw. zwei Zeitschritten rückwärts eingehen, bezeichnet man dieses Verfahren als **Mehrschrittverfahren**. Diese beiden Ausdrücke aus vorherigen Zeitschritten erhält man wieder aus einer Taylor-Reihenentwicklung und zeitlicher Ableitung:

$$\begin{aligned}\dot{x}_{i-1} &= \dot{x}_i - \ddot{x}_i \cdot \Delta t + \frac{\ddot{x}_i \cdot \Delta t^2}{2!} - \frac{\ddot{x}_i \cdot \Delta t^3}{3!} + \dots \\ \dot{x}_{i-2} &= \dot{x}_i - \ddot{x}_i \cdot 2\Delta t + \frac{\ddot{x}_i \cdot 4\Delta t^2}{2!} - \frac{\ddot{x}_i \cdot 8\Delta t^3}{3!} + \dots\end{aligned}$$

Setzt man diese Ausdrücke in die Mehrschrittformel der Gl. (8.6) ein, so erhält man durch Koeffizientenvergleich mit der Taylor-Reihe in Gl. (8.4) folgendes Gleichungssystem zur Bestimmung der Koeffizienten b_i :

$$b_0 + b_1 + b_2 = 1 \quad , \quad -b_1 - 2b_2 = \frac{1}{2} \quad , \quad b_1 + 4b_2 = \frac{1}{3} \quad .$$

Daraus kann man die Werte der Koeffizienten errechnen

$$b_0 = \frac{23}{12} \quad , \quad b_1 = -\frac{16}{12} \quad , \quad b_2 = \frac{5}{12} \quad .$$

die in die Gl. (8.6) eingesetzt die endgültige Integrationsformel ergeben:

$$x_{i+1} = x_i + \left(\frac{23}{12} \dot{x}_i - \frac{16}{12} \dot{x}_{i-1} + \frac{5}{12} \dot{x}_{i-2} \right) \Delta t \quad . \quad (8.7)$$

Diese Integrationsformel stimmt nun bereits bis zur dritten Ableitung mit der Taylor-Reihe überein. Der Abbruchfehler ist damit proportional zur vierten Ableitung und der vierten Potenz der Schrittweite:

$$R_n \approx \ddot{x}_i \cdot \Delta t^4 \quad .$$

Bei extremen Steigerungsänderungen wird diese Integrationsformel bei kleinen Schrittweiten Δt noch gute Ergebnisse liefern. Die Genauigkeit ist demnach beträchtlich gesteigert worden, ohne dass der Rechenaufwand wesentlich erhöht wurde. Die Ableitungen $\dot{x}_{i-1}$ und $\dot{x}_{i-2}$ kennt man bereits aus den zurückliegenden Rechenschritten, aber ein großer Nachteil ist nun dazugekommen. Für den Startpunkt $t = 0$ müssen jetzt zusätzlich zum Anfangswert x_0 noch die Werte x_{i-1} und x_{i-2} bekannt sein, was normalerweise nicht der Fall ist. Die Integration nach Gl. (8.7) kann also nur mit einer **Startrechnung** begonnen werden, die mit einem **Einschrittverfahren** durchgeführt werden muss. Nach dem oben beschriebenen Schema lassen sich natürlich viele unterschiedliche Mehrschrittformeln ableiten.

Der Grund dafür, dass beim Euler-Cauchy-Verfahren ein relativ hoher Fehler auftreten kann, ist die Tatsache, dass in der Iterationsformel Gl. (8.5) die Steigung $\dot{x}_i$ der Funktion zum gegenwärtigen Zeitpunkt t_i verwendet wird, um den zukünftigen Funktionswert x_{i+1} zu errechnen. Diese Gleichung kann man modifizieren, indem man den Term für die Steigung der Kurve durch einen Term ersetzt, in dem verschiedene Steigungswerte, bewertet mit einem Faktor α , verwendet werden:

$$x_{i+1} = x_i + (\alpha \cdot \dot{x}_i + (1 - \alpha) \dot{x}_{i+1}) \Delta t \quad . \quad (8.8)$$

Dabei kann der Faktor α folgende Werte annehmen:

$\alpha = 1$ ergibt die Iterationsformal nach Gl. (8.5) des expliziten (forward) Euler-Cauchy-Verfahrens

$\alpha = 0,5$ wird als Trapez-Methode bezeichnet

$\alpha = 0$ wird als rückwärts (backward) oder implizites Euler-Cauchy-Verfahren bezeichnet.

Der letzte Fall wird als **implizites** Verfahren bezeichnet, weil zum gegenwärtigen Zeitpunkt t_i der Steigungswert $\dot{x}_{i+1}$ nicht explizit bekannt ist, sondern nur implizit, d. h. bei jedem Iterationsschritt muss erst eine Rechnung zur Bestimmung der aktuellen Steigung vorgenommen werden. Die Vorgehensweise und die Unterschiede der drei Iterationsformeln sollen nun an folgendem Beispiel einer einfachen DGL 1. Ordnung demonstriert werden:

$$\dot{x} = -5x : \text{analytische Lösung} \quad x(t) = x_0 \cdot e^{-5t} . \quad (8.9)$$

Unter Verwendung von Gl. (8.8) kann die Lösung folgendermaßen geschrieben werden:

$$\begin{aligned} x_{i+1} &= x_i - (5\alpha \cdot x_i + 5(1-\alpha)x_{i+1})\Delta t \\ &= x_i - (5\alpha \cdot x_i + 5 \cdot x_{i+1} - 5\alpha \cdot x_{i+1})\Delta t . \end{aligned} \quad (8.10)$$

Nun liegt eine implizite Gleichung vor, da der Funktionswert x_{i+1} auf beiden Seiten der Gleichung auftaucht. Diese kann man aber folgendermaßen umschreiben:

$$\begin{aligned} x_{i+1}(1 + 5(1-\alpha)\Delta t) &= x_i - 5\alpha \cdot x_i \cdot \Delta t \Rightarrow \\ x_{i+1} &= \frac{x_i - 5\alpha \cdot x_i \cdot \Delta t}{(1 + 5(1-\alpha)\Delta t)} \end{aligned} \quad (8.11)$$

Für die unterschiedlichen Werte von α ergibt sich:

$$\alpha = 1 \Rightarrow x_{i+1} = x_i - 5x_i \cdot \Delta t \quad \text{Euler explizit}$$

$$\alpha = 0,5 \Rightarrow x_{i+1} = \frac{x_i - 2,5x_i \cdot \Delta t}{1 + 2,5 \cdot \Delta t} \quad \text{Euler Trapez}$$

$$\alpha = 0 \Rightarrow x_{i+1} = \frac{x_i}{1 + 5 \cdot \Delta t} \quad \text{Euler implizit}$$

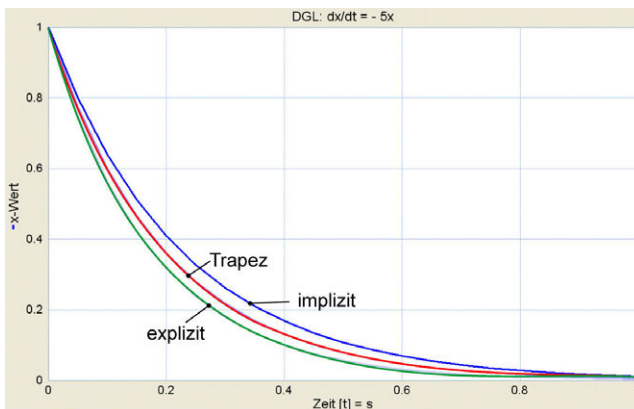


Bild 8-5: Nach verschiedenen Iterationsverfahren errechnete Lösung für die DGL $dx/dt = -5 \cdot x$

Die mit diesen Gleichungen ermittelten Funktionsverläufe sind im Bild 8-5 für den Startwert $x_0 = 1$ und eine Schrittweite $\Delta t = 0,05$ s dargestellt. Die Kurvenverläufe für das explizite und das implizite Euler-Cauchy-Verfahren liegen unterhalb bzw. oberhalb der Kurve für die Trapez-Methode. Diese liefert deshalb die beste Übereinstimmung mit der analytischen Lösung. In der Tabelle 8-1 wird das

bestätigt, in der die ersten 10 Funktionswerte für die Schrittweite $\Delta t = 0,1$ s für alle drei Integrationsmethoden, sowie die analytische Lösung errechnet wurden. Man sieht, dass sich die geringste Abweichung zwischen der analytischen Lösung und den iterativ nach der Trapez-Methode ermittelten Funktionswerten ergibt.

Auch für Einschrittverfahren gibt es Möglichkeiten die Genauigkeit gegenüber der Euler-Cauchy-Regel zu steigern. Bei dem Integrationsverfahren nach **Runge-Kutta** werden Ausdrücke zur Verbesserung des Zuwachsgliedes in die Integrationsformel eingefügt, die eine Linearkombination von Steigungswerten an mehreren Stellen in der Nähe von x_i (z. B. Index $i+1/2$) verwenden (Bild 8-6):

$$x_{i+1} = x_i + \frac{1}{6}(\dot{x}_i + 2\dot{x}_{i+1/2} + 2\dot{x}_{i+1/2} + \dot{x}_{i+1})\Delta t \quad (8.12)$$

Die Ausdrücke zur Verbesserung des Zuwachsgliedes in der Klammer kann man folgendermaßen berechnen:

$$\dot{x}_{i+1/2} = x_i + \dot{x}_i \frac{\Delta t}{2} \quad \ddot{x}_{i+1/2} = x_i + \dot{x}_{i+1/2} \frac{\Delta t}{2} \quad \dot{x}_{i+1} = x_i + \dot{x}_{i+1/2} \cdot \Delta t$$

Der Abbruchfehler beträgt bei dieser Runge-Kutta-Formel 4. Ordnung :

$$R_n \approx \ddot{\ddot{x}}_i \cdot \Delta t^5$$

Die Genauigkeitssteigerung gegenüber den anderen besprochenen Verfahren erfolgt hier dadurch, dass die Ableitung $\dot{x}$ im Intervall Δt mehrfach abgeschätzt wird und dann mit einem mittleren Wert integriert wird. Die Ableitung muss aus der Differentialgleichung viermal bestimmt werden, wodurch der Rechenaufwand beträchtlich steigt. Dafür ist eine Rechnung mit unterschiedlichster Schrittweite zur Anpassung der Genauigkeit an die Lösung möglich. Schwierigkeiten beim Start dieses Verfahrens treten nicht auf.

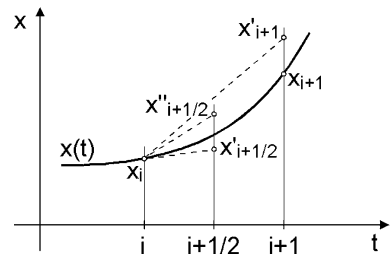


Bild 8-6: Runge-Kutta-Verfahren

Bei der Wahl der Schrittweite des Integrationsverfahrens ist auch die Frage der nu-

Tabelle 8-1: Vergleich verschiedener Berechnungsmethoden der DGL $dx/dt = -5 \cdot x$

Zeit t	analytisch	explizit	Trapez	implizit
0	1	1	1	1
0,1	0,6065	0,5	0,5950	0,6666
0,2	0,3678	0,25	0,3540	0,4444
0,3	0,2231	0,125	0,2110	0,2963
0,4	0,1353	0,0625	0,1253	0,1975
0,5	0,0820	0,0312	0,0746	0,1317
0,6	0,0497	0,0156	0,0380	0,0878
0,7	0,0302	0,0078	0,0264	0,0585
0,8	0,0183	0,0039	0,0157	0,0390
0,9	0,0111	0,0019	0,0093	0,0260
1,0	0,0067	0,0009	0,0056	0,0173

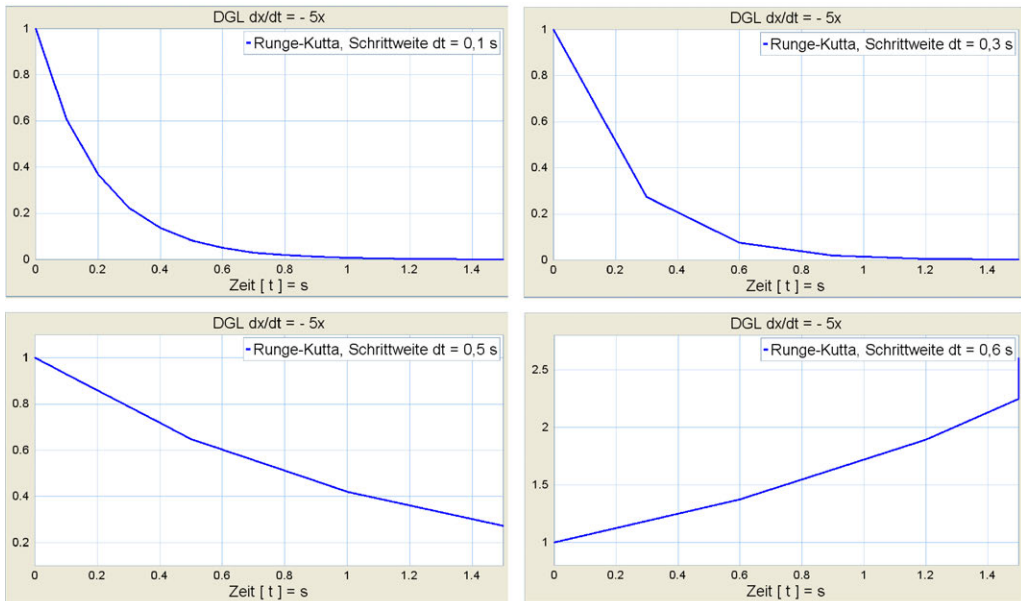


Bild 8-7: Numerische Integration der DGL $\frac{dx}{dt} = -5 \cdot x$ nach dem Runge-Kutta-Verfahren mit unterschiedlichen Schrittweiten

merischen Stabilität zu beachten. Einerseits soll die Schrittweite möglichst groß sein, um Rechenaufwand zu sparen, andererseits auch möglichst klein um eine hohe Genauigkeit zu erhalten. Wählt man die Schrittweite zu groß, so kann die numerische Integration sogar instabil sein. Im Bild 8-7 ist für verschiedene Schrittweiten die DGL (8.9) mit einem Simulationssystem nach dem Runge-Kutta-Verfahren integriert worden.

Der Anfangswert betrug dabei $x_0 = 1$. Bei der Schrittweite $\Delta t = 0,1$ s entspricht der Verlauf von $x(t)$ genau der analytischen Lösung. Bei der Schrittweite 0,3 ist der prinzipielle Verlauf noch vorhanden, es treten aber in $x(t)$ bereits Knickstellen auf. Bei der Schrittweite 0,5 weicht die Lösung bei Werten $t > 0,5$ s schon stark ab und bei 0,6 ist sie instabil und steigt zu unendlichen hohen Werten auf.

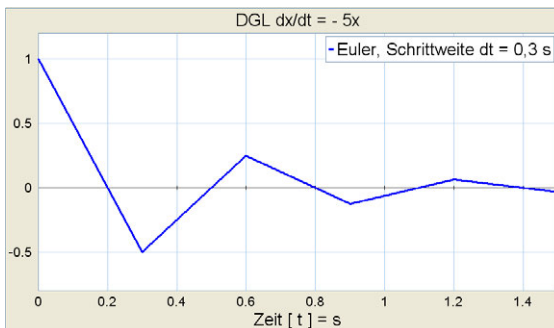


Bild 8-8: Numerische Integration der DGL $\frac{dx}{dt} = -5 \cdot x$ nach dem Euler-Verfahren mit Schrittweite $\Delta t = 0,3$ s

Wie zu erwarten, reagiert das Euler-Verfahren bei großer Schrittweite noch instabiler. So zeigt die Integration der DGL (8.9) in Bild 8-8 bereits bei der Schrittweite von 0,3 ein deutlich oszillierendes Verhalten.

Um sicher zu gehen, dass eine Lösung durch numerische Integration stabil ist, ist es daher sinnvoll, das Verfahren außer für die vorgesehene Schrittweite auch für die halbe und die doppelte Schrittweite zu prüfen. Stimmen alle einigermaßen

überein, so ist die Wahl der Schrittweite in Ordnung.

Auch die Lösung nichtlinearer Differentialgleichungen ist mit der numerischen Integration kein Problem mehr. Betrachtet man als nichtlineares Beispiel das mathematische Pendel aus Kapitel 2.2 mit der DGL

$$\ddot{\varphi} + \frac{g}{l} \sin \varphi = 0$$

für große Ausschläge φ , so kann man sie in folgende Zustandsraumdarstellung überführen:

$$\dot{\varphi} = \omega, \quad \dot{\omega} = -\frac{g}{l} \sin \varphi.$$

Hierfür kann man das in Bild 8-9 dargestellte Simulationsdiagramm aufstellen und wieder mit Hilfe des Simulationssystems numerisch integrieren. Für die angegebenen Parameter ergibt sich der in Bild 8-10 gezeigte Wegverlauf und zwar integriert mit dem Adams-Bashforth-Verfahren bei einer Schrittweite von 0,1. Der ermittelte Verlauf ist die zu erwartende ungedämpfte Sinusschwingung. Aufgrund der großen Schrittweite treten leichte Knickstellen im Funktionsverlauf auf, aber das Integrationsergebnis ist stabil. Nichtlineare Vorgänge reagieren bei Verfahren mit geringer Genauigkeit besonders empfindlich auf eine zu geringe Schrittweite. In Bild 8-10 ist ebenfalls das Simulationsergebnis der Pendelschwingung dargestellt, erzeugt mit dem Euler-Verfahren bei einer Schrittweite von 0,01. Obwohl die Schrittweite bei diesem Einschrittverfahren

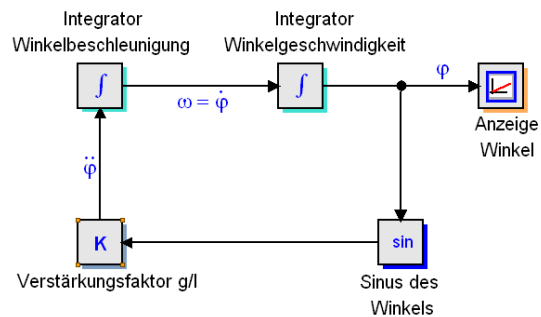


Bild 8-9: Simulationsdiagramm für die nichtlineare DGL des mathematischen Pendels

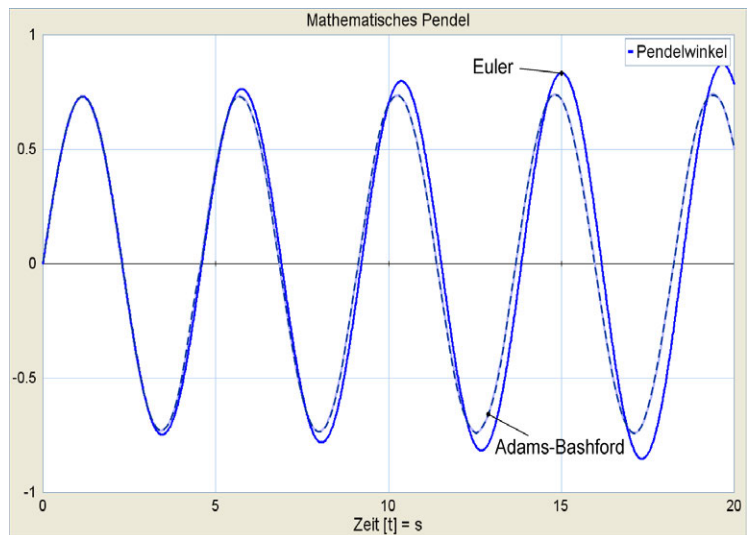


Bild 8-10: Numerische Integration der nichtlinearen DGL des mathematischen Pendels mit dem Adams-Bashforth-Verfahren und mit dem Euler-Verfahren bei unterschiedlicher Schrittweite

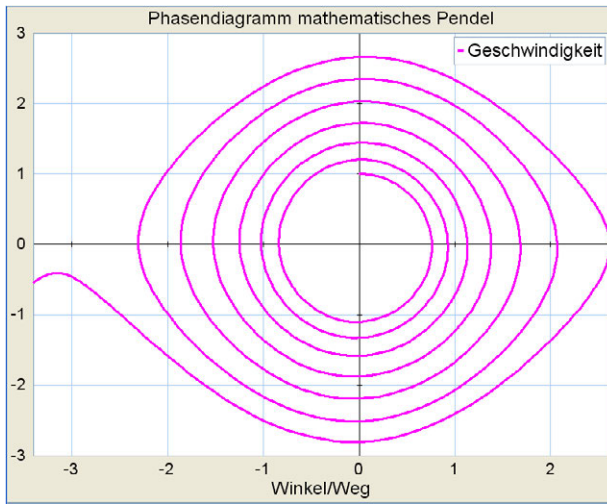


Bild 8-11: Phasendiagramm des mathematischen Pendels bei numerisch instabiler Simulation mit dem Euler-Verfahren

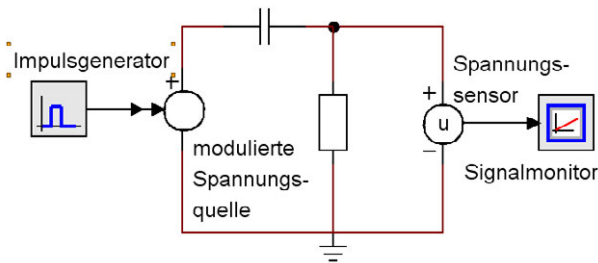


Bild 8-12: Simulationsbild aus 20-sim eines elektrischen Hochpasses

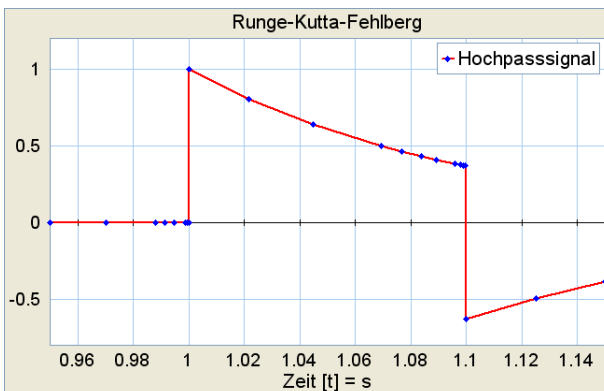


Bild 8-13: Signal eines hochpassgefilterten Rechteckimpulses interpoliert mit dem Runge-Kutta-Fehlberg-Verfahren

ren viel kleiner als beim Adams-Bashforth-Mehrschrittverfahren ist, ist der Funktionsverlauf zwar glatt aber fehlerhaft und numerisch instabil. Dies sieht man auch sehr deutlich am Phasendiagramm in Bild 8-11, in dem die Geschwindigkeit des Pendels über dem Pendelwinkel dargestellt ist. Hier sieht es so aus, als würde die Pendelschwingung aufklingen und letztendlich überschlagen.

Ein weiteres Problem der numerischen Integration stellt die Verwendung einer festen Schrittweite bei der Iteration dar. Bei dynamischen Vorgängen können sehr schnelle Signalwechsel mit steilen Flanken auftreten, während andererseits große Bereiche des Signals nur geringe Steigungen aufweisen. In Bild 8-12 ist eine Messeinrichtung dargestellt, die die Ausgangsspannung an einem elektrischen Hochpass aus einer Kapazität und einem Widerstand misst. Dieses Bild ist ein Modell aus dem Simulationssystem 20-sim (s. Kap.8.3).

Auf den Hochpass wurden Signale eines Impulsgenerators gegeben, die wie in Bild 8-13 dargestellt, durch den Hochpass verändert werden. Dieses Signal zeigt an zwei Stellen sehr steile Flanken, im übrigen Bereich ist die Steigung gering. Verwendet man nun ein Verfahren wie das explizite Euler-Verfahren mit fester Schrittweite, so erhält man die Darstellung in Bild 8-14. Im Kurvenverlauf sind die Interpolationsstützpunkte durch Punkte markiert. Im Bereich der Flanken ist nun

kein steiler Anstieg mehr vorhanden, da der letzte Stützpunkt auf dem unteren Spannungsniveau liegt und der nächste schon auf dem hohen Niveau, das nach der Flanke vorhanden ist. Der Kurvenverlauf der Flanke ist daher die Verbindungslinie dieser beiden Stützpunkte. Im Bereich geringer Steigung des Signals, reicht der Stützpunktabstand aus, um den wahren Verlauf einigermaßen genau wieder zu geben.

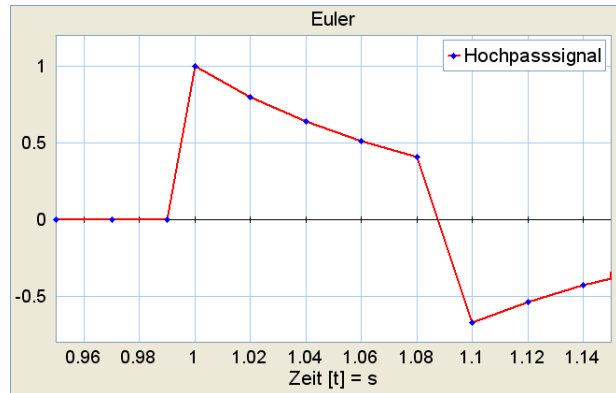


Bild 8-14: Signal eines hochpassgefilterten Rechteckimpulses interpoliert mit dem Euler-Verfahren

Man könnte jetzt den Stützpunktabstand deutlich kleiner wählen, was die Wiedergabe der steilen Flanke verbessern, den Bereich der geringen Steigung aber kaum beeinflussen würde. Gleichzeitig wird jedoch durch diese Maßnahme der Simulationsaufwand und die Rechenzeit deutlich erhöht.

Daher gibt es so genannte Verfahren mit **Zeitschrittweitensteuerung** (adaptive Verfahren), bei denen gleichzeitig mit unterschiedlichen Schrittweiten interpoliert wird. Durch eine Fehlerabschätzung wird die Zeitschrittweite zwischen einem Minimal- und einem Maximalwert automatisch verändert, d. h. bei Signalen hoher Steigung entsprechend verringert. So zeigt das Bild 8-13 den Signalverlauf, der mit dem Runge-Kutta-Fehlberg-Verfahren interpoliert wurde. Dies ist ein modifiziertes Runge-Kutta-Verfahren, bei dem die Schrittweite abhängig von der Steigung des Signals variabel gehalten wird. Dies sieht man daran, dass die im Kurvenverlauf durch Punkte dargestellten Stützstellen im Bereich der Signalfanken näher zusammenrücken. Dadurch erhält man ein Optimum zwischen exakter Signalwiedergabe und Rechenzeit. Neben diesem Verfahren gibt es noch weitere, die mit einer variablen Stützstellenweite arbeiten.

8.2 Modellbildung mit Bondgraphen

Viele Modellbildungsmethoden versuchen ein analytisches Modell in Form eines differentialalgebraischen Ausdrucks aufzustellen. Die grafische Darstellung und die Simulation benutzen dann ein Blockschaltbild (Bild 8-1) in dem die Blöcke jedoch nicht den Bauelementen des modellierten Systems entsprechen, d. h. die **Modellbildung** ist nicht **objektorientiert**. Besonders problematisch wird die Modellbildung, wenn wie in der Mechatronik üblich, verschiedene **Domänen** (Elektrotechnik, Mechanik, Hydraulik, usw.) in einem System zusammenwirken. Die aus unterschiedlichen Domänen stammenden Systemelemente werden mit unterschiedlichen Beschreibungsmethoden, Formelsätzen, Darstellungen und Bezeichnungen behandelt, was nicht nur die Ausbildung multidisziplinärer Ingenieure behindert.

Die Bondgraph-Methode stellt den Versuch dar, eine einheitliche Beschreibungsförm aller physikalischen Systeme zu ermöglichen und zwar unabhängig von einer

bestimmten Fachgebietssicht. Bei der Modellbildung komplexer Systeme zerlegt man das zu untersuchende System zuerst in Teilsysteme, ermittelt die Relationen untereinander und setzt diese Vorgehensweise mit den Teilsystemen fort, bis man eine Stufe erreicht, in der nur noch sehr elementare Teilsysteme auftreten, deren Modellcharakter und Modellgesetze einfach und überschaubar sind (s. Kap. 2). Am mehrfach behandelten Modell des Einmassenschwingers wären dies Objekte wie Punktmasse, Feder und Dämpfer.

Untersucht man ein komplexes System, so kann es solche elementaren Objekte aus den verschiedensten Fachgebieten der Ingenieurwissenschaften enthalten wie Festkörpermechanik, Fluidmechanik, Elektrotechnik, Magnetismus, Halbleiterphysik und Thermodynamik. Jedes dieser Ingenieurfachgebiete hat ganz spezielle Methoden Probleme zu beschreiben und zu lösen. Hat man ein Problem, das eindeutig nur einem dieser Fachgebiete zugeordnet ist, so ist es natürlich, genau die für dieses Fachgebiet üblichen Methoden und Lösungsstrategien, einschließlich der dort verwendeten Softwaretools, zu benutzen. Diese Vorgehensweise kann man auch in fachgebietsübergreifenden Problemstellungen verwenden, wenn die Interaktion zwischen den Fachgebieten nur sehr schwach ausgeprägt ist. Insbesondere in mechatronischen Systemen ist dies aber gerade nicht der Fall, da hier in der Regel ein sehr enges Zusammenspiel von mechanischen, elektrischen und informationsverarbeitenden Systemen vorliegt.

8.2.1 Elemente von Bondgraphen

Ein günstiger Ansatz, solche fachgebietsübergreifenden Problemstellungen zu behandeln, ist die von Henry Paynter 1959 am Massachusetts Institute of Technology entwickelte **Bondgraph-Methode**. Sie geht von der Überlegung aus, dass das grundlegende Merkmal dynamischer Systeme das Fließen von Energie (Leistung) ist und an jeder Stelle eines Systems der Energieerhaltungssatz gilt [8.2], [8.3].

In allen Fachgebieten ist die **Leistung** P das Produkt zweier charakteristischer Größen des Fachgebietes:

$$\begin{aligned} \text{Elektrisch } P &= U \cdot I, \text{ mechanisch (Translation) } P = F \cdot v, \\ \text{mechanisch (Rotation) } P &= M \cdot \omega, \text{ hydraulisch } P = p \cdot V \\ \text{thermodynamisch } P &= T \cdot dS / dt \end{aligned}$$

Die beiden Größen, deren Produkt die Leistung ergeben, werden in der Methodik der Bond-Graphen als **Effort** “ e ” (engl. Anstrengung) und **Flow** “ f ” (engl. Fluss) bezeichnet. Andere im Deutschen üblichere Begriffe wären Effort = Ursache und Flow = Wirkung.

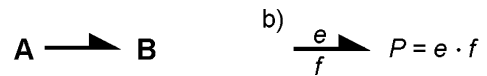
Die Tabelle 8-2 führt die in den unterschiedlichen Fachgebieten verwendeten Effort- und Flowgrößen auf. Ein Bondgraph besteht aus grafischen Elementen, die jeweils ein System repräsentieren, die untereinander mit Leistungsflüssen verbunden sind. Im einfachsten Fall werden die Systeme als **Wortmodell** dargestellt, welches ein oder mehrere Leistungsports beinhaltet. **Ports** sind Schnittstellen des Systems, über welche Leistung mit anderen Systemen oder der Umwelt ausgetauscht wird.

Zwei Wortmodelle (System A und B) in Bild 8-15 a) werden mit einem Halbpfel ver-

Tabelle 8-2: Effort- und Flowgrößen der unterschiedlichen Fachgebiete

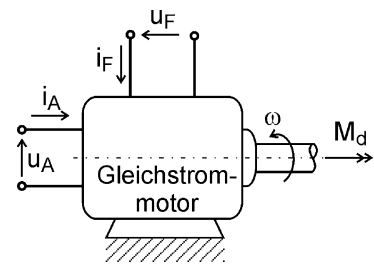
Fachgebiet	Effort	Flow
Elektrotechnik	Potential (Spannung)	(elektrischer) Strom
Mechanik	Kraft	Geschwindigkeit
Fluidtechnik	Druck	Volumenstrom
Chemie	Chemisches Potential	Molarer Fluss
Thermodynamik	Temperatur	Entropiefluss

bunden, der den Leistungsfluss zwischen den Ports der Systeme symbolisiert. Ein solcher Halbpfail wird **Leistungsbond** genannt, der die Richtung des Leistungsflusses von einem Ausgang eines Systems (A) zu einem Eingang eines anderen Systems (B) vorgibt und dadurch die Systeme miteinander verbindet (engl. bonding = verbinden). Der Leistungsfluss durch einen Leistungsbond wird durch zugehörige Effort- und Flowvariablen charakterisiert, die oberhalb und unterhalb des Halbpfails notiert werden (Bild 8-15 b). Der Leistungsbond wird als Halbpfail dargestellt, um ihn vom sonst üblichen Vektorpfail zu unterscheiden. Dieser wiederum wird auch im Bondgraphen dort verwendet, wo nur Informations ohne wesentliche Leistungsübertragung stattfindet

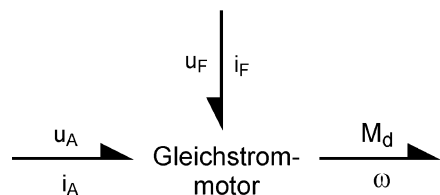
**Bild 8-15:** a) Bondgraph-Elemente

b) Leistungsbond mit zugehöriger Notation von Effort und Flow

Ein Beispiel für ein typisches elektromechanisches System ist der in Bild 8-16 schematisch dargestellte Gleichstrommotor. Er besitzt 3 Leistungsschnittstellen (Leistungsports) und wird daher in der Bondgraph Methodik auch als **Multiport** bezeichnet. Zwei elektrische Leistungsports, nämlich die Ankerwicklung und die Feldwicklung, führen dem System von außen Energie zu und ein mechanischer Leistungsport lässt die Leistung in Form mechanischer Energie wieder abfließen. Dabei sind die Effortgrößen die Spannungen u_A , u_F und das Drehmoment M_d und die Flowgrößen die Ströme i_A , i_F und die Winkelgeschwindigkeit ω .

**Bild 8-16:** Darstellung des Gleichstrommotors als Multiport

In Bild 8-17 ist das gleiche System als Bondgraph Wortmodell dargestellt, das im Zentrum das Wort "Gleichstrommotor" stellvertretend für das System enthält, in das zwei Leistungsbonds die elektrische Leistung hineinführen und ein Leistungsbond die mechanische Leistung nach außen abgibt.

**Bild 8-17:** Bondgraph Wortmodell mit gerichteten Leistungsbonds

Ziel der Modellbildung ist ja immer, ein ma-

thematisches Modell des Systems zu gewinnen, um mit dessen Hilfe das System simulieren zu können. Das in Bild 8-17 dargestellte grafische Modell ist zwar einfach und enthält die wesentlichen dynamischen Größen des Systems, aber bietet noch keine Grundlage für ein mathematisches Modell. Das liegt daran, dass das Wortmodell "Gleichstrommotor" nicht sehr elementar ist und sich dahinter eine Vielzahl sehr einfacher Subsysteme verbergen. Daher muss das grafische Modell weiter in deutlich einfachere Subsysteme untergliedert werden.

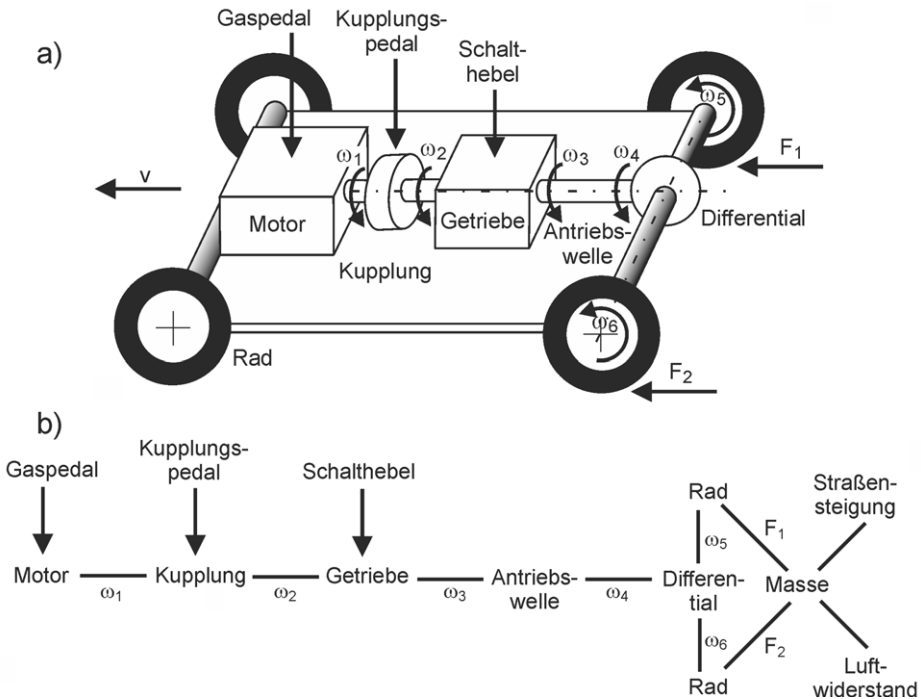


Bild 8-18: Darstellung des Antriebsstrangs eines PKW a) Schematische Darstellung b) Wort-Bondgraph

Trotzdem sind Wortmodelle bei komplexen Systemen nützlich, um zuerst einmal einen Überblick zu gewinnen. Bild 8-18 a) zeigt eine Skizze des Antriebsstranges eines PKW mit dem eigentlichen Antrieb und dem Fahrwerk. Aus einer solchen Skizze kann man direkt ein Wortmodell (Bild 8-18 b) gewinnen [8.3], [8.11] in dem alle Teile (Objekte) der Skizze als Worte vorhanden sind. Dort wo Leistung von einem Wortelement zum anderen fließt, besteht im Modell ein Bond, der bei nicht bekannter Richtung des Leistungsflusses auch ohne Halbpfeil dargestellt werden kann. Im rein objektorientierten Bondgraphen sind noch zusätzliche Leistungsflüsse durch den Luftwiderstand und die Straßensteigung eingetragen. Teile ohne Bedeutung für die Fahrdynamik wie die Vorderräder sind weggelassen. Reine Stellgrößen, die für das dynamische Fahrverhalten keine Bedeutung haben, weil durch sie keine Leistung übertragen wird, sind als Vektorpfeile mit reiner Informationsübertragung eingezeichnet.

In jedem physikalischen System gibt es typische Komponenten, die Energieströme aufteilen, speichern, dissipieren oder umwandeln. Solche elementaren Komponenten

ten werden in Bondgraphen mit speziellen Buchstaben dargestellt. Im einzelnen sind dies:

- Quellen (Source) für Effort und Flow : **SE**, **SF**
- Speicher für Effort und Flow : **L**, **C**
- Leistungsverlust : **R**
- Umwandlung von Leistung : **TF** (Transformer), **GY** (Gyrator)
- Verzweigung von Effort und Flow : **1 (e)**, **0 (f)**

Im folgenden Abschnitt werden deren Eigenschaften und mathematischen Modelle genauer erläutert.

Die Leistung, die in ein Multiport hinein- oder herausfließt ist bei dynamischen Prozessen natürlich von der Zeit abhängig, sodass die Leistung **P** mit den **generalisierten** (verallgemeinerten) **Variablen** der Bondmethodik folgendermaßen geschrieben werden kann:

$$P(t) = e(t) \cdot f(t) \quad (8.13)$$

Die Größen $e(t)$ und $f(t)$ werden in der Sprache der Bondmethodik als generalisierte Variablen bezeichnet, weil sie für alle entsprechenden Größen der unterschiedlichen Domänen stehen. Das spezielle Verhalten dynamischer Systeme wird durch die Art und Anzahl der Speicherelemente, in die Leistung hinein- oder herausfließt, bestimmt.

Das Ansammeln von Leistung in einem Speicher entspricht der Integration von Leistung, die zu einem bestimmten Energieinhalt führt. Deshalb gilt für die **Energie** E in einem Speicher:

$$E(t) = \int_t P(t) dt = \int_t e(t) \cdot f(t) dt \quad (8.14)$$

Der Vorgang der Integration von Leistung in einem Speicher ist auch noch von dem Speichertyp abhängig, da entweder Effort oder Flow im Speicher aufintegriert wird. Zur Beschreibung dieser Unterschiede werden noch zwei weitere generalisierte Variablen aus dem Bereich der Energie oder Arbeit benötigt. Diese werden als **generalisierter Impuls** $p(t)$ und **generalisierte Verschiebung** $q(t)$ bezeichnet. Der generalisierte Impuls (engl. momentum) ist folgendermaßen definiert:

$$p(t) = \int_t e(t) dt = p_0 + \int_{t_0}^t e(t) dt \quad (8.15)$$

d. h. der Impuls ist gleich dem Zeitintegral über den Effort, oder gleich dem bestimmten Integral vom Zeitpunkt t_0 mit Anfangsimpuls p_0 bis zu einem Zeitpunkt t .

Die generalisierte Verschiebung (engl. displacement) ist folgendermaßen definiert:

$$q(t) = \int_t f(t) dt = q_0 + \int_{t_0}^t f(t) dt \quad (8.16)$$

d. h. die Verschiebung ist gleich dem Zeitintegral über den Flow, oder gleich dem bestimmten Integral vom Zeitpunkt t_0 mit der Anfangsverschiebung q_0 bis zu einem

Zeitpunkt t .

Man findet die Darstellung des Zusammenhangs dieser Größen auch in der folgenden Form:

$$\frac{dp(t)}{dt} = e(t), \quad dp = e \cdot dt \quad \frac{dq(t)}{dt} = f(t), \quad dq = f \cdot dt \quad . \quad (8.17)$$

8.2.2 1-Port Bauelemente

Zu dieser Art von Grundelementen für Bondgraphen zählen Quellen, Speicher und Senken (Dissipation) von Energie.

8.2.2.1 1-Port R-Element

Das erste grundlegende Element, das in der Regel Energie der jeweiligen Domäne dissipiert, ist das **R-Element**. In Bondgraphen wird es anstelle eines Wortes durch den Großbuchstaben "R" repräsentiert, wobei diese Bezeichnung sich von dem typischen dissipativen Element der Domäne Elektrotechnik, dem elektrischen Widerstand (**R**esistor) herleitet. In diese 1-Ports fließt Energie der jeweiligen Domäne und wird in der Regel in Form von Wärme oder auch in anderen Energieformen abgegeben,

da das R-Element keine Energie speichern kann. Bild 8-19 zeigt Beispiele für R-Elemente aus zwei verschiedenen Domänen. Im Teilbild a) ist das in Bondgraphen verwendete Symbol dargestellt, ein R, in das ein Bond als Halbpfel in sehr einfach stilisierter Form hineinführt. Neben dem R kann eine erläuternde Bezeichnung (hier :R1 oder :d) stehen.

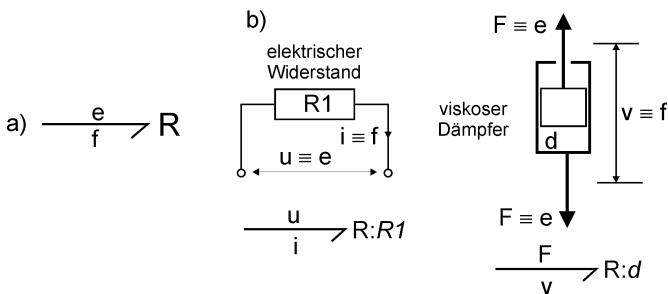


Bild 8-19: a) Bondgraph-Symbol für R-Elemente b) elektrischer Widerstand und viskoser Dämpfer als R-Element

Wie bereits erwähnt, werden Effort und Flow über und unter dem Halbpfel notiert. Für R-Elemente lautet die Definitionsgleichung:

$$e = R \cdot f \quad (8.18)$$

In Teilbild 8-19 b) ist als Beispiel aus der Domäne Elektrotechnik ein elektrischer Widerstand abgebildet, an dem die Spannung U anliegt, was den Strom i zur Folge hat. Für einen reinen Ohmschen Widerstand gilt das Ohmsche Gesetz,

$$U = R \cdot i = e = R \cdot f$$

in dem man leicht die elektrotechnische Variante der Gleichung (8.18) für ein R-Element wiedererkennt.

Ebenfalls in 8-19 b) ist als Beispiel aus der Mechanik der viskose Dämpfer dargestellt, für den gilt, dass die Dämpferkraft der Geschwindigkeit der Bewegung proportional (Proportionalitätsfaktor d : Dämpfungskonstante) ist.

$$F = d \cdot v = e = R \cdot f$$

Aufgrund der Kenntnis, welche Variablen dieser Domäne den Effort und den Flow darstellen, lässt sich auch hier leicht die Identität der charakteristischen Gleichung des viskosen Dämpfers mit der Gleichung (8.18) des R-Elementes feststellen.

Da, wie oben erwähnt, das ideale R-Element keine Energie speichert, stellt sich die Frage, wo die hineingeflossene Energie verbleibt. Diese Frage ist leicht zu beantworten: sie wird in der Regel in Wärmeenergie umgewandelt. Hier zeigt sich wiederum der Vorteil der am Leistungsfluss orientierten Bondgraphen-Methodik. Interessiert bei der Modellbildung

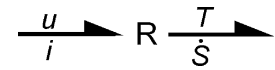


Bild 8-20:

Erweiterung des R-Elementes zum 2-Port

die abfließende Wärmeenergie nicht, so gilt die oben gezeigte Darstellung des R-Elementes als 1-Port. Ist die Wärmeenergie jedoch von Bedeutung, so kann man das R-Element (Bild 8-20) problemlos zu einem 2-Port erweitern, aus dessen 2. Port die Wärmeenergie bei der Temperatur T in Form eines Entropiestroms $\dot{S}$ abfließt.

8.2.2.2 1-Port C-Elemente

Ein weiteres 1-Port Element ist die **Kapazität** oder **Nachgiebigkeit** (engl. **Capacity** oder **Compliance**), die im Bondgraphen durch den Großbuchstaben "C" symbolisiert wird. Die Bezeichnung **C-Element** stammt von der englischen Bezeichnung und steht für ein Element, das Energie verlustlos speichern kann, es wird also keine Leistung dissipiert. Die Darstellung innerhalb eines Bondgraphen findet sich in Bild 8-21 a). Die Definitionsgleichung des C-Element lautet:

$$e = \frac{1}{C} q \quad \text{oder} \quad f = C \frac{de}{dt} \quad (8.19)$$

d.h. es verknüpft den Effort e mit der Verschiebung q .

In Bild 8-21 b) ist als Beispiel für ein C-Element aus der Domäne Elektrotechnik der Kondensator aufgeführt. Von diesem Bauteil stammt auch die Bezeichnung "C" die hier dem Parameter Kapazität (engl.

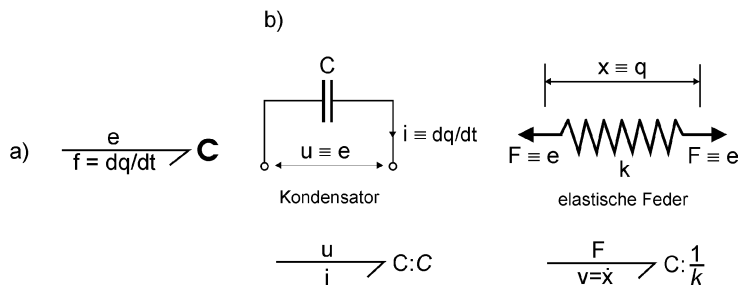


Bild 8-21: a) Bondgraph-Symbol für C-Elemente

b) Kondensator und Feder als C-Element

Capacity) entspricht. Mit der Kenntnis über Effort und Flow in der Domäne Elektrotechnik findet man leicht, wie die bekannte Beziehung zwischen Strom und Span-

nung beim Kondensator mit Gl. (8.19) übereinstimmt:

$$u = \frac{1}{C} \int i \cdot dt \equiv e = \frac{1}{C} \cdot q \quad \text{da} \quad i = f = \frac{dq}{dt}.$$

Der Flow (Strom) wird im Kondensator zur Verschiebung (Ladung) aufintegriert und führt dadurch in reziproker Abhängigkeit von der Nachgiebigkeit (Kapazität) zum Effort (Kondensatorspannung) des C-Elementes.

Die ideale mechanische Schraubenfeder (Bild 8-21 b)), die die Steifigkeit k besitzt und dem Hooke'schen Gesetz gehorcht, ist ebenfalls ein C-Element:

$$F = \frac{1}{N} \int v \cdot dt = k \cdot x \equiv e = \frac{1}{C} \cdot q \quad \text{mit} \quad N = \frac{1}{k} \quad N: \text{Nachgiebigkeit}$$

In der Feder wird Flow (Geschwindigkeit) proportional zur Steifigkeit zur Verformung (Verschiebung) aufintegriert, wodurch in der Feder ein Effort (Federkraft) auftritt.

8.2.2.3 1-Port I-Elemente

Das zweite 1-Port Element, das verlustlos Energie speichern kann, ist das **I-Element**. Die Bezeichnung leitet sich von den englischen Begriffen Induktivität (Inductance) und/oder Trägheit (Inertia) ab. Die Definitionsgleichung des I-Elementes verknüpft den Impuls über die Trägheit mit dem Flow:

$$p = I \cdot f \quad \text{oder} \quad e = I \cdot \frac{df}{dt}. \quad (8.20)$$

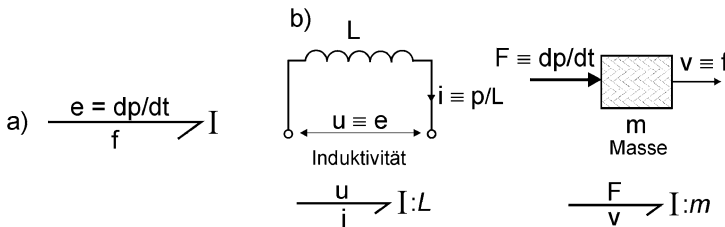


Bild 8-22: a) Bondgraph-Symbol für I-Elemente

b) Spule und Masse als I-Element

Die Darstellung des I-Elementes innerhalb eines Bondgraphen findet sich in Bild 8-22 a). In Bild 8-22 b) ist als Beispiel für ein I-Element aus der Domäne Elektrotechnik die Spule (Induktivität) aufgeführt. Betrachtet

man wieder die aus der Elektrotechnik bekannte Gleichung über Induktivitäten

$$U = L \frac{di}{dt} \quad \text{oder} \quad \int U \cdot dt = L \cdot i \equiv p = I \cdot f \quad \text{bzw.} \quad \lambda = I \cdot f,$$

so ist leicht zu erkennen, dass diese der Gl. (8.20) des I-Elementes entspricht. Der Windungsfluss λ (Zeitintegral über der Spannung) ist dem Flow (Strom) proportional, mit dem Proportionalitätsfaktor Induktivität.

Das mechanische Beispiel in Bild 8-22 b) ist einfach eine Masse, die die mechanische Eigenschaft "Trägheit" besitzt. Für sie gilt das Newton'sche Bewegungsgesetz, dessen Übereinstimmung mit der Gl. (8.20) des I-Elementes wieder leicht her-

zuleiten ist:

$$F = m \cdot a = m \cdot \dot{v} \equiv p = I \cdot f \quad \text{bzw.} \quad \frac{dp}{dt} = e = I \cdot \frac{df}{dt}.$$

Dem Parameter I entspricht hier die Masse m des bewegten Körpers.

8.2.2.4 1-Port-Quellen

Bis jetzt wurden Elemente behandelt, die Energie dem System entnehmen (R-Element) oder Energie speichern (C- und I-Element). Es fehlen nun noch Elemente, die Energie in das System hineinführen. Fließt Leistung in das System, so erfolgt dies dadurch, dass entweder Effort oder Flow dem System zugeführt werden und

sich die jeweils andere Größe aufgrund der Systemeigenschaften ergibt. Entsprechend werden in Bondgraphen **Quellen** (engl. Source) für Effort und Flow verwendet, die mit den Großbuchstaben "**SE**" und "**SF**" symbolisiert werden. Während bei den R-, C- und I-Elementen der Leistungsbond immer in Richtung des Elementes zeigt, ist bei den Quellen die Richtung natürlich umgekehrt, weil die Leistung von der Quelle in das System hineinfließt (Bild 8-23).

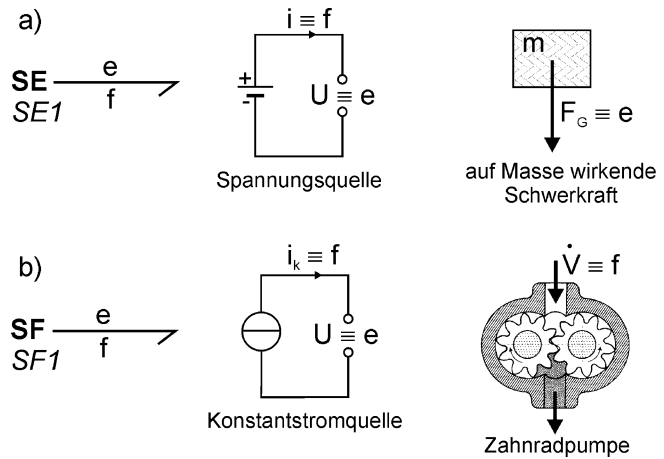


Bild 8-23: Quellen von Effort und Flow verschiedener Domänen a) Effortquellen b) Flowquellen

Bei einer Effortquelle ist der Effort vorgegeben und der Flow ergibt sich aus den Eigenschaften des Systems. So wird bei einer Spannungsquelle die Spannung vorgegeben und der Innenwiderstand des gespeisten Systems bestimmt die Höhe des Stroms. Bei einer Masse im Schwerfeld der Erde ist die Schwerkraft vorgegeben und die Geschwindigkeit zu einem bestimmten Zeitpunkt ergibt sich aus der Masse des Körpers.

Bei einer Flowquelle ist der Flow vorgegeben und der Effort ergibt sich aus den Eigenschaften des Systems. So wird bei einer Konstantstromquelle der Strom vorgegeben und die Spannung ergibt sich aus dem Innenwiderstand des gespeisten Systems, bei einer Zahnradpumpe wird ein vorgegebener Volumenstrom gefördert und der Pumpendruck am Ausgang hängt vom hydraulischen Widerstand des gespeisten Systems ab.

8.2.3 2-Port Bauelemente

2-Port Bauelemente innerhalb eines Systems nehmen aus einer anderen Systemkomponente Leistung am Eingangsport auf und übertragen diese zum Ausgangsport, wo sie verlustlos wieder an das System abgegeben wird. Dies ist natürlich wie auch schon bei den I- und C-Elementen mit nur einem Port ein idealisiertes Verhalten. Sollten im realen System in den Bauelementen Verluste auftreten, so sind diese durch zusätzliche R-Elemente zu modellieren. Es gibt zwei wichtige 2-Port Bauelemente, nämlich den **Transformer** und den **Gyrator**.

8.2.3.1 2-Port Transformer

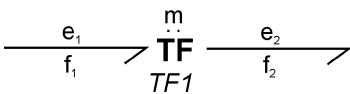


Bild 8-24: Bondgraph Symbol
für den Transformer

Der Transformer ist ein Bauelement (Bild 8-24), in das an einem Port Leistung hineinfließt, im System verlustlos umgewandelt wird und am zweiten Port wieder herausfließt. Die idealisierte Tatsache der verlustlosen Umwandlung wird durch folgende Gleichung beschrieben, bei der die Indizes 1 und 2 sich

auf Bild 8-24 beziehen:

$$e_1(t) \cdot f_1(t) = e_2(t) \cdot f_2(t) \quad . \quad (8.21)$$

Bei einem Transformer stehen der Eingangseffort und der Ausgangseffort in einem bestimmten Verhältnis, das durch einen **Transformerfaktor "m"** (engl. transformer modulus) gekennzeichnet ist. Das gleiche Übersetzungsverhältnis gilt zwischen Eingangs- und Ausgangsflow. Dieser Übertragungsfaktor, der wie in Bild 8-24 gezeigt über dem Transformersymbol "TF" durch einen waagerecht liegenden Doppelpunkt abgetrennt wird, beeinflusst die Ein- und Ausgangeigenschaften wie folgt:

$$e_1 = m \cdot e_2 \quad \text{und} \quad f_2 = m \cdot f_1 \quad . \quad (8.22)$$

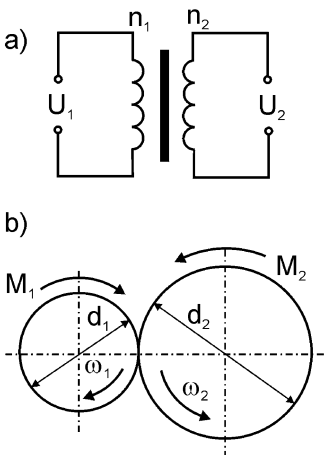


Bild 8-25: Transformer-Bauelemente

a) elektrischer Transformator

b) mechanisches Getriebe

Das Transformer-Bauelement tritt in vielen Systemen auf. In Bild 8-25 ist jeweils ein Beispiel für einen Transformer aus der Domäne Elektrotechnik und Mechanik aufgeführt. Das Beispiel aus der Elektrotechnik (Bild 8-25 a)) ist der ideale elektrische Transformator (engl. Transformer), von dessen englischer Bezeichnung sich auch der Elementname ableitet. Von diesem System weiß man, dass die Eingangsspannung U_1 und die Ausgangsspannung U_2 im Verhältnis der Windungszahlen der Spulen auf der Eingangsseite n_1 und Ausgangsseite n_2 stehen und dass die zugehörigen Ströme sich umgekehrt proportional verhalten.

Da in der Domäne Elektrotechnik der Effort eine Spannung und der Flow ein Strom ist, lassen sich die Definitionsgleichungen (8.22) folgendermaßen schreiben:

$$U_1 = m \cdot U_2 \quad \text{und} \quad i_2 = m \cdot i_1 \quad .$$

Daraus kann man bei Kenntnis des Transformatorgesetzes den Wert des Transformfaktors m ableiten:

$$m = \frac{n_1}{n_2} \quad m: \text{Transformerfaktor}$$

Ein Beispiel aus der Mechanik ist das in Bild 8-25 b) dargestellte Radgetriebe. Bei einem rotatorischen System ist der Effort ein Moment und der Flow eine Winkelgeschwindigkeit. Von einem solchen Getriebe weiß man, dass die Winkelgeschwindigkeiten (Drehzahlen) auf der Eingangsseite ω_1 und der Ausgangsseite ω_2 umgekehrt proportional zum Verhältnis der Durchmesser d_1 und d_2 der Räder sind, während die Drehmomente direkt proportional zu diesem Verhältnis sind. Für das mechanische Getriebe kann man die Definitionsgleichungen (8.22) dann folgendermaßen schreiben:

$$M_1 = m \cdot M_2 \quad \text{und} \quad \omega_2 = m \cdot \omega_1 \quad ,$$

woraus sich der Transformfaktor m wie folgt berechnet:

$$m = \frac{d_1}{d_2}$$

8.2.3.2 2-Port Gyrator

Während bei einem Transformer jeweils zwischen Eingangseffort und Ausgangseffort, bzw. dem Eingangsflow und dem Ausgangsflow eine Beziehung über den Transformfaktor hergestellt wird, bildet der **Gyrator** eine Beziehung zwischen Eingangseffort und Ausgangsflow, bzw. zwischen Ausgangseffort und Eingangsflow. Die zugehörigen Definitionsgleichungen lauten daher:

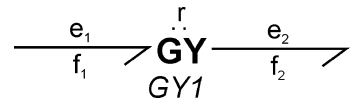


Bild 8-26: Bondgraph Symbol für den Gyrator

$$e_1 = r \cdot f_2 \quad \text{und} \quad e_2 = r \cdot f_1 \quad . \quad (8.23)$$

Der Gyratorfaktor r (engl. gyrator ratio, gyrator modulus)) ist wie beim Transformer die Verknüpfungsgröße zwischen Ein- und Ausgang.

Für den idealen Gyrator gilt ebenso wie beim Transformer, dass er weder Leistung dissipiert noch speichert, so dass auch für ihn die Gl. (8.21) gilt, d. h. Ein- und Ausgangsleistung sind gleich (Bild 8-26). Im Bild 8-27 sind wieder zwei Beispiele für Gyratoren aus den Domänen Elektrotechnik und Mechanik dargestellt. Teilbild a) zeigt das Funktionsschema eines Gleichstrommotors. Bei diesem Motor wird elektrische in mechanische Energie umgewandelt. Auf der elektrischen Eingangsseite wird die Leistung durch die innere Motorspannung U_i und den Ankerstrom I_{mot} bestimmt. Mit Bezug auf Gl. (8.23) gilt dann:

$$e_1 = U_i \quad , \quad f_1 = I_{\text{Mot}} \quad .$$

Auf der mechanischen Ausgangsseite wird die Leistung durch das Motormoment und die Winkelgeschwindigkeit ω bestimmt:

$$e_2 = M_{\text{Mot}} \quad , \quad f_2 = \omega \quad .$$

Treffen auf dieses System die konstituierenden Gleichungen des Gyratoren zu, so

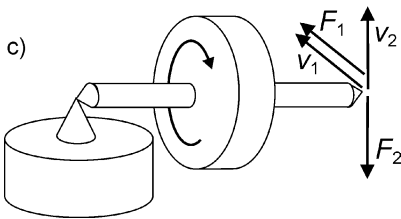
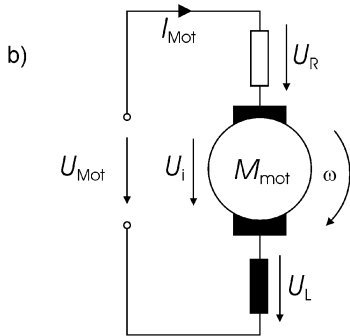
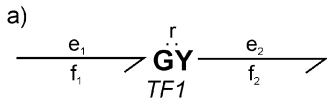


Bild 8-27: Gyrator-Bauelemente

- a) Gleichstrommotor
b) Kreisel

lauten diese:

$$U_i = r \cdot \omega \quad \text{und} \quad M_{\text{Mot}} = r \cdot I_{\text{Mot}} \quad .$$

Der Gyratorfaktor r muss dann aufgrund der Gesetzmäßigkeiten über den Gleichstrommotor folgenden Wert besitzen:

$$r = c_{\text{Mot}} \cdot \Phi$$

d.h. er ist gleich dem Produkt zweier Motorkenngrößen, der Motorkonstanten c_{Mot} und dem magnetischen Fluss in der Feldwicklung Φ .

Das mechanische Beispiel in Teilbild 8-27 b) stellt einen Kreisel (engl. Gyro) dar, von dessen englischer Bezeichnung sich auch der Elementname ableitet. Die Efforts sind hier senkrecht aufeinander stehende Kräfte, die am Kreiselende angreifen, die Flows sind senkrecht aufeinander stehende Geschwindigkeiten, mit denen sich das Kreiselende im Raum bewegen kann. Wenn sich der Kreisel mit hoher Winkelgeschwindigkeit ω dreht, so wird ein in Richtung der Kraft F_1 aufgebracht leichter Stoß eine Bewegung des Kreiselendes in Richtung der Geschwindigkeit v_2 zur Folge haben. Da bei einem realen Kreisel die Schwerkraft in Richtung F_2 wirkt, führt der Kreisel eine Präzessionsbewegung in Richtung von v_1 aus. Zeigen die Kräfte in x- und y-Richtung eines kartesischen Koordinatensystems, so ist die Drehachse des Kreisels die z-Richtung. Wenn das Trägheitsmoment des Kreisels um die z-Achse mit J_{zz} bezeichnet wird, so gilt entsprechend den Definitionsgleichungen (8.19):

$$M_x = J_{zz} \cdot \omega_z \cdot \omega_y \quad \text{und} \quad M_y = J_{zz} \cdot \omega_z \cdot \omega_x \quad .$$

Daraus kann man den Gyratorfaktor r ablesen:

$$r = J_{zz} \cdot \omega_z$$

8.2.4 Multiports

Die bis jetzt behandelten grundlegenden Bauelemente kann man beliebig in Reihe schalten (Serienschaltung), d. h. die Leistung fließt entweder durch die Bauelemente hindurch oder wird in ihnen verbraucht oder gespeichert. Schaut man sich in den verschiedenen Domänen komplexere Zusammenschaltungen von grundlegenden Bauelementen an, so sieht man schnell, dass sich häufig die Leistung auf mehrere Zweige aufteilt, die dann parallel geschaltet sind. Daher braucht man auch für Bondgraphen Elemente, die die Leistung auf verschiedene andere Modellelemente aufteilen. Solche **Multiports** haben mindestens 3, können aber beliebig viele Ports

besitzen. Weil sie andere konzentrierte Bauelemente miteinander verbinden, werden sie auch als Verknüpfungen oder **Junctions** bezeichnet und haben immer mindestens 3 Ports.

Genauso wie beispielsweise in elektrischen oder hydraulischen Schaltplänen Komponenten in Reihe oder parallel geschaltet werden können, gibt es im Bondgraphen zwei unterschiedliche Typen von Junctions. Allen ist gemeinsam, dass sie keine Leistung verbrauchen oder speichern, d. h. die an 1 oder 2 Ports hineinfließende Leistung fließt unverändert an 1 oder 2 Ports wieder hinaus. Daraus folgt für 3 Ports:

$$e_1 \cdot f_1 + e_2 \cdot f_2 + e_3 \cdot f_3 = 0 \quad . \quad (8.24)$$

Der erste Typ von Junction wird als **Flow-Junction**, Parallel-Junction oder 0-Junction, manchmal auch "Common Effort-Junction" bezeichnet. Die Bedeutung dieser unterschiedlichen Namen erhellt aus den im Folgenden erläuterten Eigenschaften und ihren Definitionsgleichungen. Ein solches 3-Port ist in Bild 8-28 dargestellt, wobei eine Null (0) als Symbol im Graphen dient, was zu der alternativen Bezeichnung 0-Junction führt. Von den drei Leistungsbonds muss immer mindestens einer in die Junction hinein- und einer herausführen. Im dargestellten Fall (8-28 a)) führt ein Bond hinein und zwei führen heraus, was dann folgende Definitionsgleichungen zur Folge hat:

$$e_1(t) = e_2(t) = e_3(t) \quad . \quad (8.25)$$

Damit Gleichung (8.20) erfüllt werden kann, muss dann für die Flows folgendes gelten:

$$f_1(t) - f_2(t) - f_3(t) = 0 \quad \text{oder} \quad f_1(t) = f_2(t) + f_3(t) \quad (8.26)$$

Es gilt also für die 0-Junction, dass alle Efforts der Junction gleich (Common Effort-Junction) und die algebraische Summe aller Flows gleich Null sein muss. Entsprechend ändern sich dann für die Richtungen der Leistungsbonds in Bild 8-28 b) die Vorzeichen in Gl. 8.26. Betrachtet man die Domäne Elektrotechnik, so erkennt man, dass die Aussage aus Gl. 8.26 bezogen auf einen Verbindungsknoten in einem elektrischen Netzwerk als 1. Kirchhoffsches Gesetz bekannt ist, das besagt, dass die Summe aller Ströme im Knoten gleich Null sein muss.

Vertauscht man nun die Rolle von Effort und Flow, so erhält man eine 3-Port-Verknüpfung oder Junction, die als **Effort-Junction**, Reihen (Serien) -Junction oder 1-Junction bezeichnet wird, oder wegen seiner Eigenschaften auch als "Common Flow-Junction".

Ein solches 3-Port ist in Bild 8-29 dargestellt, wobei eine Eins (1) als Symbol im Graphen dient, was zu der alternativen Bezeichnung 1-Junction führt. Die Definitionsgleichungen für die Effort-Junction in Teilbild 8-29 a) lauten:

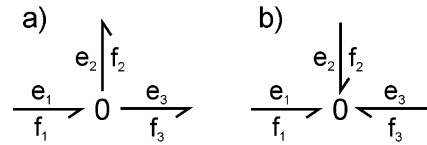


Bild 8-28: Flow-Junction mit unterschiedlicher Orientierung der Leistungsbonds

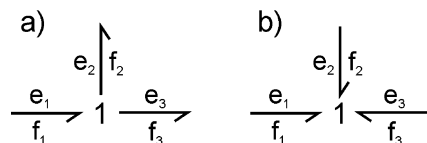


Bild 8-29: Effort-Junction mit unterschiedlicher Orientierung der Leistungsbonds

$$f_1(t) = f_2(t) = f_3(t) \quad . \quad (8.27)$$

$$e_1(t) - e_2(t) - e_3(t) = 0 \quad \text{oder} \quad e_1(t) = e_2(t) + e_3(t) \quad (8.28)$$

Die Gl. (8.28) gilt für die Richtung der Leistungsbonds in Teilbild 8-29 a). Die beiden Definitionsgleichungen besagen, dass in der Effort-Junction alle Flows gleich sind und die algebraische Summe aller Efforts gleich Null ist. Die Aussage der Gl. 8.28 ist in der Elektrotechnik als 2. Kirchhoff'sches Gesetz bekannt, das besagt, dass bei einem Maschenumlauf in einem elektrischen Netz die Summe aller Spannungen gleich Null sein muss. Auch in der Mechanik ist eine Entsprechung in Form des d'Alembert'schen Prinzips gegeben, das besagt, dass die Summe aller an einer Masse angreifenden Kräfte gleich Null sein muss.

8.2.5 Erstellung von Modellen komplexerer Systeme

Um nun Bondgraph Modelle komplexerer Systeme mit Hilfe der oben beschriebenen einfachen Bauelemente vorzunehmen, gibt es für unterschiedliche Domänen standardisierte Vorgehensweisen. Die Vorgehensweise ist bei elektrischen Systemen wie folgt:

- Man identifiziert im Netzwerk alle Positionen mit unterschiedlichem Potential und sieht dafür im Bondgraphen eine 0-Junction vor.
- Dann fügt man für konzentrierte Bauelemente als Modell das entsprechende 1-Port Bondgraph Symbol zwischen den 0-Junctions mit Hilfe einer 1-Junction ein.
- Ist das Spannungsnullpotential bekannt, so kann man die zugehörige 0-Junction und alle zugehörigen Leistungsbonds entfernen.
- Danach kann man den Graph minimieren, indem man alle Junctions, welche nur noch zwei Ports besitzen, durch einen einzigen Leistungsbond ersetzt.

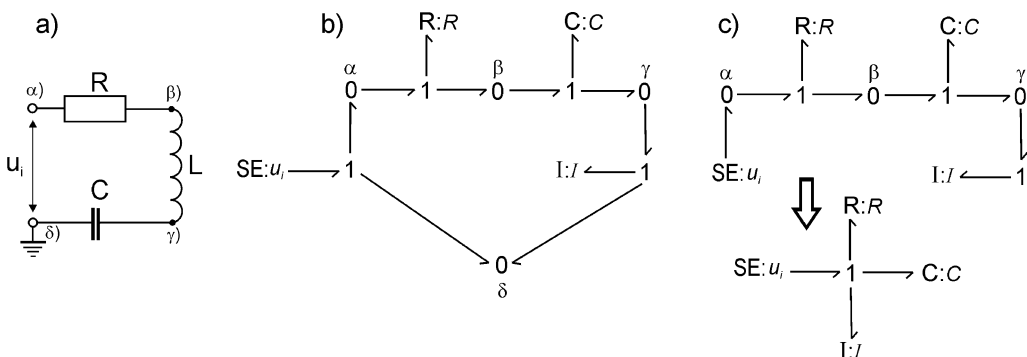


Bild 8-30: Erstellung des Bondgraphen eines Serienschwingkreises

a) Schaltplan b) Bondgraph c) minimierter Bondgraph

Die Vorgehensweise soll an einem einfachen RLC-Netzwerk (Bild 8-30), einem Serienschwingkreis, erläutert werden. In Teilbild a) ist der Schwingkreis dargestellt und die 4 Punkte im Netzwerk, die unterschiedliche Potentiale besitzen, sind mit

griechischen Buchstaben α , β , γ , δ bezeichnet. Für diese 4 Punkte wird im Bondgraphen in Teilbild b) je eine 0-Junction eingezeichnet. Die Bauteile Widerstand, Spule und Kondensator werden durch je ein R-, L- und C-Element modelliert und zwischen den Potentialpunkten (0-Junction) über eine 1-Junction eingefügt. Die Spannungsquelle wird durch ein SE-Element modelliert und ebenfalls über eine 1-Junction zwischen den zugehörigen Potentialen eingefügt. Die Richtung der Leistungsbonds ist für die konzentrierten Bauelemente vorgegeben, die übrigen Leistungsbonds können prinzipiell beliebig orientiert werden. Die gewählte Orientierung wird bei Erstellung der Systemgleichungen durch entsprechende Vorzeichenwahl berücksichtigt.

Dieser erste Bondgraph kann weiter vereinfacht werden, indem, wie in Teilbild 8-30 c) oben gezeigt, die 0-Junction für das Nullpotential (δ) und die zugehörigen Leistungsbonds weggelassen werden. Die 0-Junctions α , β , γ haben jetzt nur jeweils 2 Ports mit entsprechenden hinein- und herausführenden Leistungsbonds und können daher durch einen einfachen Leistungsbond ersetzt werden. Zwei benachbarte, nur mit einem Leistungsbond verbundene 1-Junctions, können zu einer 1-Junction verschmolzen werden. Das führt auf das minimierte Modell in Teilbild 8-30 c) unten.

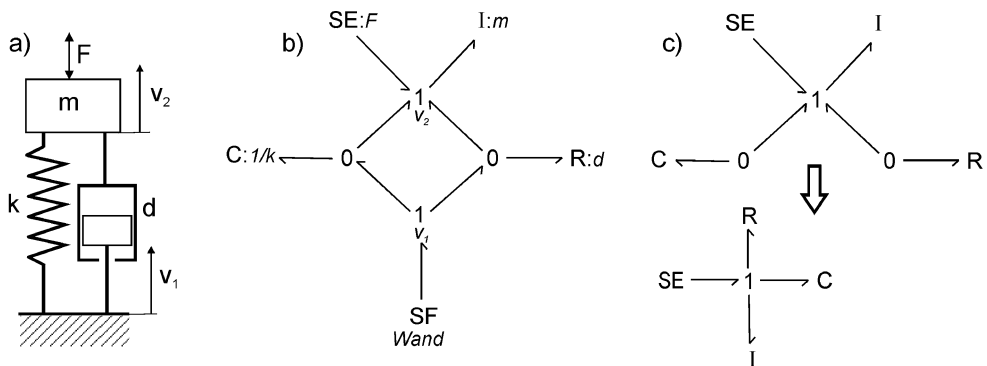


Bild 8-31: Erstellung des Bondgraphen eines Einmassenschwingers

a) Skizze b) Bondgraph c) minimierter Bondgraph

Als zweites Beispiel aus der Domäne Mechanik soll der bereits mehrfach angesprochene Einmassenschwinger (Bild 8-31) modelliert werden. Die Vorgehensweise bei mechanischen Systemen ist wie folgt:

- Man identifiziert alle Punkte des Systems mit unterschiedlicher translatorischer Geschwindigkeit oder Winkelgeschwindigkeit (auch Geschwindigkeit Null) und sieht dafür im Bondgraphen eine 1-Junction vor.
- Dann fügt man für R- und C-Bauelemente als Modell das entsprechende 1-Port Bondgraph Symbol zwischen den 1-Junctions mit Hilfe einer 0-Junction ein. I-Elemente (Massen) werden direkt mit einem Leistungsbond an die seiner Geschwindigkeit entsprechende 1-Junction angehängt.
- Alle 1-Junctions, die der Geschwindigkeit Null entsprechen und die zugehörigen Leistungsbonds, können dann entfernt werden.

- Danach kann man den Graph minimieren, indem man alle Junctions, die jetzt nur noch zwei Ports besitzen, durch einen einzigen Leistungsbond ersetzt.

In Teilbild 8-31 a) ist der Einmassenschwinger und die beiden unterschiedlichen Geschwindigkeiten v_1 der Wand und v_2 der Masse eingezeichnet. Für diese Geschwindigkeiten ist im Bondgraphen Teilbild b) je eine 1-Junction eingezeichnet. Die Feder (C-Element) und der Dämpfer (R-Element) werden mit jeweils einer 0-Junction zwischen den beiden Geschwindigkeiten (1-Junction) eingefügt, die Masse (I-Element) wird direkt an die 1-Junction (Geschwindigkeit) angehängt. Die auf die Masse einwirkende Kraft F ist ein SE-Element, das direkt an die 1-Junction (v_2), an der die Masse angebunden ist, angefügt werden kann. Die Geschwindigkeit v_1 muss durch eine Flowquelle (SF) erzeugt werden.

Dieser erste Bondgraph kann weiter vereinfacht werden, indem, wie in Teilbild 8-31 c) oben gezeigt, die 1-Junction für die Geschwindigkeit Null (v_1) und die zugehörigen Leistungsbonds, sowie die Flowquelle weggelassen werden (Bild 8-31 c) oben). Die beiden 0-Junctions, die nun nur noch 2 Leistungsbonds besitzen, können durch einen einzelnen Leistungsbond ersetzt werden. Dies führt auf den minimierten Bondgraphen für den Einmassenschwinger in Teilbild 8-31 c) unten.

Betrachtet man nun die beiden minimierten Bondgraphen der Bilder 8-30 und 8-31, so sieht man, dass diese identisch sind. Diese Erkenntnis der gleichartigen Systemstruktur wurde bereits in Kapitel 1.4 gewonnen, wo gezeigt wurde, dass die beiden Systeme gleichartige DGL's 2. Ordnung besitzen. Da die Bondgraphmethodik unabhängig von der betrachteten Domäne ist, ergeben sich deshalb natürlich identische Bondgraphen.

8.3 Simulationssysteme

Für die Simulation der verschiedenen Komponenten eines mechatronischen Systems gibt es verschiedene, den Aufgaben angepasste Simulationssysteme. Mechanische Systeme werden am besten mit Mehrkörperprogrammen wie **ADAMS**, **NEWEUL** oder **CAMEL-View** modelliert, elektrische Schaltungen mit **SPICE**, regelungstechnische Komponenten mit Blockschaltbildsystemen wie **MATLAB/SIMULINK** und allgemeine physikalische Effekte mit allgemeinen Simulationssprachen wie **ACSL**. Darüber hinaus gibt es sogenannte objektorientierte Simulationssprachen wie **DYMOLA** und **20-sim**, die eine vereinheitlichte Eingabe der Simulationsgleichungen aller Fachgebiete ermöglichen [8.1], [8.2]. Weitere Angaben und Literaturhinweise finden sich in [8.1], [8.2], [8.8] und [8.9].

8.3.1 Simulationssprachen

Seit längerem gibt es schon allgemeine Werkzeuge zum fachübergreifenden Modellieren in Form von allgemeinen Simulationssprachen, wie beispielsweise das weit in der Industrie verbreitete System ACSL (Advanced Continuous Simulation Language). Allgemeine Simulationssprachen setzen voraus, dass die Gleichungen, die die Komponenten beschreiben, gegeben sind. Der Vorteil dieser Sprachen besteht darin, dass die Gleichungen sehr direkt eingegeben werden können. Die Modellein-

gabe ist stark an den Simulationsablauf geknüpft. Mit speziellen Anweisungen können auch diskrete Ereignisse (Schaltvorgänge, u. ä.) behandelt werden.

Nach der Modelleingabe in der Simulationssprache wird die Modelldefinition meist in ein Unterprogramm einer allgemeinen Programmiersprache umgesetzt und danach kompiliert. Das übersetzte Unterprogramm wird in eine Simulationsumgebung eingebunden, mit der Simulationsexperimente durchgeführt und die Ergebnisse eines Simulationslaufes grafisch veranschaulicht werden können.

Die Verwendung einer allgemeinen Simulationssprache erleichtert die Erstellung eines Modells auf einem Rechner, sowie das Durchführen von Simulationsexperimenten. Es ist jedoch nachteilig, dass die Gleichungen der zu simulierenden Komponenten so eingegeben werden müssen, dass ein Modell durch einfaches Sortieren der Anweisungen in eine Zustandsraumdarstellung überführt werden kann. Bei größeren mechatronischen Systemen mit Komponenten aus den Fachdisziplinen Mechanik und Elektrotechnik, besteht aber eine Hauptschwierigkeit gerade darin, diese Gleichungen aus der Modelldefinition abzuleiten.

8.3.2 Simulation elektrischer Schaltungen

Aufgrund ihrer technischen Bedeutung gab es schon sehr früh spezielle Systeme zur Simulation elektrischer Schaltungen. Das auch heute noch wichtigste System für die analoge Schaltungssimulation ist SPICE, das 1975 von Nagel entwickelt wurde und kostenlos als "public domain" Programm weitergegeben wird. Die Bedeutung von Spice liegt vor allem darin, dass die Hersteller von integrierten Schaltungen zu ihren Hardwarekomponenten entsprechende Spice-Modelle liefern, wodurch eine solche Schaltung sofort in Spice simuliert werden kann.

Spice-Modelle werden durch Angabe vordefinierter Schaltungselemente, deren Parametern, sowie deren Verschaltung definiert. Als Beispiel ist in Bild 8-32 die Definition eines einfachen elektrischen Schaltkreises im SPICE-Format aus [8.1] angegeben.

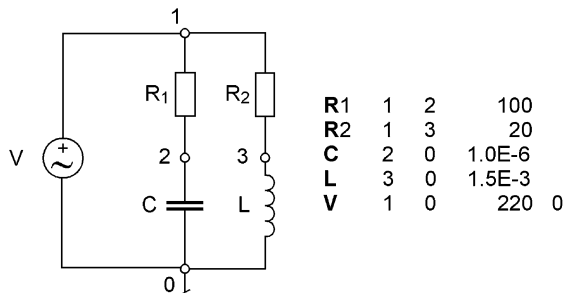


Bild 8-32: Einfache elektrische Schaltung und Modellbeschreibung im SPICE-Format

In der ersten Spalte steht der Elementname. Dieser muss mit einem Kennbuchstaben für den Elementtyp beginnen und ist fettgedruckt.

In der zweiten und dritten Spalte ist durch Knotennummern angegeben, zwischen welchen Knoten das Element angebracht ist. In der letzten Spalte sind die Parameter der Bauelemente wie etwa der Widerstandswert eines Widerstandes angegeben.

Zu Spice werden viele Pre- und Postprozessoren angeboten. Das sind Programme, mit denen beispielsweise ein elektrischer Schaltkreis grafisch definiert werden kann, indem Symbolbilder (Icons) für elektrische Bauteile auf dem Bildschirm platziert und entsprechend verbunden werden. Während Bildschirmdeditoren wie SIMU-

LINK jedoch von der Voraussetzung ausgehen, dass die in den Blöcken vorhandenen Systeme rückwirkungsfrei sind, sind hier die zu verbindenden Komponenten elektrische Bauelemente mit Energieflüssen, die auch Rückwirkungen erzeugen können. Ausgehend von der grafischen Eingabe wird eine Datei erzeugt, die das Modell im SPICE Format enthält. Nach einer Simulation mit Spice können die Simulationsergebnisse mittels Postprozessoren komfortabel angezeigt werden.

Simulationsprogramme für elektrische Schaltungen sind hervorragend für die Modellierung und Simulation rein elektrotechnischer Elemente geeignet. In beschränktem Umfang ist es auch möglich nicht-elektrische Komponenten zu modellieren, wenn es gelingt, das zugehörige Modell in einen äquivalenten elektrischen Schaltkreis zu überführen (s. Kap. 1), der dann wieder auf dem System simuliert werden kann. Regelungstechnische Ein-/Ausgabeblocke können nicht modelliert werden. Außerdem können solche Systeme in der Regel auch keine ereignisabhängigen Vorgänge, wie die Verschaltung eines digitalen Reglers mit einer kontinuierlichen Strecke, simulieren. Digitale Bauelemente, die Schaltvorgänge durchführen wie Dioden oder Thyristoren, werden durch sehr genaue stetige Modelle beschrieben, bei denen das ideale Schaltverhalten durch das reale Schaltverhalten ersetzt wird (s. Kap. 7.2.2). So wird zum Beispiel eine Diode als Widerstand mit nichtlinearer Kennlinie beschrieben, die in Durchlassrichtung einen kleinen und in Sperrichtung einen großen Widerstandswert besitzt. Dies kann schon bei relativ einfachen Schaltungen zu komplizierten Differentialgleichungen und damit langen Rechenzeiten für die Simulation führen.

8.3.3 Simulation mechanischer Systeme

Mechatronische Systeme werden häufig durch starre Mehrkörpersysteme modelliert, d. h. Verformungen der Körper werden vernachlässigt. Es gibt eine Vielzahl von Mehrkörper-Simulationsprogrammen, die sich in der Benutzerfreundlichkeit, der Funktionalität und in der Art der Erstellung der Modellgleichungen unterscheiden. Bekannte kommerzielle MKS-Programmpakete sind ADAMS, DADS, NEWEUL, SD-FAST, SIMPACK [8.1], CAMEL-View und 20-sim.

Im Gegensatz zu elektrischen Schaltungsprogrammen, hat sich bei MKS-Programmen keine Standardeingabe durchgesetzt. Die zugrundeliegende Modellierung ist jedoch bei allen Programmen entsprechend der Definition von Mehrkörpersystemen. Ein MKS wird in starre Körper, Gelenke und Kraftelemente aufgeteilt. Die starren Körper repräsentieren die Trägheitseigenschaften des Systems. Auf diesen Körpern können spezielle Punkte angegeben werden, an denen Gelenke oder Kraftelemente angeordnet werden können. Gelenke sind dabei idealisierte Elemente, die die Bewegungsmöglichkeiten von Körpern einschränken (s. Kap. 3). Kraftelemente sind ideale Elemente, die aufgrund von Relativbewegungen eine Kraft oder ein Moment erzeugen, die zwischen zwei Punkten des Körpers angreifen. Ein Beispiel für ein solches Kraftelement ist eine Feder.

Zur Modellierung wird ein mechanisches System in Elemente dieser drei Grundtypen aufgeteilt. Die Elemente werden dann entsprechend dem Aufbau der realen Struktur zusammen geschaltet. Mit Hilfe mechanischer Prinzipien wie den Lagrange'schen Gleichungen oder dem Prinzip der virtuellen Arbeit, wird dann das

MKS auf ein Gleichungssystem bestimmter Struktur abgebildet. Dabei können **differential algebraische Gleichungssysteme** (DAE: Differential-Algebraic-Equations) entstehen, die nicht mit den Standardintegrationsmethoden gelöst werden können. Je nach Detaillierungsgrad der Modellbildung kann die Simulation auf solchen Systemen zu sehr langen Rechenzeiten führen. Schwierigkeiten treten auch meist auf, wenn unstetige Elemente, wie digitale Regler, Reibung oder Lose, behandelt werden müssen, oder wenn Stöße zu modellieren sind.

Wie bei elektrischen Schaltungsprogrammen, ist es auch mit MKS-Programmen nur bedingt möglich, Komponenten aus anderen Fachgebieten zu modellieren und zu simulieren. Systeme wie CAMEL-View und 20-sim sehen daher die Möglichkeit vor, für regelungstechnische Probleme Modelle aus dem Blockschaltbild-Editor System Matlab/Simulink mit einzubinden.

8.3.4 Modellbeschreibung mit Blockschaltbild-Editoren

Ähnlich wie allgemeine Simulationssprachen unterstützen Blockschaltbild-Editoren und Simulatoren eine fachübergreifende Modellbildung und Simulation. Bei solchen Systemen ist mehr Wert auf Benutzerfreundlichkeit und leichte Bedienbarkeit gelegt als bei allgemeinen Simulationssprachen. Eines der bekanntesten Systeme ist SIMULINK, mit dem einige Simulationen in diesem Buch durchgeführt wurden [8.4].

Blockschaltbild-Editoren gehen von der in der Regelungstechnik üblichen Beschreibung eines Systems als rückwirkungsfreie Ein-/Ausgangsblöcke aus. Üblicherweise wird von den Systemen eine große Anzahl vordefinierter Funktionsblöcke zur Verfügung gestellt, z. B. Blöcke, die durch eine Übertragungsfunktion beschrieben werden, Signalgeneratorblöcke und Blöcke für unstetige Funktionszusammenhänge und nichtlineare Funktionen. Bei manchen Systemen ist es möglich, dass Anwender neue Blocktypen zur Verfügung stellen können, indem die Gleichungen des Blocks in einer Programmiersprache, wie C, oder einer speziellen Sprache, wie MATLAB, formuliert werden.

Funktionell sind allgemeine Simulationssprachen und Blockschaltbild-Editoren sehr ähnlich. Ein Block entspricht einem Satz von Gleichungen mit definierten Ein- und Ausgangsgrößen, sowie Zustandsvariablen. Bevor simuliert werden kann, wird die Abarbeitungsreihenfolge der Blöcke durch teilweise automatisches „Sortieren“ ermittelt. Dies entspricht dem Sortieren der Gleichungen bei den Simulationssprachen. Mit dem Sortieren wird für jeden vorhandenen Block gewährleistet, dass für die Berechnung des Ausgangswertes der dazu erforderliche Eingangswert bekannt ist.

Dies kann dadurch gewährleistet werden, dass sie als Ausgangswerte von voranstehenden Blöcken bereits berechnet, oder aber „vorbekannt“ sind. Als vorbekannt werden solche Blöcke bezeichnet, die einen Anfangswert besitzen, der entweder beim Start der Simulation vorgegeben wird, oder aber in vorangegangenen Rechenschritten bestimmt wurde. Derartige Elemente werden auch als **History-Blöcke** bezeichnet im Gegensatz zu den **Not-History-Blöcken**. Dazu gehören beispielsweise alle Blöcke, mit denen Integrationen durchgeführt werden, sowie Totzeitglieder.

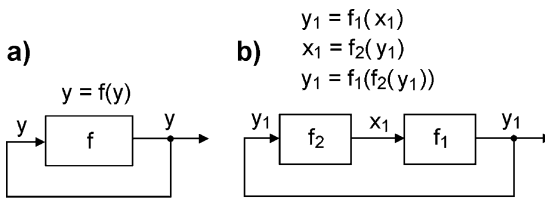


Bild 8-33: Algebraische Schleifen in einem Blockschaltbild: Rückkopplungsschleife über
a) einen Block b) mehrere Blöcke

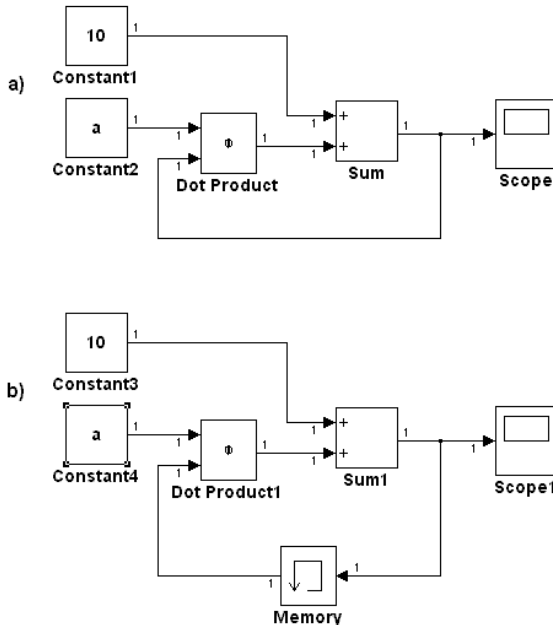


Bild 8-34: Blockstruktur mit algebraischer Schleife
a) nicht berechenbar b) berechenbar
durch Einschrittverzögerung in der Rückkopplung

Multiplikationsblock trägt die Bezeichnung „Dot Product“, die Summation wird durch einen Summierblock „Sum“ erzeugt und die Ausgabe erfolgt durch einen Anzeigeblock „Scope“.

Eine Umformung der Gl. (8.29) auf

$$y = \frac{10}{1 - a}$$

würde das Problem beseitigen. Ebenso würde, wie in Bild 8-34 b) gezeigt, das Einfügen einer Einschritt-Verzögerung „Memory“ das Problem lösen. Das liegt daran, dass bei dieser Verschaltung der rückgeführte Ausgang jeweils um eine Schrittweite des Simulationszyklus verzögert wird und somit der Funktionswert iterativ berechnet wird :

Schwierigkeiten treten auf, wenn in der Blockstruktur Rückkopplungsschleifen auftreten, in denen keine vorbekannten Werte vorliegen. Da solche Rückkopplungsschleifen keine Integrierer oder Totzeitblöcke enthalten, bezeichnet man sie wegen der rein algebraischen Struktur des Funktionszusammenhangs auch als **algebraische Schleifen**. Innerhalb einer solchen Schleife ist keine zulässige Rechenreihenfolge mehr möglich. In Bild 8-33 ist ein Beispiel für eine solche algebraische Schleife dargestellt, die sich auf nur einen aber auch über mehrere Blöcke erstrecken kann. In solchen Schleifen liegt ein mathematisches Modell vor, das eine Funktion enthält, deren Funktionswert gleichzeitig Argument ist. Solche algebraischen Schleifen müssen entweder aus der Blockstruktur entfernt werden, oder man kann versuchen mit Hilfe eines iterativ arbeitenden Gliedes die Schleife zu berechnen. In Bild 8-34 a) ist das Beispiel eines SIMULINK Blockschaltbildes dargestellt, das eine algebraische Schleife für die Gleichung

$$y = a \times y + 10 \quad (8.29)$$

zeigt. Die Zahlen werden über Konstantenblöcke mit den Bezeichnungen „Constant1“, usw. zugeführt, der

$$y(n+1) = a \cdot y(n) + 10 \quad .$$

Dieses Vorgehen ist nur unter eingeschränkten Bedingungen möglich, weil es bei zu großer Kreisverstärkung in der algebraischen Schleife zu numerischen Instabilitäten kommen kann. Der kritische Wert tritt hier für $a \geq 1$ auf.

Der Vorteil von Blockschaltbild-Editoren ist der Komfort bei der Eingabe und bei der Simulation. Nachteilig ist, ähnlich wie bei den allgemeinen Simulationssprachen, dass es meist schwierig und aufwendig ist, ein physikalisches Systemmodell in ein Blockschaltbild zu überführen. Da regelungstechnische Modelle meist schon als Blockschaltbilder vorliegen, entfällt dieses Problem und deshalb ist die Regelungstechnik auch einer der Hauptanwendungsbereiche von Blockschaltbild-Editoren. Sehr große Modelle können schnell unübersichtlich werden, weshalb man in SIMULINK die Möglichkeit besitzt, Teilmodelle in einen Makro zu überführen, der dann in der Grafik nur noch als ein Block erscheint.

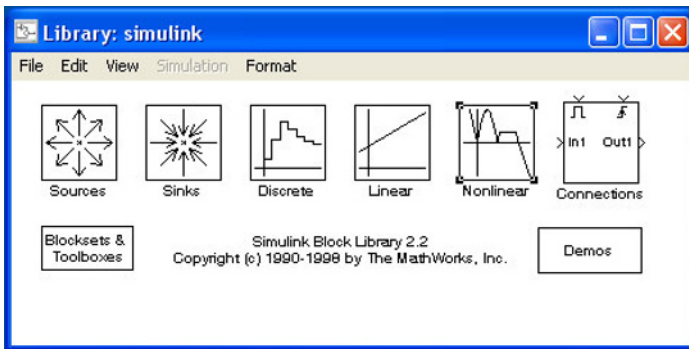


Bild 8-35: Blockbibliothek von SIMULINK

Normalerweise werden die Modellgleichungen interpretativ ausgeführt, was bei großen Modellen zu langen Rechenzeiten führt. SIMULINK besitzt deswegen die Möglichkeit der Compilierung des Modells mittels eines **Accelerators** in einen Code der Programmiersprache C. Dadurch wird ein in Echtzeit lauffähiges Modell erzeugt.

Im folgenden sollen die Möglichkeiten von Bildschirm-Editoren am Beispiel von SIMULINK kurz vorgestellt werden. SIMULINK ist eine Erweiterung des Mathematikprogramms MATLAB (MATrix-LABoratory), das seinen Schwerpunkt im Bereich der Matrizenberechnung und der numerischen Verfahren besitzt. Beim Aufruf von SIMULINK aus der MATLAB-Oberfläche erscheint als erstes ein Bildschirm mit einer Blockbibliothek (Bild 8-35).

Hinter jedem Blocksymbol verbirgt sich wiederum ein Auswahlménü von Funktionen. Man sieht, dass es sowohl Blöcke für lineares Übertragungsverhalten, solche für nichtlineares Verhalten und auch Blöcke für diskrete Funktionen gibt.

In Bild 8-36 ist das Fenster zu sehen, das sich öffnet, wenn man das „Linear-

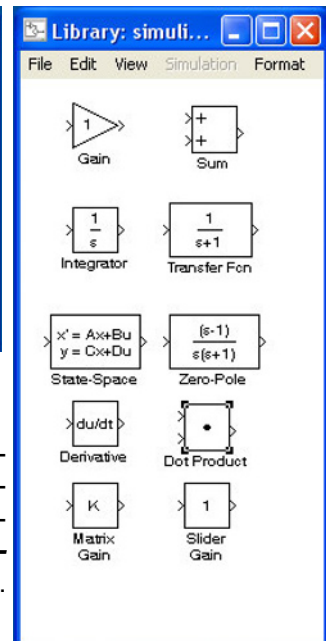


Bild 8-36: Auswahlmenü für lineare Funktionen

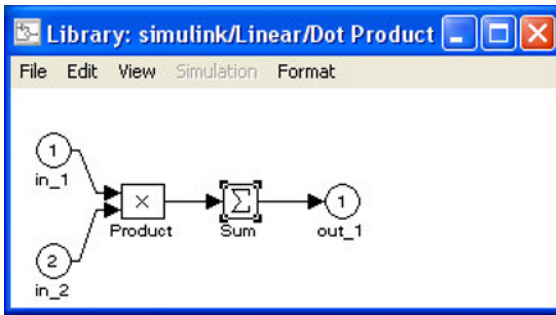


Bild 8-37: Demaskierte Form der Funktion „Dot Product“

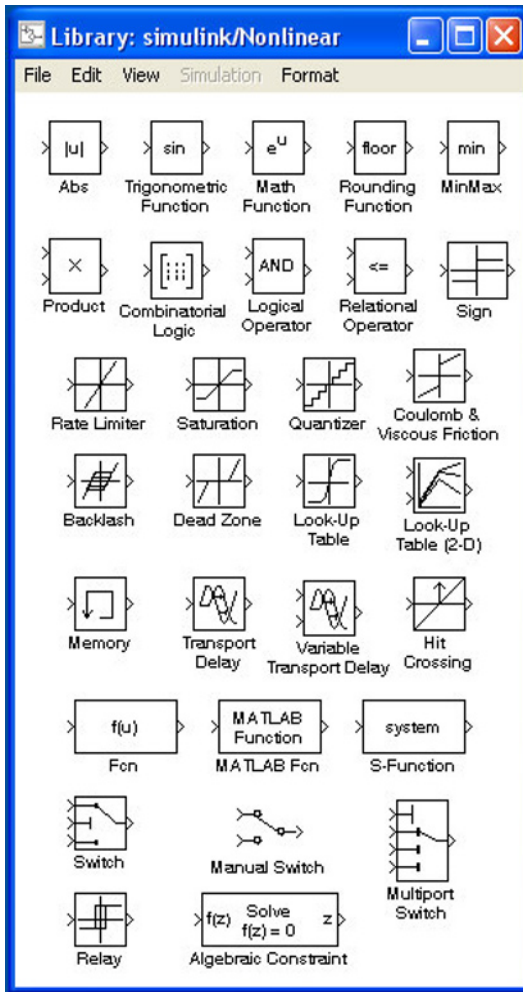


Bild 8-38: Auswahlfenster für nichtlineare Funktionen

Symbol“ anklickt. Es enthält Blöcke für Summation und Multiplikation, Integration und Bildung der Ableitung, Verstärker, Verzögerungsglieder und die Möglichkeit der Zustandsraumdarstellung. Viele Blöcke enthalten Makros, d. h. eine Anzahl von Blöcken, die **maskiert** sind und eine komplexere Funktion beinhalten. **Demaskiert** man beispielsweise den Block „Dot Product“ (Bild 8-37), so erhält man die Blockschaltbilddarstellung aller im Makro zusammengefassten Blöcke. Klickt man einen Block an, so erhält man eine Eingabemaske, die die Blockfunktion erläutert und die Möglichkeit bietet Parameter eines Blockes zu verändern. Bei der Aktivierung eines Blockes (Sum), wird dieser durch Anzeige von vier Marken an den Ecken, mit denen die Blockgröße und Position verändert werden kann, markiert.

In Bild 8-38 ist das Fenster dargestellt, das Blöcke nichtlinearer Funktionen enthält. Hier findet sich typische Nichtlinearitäten wie die Vorzeichenfunktion (Sign), Spiel (Backlash), Sättigung (Saturation), coulomb'sche und viskose Reibung (Coulomb & Viscous Friction), aber auch Relatoren, Boole'sche Operationen und mit dem Block Funktion (Fcn) kann man beliebige mathematische Funktionen wie die Sinusfunktion o. ä. eingeben. Auch Zusammenhänge, die nur in tabellarischer Form angegeben werden können sind verfügbar (Look-Up Table). Ein Funktionsblock wie beispielsweise die coulomb'sche und viskose Reibung, hat dabei den im Eingabefenster in Bild 8-39 dargestellten funktionalen Zusammenhang zwischen Ein- und Ausgang und kann, wie dort angegeben, parametrisiert werden. Bei dem im

Fenster in Bild 8-38 ebenfalls dargestellten Block „MATLAB Function“ handelt es sich um die Möglichkeit, mit der Programmiersprache MATLAB erzeugte Funktionen einzubinden.

Um nun eine Modellstruktur einzugeben, die wie am Anfang des Kapitels erläutert, am besten in einer Zustandsraumdarstellung vorliegt, zieht man aus den Bibliotheksfenstern Blöcke in

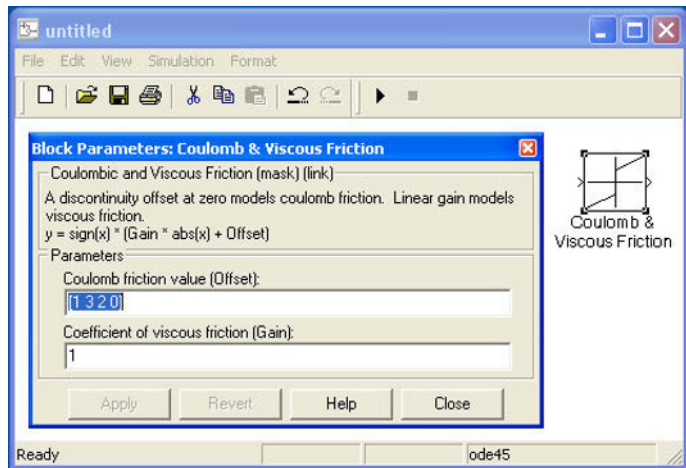


Bild 8-39: Eingabefenster für die nichtlineare Funktion *Coulomb'sche Reibung*

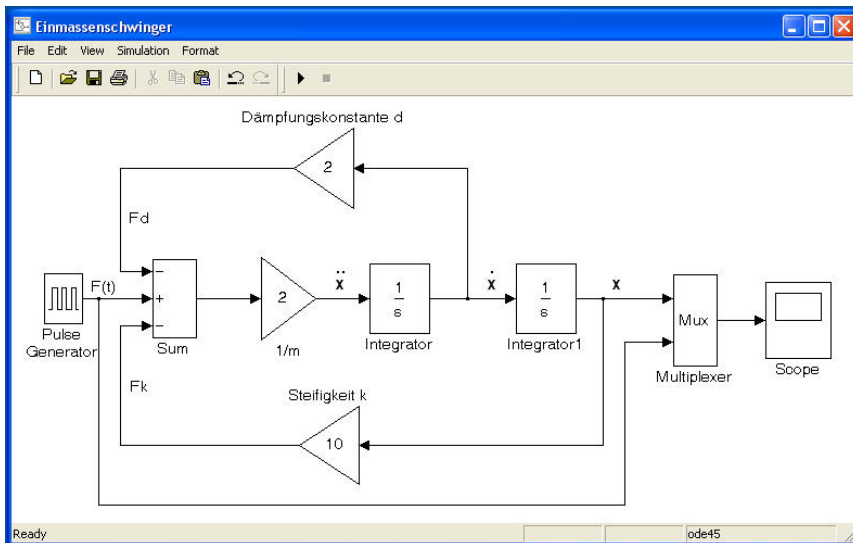


Bild 8-40: Simulationsmodell für den Einmassenschwinger in SIMULINK

ein Arbeitsfenster, ordnet sie dort in der richtigen Reihenfolge an und kann dann aus den Ausgängen der Blöcke Verbindungspfeile herausziehen. Diese kann man beliebig mit Eingängen anderer Blöcke verbinden. Blöcke, die keinen Eingang haben, werden als Source (Quelle) bezeichnet. Dies sind Blöcke wie Sprungfunktionen, Signalgeneratoren und Konstanteneingaben. Blöcke, die keinen Ausgang besitzen, werden als Sink (Senke) bezeichnet. Dies sind Ausgaben wie Oszilloskop, Grafikausgabe, Dateiausgabe, Animation. Ist die Struktur korrekt und wurden die Simulationsparameter richtig eingestellt, so kann sofort eine Simulation beginnen und beliebig oft mit gleichen oder geänderten Blockparametern wiederholt werden.

In Bild 7-62 war das Blockschaltbild des Einmassenschwingers hergeleitet worden. Hieraus kann man direkt ein SIMULINK-Modell zur Simulation des Schwingverhal-

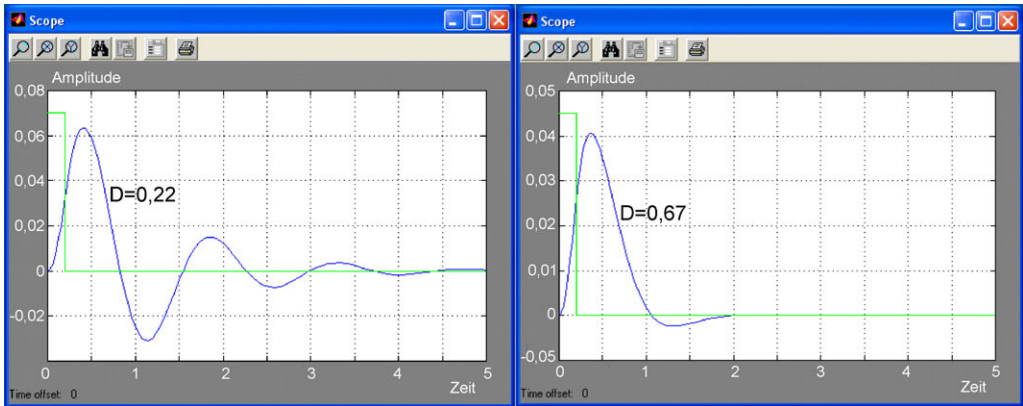


Bild 8-41: Scope-Anzeigen der Simulation des Einmassenschwingers bei unterschiedlicher Wahl des Lehr'schen Dämpfungsmaßes D

tens generieren (Bild 8-40). Beaufschlagt man das Modell beispielsweise mit einem impulsförmigen Eingangssignal, so ergeben sich je nach Dämpfung die in Bild 8-41 dargestellten Schwingungsverläufe für $D = 0,22$ und $D = 0,67$.

Außer den genannten Blockschalteditoren gibt es unter der MATLAB-Oberfläche noch weitere TOOLBOXES, mit denen man Signalverarbeitung, Systemidentifikation, Regelungstechnik (Analyse und Reglersynthese), Fuzzy-Technologie und neuronale Netzwerke behandeln kann.

Die Simulation komplexerer mechanischer Systeme bereitet mit den bis jetzt vorgestellten Möglichkeiten Probleme, weil die Modellierung beispielsweise von MKS so nicht möglich ist. Hier müssen die Rückwirkungen der Elemente untereinander berücksichtigt werden. Unter den nun folgenden objektorientierten Methoden wird auch das Simulationssystem CAMEL-View behandelt, das speziell zur Modellierung von MKS-Systemen ausgelegt ist.

8.3.5 Objektorientierte Modellbildung

Bei den bis jetzt besprochenen Simulationsmethoden ist es meist so, dass man einen Satz von differential-algebraischen Gleichungen als mathematisches Modell aufstellt und versucht, diese durch Integration und andere Berechnungsmethoden zu lösen. Hat man beispielsweise ein Simulationsdiagramm wie das in Bild 8-40 gezeigte, mit dessen Hilfe man einen Einmassenschwinger mit Dämpfung simulieren kann, so gibt es für die einzelnen physikalischen Objekte Masse, Feder, Dämpfer keine direkten Entsprechungen als einzelne Teilsysteme im Diagramm. Es wäre jedoch wünschenswert, die physikalischen Objekte eines Systems mit ihren Eigenschaften einzeln zu modellieren und sie dann so, wie sie in der Realität miteinander in Verbindung stehen, zusammenzufügen. Dabei sollte es dann auch möglich sein, Objekte der verschiedenen Fachdisziplinen (Elektrotechnik, Mechanik usw.) zusammen zu modellieren, wie sie in mechatronischen Systemen enthalten sein können. Diese Möglichkeiten bieten objektorientierte Modellwerkzeuge und Simulationssys-

teme wie beispielsweise DYMOLA² (**D**ynamic **M**odeling **L**anguage), 20-sim³ und CAMEL-View⁴.

Das Grundprinzip von DYMOLA und 20-sim besteht darin, physikalische Systeme möglichst direkt auf Modellobjekte abzubilden. Dabei wird kein Wirkzusammenhang festgelegt, sondern es werden nur Beziehungen zwischen Objekten, bzw. beschreibenden Variablen, definiert.

Große Systeme werden dazu hierarchisch in kleinere Submodelle zerlegt. Diese Submodelle sind in Klassen geordnet, die gleichartige Objekte enthalten. Es gibt eine umfangreiche Bibliothek solcher Klassen und Objekte, so dass bei der Bildung neuer Modelle auf diese zurückgegriffen werden kann. Objekte können dann entsprechend der physikalischen Verkopplung miteinander verschaltet werden. Aus Anwendersicht ist damit eine „natürliche“ Modellierung möglich, die dem jeweiligen Fachgebiet angepasst ist. Diese Art der Darstellung ist besonders gut dem Anliegen der Mechatronik angepasst, alle unterschiedlichen Fachdisziplinen mit fachübergreifenden gemeinsamen Konzepten dazustellen und zu beschreiben. Ein Konzept, das es in besonderer Weise ermöglicht, Probleme aus den unterschiedlichen Fachgebieten mit einer einheitlichen Darstellung zu behandeln, ist die Methode der **Bondgraphen** [8.2], [8.3], [8.11] (s. Abschnitt 8.2).

8.3.5.1 20-sim

Dieses Simulationssystem ist ein Nachfolger des an der Universität Twente entwickelten Programmersystems TUTSIM. Neben anderen Anwendungsmöglichkeiten stellt es eine ausgereifte und intuitiv zu bedienende Software für die Modellbildung und Simulation dar, insbesondere für Bondgraph Modelle mechatronischer Systeme [8.9].

20-sim verwendet für die Modellbildung hierarchische Modelle, die durch sehr unterschiedliche Darstellungsformen modelliert werden können. Dazu gehören Bondgraphen, Diagramme mit grafische Grundelementen (Icons), Blockschaltbildsymbole, mathematische Gleichungen oder auch Kombinationen aller dieser Eingabemöglichkeiten. Alle Modellformen sind in der in Bild 8-42 dargestellten Struktur aufgebaut, d. h. jedes Modell besteht aus einer hierarchisch übereinander angeordneten Struktur von Submodellen. Auf jeder Stufe der Struktur besteht ein Submodell wiederum aus Submodellen der darunter liegenden Ebene. Auf der untersten Ebene besteht das elementarste Submodell dann immer aus Gleichungen in der Programmiersprache SIDOPS+ (**S**tructured **I**nterdisciplinary **D**escription **O**f **P**hy-

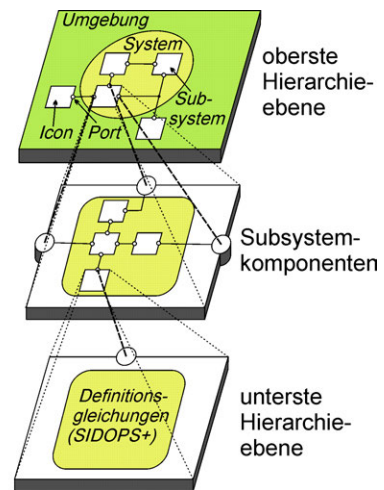


Bild 8-42: Hierarchieebenen von 20-sim

2 DYMOLA wurde von H. Elmquist entwickelt und ist ein Produkt der Fa. Dynasim AB/Schweden

3 20-sim ist ein Produkt der Fa. Controllab Products B.V./Niederlande

4 CAMEL-View ist ein Produkt der Fa. IXtronics GmbH, Paderborn

sical Objecta) [8.10], in der Formulierungen der normalen mathematischen Notation sehr stark ähneln. Die Programmiersprache kann viele mathematische Funktionen durch „built-in“ Funktionen realisieren, deren Anwendung intuitiv ist.

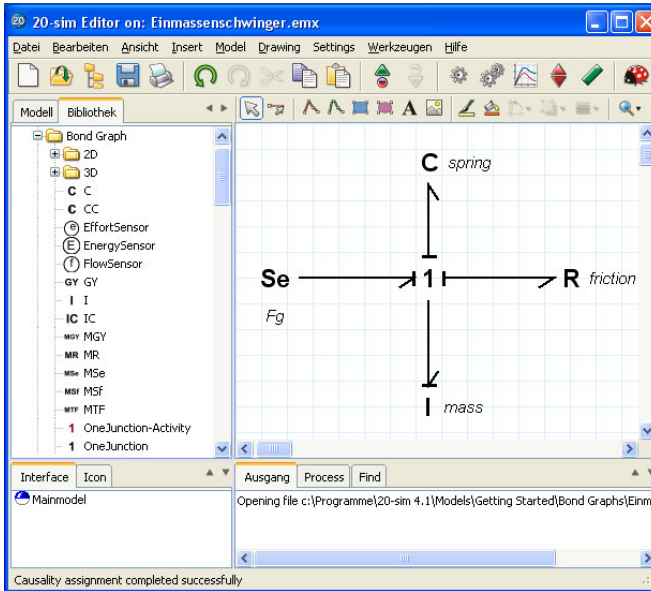


Bild 8-43: 20-sim Editor mit Bondgraph Modell des Einmassenschwingers

schwingers dargestellt (s. a. Bild 8-31), in dem die Leistungsbonds noch mit sog. Kausalitätsquerstrichen versehen sind. Diese dienen zur Vorgabe eines eindeutig integrierbaren Modellgleichungssystems durch Vorgabe der Richtung des Efforts. Dieses Konzept kann hier nicht eingehend behandelt werden.

Da sich das dargestellte Modell nicht noch aus Submodellen aufbaut, ist im Interfacefenster links unten nur das „Mainmodel“ angegeben. Würden darin „Submodels“ verwendet, so würde dort die hierarchische Struktur dargestellt werden. Im Ausgabefenster rechts unten werden Systemmeldungen über benutzte Files und Compilermeldungen dargestellt.

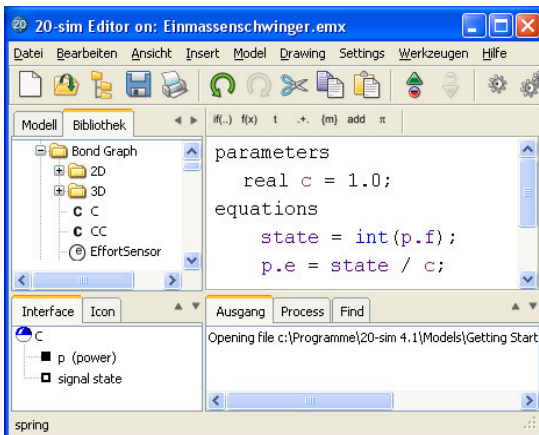


Bild 8-44: Darstellung des C-Elementes des Einmassenschwingers in SIDOPS+

Ähnlich wie in Simulink arbeitet man bei der Eingabe von 20-sim Modellen mit einem grafischen Editor, mit dem man beispielsweise die in Abschnitt 8.2 vorgestellte Symbolik von Bondgraphen direkt eingeben kann (Bild 8-43). Die Symbole werden links oben in einem Bibliotheksfenster angeboten und können durch „drag and drop“ in das rechte grafische Editorfenster gezogen und angeordnet werden. Mit einem Verbindungswerkzeug werden dann die Bondgraph-Symbole durch Leistungsbonds untereinander verbunden. In Bild 8-43 ist das Modell des Einmassenschwingers dargestellt

(s. a. Bild 8-31), in dem die Leistungsbonds noch mit sog. Kausalitätsquerstrichen versehen sind. Diese dienen zur Vorgabe eines eindeutig integrierbaren Modellgleichungssystems durch Vorgabe der Richtung des Efforts. Dieses Konzept kann hier nicht eingehend behandelt werden. Da sich das dargestellte Modell nicht noch aus Submodellen aufbaut, ist im Interfacefenster links unten nur das „Mainmodel“ angegeben. Würden darin „Submodels“ verwendet, so würde dort die hierarchische Struktur dargestellt werden. Im Ausgabefenster rechts unten werden Systemmeldungen über benutzte Files und Compilermeldungen dargestellt.

Klickt man nun eines der Modellelemente, beispielsweise das C-Element (hier die Feder des Einmassenschwingers an, so kann man sich die automatisch mit dem grafischen Symbol verbundenen Simulationsgleichungen in SIDOPS+ ansehen und editieren (Bild 8-44). Das Mo-

de layout wird durch sog. „Keywords“ gegliedert, wovon in diesem einfachen Modellbereich nur die Worte „Parameter“ und „Equations“ auftauchen. Unter Parameter wird der Wert von C (Nachgiebigkeit der Feder) angegeben. Für die Feder muss die Definitionsgleichung (8.19) modelliert werden. Dazu muss der Flow integriert werden, was in der ersten Gleichung unter „Equations“ geschieht. Da Bondgrafelemente mehrere Ports besitzen können, wird einer Effort- oder Flow-Variablen die Portbezeichnung getrennt durch einen Punkt (Dot) vorangestellt. Da das C-Element nur ein Port besitzt, heißt die zugehörige Flowvariable „p.f“. Die Zustandsvariable „state“ wird dann durch Integration ($\text{int}(p.f)$) der

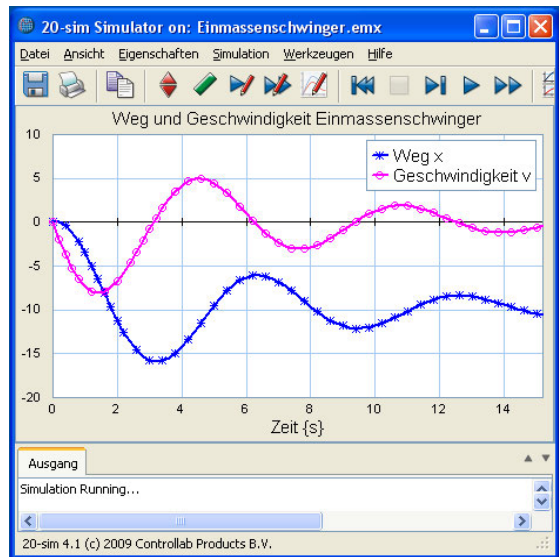


Bild 8-45: Simulationsergebnisse des Einmassenschwingers

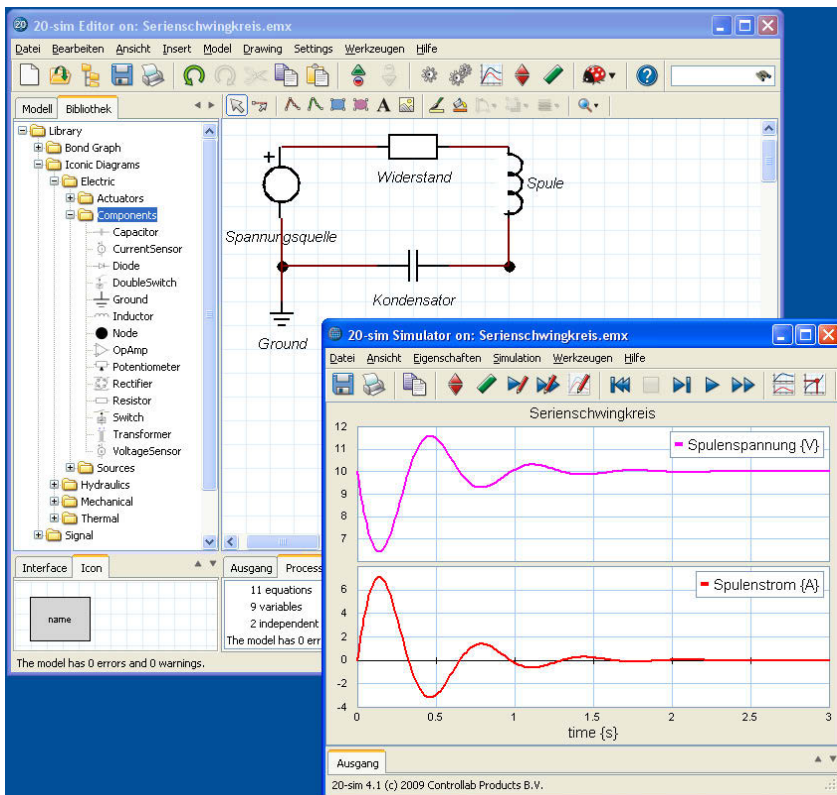


Bild 8-46: Eingabe eines Serienschwingkreises mit graphischen Icons und Simulationsergebnis

Flow-Variablen gebildet. Die zweite Modellgleichung liefert dann die Effort-Variable durch Division der Zustandsvariablen „state“ durch die Nachgiebigkeit „C“. Dies entspricht der Gl. (8.19).

Nach Parametervorgabe und Überprüfung des automatisch aus dem Bondgraphen erstellten SIDOPS+-Modells, kann ein Simulationslauf durchgeführt und grafisch dargestellt werden. Bild 8-45 zeigt den Verlauf von Weg und Geschwindigkeit der Masse für ein schwach gedämpftes System.

Die grafische, objektorientierte Modelleingabe kann aber mit 20-sim noch weiter vereinfacht werden, da es die Möglichkeit gibt, technische Objekte unterschiedlicher Domänen ohne Umweg über den Bondgraphen direkt einzugeben. Dazu kann man „Icons“ aus unterschiedlichen Domänenbibliotheken (Elektrisch, mechanisch, hydraulisch, thermisch) wieder in den Bereich des grafischen Editors ziehen und diese untereinander verbinden. Im Bild 8-46 ist so der elektrische Serienschwingkreis aus Bild 8-30 modelliert worden. Im Bildfenster im Vordergrund, dem Simulationsfenster, ist das dynamische Verhalten des gedämpften Schwingkreises dargestellt.

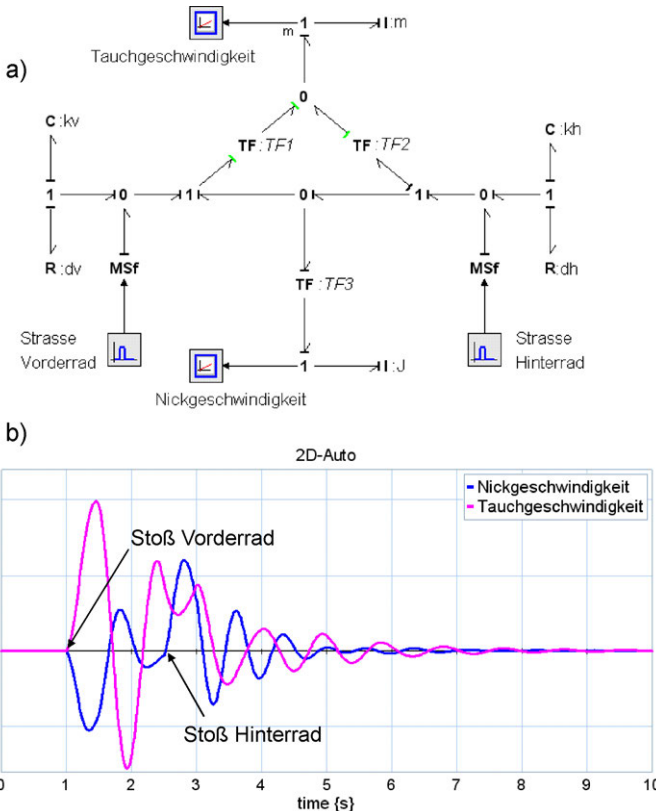


Bild 8-47: Simulation eines 2D-PKW mit einem Nick- und einem Tauchfreiheitsgrad a) Bondgraph b) Simulationsergebnis beim Überfahren einer Bodenwelle

Das Simulationssystem 20-sim ist auch gut zur Modellierung von Mehrkörpersystemen geeignet, die man als Bondgraphen oder Icongraphen direkt objektorientiert vornehmen kann. In Bild 8-18 war bereits der Wort-Bondgraph eines Mehrkörpersystems in Form eines PKW-Antriebsstrangs dargestellt worden. Für die Simulation wurde dieses Modell des ganzen Autos nochmals, wie in Bild 2-16 b) dargestellt, auf ein halbes Auto mit den zwei Freiheitsgraden **Tauchen** und **Nicken** reduziert. Das Simulationsmodell ist in Bild 8-47 a) dargestellt, das Simulationsergebnis in Bild 8-47 b) für eine Anregung durch eine Bodenwelle in der Fahrbahn.

Weitere Tools, die 20-sim zur Verfügung stellt, sind beispielsweise die Durchfüh-

nung einer FFT-Analyse. Bild 8-48 zeigt eine solche Analyse für ganz bestimmte Simulationsparameter. Aus dem FFT-Diagramm kann man ablesen, dass die Taucheneigenfrequenz $F_{Tmax} = 1,1$ Hz und die Nickeigenfrequenz bei $F_{Nmax} = 1,4$ Hz liegt. Die verwendeten Simulationsparameter (s. Bild 2-16 b) und Bild 4-22) waren dabei:

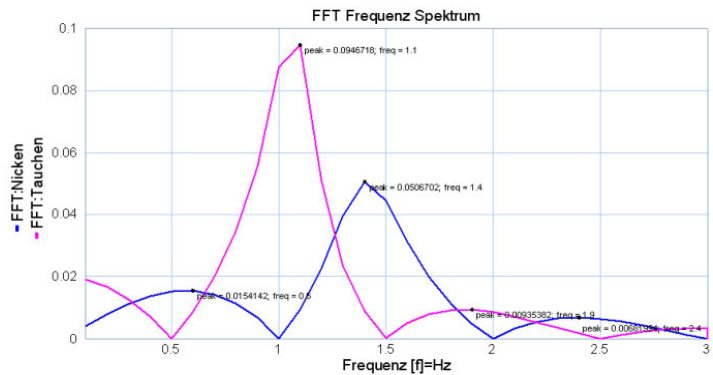


Bild 8-48: FFT-Analyse der Signalverläufe des Tauchens und Nickens beim Fahren über eine Bodenwelle

$$\Theta_S = 2000 \text{ kgm}^2 \quad M = 1500 \text{ kg} \quad k_1 = 35000 \text{ N/m} \quad k_2 = 3500 \text{ N/m}$$

$$a = 1,5 \text{ m} \quad b = 1,5 \text{ m} \quad d_1 = 1000 \text{ Ns/m} \quad d_2 = 1000 \text{ Ns/m}$$

Die Nachrechnung der Eigenfrequenzen ergibt dann folgendes:

Translation (Tauchen): $\omega_e = \sqrt{\frac{k}{m}} = \sqrt{\frac{k_1 + k_2}{M}} = 6,83 \text{ rad/s} \Rightarrow f_T = \frac{\omega_e}{2\pi} = 1,09 \text{ Hz}$

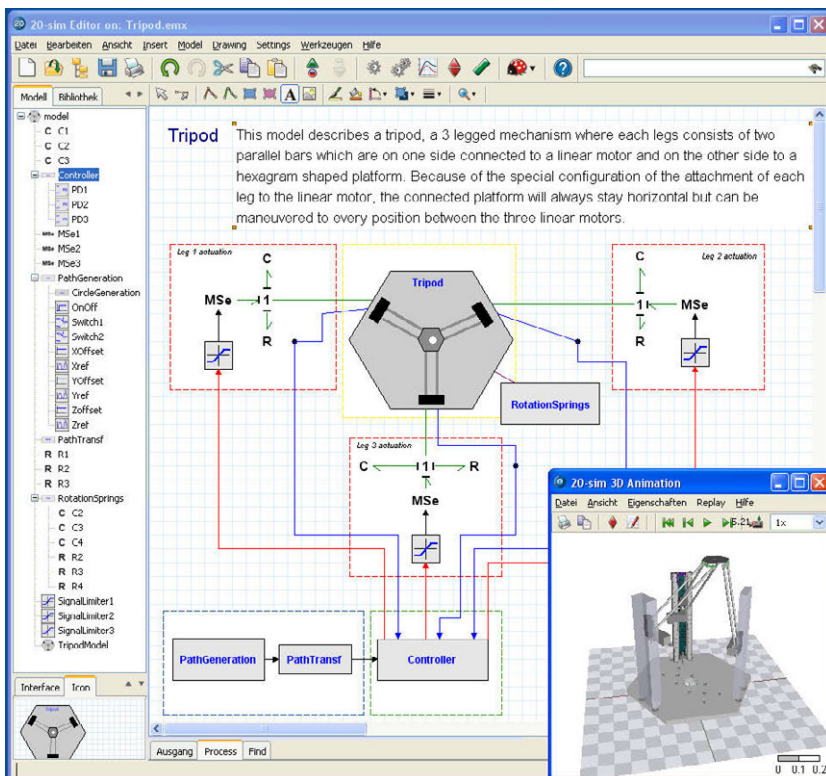


Bild 8-49: Simulationsmodell des mechatronischen Bewegungssystems Tripod

Rotation /Nicken): $\omega_e = \sqrt{\frac{k \cdot r^2}{\Theta_s}} = \sqrt{\frac{(k_1 + k_2) \cdot a^2}{\Theta_s}} = 8,87 \text{ rad/s} \Rightarrow f_N = \frac{\omega_e}{2\pi} = 1,41 \text{ Hz}$

Dies zeigt, dass die FFT-Analyse sehr gut mit den nachrechenbaren Eigenfrequenzen übereinstimmt.

Für die Simulation komplexer mechatronischer Bewegungssysteme können in die Modelle, neben den bis jetzt behandelten Elementen, auch dreidimensionale Volumenmodelle der Bewegungseinrichtung eingebunden werden. Das Bild 8-49 zeigt das Simulationsmodell eines sog. **Tripod**, eine dreiachsige Bewegungseinrichtung, die ihre Arbeitsplattform dreidimensional im Raum bewegen kann. Die Bewegungen des Volumenmodells können mit Hilfe einer 3D-Animation visualisiert werden. Im Bild sind im linken Fenster die hierarchische Struktur des Tripods und die zugehörigen Elemente dargestellt. Im grafischen Editor ist nur die oberste Ebene der Struktur dargestellt, wobei die Submodelle durch Anklicken geöffnet werden können. Im Bild rechts unten ist eine Szene aus der 3D-Animation abgebildet, mit

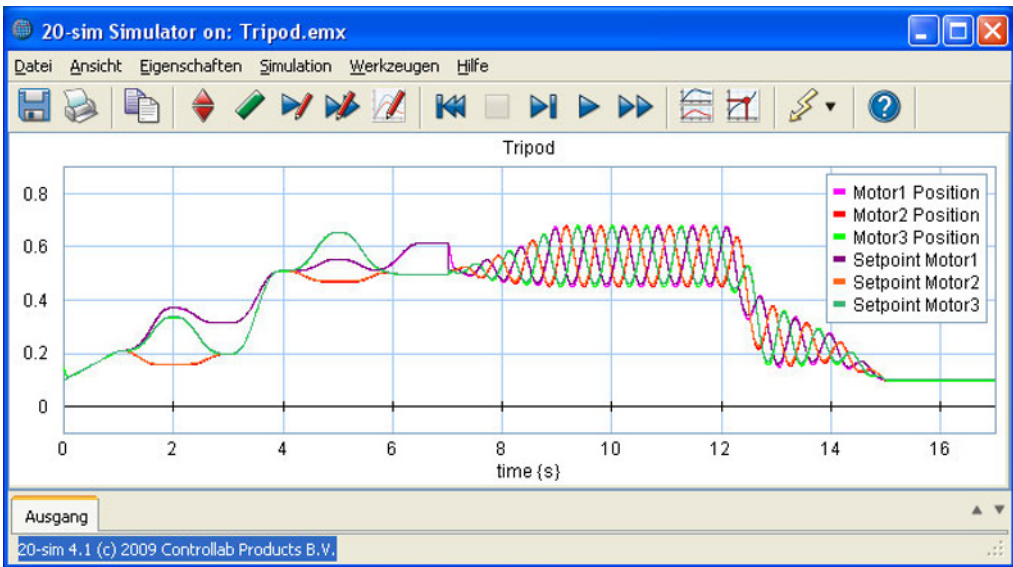


Bild 8-50: Zeitliche Bewegung der drei Achsen und Verlauf der Sollwerte (Setpoint)

der man auch Kollisionsuntersuchungen durchführen kann. Natürlich kann man auch die Bewegungen der einzelnen Achsen und den Verlauf der Sollwerte in der üblichen Darstellung abbilden (Bild 8-50).

Neben diesen Methoden können auch wie in SIMULINK Blockschaltbilder eingegeben werden und 20-sim-Modelle können nach MATLAB/SIMULINK exportiert werden. Im Bereich der Regelungstechnik können auch neuronale Netzwerke verwendet werden. Das Programmsystem 20-sim kann zu Erprobungszwecken von der Webseite www.20sim.com herunter geladen werden. Mit dieser Version können alle Module erprobt und Beispiele ausgeführt, jedoch eigene Modelle nicht gespeichert werden.

8.3.5.2 CAMEL-View

Ein speziell für die Modellierung von MKS-Systemen entwickeltes Simulationssystem ist CAMEL-View. Dieses Programmsystem ist ein Entwurfswerkzeug, das die Modellierung, Analyse und Synthese mechatronischer Systeme innerhalb einer Entwicklungsumgebung ermöglicht. Hierin werden für die Modellierung disziplinspezifische Bauteile, wie z. B. starre Körper, Federn und Gelenke, verwendet. Dabei erfolgt die Modellierung anders als bei einem Blockschaltbildeditor auf physikalischer Ebene. Für die Modellierung steht ein 2D/3D-Modellierer (Modelling Workbench) zur Verfügung, in dem Modelle erstellt und geändert werden können [8.5].

Die Ermittlung der Modellgleichungen erfolgt aufgrund des physikalischen Modells automatisch. Für die Analyse von Modellen wird ebenfalls automatisch ein C-Programmcode erzeugt, übersetzt und gebunden. Die Analyse ermöglicht nichtlineare Simulation, die Linearisierung nichtlinearer System, sowie die Berechnung von Eigenwerten und Frequenzgängen. Weitere Module sind die grafische Darstellung der Analyseergebnisse (3D-Animation, 2D-Plots) und die automatische Dokumentation von Modellen.

CAMEL-View verwendet die Beschreibungssprache **Objective DSS** (Objective-Dynamic System Structure) für die Modellbildung mechatronischer Systeme [8.6]. Der Anwender kann die Modellbildung vollständig mit Hilfe der Beschreibungssprache durchführen oder aber seine Modelle grafikunterstützt aufbauen. Wird der zweite Weg gewählt, dient Objective-DSS als textuelles, lesbares Speicherformat von CAMEL-View-Modellen.

Objective-DSS stellt sogenannte **Metaklassen** zur Verfügung. Metaklassen sind Klassen, deren Objekte wiederum Klassen darstellen. Sie beinhalten z. B. die Informationen über die möglichen Schnittstellen und die automatische Generierung des mathematischen Modells. Während die Metaklassen nicht vom Anwender geändert werden können, sondern lediglich das Gerüst bilden, können die daraus gebildeten „Baelementklassen“ beliebig geändert und erweitert werden. Die aus den Baelementklassen erzeugten Objekte werden Baelemente genannt. Von einer Baelementklasse können mehrere Baelemente erzeugt werden, d. h. Baelementklassen können beliebig oft verwendet werden. Welchen Sprachumfang eine Objective-DSS-Klasse besitzt, ist durch die Wahl der Metaklasse festgelegt. Vorteil dieser Vorgehensweise ist eine an die jeweilige Domäne (Mechanik, Regelungstechnik usw.) angepasste Syntax.

Teilbaugruppen können in der Entwurfsphase aus einer Bibliothek entnommen, modifiziert und in der 2D- oder 3D-Darstellung eingefügt, untereinander durch Koppelpunkte verbunden und die Ein-/Ausgangsverbindungen hergestellt werden. Die Linien zwischen den Baugruppen werden automatisch aufgrund der Eingabe von Verbindungsbeziehungen erstellt. Bei Mehrkörpersystemen ist vor allem die Positionierung in der 3D-Ansicht günstig, weil man hier die Lage der Teile zueinander und ihre Berührungspunkte exakt festlegen kann. Bild 8-51 zeigt die Bildschirmdarstellung der Gesamtstruktur des Modells „Halbachse“ mit allen Untermodellen und im rechten Bildteil die 3D-Darstellung des Modells. Man kann sich in diesem Fenster auch alle Teilbaugruppen bis herunter zur Bauteilebene darstellen.

Vorteil der Modellbildung mit Hilfe fachspezifischer Bindungen ist, dass die baeuele-

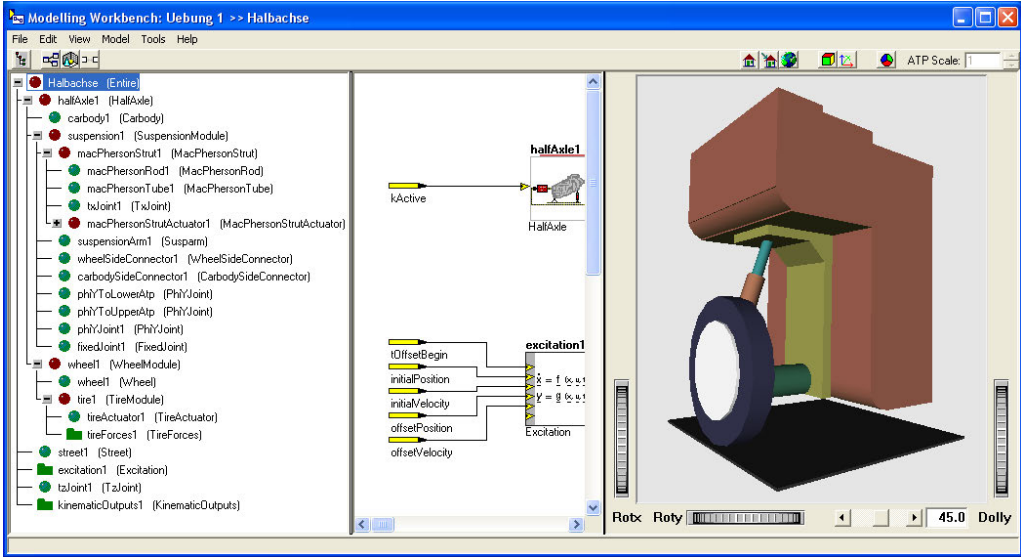


Bild 8-51: Struktur des Modells Halbachse und grafische Darstellung der Baugruppe im rechten Fenster

ment- und baugruppenorientierte Sicht auf ein System erhalten werden kann. Das Ziel, die Baugruppenstruktur des Systems auf der mathematischen Ebene zu erhalten, ist in den einzelnen Modellbildungsdomänen nur mit Hilfe speziell angepasster Formalismen zu erreichen. Insbesondere in der Mehrkörpersystemdynamik ist die Erhaltung der Baustuktur im mathematischen Modell mit einem nicht geringen Overhead hinsichtlich des Laufzeitverhaltens der numerischen Berechnungen verbunden.

Ein weiterer Vorteil der Modellbildung mit Hilfe fachspezifischer Kopplungen ist, dass sich der topologische Aufbau durch das Hinzufügen von Bauelementen und Baugruppen ausschließlich additiv ändert, unabhängig davon, welcher Formalismus zur Ableitung des mathematischen Modells verwendet wird.

Mit Hilfe dieser auf die Mechatronik zugeschnittenen Vorgehensweise gelingt es komplette mechatronische Systeme zu simulieren und ihre Dynamik zu untersuchen.

8.3.6 Hardware-in-the-Loop, Software-in-the-Loop

Im Laufe des Entwicklungsprozesses von mechatronischen Systemen wird man nicht alle Baugruppen ausschließlich auf einem Digitalrechner simulieren wollen, da dies aufgrund stets notwendiger Modellvereinfachungen in komplexen Systemen zu falschen Aussagen führen kann. Daher benutzt man auch die Möglichkeiten entweder körperlich Hardwarekomponenten in eine Simulation mit einzubeziehen (**Hardware-in-the-Loop: HIL**) oder eine entwickelte Simulation in ein mechatronisches System einzubinden (**Software-in-the-Loop: SIL**).

8.3.6.1 Hardware-in-the-Loop

Hierunter versteht man die Integration von realen Komponenten (Bauteilen und Systemmodellen in eine gemeinsame Simulationsumgebung [8.8]. Die HIL-Nachbildung (Simulation) dynamischer Systeme durch physikalische und mathematische Modelle muss dabei in Echtzeit und unter Nachbildung der physikalischen Randbedingungen erfolgen. Ein Beispiel ist die Simulation eines Gesamtfahrzeuges am Rechner mit der Anbindung eines realen Steuergerätes und der Aktorik für eine Funktionsregelung zur Fahrstabilitätsregelung. Ein entscheidender Vorteil der HIL ist der Funktionstest des Steuergerätes unter realen Bedingungen bei gleichzeitiger Einsparung von zeit- und kostenintensiven Fahrmanövern. Simulationssysteme, die diese Art der Echtzeit-Simulation erlauben, sind **CAMEL-View** und **dSPACE**. Das letztgenannte System verwendet MATLAB/SIMULINK Modelle und erzeugt einen echtzeitfähigen Code, der auf spezieller Hardware lauffähig ist.

8.3.6.2 Software-in-the-Loop

Hierunter versteht man die Integration von Systemmodellen in eine gemeinsame Simulationsumgebung mit dem modellierten Prozess (Regelstrecke); sowohl die zu entwickelnde Funktion als auch der Prozess, auf den die Funktion einwirkt, werden modelliert. Die SIL-Nachbildung (Simulation) dynamischer Systeme durch physikalische und mathematische Modelle muss dabei nicht in Echtzeit erfolgen. Ein entscheidenden Vorteil der SIL ist der Funktionstest unter simulierten Bedingungen bei gleichzeitiger Einsparung von zeit- und kostenintensiven Experimente (z.B. Fahrmanöver). Ausgehend von der SIL-Umgebung können entweder die Funktion, der Prozess oder beide Teile physikalisch realisiert und im geschlossenen Kreis hinsichtlich ihres Verhaltens analysiert werden.

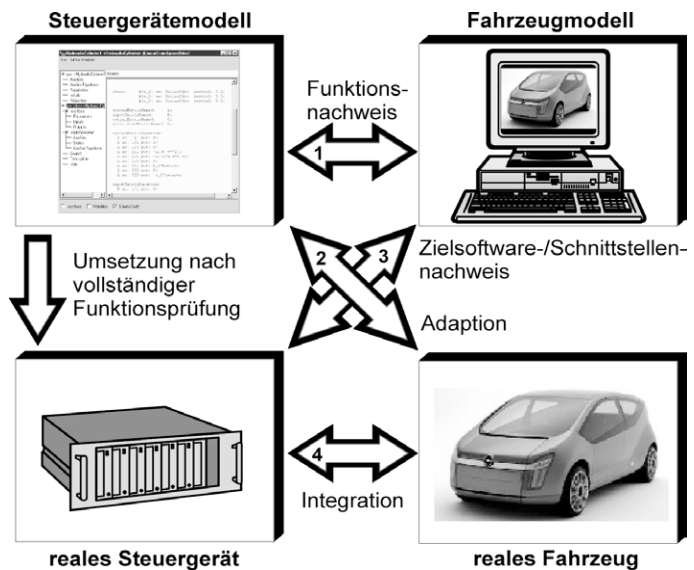


Bild 8-52: Arbeitsschritte der Eigenschaftsabsicherung einer als Simulationsmodell entwickelten Komponente am Beispiel eines Steuergerätes für ein Kraftfahrzeug

8.3.6.3 Kopplung von Modellen und Prototypen

Will man eine Komponente eines mechatronischen Systems unter Verwendung von HIL und SIL entwickeln, so muss man verschiedene Arbeitsschritte durchlaufen, um

die Eigenschaften der zu entwickelnden Komponente abzusichern. In Bild 8-52 ([8.8]) sind die einzelnen Arbeitsschritte am Beispiel der Entwicklung eines Steuergerätes für einen PKW dargestellt. Dies sind im einzelnen:

- 1.) **Funktionsnachweis:** Eine neue oder veränderte Funktionalität eines Steuergerätes wird als Modell in einem geschlossenen Regelkreis mit einem Streckenmodell (Prozessmodell) getestet. Diese Untersuchung wird als Software-in-the-Loop (SIL) bezeichnet.
- 2.) **Adaption:** Die am Streckenmodell überprüfte Funktion kann dann an dem realen Prozess abgestimmt werden (so genannte Applikation).
- 3.) **Zielsoftware-/Schnittstellennachweis:** Durch die Kopplung des realen Steuergerätes mit dem Streckenmodell in einer HIL-Umgebung kann die Fehlerfreiheit der Zielsoftware und der Schnittstellenkommunikation überprüft werden.
- 4.) **Integration:** Die Integration des mit einer neuen Funktionalität ausgestatteten Steuergerätes in den realen Prozess erlaubt die Erprobung des Gesamtsystems und die Anpassung aller relevanten Signal- und Steuerdaten.

Eine solche Kombination aus virtuellen und realen Tests neuer Komponenten eines mechatronischen Systems verkürzt die früheren Entwicklungs- und Erprobungszeiten erheblich.

8.3.7 Simulationssysteme für Industrieroboter

Für bestimmte Technologien gibt es spezialisierte Simulationssysteme, die auf die Erfordernisse der jeweiligen Anwendungen zugeschnitten sind. Bei Industrierobotern ist für die Planung komplexer Arbeitszellen, die räumliche Anordnung der physikalischen Objekte, einschließlich des Roboterarms innerhalb der Zelle, ein wichtiges Planungsziel. Ein anderes Ziel ist die Festlegung der Bewegungsbahnen des Roboters für die Arbeitsaufgabe, ohne dass die physikalische Arbeitszelle bereits zur Verfügung steht und die Bestimmung der Prozesszeit. Für diese Aufgaben reicht eine Simulation der Geometrie der Objekte in der Zelle und die Modellierung der Kinematik der bewegten Objekte aus. Ziel eines solchen Systems ist es, automatisch aus einer grafischen Simulation das Bewegungsprogramm des Roboters zu erzeugen.

Robotersimulationssysteme gehen in ihren Möglichkeiten weit über die reine Programmerzeugung hinaus. Sie schaffen die Voraussetzungen, eine komplette Anwendung inklusive Umgebung des Roboters (Peripherie), vom Effektor bis hin zu Fördereinrichtungen, Sensoren und Sicherheitseinrichtungen zu gestalten und in bewegten Bildern zu simulieren. Dabei kann nicht nur die Anordnung der Anlagenkomponenten entwickelt, simuliert und optimiert werden, sondern auch deren Funktion. In Bild 8-53 ist der allgemeine Aufbau eines Roboter-Simulationssystems mit den wichtigsten Programmmodulen dargestellt. Neben der Möglichkeit die Geometrie des Roboters und der gesamten Applikationsumgebung zu modellieren, können auch Geometriedaten von beispielsweise Werkstücken direkt aus anderen CAD-Programmen übernommen werden. Bei manchen Systemen kann außer der Kinematik auch die Kinetik des Robotersystems modelliert werden, so dass sich auch

das dynamische Verhalten direkt simulieren lässt.

Grafische Robotersimulationssysteme sind in der Regel vom Robotersystem unabhängig, so dass verschiedenartige Robotertypen unterschiedlicher Hersteller für eine bestimmte Anwendung getestet werden können, um in der Simulation den für die Aufgabe am besten geeigneten Typ zu bestimmen. Roboterhersteller bieten auch für ihre Produkte entwickelte Simulationssysteme an. Durch Untersuchungen des Arbeitsbereiches und durch Kollisionsbetrachtungen kann die kinematische Tauglichkeit eines Robotersystems beurteilt werden. Hierfür stehen außerdem Funktionen zur Simulation des Geschwindigkeits- und Beschleunigungsverhaltens zur Verfügung, mit denen das Zeitverhalten eines Robotersystems in Abhängigkeit vorgegebener Randbedingungen untersucht werden kann.

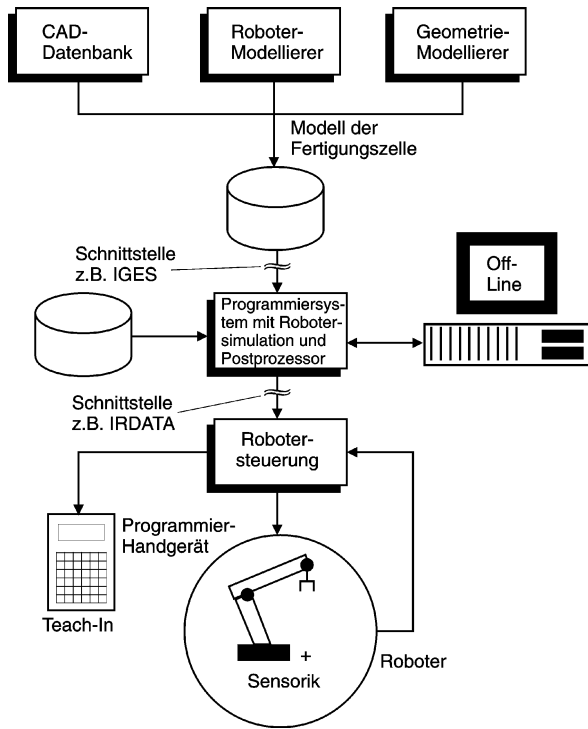


Bild 8-53: Allgemeiner Aufbau eines Roboter-Simulationssystems

Am Markt eingeführte, herstellernunabhängige Systeme zur grafischen Simulation von Roboteranwendungen sind:

- CATIA-Robotics System (von Dassault/IBM),
- GRASP (Grafical Robot Applikation Simulation Package von BYG),
- ROBCAD (Robot Computer Aided Design von Tecnomatix),
- WORKSPACE (Robot Simulations Ltd.)

Die führenden Roboterhersteller bieten ebenfalls eigene Simulationssysteme an, die dann in der Regel die Robotergeometrien bereits beinhalten. Mit Hilfe des inte-

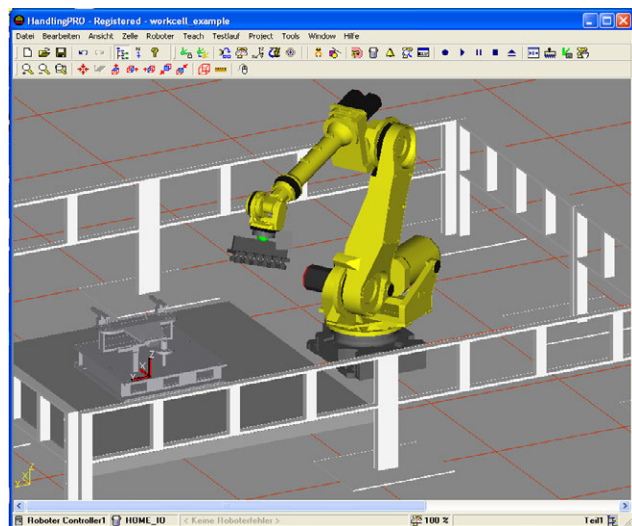


Bild 8-54: Roboterzelle des Simulationssystems Roboguide

grierten CAD-Systemen kann außerdem die Geometrie von anderen Robotern sowie der Prozessperipherie wie Werkstücke, Greifer, Einhausungen modelliert werden. Bild 8-54 zeigt eine einfache Roboterzelle, die mit dem Simulationssystem "Roboguide" der Fa. Fanuc Robotics modelliert wurde. Sie enthält einen 6-achsigen Knickarmroboter mit einem Greifer, sowie ein in einer Vorrichtung eingespanntes Werkstück. Die Zelle ist von einem einfachen Zaun umgeben.

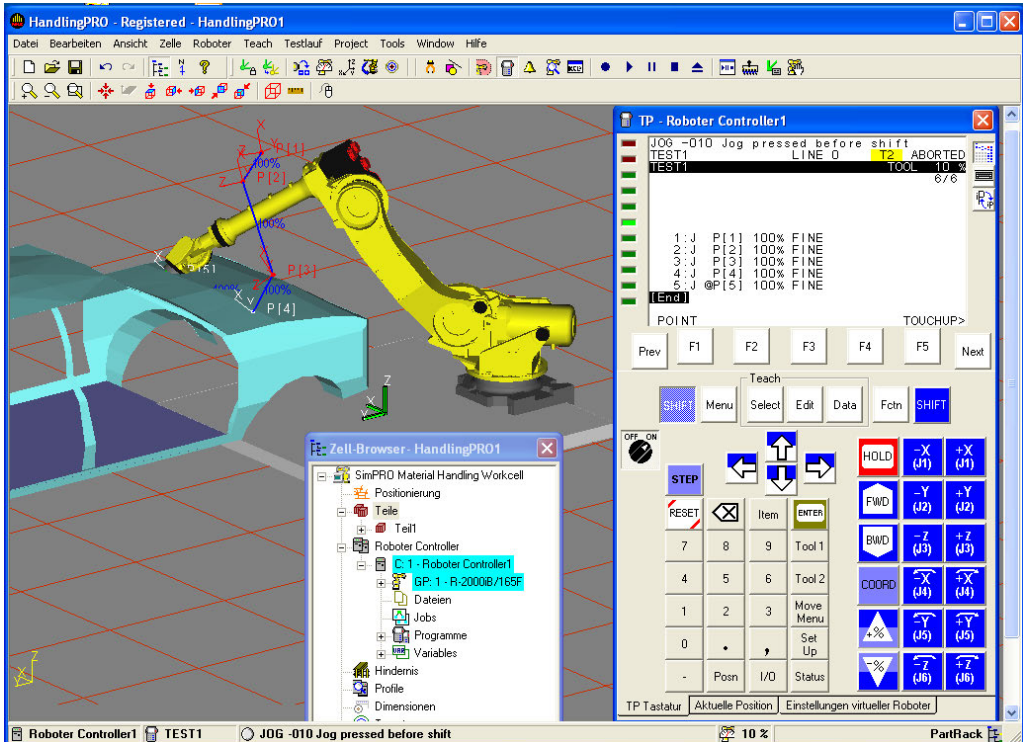


Bild 8-55: Roboterzelle mit eingblendetem Teachpanel und programmierter Bewegungsbahn

In der modellierten, virtuellen Umgebung kann anschließend der Roboter wie in der Realität mit einem virtuellen Teachpanel bewegt werden. In Bild 8-54 ist ein weiteres mit Roboguide erstelltes Modell abgebildet. Außerdem ist das virtuelle Teachpanel eingblendet, das in seinem Aussehen genau dem tatsächlichen Panel des Roboters entspricht. Mit dem Panel kann man die Achsen manuell in verschiedenen Koordinatensystemen bewegen und dabei Bahnstützpunkte entlang der zu programmierenden Bewegungsbahn aufnehmen. Die Punkte und die Bewegungsbahn werden ebenfalls im Bildschirm dargestellt. Danach kann man die Bewegungen in Echtzeit ablaufen lassen. Das System überprüft dabei, ob es zu Kollisionen mit der Peripherie kommt und ermittelt aufgrund der Geschwindigkeitsvorgabe die Zykluszeit. Damit hat man ein sehr aussagekräftiges Instrument in der Hand, eine Anwendung zu planen und die Bewegungsabläufe zu optimieren und zu visualisieren, ohne dass die Hardware bereits zur Verfügung steht.

9 Mechatronische Systeme

In diesem abschließenden Kapitel werden nun einige Systeme vorgestellt, die typischerweise in der Mechatronik behandelt werden. Anhand der Beispiele soll nochmals aufgezeigt werden, wie die Prinzipien und Methoden, die in den vorhergehenden Kapiteln erläutert wurden, zur Anwendung kommen und bei realen technischen Produkten zu verbesserten Problemlösungen führen.

9.1 Wann ist der Einsatz der Mechatronik sinnvoll?

Manchmal werden heute Produkte, die relativ einfach sind, zu mechatronischen Systemen „aufgerüstet“, um neue Verkaufsargumente für diese Produkte zu finden. Beispielsweise waren früher Staubsauger mit Motoren ausgerüstet, die nur über eine Saugstufe verfügten. Sollten spezielle Saugprobleme abgedeckt werden, so verwendete man spezielle Ansaugdüsen. Ich besitze heute einen Staubsauger mit stufenlos regelbarer Saugleistung und elektronischer Füllgradanzeige. Sogar Staubsauger mit Fuzzy-Reglern wurden schon angepriesen. Bei der Anwendung stelle ich fest, dass ich aus Bequemlichkeit immer die gleiche, nämlich die maximale Saugleistung verwende und dass die anderen Stufen von mir eigentlich nicht gebraucht werden. Eine wesentliche Erleichterung durch die elektronische Überwachung des Füllgrades des Filterbeutels kann ich ebenfalls nicht feststellen. Daher besteht für mich der wichtigste Vorteil gegenüber alten Staubsaugern in einer erhöhten Saugleistung des Motors. Hier bewirkt der hohe Einsatz mechatronischer Komponenten im wesentlichen Eigenschaften, die im amerikanischen Sprachgebrauch als „nice to have“ bezeichnet werden und die einen höheren Verkaufspreis rechtfertigen sollen.

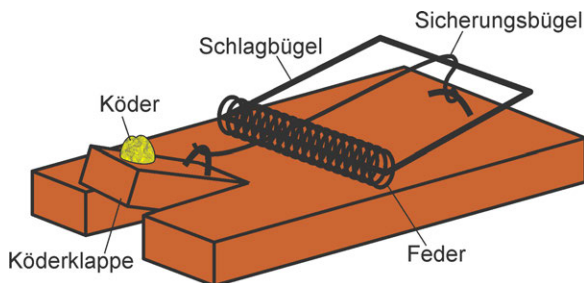


Bild 9-1: Mausefalle als einfaches mechanisches System mit Sensor (Köderklappe), Aktor (Feder und Schlagbügel) und Steuerung

Wie die Mechatronik sinnvoll und nutzbringend eingesetzt werden kann, soll an einem Beispiel verdeutlicht werden, das einfach ist und in einer satirisch überspitzten Art den Sachverhalt deutlich herausstellt. Ein solches einfaches, lange bekanntes System, das viele wesentliche Eigenschaften eines mechatronischen Systems aufweist, ist die in Bild 9-1 dargestellte Mausefalle. Sie verfügt über einen Sensor (bewegliche

Köderklappe), der eine bestimmte Belastung des beweglichen Teils der Mausefalle durch das Gewicht der Maus registriert. In einem mechanischen Speicher (Feder) wird Energie für den Antrieb des Aktors (Schlagbügel) zur Verfügung gestellt. Sensor und Aktor sind über eine mechanische Steuerung (Sicherungsbügel) miteinander verbunden, die ein Auslösen des Aktors bei Überschreiten eines Grenzwinkels des Sensors bewirkt. Mit diesem sehr einfachen mechanischen System werden seit vielen Jahrzehnten erfolgreich Mäuse „gefangen“, nämlich getötet. Es kann im Handel für 1 Euro gekauft werden.

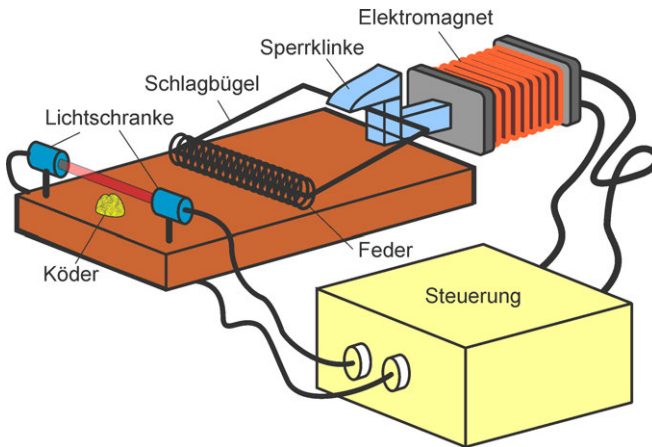


Bild 9-2: „Mechatronische“ Mausefalle mit Lichtschranke und elektromagnetischer Verriegelung

Nun kann man versucht sein, dieses in seiner Einfachheit gar nicht zu über-treffende mechanische System mechatronisch aufzu-rüsten. Man ersetzt den mechanischen Sensor durch eine Lichtschranke und arretiert den Schlagbügel über eine elektromagnetisch be-tätigte Sperrklinke (Bild 9-2). Beide Einrichtungen werden über eine elektronische Steuerung miteinander verbunden, die beim Unterbrechen der Licht-schranke durch eine Maus

die Sperrklinke entriegelt, so dass der Schlagbügel seine aus dem rein mechani-schen Modell bekannte Aufgabe durchführen kann. Die Nutzenfunktion der mecha-tronischen Mausefalle ist demnach genau die gleiche, nur mit dem Unterschied, dass nun die Mausefalle sicher nicht unter 30 Euro hergestellt werden kann, wäh-rend sie vorher 1 Euro kostete.

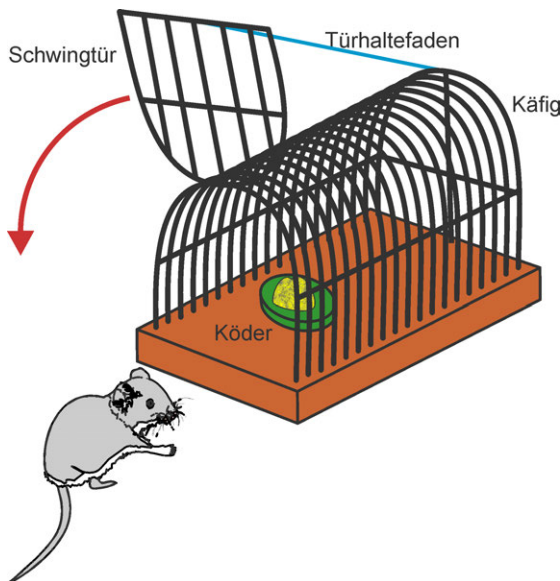


Bild 9-3: „Mechanische“ Mausefalle zum Lebendfangen von Mäusen

Kann man denn nun für die Aufgabe des Mäusefangens Mechatronik auch sinnvoll einsetzen? Dies wird mög-lich, wenn man von der Falle eine hö-herwertige Funktion erwartet, nämlich dass die Maus unbeschädigt, lebend gefangen wird. Auch für diese Auf-gabe gibt es natürlich rein mechanische Vorläufer, die diese Funktion erfüllen können (Bild 9-3). Sie sind schon et-was aufwendiger gebaut und fangen das Tier dadurch in einem Käfig, dass durch das Gewicht der Maus an der Köderstelle die Eingangsklappe (Schwingtür) geschlossen wird. Alle Komponenten funktionieren auch wieder einfach und mechanisch. Auch hier würde die mechatronische Aufrüstung (Lichtschranke, elektro-magnetische Verriegelung) nichts weiter bringen als einen höheren Herstellungspreis.

Erst wenn man nun noch fordert, dass die Falle hintereinander mehrere Mäuse fan-gen kann, ohne dass ein Mensch eingreift, würde das rein mechanische System versagen oder sehr aufwendig werden. Ausgestattet als mechatronisches System

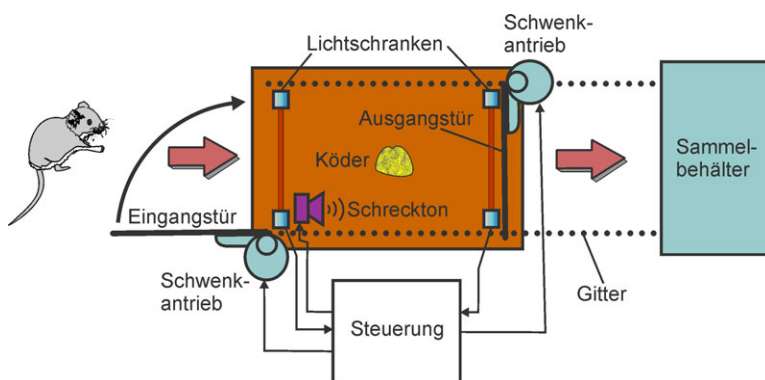


Bild 9-4: „Mechatronische Mausefalle“ für das Lebendfangen von Mäusen (Ansicht von oben)

Köder angelockt und tritt in die Falle ein, so wird eine Lichtschranke betätigt, deren Signal ein Schließen der Tür durch den Schwenkantrieb auslöst. Gleichzeitig öffnet sich die Ausgangstür, die zu einem Sammelbehälter führt. Ein elektrisch betätigter Schreckton sorgt dann dafür, dass die Maus die Falle durch die Ausgangstür verlässt. Danach wird wieder, kontrolliert durch eine weitere Lichtschranke, die Ausgangstür geschlossen und die Eingangstür geöffnet. Nun ist die Falle wieder im Grundzustand und kann weitere Mäuse lebend fangen.

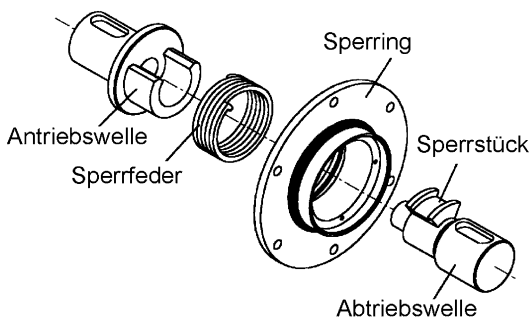


Bild 9-5: Einzelteile einer Last-Drehmomentsperre

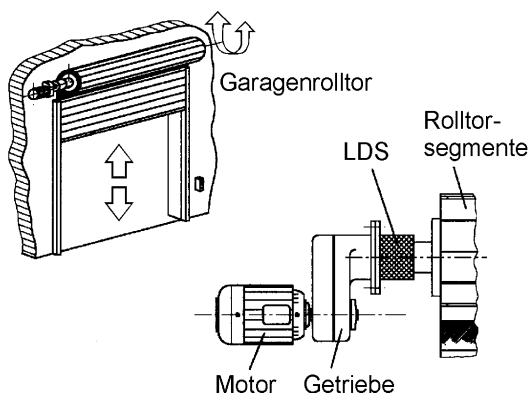


Bild 9-6: Einbau einer Last-Drehmomentsperre in ein Garagenrolltor

könnte man jetzt Vorder- und Rückseite der Falle mit einer Tür versehen, die durch einen Schwenkantrieb geschlossen oder geöffnet werden können (Bild 9-4). Zu Anfang ist die Eingangstür geöffnet und die Ausgangstür geschlossen. Wird eine Maus vom

Hier ist nun ein mechatronisches System entstanden, das durch eine einfache mechanische Lösung nicht ersetzt werden kann, da es zusätzlich geforderte Funktionen realisiert.

Um diese Gedanken nun wieder weg von der Satire in den Alltag der Automatisierungspraxis zu transportieren, möge ein weiteres einfaches Beispiel dienen. In Bild 9-5 ist eine so genannte **Last-Drehmomentsperre** (LDS) gezeigt. Dabei handelt es sich um eine rein mechanische Vorrichtung, die man beispielsweise in den Antriebsmechanismus eines Garagenrolltors (Bild 9-6) einbauen kann. Bei einem solchen durch einen Elektromotor zu öffnenden Rolltor ist es erwünscht, dass das Rolltor nach Einschalten des Motors in jeder beliebigen Zwischenstellung anhalten kann und dort in stabiler Lage verbleibt. Dies kann man beispielsweise so erreichen, dass man in den Antriebszweig des Rolltors eine elektrisch

betätigte Bremse einbaut, die bei Abschalten des Motors anzieht. Bei Wiederanlauf des Motors muss die Bremse dann wieder gelöst werden. Hierzu ist dann noch eine einfache Steuerung notwendig, die sicherstellt, dass die Betätigung von Motor und Bremse in der richtigen Reihenfolge erfolgt. So muss beim Anhalten zuerst der Motor und dann die Bremse geschaltet werden und beim Wiederanlauf zuerst die Bremse und dann der Motor.

Baut man eine LDS in den Kraftfluss des Motors ein, so wird bei dessen Abschalten das Rolltor sicher arretiert und kann beim Wiederanlauf ohne Belastung des Motors wieder gelöst werden. Dies geschieht dadurch, dass bei Wegfall des Antriebsmomentes die Sperrfeder (Bild 9-5) durch das Drehmoment der Last zusammengezogen wird und dadurch ein Weiterdrehen der Abtriebswelle durch die Last blockiert. Daher kann das Rolltor nicht zurücklaufen. Sobald über die Abtriebswelle vom Motor wieder ein Drehmoment in die LDS eingeleitet wird, entspannt sich die Sperrfeder, wodurch die Last über die Abtriebswelle erneut bewegt werden kann. Dieses einfache mechanische System leistet also das Gleiche wie die oben beschriebene elektromechanische Einrichtung, ist sicher preiswerter und auf die Dauer zuverlässiger und wartungsfrei.

Die angeführten Beispiele zeigen, dass nicht in jedem Fall sofort eine komplizierte, aus mehreren Technologien bestehende Lösung einer technischen Problemstellung anzustreben ist. Es sollte immer auch geprüft werden, ob einfache mechanische Systeme eingesetzt werden können, was wegen ihrer Wirtschaftlichkeit wünschenswert ist. Ein weiteres Beispiel dafür wurde schon im Kapitel 6 genannt, die Formgedächtnislegierung. Mit einer solchen FGL kann man Stellantriebe bauen, die gleichzeitig eine Temperatur überwachen und auf das Über- oder Unterschreiten einer Grenztemperatur mit einer Stellbewegung reagieren. Solche Einfachsysteme sollte man immer bei der Lösung mechatronischer Probleme im Auge behalten.

9.2 Entwicklung mechatronischer Systeme

In den vorausgegangenen Kapiteln sind die einzelnen erforderlichen Arbeitsschritte zur Entwicklung eines mechatronischen Systems und die mit den unterschiedlichen Teilsystemen verbundenen Problemstellungen behandelt worden. Nun soll nochmals in der Zusammenschau für das einfache Beispiel eines **invertierten Pendels** die Vorgehensweise bei der Entwicklung dargestellt werden.

In Bild 9-7 ist die Problemstellung in mehreren Varianten dargestellt. Teilbild 9-7 a) zeigt eine einfache technische Realisierung bei der ein gelenkig auf einem Wagen gelagerter Stab durch gezielte Vor- und Rückwärtsbewegungen in Richtung x so ausbalanciert werden soll, dass die Stabneigung um den Winkel α möglichst klein bleibt und dadurch der Stab nicht umklappen kann. Für den Stab gibt es also nur eine dynamische Gleichgewichtsbedingung, ein statisches System wäre nicht stabil.

Diese Problemstellung kennt jeder, der schon einmal versucht hat einen Stab auf der Fingerspitze stehend auszubalancieren (Bild 9-7 b)). Sie ist aber nicht nur für Jongleure relevant, sondern tritt auch als ganz reale technische Problemstellung auf. Wie in Bild 9-7 c) dargestellt, ist die Problemstellung beim Start einer Rakete

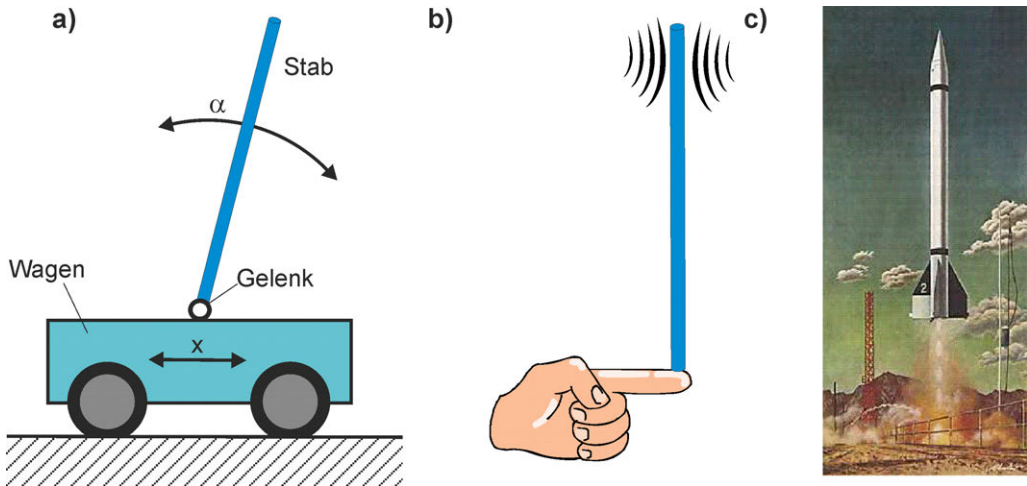


Bild 9-7: Problemstellung des inversen Pendels a) einfache technische Ausführung b) Balancieren eines Stabes auf der Hand c) Balancieren einer startenden Rakete auf dem Antriebsstrahl

von Bedeutung, wo durch Steuerbewegungen von Strahlrudern der schlanke Raketenkörper auf dem aus dem Heck austretenden Antriebsstrahl ausbalanciert werden muss, um ein Umkippen zu vermeiden. Seit einiger Zeit gibt es auch bereits ein als **Segway**® bekanntes zweirädriges Fahrzeug (Bild 9-8), auf dem eine Person in senkrechter Haltung durch Verlagerung des Körpergewichtes dieses steuern kann. Das Fahrzeug muss, da es ebenfalls ein invertiertes Pendel darstellt, durch eine entsprechende Regelung in der aufrechten Lage gehalten werden.

Es soll nun der Fall aus Bild 9-7 a) behandelt werden, wo der Stab gelenkig auf einem Wagen angebracht ist und durch geregelte Bewegungen des Wagens in $\pm x$ -Richtung im labilen Gleichgewicht gehalten werden soll. Bevor man sich Gedanken über die Ausstattung dieses mechatronischen Systems mit Sensoren, Aktoren und Steuerungen macht, muss man hier wie bei den meisten Problemstellungen versuchen, ein mathematisches Modell des Systems zu erstellen, um es hinterher simulieren und dabei optimal auslegen zu können. Bei diesem System ist es ohnehin sofort einsehbar, dass der Versuch, einen Hardwareprototypen zu erstellen und durch Probieren zu einer Lösung zu kommen, von vornherein aussichtslos ist.



Bild 9-8: Segway-Roller zur Personenbeförderung

In Bild 9-9 ist die Konfiguration aus Bild 9-7 a) nochmals mit allen Einflussgrößen und den

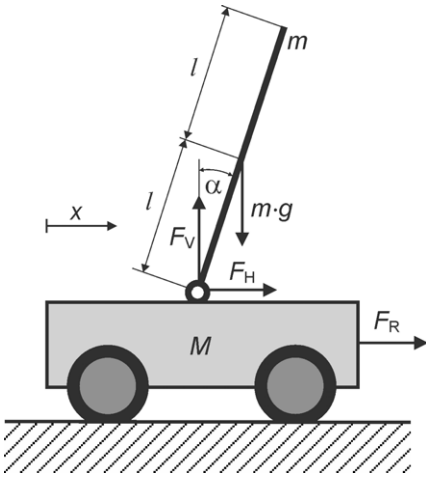


Bild 9-9: Invertiertes Pendel mit angreifenden Kräften

für das dynamische Problem relevanten Kräften dargestellt. Auf den Wagen der Masse M soll die Regelkraft F_R einwirken, mit der dem Wagen eine Gegenbewegung beim Kippen des Pendels aufgeprägt werden soll. Seine Lage wird durch die Wegkoordinate x beschrieben. Der Stab der Masse m und der Länge $2l$ erzeugt im Gelenk am Wagen die beiden Auflagerkräfte F_H und F_V . In seinem Schwerpunkt greift die Gewichtskraft $F_G = m \cdot g$ an, die den Stab um den Winkel α auslenkt. Um die Bewegungs- Differentialgleichung des Systems aufzustellen, werden zuerst einfache bekannte Beziehungen über die angreifenden Kräfte aufgestellt, und zwar aufgeteilt nach den Teilsystemen Wagen und Stab:

$$F_R - F_H = M \frac{d^2 x}{dt^2} \quad \text{Kräfte am Wagen} \quad (9.1)$$

Für die Kräfte am Stab benötigt man die Koordinaten des Stabschwerpunktes:

$$x + l \cdot \sin \alpha \quad \text{horizontale Richtung} \quad l \cdot \cos \alpha \quad \text{vertikale Richtung}$$

Damit ergeben sich die Kräfte am Stab für die beiden Richtungen zu:

$$F_H = m \frac{d^2}{dt^2} (x + l \cdot \sin \alpha) \quad (9.2)$$

$$F_V - m \cdot g = m \frac{d^2}{dt^2} (l \cdot \cos \alpha) \quad (9.3)$$

Da es sich bei der Bewegung des Stabes um eine Rotationsbewegung handelt, kann man auch noch die Momentensumme um den Stabschwerpunkt als Bewegungsgleichung verwenden:

$$F_V (l \cdot \sin \alpha) - F_H (l \cdot \cos \alpha) = J \frac{d^2 \alpha}{dt^2} \quad (9.4)$$

Das Massenträgheitsmoment des Stabes für eine Drehung um den Schwerpunkt beträgt:

$$J = \frac{m \cdot l^2}{3}$$

Man hat nun 4 Bestimmungsgleichungen für die beiden Wegkoordinaten x und α . Nach Bildung der Ableitungen in den Gleichungen (9.1 - 9.4) erhält man folgende Formen Gl. (9.5 - 9.8) :

$$F_R - F_H = M \cdot \ddot{x} \quad (9.5)$$

$$\begin{aligned}
 F_H &= m \cdot \ddot{x} + m \cdot l \cdot \frac{d}{dt}(\cos \alpha \cdot \dot{\alpha}) \\
 &= m \cdot \ddot{x} + m \cdot l \cdot (-\sin \alpha \cdot \dot{\alpha}^2 + \cos \alpha \cdot \ddot{\alpha}) \\
 &= m \cdot \ddot{x} + m \cdot l \cdot (\ddot{\alpha} \cdot \cos \alpha - \dot{\alpha}^2 \cdot \sin \alpha)
 \end{aligned} \tag{9.6}$$

$$\begin{aligned}
 F_V - m \cdot g &= m \cdot l \cdot \frac{d}{dt}(-\sin \alpha \cdot \dot{\alpha}) \\
 &= m \cdot l (-\dot{\alpha}^2 \cdot \cos \alpha - \ddot{\alpha} \cdot \sin \alpha)
 \end{aligned} \tag{9.7}$$

$$J \cdot \ddot{\alpha} = F_V \cdot l \cdot \sin \alpha - F_H \cdot l \cdot \cos \alpha \tag{9.8}$$

In den Gl. (9.6 - 9.8) tritt die Wegkoordinate α in nichtlinearen Funktionen auf, weshalb diese Gleichungen sehr unübersichtlich sind und die Berechnung schwer fällt. Wie aber bereits beim mathematischen Pendel durchgeführt, lassen sich diese Nichtlinearitäten linearisieren, weil insbesondere für das Ausbalancieren des Pendels gilt, dass der Auslenkungswinkel α stets sehr klein sein muss, um ein Kippen zu verhindern. Für kleine Winkel α gilt:

$$\sin \alpha = \alpha - \frac{\alpha^3}{6} + \dots \approx \alpha \quad \cos \alpha = 1 - \frac{\alpha^2}{2} + \dots \approx 1$$

Dadurch vereinfachen sich die Gl. (9.6 - 9.8) zu den Gl. (9.9 - 9.11):

$$F_H \approx m \cdot \ddot{x} + m \cdot l (\ddot{\alpha} - \dot{\alpha}^2 \cdot \alpha) \tag{9.9}$$

$$F_V - m \cdot g \approx -m \cdot l (\dot{\alpha}^2 + \ddot{\alpha} \cdot \alpha) \tag{9.10}$$

$$J \cdot \ddot{\alpha} \approx F_V \cdot l \cdot \alpha - F_H \cdot l \tag{9.11}$$

Da für kleine Winkel α Glieder höherer Ordnung ($\dot{\alpha}^2, \ddot{\alpha} \cdot \alpha$) sehr klein werden, können diese vernachlässigt werden, was zu dem endgültigen, linearisierten Gleichungssystem Gl. (9.12 - 9.15) führt:

$$F_R - F_H = M \cdot \ddot{x} \tag{9.12}$$

$$F_H \approx m \cdot \ddot{x} + m \cdot l \cdot \ddot{\alpha} \tag{9.13}$$

$$F_V - m \cdot g \approx 0 \tag{9.14}$$

$$J \cdot \ddot{\alpha} \approx F_V \cdot l \cdot \alpha - F_H \cdot l \tag{9.15}$$

Aus diesem Gleichungssystem lassen sich die für die Regelung unwichtigen Reaktionskräfte im Lager durch Substitution eliminieren, was auf die beiden folgenden Bewegungsgleichungen führt:

$$m \cdot \ddot{x} + m \cdot l \cdot \ddot{\alpha} = F_R - M \cdot \ddot{x} \Rightarrow m \cdot l \cdot \ddot{\alpha} + (m + M) \ddot{x} = F_R \tag{9.16}$$

$$J \cdot \ddot{\alpha} = m \cdot g \cdot l \cdot \alpha - m \cdot l \cdot \ddot{x} \Rightarrow (J + ml^2) \ddot{\alpha} + m \cdot l \cdot \ddot{x} - m \cdot g \cdot l \cdot \alpha = 0 \tag{9.17}$$

Aus den Gl. (9.16) und (9.17) findet man durch Auflösung nach den höchsten Ableitungen und durch Verwendung folgender Abkürzung

$$A = J(m + M) + m \cdot M \cdot l^2,$$

das folgende Gleichungssystem:

$$\begin{aligned}\ddot{\alpha} &= \left[\frac{g(m+M)m \cdot l}{A} \right] \cdot \alpha - \left(\frac{m \cdot l}{A} \right) \cdot F_R \\ \ddot{x} &= \left(\frac{-gm^2 l^2}{A} \right) \cdot \alpha + \left(\frac{J + m \cdot l^2}{A} \right) \cdot F_R\end{aligned}\quad (9.18)$$

Führt man für die konstanten Koeffizienten weitere Abkürzungen in der Form

$$a = \left[\frac{g(m+M)m \cdot l}{A} \right] \quad b = -\left(\frac{m \cdot l}{A} \right) \quad c = \left(\frac{-gm^2 l^2}{A} \right) \quad d = \left[\frac{(J + m \cdot l^2)}{A} \right]$$

ein, erhält man die Systemgleichungen in sehr übersichtlicher Form:

$$\begin{aligned}\ddot{\alpha} &= a \cdot \alpha + b \cdot F_R \\ \ddot{x} &= c \cdot \alpha + d \cdot F_R\end{aligned} \quad \text{oder} \quad \begin{pmatrix} \ddot{\alpha} \\ \ddot{x} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ F_R \end{pmatrix} \quad (9.19)$$

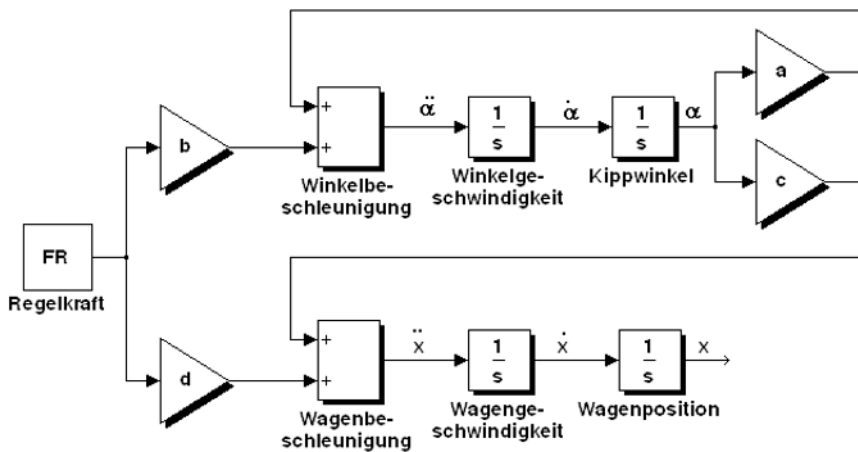


Bild 9-10: Blockschaltbild der linearisierten Zustandsgleichungen des invertierten Pendels

Die Gl. (9.19) kann man auch als Blockschaltbild darstellen (Bild 9-10). Dies wäre bereits das Diagramm, mit dem man eine Simulation des Systems mit einem Blockschaltbildeditor wie Simulink durchführen könnte. Wie man jedoch an dem Bild sieht, ist es zum jetzigen Zeitpunkt unklar, wie der Regler die erforderlicher Regelkraft F_R aufbringen soll, um ein Kippen des Pendels zu vermeiden, d. h. den Kippwinkel α möglichst klein zu halten. Der Winkel ist deshalb eine wichtige Zustandsgröße des Systems, die auf jeden Fall in die Regelkraft eingreifen muss. Wie man sieht, wird die Wagenposition x aus dem Kippwinkel durch zweifache Integration abgeleitet und stellt ebenfalls eine wichtige Zustandsgröße des Systems dar, die in die Regelung eingreifen sollte.

Aufgrund der Ermittlung von α und x durch Integration von $\dot{\alpha}$ und $\dot{x}$ stehen auch die

Zustandsgrößen Winkelgeschwindigkeit und Wagengeschwindigkeit zur Verfügung. Wie wir bereits bei der Kaskadenregelung in einem Lageregelkreis gesehen haben (Kap. 7.3), führt die Hinzunahme zusätzlicher Zustandsgrößen wie der Geschwindigkeit zu einer deutlichen Verbesserung der Regeldynamik. Da das Gleichungssystem Gl. (9.19) linear ist, sollte eine Linearkombination der vier genannten Zustandsgrößen zur Erzeugung der Regelkraft eine Stabilisierung des Pendels ermöglichen. Sie lautet:

$$F_R = k_\alpha \cdot \alpha + k_{\dot{\alpha}} \cdot \dot{\alpha} + k_x \cdot x + k_{\dot{x}} \cdot \dot{x} \quad (9.20)$$

wobei $k_\alpha, k_{\dot{\alpha}}, k_x, k_{\dot{x}}$ noch zu bestimmende Konstanten sind. Eine solche Regelfunktion wird als **Zustandsregelung** bezeichnet, die hier nicht vertieft behandelt werden kann. Näheres hierzu findet sich in [6.2].

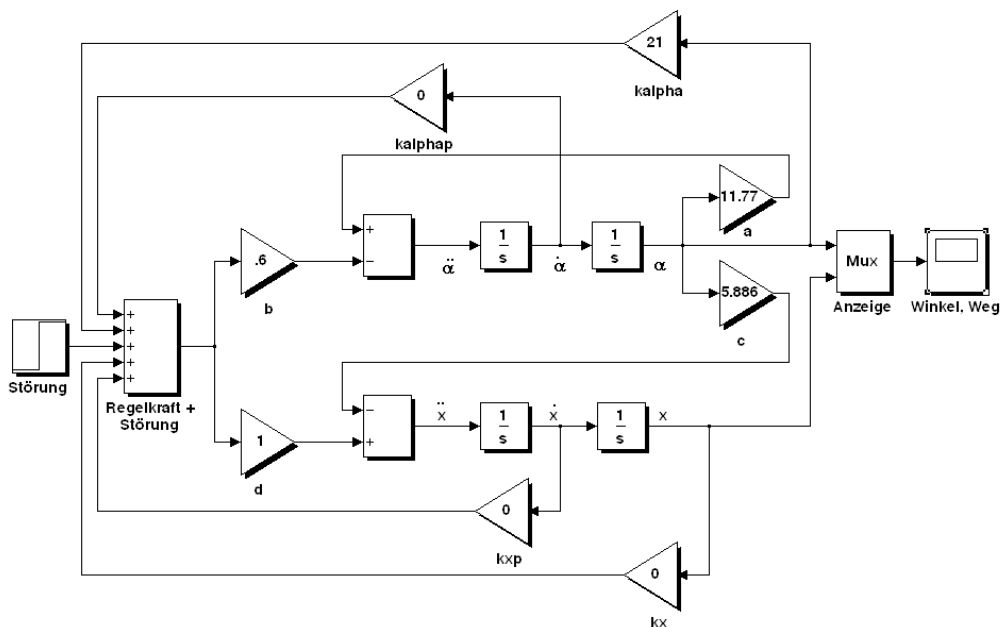


Bild 9-11: Blockschaltbild der Simulation des invertierten Pendels mit Zustandsregelung

Das vollständige Simulationsblockschaltbild mit der Erzeugung der Reglerfunktion ist in Bild 9-11 dargestellt. Zusätzlich kann auf die Regelfunktion eine sprungförmige Störung aufgeschaltet werden. In der Simulation kann man nun unterschiedliche Regelfunktionen durch Variation der Konstanten in der Gl. (9.20) testen und so ein optimales Reglerverhalten ermitteln. Man kann beispielsweise damit beginnen, alle Regelparameter außer k_α gleich Null zu setzen, was einem Proportionalregler für den Kippwinkel entspricht. Bei kleinen Werten von k_α kippt das Pendel nach Aufbringen der Störung sofort um, da der Regeleingriff nicht ausreicht. Ab einem bestimmten Wert ist dann der P-Regler in der Lage das Pendel zu stabilisieren. Bild 9-12 zeigt den Verlauf von Wagenweg und Kippwinkel bei einem Wert von $k_\alpha = 21$. Die Verläufe zeigen, dass nach dem Störungssprung zum Zeitpunkt $t = 1\text{s}$ der Wagen versucht die Störung auszugleichen, entfernt sich dabei aber immer weiter aus

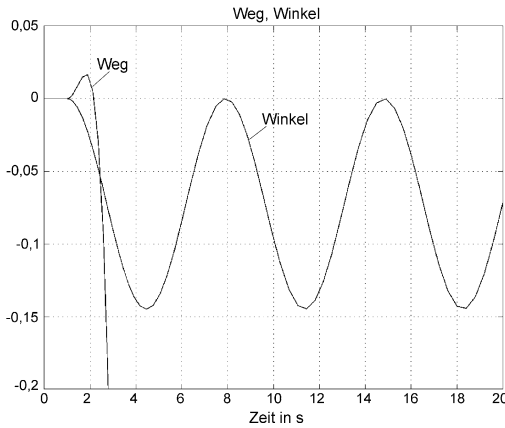


Bild 9-12: Verlauf von Weg und Winkel bei P-Regelung des Winkels ($k_\alpha = 21$)

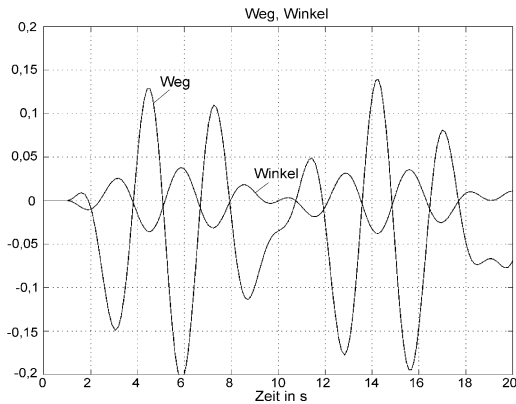


Bild 9-13: Verlauf von Weg und Winkel bei P-Regelung von Weg und Winkel ($k_\alpha = 42$, $k_x = 3$)

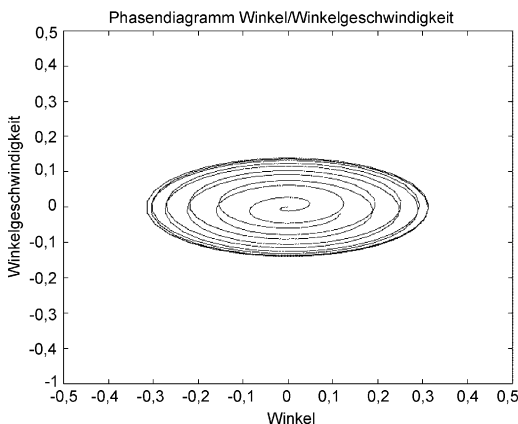


Bild 9-14: Phasendiagramm des invertierten Pendels bei P-Regelung, $k_\alpha = 42$, $k_x = 3,175$

seiner Ausgangslage. Das Pendel verbleibt nicht in seiner Ruhelage, sondern führt eine Dauerschwingung um einen von Null verschiedenen Winkel aus.

Die nächste Möglichkeit der Reglervariation müsste daher die Hinzunahme der Regelung des Wagenweges durch $k_x \neq 0$ sein. Bild 9-13 zeigt das Ergebnis für $k_\alpha = 42$, $k_x = 3$. Nun versucht der Regler sowohl das Pendel zu stabilisieren, als auch den Wagen wieder in die Ausgangsposition zu bringen. Dies gelingt jedoch auch nur für beide Systemkomponenten in Form einer Schwingung. Dabei befindet sich das System bereits an der Stabilitätsgrenze. So zeigt das Bild 9-14 das **Phasendiagramm** von Kippwinkel und Winkelgeschwindigkeit bei $k_\alpha = 42$, $k_x = 3,175$. Wie man sieht, führt das Pendel eine Dauerschwingung aus, die aber einen bestimmten Grenzwinkel nicht überschreitet. Bei den Reglereinstellungen $k_\alpha = 42$, $k_x = 3,2$ in Bild 9-15, zeigt das Phasendiagramm eine dauerhaft aufklingende Schwingung an. Wenn hierbei der Kippwinkel große Werte annimmt, gilt natürlich das linearisierte Modell des invertierten Pendels nur noch sehr eingeschränkt.

Zieht man nun die Ableitungen von Winkel und Weg nämlich die Winkelgeschwindigkeit und die Wagengeschwindigkeit in die Regelung mit ein, indem man $k_{\dot{\alpha}}, k_{\dot{x}} \neq 0$ wählt, so erhält man für $k_\alpha = 100$, $k_x = 3$, $k_{\dot{\alpha}} = 30$, $k_{\dot{x}} = 10$ das in Bild 9-16 dargestellte Regelverhalten. Es liegt nun wegen der Verwendung der Ableitungen ein PD-Regler vor, der schon reagiert, wenn der Kippwinkel noch sehr klein ist, die Winkelgeschwindigkeit aber schon messbare Werte erreicht hat. Dadurch bleibt der Stab sehr stabil in seiner Ausgangslage und zeigt nur minimale Veränderungen.

gen beim Auftreten einer sprungförmigen Störung. Der Wagen jedoch kann nicht in seine Ausgangslage zurückkehren, da ein PD-Regler eine bleibende Regelabweichung besitzt.

Wie bereits in Abschnitt 7.3.1.7 über Regler festgestellt wurde, lässt sich eine solche bleibende Regelabweichung nur vermeiden, wenn der Regler eine integrale Komponente besitzt. In der Simulation kann man dies leicht erreichen, indem man den Wagenweg nochmals integriert und dessen Ausgang, mit einem Koppelkoeffizient bewertet, ebenfalls zur Regelfunktion hinzufügt. Das Ergebnis ist in Bild 9-

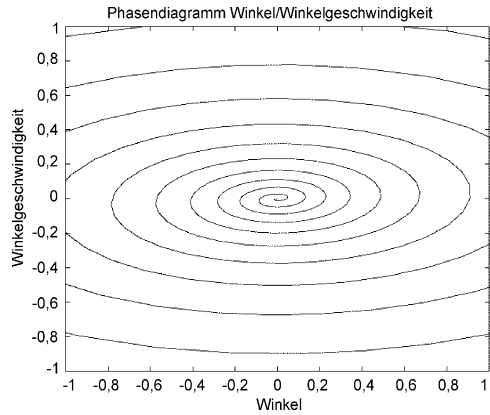


Bild 9-15: Phasendiagramm des invertierten Pendels bei P-Regelung, $k_\alpha = 42$, $k_x = 3,2$

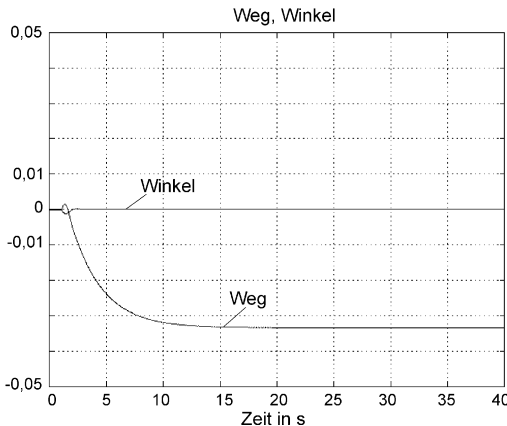


Bild 9-16: Verlauf von Weg und Winkel bei PD-Regelung von Weg und Winkel ($k_\alpha = 100$, $k_x = 3$, $k_{\dot{\alpha}} = 30$, $k_{\dot{x}} = 10$)

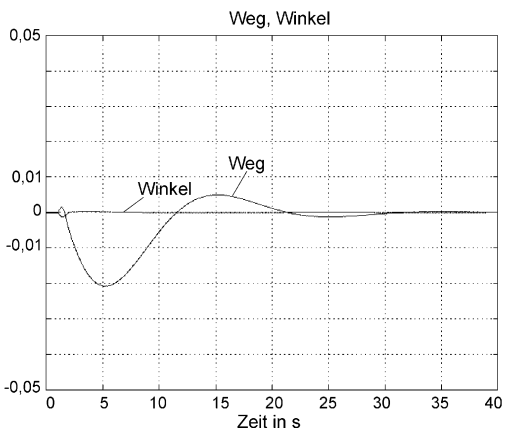


Bild 9-17: Verlauf von Weg und Winkel bei PID-Regelung des Weges ($k_\alpha = 100$, $k_x = 3$, $k_{\dot{\alpha}} = 30$, $k_{\dot{x}} = 10$)

17 zu sehen, in dem man sieht, dass der Wagen nach einer sprungförmigen Störung durch den integralen Anteil (I-Anteil) nach einem Einschwingvorgang wieder in die Ausgangsposition zurückkehrt. Durch Variation des I-Anteils kann man nun noch die erforderliche Regelleistung P_R optimieren, die das Produkt aus Regelkraft und Fahrgeschwindigkeit darstellt. In Bild 9-18 und 9-19 ist der Unterschied bei bestimmtem I-Anteil und ansonsten gleichen Regelfaktoren aber unterschiedlichem Winkel-Rückkopplungsfaktor k_α dargestellt. In Bild 9-18 ist $k_\alpha = 100$, wobei die Auswirkung der Störung nach 20 Sekunden zwar ganz abgeklungen ist, aber die maximale Wagengeschwindigkeit und damit die maximale Regelleistung hohe Werte annimmt. Im Bild ist zusätzlich das Integral über das Quadrat der Regelleistung dargestellt, als Maß für die aufzubringende Gesamtleistung. Verglichen mit Bild 9-19 sieht man, dass für die Gesamtleistung zur Ausregelung der Störung deutlich geringer ist, der Regelvorgang aber länger als 20 Sekunden andauert. Hier muss der

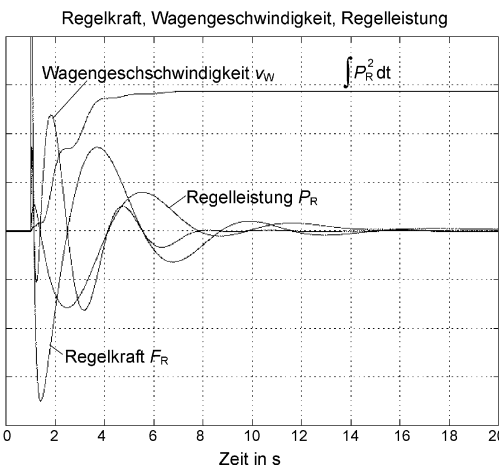


Bild 9-18: Verlauf von F_R , v_W , P_R , $\int P_R^2 dt$ bei $k_\alpha = 100$

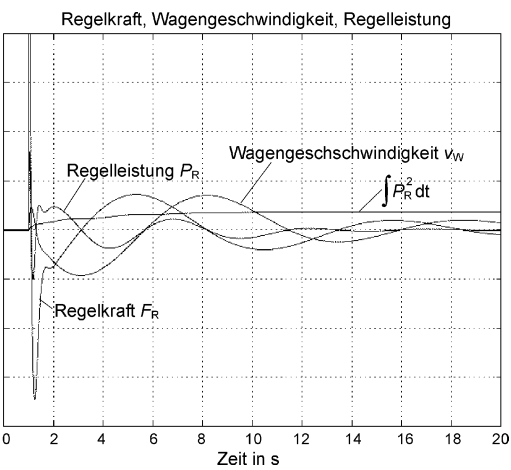


Bild 9-19: Verlauf von F_R , v_W , P_R , $\int P_R^2 dt$ bei $k_\alpha = 200$

Entwickler entscheiden, welche der Regelvorgänge für sein Projekt das Optimum darstellt. So wäre es für ein Fahrzeug mit Batterieversorgung möglicherweise besser die Regelleistung zu minimieren, bei ausreichend vorhandener Antriebsenergie wäre eher ein möglichst schnell stabilisierendes Regelverhalten wünschenswert.

		Benötigte Werte für Regelung				
		α	$\dot{\alpha}$	x	$\dot{x}$	$\int x \cdot dt$
Messbare Zustandsgrößen	α	direkte Messung	$(\alpha)'$	-	-	-
	$\dot{\alpha}$	-	direkte Messung	-	-	-
	x	-	-	direkte Messung	$(x)'$	$\int x \cdot dt$
	$\dot{x}$	-	-	-	direkte Messung	-
	$\ddot{\alpha}$	-	$\int \ddot{\alpha} \cdot dt$	-	-	-
	$\ddot{x}$	-	-	-	$\int \ddot{x} \cdot dt$	-

Bild 9-20: Mögliche Ermittlung der benötigten Zustandsgrößen aus messbaren Zustandsgrößen

Nach Ermittlung einer optimalen Reglerfunktion kann man nun an die Realisierung eines Prototyps des invertierten Pendels gehen. Außer der Hardware des Wagens und des Pendels benötigt man einen Aktor für die Wagenbewegung. Als Kernstück ist ein Mikrorechner erforderlich für die Aufnahme von Sensorsignalen, die Erzeugung der Reglerfunktion und die Ausgabe des Stellsignals an den Wagen-

antrieb. Da nun für die Regelung Informationen über den Kippwinkel, die Winkelgeschwindigkeit des Pendels, die Wagenposition und die Wangengeschwindigkeit benötigt werden, braucht man prinzipiell zuerst einmal Sensoren, die diese vier Zustandsgrößen messen. In Bild 9-20 sind den für die Regelung benötigten Zustandsgrößen die Möglichkeiten gegenübergestellt, diese aus messbaren Größen entweder durch direkte Messung oder durch Berechnungen wie Ableiten oder Integrieren zu erzeugen. Da die beiden Zustandsgrößen "Geschwindigkeit" die Ableitungen der "Wege" sind, können diese theoretisch auch durch Differenzieren der Wegsignale erzeugt werden. Ob dies sinnvoll ist, hängt von der Rechenleistung und der daraus resultierenden Rechenzeit des Mikrorechners für den Rechengang ab. Dieser muss so schnell sein, dass die daraus entstehende Zeitverzögerung zwi-

schen einer Wegänderung und der dann berechneten Geschwindigkeit kleiner als die entsprechenden Zeitkonstanten der Regelung sind. Kann dies durch den Mikrorechner gewährleistet werden, so kommt man mit nur zwei Sensoren für den Kippwinkel und die Wagenposition aus. Dies könnten beispielsweise inkrementale Drehgeber sein, die direkt oder über ein Untersetzungsgetriebe an die Antriebsachse des Wagens und an das Gelenk des Stabes angebracht werden. Sind die Genauigkeitsansprüche an die Wegmessung nicht so hoch, so könnten auch einfache elektrische Drehpotentiometer mit und ohne Getriebeuntersetzung ausreichen. Auch die Ermittlung der "Geschwindigkeiten" durch Messung der "Beschleunigungen" beispielsweise mit sehr klein bauenden piezoelektrischen Beschleunigungssensoren und anschließende Integration im Mikrorechner wäre unter den oben genannten Einschränkungen eine Möglichkeit. Bei Messung aller Zustandsgrößen würde sich für die Wangengeschwindigkeit ein Tachogenerator anbieten. Ein weiterer Weg, ohne die Rechenkapazität des Mikrorechners auszunutzen, die inkrementellen Weginformationen in Geschwindigkeitsinformationen umzusetzen, ist die Verwendung eines separaten Messwertwandlers in Form eines Frequenz-/ Spannungswandlers.

Als Aktor würde sich für den Wagenantrieb ein Gleichstrommotor mit Getriebe oder ein Schrittmotor anbieten. Beide könnten ebenso wie der Mikrorechner und die Sensoren aus einer Batterie auf dem Wagen gespeist werden. Für die Auslegung des Motors und des Getriebes ist zu beachten, dass dieser die Beschleunigung zur Erzeugung der erforderlichen Regelkraft auch aufbringen kann.

Die Auswahl des Mikrorechners, der in diesem Fall sicher ein Einchip-Mikrocontroller wäre, muss sich vor allem an den erforderlichen Signaleingängen für die gewählten Sensoren und den sich daraus ergebenden Rechenleistungen richten. Der Regelalgorithmus wird als digitale Regelung ebenfalls auf dem Mikrocontroller ausgeführt. Bild 9-21

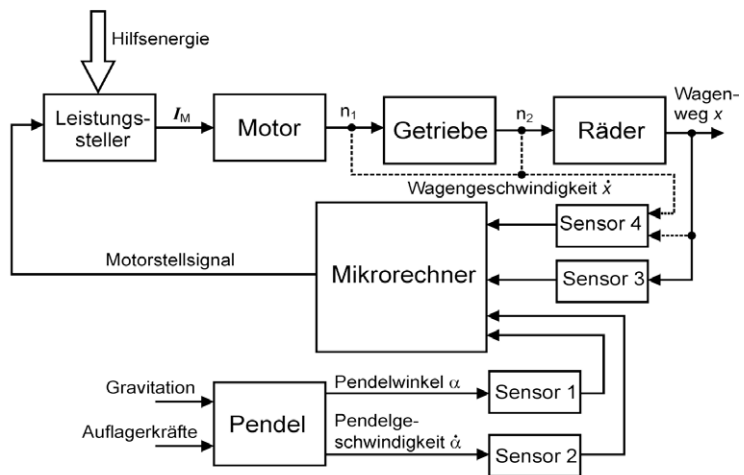


Bild 9-21: Blockschaftbild des Antriebssystems des Wagens

zeigt ein Blockschaftbild des gesamten Antriebssystems einschließlich Mikrorechner und Sensorik.

9.3 Mechatronische Teilsysteme

In vielen Fällen sind in der Vergangenheit Teilfunktionen größerer Systeme, die bis dahin mit rein mechanischen oder elektromechanischen Mitteln realisiert wurden, durch mechatronische Teilsysteme ersetzt worden. Dies geschieht häufig aus dem

Grund, dass die technischen Anforderungen an das Teilsystem zu einem bestimmten Zeitpunkt gestiegen sind. Die gesteigerten Anforderungen konnten aber nicht durch Verbesserung der vorhandenen mechanischen oder elektromechanischen Komponenten gelöst werden. Die Fahrzeugtechnik ist hier ein Bereich, in dem immer mehr Funktionen, die früher ausschließlich durch den Fahrer kontrolliert wurden, durch mechatronische Teilsysteme durchgeführt oder unterstützt werden. Beispiele hierfür sind das **ABS (Anti-Blockier-System)** und das **ESP (Elektronisches Stabilitäts-Programm)**

Anhand einiger Beispiele solcher Teilfunktionen von größeren Systemen, die in den verschiedensten Technologiebereichen durch mechatronische Ansätze gelöst werden können, soll diese Vorgehensweise erläutert werden.

9.3.1 Magnetlager

Diese Art der Lagerung ist in der Vergangenheit durch den Wunsch entstanden, Werkzeugmaschinenspindeln mit immer höheren Drehzahlen laufen zu lassen. Dabei gerät man in einen Drehzahlbereich, bei dem konventionelle Lagerungen wie hydrodynamische Gleitlager oder Kugellager versagen. Die Anforderung hoher Drehzahlen ist deshalb entstanden, weil man festgestellt hat, dass bei spanabhebenden Verfahren wie etwa dem Fräsen, eine drastische Steigerung der Schnittgeschwindigkeit eine Steigerung der Abtragleistung und der Wirtschaftlichkeit zur Folge hat. Das Zeitspannungsvolumen, die Oberflächenqualität und der Standweg der Werkzeuge steigen bei gleichzeitiger Reduktion der Zerspanungskräfte. Die Schnittgeschwindigkeit ist in erster Näherung die Relativgeschwindigkeit zwischen Schneide und Werkstück, d. h. beim Fräsen die Umfangsgeschwindigkeit des Fräasers. Bei den geforderten hohen Schnittgeschwindigkeiten wird die Wärme im wesentlichen über den Span abgeführt, wodurch das Werkstück kalt bleibt. Da gleichzeitig die Zerspankräfte klein sind, können mit so betriebenen Frässpindeln sehr dünnwandige Strukturen bearbeitet werden.

Bei Drehzahlen über 20000 Umdrehungen pro Minute wird daher eine magnetisch arbeitende, berührungsfreie Lagerung einer solchen Frässpindel interessant, bei

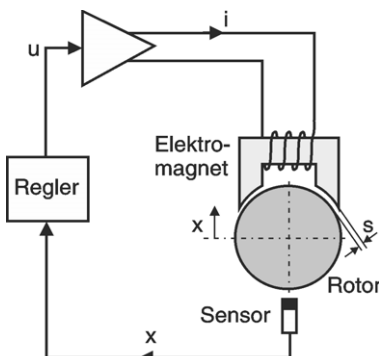


Bild 9-22: Funktionsprinzip einer aktiven elektromagnetischen Lagerung

der die Lagerkräfte durch Magnetkräfte realisiert werden. Dabei wird zwischen zwei grundsätzlich verschiedenartigen Lagerungen unterschieden [9.1]. Die eine ist die **passive magnetische Lagerung**, bei der die abstoßenden Kräfte von Permanentmagneten ausgenutzt werden. Eine Regelung der Lagerungseigenschaften ist nicht möglich.

Die zweite Lagerungsart ist die **aktive elektromagnetische Lagerung**, die unter der Bezeichnung **Magnetlager** oder im englischen Sprachbereich **Active Magnetic Bearing** kurz **AMB** bekannt ist. Das Grundprinzip einer solchen Lagerung ist in Bild 9-22 dargestellt [9.1]. Ein Sensor misst die Abweichung des Rotors von seiner Soll-Lage. Aus dem Messwert erzeugt der Regler ein Stellsignal,

das über einen Leistungsverstärker den Strom des Elektromagneten und damit die Magnetkräfte so steuert, dass der Rotor stabil schwebt. Elektromagnet und Leistungsverstärker bilden demnach in einem solchen System den Aktor. Das gesamte Wirkprinzip ist ähnlich wie bei der in Kapitel 7.3 besprochenen magnetischen Wägung.

Als Sensoren kann man Wirbelstromsensoren oder induktive Sensoren einsetzen. Es besteht prinzipiell auch die Möglichkeit aus Strom und Spannung in der Magnetwicklung, sowie aus der Flussdichte im Luftspalt die Rotorposition direkt zu bestimmen.

Die Regelung kann digital oder analog realisiert werden, wobei moderne digitale Regelungen die größere Flexibilität aufweisen. Für eine ausreichende Dynamik sind dabei Abtastfrequenzen über 10 kHz erforderlich. Da eine komplette Lagerung einer Welle in der Regel mehrere Lagerstellen benötigt, tritt diese Anordnung mehrfach auf. Für solche Komplettlagerungen gibt es Beispiele mit einem Gesamtregler oder auch solche mit einem Teilregler für jede Lagerstelle. Da erhebliche elektrische Leistungen im Verstärker umgesetzt werden, werden wegen der wesentlich kleineren Verluste vorwiegend gepochte Verstärker eingesetzt.

Für eine konkrete Lagerung werden als Lagermagnete üblicherweise paarweise gegenüberliegende Magnete eingesetzt (Bild 9-23). Ein komplettes Radiallager besteht dann aus zwei solcher Magnetpaare für die x -Richtung und für die y -Richtung. Als Axiallager verwendet man ein Paar Topfmagneten. Die Minimallagerung einer Welle umfasst zwei Radial- und ein Axiallager. Entsprechend müssen bei einer vollständigen, elektromagnetischen Lagerung die fünf Bewegungsrichtungen (x_1, y_1, x_2, y_2, z) geregelt werden.

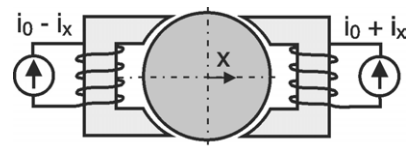


Bild 9-23: Paarweise Anordnung der Lagermagnete für eine Lagereinrichtung

Wie in Bild 9-23 dargestellt, werden in der paarweisen Anordnung die Lagermagnete mit einem konstanten Vormagnetisierungsstrom $i_0 = i_{\max}/2$ betrieben, dem der Regelstrom i_x differenziell überlagert wird. Dies ergibt eine sehr gute Linearisierung der Strom-Kraft-Kennlinie.

In Bild 9-24 ist eine komplette AMB-Lagerung dargestellt, die aus 10 Elektromagneten, 10 Leistungsverstärkern, einem Minimum von 5 Sensoren und einem Regler mit 5 Reglerkanälen besteht.

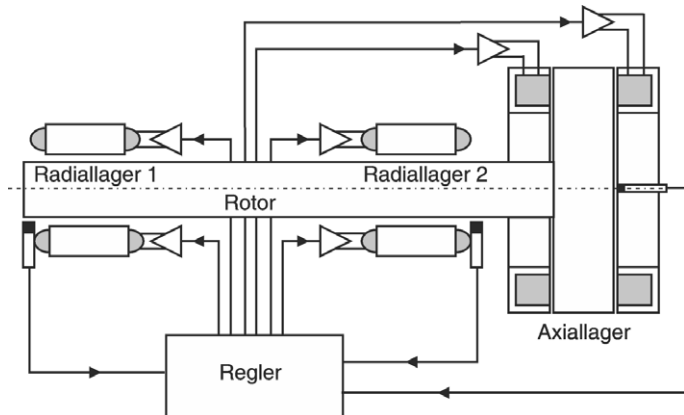


Bild 9-24: Vollständige AMB-Lagerung einer Welle durch zwei Radial- und ein Axiallager

Die berührungsfreie Lagerung mit AMB's ergibt

in verschiedenartigen Anwendungen wesentliche Vorteile gegenüber konventionellen Lagerungen. Sie ist verschleiß-, wartungs- und schmiermittelfrei, wodurch die maximal zulässigen Drehzahlen nur durch die Fliehkraftbeanspruchungen des Rotors begrenzt sind. Die Lagercharakteristik kann über den elektronischen Regler vorgegeben werden und ermöglicht die Dämpfung elastischer Rotoren. Außerdem ist die aktuelle Position des Rotors in den Lagerstellen bekannt, so dass diese zur Prozesskontrolle herangezogen werden kann.

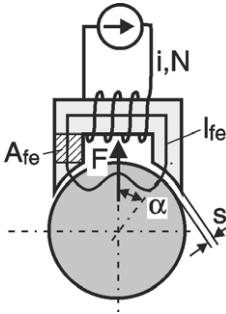


Bild 9-25: AMB-Elektromagnet und

Die maximal erreichbare Steifigkeit und Dämpfung einer AMB-gelagerten Welle wird wesentlich durch das Übertragungsverhalten des AMB-Aktors bestimmt. In Bild 9-25 ist dieser nochmals schematisch mit wichtigen Kenngrößen dargestellt. Es lässt sich folgende quadratische Beziehung zwischen dem Strom i und der Kraft F bzw. dem magnetischen Fluss Φ ableiten:

$$F = \frac{\cos \alpha}{(1 + \Delta)^2} \frac{N^2 \cdot A \cdot \mu_0}{4} \left(\frac{i}{s - x} \right)^2 = k \cdot \left(\frac{i}{s - x} \right)^2 \quad (9.21)$$

$$F = \frac{\Phi^2}{\mu_0 \cdot A} \quad (9.22)$$

Darin sind N die Windungszahl des Elektromagneten, A der Durchflutungsquerschnitt im Luftspalt, μ_0 die Permeabilität von Luft und x die Verschiebung des Rotors vom Luftspalt s aus gegen das Joch. Der Korrekturfaktor

$$\Delta = \frac{l_{fe} \cdot A}{2\mu_r(s - x)A_{fe}} \approx \frac{l_{fe} \cdot A}{2\mu_r \cdot s \cdot A_{fe}}$$

beschreibt den Einfluss des Eisens auf die Lagerkraft, wobei A_{fe} der Durchflutungsquerschnitt im Eisen, l_{fe} die Länge des magnetischen Kreises im Eisen und μ_r die relative Permeabilität des Eisens sind. Mit der Maxwell-Gleichung

$$u = N \frac{\partial \Phi}{\partial t}$$

zwischen dem magnetischen Fluss Φ und der Spannung u über der Wicklung, sowie den Gl. (9.21) und (9.22) ergibt sich:

$$u = N \frac{\partial F}{\partial t} = \frac{N^2 \cdot A \cdot \mu_0}{2(1 + \Delta)} \frac{\partial}{\partial t} \left(\frac{i}{s - x} \right) = 2 \frac{(1 + \Delta)}{\cos \alpha} \cdot k \cdot \frac{\partial}{\partial t} \left(\frac{i}{s - x} \right) \quad (9.23)$$

In den obigen Beziehungen werden die Sättigungseffekte des Eisens und die Einflüsse durch Streufelder nicht berücksichtigt. Die Eisenverluste, die durch den Korrekturfaktor Δ berücksichtigt werden, sind oft sehr klein und können vor allem bei großen Luftspalten vernachlässigt werden. Die Induktivität L des Elektromagneten lässt sich aus

$$L = \frac{N \cdot \Phi}{i}$$

und den Gl. (9.21) und (9.22) bestimmen:

$$L = \frac{N^2 \cdot A \cdot \mu_0}{2(1+\Delta)} \frac{1}{(s-x)} = 2 \cdot k \cdot \frac{(1+\Delta)}{\cos \alpha} \frac{1}{(s-x)} \quad (9.24)$$

und ist vom Luftspalt s abhängig.

Bei einer paarweisen Anordnung der Elektromagneten gemäß Bild 9-23 ergibt sich die ganze Lagerkraft F_t aus der Differenz der Einzelkräfte der sich gegenüberliegenden Magneten. Der Steuerstrom i_x und die Verschiebung x aus der Ruhelage muss dabei in den Beziehungen für das Lager in der positiven x -Richtung positiv und auf der Gegenseite negativ eingesetzt werden:

$$F_t = k \left(\left(\frac{i_0 + i_x}{s-x} \right)^2 - \left(\frac{i_0 - i_x}{s+x} \right)^2 \right) . \quad (9.25)$$

Für die jeweiligen Spannungen u^+ und u^- über den Wicklungen der beiden Seiten gilt dann entsprechend:

$$u^+ = 2 \frac{(1+\Delta)}{\cos \alpha} \cdot k \cdot \frac{\partial}{\partial t} \left(\frac{i_0 + i_x}{s-x} \right) , \quad u^- = 2 \frac{(1+\Delta)}{\cos \alpha} \cdot k \cdot \frac{\partial}{\partial t} \left(\frac{i_0 - i_x}{s+x} \right) . \quad (9.26)$$

Die Gl. (9.25) kann für $x \ll s$ und $i_x \ll i_0$ linearisiert werden:

$$F_t = \frac{4k \cdot i_0}{s^2} i_x + \frac{4k \cdot i_0^2}{s^3} x = k \cdot i_x + k_s \cdot x . \quad (9.27)$$

Die Bedingung, dass die Verschiebung x im Lager bedeutend kleiner als der Luftspalt s ist, ist meist gut erfüllt. Die Bedingung, dass der differentielle Strom i_x bedeutend kleiner als der Vormagnetisierungsstrom i_0 ist, muss durch eine sinnvolle Reglerauslegung erreicht werden.

Bei Verwendung von weichmagnetischem Eisen im Elektromagneten eines AMB-Lagers liegt die maximal nutzbare Flussdichte $B_{\max} = \Phi / A$ bei etwa 1,5 Tesla. Für einen typischen Winkel $\alpha = 22,5^\circ$ kann man daher die maximale Kraft $F_{\max}$ bezogen auf den Durchflutungsquerschnitt A aus der Gl. (9.22) bestimmen:

$$\frac{F_{\max}}{A} = \frac{B_{\max}^2}{\mu_0} \cos \alpha = 1,65 \cdot 10^6 \frac{N}{m^2} \quad \text{mit} \quad \mu_0 = 1,26 \cdot 10^6 \frac{V \cdot s}{A \cdot m} .$$

Bezieht man die Lagerkraft auf die Lagerdimensionen Lagerbreite b und Lagerdurchmesser d und setzt als realistisches Verhältnis von Rotorlängsschnittfläche A_R und Durchflutungsquerschnitt A den Wert 5 an, so gilt für die Maximalkraft bezogen auf die Rotorlängsschnittfläche im Lager:

$$\frac{F_{\max}}{A_R} = 33 \frac{N}{cm^2} .$$

Mit Hilfe der Gl. (9.21) kann man die zur Erreichung der maximalen Lagerkraft erforderliche Durchflutung berechnen:

$$\Theta_{\max} = N \cdot i_{\max} = \frac{2(1+\Delta)B_{\max}}{\mu_0} (s-x) = 2,38 \cdot 10^6 (1+\Delta)(s-x) .$$

Somit ist der benötigte Maximalstrom $i_{\max}$, der vom Leistungsverstärker geliefert werden muss, von der gewählten Windungszahl und dem Luftspalt abhängig.

Die Dynamik, mit der die Lagerkraft als Reaktion auf Schwankungen der Lagerbelastung geändert werden kann, kann durch die Ableitung der Kraft nach der Zeit charakterisiert werden:

$$\frac{\partial F}{\partial t} = \frac{\cos \alpha}{2(1 + \Delta)} \left(\left(\frac{i_0 + i_x}{s - x} \right) u^+ - \left(\frac{i_0 - i_x}{s + x} \right) u^- \right) .$$

Für die maximale Dynamik um die Ruhelage ($x = 0$) ergibt sich somit unter der Voraussetzung $u = u^+ = u^-$ und unter Vernachlässigung des Kupferwiderstandes R_{Cu} der Wicklungen, sowie mit den Werten $\alpha = 22,5^\circ$ und $\Delta \approx 0$:

$$\frac{\partial F}{\partial t} = 1,84 \frac{i_0 \cdot u}{s} .$$

Der Term $i_0 \cdot u$ dieser Gleichung entspricht der verfügbaren Leistung $P_{\max}$ des Verstärkers in der Ruhelage. Mit $i_0 = i_{\max}/2$ erhält man eine direkte Beziehung zwischen der maximalen Verstärkerleistung $P_{\max}$ und der Dynamik der Lagerkraft:

$$\frac{\partial F}{\partial t} = 0,92 \frac{P_{\max}}{s} .$$

Diese Gleichung kann dann zur Auslegung des Verstärkers herangezogen werden.

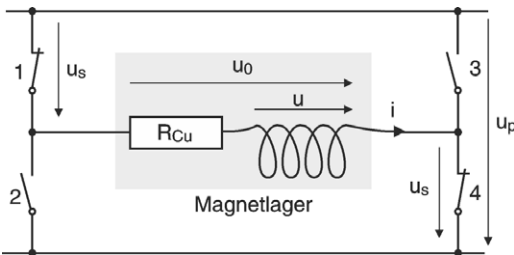


Bild 9-26: Prinzipbild des gepulsten Leistungsverstärkers

Für ein in [9.1] beschriebenes Beispiel der Lagerung einer Hochgeschwindigkeitsfrässpindel kommt als Leistungsverstärker ein gepulster Verstärker zur Speisung der Lagerspulen zum Einsatz, dessen Taktfrequenz im Bereich 10-100 kHz liegt. Dieser Verstärker (Bild 9-26) arbeitet nicht mit Pulsweitenmodulation zur Stromregelung, sondern die Steuerlogik vergleicht in jedem Verstärkertaktzyklus den vorgegebenen

Soll-Strom mit dem gemessenen Ist-Strom und schaltet davon abhängig die Spannung u_0 in der einen oder der anderen Richtung über die Lagerwicklung. Dazu werden entweder die Schalter 1 und 4 bzw. die Schalter 2 und 3 paarweise geöffnet oder geschlossen, je nach dem, ob i_{ist} größer oder kleiner als i_{soll} ist. Als Schalter werden schnelle Leistungstransistoren mit kleinem Spannungsabfall u_s bei geschlossenem Schalter verwendet, an denen bei Verstärkerleistungen von 1 kW nur Verlustleistungen von 50 W auftreten. Der Verstärker wird mit einer Versorgungsspannung von $u_p = 310$ V beaufschlagt.

Im Beispiel wird die Auslegung einer Lagerung einer Frässpindel mit zwei Radial- und einem Axiallager dargestellt. Für die Auslegung wurde eine maximale Schnittkraft am Spindelende von 1000 N angenommen. Dies führt bei der Auslegung der Spindel zu Lagerkräften von 1000 N im Radiallager 1 und 2000 N im Radiallager 2 (Bild 9-24). Nach den oben aufgeführten Gleichungen und einem Lagerdurchmes-

ser von 80 mm im Lager 1 und 96 mm im Lager 2 ergeben sich folgende Lagerlängen:

$$L_1 = 34 - 40 \text{ mm} ; L_2 = 65 - 70 \text{ mm} .$$

Um die 1. elastische Eigenfrequenz der Spindel noch gut zu dämpfen, ermittelt sich aus der Kraftdynamik ein Leistungsbedarf der Verstärker bei einem Luftspalt von 0,35 mm von:

$$\text{Lager 1: } P_{\max} = 0,85 \text{ kW} \quad \text{Lager 2: } P_{\max} = 1,7 \text{ kW}$$

Mit Hilfe der beschriebenen Technik lassen sich angepasste aktive, elektromagnetische Lagerungen für die verschiedensten Anwendungsfälle optimieren.

9.3.2 Anti-Blockier-System (ABS)

Heute sind in fast allen Autos **Anti-Blockier-Systeme** (ABS) eingebaut, dass den Fahrer beim Bremsen so unterstützt, dass der kürzest mögliche Bremsweg auftritt. Aufgrund der physikalischen Gesetze tritt beim Bremsvorgang dann die optimale Bremswirkung auf, wenn der Reifen auf der Fahrbahn gerade noch abrollt und nicht blockiert, d.h. zum Stillstand gekommen ist. Dazu müsste der Fahrer sehr gefühlvoll und doch bestimmt das Bremspedal treten, was unter Stress in einer Gefahrensituation oft nicht möglich ist. Der Normalfall bei Notbremsungen ist, dass durch Vollbremsung die Räder blockieren, was besonders bei kritischem Straßenzustand (Nässe, Eis) zu deutlichen Verlängerungen des Bremswegs führt.

Dieses Bremsverhalten kann man quantitativ durch die Reibzahl μ für den jeweiligen Bremsvorgang charakterisieren. Bild 9-27 zeigt die unterschiedlichen Reibzahlen für Bremsvorgänge auf verschiedenen Untergründen in Abhängigkeit vom **Schlupf** s der Räder. Der Schlupf beim Bremsen ist die Differenz zwischen der Drehgeschwindigkeit des Rades v_R und der Geschwindigkeit des Radkontaktpunktes zur Straße v_K , bezogen auf v_K . Blockiert das Rad, so ist der Schlupf $s = 1$, nähert sich die Raddrehgeschwindigkeit v_R der Horizontalgeschwindigkeit v_K , so sinkt der Schlupf und nimmt kleine Werte an. Bei einem Schlupf von ca. 0,15 tritt dann der optimale Reibwert, d. h. die bestmögliche Bremsverzögerung auf. Diese ist auf trockenem Asphalt besser als auf nassem und diese wiederum besser als auf Schnee. Die unterschiedlichen Situationen zu beherrschen und den optimal kurzen Bremsweg zu ermöglichen ist Aufgabe eines ABS-Systems.

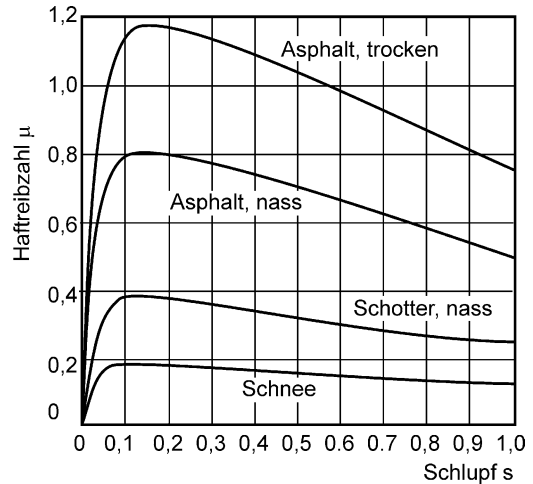


Bild 9-27: Haftreibungszahl μ in Abhängigkeit von Radschlupf s und Fahrzustand

Für ein ABS-System werden die Räder mit Raddrehzahl-Sensoren ausgestattet,

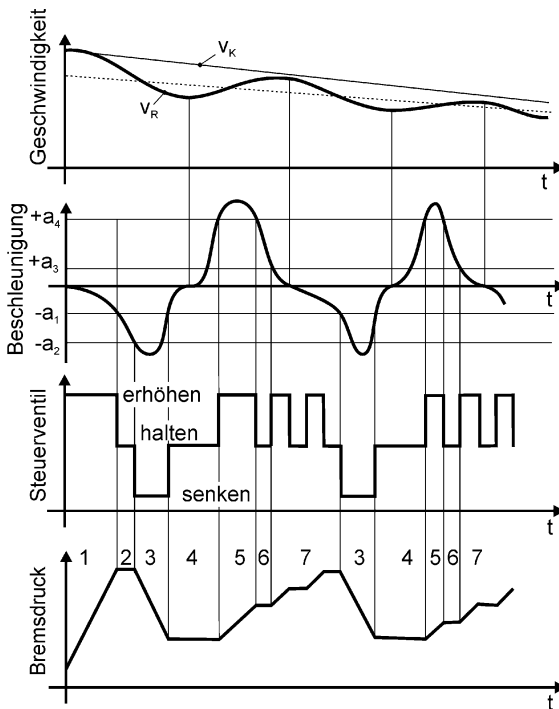


Bild 9-28: Ablauf eines Bremsvorgangs bei einem ABS-System

deren Messwerte einer Steuer-elektronik zugeleitet werden. Stellt diese aufgrund der Messwerte fest, dass gebremst wird, die Räder aber nicht rollen, so wird automatisch die Bremse unabhängig vom Pedaldruck wieder gelöst und zwar so weit, dass die Räder sich gerade wieder zu drehen beginnen. Das Stellsignal zur Erzeugung des Bremsdrucks wird daher nicht nur durch den Druck auf das Bremspedal, sondern auch durch ein Steuersignal beeinflusst, das aus dem Signal des Raddrehzahlsensors abgeleitet wird. Der genaue Regelvorgang ist in Bild 9-28 dargestellt.

Zur Steuerung des Bremsdrucks wird ein Stellventil verwendet, das es erlaubt, den Bremsdruck zu halten, zu erhöhen oder zu erniedrigen. Als eigentliche Steuergröße wird die Raddrehverzögerung a_R (negative Beschleunigung) benutzt, die rechnerisch aus der Raddrehgeschwindigkeit v_R bestimmt wird.

Zuerst tritt bei einer starken Bremsung durch maximalen Pedaldruck ein stark ansteigender Bremsdruck auf (Intervall 1). Wird a_R kleiner als der negative Schwellwert a_1 , so wird die maximal mögliche Reibung zwischen Rad und Straße unterschritten. Im Intervall 2 wird dann zur Störungsunterdrückung bis zum Unterschreiten des negativen Schwellwerts a_2 der Bremsdruck konstant gehalten. Sinkt a_R unter den Schwellwert a_2 , so wird der Bremsdruck vermindert, wodurch das Rad wieder schneller werden kann (3). Zu Beginn von Intervall 4 überschreitet a_R wieder den Schwellwert a_1 und die Abnahme des Bremsdrucks wird beendet. Nun wird der Bremsdruck vorerst konstant gehalten. Dadurch beschleunigt das Rad wieder und der Radschlupf nimmt ab, d.h. der Reibbeiwert und die Bremswirkung nehmen zu. Im Intervall 5 wird beim Überschreiten des positiven Schwellwertes a_4 der Radbeschleunigung, durch Erhöhen des Bremsdrucks das Rad erneut abgebremst, damit die Bremswirkung aufgrund zu kleinen Radschlupfes nicht zu klein wird. Im Intervall 6 zwischen den Schwellwerten a_4 und a_3 wird der Bremsdruck wieder konstant gehalten, im Intervall 7, nach Unterschreiten von Schwellwert a_3 wird der Bremsdruck leicht erhöht.

Wenn a_R den Schwellwert a_1 erneut unterschreitet, beginnt ein zweiter Regelungszyklus. Nun wird jedoch der Bremsdruck sofort erniedrigt, ohne auf das Unterschreiten des Schwellwerts a_2 zu warten (zweites Intervall 3). Beim Durchlaufen weiterer Bremszyklen wird die Radrehverzögerung in einem Bereich gehalten, in dem der

Schlupf die maximal mögliche Reibung zwischen Rad und Straße erlaubt und dadurch der Bremsweg minimiert.

Ein solches mit ABS ausgestattetes Bremssystem leistet etwas, wozu der Mensch nicht in der Lage wäre, nämlich einen physikalischen Vorgang im Grenzbereich des Möglichen zu betreiben. Nach ähnlichen Prinzipien arbeitet auch ein ESP-System, dessen Aufgabe es ist, den Fahrer in schwierig zu beherrschenden Lenksituationen zu unterstützen. Deshalb spricht man auch von einem **Fahrer-Assistenzsystem**. Ein solches Fahrer-Assistenzsystem versucht durch gezieltes Bremsen von einzelnen der vier Räder, wie beim ABS, ein Schleudern des Fahrzeugs bei kritischen Kurvenfahrten zu vermeiden und ohne Über- oder Untersteuern die Spur zu halten. Dadurch kann der Fahrer auch in kritischen Situationen die Kontrolle über sein Fahrzeug behalten.

9.3.3 Aktives Fahrwerk

Die heute noch bei den meisten Kraftfahrzeugen im Einsatz befindlichen passiven Feder-/Dämpfersysteme haben einen Entwicklungsstand erreicht, der die Möglichkeiten für die gleichzeitige Verbesserungen von Fahrsicherheit und Fahrkomfort nahezu ausschöpft. Jede gewählte Fahrwerksabstimmung stellt dabei immer einen Kompromiss zwischen diesen beiden Kriterien dar, je nach dem, ob eine mehr sportlich- sicherheitstechnische oder eine komfortbetonte Fahrphilosophie beim Fahrzeug im Vordergrund steht. Eine wesentliche auch vom Fahrzeugnutzer spürbare Verbesserung beider Eigenschaften, über das Optimum der passiven Abstimmung hinaus, kann nur durch eine sich aktiv an die äußeren Randbedingungen anpassende Feder-/Dämpfercharakteristik erreicht werden. Im folgenden werden die Anforderungen an eine aktive Fahrwerksfederung präzisiert [9.2]:

- **Verbesserung der Fahrsicherheit** durch Verminderung der dynamischen Radlaständerungen: Durch Fahrbahnanregungen wird das Rad zu Schwingungen angeregt, die in extremen Situationen bis zum Abheben des Rades von der Fahrbahn führen können. In einem solchen Grenzfall können die Räder keine Kraft mehr übertragen, was einen beträchtlichen Sicherheitsverlust beim Bremsen oder bei der Kurvenfahrt hervorruft. Die dynamischen Radlastschwankungen dienen daher als Bewertungsgröße für die aktive Sicherheit.
- **Verbesserung der Fahrsicherheit und Lenkcharakteristik** durch fahrzustandsabhängige Beeinflussung der Radlastverteilung: Bei Kurvenfahrt werden durch Fliehkräfte Wankmomente erzeugt, die Radlaständerungen verursachen. Unter **Wanken** versteht man Schwingungen um die Fahrzeuglängsachse. Die Aufteilung dieser Wankmomente auf Vorder- und Hinterachse wird in konventionellen Fahrwerken durch mechanische Stabilisatoren beeinflusst. Je härter der Stabilisator der Hinterachse bei unveränderter Vorderachse, um so größer sind die relativen Radlaständerungen an den Hinterrädern und um so kleiner wird das Seitenkraftpotential der Hinterachse. Verwendet man als aktive Komponente einen Hydraulikzylinder, so werden Stabilisatoren überflüssig, da deren Funktion vom aktiven Stellglied übernommen werden kann. Eine sich anpassende Wankstabilisierung kann durch gezielte Radlastverteilung das Seiten-

kraftpotential erhöhen, sowie im gesamten Einsatzbereich ein definiert wählbares Eigenlenkverhalten (Unter- oder Übersteuern) ermöglichen. Entsprechendes gilt bei Brems- und Beschleunigungsvorgängen für die Nickstabilisierung (Nicken: Schwingungen um die Querachse).

- **Verbesserung des vertikalen Federungskomforts:** Anregungen durch die Fahrbahn werden über die Aufhängung in die Fahrgastzelle weitergeleitet und führen zu Komforteinbußen der Insassen (Stößen, Schütteln). Die sogenannte Aufbaubeschleunigung hat sich hierfür als objektive Bewertungsgröße erwiesen. Für eine gegebene Fahrbahn ist der Komfort um so größer, je kleiner die Beschleunigungsamplituden sind. Die oben beschriebenen Wank- und Nickstabilisierungen führen zu minimalen Wank- und Nickwinkeln und -Beschleunigungen, die ebenfalls eine Komfortverbesserung bedeuten.
- **Verbesserte Niveauregelung:** Eine quasistationär arbeitende Niveauregulierung gehört schon bei vielen Fahrzeugen zur Serien- oder zumindest zur Sonderausstattung. Eine aktive Federung soll darüber hinaus eine verbesserte Höhenstandsregelung beinhalten, um die bekannten Vorzüge von beladungsunabhängigen Federwegen und von lastunabhängiger Achskinematik voll auszuschöpfen. Man erreicht dies durch den geringen Radsturz, der im Grenzbereich

ein verbessertes Längs- und Seitenkraftpotential bietet.

Bei konventionellen Fahrwerken verrichten Feder-/Dämpferelemente die Aufgabe, Rad und Karosserie zu führen und zu dämpfen. Bei einem aktiven Fahrwerk werden die Feder-/Dämpferelemente durch aktive Kraftstellglieder, in der Regel Hydraulikzylinder mit elektrohydraulischem Ventil, ersetzt. Um einen geschlossenen Regelkreis herzustellen, benötigt das System außerdem Sen-

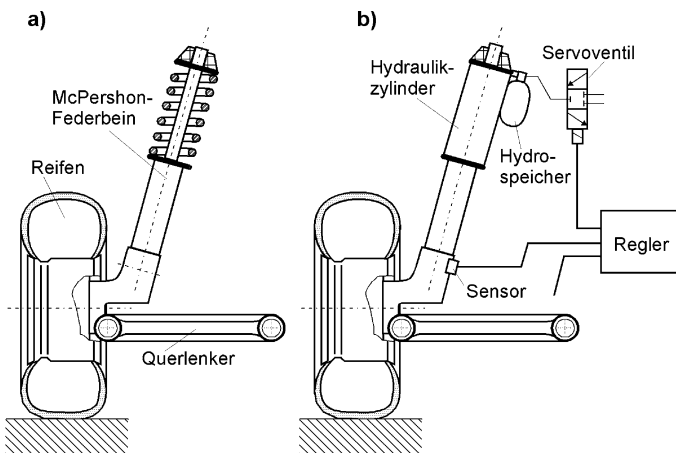


Bild 9-29: Verschiedenartig gefederte Fahrwerke a) konventionell mit Feder-/Dämpfer-Kombination b) aktive Federung mit Hydraulikzylinder und Servoventil

soren zur Erfassung der Federwege und Zylinderdrücke. Diese und weitere Informationen werden dann im Regler zu Stellsignalen für die Ventile in der Art verknüpft, dass Fahrkomfort und Fahrsicherheit des Fahrzeugs in jeder Fahrsituation optimal sind. Bild 9-29 a) zeigt ein einzelnes Rad mit konventionellem Feder-/Dämpfer-Element (McPershon Federbein) und Bild 9-29 b) ein Rad mit aktivem Federungssystem [9.3].

Um die Auswirkung der aktiven Federung zu untersuchen, wurde ein Viertel des Fahrwerks und das Federungssystem auf einem Simulationssystem modelliert. Bild

9-30 zeigt das Modell des Viertelfahrzeugs, in dessen Zentrum zwischen Aufbau-masse und Radmasse das aktive Federungssystem eingefügt werden kann.

9.3.3.1 Aktive Federung mit Hydrozylinder

Das aktive Federungssystem besteht aus einem Hydraulikzylinder in Plungerbauweise, der von einem Servoventil mit Drucköl versorgt wird. Bild 9-31 zeigt eine schematische Darstellung des Systems und die für die Simulation erforderlichen Größen. Mit dem Servoventil kann durch Zu- oder Abführung von Hydrauliköl ein vorgegebener Druck im Hydraulikzylinder und damit eine gewünschte Zylinderkraft eingestellt werden. Der Hydrospeicher übernimmt bei hohen Kolbengeschwindigkeiten die Ölströme, die nicht vom Ventil gestellt werden können und wirkt somit entlastend für das Ventil. Die Kombination Hydrospeicher/Drossel bestimmt die Grundsteifigkeit, also die Federsteifigkeit bei ausgeschaltetem Regler und geschlossenen Ventilen. Diese Federsteifigkeit ist entscheidend für das Systemverhalten außerhalb des Regelbereichs und wird deshalb hydraulisch weich und komfortabel gewählt. Mit einer weichen Grundabstimmung weist das System auch Notlaufeigenschaften auf, so dass bei Reglerausfall das Fahrzeug weiterhin gute Fahreigenschaften behält. Komfort und Sicherheit sind dann so abgestimmt wie bei einem Fahrzeug mit konventioneller Federung, so dass ein problemloses Weiterfahren auch ohne Reglerbetrieb möglich ist.

Gegenüber einem konventionell gefederten Vergleichsfahrzeug können Fahrkomfort und Fahrsicherheit durch folgende Eigenschaften der aktiven Federung verbessert werden:

- Erhöhung der Aufbaudämpfung,
- Senkung der Aufbaubeschleunigung bis zu 38% (RMS),
- Kompensation von Wank- und Nickbewegungen.

Das Federungssystem wurde im Labor mittels einer **Hardware-in-the-loop- Simulation** erprobt. Dabei werden die Hardwarekomponenten der aktiven Federung auf einem Prüfstand aufgebaut und alle anderen Komponenten simuliert. Zwei zusätzliche Hydraulik-

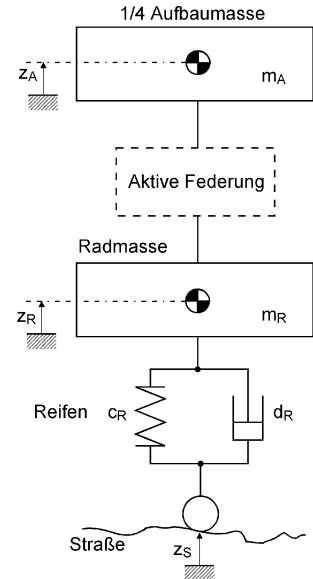


Bild 9-30: Ersatzmodell eines Viertelfahrzeugs

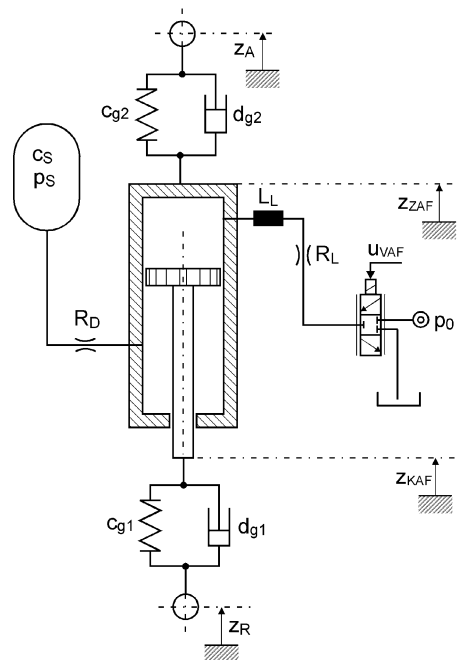


Bild 9-31: Simulationsmodell des aktiven Federungssystems

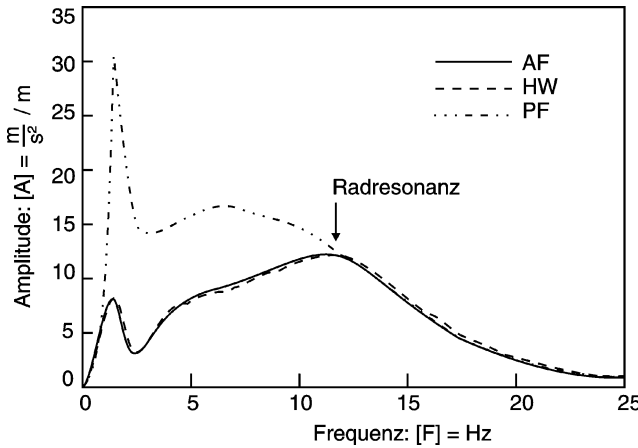


Bild 9-32: Betragsfrequenzgänge von aktiver und konventioneller (passiver) Federung **AF:** aktiv **PF:** passiv

likzylinder bewegen im Simulationsaufbau den Federungszyylinder so, als wäre er im Fahrzeug eingebaut. Über einen der Zusatzzyylinder werden die vom Rad weitergeleiteten Stöße des Straßenprofils simuliert.

In Bild 9-32 sind die Betragsfrequenzgänge vom Eingang des Anregungsmodells (Straßenformfilter) zur bewerteten Aufbaubeschleunigung des Fahrzeugs für verschiedene Fälle aufgetragen. Die mit **AF** bezeichnete Kurve stellt das Ergebnis der Simulation der

aktiven Federung, die mit **HW** bezeichnete Kurve gemessene Ergebnisse dar. Zusätzlich ist die mit **PF** bezeichnete Kurve der Simulation eines konventionellen Fahrwerks dargestellt. An diesen Kurven sieht man, dass die Karosserie mit der konventionellen Federung unterhalb der Radresonanz, die bei etwa 12 Hz liegt, bei gleicher Anregung deutlich höhere Amplituden aufweist. Dabei stimmen die Kurven für die Simulation des aktiven Fahrwerks und die durch Messung am aktiven Fahrwerk (HW) ermittelte sehr gut überein. Daher muss das Modell für den untersuchten Frequenzbereich gut mit dem realen Fahrwerk übereinstimmen. Oberhalb der Radresonanzstelle verlaufen alle Kurven gleich, so dass sich hier keine Verbesserung durch das aktive Fahrwerk mehr ergibt. Weitere Verbesserungen sind nur noch mit den aktiven Schwingungstilgern zu erreichen, die im nächsten Abschnitt beschrieben werden.

Das hier vorgestellte aktive Federungssystem ist inzwischen erfolgreich in Kraftfahrzeugen eingebaut und erprobt worden und hat Serienreife erreicht.

9.3.3.2 Aktive Federung mit Hydrozylinder und aktivem Tilger

Um allen für eine aktive Federung gestellten Anforderungen einigermaßen gerecht zu werden, muss man auch bei dem im vorherigen Abschnitt besprochenen aktiven Federungssystem Kompromisse eingehen. So lassen sich Aufbau- und Raddämpfung des Fahrzeugs nicht unabhängig voneinander einstellen. Daher tritt eine ausgeprägte Amplitudenüberhöhung im Bereich der Radresonanz auf. Um also diese Resonanz gezielt zu dämpfen, kann man das Prinzip des **Hilfsmassendämpfers** oder **Schwingungstilgers** anwenden.

Die Wirkung eines einfachen Tilgers kann man anhand des in Bild 9-33 a) dargestellten Zwei-Massen-Systems zeigen. Die Masse m_1 sei die Masse, die einer harmonisch veränderlichen Erregerkraft $F_0 \cos \Omega_1 t$ unterworfen ist und ohne die Masse m_2 starke Schwingungen ausführt, wenn Ω_1 in der Nähe der Eigenfrequenz ω_1 des

Feder-Masse- Systems liegt. Die zweite Masse m_2 stellt zusammen mit der Koppel- feder k_2 den Schwingungstilger dar. Die beiden Massen mögen sich reibungsfrei ohne Dämpfung auf der Unterlage bewegen. Um die Amplituden der beiden Massen in Ab- hängigkeit von der Erregerfrequenz Ω_1 zu berechnen, muss das folgende aus der Skizze in Bild 9-33 b) hergeleitete Differenti- algleichungssystem gelöst werden:

$$\begin{aligned} m_1 \ddot{x}_1 + k_1 x_1 - k_2 (x_2 - x_1) &= F_0 \cos \Omega_1 t \\ m_2 \ddot{x}_2 + k_2 (x_2 - x_1) &= 0 \end{aligned}$$

Zur Untersuchung des stationären Schwin- gungsverhaltens ist für die Lösung des inho- mogenen Gleichungssystems nur eine partik- uläre Lösung erforderlich. Mit dem Lö- sungsansatz

$$x_1 = B_1 \cos \Omega_1 t$$

$$x_2 = B_2 \cos \Omega_1 t$$

ergibt sich folgendes Gleichungssystem:
$$\begin{aligned} [(k_1 + k_2 - m_1 \Omega_1^2) B_1 - k_2 B_2 - F_0] \cos \Omega_1 t &= 0 \\ [(k_2 - m_2 \Omega_1^2) B_2 - k_2 B_1 - F_0] \cos \Omega_1 t &= 0 \end{aligned}$$

Aus diesem Gleichungssystem kann man die unbekannten Amplituden der beiden Massen berechnen:

$$B_1 = \frac{\frac{F_0}{m_1} \left(\frac{k_2}{m_2} - \Omega_1^2 \right)}{(\Omega_1^2 - \omega_1^2)(\Omega_1^2 - \omega_2^2)}, \quad B_2 = \frac{\frac{F_0}{m_1} \cdot \frac{k_2}{m_2}}{(\Omega_1^2 - \omega_1^2)(\Omega_1^2 - \omega_2^2)}.$$

Aus diesen Amplitudengleichungen kann man folgendes ablesen: Die Amplitude B_2 des Tilgers, auf den die Anregungskraft nicht direkt, sondern nur über die Feder k_2 wirkt, wird für alle Kombinationen der Massen, Federkonstanten und der Anre- gungsfrequenz von Null verschieden sein, d. h. der Tilger führt immer eine Schwin- gung aus. Die Amplitude B_1 der erregten Hauptmasse wird jedoch für den Fall $\Omega_1^2 = k_2 / m_2$ Null werden, d. h. die Masse bleibt in Ruhe. Die Eigenfrequenz des Til- gers beträgt aber gerade $\omega_2 = \sqrt{k_2 / m_2}$ bzw. $\omega_2^2 = k_2 / m_2$, so dass man folgern kann: Wenn die Eigenfrequenz des Tilgers mit der Erregerfrequenz übereinstimmt, so bleibt die Hauptmasse in Ruhe. Wird also ein schwingungsfähiges System mit einer bestimmten Frequenz angeregt (z. B. durch Unwucherregung), so kann man durch Anbringung eines auf diese Frequenz abgestimmten Tilgers, die Schwingung des Systems zur Ruhe bringen. Bild 9-34 zeigt als numerische Simulation die Schwin- gung der Masse eines Einmassenschwingers, der eine Eigenfrequenz $\omega_1 = 10$ rad/s besitzt. Bei Anregung durch eine periodische Kraft mit einer Erregerfrequenz von $\Omega_1 = 9$ rad/s führt die Masse die im Bild oben links gezeigte Schwingung aus. An

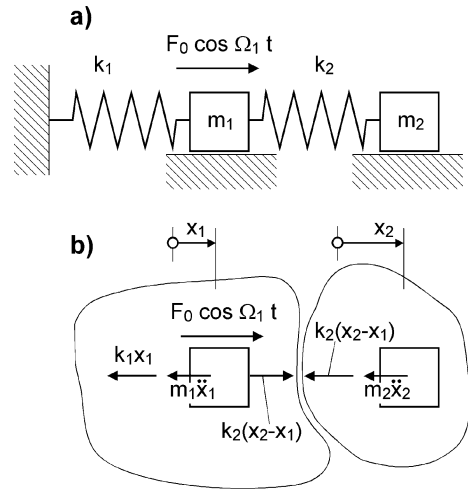


Bild 9-33: Zwei-Massen-System aus ange- regtem Hauptschwinger und Schwingungstilger
a) Anordnung b) freigeschnittene Massen

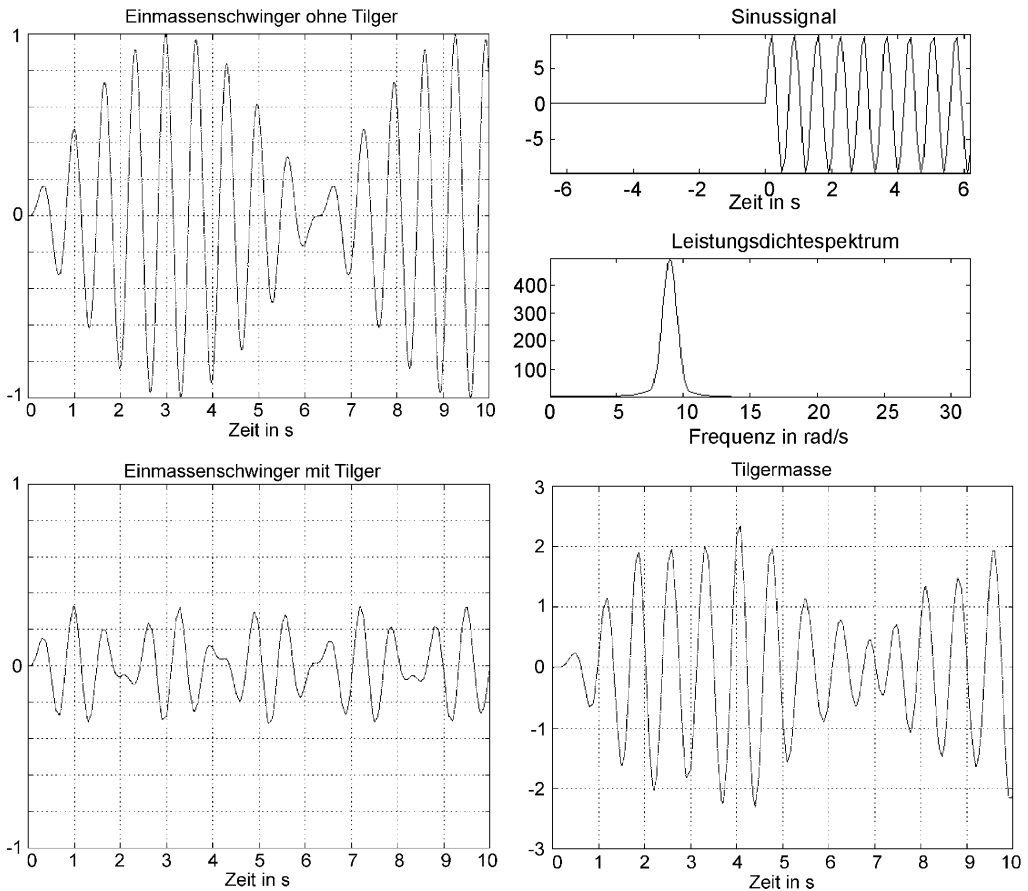


Bild 9-34: Schwingung eines Einmassenschwingers ohne und mit Tilger $\omega_1 = 10 \text{ rad/s}$, $\omega_2 = 9 \text{ rad/s}$, $\Omega_1 = 9 \text{ rad/s}$

die Masse des Einmassenschwingers wurde dann eine Tilgermasse über eine Feder angekoppelt und die Eigenfrequenz des Tilgers auf die Erregerfrequenz ausgelegt. Bild 9-34 zeigt unten links den Amplitudenverlauf der Masse des Einmassenschwingers und rechts den Amplitudenverlauf der Tilgermasse. Die Anregung erfolgte hier mit gleicher Frequenz und Kraftamplitude. Man sieht, dass die Schwingungen der Hauptmasse stark zurückgegangen sind.

Da reale Systeme nie dämpfungsfrei sind, klingen die Eigenvorgänge auf jeden Fall ab und man muss nur noch die stationären Vorgänge betrachten. Für gedämpfte Schwinger bleibt im stationären Zustand immer eine Restschwingung, so dass es günstig ist die Dämpfung des Tilgers möglichst klein zu wählen.

Diese Art der Schwingungstilgung funktioniert natürlich nur für eine ganz bestimmte Erregerfrequenz. Dies ist auch der Grund warum die Schwingung des Einmassenschwingers, wie theoretisch gefordert, nicht vollständig verschwindet (Bild 9-34). Durch die Abstimmung des Tilgers auf die Anregefrequenz entstehen zusätzliche Resonanzen unterhalb und oberhalb der Erregerfrequenz. Ist das Anregesignal nicht rein sinusförmig, wie in Bild 9-34 als Leistungsdichtespektrum dargestellt, so

hat in der Simulation das Sinus-signal eine gewisse Bandbreite. Dadurch werden Frequenzen neben der Eigenfrequenz des Tilgers das System wiederum anregen.

Im vorherigen Abschnitt war das aktive Fahrwerk mit Hydrozylinder und Servoventil beschrieben worden. Durch Änderung der Drosselung des Ölstroms durch das Servoventil wurden bei ihm Federkonstante und Dämpfung den jeweiligen Straßenverhältnissen angepasst. Das gleiche Prinzip kann man natürlich auch für einen in der Eigenfrequenz verstellbaren Schwingungstilger verwenden. Mit ihm können Schwingungen

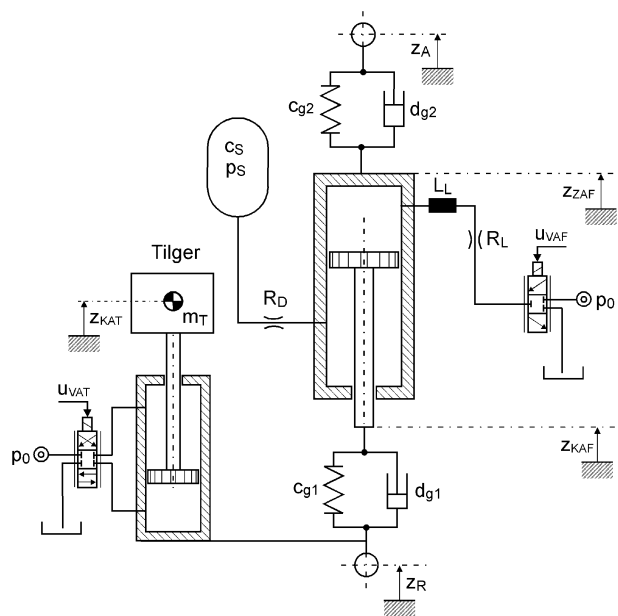


Bild 9-35: Aktive Federung mit aktivem Tilgersystem

aufgrund der Radresonanz gezielt gedämpft werden. In Bild 9-35 ist der schematische Aufbau einer aktiven Federung mit zusätzlichem Schwingungstilger dargestellt. Der Tilger besteht aus einer Tilgermasse m_T und einem zusätzlichen Hydrozylinder mit Servoventil, das vom Regler je nach Straßenzustand verschieden angesteuert wird. Diese Anordnung wird in das mit "aktive Federung" bezeichnete gestrichelte Feld in Bild 9-30 eingefügt.

Die aktive Federung des vorherigen Abschnitts brachte nur Verbesserungen von Fahrsicherheit und Fahrkomfort unterhalb der Radresonanz. Da diese jetzt durch den aktiven Tilger gezielt bedämpft wird, werden hierdurch weitere Verbesserungen der Fahreigenschaften erzielt. Bild 9-36 zeigt wiederum die Betragsfrequenzgänge der Hardware-in-the-loop-Simulation. Darin bedeuten die Kurven **PF** ein konventionelles Fahrwerk mit passiver Federung (Simulation), **HW** die aktive Federung (Messung) und **HWAT** die aktive Federung mit aktivem Tilger (Messung). Man sieht, dass im Vergleich zu Bild 9-32 die Radresonanz deutlich gedämpft wurde und daher die Amplitude im Bereich 2-20 Hz nochmals stark abgenommen hat.

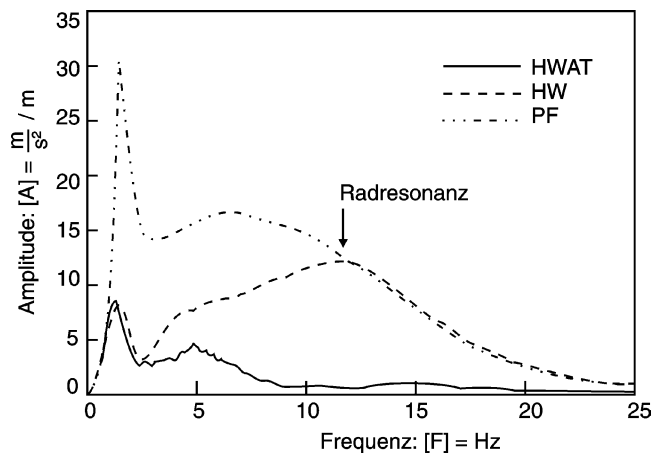


Bild 9-36: Betragsfrequenzgänge der aktiven Federung mit und ohne aktiven Tilger

Bild 9-36 zeigt wiederum die Betragsfrequenzgänge der Hardware-in-the-loop-Simulation. Darin bedeuten die Kurven **PF** ein konventionelles Fahrwerk mit passiver Federung (Simulation), **HW** die aktive Federung (Messung) und **HWAT** die aktive Federung mit aktivem Tilger (Messung). Man sieht, dass im Vergleich zu Bild 9-32 die Radresonanz deutlich gedämpft wurde und daher die Amplitude im Bereich 2-20 Hz nochmals stark abgenommen hat.

9.3.4 Mechatronische Anwendungen bei Industrierobotern

Anfänglich tauchte der Begriff Mechatronik vor allem im Zusammenhang mit Industrierobotern und anderen autonomen Robotersystemen auf. Eine der ersten Buchveröffentlichungen [9.4] aus dem Jahr 1991 trägt den Titel "Mechatronics & Robotics" und ist eine Sammlung von Vorträgen einer internationalen Konferenz. Die Robotertechnologie ist seitdem ein wichtiges Anwendungsfeld der Mechatronik.

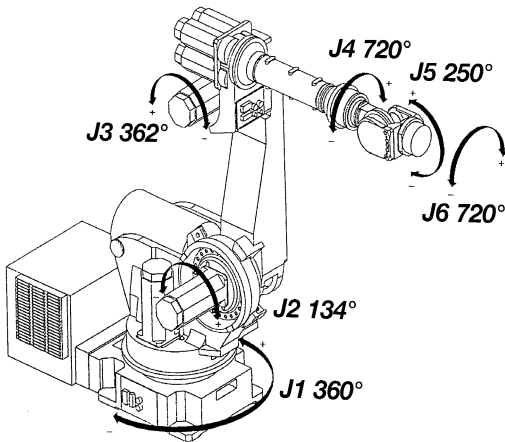


Bild 9-37: 6-achsiger Vertikal-Knickarmroboter
Typ S430 der Fa. Fanuc Robotics

Nun ist ein "normaler" Industrieroboter, wie er in Bild 9-37 dargestellt ist, noch nicht von vornherein ein mechatronisches System. Vergleicht man ein solches Gerät, das eine universelle Handhabungseinrichtung für Werkstücke und Werkzeuge darstellt, mit dem Strukturbild eines mechatronischen Systems (Bild 1-24), so sieht man, dass alle wesentlichen Bestandteile vorhanden sind, bis auf Sensoren, die Informationen aus der Umwelt aufnehmen. Für Standard-Handhabungsaufgaben ist dies auch nicht erforderlich, da man hier von festen Positionen der Handhabungsobjekte ausgehen kann, die dem Roboter bei der Erstellung des Bewegungspro-

gramms gezeigt (geteacht) werden. Ein solches Bewegungsprogramm kann dann immer wieder automatisch abgefahren werden.

Häufig treten jedoch Umstände auf, die eine mehr oder weniger große Unbestimmtheit in einen Handhabungsprozess einbringen, die durch einfaches Zeigen der Positionen und anschließendes automatisches Abfahren der Bewegungsbahnen nicht gelöst werden können. Einfache Problemstellungen dieser Art liegen vor, wenn Objekte, die gehandhabt werden sollen, nicht immer reproduzierbar an der gleichen Position dem Roboter zur Handhabung übergeben werden. Um trotzdem eine automatische Handhabung zu ermöglichen, benötigt der Roboter zusätzliche Sensoren, die die erforderlichen Informationen über Abweichungen in Lage, Form, Gewicht oder ähnlichem erfassen und daraufhin eine zielgerichtete Beeinflussung des Bewegungsprogramms vornehmen. Bild 7-6 zeigt eine solche Anwendung, bei der Kleinteile, die auf einem Förderband in das Sichtfeld eines Vision-Systems befördert werden, dort in geometrisch unbestimmter Lage vom Roboter aufgenommen werden. Die Greifposition wird dabei von der Robotersteuerung aufgrund des vom Vision-System aufgenommenen Bildes der Teileanordnung errechnet. Dies entspricht dem oberen Zweig in dem Strukturbild über Funktionalitäten von mechatronischen Systemen in Bild 1-26, in dem aufgenommene Informationen aus der Umwelt so interpretiert werden, dass ein zielorientiertes Verhalten für die Ausführung der Handlungen vorliegt.

Der Vorgang der Entwicklung eines Sensorsystems zur Erfassung von Unbestimmtheiten in einem Werkzeug-Handhabungsprozess und die dazu erforderliche Korrektur

tursoftware, wurde in einem studentischen Entwicklungsprojekt im Rahmen der Lehrveranstaltung "Industrieroboter" für Studierende der Mechatronik durchgeführt. Dies wird im nachfolgenden Abschnitt beispielhaft dargestellt.

9.3.4.1 Nachführen eines Roboterarms an einer Freiformfläche

Bei vielen Automatisierungsaufgaben, bei denen ein Industrieroboter (IR) ein Werkzeug handhaben soll, besteht die Aufgabe, das Werkzeug entlang einer ebenen oder räumlichen Kontur mit konstantem Abstand zum Werkstück zu bewegen. Beispiele hierfür sind das Laser- und Wasserstrahlschneiden oder das Auftragen einer Kleberraupe. Im Normalfall wird der Verlauf der erforderlichen Werkzeugbahn durch Teachen von Bahnstützpunkten dem IR mitgeteilt. Vielfach sind bei solchen Verfahren die Werkstücke (Blechteile, Kunststoffteile) aufgrund der inneren Instabilität oder durch Toleranzen in ihrer Geometrie nicht sehr genau. Reagiert dann das Bearbeitungsverfahren empfindlich auf Abstandsveränderungen zwischen Werkzeug und Werkstück, so ist häufiges Nachteachen der Bahnstützpunkte erforderlich oder die Verfahrensqualität schwankt sehr stark. Bild 9-38 zeigt eines der Strahlverfahren. Da der Schneidstrahl (Laser, Wasser) nicht zylindrisch sondern konisch ist, kommt es bei Abstandsänderungen der Schneiddüse zum Werkstück zu Veränderungen des Schneidspaltes.

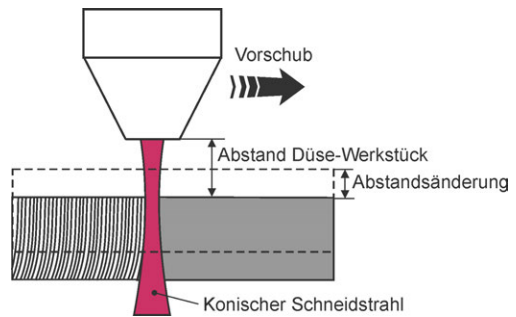


Bild 9-38: Veränderung des Schneidspaltes im Werkstück bei Abstandsänderungen aufgrund von Welligkeit des Werkstücks

Dieses Problem könnte gelöst werden, wenn der Roboter aufgrund von Abstandsmessdaten einer Sensorik den Bahnverlauf in Abhängigkeit von den aktuellen Werkstückschwankungen selbsttätig korrigieren würde. Die Projektidee für das studentische Entwicklungsprojekt bestand daher darin, eine Sensorik zu entwickeln und mit der Steuerung des Roboters zu verbinden. Die Messwerte sollten in die Steuerung übernommen, ausgewertet und die Bahndaten eines Bewegungsprogramms aufgrund dieser Messwerterfassung automatisch korrigiert werden.

Da das Problem **Abstandsänderung** bei den oben genannten Bearbeitungsverfahren seit längerem bekannt ist, gibt es natürlich kommerzielle Produkte, die das Problem lösen. Beispielsweise wird beim Laserschneiden eine kapazitive Abstandssensorik benutzt, deren Messwert zur Erzeugung eines Stellsignals für eine separate Zustellachse benutzt wird. Wegen dieses Aufwandes und verschiedener technologischer Zusatzfunktionen ist ein solches System sehr teuer und liegt in einem Kostenrahmen von ca. 50 T Euro. Die zusätzliche Stellachse in z-Richtung des Koordinatensystems ist beispielsweise bei einer Laser-Bearbeitungsanlage für ebene Bleche ohnehin erforderlich, da die Bewegungskinetik einer solchen Anlage nur Bewegungen in der x-y-Ebene vorsieht. Hierdurch wird beispielsweise die Welligkeit einer ansonsten ebenen Blechtafel ausgeglichen. Wendet man das Verfahren an

räumlich gekrümmten Oberflächen an, indem man den Laserschneidkopf von einem 6-achsigen Industrieroboter (Bild 9-37) führen lässt, so enthält die Bewegungskinematik schon alle Freiheitsgrade die erforderlich wären, um Abweichungen von Form und Lage auszugleichen. Trotzdem verwendet man in kommerziellen Produkten auch ein komplettes Regelungssystem mit zusätzlicher Stellachse, um vom Robotertyp und seinen steuerungstechnischen Fähigkeiten unabhängig zu sein.

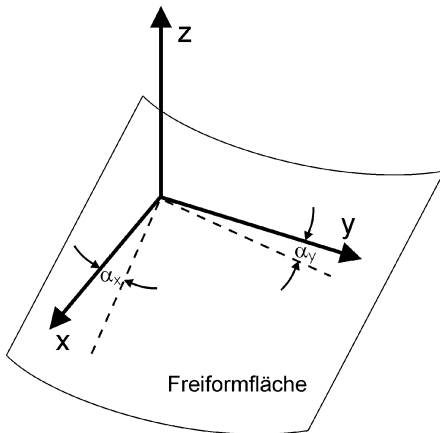


Bild 9-39: Neigungswinkel des Koordinatensystems gegenüber der Freiformfläche

Um den Abstand zu der Freiformfläche zu messen, würde eine Abstandsmessung in einem Punkt ausreichen. Um jedoch die Orientierung des Roboterhandgelenkes und damit die des Werkzeug-Koordinatensystems konstant senkrecht zur Fläche zu halten, müssen die beiden Neigungswinkel α_x und α_y der Achsen x und y des Koordinatensystems zur Freiformfläche erfasst und in einer Korrekturstrategie berücksichtigt werden (Bild 9-39). Im Soll-Zustand müssten die beiden Winkel den Wert Null haben.

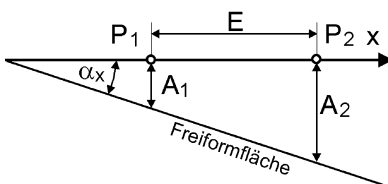


Bild 9-40: Winkelbestimmung zwischen Koordinatenachse und Freiformfläche

Um den Neigungswinkel zwischen zwei Geraden (Tangente an die Freiformfläche/Koordinatenachse) zu bestimmen, muss man, von zwei Punkten einer Parallelen zur Koordinatenachse aus, eine Abstandsmessung zur Tangente an die Freiformfläche vornehmen (Bild 9-40). Sind die beiden Punkte nicht zu weit voneinander entfernt und die Fläche nicht zu stark gekrümmt, so kann man die Freiformfläche zwischen den beiden Messpunkten durch eine Gerade ersetzen und der Winkel α ergibt sich dann zu:

$$\alpha = \arctan\left(\frac{A_2 - A_1}{E}\right)$$

Die minimale Anzahl von Sensoren für die Messung von zwei Winkeln beträgt drei, da man den Punkt A_1 für beide Winkelmessungen gleich wählen kann. Ordnet man

die Sensoren symmetrisch auf einem Kreis an, so sollte eine Messung der Neigung der xy-Ebene des Werkzeugkoordinatensystems gegenüber der Freiformfläche immer möglich sein.

Für die Abstandmessung sollten preiswerte Lichttaster, die in einem Gehäuse integriert eine Infrarot-Leuchtdiode als Sender und einen Fototransistor als Empfänger enthalten, verwendet werden. Die Vermessung der Kennlinie mit Hilfe eines 12-Bit A/D-Wandlers (Wertebereich 0-2047) in der Robotersteuerung ergab die in Bild 9-41 dargestellte Kennlinie. Wie zu erwarten, ergab sich ein stark nichtlineares Messverhalten, das sich im Bild

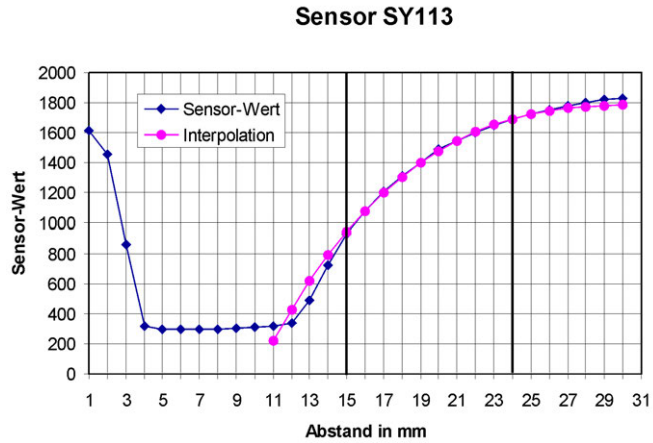


Bild 9-41: Amplitudenkennlinie des Abstandsensors SY113 und Verlauf der interpolierten Funktion

ablesen lässt. Als erstes kann man erkennen, dass sich deutlich unterscheidbare Messwertänderungen nur im Bereich zwischen 12 mm und 30 mm ergeben. Um die dort messbaren Werte zur Berechnung der Entfernung auswerten zu können, müssen sie jedoch linearisiert werden, bzw. die nichtlineare Kennlinie muss durch eine bekannte rationale Funktion approximiert werden. Dies kann man durch die Newton'sche Interpolation der Kennlinienfunktion im interessierenden Bereich berechnen. Hier ergab sich folgende interpolierte Funktion:

$$W = -0,158(32,5 - A)^3 + 1789$$

wobei W der Messwert am Ausgang des A/D-Wandlers und A der wahre Abstand in mm ist.

Die Anschmiegung dieses Polynoms dritter Ordnung an die Kennlinie des Sensors passt nur im Bereich zwischen 15 mm und 25 mm, da hier die Abweichung, als Fehler in mm berechnet, kleiner oder gleich 0,1 mm beträgt. Man könnte durch Einbeziehung von mehr Stützstellen und einem daraus resultierenden Polynom höherer Ordnung zwar eine bessere Anschmiegung in einem größeren Intervall erreichen, würde aber dadurch den Rechenaufwand bei der späteren Positionskorrektur stark erhöhen. Der brauchbare Messbereich beträgt demnach 10 mm, der Sollabstand für die Abstandsregelung sollte in der Mitte des Messbereichs bei 20 mm Abstand von der Oberfläche liegen. Um noch genügend Sicherheit für die Regelung zu haben, wurde daher eine maximale Regelabweichung von ± 3 mm festgelegt. Die drei erforderlichen Sensoren wurden dann in einen nach unten in Tastrichtung offenen Sensorhalter (Bild 9-42) eingebaut, der so an der Flansch-

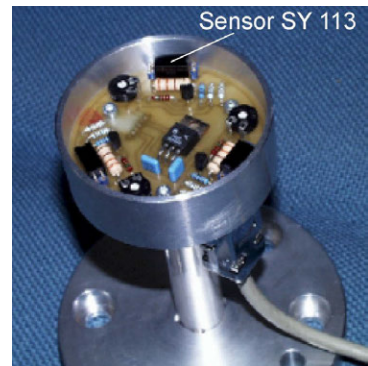


Bild 9-42: Sensorhalter mit drei Sensoren SY113

platte befestigt wurde, dass die z-Achse des Werkzeug-Koordinatensystems in Tastrichtung zeigt.

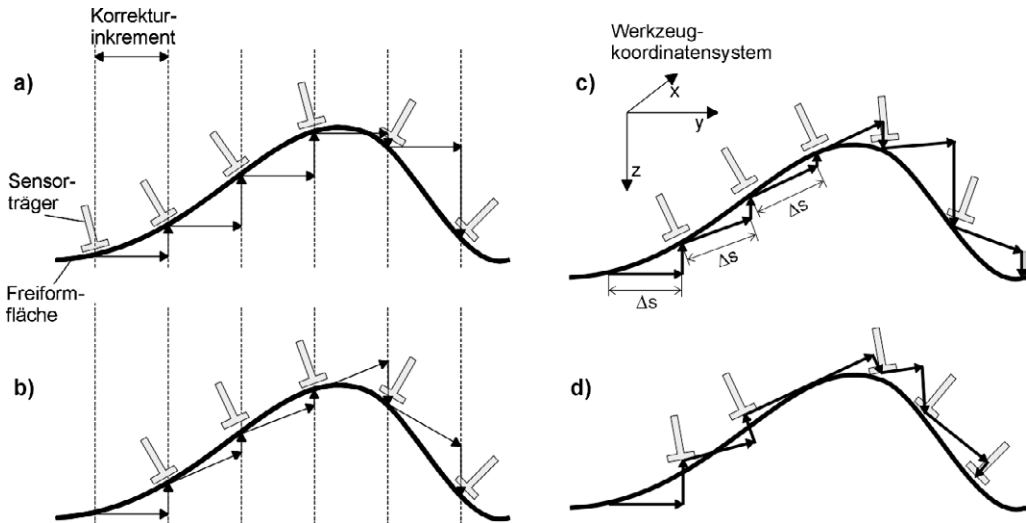


Bild 9-43: Verschiedene Möglichkeiten des Verfahrens entlang einer Freiformkontur mit den zugehörigen Korrekturverfahren

Will man das geplante Abstands-Korrektursystem bei einer Werkzeughandhabung über einer Freiformfläche einsetzen, so gibt es eine Soll-Werkzeugbahn, die auf übliche Art und Weise geteacht wurde. Um nun Korrekturen beim Abfahren dieser Bahn zu berücksichtigen, sind die in Bild 9-43 dargestellten unterschiedlichen Strategien möglich. Prinzipiell gibt es die beiden folgenden Möglichkeiten:

- Ereignisüberwachung beim Abfahren der Soll-Bahn und Korrektur bei Bedarf
- Inkrementelles Fortschreiten auf der Soll-Bahn und Korrektur nach jedem Weginkrement.

Die in Teilbild 9-43 a) dargestellte Vorgehensweise geht davon aus, dass der Soll-Weg in der xy-Ebene in feste Inkremente aufgeteilt wird. Der Verfahrensbefehl für das Zurücklegen des Weges wird nur in der xy-Ebene ausgeführt. Nach dem Verfahren eines solchen Weginkrementes erfolgt eine Messung des Abstandes und der Orientierung mit anschließender Korrektur des Abstandes in z-Richtung und der Orientierung des an der Roboterflanschplatte montierten Sensorhalters normal zur Fläche. Bei dieser Vorgehensweise treten Probleme bei großen positiven oder negativen Steigungen der Freiformfläche auf, weil die Abstands- und Orientierungsänderungen in den Korrekturpunkten sehr groß werden können. Wie oben gesagt, ist aber nur ein Arbeitsbereich von 10 mm mit einer maximalen Regelabweichung von $\pm 3\text{mm}$ vorgesehen, so dass es bei dieser Methode leicht zum Verlassen des Arbeitsbereichs kommen kann.

In Teilbild 9-43 b) wird wie bei der Vorgehensweise in Teilbild a) der Weg über der Fläche in Inkremente zerlegt. Das Verfahren erfolgt dann jedoch nicht nur in der xy-Ebene, sondern an den Korrekturpunkten wird eine neue z-Koordinate für den nächsten Verfahrensweg berechnet. In den Korrekturpunkten wird dann auch die Ori-

entierung korrigiert. Bei gleichmäßigen Steigungen eines beliebigen Winkels erreicht man eine recht geringe Abweichung von der Freiformfläche. Jedoch treten hier Probleme bei kleinen Krümmungsradien nach großen Steigungen auf.

Bei der Vorgehensweise in Teilbild 9-43 c) erfolgt eine Zerlegung des Weges in feste Weginkremente Δs im Raum. Der Abstand und die Orientierung werden nach jedem Schritt korrigiert. Probleme treten hierbei auch nach kleinen Krümmungsradien der Fläche auf.

Bei der Methode in Teilbild 9-43 d) wird auf die Ereignisüberwachung innerhalb von Fahrbefehlen zurückgegriffen, d. h. die Sensoren kontrollieren während eines Fahrbefehls ständig Abstand und Orientierung und korrigieren gegebenenfalls. Hierbei besteht das Problem, dass man insbesondere bei geringen Krümmungen der Fläche kaum eine Aussage machen kann, wann und wie stark korrigiert werden muss.

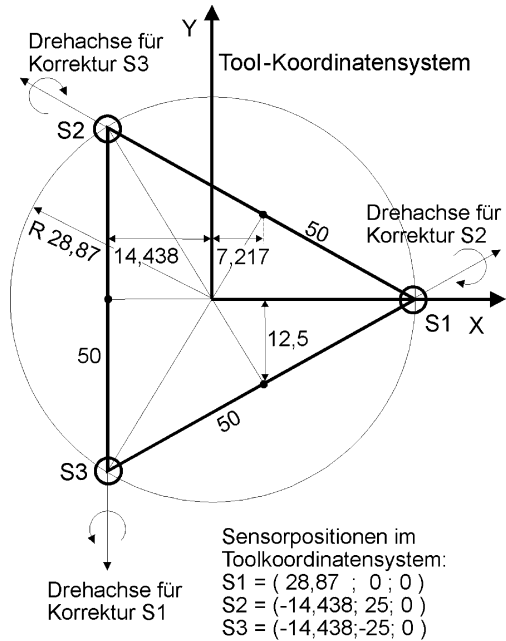


Bild 9-44: Geometrische Anordnung der drei Abstandssensoren im Sensorträger

Im Projekt entfiel die Entscheidung auf die Korrekturmethode in Teilbild 9-43 c), da sie bei nicht zu kleinen Krümmungsradien die besten Korrekturergebnisse liefern sollte. In Bild 9-44 ist nun die genaue geometrische Anordnung der Sensoren im Sensorhalter dargestellt. Die eigentliche Abstands- und Orientierungskorrektur an den Stützpunkten der Soll-Bewegungsbahn erfolgt dann so, dass man drei Drehachsen S1, S2 und S3 durch jeweils 2 der Sensoren definiert, um die eine Drehung des Sensorkopfes ausgeführt wird. Diese Drehung erfolgt abgeleitet vom Abstandsmesswert, den der jeweils dritte Sensor liefert. Unter Kenntnis der Geometrie der Sensoranordnung kann man wie in Bild 9-40 gezeigt, den Wert für den Korrekturwinkel $\Delta\alpha$, um den um die jeweilige Achse gedreht werden muss, ermitteln. Damit sich wieder

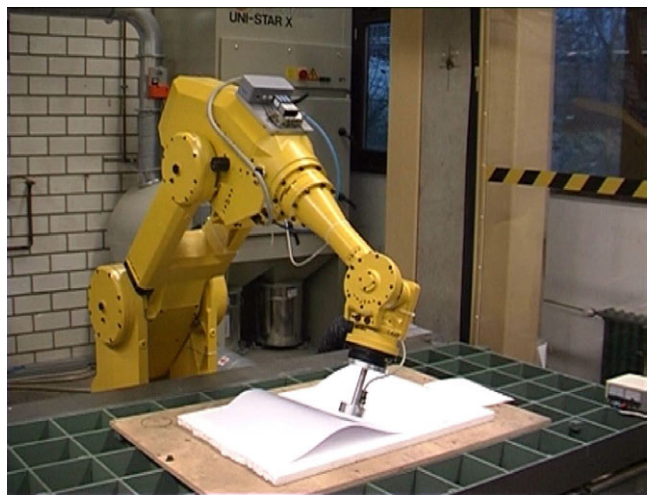


Bild 9-45: Fahren eines Industrieroboters über eine Freiformfläche mit Sensorführung

am Messpunkt der Soll-Abstand einstellt, muss $\Delta\alpha = 1,324^\circ$ je mm Abweichung vom Soll-Abstand betragen. Führt man hintereinander für alle drei Messpunkte eine Drehung um die gegenüberliegende Drehachse aus, so ist an allen drei Punkten wieder der Soll-Abstand eingestellt und die z-Achse des Tool-Koordinatensystems muss erneut senkrecht zur Freiformfläche orientiert sein. Aus den Messwerten wird auch der neue Verfahrvektor im Raum Δs mit Hilfe einer Koordinatentransformation, die Bestandteil des Sprachumfangs der Roboter-Software ist, berechnet.

Mit Hilfe dieses Konzeptes war es möglich, wie in Bild 9-45 gezeigt, in der Größenordnung der Prozessgeschwindigkeit einer Strahlbearbeitung mit konstantem Abstand und stets senkrechter Orientierung des Schneidstrahls über einer beliebig geformten Freiformfläche mit Krümmungsradien > 100 mm, eine vorgegebene geschlossene Sollbahn abzufahren. Veränderungen von Lage und Form der Freiformfläche können dabei vollständig ausgeglichen werden, wodurch der IR zu einer "intelligenten" Maschine geworden ist, die die Bezeichnung "mechatronisches System" zu recht trägt.

9.3.4.2 Zusätzliche Bewegungsachsen für Industrieroboter

Heute wird die Laser-Schneidtechnologie in vielen industriellen Branchen in zunehmenden Maß eingesetzt. Benutzt man als Bewegungseinrichtung für den Laser-Schneidkopf einen Industrieroboter, so lässt sich die 3-dimensionale Laserbearbeitung großvolumiger Blechteile, wie sie typischerweise im Automobilbau vorkommen, mit einem sehr wirtschaftlichen Anlagenaufwand durchführen. Vor allem Nd:YAG-Laser eignen sich gut für solche Schneidanlagen, da ihre Strahlung, anders als bei CO₂-Lasern, durch eine Glasfaser geführt werden kann. Diese lässt sich wiederum leicht durch einen Industrieroboter handhaben und dadurch kann der Schneidstrahl in dessen Arbeitsbereich 3-dimensionale Bewegungen ausführen.

Ein Problem ist dabei die Bahngenauigkeit bei kleinräumigen Konturen, wie beispielsweise bei Kreisen mit Durchmessern von 1 mm - 30 mm.

In der Prototypenfertigung von Automobilkarosserien werden heute vielfach Laserschneidanlagen eingesetzt, weil zu diesem Zeitpunkt die Stanzwerkzeuge, die die Karosseriebleche in der Serie herstellen, noch nicht verfügbar sind. Verwendet man zur Fertigung dieser großvolumigen Bauteile Laser-Schneidanlagen für 3-dimensionale Bearbeitung in der Art eines CNC-Bearbeitungszentrums, so stellt dies ein Investment oberhalb von 0,5 Mio. Euro dar. Kombiniert man jedoch den relativ preisgünstigen Nd:YAG-Laser mit einem Vertikal-Knickarmroboter als Strahlführungseinrichtung, so liegt man höchstens

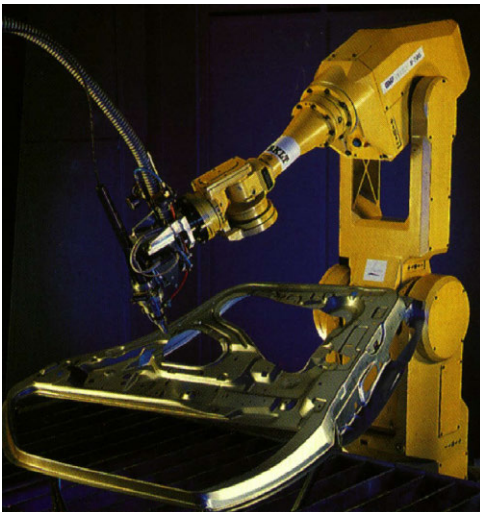


Bild 9-46: Knickarmroboter mit Ausrüstung zum Laserschneiden (Werkbild BKLT)

bei der Hälfte des genannten Preises (Bild 9-46). Dies ist deshalb möglich, weil die Strahlung des Nd:YAG-Lasers (Wellenlänge $1,06 \mu\text{m}$) sich mit vertretbaren Verlusten durch eine Glasfaser leiten lässt. Die Strahlung des in der Materialbearbeitung häufig eingesetzten CO_2 -Lasers (Wellenlänge $10,6 \mu\text{m}$) hingegen lässt sich nur durch Spiegel umlenken, da ihre Verluste in einer Glasfaser sehr hoch sind. Die Positionier- und Bahnengenauigkeit eines Industrieroboters der erforderlichen Tragkraft ($> 25 \text{ kg}$) liegt im Bereich der in der Blechbearbeitung erforderlichen Toleranz ($0,1 \text{ mm}$), ist jedoch im allgemeinen schlechter als bei den vorher genannten Strahlführungsmaschinen.

Dies liegt zum einen an der mangelnden Steifigkeit der kinematischen Kette eines 6-achsigen Knickarmroboters verglichen mit einem konventionellen Aufbau, ähnlich einer 5-achsigen Fräsmaschine. Ein Vergleich der Kinematik dieser beiden Einrichtungen ist in Bild 9-47 dargestellt. Während beim Knickarmroboter alle Achsen und deren kinematische Fehler hintereinander geschaltet sind, sind die 5 Bewegungsachsen des werkzeugmaschinenähnlichen Aufbaus in zwei Bewegungsgruppen mit maximal 3 Achsen aufgespalten. Zusätzlich

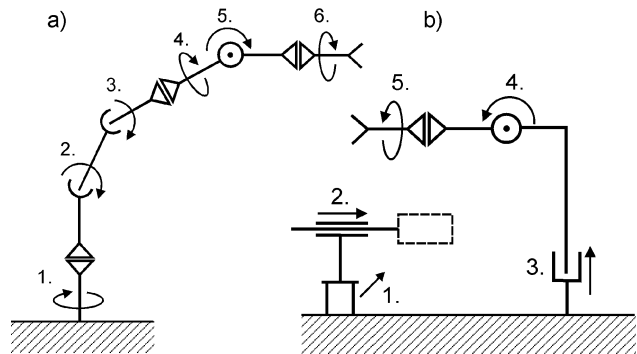


Bild 9-47: Kinematik unterschiedlicher Strahlführungseinrichtungen für das Laserschneiden a) 6-achsiger Knickarmroboter b) 5-achsige Bauweise in der Art einer CNC-Fräsmaschine

treiben die ersten beiden gekreuzten Linearachsen einen sehr steifen Werkstücktisch an, deren Steifigkeit deutlich höher als beim Knickarm liegt.

Ein zweiter Grund für die mangelhafte Bearbeitungsgenauigkeit von Kreisausschnitten kleinen Bahndurchmessers ist das bei Industrierobotern verwendete Verfahren der **Kreisinterpolation**. Dies ist ein Rechenverfahren, mit dem auf einer Bahn vorgegebener Form viele Bahnstützpunkte berechnet werden, zwischen denen sich der IR bewegen soll. Bei spanenden Werkzeugmaschinen wie einer Fräsmaschine wird die Kreisinterpolation zur Bearbeitung von Kreissegmenten in der eher niedrigen Vorschubgeschwindigkeit benötigt, so dass der Rechenvorgang für die Interpolation nicht zeitkritisch ist. Bei IR wird die Kreisinterpolation jedoch eher dafür benötigt, Punkte einer Bewegungsbahn bei großen Durchmessern mit Kreissegmenten zu verbinden, die auch mit hoher Bahngeschwindigkeit durchfahren werden müssen und nicht für kleine Vollkreise. In Bild 9-48 ist

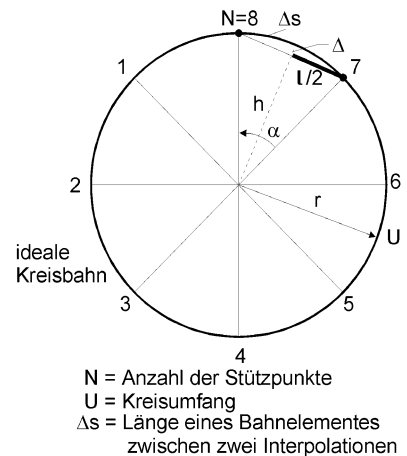


Bild 9-48: Aufteilung einer Kreisbahn für die Kreisinterpolation

ein Kreis mit Radius r und Umfang U dargestellt, der mit einer vorgegebenen Bahngeschwindigkeit v durchfahren werden soll. Aufgrund der durch das Interpolationsverfahren vorgegebenen Interpolationstaktzeit T wird der Umfang in eine feste Anzahl N von Bahnstützpunkten aufgeteilt. Die Länge eines Bahnelementes zwischen zwei Interpolationstakten Δs ergibt sich dann zu:

$$\Delta s = \alpha \cdot r \approx l = v \cdot T \quad \text{mit } l: \text{Länge der Sekante} \quad (9.28)$$

Der Bahnverlauf der interpolierten Bahn führt entlang der Sekante, da die Bahnstützpunkte auf dem Kreis mit Geradenstücken verbunden werden. Zwischen den schon bekannten Größen bestehen folgende Zusammenhänge:

$$U = 2\pi \cdot r = N \cdot \Delta s \quad (9.29)$$

$$l/2 = r \cdot \sin \alpha/2 \quad (9.30)$$

Durch das Ersetzen des Bogens durch die Sekante der Länge l tritt ein Radiusfehler Δ auf. Um diesen Fehler zu bestimmen berechnet man zuerst die Höhe h des Dreiecks mit Hypotenuse r und Kathete $l/2$:

$$h = \sqrt{r^2 - (l/2)^2} \quad (9.31)$$

Den Fehler Δ kann man dann errechnen:

$$\Delta = r - h \quad (9.32)$$

Führt man nun folgendes durch $[(9.30) \rightarrow (9.31)] \rightarrow (9.32)$, so erhält man eine Näherung für den Fehler:

$$\Delta = r - (1 - \cos \alpha/2) \approx r/2 \cdot (\alpha/2)^2 \quad \text{für } \alpha/2 < 1 \quad (9.33)$$

Setzt man nun noch Gl. (9.28) in Gl. (9.33) ein, so erhält man den Fehler in Abhängigkeit der durch die Bahnbewegung und den Interpolator vorgegebenen Größen:

$$\Delta = \frac{1}{2r} \left(v \frac{T}{2} \right)^2 \quad (9.34)$$

Die Anzahl der dabei verwendeten Bahnstützpunkte beträgt:

$$N = 2\pi \left(\frac{2}{v \cdot T} \right) \quad (9.35)$$

Im Bild 9-49 sind nun für eine Interpolationszeit $T=150$ ms die Verläufe der Funktionen in Gl. (9.34) und Gl. (9.35) dargestellt. Man sieht, dass beispielsweise bei einem Kreisdurchmesser von 20 mm und einer Bahngeschwindigkeit (Laserschneiden) von 30 mm/s bei der Interpolation nur wenige Stützpunkte gewählt werden können und dadurch leicht ein Radiusfehler von 0,3 mm auftreten kann.

Aus den beiden genannten Gründen treten daher beim Laserschneiden von kleinen Kreisen mit einem IR die in Bild 9-50 a) dargestellten Konturabweichungen von der idealen Kreisform auf, die bei Löchern von 20 mm Durchmesser im Bereich $\pm 0,5$ mm liegen können. Da aber kreisrunde Durchbrüche im Karosseriebau sehr häufig sind, gibt es bei Verwendung der genannten Anlagen Probleme bei der Fertigungs-

genauigkeit solcher Konturelemente. Da man als Anwender kaum Einfluss auf die Kinematik des IR oder auf sein Kreisinterpolationsverfahren nehmen kann, war das Ziel der hier beschriebenen Entwicklung, mit Hilfe einer Zusatzeinrichtung das Einbringen von kleinen, kreisrunden Löchern mit einer Durchmessertoleranz von $\pm 0,01$ mm zu ermöglichen.

Für das IR-Zusatzgerät **Kreisschneider** gab es mehrere Randbedingungen, die seine Konstruktion in hohem Maß beeinflussten [9.5], [9.6] :

- Das Gerät soll Durchmesser von 1mm bis 30 mm einstellen können und Schneidgeschwindigkeiten bis 30 mm/s ermöglichen.
- Die Einstellung muss programmgesteuert vorgenommen werden können und es müssen Zusatzfunktionen, wie das tangentielle Anfahren der Schneidkontur, möglich sein (Bild 9-50 b).

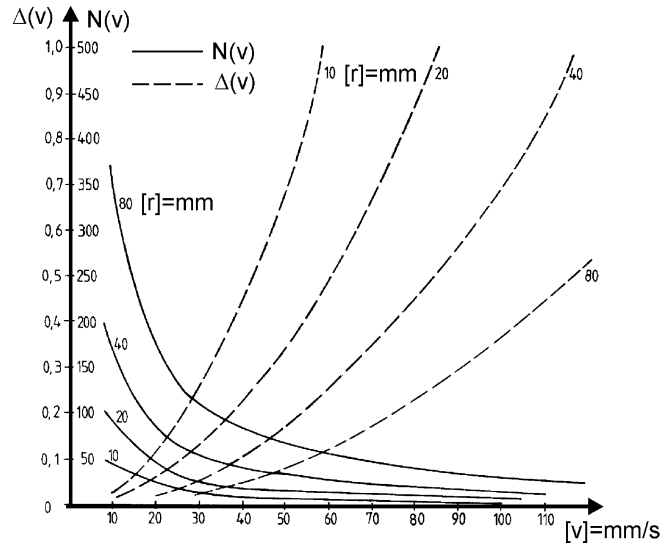


Bild 9-49: Radiusfehler Δ und Anzahl der Bahnstützpunkte N in Abhängigkeit der Bahngeschwindigkeit v bei unterschiedlichen Kreisradien r

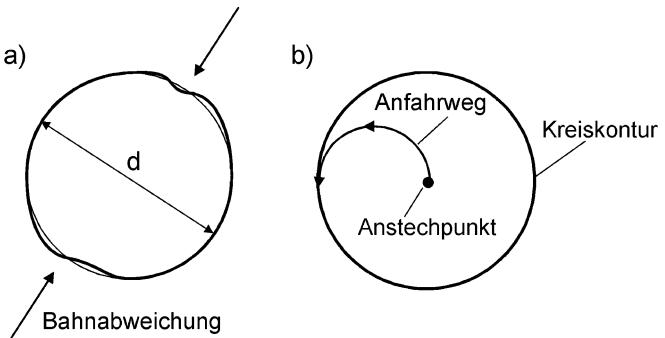


Bild 9-50: Laserstrahlschneiden von Kreisen a) Bahnabweichungen b) Bahnverlauf für Kreisschneidvorgang

- Der verwendete Roboter hat eine nominelle Tragfähigkeit von 30 kg. Daher muss das Gerät leicht und kompakt sein, weil das Gewicht des Schneidkopfes und eines sensorgeführten Stellantriebes für die Regelung des Fokusabstandes (Werkzeug Z-Achse, s. a. Kap. 9.3.4.1), sowie Zusatzkräfte, die von außen über die Glasfaser in das Handgelenk des Roboters eingeleitet werden, die Tragfähigkeit des Roboters schon zu etwa einem Drittel ausschöpfen. Diese schon vorhandenen Massen und Kräfte wirken sich durch das Zusatzgerät noch stärker aus, da es zwischen Roboter-Flanschplatte und Stellantrieb der Z-Achse montiert wird, und dadurch die Hebelarme vergrößert werden. Bild 9-51 zeigt die Anordnung des Kreisschneiders zwischen Roboterflansch und Schneidkopf.
- Bei Führung des Schneidkopfes auf einer Kreisbahn darf der Schneidkopf keine

Rotation um die Z-Achse ausführen, da sich sonst die Glasfaser verdrehen würde und somit die Gefahr der Beschädigung der Faser besteht.

- Die 2-dimensionale Bewegungsaufgabe des Zusatzgerätes muss in allen Raumlagen und Orientierungen des Schneidkopfes möglich sein. Das Drehmoment muss daher ausreichend groß sein, um das Gerät selber und den Schneidkopf mit Stellachse und Glasfaser gegen die Schwerkraft mit ausreichender Beschleunigung anheben zu können.

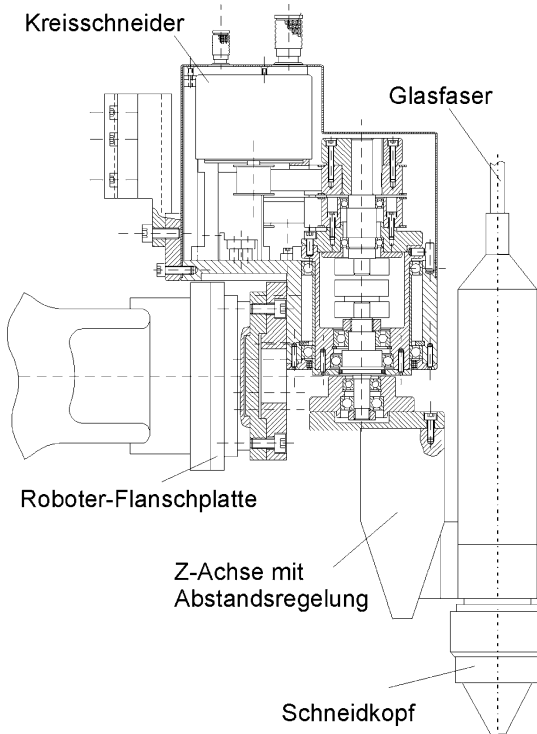


Bild 9-51: Schnitt durch den an der Roboterflanschplatte befestigten Kreisschneider mit Laser-Schneidkopf

Die Einbeziehung der letzten Rotationsachse des Roboters in die 2-dimensionale Führungsbewegung des Schneidkopfes musste von vornherein ausgeschlossen werden, weil dies die Bewegungsmöglichkeiten des Roboters zu stark eingeschränkt hätte. Der Kreisschneider musste daher über zwei im Funktionszusammenhang steuerbare Achsen verfügen.

Die einfachste Möglichkeit eine im Durchmesser verstellbare Kreisbahn zu erzeugen, ist die Verwendung einer Konfiguration aus einer Linear- und einer Drehachse. Diese scheidet aber von vornherein aus, da hierbei die mit dem Schneidkopf verbundene Glasfaser verdreht werden würde (Bild 9-52 a).

Die nächste Möglichkeit wäre die Verwendung zweier kartesischer Linearachsen, wobei hier die Kreisbewegung durch Interpolation erzeugt werden müsste (Bild 9-52 b). Diese Möglichkeit wurde ebenfalls

verworfen, da das Bauvolumen und Gewicht zweier Linearachsen für die Arbeitsaufgabe „Kreisschneiden“ ungünstig ist. Dies ergibt sich zum Teil daraus, dass der vorhandene Arbeitsbereich der Achsen bei einem Kreis nicht ausgenutzt wird. Zusätzlich ist hier für die Erzeugung der Kreisbahn eine Interpolation notwendig, die wiederum fehlerbehaftet sein kann.

Es wurde daher eine Kombination von zwei Rotationsachsen favorisiert, bei der der Kreisschneidvorgang durch Drehung einer Rotationsachse sehr exakt ausgeführt werden kann (Bild 9-52 c). Damit dabei aber nicht die Glasfaser verdreht wird, müssen die beiden Rotationsachsen auf eine Exzenter-Einrichtung wirken, die den Schneidkopf ohne Drehung um die Z-Achse auf einer Kreisbahn führt und bei der die Durchmesserstellung durch Verstellung der Exzentrizität erfolgt.

Das gesamte System, dessen Getriebschema in Seitenansicht in Bild 9-53 darge-

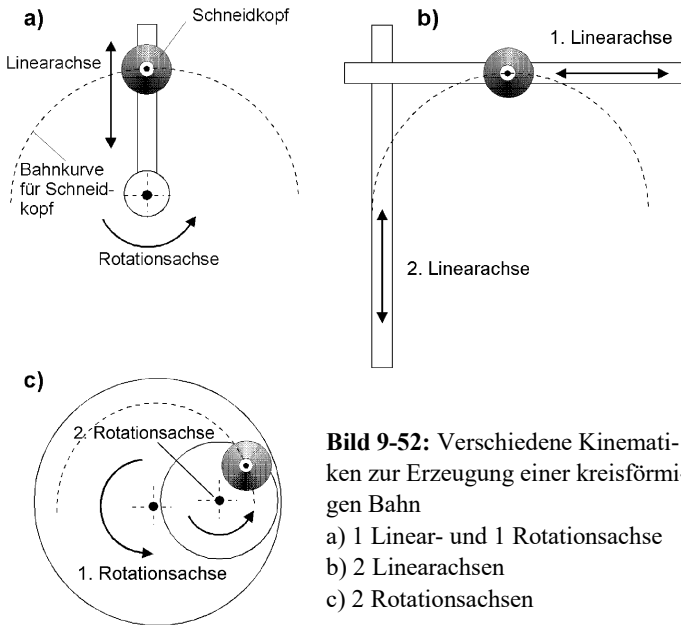


Bild 9-52: Verschiedene Kinematiken zur Erzeugung einer kreisförmigen Bahn

- a) 1 Linear- und 1 Rotationsachse
- b) 2 Linearachsen
- c) 2 Rotationsachsen

stellt ist, ist als Doppelenxentersystem ausgeführt. Es besteht aus zwei symmetrischen Teilsystemen mit jeweils einer drehbar gelagerten Hohlwelle als erste Rotationsachse und der innen in der Hohlwelle gelagerten Exzenterwelle als zweite Rotationsachse. Die jeweils entsprechenden Achsen in den Teilsystemen sind durch Zahnriemen miteinander verbunden, so dass sie sich immer synchron bewegen. Wie im Bild 9-54 in der Draufsicht erkennbar, besitzen die beiden Verstellexzenterwellen

jeweils wieder einen auf ihnen exzentrisch angeordneten Zapfen, die durch ein Verbindungsstück starr miteinander verbunden sind. An diesem Verbinders wird der Schneidkopf befestigt. Um den Achsversatz zwischen der Antriebsseite der Exzenterwelle und ihrer Abtriebsseite zu überbrücken, ist jeweils eine Parallelenkerkuppung in die Exzenterwelle eingefügt. Der Zapfen am Ende der Exzenterwellen ist auf einem Teilkreis angeordnet, dessen Radius genau dem Achsversatz von An- und Abtriebsseite der Exzenterwelle entspricht. Daher kann die Exzenterwelle in der Hohlwelle eine Position (grün) annehmen, die einem zu erzeugenden Kreis mit Durchmesser 0 mm entspricht. Werden aus dieser Stellung heraus Hohl- und Exzenterwelle mit gleicher Drehzahl angetrieben, so bleibt der Schneidkopf in seiner Position ohne sich zu bewegen. Bewegt man aus dieser „Nullposition“ heraus die beiden Rotationsachsen kurzzeitig mit unterschiedlicher Drehzahl, so findet eine

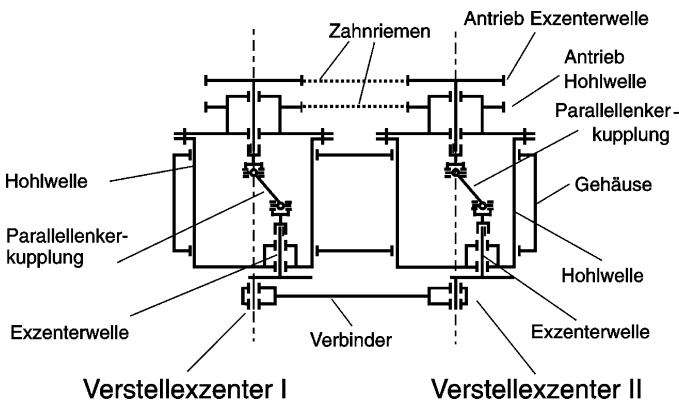


Bild 9-53: Getriebeschema des Kreisschneiders mit zwei Verstellexzentern

Verdrehung von Hohl- und Exzenterwelle zueinander statt, wodurch sich die Zapfen der Exzenterwellen aus der Mitte der Hohlwelle nach außen bewegen. Dies entspricht einer Durchmesserstellung des Kreisschneiders, der nach einer halben Drehung der Exzenterwellen zum größten Durchmesser (blau) der Schneidbahn führt. Treibt man danach

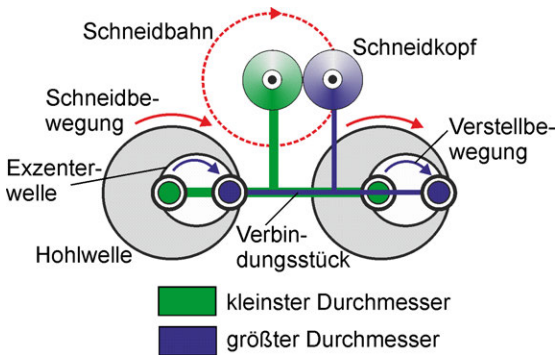


Bild 9-54: Durchmessereinstellung und Kreisbewegung der Exzentervorrichtung

Es wurden daher Schrittmotoren mit 1/10 Schritt-Verstellung verwendet und zur Erhöhung des Drehmomentes eine Drehzahluntersetzung zu den Rotationsachsen über Zahnriemen vorgenommen. Um für die geforderte Schneidgeschwindigkeit von 30 mm/s auch bei kleinen Durchmessern eine ausreichend hohe Drehzahl der Wellen erzeugen zu können, wird eine maximale Schrittfrequenz von 100 KHz verwendet; die Auflösung liegt bei 13 Schritten/ Grad Verdrehung der Rotationsachsen. Die Schrittmotoren werden von einer eigenständigen Schrittmotorsteuerung gesteuert und mit der Antriebsleistung versorgt. Die Nullpunktlage von Hohl- und Exzenterwellen wird durch feinjustierbare Mikroinitiatoren vorgegeben. Die Schneidparameter wie Kreisdurchmesser, Schneidgeschwindigkeit, maximaler Drehwinkel und Bahnform zwischen Anstich und Kreiskontur, sowie das Startsignal der Bewegung, werden von der Robotersteuerung über Binäreingänge an die Schrittmotorsteuerung übergeben. Da die beiden Rotationsachsen in einem beliebigen Funktionszusammenhang bewegt werden können, können natürlich auch andere Bahnverläufe als Kreise, etwa Ellipsen, Langlöcher oder sogar Rechtecke erzeugt werden. Dazu bedarf es dann spezieller Interpolationsalgorithmen.

Im Bild 9-55 sieht man die komplette Installation des Kreisschneiders am Arm des Roboters. Er trägt an seiner Vorderseite den Laser-Schneidkopf mit dem Lichtwellenleiter und der zusätzlichen Linearachse für die Fokus-Abstandsregelung. Die Arbeitsweise der Einrichtung ist nun so,

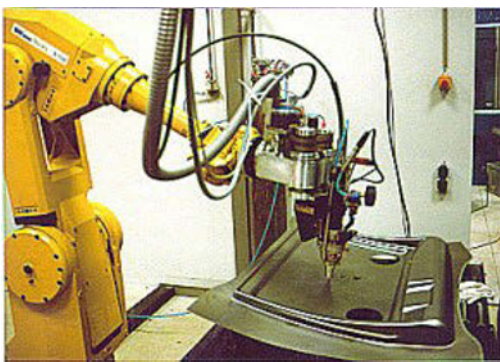


Bild 9-55: IR mit Kreisschneider und Laser-Schneidkopf

Hohl- und Exzenterwelle wiederum mit gleicher Drehzahl an, so beschreibt der Schneidkopf, der über den Verbinders mit den Exzenterwellen verbunden ist, eine Kreisbahn, deren Radius dem Abstand der Exzenterzapfen von der Mitte der Hohlwelle entspricht. Der Schneidkopf führt dabei, entsprechend der oben genannten Forderung, keine Drehung um die Z-Achse aus.

Das Antriebssystem sollte möglichst einfach und leicht sein, bei gleichzeitig ausreichendem Drehmoment.

beitsweise der Einrichtung ist nun so, dass der Roboter den am Kreisschneider befestigten Schneidkopf an der Einstichstelle des Bleches positioniert, den Laser einschaltet und den Bewegungsbefehl an den Kreisschneider ausgibt. Dieser führt dann bei stillstehendem Roboter die Schneidbewegung für das zu erzeugende Konturelement aus und gibt bei Beendigung der Bewegung eine Rückmeldung an den Roboter. Danach führt der Roboter dann die weiteren Bewegungsaufgaben durch.

Mit der hier beschriebenen Einrichtung können mit einem Laser-Roboter auch Kreise der geforderten Toleranz geschnitten werden. Durch diese mechatronische Lösung ist es möglich mit einer viel kostengünstigeren Anlagenlösung als einer konventionellen 3 D-Laseranlage mit gleicher Genauigkeit zu arbeiten. Zur Entwicklung des Gesamtkonzeptes waren umfangreiche Kenntnisse aus vielen Bereichen des Maschinenbaus und der Elektrotechnik erforderlich, die einem Ingenieur, der nur in seiner Fachdisziplin ausgebildet ist, nicht zur Verfügung stehen. In Aufgabenbereichen, die an der Lösung solcher Problemstellungen arbeiten, findet der Mechatronik-Ingenieur beispielsweise sein Arbeitsfeld.

9.4 Mechatronische Gesamtsysteme

9.4.1 Hexapodenkonzepte

Neben den weit verbreiteten Kinematiken des Knickarmroboters (Bild 9-37), bei dem alle Bewegungsachsen in Reihe hintereinander geschaltet sind, ist in den letzten Jahren für Anwendungen, die eine hohe mechanische Steifigkeit erfordern, die sogenannte Parallelkinematik entwickelt worden. Bei einem solchen Roboter (Bild 9-56), der ebenfalls sechs Freiheitsgrade besitzt, sind die 6 Bewegungsachsen parallel angeordnet, was die Tragfähigkeit, die Möglichkeit der Momentausübung für Werkzeuganwendungen und die Steifigkeit der kinematischen Struktur deutlich erhöht.

Eine solche auch als **Hexapode** (Sechsfüßler) bezeichnete Parallelkinematik besteht prinzipiell aus zwei gegeneinander um 60° gedrehten, übereinander liegenden Dreiecken, bei denen jede Ecke mit zwei benachbarten unterhalb liegenden Ecken verbunden ist (Bild 9-57). Diese Verbindungen haben eine variable Länge, so dass das obere Dreieck sechs Freiheitsgrade erhält. Es kann sich heben und senken, neigen und drehen. Diese kinematische Struktur besitzt auch der in Bild 8-49 dargestellte „Tripod“.

Durch die Querverstrebungen wird die Hexapode-Struktur zu einem sehr stabilen Gebilde und kann bei definierten „Beinlängen“ nur genau eine Position im Raum einnehmen. Üblicherweise ist das untere Dreieck am Boden verankert und wird Basis genannt, während das obere, frei bewegliche Dreieck den Endeffektor trägt. Das Hauptanwendungsgebiet heutiger Hexapode-Strukturen liegt im Bereich der Werkzeugmaschinen und Simulatoren.

Normalerweise sind die Stellelemente zur Veränderung der Beinlänge wie im Beispiel in Bild 9-56 Spindel/Mutter-Systeme, die von einem Elektromotor angetrieben werden. Die Idee der im folgenden dargestellten Entwicklung [9.7] „Doppel-Hexagone-Struktur basierend auf Fluidmuskeln“ hatte zum Ziel, den Aufwand für



Bild 9-56: IR (Typ F-204i) mit Parallelkinematik (Werkbild Fanuc Robotics)

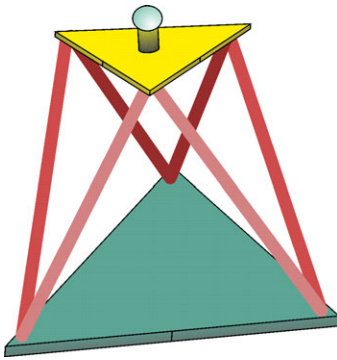


Bild 9-57: Schematische Darstellung der Hexapodenstruktur

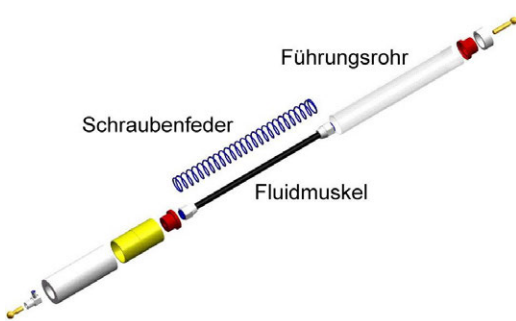


Bild 9-58: Antriebskomponente für Hexapoden aus Fluidmuskel, Druckfeder und Teleskopführung

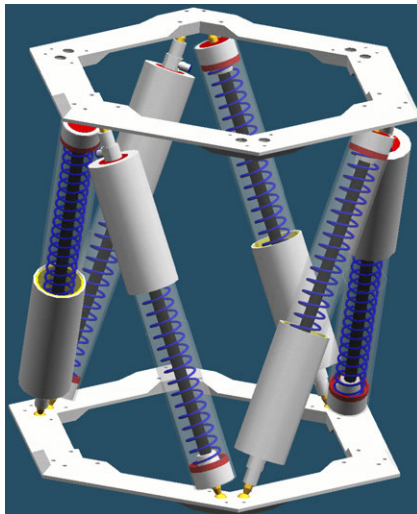


Bild 9-59: Hexapodenstruktur mit 6 fluidischen Antriebseinheiten

eine Hexapodenstruktur mit Hilfe neuartiger Aktoren soweit zu senken, dass man mehrere Hexapodenelemente hintereinander schalten kann. Übereinander montiert, könnten auch verwinkelte Ecken erreicht oder Hindernisse umgangen werden. Ein Vergleich dieser Roboterstruktur mit der Natur lässt hoffen, sich der Flexibilität und Bewegung einer Schlange nähern zu können.

Antreibendes Element der Hexapodenstruktur sollte der in Kapitel 6.2 beschriebene neuartige Aktor "Fluidmuskel" sein, ein druckluftbeaufschlagter Zugaktor, der dem biologischen Muskel nachempfunden ist und im Vergleich zum Pneumatikzylinder hohe Anfangszugkräfte erzeugen kann. Damit ist sowohl ein

kraftvolles Beschleunigen als auch ein sanftes Einfahren in die Sollposition möglich. Die Zugkraft hat ihr Maximum zu Beginn der Kontraktion und fällt nahezu linear mit dem Hub auf Null ab. Da der Fluidmuskel ein reiner Zugaktor ist, kann er keine Druckkräfte übertragen und verfügt über keine Führung. Weil der Fluidmuskel nur Kraft bei Kontraktion aufbringt, muss er mit einer Druckfeder gepaart werden, die die benötigte Kraft zur Expansion liefert. In der Regel ist der Fluidmuskel langgestreckt, wodurch auch die verwendeten Schraubenfedern relativ lang sind und bei Kontraktion zur

Knickung neigen. Daher wurde zur Führung der Feder ein umschließendes Teleskoprohr verwendet, in das, um Reibungsverluste zu minimieren, ein Kunststoff-Lineargleitlager eingesetzt wurde. Die gesamte Antriebseinheit für eine Achse des Hexapoden zeigt das Explosionsbild Bild 9-58. Der Fluidmuskel liegt im Inneren und wird von der Druckfeder umschlossen. Ein Plexiglasrohr umschließt die Feder, so dass diese nicht knicken kann und der innere Mechanismus sichtbar bleibt.

Setzt man nun sechs dieser Antriebselemente mit 2 sechseckigen Kopf- und Basiselementen zusammen, so entsteht ein Hexagone (Bild 9-59). Da die Antriebselemente an den Verbindungsstellen zu den Verbindungsrahmen gedreht und geneigt werden müssen, werden an diesen Stellen Kugelgelenke verwendet. Das

ganze System ist symmetrisch aufgebaut, so dass man mehrere Systeme hintereinander montieren kann.

Steuert man nun die fluidischen Muskeln mit einem Drucksignal an, so führen diese eine Längenänderung zusammen mit den anderen Bauelementen aus. Die Kennlinie dieser Antriebseinheit muss aber nicht unbedingt einen linearen Verlauf haben.

Bild 9-60 zeigt den Zusammenhang zwischen dem Eingangsdruck in den fluidischen Muskel und der Längenänderung des Antriebsstrangs. Die fluidischen Muskeln im Antriebsstrang haben dabei eine Nennlänge von 350 mm und können eine Maximalkraft von 400 N aufbringen. Die Kennlinien wurden für eine Antriebseinheit auf einem Prüfstand aufgenommen und zeigen eine Hysterese für das Kontrahieren und Entspannen der Antriebseinheit. Im Diagramm ist ebenfalls eine gemittelte Gesamtfunktion zu erkennen, die für die Wegansteuerung der Hexapodenstruktur verwendet wird. Diese Funktion wurde durch folgendes Polynom approximiert:

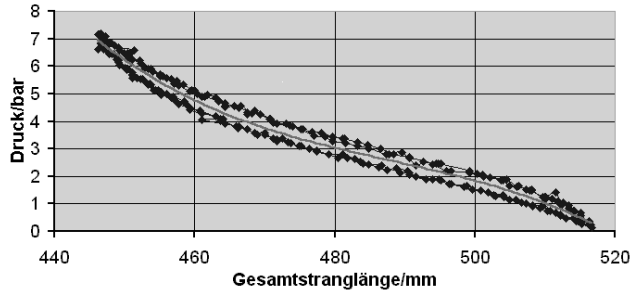


Bild 9-60: Längenänderung eines Antriebsstrangs in Abhängigkeit vom Ansteuerdruck (Messwerte und gemittelte Gesamtfunktion)

$$f(x) = 28,3 \cdot 10^{-6} x^3 + 41,5 \cdot 10^{-3} x^2 - 20,3x + 3327,5 \quad \text{mit } 455 < x < 516$$

Damit die Struktur auch bei unterschiedlichen Stranglängen gleichmäßig fährt, z.B. bei der Einstellung des maximalen Drehwinkels einer Elementebene, wird die Ventilsteuerung pulsweitenmoduliert. Die Steuerventile können über ein „Speed Up“-Modul mit einer Frequenz von bis zu 200 Hz geschaltet werden. Über die Frequenz wird die Fahrgeschwindigkeit der Hexagone-Struktur geregelt und über die Pulsweite der Luftdurchfluss. Muss der Fluidmuskel eine große Entfernung überbrücken, wird sein Druck schnell auf oder abgebaut. Dazu wird unter Beibehaltung der Frequenz die Pulsdauer verlängert. Die Vorgabe für die Druck-Sollwerte der einzelnen Antriebsstränge erfolgt aus einem Digitalrechner, auf dem auch die erforderliche Koordinatentransformation durchgeführt wird, die für das Einnehmen einer be-

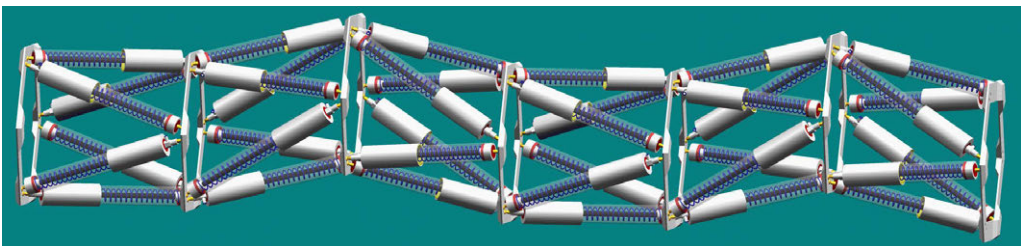


Bild 9-61: Hexapodische Schlange aus 6 Hexapod-Strukturen

stimmten Raumlage des oberen Endes der Hexapoden-Struktur die erforderlichen Längen aller Antriebsstränge ermittelt. Die mathematischen Zusammenhänge zur Bestimmung der Stranglängen in Abhängigkeit der Raumposition des Endeffektors sind in [9.8] angegeben

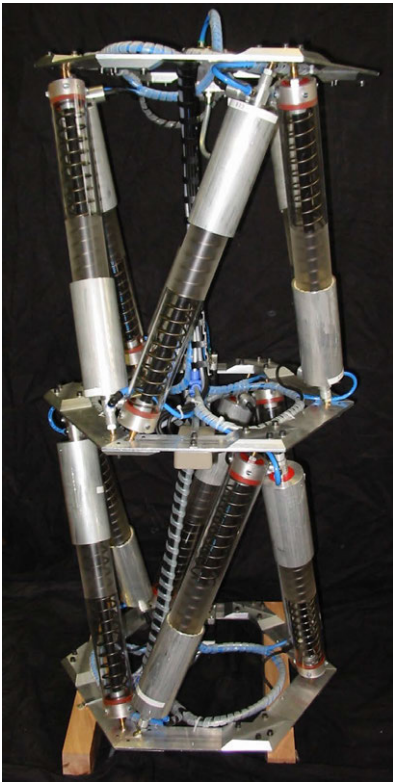


Bild 9-62: Doppel Hexapoden-Struktur

Wie oben erwähnt, war es das Ziel dieser Entwicklung, eine ganze Kette solcher Hexapoden-Systeme hintereinander zu schalten, um beispielsweise einen ganzen Greifarm daraus zu bilden (Bild 9-61). Ein Greifarm kann durch die modulare Aufbauweise aus fünf oder sechs aneinander gereihter Elemente bestehen. An einer Wand oder der Decke befestigt, ist es ein Leichtes verwinkelte Arbeitsbereiche zu erschließen.

Eine weitere Anwendungsmöglichkeit wäre die autonome Überprüfung verwinklter Rohrsysteme. Hierbei würden zehn oder mehr Hexagone-Strukturen hintereinander gekoppelt werden. Ausgerüstet mit einer in einem Mikrocontroller verankerten künstlichen Intelligenz entstünde ein schlangenartiges Mobil. Mit Werkzeug und einer Kamera ausgerüstet, wäre es in der Lage selbstständig nach Fehlern in Rohrsystem suchen und diese zu reparieren. Selbst dünne senkrecht verlaufende Rohre zu erkunden würde kein Problem darstellen. Das Mobil verkantet sich mit der hinteren Hälfte und schiebt die vordere Hälfte weiter nach oben. Die federnde Eigenschaft der Fluidmuskeln sorgt hierbei für den Anpressdruck zu den Seiten. Bei herkömmlichen Hexagone-Strukturen mit Linearantrieben müsste eine sehr präzise Kraftüberwachung integriert werden.

Bis jetzt tatsächlich realisiert wurde eine Hintereinanderschaltung von zwei Hexapoden-Strukturen (Bild 9-62). Das mathematische Problem der Stranglängenberechnung für viele hintereinander geschaltete Hexapoden-Strukturen wird natürlich immer komplexer. Da schon beim Doppel-Hexapoden die Basis des oberen Elementes nicht mehr ortsfest ist, benötigt der Transformationsalgorithmus gegenüber dem beim unteren, an der Basis befestigten System, zusätzliche Eingangsinformationen.

9.4.2 Fahrrad mit aktiver Neigetechnik

An der Hochschule Bochum sind mehrere Forschungsfahrzeuge zu dem Zweck entwickelt worden, mechatronische Fahrwerkskomponenten in der Praxis erproben zu können und den Studierenden der Mechatronik die Gelegenheit zur Anwendung moderner Entwicklungsarbeit zu geben [9.10]. Der Grundgedanke des Projektes ist die Entwicklung von schnellen, wendigen Fahrzeugen auf drei Rädern, die durch einen aktiven Kurvenneige-Mechanismus zum einen den Fahrkomfort, zum ande-

ren aber auch die Sicherheit bei Kurvenfahrten erhöhen. Die Namen der zwei Versuchsfahrzeuge VIRAGE2000 und VIRAGETTE sind aus dem Französischen abgeleitet und bedeuten die Kurve.

In den letzten Jahren haben weltweit die Aktivitäten in der Konstruktion von muskel- und motorgetriebenen Neigefahrzeugen unterschiedlicher Fahrwerksausprägung zugenommen. Diesen Trend belegen Beispiele über das gesamte Spektrum der Anwendungen, im Bereich der Konzeptstudien für Motorfahrzeuge der Mercedes-Benz LifeJet F300, sowie eine Dreiradstudie von BMW, im Bereich der Serien-Liegeräder das deutsche Tripendo (Bild 9-63). Allen Konzepten liegt mit unterschiedlicher Gewichtung der Wunsch nach Fahrspaß und Wendigkeit, aber auch nach ökonomischem Transportnutzen zugrunde.

Um komplexe mechatronische Systeme zu entwickeln benötigt man auch eine neue Entwicklungsmethodik, um in überschaubaren Zeiträumen zu guten Ergebnissen zu kommen. In [8.8] und [9.11] wird das sogenannte V-Modell als geeignete Methode für die den Entwurf mechatronischer Systeme (Bild 9-64) vorgestellt. Die im Bild genannten Begriffe haben folgende Bedeutung:

- **Anforderungen:** Ausgangspunkt bildet ein konkreter Entwicklungsauftrag. Die Aufgabenstellung wurde präzisiert und in Form von Anforderungen beschrieben. Diese Anforderungen bilden den Maßstab, anhand dessen das spätere Produkt zu bewerten ist.
- **Systementwurf:** Ziel ist die Festlegung eines fachgebietsübergreifenden Lösungskonzeptes, das die wesentlichen physikalischen und logischen Wirkungsweisen des zukünftigen Produktes beschreibt. Hierzu wird die Gesamtfunktion eines Systems in wesentliche Teilfunktionen zerlegt. Diesen Teilfunktionen werden geeignete Strukturen zugeordnet und die Funktionserfüllung wird im Systemzusammenhang geprüft.
- **fachgebietsübergreifender Entwurf:** Auf der Basis dieses gemeinsam entwickelten Lösungskonzeptes erfolgt die weitere Konkretisierung meist getrennt in den beteiligten Fachgebieten. Detaillierte Auslegungen und Berechnungen sind nötig, um insbesondere bei kritischen Funktionen die Funktionserfüllung sicherzustellen.
- **Systemintegration:** Die Ergebnisse aus den einzelnen Fachgebieten werden zu einem Gesamtsystem integriert, um das Zusammenwirken untersuchen zu können.



Bild 9-63: Beispiele für Versuchsdreiräder mit Motor- und Muskelantrieb: v. o. Mercedes-Benz LifeJet F300, BMW-Studie, Tripendo

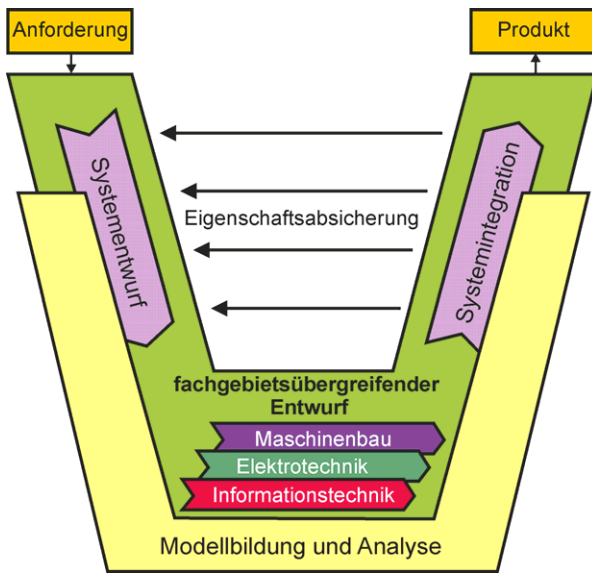


Bild 9-64: V-Modell für den Entwurf mechatronischer Systeme

lich das fertige, real existierende Erzeugnis verstanden, sondern die zunehmende Konkretisierung des zukünftigen Produktes. Reifegrade sind dabei das Labormuster, das Funktionsmuster, das Vorserienprodukt, usw.

Der oben aufgezeigt Entwicklungszyklus ist dann unter Umständen für jeden Reifegrad des Produktes erneut zu durchlaufen.

- **Eigenschaftsabsicherung:** Der Entwurfsfortschritt muss fortlaufend anhand des spezifizierten Lösungskonzepts und der Anforderungen überprüft werden. Es ist sicherzustellen, dass die tatsächlichen mit den gewünschten Systemeigenschaften übereinstimmen.
- **Modellbildung- und Analyse:** Die beschriebenen Phasen werden durch die Abbildung und Untersuchung der Systemeigenschaften mit Hilfe von Modellen und rechnerunterstützten Werkzeugen flankiert.
- **Produkt:** Ergebnis eines durchlaufenen Entwicklungszyklus ist das Produkt. Dabei wird unter Produkt nicht ausschließlich

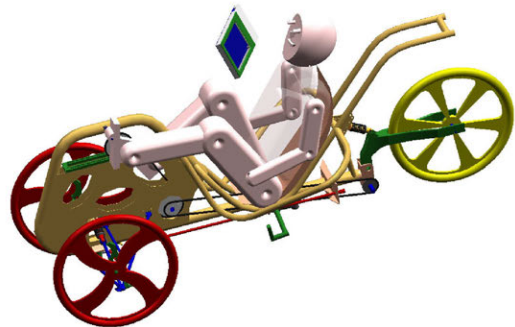
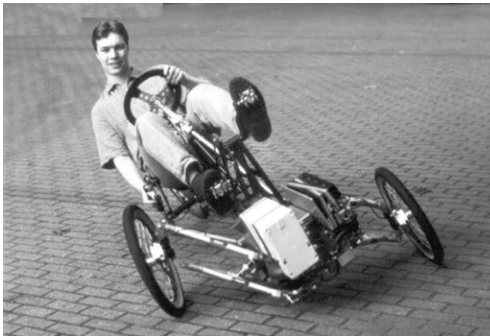


Bild 9-65: Dreiräder mit Neigetechnik, links Prototyp VIREAGE2000, rechts 3D-Modell VIRAGETTE

Es wurden nacheinander zwei Entwicklungsstufen des Dreirades mit Neigetechnik erstellt (Bild 9-65). Beim ersten Prototyp, dem VIRAGE2000, wurde die Produkterstellung noch weitgehend seriell durchgeführt. Nach der Formulierung der Produktziele folgte die Konstruktion mittels 2D-CAD. Danach wurde das Fahrzeug mit auf Prototypfertigung ausgerichteter Technologie gefertigt. Die Entwicklung des Neigesystems und dessen Integration in das fertige Fahrzeug erfolgte zuletzt, wobei vorgegebene Schnittstellen nur mit großen Mühen veränderbar waren. Bereits hierbei

wurde jedoch schon mittels der Modellbildung und Simulation von Teilkomponenten des Neigefahrzeugs eine Entwurfsiteration durchgeführt, die zu einer stark verkürzten Entwicklungszeit führte.

Mit Hilfe moderner mechatronischer Entwicklungsumgebungen konnte die Produktentwicklung des zweiten Dreirades VIRAGETTE in Form eines integrierten Entwurfsprozesses geschehen. Die mechanische Konstruktion wurde vollständig im 3D-CAD System **Unigraphics Solide Edge** erstellt, wobei Kollisionskontrolle, Anpassung an Körpermaße des Fahrers und fertigungsgeechtes Design ermöglicht wurden. Durch die durchgängige Parametrierung des konstruierten Systems ist noch bis zum Zeitpunkt der Fertigung eine Variantenkonstruktion durchführbar. Für die dynamische Simulation des 3-D-Modells wurde das System **CamelView** verwendet. Hierin konnte auch eine Aktoranpassung über einen Hardware-in-the-Loop Aufbau erfolgen. Spezielle Voruntersuchungen des gewählten Aktorverhaltens wurden mittels **WINFACT/Boris** blockorientiert simuliert.

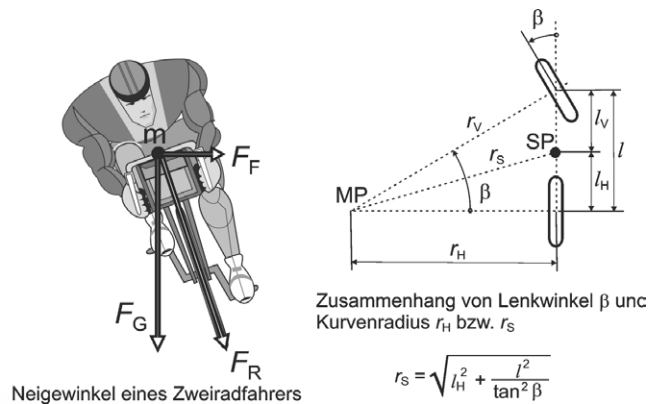


Bild 9-66: Neigung eines Fahrrades bei kräftefreier Kurvenfahrt und Zusammenhang zwischen Kurvenradius und Lenkwinkel

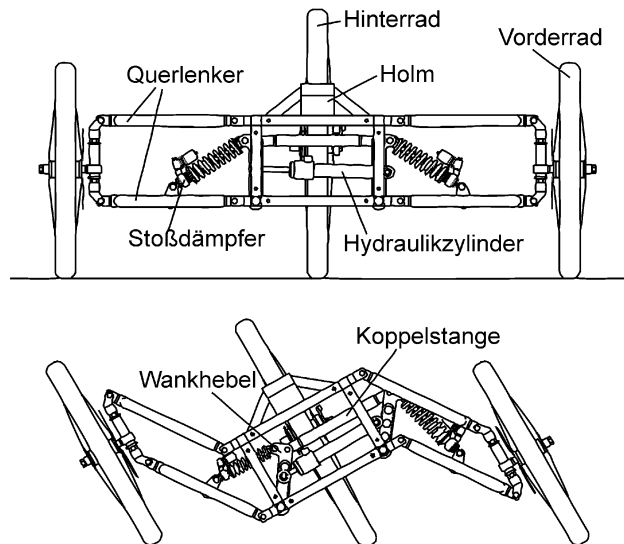


Bild 9-67: Neigung der Vorderräder des Virage2000 über Hydraulikzylinder und Koppelstange

Das Ackermann-Fahrzeugmodell [9.12] gibt den kräftefreien Fall der Kurvenfahrt wieder (Bild 9-63). Dabei ist MP der Momentanpol der Kurvenbewegung und SP der Schwerpunkt des Rades. Die Indizes V und H der Maße stehen für Vorder- bzw. Hinterrad. Diese rein kinematische Modell war Grundlage für die Ableitung des Kurvenneigalgorithmus des VIRAGE 2000. Aus dem Kräfteparallelogramm aus Fliehkraft F_F und Gewichtskraft F_G kann der Neigewinkel errechnet werden.

Um zu erreichen, dass sich die beiden Vorderräder des VIRAGE2000 bei Kurvenfahrt neigen, wurde ein Stellmechanismus über einen Hydraulikzylinder vorgesehen

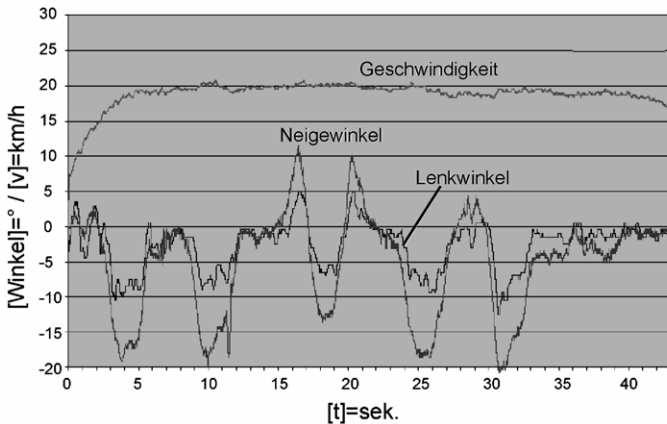


Bild 9-68: Verlauf von Geschwindigkeit, Neigewinkel und Lenkwinkel bei einer Testfahrt

(Bild 9-67). Die Bewegung des Zylinders wird über die Wankhebel und die daran befestigten Stoßdämpfer auf die Querlenker übertragen und hebt bzw. senkt eines der Vorderräder. Gleichzeitig wird über die Koppelstange das zweite Vorderrad entgegengesetzt bewegt. Zusammen ergibt sich daraus die Neigung des gesamten Fahrzeugs.

Mit einem normalen Zweirad wurde der Zusammenhang zwischen Lenkwinkel, Geschwindigkeit und Radneigung auf Versuchsfahrten gemessen. Die Fahrstrecke bildete die normalen Fahrsituationen ab. Die mit einem Microcontroller aufgezeichneten Fahrdaten ergaben folgende Erkenntnisse (Bild 9-68):

- Es besteht ein klarer Zusammenhang zwischen Geschwindigkeit, Lenkwinkel und Neigewinkel.
- Die zeitliche Verschiebung zwischen dem Lenk- und dem Neigewinkelverlauf ist klein.

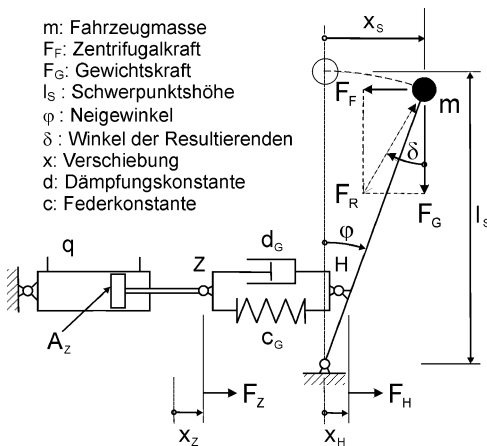


Bild 9-69: Zweidimensionales Ersatzmodell des Fahrzeugs

Mit diesen Erkenntnissen wurde ein physikalisches Ersatzmodell des zweidimensionalen Fahrzeuges erstellt (Bild 9-69). Das sehr einfache mechanische Modell des neigenden Fahrzeugs berücksichtigt weder die Masse der Achse und Querlenker, noch die genaue nichtlineare Charakteristik. Dennoch konnten aus der Simulation nützliche Erkenntnisse über das Systemverhalten gewonnen werden.

Der Hydraulikzylinder, der als Stellglied der Neigeeinrichtung dient, wird von einer Pumpe-/Speicherkombination über ein elektrisches Proportionalventil angesteuert (Bild 9-70). Als Energiequelle der Pumpe dient eine Nickel-Cadmium Batterie. Um die Pumpe, den Speicher und das Ventil

dimensionieren zu können, wurde das System einschließlich des Ersatzmodells der Neigeeinrichtung (Bild 9-69) mit dem Simulationssystem **DSH+** modelliert. Auf dieses Modell wurde als Eingangsgröße ein spezieller erforderlicher Fahrzyklus aufgegeben, so dass die Speichergröße und der Lageregler des Zylinders dimensioniert

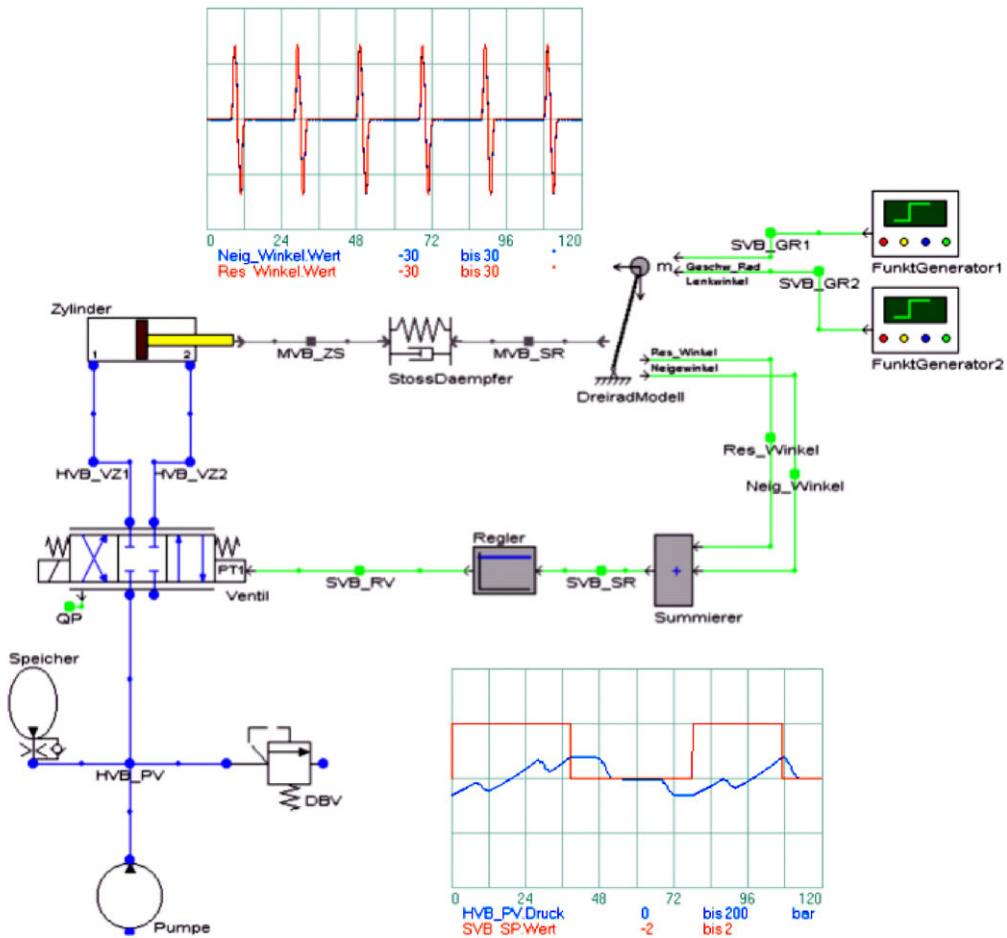


Bild 9-70: Simulation des Neigeverhaltens insbesondere des Hydrauliksystems mit DHS+

werden konnten.

Die Ergebnisse der Simulation des Neigesystems wurden als Block in das Fahrzeugmodell integriert. Dieses Gesamtmodell umfasste nur die Dynamik des Neigevorgangs und wurde mit dem Simulationssystem **Winfact 98** modelliert. Mit Hilfe dieses Modells konnten einige Fahrexperimente ähnlich der anfänglichen Messfahrt mit einem normalen Fahrrad durchgeführt werden, in denen Standardsituationen und Sicherheitsaspekte getestet werden konnten.

Für die Messfahrten wurde der Mikrocontroller **Basic-Tiger** aus dem Automotive-Bereich benutzt. Der multi-tasking-fähige Controller kann digitale und analoge Messwerte mit einer Zykluszeit entsprechend 80 Hz verarbeiten. Diese Arbeitsgeschwindigkeit reichte für die hydraulische Lageregelung aus. Die Regelung der Pumpe in Abhängigkeit vom Druck im Speicher und das Batteriemanagement waren weitere Aufgaben des Controllers.

Der Algorithmus zur Vorgabe der entsprechenden Verschiebung des Hydraulikzylinders basiert, wie in Bild 9-71 dargestellt, auf einem Polynom, das geschwindig-

keitsabhängig geändert wird und als Eingangsgröße den Lenkwinkel verwendet.

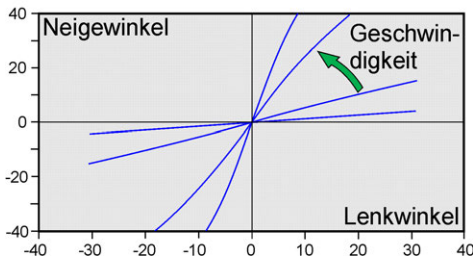


Bild 9-71: Vorgabe des Neigungswinkels in Abhängigkeit der Geschwindigkeit und vom Lenkwinkel

Das Fahrrad VIRAGE2000 (Bild 9-65 links) wurde dann aus käuflichen Standardbauteilen und aus werkstattgefertigten Aluminiumbauteilen als geschweißter Prototyp gefertigt. Ausgiebige Probefahrten zeigten die Funktionsfähigkeit des Neigemechanismus und der gesamten Konstruktion.

Das zweite Fahrrad VIRAGETTE (Bild 9-65 rechts) wurde direkt als 3D-CAD-Entwurf entwickelt und die Fertigung von einem Serienhersteller von Fahrrädern mit einer bedeutend leichteren Rohrkonstruktion durch-

geführt (Bild 9-72).



Bild 9-72: Rahmenaufbau des VIRAGETTE-Neigefahrrads

Der mit den Stoßdämpfern verbundene Neigemechanismus musste, anders als beim Vorgänger, aufgrund des schmalen Rahmens, seitlich durch Seilzüge betätigt angebracht werden. Das hydraulische Stellsystem des VIRAGE2000 hatte zu einem recht hohen Gesamtgewicht des Dreirads geführt, so dass ein leichter pneumatischer Mechanismus für das neue Dreirad vorgesehen wurde. Der Vorteil des geringeren Gewichtes von Pneumatik ge-

genüber Hydraulik wird aber mit dem Nachteil der geringeren Regelbarkeit des Stellantriebs erkaufte. Als pneumatischer Aktor wurden zwei parallel angeordnete fluidische Muskel (Kap. 6.2) mit einer Länge von 1 m verwendet, die über jeweils einen Seilzug auf einen Neigehebel an der Vorderachse wirken. Die Fluidmuskeln werden über vier pulsweitenmodulierte Regelventile angesteuert. Dabei wird die Druckluft dem Hauptrohr des Fahrradrahmens entnommen, das als Druckspeicher ausgelegt ist. Die Drucklufterzeugung wird durch eine Kolbenpumpe bewirkt, die von den Rädern angetrieben wird, oder durch einen Elektromotor. Dieser neue Typ von Dreirad mit Neigetechnik befindet sich zur Zeit in der Erprobung.

Literaturverzeichnis

- [1.1] *Schweitzer, G.* : Mechatronik. VDI-Berichte 787, VDI-Verlag, Düsseldorf, 1989, S. 1-15
- [1.2] *Isermann, R.* : Mechatronische Systeme. Springer-Verlag, Berlin-Heidelberg-New York, 2008
- [1.3] *Weck, M.* : Werkzeugmaschinen. Band 1, 4. Auflage, VDI-Verlag, Düsseldorf, 1995
- [1.4] *Fischetti, M. / Linsmeier, K.-D.* : Konzert für Triebwerk und Orchester. Spektrum der Wissenschaft, Aug. 2005, Heidelberg, S. 84-85
- [1.5] *Dankert, H. / Dankert, J.* : Technische Mechanik. 7. Aufl., Springer-Vieweg Verlag, Wiesbaden, 2013
- [1.6] *Rzevski, G.; Editor* : Perception, Cognition and Execution. Butterworth-Heinemann, Oxford-London-Boston, 1999
- [2.1] *Ropohl, G.* : Systemtechnik-Grundlagen und Anwendung. Carl Hanser Verlag, München, 1975
- [2.2] *Gasch, R. / Knothe, K. / Liebig, R.*: Strukturdynamik, 2. Aufl. , Springer Verlag, Berlin-Heidelberg, 2012
- [2.3] *Bremer, H.* : Dynamik und Regelung mechanischer Systeme. Teubner Verlag, Stuttgart, 1988
- [2.4] *Huston, R.L.* : Multibody Dynamics. Butterworth-Heinemann, Boston, 1990
- [2.5] *Ludyk, G.* : Theorie dynamischer Systeme. Elitera Verlag, Berlin, 1977
- [3.1] *Mayr, M.* : Technische Mechanik. Carl Hanser Verlag, München-Wien, 1995
- [3.2] *Waller, H. / Schmidt, R.* : Schwingungslehre für Ingenieure. BI-Wissenschaftsverlag, Mannheim/Wien/Zürich, 1995
- [3.3] *Robertson, R. E. / Schwertassek, R.* : Dynamics of Multibody Systems. Springer-Verlag, Berlin-Heidelberg-New York, 1988
- [3.4] *Jeger, M. / Eckmann, B.* : Einführung in die vektorielle Geometrie und lineare Algebra. Birkhäuser Verlag, Basel, 1967
- [3.5] *Kreuzer, E. J. / Lugtenburg, J.-B. / Meißner, H.-G. / Truckenbrodt, A.* : Industrieroboter. Springer-Verlag, Berlin-Heidelberg-New York, 1994
- [3.6] *Bögelsack, G. / Kallenbach, E. / Linnemann, G.* : Roboter in der Gerätetechnik. Hüthig Verlag, Heidelberg, 1985
- [3.7] *Bartenschlager, J. / Hebel, H. / Schmidt, G.* : Handhabungstechnik mit Robotertechnik. Vieweg+Teubner Verlag, Wiesbaden 1998
- 3.8] *Pfeiffer, F. / Reithmeier, E.* : Roboterdynamik. Teubner Verlag, Stuttgart, 1987
- [3.9] *Pfeiffer, F., Schindler, Th.* : Einführung in die Dynamik. Springer-Vieweg Verlag, Berlin-Heidelberg 3. Aufl., 2014
- [4.1] *Hesselmann, N.* : Digitale Signalverarbeitung. Vogel Verlag, Würzburg, 1983
- [5.1] *Hauptmann, P.* : Sensoren. Carl Hanser Verlag, München-Wien, 1991
- [5.2] *Heywang, W.* : Sensorik. 4. Aufl., Springer Verlag, Berlin-Heidelberg-New York, 1993
- [5.3] *Mühl, Th.* : Einführung in die elektrische Messtechnik. Springer-Vieweg Verlag, 4. Aufl. 2014
- [5.4] *Schrüfer, E.* : Elektrische Messtechnik. 8. Aufl., Hanser Verlag, München, 2004
- [5.5] *Lemme, H.* : Optische Halbleitersensoren. Elektronik Sonderheft Sensoren II, Franzis Verlag, München, 1987, S. 5 - 8

- [5.6] *Schoppnies, E. (Hrsg.)* : Lexikon der Sensortechnik. VDE-Verlag, Berlin-Offenbach, 1998
- [5.7] *Schiessle, E.* : Sensortechnik und Meßwertaufnahme. Vogel Verlag, Würzburg, 1992
- [5.8] *Warnecke, H.-J. / Schraft, R.D. (Hrsg.)* : Industrieroboter. Springer Verlag, Berlin, 2012
- [5.9] *Heuberger, A. (Hrsg.)* : Mikromechanik. Springer Verlag, Berlin-Heidelberg-New York, 1994
- [5.10] *Nehmzow, U.* : Mobile Robotik. Springer Verlag, Berlin-Heidelberg-New York, 2002
- [5.11] *Czichos, H.* : Mechatronik, Springer-Vieweg Verlag, Wiesbaden, 3. Aufl., 2015
- [6.1] *Brosch, P.F.* : Moderne Stromrichterantriebe. Vogel Verlag, Würzburg, 2007
- [6.2] *Heimann, B. / Albert, A. / Ortmaier, T. / Rissing, L.* : Mechatronik. Karl Hanser Verlag, München, 4. Aufl., 2016
- [6.3] *Weck, M.* : Werkzeugmaschinen 3. Mechatronische Systeme, Vorschubantriebe, Prozessdiagnose. 6. Auflage, Springer Vieweg, Berlin-Heidelberg, 2006
- [6.4] *Hofer, K.* : Gekrümmte Drehstrom-Linearmotoren als Antriebe in der Fahrzeugtechnik. Maschinenmarkt, Vogel Verlag, Würzburg, Nr. 35, Aug. 1996, S. 42 - 47
- [6.5] *Budig, P.-K.* : Schraubgewinde-Reluktanzmotor: ein linearer Präzisionsantrieb. Maschinenmarkt, Vogel Verlag, Würzburg, Nr. 35, Aug. 1996, S. 40 - 41
- [6.6] *Nishibori, K. / Okuma, S. / Obata, H.* : Driving Characteristics of Robot Hand with Fingers using Langevin-Type Ultrasonic Motors. Proceedings of the 3. Conference on Mechatronics and Robotics, B. G. Teubner Verlag, Stuttgart, 1995, S.94-104
- [6.7] *Halter, K.* : Intelligente Werkstoffe - Formgedächtnislegierungen. Techpress, Fachorgan der Ingenieurschule Burgdorf (CH), 1/96, S. 3-12
- [6.8] *Jendritza, D.J.* : Neue Verfahren und Produkte der Mikrosystemtechnik und Mechatronik. Automatisierungstechnische Praxis, Oldenbourg Verlag, Sept. 1994, S. 58 - 64
- [6.9] *Menz, W.* : Die Mikrosystemtechnik und ihre Anwendungsgebiete. Spektrum der Wissenschaft, Dossier: Mikrosystemtechnik, Jan. 1996, S. 32 - 39
- [6.10] *Menz, W. / Mohr, J. / Paul, O.* : Mikrosystemtechnik für Ingenieure. Wiley-VCH Verlag, 3. Aufl., 2005
- [7.1] *Stein, G.* : Automatisierungstechnik in der Maschinentechnik. Hanser Verlag München, 1994
- [7.2] *Mealy, G. H.* : A Method for Synthesizing Sequential Circuits. The Bell Syst. Techn. Journal, no. 9, 1955, S. 1045-1079
- [7.3] *Moore, E. F.* : Gedanken-Experiments on Sequential Machines. Automata Studies, Princeton University Press, Annals of Mathematical Studies no. 34, 1956
- [7.4] *Roddeck, W.* : Mehrfacher Signalwechsel in Schaltnetzen und asynchronen Schaltwerken. Dissertation, Bochum, 1979, Schriftenreihe des Lehrstuhls für Meß- und Regelungstechnik der Ruhr Universität Bochum.
- [7.5] *Drösser, C.* : Fuzzy Logic - Methodische Einführung in krauses Denken. Rowohlt, 1994
- [7.6] *Kahlert, J. / Frank, H.* : Fuzzy-Logik und Fuzzy-Control. Vieweg Verlag, Braunschweig, 1994
- [7.7] *Hirota, K. (Hrsg.)* : Industrial Applications of Fuzzy Technology. Springer Verlag, Tokyo, 1993
- [7.8] *Fasol, K. H.* : Skript zur Vorlesung Regelungstechnik 1, Ruhr-Universität Bochum Lehrstuhl für Mess- und Regelungstechnik der Ruhr Universität Bochum
- [7.9] *Göldner, K.* : Mathematische Grundlagen der Systemanalyse. Verlag Harri Deutsch, Frankfurt/Main, 1983

-
- [7.10] Weck, M. : Werkzeugmaschinen, Band 3, VDI-Verlag, Düsseldorf, 1995
- [7.11] Olsson, G. / Piani, G. : Steuern, Regeln und Automatisieren. Hanser Verlag, München, 1993
- [7.12] Isermann, R. : Identifikation dynamischer Systeme , Springer Verlag, Berlin, 2012
- [7.13] Roddeck, W. / Rehbein, H.-J. : Ordnung muss sein. Maschinenmarkt, Nr. 22, 2002, S. 106-109
- [7.14] Phoenix Contact (Hrsg.) : Grundkurs Sensor/Aktor-Feldbustech. Vogel Verlag, Würzb., 1997
- [7.15] Schmid, D. u. A. : Steuern und Regeln für Maschinenbau und Mechatronik. Verlag Europa-Lehrmittel, Haan-Gruiten, 2014
- [7.16] Rembold, U / Levi, P. : Realzeitsysteme zur Prozeßautomatisierung. Hanser Verlag, 1994
- [7.17] Bähring, H. : Mikrorechner-Technik. Springer Verlag, Berlin-Heidelberg, 2005
- [7.18] Seraphin, M. : Neuronale Netze & Fuzzy-Logik, Franzis Verlag, München, 1998
- [7.19] Braitenberg, V. : Vehikel: Experimente mit kybernetischen Wesen, Rowohlt, 1993
- [8.1] Otter, M. : Objektorientierte Modellierung mechatronischer Systeme am Beispiel geregelter Roboter. Fortschr.-Ber. VDI Reihe 20 Nr. 147, VDI-Verlag, Düsseldorf, 1995
- [8.2] Damic, V. / Montgomery, J. : Mechatronics by Bond Graphs. Springer Verlag, Berlin, 2003
- [8.3] Karnopp, D. / Margolis, D. / Rosenberg, R. : Modeling and Simulation of Mechatronic Systems. John Wiley & Sons, Hoboken, New Jersey, 1995
- [8.4] Hoffmann, J. : Matlab und Simulink. Addison Wesley Longman Verlag, Bonn, 1999
- [8.5] iXtronics GmbH (Hrsg.) : Benutzerhandbuch CAMEL-View R 4.0, Paderborn, 2001
- [8.6] Hahn, M. : OMD-Ein Objektmodell für den Mechatronikentwurf Fortschr. Ber. VDI, Reihe 20, Nr. 299, VDI-Verlag, Düsseldorf, 1999
- [8.7] Pelz, G. : Modellierung und Simulation mechatronischer Systeme. Hüthig Verlag, Heidelberg, 2001
- [8.8] VDI-Richtlinie 2206 Entwicklungsmethodik mechatronischer Systeme.
- [8.9] Borutzky, W. : Bond Graph Methodology. Springer Verlag London, 2010
- [8.10] Breunese, A. / Broenink, J. : Modeling Mechatronic Systems using SIDOPS+ Language, Int. Conference on Bond Graph Modeling and Simulation, S. 301-306, SCS Publishing, 1997
- [8.11] Roddeck, W. : Grundprinzipien der Mechatronik, Springer Vieweg Verlag, Wiesbaden, 2013
- [9.1] Siegwart, R. / Traxler, A. : Möglichkeiten und Grenzen schneller Aktuatoren am Beispiel einer magnetisch gelagerten Hochgeschwindigkeits-Frässpindel. VDI-Berichte, Nr. 787, VDI-Verlag Düsseldorf, 1989, S. 101-118
- [9.2] Plitt, K. / Thomson, B. / Truckenbrodt, A. : Hydraulische Stellzylinder und aktive Rad-Feder-Systeme am PKW. VDI-Berichte, Nr. 787, VDI-Verlag Düsseldorf, 1989, S. 85-100
- [9.3] Lückel, J. : Interne Mitteilung des Mechatronik Laboratoriums Paderborn, Universität-GH Paderborn über Aktive Fahrwerkstechnik.
- [9.4] MacConaill, P.A. / Drews, P. / Robrock, K.-H. (Hrsg.) : Mechatronics & Robotics, IOS Press Amsterdam, Oxford, Washington, Tokio, 1991
- [9.5] Roddeck, W. / Rehbein, H.-J. : Zusatzachsen am Handgelenk bewegen Laserschneidkopf. Maschinenmarkt, Würzburg, Nr. 23 Juni 1995, S. 42 - 47
- [9.6] Roddeck, W. / Rehbein, H.-J. : Autonomous Tool-Mover for Laser-Cutting with Industrial

Robots. Proc. of the 3. Conf. on Mechatron. and Robotics, Teubner Verlag, Stuttgart, 1995, S. 226-234

- [9.7] *Beßler, T. / Henschel, F.* : Aufbau und Steuerung einer Doppel-Hexapode-Struktur basierend auf Fluidmuskeln. Diplomarbeit, FH-Bochum, 2002
- [9.8] *Hahn, H. / Klier, W. / Fürst, D.* : Theoretische Modellbildung räumlich bewegter servopneumatischer und elektrischer Mehrachsenprüfstände und Hexapoden. Forschungsbericht zum DFG-Vorhaben Ha-1666/6-1/2, Universität-Gh Kassel, 2001
- [9.10] *Pohl, M.* : Mechatronic Design of Tilting Threewheelers. Proceedings of 2. IFAC-Conference on Mechatronic Systems, Berkley, 2002
- [9.11] *Gausemeier, J. / Lückel, J.* : Entwicklungsumgebungen Mechatronik. HNI-Verlagsschriftenreihe 2001
- [9.12] *Henker, E.* : Fahrwerktechnik Grundlagen, Bauelemente, Auslegung. Vieweg-Verlag, 1996

Sachwortverzeichnis

Abbruchfehler 400
ABS 461 ff.
Absolutbewegung 59 ff.
Abstandsmessung 194, 471 ff.
Abtastung 148, 350 ff.
Adresse 384
Ähnlichkeitsmodell 29
Agent
-, intelligenter autonomer 267 ff.
Airbag 23
Akkumulator 387
Aktor 21, 212 ff.
-, fluidischer 238 ff.
-, hydraulischer 239 ff.
-, klassischer 214 ff.
-, mikrostrukturierte 257 ff.
-, neuartiger 248 ff.
-, piezoelektrische 250 ff.
-, pneumatischer 238
Aliasing 149
Amplitude 116
Amplitudenfunktion 173
Amplitudengang 172, 334
Amplitudenkennlinie 126
Amplitudenkorrektur 331
Amplitudenverhältnis 173
Analog-/Digital-Wandler 148, 166 ff.
Analyse 370
Anfangsbedingung 25, 315
Anfangsphase 116
Ankerverstellung 218, 231
Anregelzeit 372
Antivalenz 287
Antriebe 214 ff.
Äquipotentiallinien 187
aperiodischer Grenzfall 119
Arbeit 97
-, virtuelle 100
Arbeitskontakt 285
Arbeitssatz 98
Arbeitsspeicher 383
ASIC 393
Assemblersprache 392
Assoziativspeicher 303
Asynchronmaschine 215, 227, 229 ff.
Attribut 26
Ausgangsgleichung 342
AusgangsvARIABLE 280
Ausgleichszeit 374
Ausregelzeit 372
Aussetzbetrieb 220

Automat 280 ff.
-, asynchroner 281
-, synchroner 281
Automatisierungstechnik 260 ff.
autonom 266
Axialkolbenmotor 242

Backpropagation-Netzwerk 301, 305 ff.
Beschleunigung 51 f.
-, des starren Körpers 56
Beschleunigungssensoren 206 ff.
Bewegung
-, allgemeine, des starren Körpers 64, 91
-, ebene, des starren Körpers 56
Bildbereich 146, 310
Bildverarbeitungssystem 263
Bindung 77 ff.
-, holonome 78
-, skleronome 78
-, rheonome 78
Binormalenrichtung 52
Bitbus 276
Black Box 9, 26
Blockschaltbild 318 ff.
Blockschaltbild-Editoren 425 ff.
Bode-Diagramm 126, 333
Bondgraph 408 ff., 431
Boole'sche Algebra 281 ff.
Bus 272, 383

CAD, CAE 15
CAN-Feldbus 166, 273, 276
CCD-Element 200 ff.
CCD-Kamera 263
C-Element 413
Coriolisbeschleunigung 60, 66
Coulomb-Reibung 33
CPU 383, 385
Cross-Compiler 392

D'Alembertsches Prinzip 89 ff.
D-Feld-Sensor 195
D-Glied 332
D-Verhalten 332, 334
Dämpfung
-, überkritische 119
-, viskose 33
Dämpfungskörper 13
Dämpfungsmaß
-, elektrisches 11
-, mechanisches 11
Dämpfungswinkel 118
Dauerbetrieb 220
Debugger 392

- Dehnungsmessstreifen 180 ff., 208
- Delta-Dirac-Impuls 332
- De Morgan'sches Theorem 283
- Denavit-Hartenberg Methode 80 ff.
- Deviationsmomente 94
- Dezibel 334
- Dielektrizitätskonstante 188
- Differentialfeldplatte 206
- Differentialgleichung 10, 25, 113 ff.
 - , gewöhnliche 32, 113
 - , homogene 113
 - , nichtlineare 34
 - , partielle 32
- Differentialtransformator 184
- Differentiationssatz 315
- Differenzverfahren 44
- Digitalrechner 21
- Diode 224
- Dirac-Impuls 148, 151, 314
- Direktantrieb 224
- Disjunktion 282
- Domäne 407
- Drall 93
- Drallsatz 93
- Drehimpuls 93
- Drehmelder 196 ff.
- Drehzahlmesser 204
- Druckmessung 210
- Durchschnitt 298
- Durchtrittsfrequenz 358
- Dynamik 22, 50 ff.
- Echtzeitbetrieb** 380
- Eckfrequenz 335
- Effekt
 - , kapazitiver 178, 188
 - , magnetische 184 ff.
 - , magnetoelastischer 187
 - , magnetostriktiver 187, 209
 - , optische 190 ff.
 - , piezoelektrischer 11, 189 ff.
 - , piezoresistiver 174, 178
 - , pyroelektrischer 178, 189 ff.
 - , reziproker piezoelektrischer 12, 189
- Effizienz 36
- Effort 408 ff.
- Eigenfrequenz 11, 116, 124
- Eigenkreisfrequenz 115 f.
- Eigenvorgänge 113
- Eigenwert 114
- Eingangsvariable 280
- Eingangssignal
 - , zeitdiskret 350
 - , zeitkontinuierlich 350
- Einmassenschwinger 112 ff., 321, 421
- Einschrittverfahren 401
- Einstellregeln 373 ff.
 - , nach Chien/Hrones/Reswick 374
 - , nach Ziegler/Nichols 374
- Einstellzeit 172
- Einwegformgedächtnis 254
- emergent 268
- Empfindlichkeit 170
- Endschalter 289
- Energie 411
 - , kinetische 98
 - , potentielle 97
- Energiesatz 97 ff.
- Energiesteller 212
- Energiewandler 212
- Entwicklungssystem 392
- EPROM 383, 393
- Erkennen 23
- Erregung
 - , deterministische 120
 - , harmonische 120, 122 ff.
 - , periodische 120
 - , stochastische 120
 - , transiente 120
- Eulerrelation 115, 123, 138
- Eulerwinkel 69
- Expertensystem 266
- Exzentrizität 242, 481
- Fahrer-Assistenzsystem** 463
- Fahrwerk
 - , aktives 463 ff.
- Faltung 151, 314, 316
- Fast-Fourier-Transformation 154, 435
- Fehler
 - , absoluter 174
 - , relativer 174
 - , systematischer 175
- Fehlerfortpflanzung 176
- Fehlerklasse 175
- Fehlerkurve 174
- Feldbus 166, 273
- Feldplatte 180, 187, 198, 205
- Feldschwächung 219
- Feldverstellung 218, 231
- Fensterfunktion 150
- Finite-Elemente-Methode 15, 42, 44 ff.
- Finite-Segment-Methode 42 ff.
- Fliehkraftregler 261
- Fließbandprinzip 388
- Flip-Flop 289 ff.

- Flow 408 ff.
- Flügelzellenmotor 241
- Fluss
 - , magnetischer 184, 216
- Flüssigkeit
 - , elektrorheologische 256
 - , magnetorheologische 256
- Formgedächtnislegierung 253 ff.
- Fotodiode 190 ff.
- Fotoelement 191
- Fototransistor 192
- Fotowiderstand 180, 193
- Fourier-Reihe 140 ff.
- Fourier-Analyse 141, 435
- Fourierkoeffizient 139
- Fourier-Reihe 140 ff.
- Fourier-Transformation 146 ff.
 - , diskrete 148 ff.
- FPGA, FPLA 394
- Freiheitsgrad 37 f. , 53, 77
- Freilaufdiode 226
- Frequenzbereich 120, 139
- Frequenzgang 125, 172, 310, 324 ff.
- Frequenzkennlinien 126, 333 ff.
- Frequenzumrichter 215, 228, 232 ff.
- Frequenzverhältnis 124
- Führungs-Übertragungsfunktion 324
- Führungsbewegung 59, 60, 61
- Führungsgröße 265
- Funktion
 - , gerade 143, 146
 - , ungerade 143, 146
- Funktionshazard 292
- Funktionsplan 280, 286
- Fuse 394
- Fuzzifizierung 297
- Fuzzy-Inferenz 299
- Fuzzy-Logic 253, 294 ff.
- Fuzzy-Menge 295 ff.
- Fuzzy-Set-Theorie 294
- G**age-Faktor 182
- Galoping-Schwingung 352
- Gegenkopplung 319
- Geräte
 - , mechatronische 6
- Geschwindigkeit 51
 - , des starren Körpers 55
 - , generalisierte 107
- Geschwindigkeitsfehler 371
- Geschwindigkeitsensor 204 ff.
- Gewichtsfunktion 318, 367
- Gleichanteil 140, 142
- Gleichspannungszwischenkreis 224
- Gleichstrommotor 215 ff.
- Gleichstromsteller 226
- Gleichung
 - , charakteristische 114
- Glockenläufermotor 223
- Grenzfrequenz 173
- Gültigkeitsbereich 34, 36
- Güte
 - , mechanische 14
- Gütekriterien 372
- Gyrator 417 ff.
- H**all-Effekt 185
- Hallgenerator 186
- Hann-Fenster 153
- Hardware-in-the-Loop 438, 459, 465, 489
- Hauptachse 94 ff.
- Hazard 281, 290
- Heisenbergsche Unschärferelation 31
- Hexapode 483 ff.
- Hierarchien 270 ff.
- Hilfsenergie 212
- Hilfsmassendämpfer 466
- Historie-Blöcke 425
- Hooke'sches Gesetz 17, 112
- Host-Rechner 392
- Hurwitz-Kriterium 355
- Hydraulikzylinder 239
- Hydromotor 240 ff.
- Hydropumpe 240
- I**-Element 414
- I-Glied 331
- I-Regler 346 ff.
 - Verhalten 331, 333, 345
- Identifikation 138, 365 ff.
- Identifikationsverfahren 30, 366
- Identität 286
- Impuls 89
 - , generalisierter 411
- Impulssatz 88 ff.
- In-circuit-Emulator 392
- Induktion
 - , magnetische 185
- Induktionsgesetz 184
- Induktivität 414
- Induktor 236
- Industrieroboter 78, 470 ff.
- Inertialsystem 51, 64, 72, 80
- Inferenzsystem 266
- Integralkriterien 373
- Integrationsformel
 - , nach Adams- Bashfort 400

- , nach Euler-Cauchy 400
- , nach Runge-Kutta 402
- Interbus 277
- Interferenz
- , destruktive 20
- Istwert 265
- Junction** 419 ff.
- Kalibrierung** 171
- Kammantrieb 259
- Kapazität 413
- Kardan-Winkel 69
- Kardanwinkelgeschwindigkeit 74
- Kaskadenregelung 376
- Kenngößen
- , dynamische 171 ff.
- , statische 169 ff.
- Kennlinie
- , statische 170, 309
- Kinematik 22, 41, 50 ff.
- Kinetik 22, 50, 88 ff.
- Kippmoment 228, 230, 235
- Klasse 437
- Knickarmroboter 470
- Koeffizient
- , piezoresistiver 183
- Kommutator 46, 216, 221
- Kommutierungsgrenze 220
- Komparator 167, 300
- Komplement 267
- Konfiguration 88
- Konjunktion 282
- Kontaktplan 280, 286
- Kontinuitätsgleichung 27
- Koordinaten
- , gelenknatürliche 79
- , generalisierte 102
- , verallgemeinerte 79
- Koordinatensystem 52, 80 ff.
- , körpereigenes 62
- , körperfestes 64 ff.
- , natürliches 52
- , raumfestes 64 ff.
- Körper
- , elastischer 41
- , starrer 41, 53 ff.
- Korrespondenztabelle 314
- Kräfte
- , eingeprägte 41
- , generalisierte 105
- , konservative 97 f.
- Kraftmessung 209 ff.
- Kreisfrequenz 123
- Kreisinterpolation 477 ff.
- Kreisschneider 479 ff.
- Kreisverstärkung 346
- Krümmungsmittelpunkt 53
- Krümmungsradius 53, 56
- Lagemessung** 194
- , absolute 197
- , inkrementale 197
- , zyklisch absolute 197
- Lageregelung 329, 376 ff.
- Lagerung
- , aktive elektromagnetische 456 ff
- , kardanische 69
- , passive magnetische 456
- Lagrangesche Bewegungsgl. 105 ff.
- Lagrangesche Funktion 108
- LAN 271, 273 ff.
- Langsamläufer 224
- Laplace-Transformation 9, 146, 310 ff.
- Laserradar (Lidar) 192, 202 ff.
- Laser-Schneidtechnologie 471 ff.
- Last-Drehmomentsperre 445
- Lateraleffekt-Fotodioden 202
- Layer 277
- , hidden 301
- Leckverlust 242
- Leerlauf 218
- Lehr'sches Dämpfungsmaß 118, 124
- Leistung 408
- Leistungsbond 409
- Leistungsdichtespektrum 155
- Lernen 268 ff., 299
- Leuchtdiode 192
- Lichtschranke 23, 192
- Lichttaster 192
- LIGA-Verfahren 257 ff.
- Linearantrieb 330
- Linearisierung 34, 308, 449
- Linearität 39
- Linearmotor 236
- Linearschrittmotor 237
- Logikplan 286
- Lorentzkraft 185, 216
- Lösung
- , der Differentialgleichung 113 ff.
- , homogene 113
- , klassische 114
- , mechanische 6
- , mechatronische 6
- , partikuläre 113
- Magnetlager** 456 ff.
- Magnetostriktion 187, 209

- Maschine
 - , hydrodynamische 240
 - , hydrostatische 240
 - , intelligente 261 ff.
 - , von Neumann- 383 ff.
- Masse
 - , seismische 207
 - , träge 97
- Massenmittelpunkt 50, 92
- Massenpunkt 41, 50 ff., 56
- Massenträgheitsmomente 93 ff.
- , axiale 94
- Maßverkörperung 161, 198
- , kodierte 199
- Matrix
 - , antisymmetrische 73
 - , inverse 69, 72
 - , orthogonale 69, 72
 - , transponierte 69 ff.
- Maxterm 284
- Mechanik
 - , analytische 100
 - , Prinzipien der 100 ff.
 - , technische 41
- Mechatronik 6, 21
- Mehrkörpersysteme 41, 77 ff., 424, 436
- Mehrmassenschwinger 130 ff.
- Mehrschrittverfahren 400
- Memory-Metall 253
- MEMS 257
- Messbereich 170
- Messeffekt 177 ff.
- Messfehler 161
- Messfühler 158, 163 f.
- Messgröße 160, 162
- Messkette 162
- Messnormal 161
- Messprinzip 163
- Messsignal 162
- Messtechnik 160 ff.
- Messwert 161
- , wahrer 173
- Messwertaufnehmer 159
- Messwertwandler 162 ff.
- Metaklasse 437
- Methode nach Ritz 43
- Mikrocontroller 382, 390 ff., 455, 490
- Mikrorechner 279, 381 ff., 455
- Mikroprozessor 381 f.
- Mikrosystemtechnik 180, 248
- Minimalform 285
- Minterm 283
- Mitkopplung 319
- Mittelwert
 - , arithmetischer 175
- Modalanalyse 15
- Modalkoeffizient 131
- Modellbildung 28 ff.
- , objektorientierte 430 ff.
- , theoretische 366
- Modell
 - , analytische 29
 - , mathematische 28
 - , nichtparametrische 39
 - , parametrische 38
 - , physikalische 28
- Modellgültigkeit 36
- Modulation 19
- Momentanpol 57 f., 95, 102, 489
- Momente
 - , generalisierte 108
- Momentensatz 94
- MOS 200
- Motormoment 215 ff.
- Motorverstellung 243
- Multiport 409, 418 ff.
- Muskel
 - , fluidischer 256 ff., 484
- Mustererkennung 269, 293
- Nachgiebigkeit 413**
- Näherungsverfahren 43
- NAND 287, 289
- Nebenschluss 215
- Negation 284, 286
- Neigetechnik 486 ff.
- Nenn
 - , kennlinie 171
 - , widerstand 171
 - , übertragungsfaktor 171
- Netzwerk 271 ff.
- , feed-forward 304
- , neuronales 269, 299 ff.
- Neuron 269
- , McCulloch-Pitts 300 ff.
- Newton'sche Axiome 31, 88
- nichtlinear 32
- Nicken 137, 434, 464
- NOR 287, 289
- Normalbeschleunigung 53
- Normalenrichtung 52
- Normalform
 - , disjunktive 284
 - , konjunktive 284
- Normalverteilung 176
- Numerische Integration 399 ff.

- Nyquist-Kriterium 357
- Ö**DER-Funktion 282
- Öffner 285
- Operationsverstärker 165, 319 ff.
- Orientierung 62
- Originalbereich 146, 310, 316
- orthogonal 70
- Ortskurve 126, 310, 324 ff.
- Ortskurvenkriterien 355
- Ortsvektor 30, 51, 54
- OSI-Referenzmodell 274 ff.
- P**-, PI-Regler 345 ff.
- PD₁-Glied 333
- PID-Regler 348 ff., 364
- P-Verhalten 329
- PAL 393 f.
- Parallelschaltung 318
- Parallelumsetzung 167
- Parameter
- , konzentrierte 32, 39
- , verteilte 32, 39
- Pendel 34
- , invertiertes 446 ff.
- perceptron 300, 302 ff.
- Permanentmagnet 215, 221
- Permeabilität 184
- Phasenanschnittsteuerung 225
- Phasendiagramm 452
- Phasenfunktion 128
- Phasengang 172, 334
- Phasenkenlinie 126
- Phasenkurve 128, 406
- Phasenverschiebung 123
- Phasenwinkel 123
- Piezoeffekt 12, 208
- Piezomaterialien 14, 250
- Piezomotor (Langevin) 253
- Piezoventil 246 ff.
- Piezowiderstandseffekt 178, 189 ff.
- Pilotmodell 28
- Pipelining 388 ff.
- Pixel 200
- Plattenkondensator 179
- PLA, PLD, PLS 393 f.
- Pneumatikzylinder 238 ff.
- Polarisation 250
- Polarkoordinaten 54
- Polpaarzahl 228
- Port 408
- Position 62
- Positionsfehler 371
- Positionssensoren 194 ff.
- Potentiometer 163, 180
- Pressduktor 209
- Prinzip
- , der virtuellen Arbeit 101 ff.
- , vom Minimum des Potentials 101
- . d'Alembert 89
- , von Hamilton 101
- Profibus 276
- Programm 261
- Programmsprung 385
- Programmzähler 385
- Proportionalventil 244 ff.
- Prozessdatenverarbeitung 260
- mit Mikrorechnern 380 ff.
- Prozessleitsystem 273
- Prozessrechner 380
- PROM 383, 393
- Pototypmodell 28
- PT₁-Glied 329
- Pulsbreitensteuerung 226
- Pumpenverstellung 242
- Punktmechanik 50
- Q**uellen 415
- Radialkolbenmotor 242
- RAM 383, 388
- Rampenfunktion 312, 318
- Rampenverfahren 168
- Rattern
- , regeneratives 16
- Reaktionsschiene 236
- Rechenwerk 386
- Rechner
- , RISC 389
- , Vektor 389
- Rechnerarchitektur 385
- Regeldifferenz 265
- Regelfaktor
- , statischer 346
- Regelgröße 265
- Regelkreis
- , aufgeschnittener o. offener 324, 353
- , mehrschleifiger 376 ff.
- Regelstrecke 265
- Regelung 264 f., 307
- Regelungstechnik 49, 260, 307 ff.
- Register 386
- Regler 265, 345 ff.
- Reibung
- , coulombsche 33
- , viskose 33
- Reihenschaltung 318
- Reihenschluss 215

- Relais 285
- Relation 26, 286
- Relativbewegung 59 ff.
- R-Element 412
- Reluktanz 249
- Resolver 196 ff.
- Resonanzstelle 125
- Robotersimulationssystem 440
- ROM 383
- Rotation 54 ff.
- Ruck 194
- Rückkopplung 265, 319
- Rücksetzsignal 289
- Rückwärtstransformation 86
- Rückwirkung 337
- Ruhekontakt 285
- Rundungsfehler 400
- Sättigung** 166
- Schaltalgebra 281
- Schaltfolgediagramm 290
- Schaltnetz 280
- Schaltwerk 280
- Scheibenläufermotor 223
- Schichtdickenmessung 159
- Schleife
 - , algebraische 426 ff.
- Schleifringläufermotor 229
- Schleppabstand 380
- Schließer 285
- Schluckvolumen 240
- Schlupf 229, 461
- Schlussfolgern 266 ff.
- Schmiegungsebene 52
- Schnittprinzip 90
- Schraubgewinde-Reluktanzmotor 248
- Schrittmotor 215, 234 ff.
- Schwellwertfunktion 300
- Schwerpunkt 34, 90
- Schwerpunktsatz 93
- Schwinger
 - , piezoelektrischer 13
 - , nichtlinearer 128 ff.
- Schwingungstilger 466 ff.
- Selbsthalteschaltung 288
- Selbstregulierung 261
- Sensor 21, 155, 157 ff.
 - , integrierter 160
 - , intelligenter 160
 - , rezeptiver 178
 - , signalbearbeitender 178
- Servoventil
 - , elektrohydraulisches 244 ff.
- Setzsignal 289
- Sigmoid-Funktion 305
- Signalflussplan 264
- Signalprozessor 387
- Signum-Funktion 303
- SI-System 161
- Simulation 397 ff.
- Simulationsmodell 29
- Simulationssprache 422 ff.
- Simulationssystem 15, 35, 134, 367, 397, 422 ff.
 - , Roboter- 440 ff.
- Singleton 296
- Software-in-the-Loop 438 f.
- Solarzelle 191
- Soll-Istwert-Vergleich 265
- Sollwert 265, 323
- Sonar-Prinzip 202
- Spaltfunktion 143, 150
- Spannungsfolger 320
- Speicher 289
- Spektrum 13, 140, 152
- Spezifikation 371 ff.
- Sprungantwort 172, 318, 367
- Sprungfunktion 314
- Stabankermotor 222
- Stabilität 352 ff.
- Stabilitätskriterien
 - , algebraische 355
- Stabilitätsreserve 356
- Stapeltranslator 251
- Statik 22, 89
- Statistik 175
- Standardabweichung 175
- Steller (Chopper) 224
- Stellgröße 278
- Stellsignal 212 ff., 264
- Stellwert 265
- Stern-Dreieck-Umschaltung 230
- Steuerspiegel 242
- Steuerung 264, 278 ff., 307
 - , kombinatorische 280, 285 ff.
 - , sequentielle 280, 288 ff.
 - , speicherprogrammiert 279
 - , verbindungsprogrammiert 279
- Steuerungstechnik 264, 278 ff.
- Steuerwerk 386
- Stoff
 - , piezoelektrischer 11
- Stör-Übertragungsfunktion 323
- Störgröße 323
- Störverhalten 323

- Strecke
 - , mit Ausgleich 329
 - , ohne Ausgleich 330
- Stromlaufplan 286
- Stromrichter 224
- Struktur 26
- Strukturdynamik 22
- Subsystem 26
- Summationsstelle 319
- Superelastizität 253
- Superpositionsprinzip 39, 128, 308
- Synchronisierung 290, 293
- Synchronmaschine 215, 227 ff.
- Syntheseverfahren 308, 370 ff.
- System 25
 - , determiniertes 38
 - , dynamisches 38
 - , elektrisches 10
 - , kontinuierliches 39, 43
 - , mechanisches 10
 - , mechatronisches 22, 443 ff.
 - , stochastisches 39
 - , zeitinvariantes 40
- Systemanalyse 308
- Systembegriff 25 ff.
- Systemidentifikation 138, 365 ff.
- Tachogenerator 204 ff.
- Taktsignal 290
- Tangentenrichtung 52
- Tangentialbeschleunigung 53
- Tastverhältnis 141
- Tauchen 137, 434
- Taylorreihe 304, 399
- TCP 79
- Teach-In-Methode 268, 442
- Teilträder 7
- Theorem von Shannon 149
- Thyristor 224
- Token-Passing 276
- Torque-Motoren 224
- Totzeit 314, 333
- Trägheit 50, 414
- Trägheitsmoment 93
- Transferfunktion 303
- Transformation 66 ff.
 - , inverse 70
 - , orthogonale 70
- Transformationsmatrix 67 ff.
- Transformer 416 ff.
- Transparenz
 - , physikalische 36
- Übergangsbereich 292
- Übergangsfunktion 172, 318, 367
- Überschwingen 368 f.
- Überschwingweite 369
- Übertragungsfunktion 9 f., 315, 324
 - , des geschlossenen Regelkr. 324, 353
 - , des offenen Regelkreises 324, 353
- Übertragungsglied 1. Ordnung 327
- Übertragungskonstante 308, 374
- Ultraschall 12 ff., 201
- UND-Funktion 282
- Universalbus 385
- Unschärfe 295
- Variable**
 - , generalisierte 411
 - , linguistische 297
 - , logische 281
 - , von Neumann 384
- Venn-Diagramm 298
- Verbindungsnormale 80
- Verdrängerräume 240
- Verdrehung 68
- Vereinigung 298
- Verfahren
 - , implizit 402
 - , explizit 399
- Verformung
 - , elastische 41
 - , plastische 41
- Vergrößerungsfunktion 125
- Verhalten
 - , integrierendes 330
 - , proportionales 329
 - , reaktives 23
 - , zielorientiertes 23
- Verschiebung 68, 194
 - , generalisierte 411
 - , virtuelle 100
- Verschiebungssatz 314
- Verstärker
 - , invertierender 347
- Verstärkung 165, 318, 329
- Verstärkungsprinzip 39, 308
- Vertrauensbereich 175
- Verzögerungsglied 1. Ordnung 328
- Verzugszeit 374
- V-Modell 487
- Vorwärtstransformation 85
- Wahrheitstabelle** 282
- Wahrnehmen 23
- Wahrscheinlichkeit 176, 296
- Wanken 463
- Wechsler 285

Wegmessung
-, absolute 200
Werkzeugmaschine 4, 16
Weltkoordinaten 79
Wheatston'sche Brücke 182
Widerstand
-, spezifischer 180
Widerstandsthermometer 180
Winkelgeschwindigkeit 56
-, momentane 73
Wissen
-, regelbasiertes 266 ff.
Wortmodell 408
Wurzelortskurve 356

Z -Transformation 351
Zahl
-, komplexe 123
Zahnradmotor 240
Zeitbereich 120, 138, 308, 311
Zeitkonstante 318, 327
Zeitschaltung 288
Zeitschrittweitensteuerung 407
Zeitvarianz 40
Zentrifugalmomente 94
Zentripetalbeschleunigung 53
Zugehörigkeitsfunktion 295 ff.
Zugehörigkeitsgrad 296
Zustand 26
-, stationärer 27
Zustandsgleichungen 342, 450
Zustandsgröße 341, 338
Zustandsraumdarstellung 340 ff.
Zustandsregelung 451
Zwangsbedingung 78
Zweiwegeffekt 24