FORSCHUNGS FORUM

PADERBORN

Paderborner Universitätsmagazin 4-2001

Mit Beiträgen aus Paderborn, Höxter, Meschede und Soest

- Umsonst durchs All: GENESIS startet 2001
- Exzitonen im Gleichtakt
- Huminstoffe und Trinkwasser
- Strategien in Informations- und Kommunikationsbranchen
- Mikroelektronik neuronaler Netze
- Flugplanung mit Informatik-Methoden

IMPRESSUM

Herausgeber
Prof. Dr. Wolfgang Weber
Rektor der Universität Paderborn

Konzeption und Redaktion
Ramona Wiesner
Referentin für Öffentlichkeitsarbeit

Referat Hochschulmarketing und Universitätszeitschrift
Warburger Str. 100, 33098 Paderborn
Tel.: 05251/60 2553, 3880
E-Mail: wiesner@zv.uni-paderborn.de
http://hrz.uni-paderborn.de/hochschulmarketing

ForschungsForum Paderborn (ffp) im Internet
http://www.uni-paderborn.de/ffp/

Wissenschaftlicher Beirat
Prof. Dr. Gitta Domik
Prof. Dr. Wilfried B. Holzapfel
Prof. Dr. Jörg Jarnut
Prof. Dr. Klaus Meerkötter
Prof. Dr. Winfried Reiß
Prof. Dr. Heinrich Schulte-Sienbeck
Prof. Dr. Jürgen Voß
Prof. Dr. Jörg Wallaschek

Drucklegung
Dezember 2000

ISSN (Print) 1435-3709

Layout
PADA-Werbeagentur
Heierswall 2, 33098 Paderborn

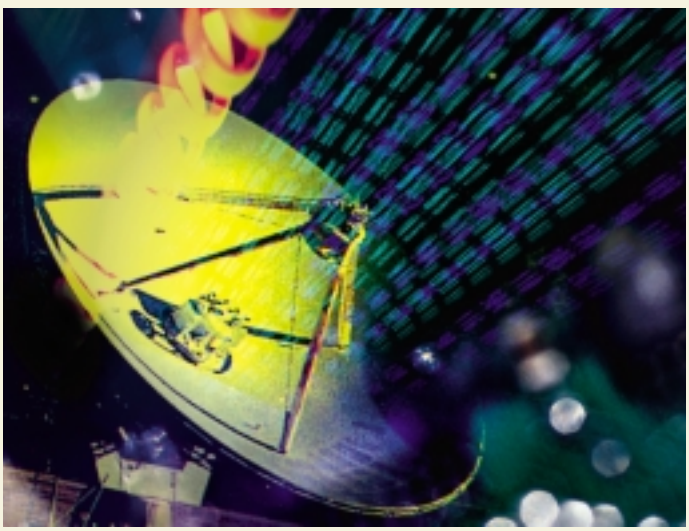
Druck und Anzeigenverwaltung
Verlag für Marketing und Kommunikation
D-67547 Worms, Hafenstraße 99

Auflage
5 000

Das ForschungsForum Paderborn erscheint weitestgehend auf der Grundlage der neuen amtlichen Rechtschreibregeln.

Titel

„Die Universität der Informationsgesellschaft“



Editorial

Das Raumfahrzeug der Mission GENESIS startet ins All, Chips imitieren Funktionen des Gehirns, Informatiker lösen Planungsprobleme von Fluggesellschaften, Fischstäbchen gehen auf eine lange Reise ... In Paderborn wird vielfältig und erfolgreich geforscht. Auf unterschiedlichsten Gebieten leisten Wissenschaftlerinnen und Wissenschaftler Forschung auf höchstem Niveau. Daher ist es uns immer wieder ein Leichtes, das Wissenschaftsmagazin mit interessanten und informativen Beiträgen zu füllen.

Insgesamt über 400 Seiten Forschung zum Anfassen sind so im Laufe der letzten vier Jahre entstanden. Der Zuspruch aus dem In- und Ausland spornt uns an. Dank der Beiträge aller beteiligten Wissenschaftlerinnen und Wissenschaftler kann der Anspruch des ForschungsForums Paderborn, den bestehenden Dialog zwischen Hochschule und Öffentlichkeit zu fördern, die Transparenz der Forschung an der Universität für die Wirtschaft und auch für Außenstehende zu erhöhen sowie Anknüpfungspunkte für Unternehmen erkennbar zu machen, erfüllt werden.

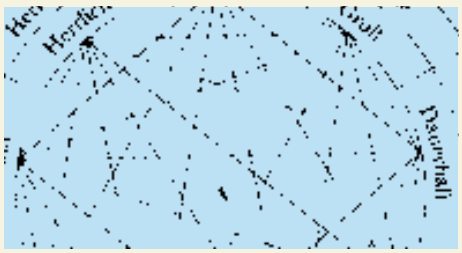
Und es wird einmal mehr deutlich: Die Universität Paderborn ist eine anerkannte wissenschaftliche Forschungseinrichtung. Das Drittmittelaufkommen beträgt derzeit rund 55 Millionen Mark. Es umfasst Forschungsmittel aus Institutionen der Forschungsförderung, insbesondere der Deutschen Forschungsgemeinschaft (DFG), aus der Wirtschaft, aus Stiftungen und aus verschiedenen Ministerien. Mit einem Sonderforschungsbereich, einer DFG-Forschergruppe sowie drei DFG-Graduiertenkollegs und hervorragenden Platzierungen in verschiedenen Forschungsrankings kann die Hochschule ihre Position als erfolgreiche Forschungseinrichtung eindrucksvoll belegen.

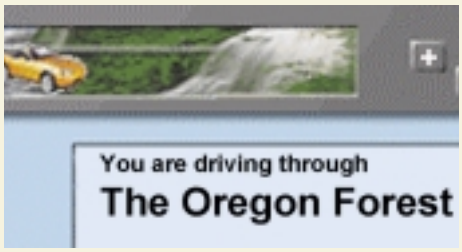
Die Herstellung des vorliegenden Forschungsmagazins wurde überwiegend durch Anzeigen finanziert. Wir danken allen Firmen und Organisationen, die unser Projekt unterstützten.

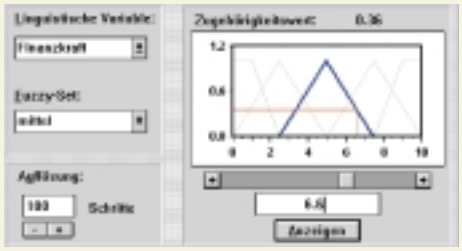
Als feste Komponente der Öffentlichkeitsarbeit und Außendarstellung der Universität erscheint jeweils im Wintersemester eine weitere Ausgabe des ForschungsForums.

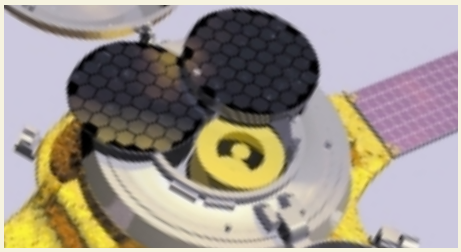
Wir bedanken uns bei Ihnen, liebe Leserinnen und Leser, für das Interesse, das Sie dem ForschungsForum Paderborn entgegenbringen, und wünschen Ihnen beim Lesen unserer vierten Ausgabe viele neue Erkenntnisse.

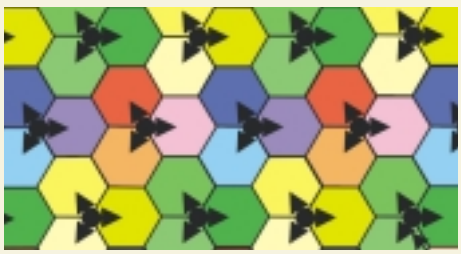
Ihre Ramona Wiesner
Referentin für Öffentlichkeitsarbeit

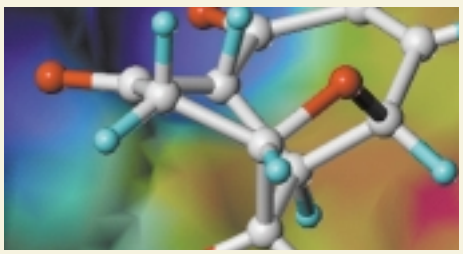


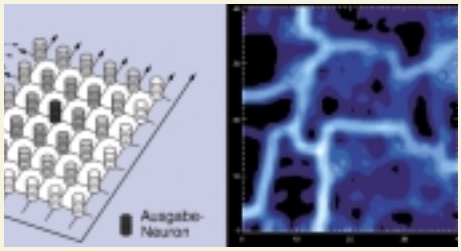


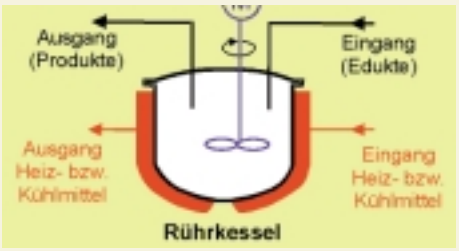


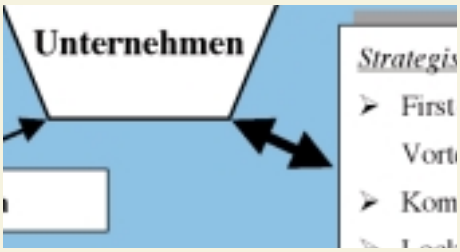












Mineraldünger	4,75
Summe Dünger	6,90
Maschinen	2,50
Input-Fossil	10,38
Input-Regenerativ	5,56
∑ Input	15,94
Output	115,14

Vorwort

Längst haben wir uns daran gewöhnt, dass sich das Wissen der Welt in immer kürzerer Zeit verdoppelt. Immer mehr wissenschaftliche Publikationen und Patentschriften dokumentieren den Fortschritt der Forschung. Die Folgen dieser an sich erfreulichen Entwicklung zeigen sich indes immer deutlicher. Es geht nicht ohne Spezialisierung. Wer an der Spitze der Forschung originäre Ergebnisse erzielen will, kann dies nur, indem er sich auf ein eng begrenztes Fach konzentriert. Gleichzeitig finden aber die wichtigen wissenschaftlichen Entdeckungen vielfach zwischen den Fächern statt.

In dem sich immer stärker herausbildenden Wettbewerb zwischen den Hochschulen haben nur die eine Chance, denen es gelingt, ein markantes Profil zu entwickeln. Auch die Universität Paderborn und ihre Fachhochschulabteilungen in Höxter, Soest und Meschede sind auf der Suche nach ihrem jeweils charakteristischen Profil. Insbesondere die Universität steht in harter Konkurrenz um Forschungsmittel. Was also ist das Charakteristische an der Paderborner Forschung? Wenn man die einzelnen Fächer für sich alleine betrachtet, so liegt Paderborn sicher nicht überall an allererster Stelle. Schaut man jedoch dahin wo sich die Fächer berühren oder überlappen, dahin wo Neues entsteht, so zählen Paderborner Forschergruppen erstaunlich oft zur Spitzengruppe. Die ausgeprägte Kommunikation und Kooperation zwischen den Disziplinen ist zweifelsohne eine Stärke unserer Hochschule. Diese Interdisziplinarität scheint auch für das Umland bedeutsam zu sein. Mit über 120 Unternehmensgründungen, die in den letzten Jahren aus der Universität hervorgegangen sind, entstanden einige Tausend neue Arbeitsplätze, meist in attraktiven Zukunftsfeldern. Es wäre jedoch falsch, den Vorteil der interdisziplinären Zusammenarbeit nur in ihrem praktischen Nutzen zu sehen.

Jedes Fach erfasst stets nur einen speziellen Aspekt der Wirklichkeit, nie die Wirklichkeit selbst. Dieses Grundproblem wird uns heute öfter bewusst denn je. Die Probleme und Herausforderungen, denen wir uns stellen müssen, sind selten rein technischer, mathematischer, soziologischer, politischer oder ökonomischer Natur, sondern gehen, öfter als uns lieb ist, über die Fächergrenzen hinweg. Es ist wichtig, sich immer wieder klarzumachen, dass die Wissenschaften, wie Karl Jaspers formulierte, „nicht wie säuberlich abgetrennte Fächer eines Aktenschrankes nebeneinander liegen“. Wir müssen wieder lernen, Wissenschaft als Ganzes zu begreifen. Es reicht eben nicht aus, die Einheit von Forschung und Lehre zu beschwören, und dann jedes Fach für sich alleine anzusehen. Die gegenseitige Ergänzung, das Geben und Nehmen über Fächergrenzen hinweg ist heute wichtiger denn je. Die ansatzweise zu beobachtende Entwicklung zur Abschottung der Disziplinen muss deshalb aufgehalten und umgekehrt werden.

Auch unsere Hochschule kommt nicht ohne Spezialisierung aus. Jeder von uns muss sich in irgend einer Weise spezialisieren, wenn er in seinem Fachgebiet Anschluss an den Fortschritt des Wissens halten will. Was dabei jedoch unbedingt vermieden werden muss, ist, zum Nur-Fachmann, oder um es ganz drastisch auszudrücken, zum Fachidioten zu werden. Wir brauchen Studierende, Wissenschaftlerinnen und Wissenschaftler, die ihr Fachgebiet beherrschen, die aber auch darüber hinaus in der Lage sind, Beziehungen zu anderen Fächern herzustellen, und mit den dort tätigen Menschen zu kooperieren. Zur Kooperation gehören partnerschaftliches Verhalten, gegenseitige Anerkennung und Achtung der unterschiedlichen fachlichen Qualifikation, aber auch persönliche Integrität und, vor allem, die Bereitschaft zu teilen. Nach Eric S. Raymond, einem der Gurus der Hacker-Kultur, erwirbt man Status und Anerkennung nicht durch Besitz, „sondern indem man etwas hergibt, insbesondere die eigene Zeit, die eigene Kreativität und die Produkte der eigenen Fähigkeiten“. Natürlich lässt sich diese Einstellung nicht verordnen. Es hängt von jedem Einzelnen ab, ob er sich in den wissenschaftlichen Austausch einbringt oder lieber in falsch verstandener „Einsamkeit und Freiheit“ vor sich hin forscht.

Ich würde mich freuen, wenn es den Autoren der vorliegenden Schrift gelungen ist, durch ihre Beiträge zur Kommunikation zwischen den Disziplinen beizutragen und bei Ihnen, liebe Leserinnen und Leser, die Lust auf mehr Information über die Forschung an unserer Universität und ihren Fachhochschulabteilungen zu wecken.

Prof. Dr.-Ing. Jörg Wallaschek
Prorektor für Forschung und wissenschaftlichen Nachwuchs

Exzitonen im Gleichtakt

Über eine Anwendung kohärenter Laser-Lichtpulse in der Halbleiterspektroskopie

Seit Einführung des Exzitonen-Konzepts durch Frenkel, Peierls und Wannier vor mehr als sechzig Jahren hat die Erforschung des Exzitons für den experimentell arbeitenden Spektroskopiker wie auch für den Festkörpertheoretiker nie an Reiz verloren. Mit dem Laser als Quelle ultrakurzer und intensiver Lichtpulse sind der Exzitonenspektroskopie in Kristallen heutzutage vor allem auch extrem schnelle auf einer Pikosekunden- und sogar kürzeren Zeitskala ablaufende Prozesse zugänglich. Laseranregung führt im Festkörper aber zu allererst zur Ausbildung eines kohärenten Anregungszustands, der die Kohärenzeigenschaften des anregenden Lasers in gewisser Weise widerspiegelt. Damit eröffnet sich die Möglichkeit, außer der Dynamik der Exzitonen besonders auch deren Kohärenzeigenschaften experimentell zu untersuchen. Durch die Entwicklung spezieller Messtechniken ist es gelungen, in Bereiche extrem kurzer Zeiten vorzudringen, in denen Kohärenzeffekte bei Exzitonen eine Rolle spielen. Im Folgenden werden Experimente beschrieben, mit denen die Exzitonenkohärenz in verschiedenen Halbleitern nachgewiesen und quantitativ untersucht werden konnte. Damit war es möglich, kohärenzerhaltende Raman-artige Prozesse von lumineszenzartigen Prozessen zu unterscheiden, bei denen die Kohärenz durch Energie- und Phasenrelaxation verloren geht, und ein bislang unerreicht detailliertes Bild des Exzitons als elementare Anregung des Festkörpers zu gewinnen.

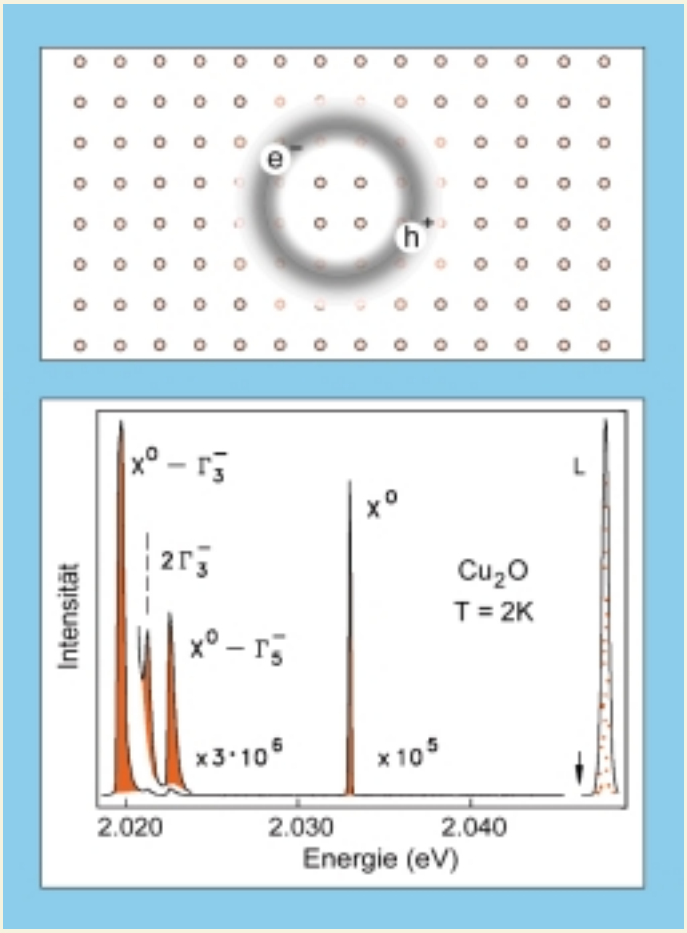


Abb. 1: Modell des Wannier-Mott-Exzitons bestehend aus Elektron (e) und positivem Loch (h; engl.: hole) im Kristallgitter (oben) und typisches Exziton-Rekombinationsspektrum (unten). Der Pfeil und L markieren die Absorptionskante und die anregende Laserlinie (s. dazu [5]). Messtemperatur: 2 Kelvin (-271°C).

Prof. Dr. rer. nat. Wolf von der Osten ist seit 1975 Professor für Experimentalphysik im Fachbereich 6/Physik an der Universität Paderborn. Arbeitsgebiet ist die optische Spektroskopie von Halbleiter- und Ionenkristallen. Einer seiner Arbeitsschwerpunkte ist die Exzitonenphysik. Professor von der Osten ist seit Ende des Wintersemesters 1999/2000 emeritiert.

Exzitonen: ein Anregungszustand des Festkörpers

Exzitonen in Halbleiter- und Ionenkristallen werden erzeugt, wenn Licht geeigneter Energie bzw. Frequenz („Farbe“) eingestrahlt und absorbiert wird (s. Ref. in [1], [2]). Dabei entstehen negative und positive elektrische Ladungen (Elektronen und positive Löcher), die sich aufgrund ihres entgegengesetzten Ladungszustands anziehen und ein insgesamt neutrales Teilchen, das Exziton, bilden (lat.: excitare = anregen). In den betrachteten Beispielen handelt es sich um das so genannte Wannier-Mott-Exziton, bei dem Elektron und Loch nur schwach aneinander gebunden sind (Abbildung 1, oben) und welches Eigenschaften ähnlich einem Wasserstoffatom besitzt.

Aus theoretischen und experimentellen Untersuchungen weiß man, dass sich Exzitonen quasi-frei durch den Kristall bewegen können. Sie transportieren dabei Energie und Impuls, die zur Erzeugung des Exzitons durch das Licht aufgebracht werden müssen. „Quasi-frei“ bedeutet, dass die Exzitonen bei ihrer Bewegung natürlich die Anwesenheit des Kristallgitters spüren. Durch inelastische Stöße, zum Beispiel mit Phononen (das sind die Elementaranregungen der Schwingungen der Kristallgitter-

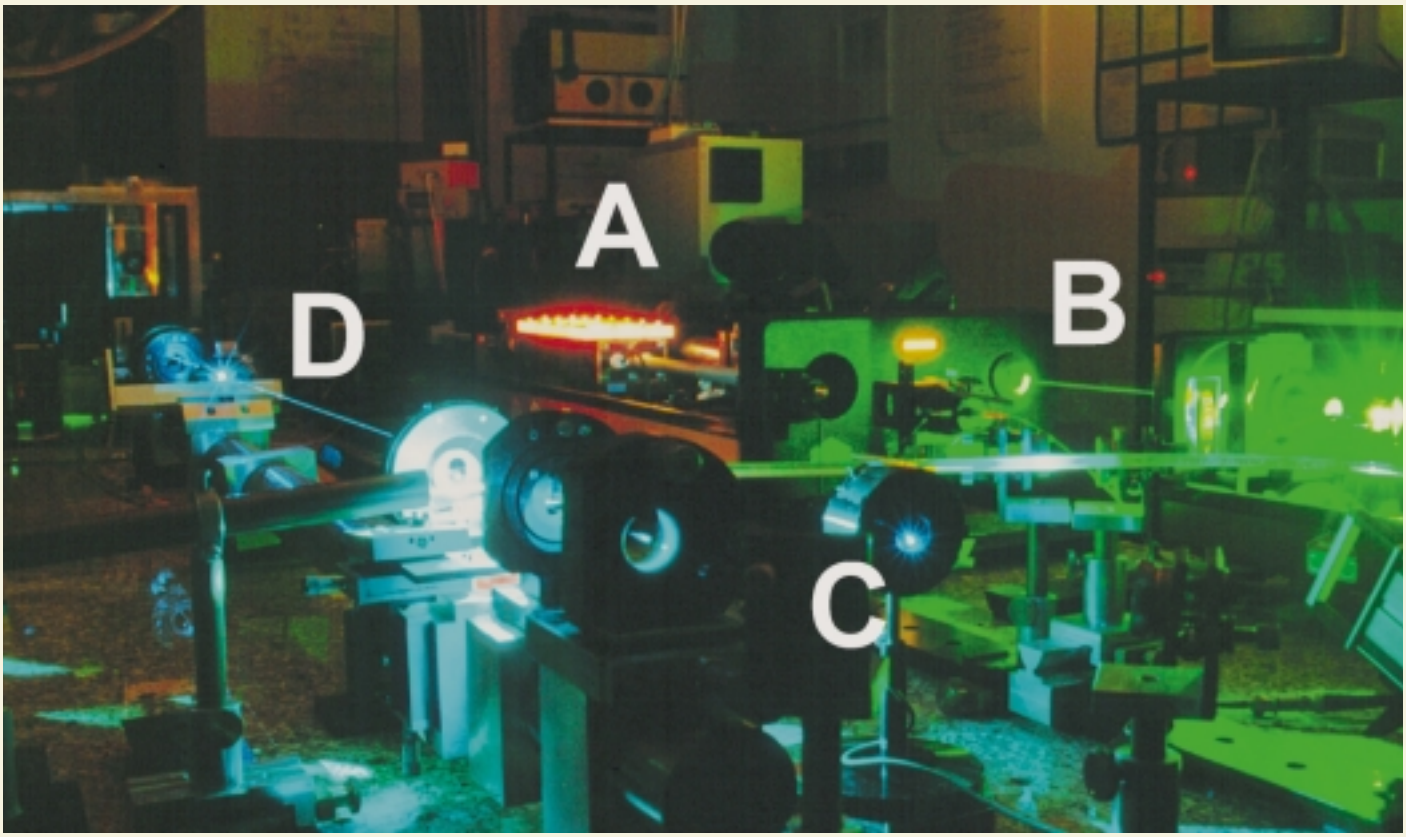


Abb. 2: Lasersystem zur Erzeugung kohärenter Pikosekunden-Lichtpulse. Infrarotstrahlung eines blitzlampengepumpten Festkörperlasers (A) wird frequenzverdreifacht und über grünes (B) in unsichtbares ultraviolettes Licht (C) umgewandelt, mit dem ein frequenzvariabler Farbstofflaser (D) gepumpt wird. Pro Sekunde verlassen den Laser 80 Millionen Pulse jeweils im Abstand von 4 Metern, die für das Auge allerdings nicht auflösbar sind, sondern als kontinuierlicher Lichtstrahl erscheinen.

bausteine), können sie einen Teil ihrer Energie und ihres Impulses verlieren. Exzitonen erleiden aber auch lediglich mit einer Impulsänderung verbundene elastische Stöße, zum Beispiel durch Streuung an unabsichtlich oder auch gezielt in den Kristallverband eingebauten Fremdverunreinigungen. Oft werden sie an solchen Verunreinigungen sogar eingefangen. Sie verlieren dann ihre Bewegungsfreiheit und werden zu gebundenen Exzitonen. Zuletzt rekombinieren Exzitonen: Elektron und Loch stürzen dann unter Emission sehr charakteristischer Strahlung und manchmal auch nichtstrahlend ineinander, wobei sie selbst vernichtet werden und der Kristall in seinen unangeregten Ausgangszustand zurückkehrt.

In vielen Kristallen liegen die Exzitonabsorptionen und -emissionen im sichtbaren Bereich des Spektrums, oft aber auch im angrenzenden unsichtbaren Infrarot- und Ultraviolettgebiet. Da Exzitonen in manchen Halbleitern und Ionenkristallen thermisch nicht stabil sind, erfordert die Messung dieser Spektren die Abkühlung der Kristalle mit flüssigem Stickstoff oder flüssigem Helium bis herab zu Temperaturen von wenigen Kelvin. Als Beispiel für ein typisches Rekombinationsspektrum zeigt Abbildung 1, unten, einen Ausschnitt der Emission der „gelben" Exzitonenserie in Kupferoxidul (Cu_2O), einer Paradesubstanz für die Exzitonenphysik mit mehreren Exzitonenserien in verschiedenen Spektralbereichen. Der Großteil der Modellvorstellung, die wir heute vom Exziton besitzen, entstammt der Analyse solcher Spektren.

Kohärenz spielt eine zentrale Rolle in der Physik

Der Begriff der Kohärenz spielt in allen Teilgebieten der modernen Physik eine zentrale Rolle. Anschaulich und verständlich wird er anhand von Wellen: Man spricht von der Kohärenz eines Wellenfeldes, wenn definierte Phasenbeziehungen zwischen sich überlagernden Teilwellen bestehen, sodass Interferenzphänomene beobachtet werden können. Ein Beispiel aus der Optik ist der Laser, mit dem Interferenzexperimente wie etwa Thomas Young's berühmter Doppelspalt-Versuch ziemlich einfach durchzuführen sind; anders als jede andere Lichtquelle liefert der Laser nämlich als Folge eines speziellen Emissionsmechanismus weitgehend kohärente Strahlung, in der alle erzeugten Wellen in Phase emittiert werden.

Statt wie in Abbildung 1 als Elektron-Loch-Paar im Teilchenbild lassen sich Exzitonen auch durch (quantenmechanische) Wellenfunktionen beschreiben. Sie sind wie andere Wellen durch Frequenz (bzw. Energie), Amplitude und Phase charakterisiert und breiten sich im Kristall aus. Durch einen kohärenten Lichtpuls in Resonanz, d.h. gleichfrequent mit dem Exziton werden die Exzitonenwellen kohärent erzeugt und laufen mit fester Phasenbeziehung. Durch die beschriebenen Stoßprozesse des Exzitons im Kristall werden die Phasen jedoch gestört und die Wellen geraten aus dem Takt. Damit geht die Kohärenz verloren. Noch bis vor etwa zehn Jahren fehlte der experimentelle Nachweis für die Kohärenz von Exzitonen. Obwohl ihre Lebensdauer in vielen Materialien bereits bekannt war, gab es keinerlei Erkenntnisse über die optische Kohärenzzeit T_2, die angibt wie lange im Mittel die Kohärenz erhalten bleibt. Auch war unbekannt, inwieweit inelastische und elastische Stöße des Exzitons, charakterisiert durch die Energierelaxationszeit („Lebensdauer") T_1 bzw. die reine Phasenrelaxationszeit T_2', zur Zerstörung der Kohärenz beitragen. Man vermutete, dass diese Zeiten extrem kurz seien und sich dem Experiment entziehen würden. Die beschriebenen (linearen) Verfahren der Quanten- und Propagati-

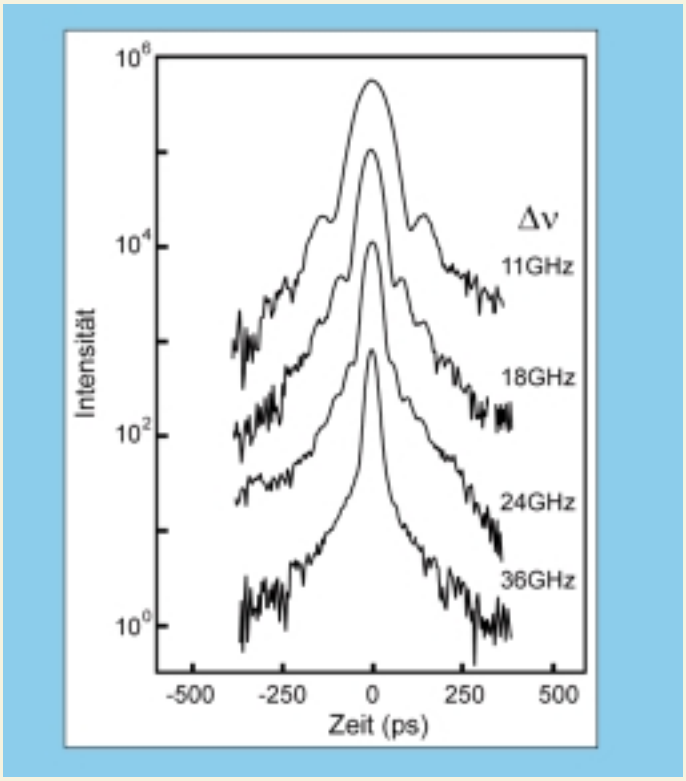


Abb. 3: Wirkung der spektralen Filterung auf das Zeitverhalten eines kohärenten Pikosekunden-Lichtpulses. $\Delta\nu$: spektrale Bandbreite des Monochromators.

onsschwebungen sowie sich gleichzeitig entwickelnde nichtlineare Methoden (siehe weitere Beiträge in [2]) ermöglichten schließlich solche Untersuchungen. Die Quantenschwebungs-Spektroskopie war bereits als Methode zur Erforschung kohärenter Prozesse in der Atom- und Molekülphysik etabliert. Allerdings liegen dort die zu messenden Abkling- und Oszillationszeiten in einem Bereich zwischen Mikro- und Nanosekunden (10^{-6}-10^{-9} Sekunden). Für exzitonische Zustände laufen diese Prozesse wesentlich schneller ab (in 10^{-9}-10^{-12} Sekunden), sodass es entscheidender Fortschritte in der Laser- und Messtechnik bedurfte, ehe Messungen auch an diesen Systemen möglich wurden.

Spektroskopie mit kohärenten Laser-Lichtpulsen

Entscheidend für die Kohärenzuntersuchungen waren die Fortschritte in der Lasertechnik. Während aus konventionell gemessenen Exzitonenspektren zum Beispiel in Verbindung mit äußeren Magnetfeldern in vielen Kristallen bereits ein komplettes Bild der Energiezustände vorlag, boten sich der Exzitonenspektroskopie – wie der Festkörperspektroskopie insgesamt – mit dem Laser völlig neue Messmöglichkeiten. Mit ihm steht eine konkurrenzlose Quelle für kurze und kohärente Lichtpulse zur Verfügung, die außerdem noch über große Spektralbereiche frequenzveränderlich sein können.

Für Untersuchungen kohärenter Prozesse und ihrer Dynamik muss die Pulsdauer im Vergleich zu den relevanten Relaxationszeiten klein sein, d.h. $\Delta t_p < T_1, T_2$. Für unsere Exzitonen bedeutet das Pulsdauern von wenigen Pikosekunden (1 ps = 10^{-12} s = 0,000 000 000 001 s). Sind schon die hundertstel und tausendstel Sekunden unglaublich kurz, die heute im Sport gemessen werden und über Sieg oder Niederlage entscheiden, so entziehen sich diese Zeiten der menschlichen Vorstellungskraft vollständig.

Durch die Fortschritte in Optik und Elektronik stehen jedoch dem Physiker im Labor die Mittel zur Verfügung, solche ultrakurzen Pulse reproduzierbar zu erzeugen (Abbildung 2) und ihren zeitlichen Verlauf sowie die Pulsdauer präzise zu bestimmen.

Bei der Spektroskopie mit Pulsen in diesem extremen Zeitbereich treten bemerkenswerte Besonderheiten zutage, die unmittelbare Auswirkungen auf die Messung haben. So legt Licht verschiedener Frequenzen in einem Monochromator, mit dem die zeitabhängigen Signale spektral analysiert werden müssen, im Allgemeinen unterschiedliche Wege zurück, die sich um einige Zentimeter Länge unterscheiden können. Aufgrund der Lichtgeschwindigkeit entspricht dies Signalverbreiterungen im Zeitbereich bis zu 100 ps und führt zur Verfälschung der Signale, die sich nur durch besonders konstruierte Monochromatoren verhindern lassen.

Eine weitere Besonderheit betrifft den Lichtpuls. Kohärenz eines Pulses bedeutet, dass ein reziproker Zusammenhang zwischen der Dauer und der spektralen Breite des Pulses besteht: je kürzer der Puls desto breiter das Spektrum. Quantitativ sind die zeitliche und die spektrale Gestalt über eine Fourier-Transformation verknüpft. Dieser Zusammenhang wird zwar ausgenutzt, um die Pulskohärenz experimentell zu überprüfen, er bedeutet aber auch eine Beschränkung in der Wahl der Bandbreiten. Während die kürzesten heute darstellbaren Laser-Lichtpulse von 4 Femtosekunden (1 fs = 10^{-15} s = 0,000 000 000 000 001 s) für die Messung extrem kurzlebiger Relaxationsvorgänge ideal wären, erstreckt sich ihr Spektrum über 1 bis 2 Elektronenvolt (eV), d.h. zum Beispiel über den gesamten sichtbaren Bereich, sodass diese Pulse für die Anregung der schmalbandigen Exzitonenübergänge nicht in Frage kommen. Pikosekunden-Lichtpulse dagegen besitzen spektrale Breiten von nur wenigen Millielektronenvolt (meV) und sind damit für die Experimente ideal geeignet.

Der Effekt der Bandbreitenbegrenzung wirkt sich natürlich auch auf die eigentlichen Messsignale aus, die sowohl im Hinblick auf ihr Spektrum als auch auf ihre Zeitabhängigkeit untersucht werden müssen [1], [2]. Abbildung 3 zeigt dafür als Beispiel den Zeitverlauf eines durch einen Monochromator laufenden Pikosekundenpulses. Mit Reduktion der Spaltbreite, d.h. der spektralen Bandbreite des Monochromators und der damit verbundenen Erhöhung der spektralen Auflösung, verlängert sich die Pulsdauer deutlich und führt gleichzeitig, konsistent mit der theoretischen Erwartung, zur Entstehung kleiner Nebenmaxima im Zeitverlauf. Als wichtige Konsequenz folgt daraus, dass die Messungen der zeitaufgelösten Emissionen bandbreitenbegrenzt erfolgen müssen, d.h. mit angepasster, jeweils optimaler zeitlicher und spektraler Auflösung.

Quanten- und Propagationsschwebungen

Schwebungen sind aus der Akustik geläufig. Sie entstehen bei der (kohärenten) Überlagerung zweier in ihren Frequenzen geringfügig verschiedenen Schwingungen oder Wellen, etwa wenn zwei wenig verstimmte Stimmgabeln in räumlicher Nähe zueinander angeschlagen werden oder ein Gitarrespieler die Saiten seines Instrumentes stimmt. Das Ohr vernimmt dann einen Ton ungefähr bei der ursprünglichen Frequenz der Saiten mit langsam an- und abschwellender Lautstärke: Der Ton ist im Takt der sehr geringen Differenzfrequenz amplitudenmoduliert.

Ganz ähnlich kommen exzitonische Quantenschwebungen

zustande. Wie anhand des 3-Niveau-Systems in Abbildung 4, oben, schematisch gezeigt ist, werden hier zwei Exzitonenzustände der Energien E_1 und E_2 (quantenmechanisch) überlagert. Die Überlagerung wird durch einen kohärenten Lichtpuls erreicht, dessen Energiebreite die beiden Zustände überdeckt, und führt in der emittierten Intensität zu einer Schwebung, die einem exponentiellen Abfall überlagert ist. Die Schwebungsperiode entspricht dabei dem reziproken Energieabstand der Zustände. Die Schwebung selbst, die dem kohärenten (Raman-artigen) Emissionsanteil entspricht, klingt mit der Kohärenzzeit T_2 des Exzitons ab, der exponentielle nicht kohärente (lumineszenzartige) Untergrund mit der Exzitonen-Lebensdauer T_1. Damit werden diese wichtigen Zeiten messbar.

Das erste erfolgreiche Quantenschwebungsexperiment an Exzitonen überhaupt wurde an freien, im Magnetfeld aufgespaltenen Exzitonenzuständen in Silberbromid (AgBr) durchgeführt [3]. Die gefundenen Oszillationen waren zwar nicht stark ausgeprägt, die Messungen lieferten aber überraschend „lange“ Kohärenzzeiten von 400 ps und bewiesen, dass die Methode der Quantenschwebungs-Spektroskopie auch für Exzitonen anwendbar ist. Abbildung 4, unten, zeigt ein Beispiel für die später an verschiedenen freien und gebundenen Exzitonen in Cu_2O und Cadmiumsulfid (CdS) beobachteten Magneto-Quantenschwebungen mit T_1- und T_2-Zeiten von einigen zehn bis hundert Pikosekunden. Aufgrund der Kristallsymmetrie und damit verbundener Auswahlregeln hängen die Schwebungen in Kristallen von der Lichtpolarisation und der Magnetfeldrichtung ab. Auf dem Umweg über die Messung der Periodendauer lassen sich selbst geringste, im Spektrum nicht auflösbare Aufspaltungen des Exzitonenzustands als Funktion der Magnetfeldstärke sehr genau bestimmen und daraus wichtige Exzitonenparameter gewinnen.

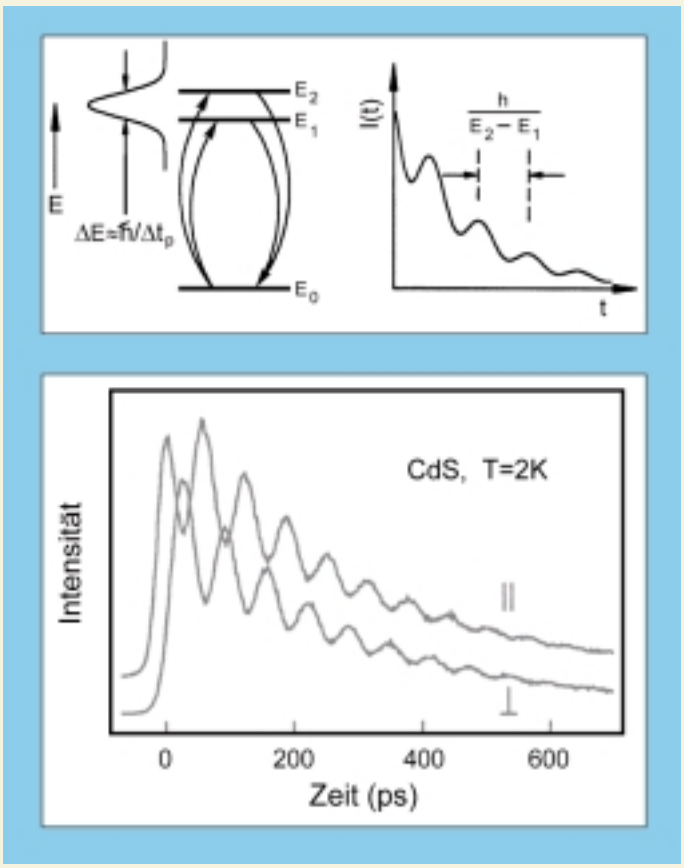


Abb. 4: Prinzip der Quantenschwebungen in Emission (oben) und Magneto-Quantenschwebung für ein gebundenes Exziton in CdS in einem Magnetfeld B = 2 Tesla bei 2 K (unten). E_0 und E_1, E_2 bezeichnen den Kristallgrundzustand und zwei Exzitonenzustände, || und ⊥ Lichtpolarisationen [4].

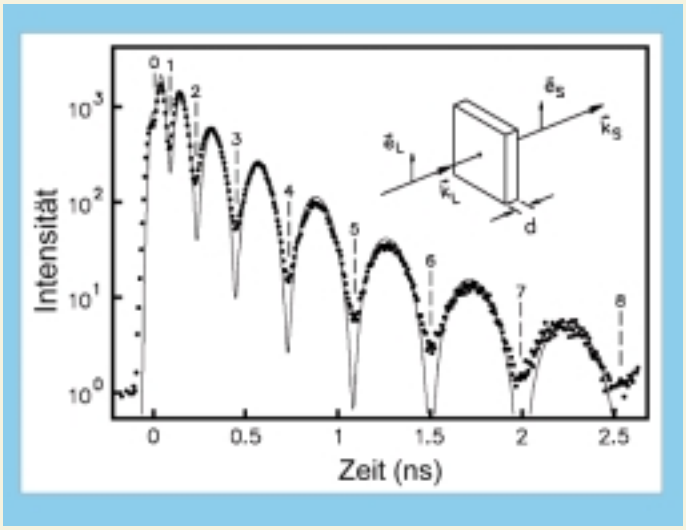


Abb. 5: Propagationsschwebung des Exziton-Polaritons in Cu_2O bei T = 2 K nach Anregung mit einem 30-ps-Lichtpuls [6]. Die ausgezogene Linie ist eine Anpassung der experimentellen Daten (Punkte) an ein Polaritonmodell.

Eine interessante Veränderung des Schwebungsverhaltens tritt zutage, wenn in den Experimenten die Wechselwirkung des Exzitons mit der Laserpuls-Lichtwelle (oder „Photonen“) eine Rolle spielt. Exziton und Photon können dann nicht mehr als unabhängige Elementaranregungen betrachtet werden. Vielmehr wird eine neue als Exziton-Polariton bekannte Anregung erzeugt, bei der – ähnlich einem gekoppelten Pendel – ein ständiger Energieaustausch zwischen Photon und Exziton stattfindet. Resultat dieser Kopplung ist eine auch ohne äußeres Magnetfeld bereits vorhandene intrinsische Aufspaltung des Exzitonzustands in zwei Polaritonenäste. Wie in Abbildung 5 am Beispiel des Polaritons in Cu_2O gezeigt ist, führt die kohärente Überlagerung in diesem Falle zu einer mit der Zeit zunehmenden Periode [6]. Ursache dafür ist die Propagation der Polaritonen. Der Laser-Lichtpuls erzeugt bei Eintritt in den Kristall (siehe Innenbild) auf beiden Ästen Polaritonenwellen, die sich kohärent durch den Kristall bewegen und nach ihrer Rückverwandlung in Licht auf der Rückseite bei Austritt aus der Probe interferieren. Wie bei jeder anderen Interferenz entsprechen die Maxima und Minima in der transmittierten zeitlich oszillierenden Lichtintensität einer Lichtverstärkung und -auslöschung („konstruktive und destruktive Interferenz“) aufgrund der Phasenverschiebungen, welche die Polaritonenwellen auf ihrem Weg durch die Probe erfahren. Das eigentlich Erstaunliche an diesen Experimenten (und zugleich der Beweis für exzellente Kristallqualität) ist, dass die Kohärenz der Polaritonen über makroskopische Probenlängen von Millimetern erhalten bleibt und durch Stöße im Kristall offenbar nur wenig beeinträchtigt wird. Im Vergleich zum einfallenden Lichtpuls ist der durchgelassene Puls zeitlich stark verlängert und selbst nach einigen Nanosekunden noch nicht abgeklungen. Dieses Ergebnis zeigt, dass die Polaritonen bis zu etwa tausendmal langsamer durch den Kristall propagieren als es ein einfacher Lichtpuls tun würde. Überdies bewegen sich nicht alle Polaritonen mit gleicher Geschwindigkeit, vielmehr hängt die Geschwindigkeit von der Energie der Polaritonenäste ab. Die quantitative Analyse dieser Ergebnisse gestattet eine sehr genaue Bestimmung dieser als Polaritonendispersion bezeichneten Abhängigkeit. Neben wichtigen anderen Größen liefert sie sehr genaue

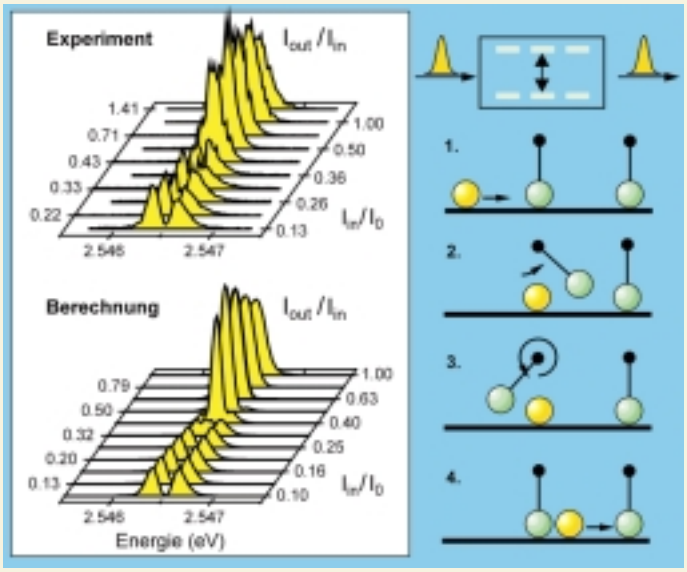


Abb. 6: Selbstinduzierte Transparenz kohärenter Lichtpulse in Resonanz mit einem gebundenen Exziton in CdS. Links: Auswirkung auf das transmittierte Pulsspektrum bei T = 2 K. Rechts: Pendelmodell zur Veranschaulichung der verlustfreien Pulspropagation.

Geschwindigkeitswerte und ermöglicht so einen detaillierten Einblick in die Polaritondynamik im Kristall [2], [6].

Selbstinduzierte Transparenz

Ein interessantes von Atomen her bekanntes, auf Kohärenz beruhendes Phänomen, das von uns erstmals bei gebundenen Exzitonen nachgewiesen werden konnte, ist die selbstinduzierte Transparenz [7], [8]. Man versteht darunter die verlustfreie Propagation eines kohärenten Lichtpulses, der die Kristallprobe in exakter Resonanz mit absorbierenden Zwei-Niveau-Systemen durchläuft: Anders als nach dem Beer'schen Absorptionsgesetz der inkohärenten Optik, wonach die Pulsintensität mit zunehmender Absorptionsstrecke exponentiell abnimmt, wird der Puls nicht abgeschwächt (Abbildung 6, rechts oben). Dieser auf den ersten Blick überraschende Effekt ist Folge einer nichtlinearen Wechselwirkung zwischen dem kohärenten Lichtpuls und dem Exzitonensystem, die bei genügend großer Intensität des Eingangspulses ins Spiel kommt.

Die bei gebundenen Exzitonen geltende Bedingung für kohärente Anregung ($\Delta t_p < T_2$) bedeutet, dass die spektrale Breite des anregenden Pulses größer ist als die (homogene) Breite der Exzitonenabsorption. Im Spektrum des transmittierten Pulses, dessen spektrale Lage genau auf die Exzitonresonanz abgestimmt ist, wird daher bei kleiner Eingangsintensität als Folge der Absorption ein „spektrales Loch" detektiert (Abbildung 6, links, I_{in}/I_0 klein). Parallel dazu kommt es zu Propagationsschwebungen ähnlich wie in Abbildung 5, welche die Kohärenzeigenschaften des Exzitons widerspiegeln (im Beispiel: T_2 = 80 ps). Bei schrittweiser Erhöhung der Eingangsintensität findet man eine scharfe Intensitätsschwelle (in Abbildung 6 bei I_{in}/I_0 = 0.43), in der eine grundlegende Änderung der Wechselwirkung sichtbar wird. Das Überschreiten dieser markanten Grenze ist im Spektrum durch das Verschwinden des spektralen Lochs und durch einen drastischen nichtlinearen Anstieg der transmittierten Intensität gekennzeichnet. Im Zeitbereich verschwinden gleichzeitig die Propagationsschwebungen und werden abgelöst durch einen „glatten" form- und flächenstabilen, gegen den Eingangspuls um einige Pikosekunden verzögerten Puls, ein sogenanntes Soliton. Obwohl in exakter Resonanz mit der Exzitonenabsorption, propagiert der Puls nahezu unbeeinflusst durch das Medium.

Abbildung 6, rechts, zeigt ein mechanisches Analogon, dass die selbstinduzierte Transparenz sehr schön veranschaulicht. Eine Kugel rollt auf ein Pendel gleicher Masse zu und überträgt durch Stoß ihre gesamte Energie, die ihr nach einer Rotation des Pendels wieder vollständig zurückgegeben wird. Die Kugel entspricht dem anregenden Puls, die Pendel den resonanten exzitonischen Zwei-Niveau-Systemen und die Rotation dem kohärenten Anregungs- und Emissionsprozess. Vernachlässigt man Reibung und Verformung (d.h. die Exzitonenrelaxation), bewegt sich die Kugel ohne Energieverlust durch die Pendelhindernisse. Durch die Rotation erfährt sie eine Verzögerung, die auch für den Puls gefunden wird und sich dort durch die kurzzeitige Speicherung der Energie in den absorbierenden Exzitonen erklärt.

Die in diesem kurzen Rückblick auf 10 Jahre Forschung auf dem Gebiet der Exzitonenkohärenz berichteten Ergebnisse sind im Rahmen der Dissertationsarbeiten von Elmar Schreiber, Volkmar Langer und Michael Jütte sowie in einer Kooperation mit meinem Dortmunder Kollegen Dietmar Fröhlich entstanden. Ihnen sei an dieser Stelle herzlich gedankt. Besonderer Dank gilt meinem langjährigen Mitarbeiter Heinrich Stolz für seine entscheidenden experimentellen und theoretischen Beiträge zur Entwicklung dieser Methode der Kohärenzspektroskopie. Er ist inzwischen einem Ruf auf eine Professur für Halbleiterphysik an der Universität Rostock gefolgt und führt die Arbeiten dort erfolgreich fort. Schließlich danke ich der Deutschen Forschungsgemeinschaft sowie der Universität Paderborn für die langjährige Förderung unserer Arbeiten.

Literatur

[1] H. Stolz: Time-Resolved Light Scattering from Excitons, Springer Tracts in Modern Physics Vol. 130 Springer, Berlin 1994.

[2] W. von der Osten, V. Langer, H. Stolz in: Coherent Optical Interactions in Semiconductors (ed. R. Philips), NATO ASI Series, Plenum Press, New York 1994, S. 111.

[3] V. Langer, H. Stolz, W. von der Osten, Phys. Rev. Lett. 64, 854 (1990).

[4] H. Stolz, V. Langer, E. Schreiber, S. Permogorov, W. von der Osten, Phys. Rev. Lett. 67, 679 (1991).

[5] V. Langer, H. Stolz, W. von der Osten, Phys. Rev. B 51, 2103 (1995-II).

[6] D. Fröhlich, A. Kulik, B. Uebbing, A. Mysyrowicz, V. Langer, H. Stolz, W. von der Osten, Phys. Rev. Lett. 67, 2343 (1991).

[7] M. Jütte, H. Stolz, W. von der Osten, J. Opt. Soc. Am. B13, 1205 (1996).

[8] M. Jütte, W. von der Osten, J. Luminesc. 83/84, 77 (1999) & phys. stat. sol. (6) 215, 59 (1999).

Ein Irrtum des Geistes ist dasselbe wie ein Irrtum im Kalkül

Gottfried Wilhelm Leibniz: Die Grundlagen des logischen Kalküls

Im Fachbereich 1/Philosophie wurde im März 2000 ein dreijähriges, von der Fritz Thyssen Stiftung (Köln) unterstütztes Forschungsprojekt zur Logik von Gottfried Wilhelm Leibniz (1646-1716) abgeschlossen. Das Projekt stand unter der Leitung von Prof. Dr. Dr. Franz Schupp. Die Ergebnisse dieser Arbeit sind in Form einer zweisprachigen kommentierten Studienausgabe veröffentlicht (vgl. Literaturangabe).
Die Beschäftigung mit der Logik von Leibniz ist nicht allein den Lehrenden und Studierenden des Faches Philosophie vorbehalten. Wer sich heute z.B. Kenntnisse über die Vorgeschichte einer modernen Disziplin wie der Informatik aneignen möchte, sei es aus rein historischem Interesse oder sei es aus der Hoffnung, auf diese Weise vielleicht auch Anregungen für die eigene gegenwärtige Arbeit zu erhalten, wird unweigerlich zur Logik von Leibniz gelangen.

Gottfried Wilhelm Leibniz war der Überzeugung, jeder Denkfehler sei als Rechenfehler zu identifizieren. Dieser für seine logischen Schriften grundlegende Gedanke soll im Folgenden exemplarisch vorgestellt werden. Da jedoch dem Leser bei der Beschäftigung mit einer Studienausgabe zur Logik von Leibniz häufig die grundsätzlichen Fragen und Schwierigkeiten, vor die sich ein Bearbeiter von Leibniz-Texten gestellt sieht, verborgen bleiben, seien zunächst einige Informationen zur Textsituation gegeben.

Leibniz-Texte und ihre Schwierigkeiten

1. Die lateinische Sprache. Die erste Schwierigkeit ist bereits durch die verwendete Sprache gegeben, da Leibniz seine logischen Schriften auf Latein verfasste. Es kann jedoch nicht vorausgesetzt werden, dass alle diejenigen, die sich für die Logik von Leibniz interessieren, auch über ausreichende Kenntnisse der lateinischen Sprache verfügen, um mit diesen Texten arbeiten zu können, und selbst ein irgendwann in der Schule erworbenes Latinum garantiert noch nicht die problemlose Lektüre. Aus diesem Grund haben wir uns entschieden, in unserer Studienausgabe den lateinischen Texten ihre deutsche Übersetzung gegenüberzustellen und sie so einem breiteren Leserkreis zugänglich zu machen. Der lateinische Text bietet darüber hinaus dem Latein-Kundigen die Möglichkeit, die Übersetzung selbst zu überprüfen.

2. Die große Fülle des Leibniz-Materials. Leibniz hat zu seinen Lebzeiten sehr viel geschrieben, fast alle Texte, die wir in unserem Projekt bearbeitet haben, blieben von ihm jedoch unveröffentlicht. Deren Veröffentlichung erfolgte erst später in verschiedenen Ausgaben, u.a. im Band VI, 4, der Akademieausgabe, der 1999 erschienen ist. Allein dieser Band, der die philosophischen Schriften von Leibniz der Jahre 1677 bis Juni 1690 enthält, hat – den zusätzlichen Registerband nicht eingerechnet – einen Umfang von 2949 Seiten, von denen 1002 Seiten auf den für uns besonders relevanten Teil der Allgemeinen Wissenschaft *(Scientia generalis)* entfallen. In Anbetracht solcher Dimensionen wird es verständlich, dass selbst ein interessierter Nicht-Spezialist ohne zusätzliche Hilfen in Form eines ausführlichen Kommentars zu den entsprechenden Texten sehr schnell an seine Grenzen stößt, und es fällt nicht schwer sich vorzustellen, dass bis zum Abschluss der kritischen Gesamtausgabe der Leibnizschen Werke noch viele Jahrzehnte vergehen werden.

Stephanie Weber war von März 1997 bis März 2000 wissenschaftliche Angestellte in dem von der Fritz Thyssen Stiftung unterstützten Forschungsprojekt „Leibniz, die Grundlagen des logischen Kalküls“ unter der Leitung von **Prof. Dr. Dr. Franz Schupp**. Sie erhielt zwischenzeitlich ein Stipendium der Weidmüller Stiftung und ist seit Oktober 2000 wissenschaftliche Mitarbeiterin (von Prof. Dr. Dr. Franz Schupp) im Fachbereich 1/Philosophie.

3. Der Fragment-Charakter der Leibniz-Texte. Leibniz schrieb nicht nur sehr viel, sondern er hatte beinahe ununterbrochen gute Einfälle, die er häufig zu Papier brachte, noch bevor der vorausgehende Gedanke zu Ende formuliert war. Zudem arbeitete er mit verschiedenen Zeichensystemen, wechselte mitunter auch mitten in einem Text von einem System auf ein bereits früher verwendetes, und es kommt durchaus vor, dass er einen Text mit einem Zeichensystem beginnt, im Laufe seiner Arbeit dieses jedoch modifiziert, ohne dabei den vorausgehenden Text entsprechend zu ändern. Und auch Leibniz „verrechnete“ sich gelegentlich in seinen Kalkülen, wobei es sich meist um kleinere Fehler handelt, die nicht den Gedanken selbst in Frage stellen. Da er jedoch nach der raschen Niederschrift seiner Gedanken den Text in der Regel beiseite legte, ohne ihn noch einmal durchzusehen, weisen viele seiner Texte einen sehr fragmentarischen Charakter auf, durch den sich die Lektüre mühsam gestaltet, für einen breiten Leserkreis sogar zum unüberwindbaren Hindernis wird.
Ziel unseres Projekts war es daher, anhand einer sinnvollen

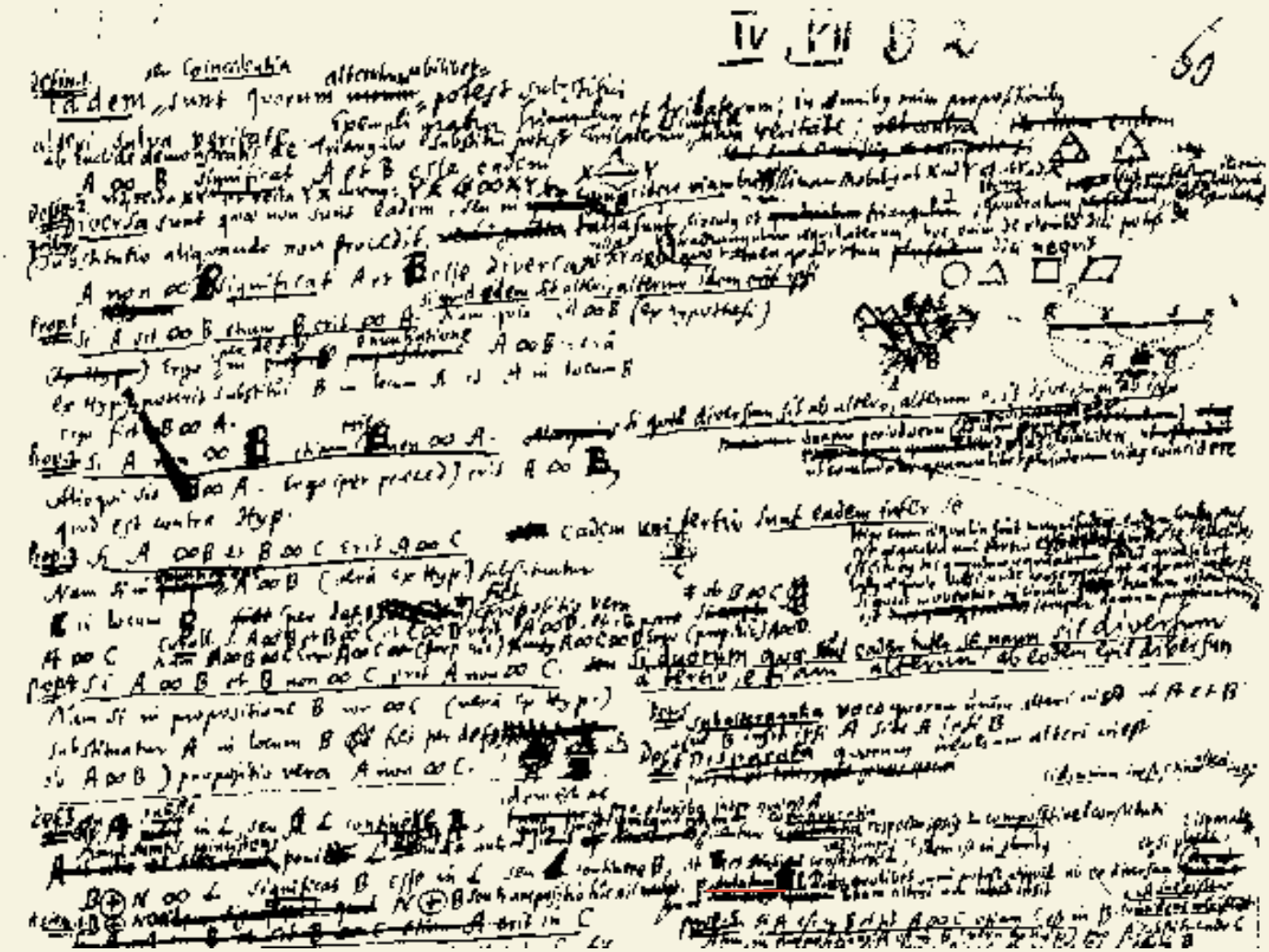


Abb. 1: Ausschnitt aus einer Leibniz-Handschrift. Hannover, Niedersächsische Landesbibliothek, LH IV 7B, 2 Bl. 60r°.

Auswahl und Zusammenstellung einiger (es sind insgesamt zehn) Texte und ihrer Übersetzungen, einer einführenden „Anleitung" und einem Kommentar die Logik von Leibniz auch dem Nicht-Spezialisten zugänglich zu machen.

Auf dem Weg zum gedruckten Text – Die Leibniz-Handschriften

Die Herausgabe von Leibniz-Texten bringt eigene Schwierigkeiten mit sich, die an dieser Stelle nur angedeutet werden können. Sieht man sich einmal einen Ausschnitt einer Handschrift an (Abbildung 1), so gewinnt man eine erste Vorstellung von der Aufgabe eines Editors. Leibniz hat sehr klein geschrieben, sicherlich auch, um Papier zu sparen, die Schrift selbst ist jedoch verhältnismäßig gut zu entziffern. Sehr viel mühsamer gestaltet sich die Zusammenstellung und Organisation des Textes. Auch Leibniz konnte bei einem solchen Text den Überblick verlieren, was ihn dann veranlasste, diesen zu organisieren. In Abbildung 1 lassen sich, besonders am linken Rand, sehr gut die späteren Hinzufügungen („Defin. 1", „Defin. 2", „Prop. 1", „Prop. 2" usw.) erkennen, durch die Leibniz den Text im Nachhinein gliederte. In der rechten Spalte hingegen fallen Hinzufügungen auf, die beinahe den gesamten Platz ausfüllen. Auf diese Weise lässt sich der Entstehungsprozess des Textes nachkonstruieren, was für die Textanalyse sehr aufschlussreich ist. Auf dieser Kopie sieht man z.B., dass das Postulat 1 nachträglich hinzugefügt wurde, was natürlich den Kalkül selbst verändert und die Frage aufwirft, an welcher Stelle der Entwicklung des Kalküls sich Leibniz veranlasst sah, eine solche Hinzufügung vorzunehmen. Auch solchen Fragen ist in der Einleitung und im Kommentar nachzugehen, um ein adäquates Verständnis des jeweiligen Kalküls zu ermöglichen, denn natürlich kann nicht alles, was interessant ist, im gedruckten Text selbst angemessen wiedergegeben werden.

Wenden wir uns nun dem Grundgedanken der leibnizschen Logik zu:

Denken als regelgeleiteter Prozess der Zeichentransformation

Leibniz war der Überzeugung, unser Denken sei ein Zeichenprozess, der durch logische Regeln geleitet wird. Wenn es uns daher gelingt, diese Zeichen in ihrem logischen, nicht in ihrem psychologischen Gehalt abzubilden und ihre Transformationsregeln anzugeben, dann lassen sich unsere Gedankenprozesse wie logische Kalküle, d.h. wie Rechenprozesse, darstellen. Auf diese Weise können alle Denkfehler als Rechenfehler nachgewiesen werden, denn dann wird Denken *(raisonner)* und Rechnen *(calculer)* ein und dasselbe sein (Akademieausgabe VI, 4, S. 1030), d.h., es gilt, wie Leibniz in Text [II] unserer Sammlung formuliert: Ein Irrtum des Geistes ist dasselbe wie ein Irrtum im Kalkül (S. 19). Dieser Gedanke liegt Leibniz' Großprojekt einer Allgemeinen Wissenschaft *(Scientia generalis)* zu Grunde, mit der er das Ziel verfolgte, eine Enzyklopädie aller Wissenschaften zu erstellen und für diese eine allgemeine, auf alle Wissenschaf-

ten anwendbare Methode zu erarbeiten. Dazu ist es notwendig, die Grundbegriffe aller Wissenschaften, d.h. die Begriffe, die sich nicht weiter in andere Begriffe zerlegen lassen, und damit das Denkbare überhaupt, aufzufinden. Über einzelne formale Probleme hinaus, mit denen sich Leibniz in seinen Kalkülen beschäftigte, war er davon überzeugt, diese Methode könnte sehr viel zum Glück des Menschengeschlechts beitragen (Text [I], S. 5). Dieses „Lebenswerk", von dem Leibniz lange Zeit annahm, es lasse sich in einem überschaubaren Zeitraum realisieren, umfasst zwei Komponenten:

I. Die allgemeine Wissenschaftssprache *(Characteristica universalis)*, das wahre Instrument *(Verum Organon)* der Allgemeinen Wissenschaft, mit deren Hilfe die begrifflichen Zusammenhänge, die in der Allgemeinen Wissenschaft zu erarbeiten sind, in einer Zeichensprache abgebildet werden sollen, bzw. eine für bestimmte Wissenschaften spezifische Sprache, durch die sich alle Grundbegriffe eines wissenschaftlichen Bereichs mittels geeigneter Zeichen ausdrücken lassen. Solche Zeichen können auch Zahlen sein, eine Möglichkeit, die Leibniz selbst als viel versprechend ansah. Nehmen wir z.B. die traditionelle Definition des Menschen als einem vernünftigen Lebewesen *(homo est animal rationale)*. Sei „Vernünftiges" $=_{def.}$ 5 und „Lebewesen" $=_{def.}$ 3, dann ist „Mensch" = 5 x 3 = 15. Mit der Aussage „Jeder Mensch ist ein Lebewesen" wird nun gesagt, dass die charakteristische Zahl für „Mensch", also 15, die charakteristische Zahl für „Lebewesen", d.h. 3, als Teiler enthalten muss, was der Fall ist. Zudem müssen in dem anderen Teiler der Division, d.h. in 5, alle diejenigen Prädikate enthalten sein, die in „Mensch" enthalten sind, jedoch über „Lebewesen" hinaus von ihm gelten. Die verwendeten Zeichen sollen nicht einfache konventionelle Merkzeichen sein, sondern sie sollen „charakterisieren", welche wesentlichen Merkmale oder Eigenschaften in einem Begriff enthalten sind. Um mit diesen Zeichen arbeiten zu können, benötigt man

II. den logischen Kalkül (*Calculus logicus* oder *Calculus ratiocinator*), durch den die Transformationsregeln für die Zeichen festgelegt werden. Dazu bemerkt Leibniz: „Ein Kalkül bzw. eine Operation besteht im Hervorbringen von Relationen, was bewirkt wird durch die Umformungen der Formeln gemäß bestimmten, den Gegebenheiten vorgeschriebenen Gesetzen (Text [II], S. 23)." Die Texte [III] bis [X] in unserer Textsammlung sind Texte bzw. Entwürfe zum logischen Kalkül. In einem solchen sind (a) **Axiome** erforderlich, durch die festgelegt wird, welche Operationen in ihm durchführbar sind, und (b) **Definitionen**, die vorgeben, wie sich deskriptive Aussagen in Kalkülformeln umformen lassen und damit im Kalkül deduktiv abgeleitet werden können. Leibniz ging in seinen Kalkülen zunächst von Begriffen aus, für die er Buchstabenvariablen verwendete, um die Kalküle möglichst allgemein zu formulieren, bis die Grundbegriffe der einzelnen Wissenschaften aufgefunden sind und festgelegt werden kann, auf welche Weise die Zeichen für die einzelnen Begriffe gebildet werden können (vgl. Text [II], S. 21). Dennoch sind die Kalküle so formuliert, dass sich für die Buchstabenvariablen auch Aussagen einsetzen lassen, mit denen sich entsprechend „rechnen" lässt (vgl. Text [III, A], (13): „Unter *A* oder *B* verstehe ich hier entweder einen Begriff oder eine Aussage.").
Im Folgenden sollen (a) und (b) anhand von Beispielen veranschaulicht werden.

(a) Ein Axiomensystem

Leibniz hat verschiedene Axiomensysteme entworfen. Eines soll an dieser Stelle vorgestellt werden (es ist dem Text [V], S. 53, unserer Sammlung entnommen), wobei hier die modernen Äquivalente der leibnizschen Symbole verwendet werden. Der Strich über einem Zeichen bedeutet die Negation desselben, die reine Aneinanderreihung von zwei Zeichen deren Konjunktion.

(1) $A = B$ ist dasselbe wie, dass $A = B$ wahr ist.
(2) $A \neq B$ ist dasselbe wie, dass $A = B$ falsch ist.
(3) $A = A$.
(4) $A \neq B\bar{A}$.
(5) $A = \bar{\bar{A}}$.
(6) $AA = A$.
(7) $AB = BA$.
(8) Dasselbe sind $A = B$, $\bar{A} = \bar{B}$, $A \# B$ („#" wird hier für das *non-non* der doppelten Aussagenverneinung verwendet).
(9) Wenn $A = B$, dann folgt: $A \neq \bar{B}$.
(10) Wenn $A = AB$ ist, dann kann ein solches Y angenommen werden, sodass gilt: $A = YB$.
(11) Wenn $A = B$ ist, dann wird $AC = BC$ sein.
(12) Es koinzidieren $A = AB$ und $\bar{B} = \bar{B}\bar{A}$.

Im Anschluss an diese zwölf Axiome geht Leibniz direkt zu den Definitionen über, wie sich deskriptive Aussagen in Kalkülformeln umformen lassen.

(b) Definitionen zur Umformung deskriptiver Aussagen in Kalkülformeln

Um mit deskriptiven Aussagen „rechnen" zu können, d.h., um einen Denkfehler als „Kalkül-" bzw. „Rechenfehler" nachweisen zu können, müssen diese zunächst in Kalkülformeln umgewandelt werden. Leibniz hat viele Definitionsversuche unternommen, von denen hier zwei vorgestellt werden sollen. Beginnen wir mit der Formelgruppe, die sich auch in Text [V] findet, und stellen wir den ganz einfachen Aussagen der Beschreibungssprache, in der wir nur mit Begriffen sowie mit „alle", „einige" (d.h. mindestens ein) und „nicht" arbeiten, die algebraischen Äquivalente gegenüber. Die jeweils zweiten Kalkülformeln stellen eine weitere, von Leibniz vorgeschlagene Formelgruppe dar, die sich u.a. in Text [IV], S.45, unserer Textsammlung findet. Angefügt werden die entsprechenden Diagramme, mit denen Leibniz selbst häufig arbeitet und aus denen die Korrektheit der leibnizschen Umformungen leicht ersichtlich ist:

Die Kalkülformeln in (3) werden jeweils durch Negation der Kalkülformeln in (2) gewonnen, denn wenn ein oder mehrere A B sind, dann ist es falsch, dass kein A B ist, bzw. dann ist AB Seiendes.
Entsprechend lassen sich die Kalkülformeln in (4) durch die Negation von (1) gewinnen, denn wenn ein oder mehrere A nicht B sind, dann ist es falsch, dass alle A B sind, bzw. dann ist $A\bar{B}$ Seiendes.

Abb. 2: Die erste Figur, welche mit A bezeichnet wird, in: Raimundus Lullus: *Ars brevis*, Hamburg 1999, S. 7.

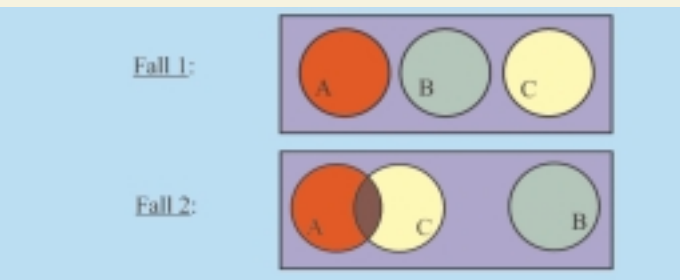
Abb. 3: Die dritte Figur, in: Raimundus Lullus: *Ars brevis*, Hamburg 1999, S. 17.

Sehen wir uns nun ein Beispiel für eine gültige Deduktion und einen Denkfehler an.

Rechnen mit Begriffen – zwei einfache Beispiele

1. „Alle Tische (A) sind braun (B), und alle Tische (A) sind aus Holz (C); folglich sind alle Tische (A) braun (B) und aus Holz (C).“ Um zu überprüfen, ob diese Folgerung gültig ist, formen wir die einzelnen deskriptiven Aussagen entsprechend der oben angegebenen ersten Formelgruppe in Kalkülformeln um. Wir erhalten A = AB und A = AC. Substituieren wir nun in A = AC für das A auf der rechten Seite AB (auf Grund von A = AB), so ergibt sich A = ABC. Setzen wir abschließend für die Buchstabenvariablen die entsprechenden Begriffe des Ausgangsbeispiels ein, so gelangen wir zu unserer Aussage „Alle Tische sind braun und aus Holz“. Die durchgeführte Deduktion zeigt, dass die angegebene Folgerung gültig ist.

2. „Kein Mensch (A) ist ein Esel (B), und kein Esel (B) ist ein Stein (C); folglich ist kein Mensch (A) ein Stein (C).“ Die einzelnen Aussagen sind für sich genommen wahr, und auf den ersten Blick könnte man möglicherweise in dieser Folgerung ein gültiges Argument sehen. Formen wir hier jedoch die einzelnen deskriptiven Aussagen – wiederum entsprechend der oben angegebenen ersten Formelgruppe – in Kalkülformeln um, so erhalten wir A = A$\bar{B}$ und B = B$\bar{C}$, aus denen sich jedoch mithilfe der Axiome des Kalküls nicht A = A$\bar{C}$ ableiten lässt, d.h., hier gelingt keine Deduktion, mit der gezeigt werden könnte, dass die angeführte Folgerung gültig ist. Tatsächlich stellt das „folglich“ einen Denkfehler dar, was sich leicht anhand einer Darstellung in Diagrammen zeigen lässt (vgl. die Diagramme unten links):



Die Diagramme zeigen, dass A = A$\bar{B}$ und B = B$\bar{C}$ nicht nur dann erfüllt sind, wenn kein A C ist (Fall 1), also wenn gilt: A = A$\bar{C}$, sondern auch dann, wenn z.B. einige A C sind, also wenn A ≠ A$\bar{C}$ gilt (Fall 2). Als Beispiel für den zweiten Fall ließe sich auch folgendes Beispiel anführen: „Kein Mensch ist ein Esel, kein Esel kann Latein, aber einige Menschen können Latein.“

Kalkül und Kabbala – Die historischen Wurzeln des Leibniz-Programms

Die historischen Aspekte des Leibniz-Programms waren zwar nicht mehr unmittelbarer Gegenstand unseres Forschungsprojekts, da sie aber für einige Leser interessant sein dürften, sei an dieser Stelle kurz auf sie eingegangen. Leibniz selbst kannte die historischen Wurzeln seines Projekts, und er hat an zahlreichen Stellen auf sie verwiesen. Als historisch frühester unter den Vorgängern ist Raimundus Lullus (1232/33-1315/16) zu nennen, ein mittelalterlicher katalanischer Philosoph, welcher in seiner „Großen Kunst“ *(Ars magna)* das Ziel verfolgte, die denkbaren Wahrheiten eines Wissensgebiets durch die Aufzählung aller möglichen Kombinationen der Grundprädikate dieses Gebiets aufzufinden. Dazu wurden die Buchstaben bzw. die den Buchstaben zugeordneten Prädikate auf konzentrischen Kreisen angeordnet (Abbildung 2) oder auch in Form einer Tabelle aufgelistet (Abbildung 3). Den Hintergrund für diese lullsche Kunst bildet die Kabbala, eine jüdisch mystische Tradition, die den gesamten Weltprozess als einen Zeichenprozess interpretiert, wobei als Zeichen Buchstaben oder Zahlen verwendet werden. Das Interesse von Leibniz an derartigen Versuchen zeigt sich auch darin, dass er selbst eine sehr viel kompliziertere Weiterentwicklung dieser *Ars* abgeschrieben hat (Akademieausgabe VI, 4, S. 1027). Bereits Lullus entwickelte mechanische Verfahren zur Herstellung solcher Listen von Buchstabenkombinationen. Dazu wurden die

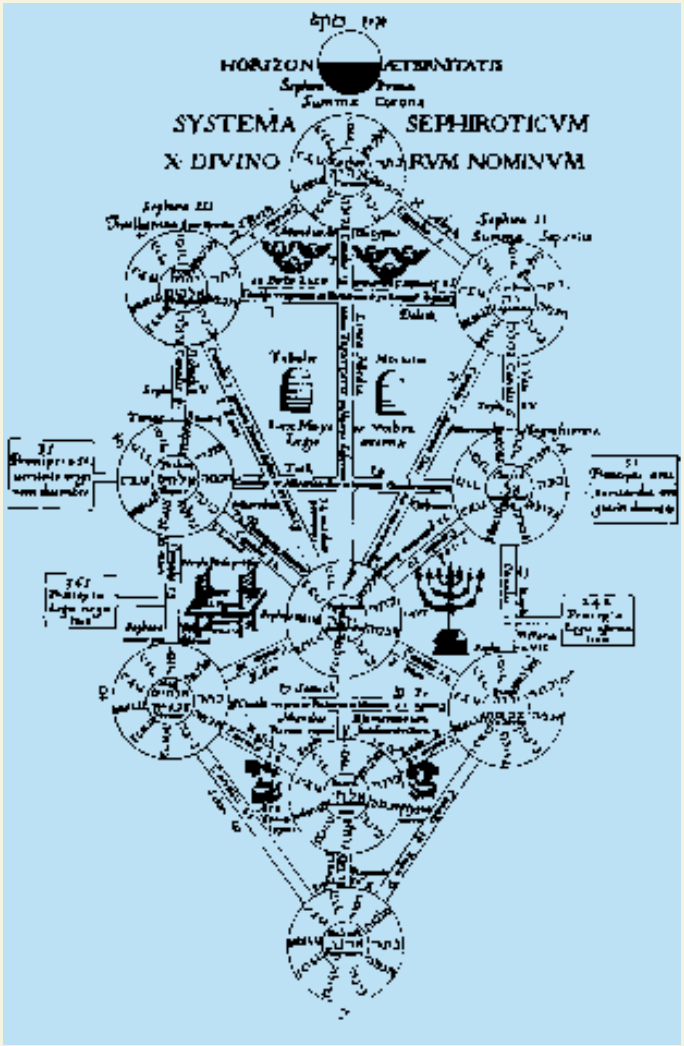


Abb. 4: Das Sefirot-System als die zentrale Glyphe der Kabbala, in: Athanasius Kircher: *Oedipus Aegyptiacus*, Rom 1652.

Buchstaben auf konzentrischen Kreisen angeordnet, von denen sich die beiden inneren auf drehbaren Scheiben bewegen ließen. Bei Leibniz führte der Versuch der Mechanisierung von Denkvorgängen dann später zur Entwicklung seiner berühmten Rechenmaschine. Verwendete Lullus die Buchstaben noch für feststehende Begriffe, so findet bei Leibniz eine Mathematisierung statt, bei der mit Buchstabenvariablen gearbeitet wird, für die sich beliebige Begriffe einsetzen lassen.

Die lullsche Kunst wurde u.a. von Giordano Bruno weitergegeben und war zu Leibniz' Zeit bekannt. **Athanasius Kircher** (1602-1680), der bedeutendste Vertreter (der in Paderborn studierte), kannte nicht nur die zeichentheoretischen Versuche der Lull-Tradition, sondern wusste auch um deren kabbalistischen Hintergrund. Bei ihm findet sich auch der *Sefirotische Baum* (Abbildung 4), eine symbolische Darstellung in Gott gegebener Grundprädikate, die gleichzeitig auch die – wiederum durch Buchstaben und Zahlen darstellbaren und kombinierbaren – Grundprädikate der Welt repräsentieren. Leibniz hat zu Schriften von Kircher umfangreiche Exzerpte angefertigt (z.B. Akademieausgabe VI, 4, S. 1201-1203). Sein Interesse galt in auffälliger Weise der Kabbala (vgl. die Indizes zu ebd., Teilband 4, S. 38), sodass der Hintergrund für die leibnizsche Algebra der Logik zu Recht in der Zahlenmystik zu suchen ist.

Literatur

G. W. Leibniz: *Sämtliche Schriften und Briefe*, Band VI, 4. Hrsg. von der Leibniz-Forschungsstelle der Universität Münster, Berlin 1999.

G. W. Leibniz: *Allgemeine Untersuchungen über die Analyse der Begriffe und Wahrheiten*. Hrsg., übers. und mit einem Kommentar versehen von Franz Schupp, 2., durchges. Aufl. Hamburg 1993.

G. W. Leibniz: *Die Grundlagen des logischen Kalküls*. Hrsg., übers. und mit einem Kommentar versehen von Franz Schupp unter der Mitarbeit von Stephanie Weber, Hamburg 2000.

Modelle für automobile Software

Objektorientierte Modellierung von eingebetteten, interaktiven Softwaresystemen im Automobil

Wie in allen technischen Geräten werden auch im Automobil immer mehr Funktionen durch Softwaresysteme realisiert bzw. gesteuert. Bei einer Entwicklung derartiger Softwaresysteme wird im Rahmen eines ingenieurmäßigen Entwicklungsprozesses zunächst ein Modell erstellt. Hierzu muss eine Modellierungssprache zur Verfügung stehen, die den Modellierer adäquat bei der Erstellung des Modells unterstützt und ein einheitliches Verständnis des Modells ermöglicht.

Bei dem Begriff „Modell“ wird bei technischen Systemen meistens an ein im Maßstab 1:20 hergestelltes Abbild des Originals gedacht. Dabei sind Modelle keineswegs nur Bastlern vorbehalten. In jeder wissenschaftlichen Disziplin werden Modelle der Realität entwickelt, um für den jeweiligen Bereich ein besseres Verständnis und einen tieferen Einblick zu entwickeln. Bei der Entwicklung von Softwaresystemen werden durch Modelle Aspekte der Struktur und des Verhaltens wiedergegeben. Dies betrifft sowohl den Ist-Zustand bei der Weiterentwicklung eines bestehenden Systems als auch die Spezifikation (Aspekte des Sollzustandes eines zu entwickelnden Systems). Modelle werden zum systematischen Entwurf, zur Überprüfung und Analyse, zur Bewertung und zur Optimierung eines Systems eingesetzt. Außerdem können sie zum einen als Vertragsgrundlage zwischen einem Auftraggeber und einem Entwickler dienen, zum anderen auch als Dokumentation eines Systems im Verlauf seiner Erstellung und spätere Wartung. In jedem Fall ist es hierbei wichtig, wenn alle Leser der Modellbeschreibung dasselbe Verständnis von dem Modell haben. Dies wird am ehesten dadurch erreicht, dass alle beteiligten Personen dieselbe Beschreibungssprache benutzen. In der Informatik wird seit einiger Zeit an solch einer einheitlichen Modellierungssprache mit dem Namen UML (Unified Modeling Language) [1, 3] gearbeitet, die sich mittlerweile auch als De-facto-Standard in Industrie, Lehre und Forschung durchgesetzt hat.

Im Folgenden sollen die Aspekte des Modellierungsprozesses und der Modellierungssprache für die Softwareentwicklung im automobilen Kontext detaillierter erläutert werden. Zunächst wird die Bedeutung eines Modells im Softwareentwicklungsprozess diskutiert. Danach wird anhand einer Fallstudie die Erstellung eines Modells mithilfe der Modellierungssprache UML erläutert, bevor auf noch bestehende Defizite der UML und Anpassungen der UML zur Modellierung von eingebetteten, interaktiven Systemen eingegangen wird. Die Fallstudie beruht auf Arbeiten im Rahmen einer Forschungskooperation mit der DaimlerChrysler AG in Stuttgart [2].

Prof. Dr. Gregor Engels (vorn) ist seit 1997 Professor für Praktische Informatik im Fachbereich 17/Mathematik – Informatik. Seine Arbeitsgebiete sind objektorientierte und visuelle Modellierungstechniken und die Entwicklung von Multimedia-Anwendungen.

Dipl.-Inform. Jens Gaulke (rechts) ist wissenschaftlicher Mitarbeiter bei Prof. Dr. Gregor Engels. Sein Forschungsgebiet ist die Entwicklung einer objektorientierten Modellierungssprache für eingebettete Systeme im Automobil auf der Basis der standardisierten objektorientierten Modellierungssprache UML (Unified Modeling Language).

Dipl.-Inform. Stefan Sauer (links) ist wissenschaftlicher Mitarbeiter bei Prof. Dr. Gregor Engels. Seine Forschungsgebiete sind die Entwicklung einer auf der UML basierenden, objektorientierten Modellierungssprache für Multimedia-Anwendungen sowie eine methodische und semantische Fundierung der UML.

Warum wird ein Modell angefertigt?

Softwaresysteme werden heutzutage in einem *verteilten* Entwicklungsprozess erstellt. Verteiltheit bedeutet hierbei unter anderem, dass die Entwicklungsarbeit auf verschiedene Personen verteilt wird, die an verschiedenen Orten aktiv sein können. Weiterhin wird in der Regel die Entwicklungsarbeit auf verschiedene Teilaktivitäten verteilt. Hierzu gehört insbesondere, dass vor der eigentlichen Programmierung ein Modell des zu realisierenden Systems erstellt wird. Ein solches Modell, in dem die Anforderungen an das System präzisiert werden, wird häufig in enger Zusammenarbeit zwischen einem Auftraggeber und dem für die Realisierung verantwortlichen Softwareentwicklungsteam angefertigt. Die Modellbeschreibung wird dazu in einer Sprache erstellt, die sowohl vom Auftraggeber als auch von Softwareentwicklern verstanden werden muss. Da man im Allgemeinen nicht davon ausgehen kann, dass der Auftraggeber (künstliche) Programmiersprachen versteht, muss die Modellierungssprache entweder eine (natürliche) Umgangssprache sein oder eine leicht verständliche Kunstsprache. Da Beschreibungen in natürlicher Sprache häufig verschiedene Interpretationen zulassen und deshalb nicht präzise genug sind, werden in der Informatik so genannte Diagrammsprachen eingesetzt. Diese Sprachen benutzen eine fest vorgegebene Anzahl grafischer Symbole mit einer präzisen Bedeutung zur Erstellung von Modellbeschreibungen. Die mithilfe dieser Sprachen erstellten Modellbeschreibungen sind in der Regel

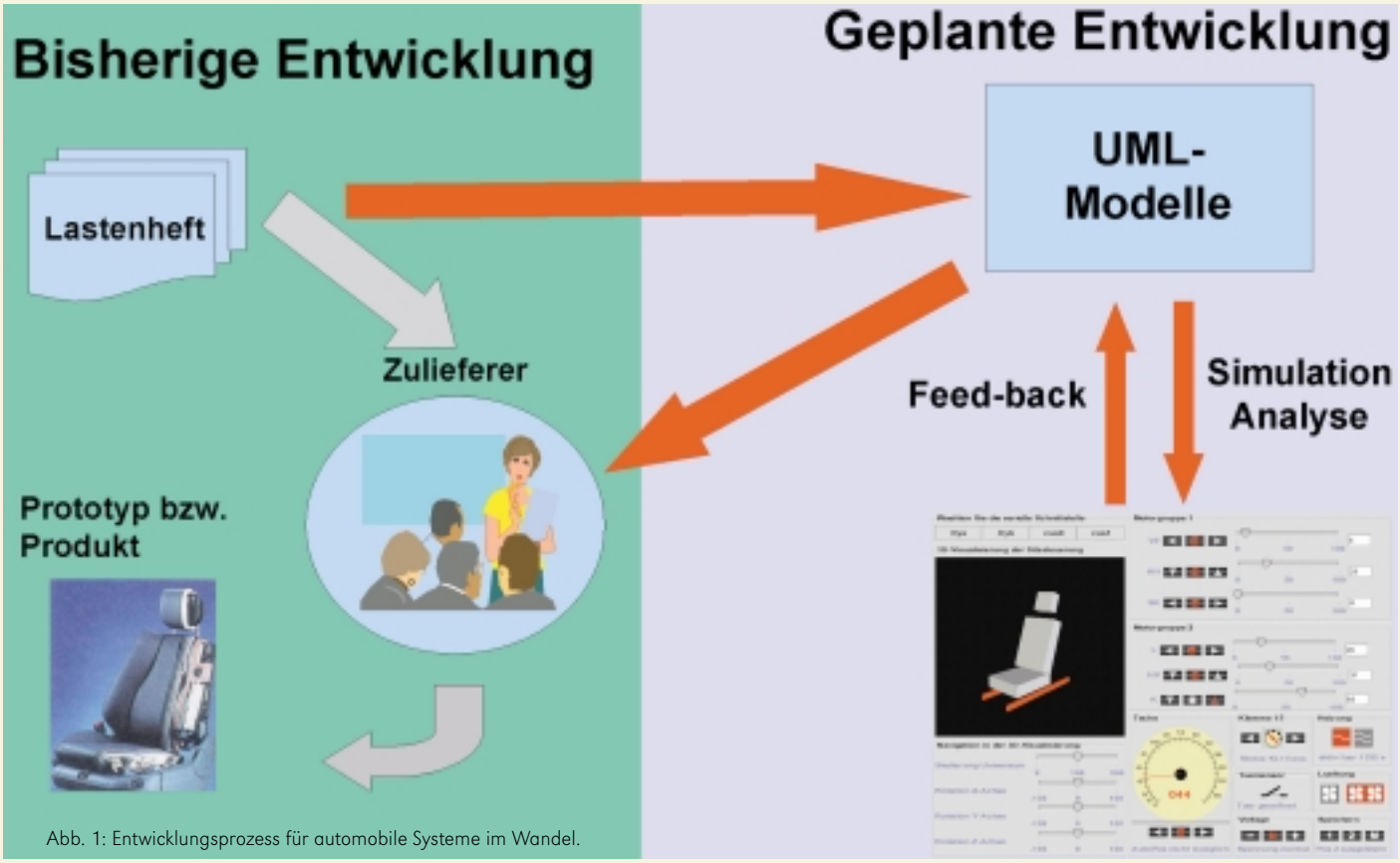


Abb. 1: Entwicklungsprozess für automobile Systeme im Wandel.

intuitiv zu verstehen und damit auch für Informatiklaien als Auftraggeber verwendbar. Das *Modell* spielt somit im Softwareentwicklungsprozess eine *zentrale Rolle*. Die in der Modellbeschreibung festgelegten Anforderungen müssen von dem zu realisierenden Softwaresystem erfüllt werden. In diesem Sinne kann das Modell als Vertragsgrundlage zwischen Auftraggeber und Softwareentwicklungsteam dienen.

Bei der Entwicklung technischer Systeme ist es heutzutage üblich, dass Teile des Systems von externen Partnern, so genannten Zulieferern, erstellt werden. Hierbei wird dem Zulieferer ein Lastenheft, in der Regel eine textuelle Form einer Modellbeschreibung, übergeben. Der Zulieferer entwickelt dann einen entsprechenden Prototypen bzw. letztendlich ein Produkt (vgl. hierzu Abbildung 1). Auf Grund der angesprochenen möglichen Fehlinterpretationen können dann allerdings mehrfache (vermeidbare und zeitraubende) Rücksprachen mit dem Auftraggeber notwendig werden oder sogar fehlerhafte Produkte erstellt werden. Eine verbesserte Systementwicklung soll deshalb dadurch erreicht werden, dass an Stelle einer umgangssprachlichen, textuellen Beschreibung eine grafische Beschreibung mithilfe von Diagrammen erstellt wird. Hierbei soll dann eine Sprache eingesetzt werden, die von allen Beteiligten gleich verstanden wird und somit wenig Interpretationsspielraum zulässt.

Unified Modeling Language (UML) stellt in der Softwareentwicklung eine solche Sprache dar. Die UML wurde in den letzten fünf Jahren mit großer weltweiter Unterstützung der Softwareindustrie entwickelt. Unter der Federführung des von der Industrie getragenen Standardisierungsgremiums OMG (Object Management Group) wurden schließlich verschiedene Vorschläge für eine derartige Sprache vereinheitlicht [1, 3].

Die UML ist eine *objektorientierte* Modellierungssprache. Damit ist gemeint, dass sowohl strukturelle Eigenschaften („WER macht etwas?“) als auch dynamische Eigenschaften eines Softwaresystems („WANN soll WAS geschehen?“) mithilfe der UML innerhalb eines integrierten Modells beschrieben werden können. Hierzu bietet die UML verschiedene Teilsprachen an, die an Beispielen im Folgenden erläutert werden.

Die Benutzung einer präzise definierten Modellierungssprache hat den weiteren Vorteil, dass die erstellten Modelle analysiert und dadurch Fehler und Missverständnisse frühzeitig erkannt werden können. Dies kann insbesondere dadurch erreicht werden, dass aus der Modellbeschreibung automatisch ein Prototyp generiert wird, der es dann dem Auftraggeber und Entwickler bzw. dem Zulieferer ermöglicht, mit dem Modell zu „spielen“. Entdeckte Fehler während einer solchen *Simulation* des Systems können als Feed-back in den Modellierungsvorgang einfließen.

Einige Aspekte der Modellierungssprache UML werden nun anhand einer Fallstudie erläutert.

Fallstudie: Sitzsteuerung

Das Auto von heute besitzt eine große Anzahl an elektronischen Komponenten – sowohl Hard- als auch Software – deren Zusammenspiel von Mikrocontrollern und Computern überwacht wird. Auf Grund dieser engen Koppelung spricht man von einem eingebetteten System. Neben Messwerten, die über Sensoren aufgenommen werden, wird auch jede Eingabe, die der Fahrer des Autos vornimmt, von einem Computer ausgewertet und an die zuständige Komponente weitergeleitet. Ein Beispiel hierfür sind die Sitze der aktuellen Baureihe der S-Klasse von Mercedes-Benz. Jeder der Sitze kann z.B. nur noch durch Elektronik positioniert werden bzw. Komfortfunktionen wie Sitzheizung und Sitzlüftung können nur elektronisch gesteuert werden,

Sitzheizung:

Die Funktion Sitzheizung ist bei Klemme 15R ein und Klemme 15 ein aktivierbar [...] Die Sitzheizung stellt 2 Heizstufen zur Verfügung, die mittels zweier Taster bedient werden können, ein Taster zum Ein- und Ausschalten von Heizstufe 1 und der andere Taster zum Ein- und Ausschalten von Heizstufe 2. Im Grundzustand ist die Sitzheizung ausgeschaltet. Durch Betätigung der Taste „Heizstufe 1"/„Heizstufe 2" wird die Heizstufe 1/Heizstufe 2 eingeschaltet. Wenn Heizstufe 1/Heizstufe 2 aktiv ist, wird durch Betätigung der Taste „Heizstufe 1"/„Heizstufe 2" die Sitzheizung ausgeschaltet. Wenn Heizstufe 1 aktiv ist, wird durch Betätigen der Taste „Heizstufe 2" in Heizstufe 2 umgeschaltet; wenn Heizstufe 2 aktiv ist, wird durch Betätigen der Taste „Heizstufe 1" in Heizstufe 1 umgeschaltet [...] Beim Eintreten eines der nachfolgenden Kriterien wird die Sitzheizung abgeschaltet. [...] Abschalten nach 15R aus: Nach Ausschalten von 15R schaltet auch die Sitzheizung ab. [...]

Abb. 2: Auszug aus dem Lastenheft für die Sitzheizung.

eine manuelle Bedienung ist ausgeschlossen. In einem Teilprojekt der Kooperation mit DaimlerChrysler wurde ein Fahrersitz und insbesondere die zugehörige Steuerungssoftware in UML modelliert. Als Basis für das Modell dienen Auszüge aus dem Lastenheft (textuelle Anforderungsbeschreibung). Der in Abbildung 2 gezeigte Text beschreibt z.B. Anforderungen an die Sitzheizungssteuerung als Teil der Sitzsteuerung, deren Modellierung im Folgenden erläutert wird. Die Angaben zu Klemme 15 bezeichnen dabei die verschiedenen Positionen des Zündschlossschalters.

Wie wird ein Modell in UML angefertigt?

Das Erstellen eines Modells geschieht in mehreren Schritten. Zunächst wird grob an die Aufgabenstellung herangegangen: Welche Strukturen können identifiziert werden? Welche (Teil-) Objekte sind enthalten? In welcher Beziehung stehen die Objekte zueinander? Zu diesem Zweck stellt die UML das Klassendiagramm zur Verfügung. Das *Klassendiagramm* selbst gibt allerdings noch keine Auskunft darüber, welches dynamische Verhalten das System haben soll.

Der nächste Schritt bei der Modellierung besteht daher in der Identifizierung und Strukturierung der ausführbaren Systemfunktionen. Diese Systemfunktionen, auch Anwendungsfälle (engl.: use cases) genannt, werden mithilfe eines UML *Use-Case-Diagramms* beschrieben. Es wird dargestellt, welche Funktionen es gibt und wie diese untereinander wechselwirken. Dabei wird nur die Funktion – also das, was sie leisten soll – beschrieben, aber noch nicht, wie sie dies leisten soll. Letzteres wird im dritten Schritt der Modellierung durch UML *Statechart-Diagramme* dargestellt. In Statecharts wird die Abfolge von Zuständen (States), die ein Objekt im Laufe seiner Existenz einnehmen kann, dargestellt. Statecharts sind bereits sehr implementierungsnah und können direkt als Grundlage zur Realisierung eines Simulationsprototypen oder des eigentlichen Softwaresystems dienen.

Die UML bietet darüber hinaus noch eine Reihe weiterer Diagramme an, um insbesondere das dynamische Verhalten eines Systems zu beschreiben. Hierzu zählen Sequenzdiagramme, Kollaborationsdiagramme oder Aktivitätendiagramme. Jede dieser Diagrammarten betont einen anderen Aspekt des dynamischen Verhaltens wie z.B. den zeitliche Ablauf oder das Kommunikationsverhalten zwischen Objekten eines Systems. Diese Diagrammarten werden deshalb alternativ oder ergänzend zu Statechart-Diagrammen bei der Modellierung eingesetzt. Der interessierte Leser sei an dieser Stelle auf die entsprechende Literatur verwiesen (z.B. [1, 3]).

Entstehung des Klassendiagramms

Für die Modellierung der Sitzsteuerung ist es wichtig, zunächst die einzelnen Teile der Steuerung zu identifizieren. Dies geschieht durch ein Klassendiagramm (vgl. Abbildung 3). Der Sitz besteht aus den Einzelteilen Kopfstütze, Lehne und Sitzkissen. Diese kann der Fahrer über Motoren positionieren, die in diesen Sitzteilen eingebaut sind. Der gesamte Sitz kann in der Höhe und der Länge verstellt werden, wozu drei weitere Motoren in den Sitz eingebaut sind. Desweiteren gibt es eine Sitzheizung und die Möglichkeit, den Sitz zu belüften. Die identifizierten Sitzteile wurden im Diagramm als Klassen dargestellt. Klassen beschreiben die grundlegende Struktur eines Objekts. Daher kann von den verschiedenen Motoren im Sitz auf „den Motor" verallgemeinert (abstrahiert) werden. Jeder Motor hat die gleiche Funktionalität: Er kann links herum oder rechts herum laufen. Farbe, Form, Größe und weitere Eigenschaften sind für die Modellierung vorläufig irrelevant. Auf Grund dieser Abstraktion von einzelnen Objekten zur Klasse ist es in der Modellbeschreibung ausreichend, die Klasse Motor nur einmal zu zeichnen. Eine Klasse wird hierbei grafisch als Rechteck dargestellt.
Objekte können zu anderen Objekten in einer strukturellen Beziehung stehen. Eine Beziehung wird im Klassendiagramm durch eine Verbindungskante zwischen den beteiligten Klassen dargestellt. Die Beziehung kann lose und temporär sein, aber auch fest und permanent. Ein Beispiel für letztere ist die „besteht aus"-Beziehung zwischen zwei Objekten. So besteht ein Sitz z.B. aus einer Kopfstütze, einer Lehne, einem Sitzkissen und drei Motoren zur Positionierung des Sitzes. Derartige Beziehun-

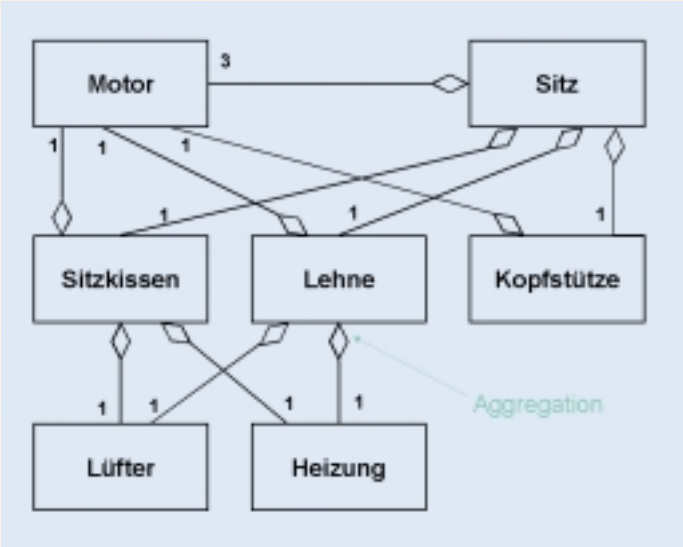


Abb. 3: Klassendiagrammausschnitt für die Sitzsteuerung.

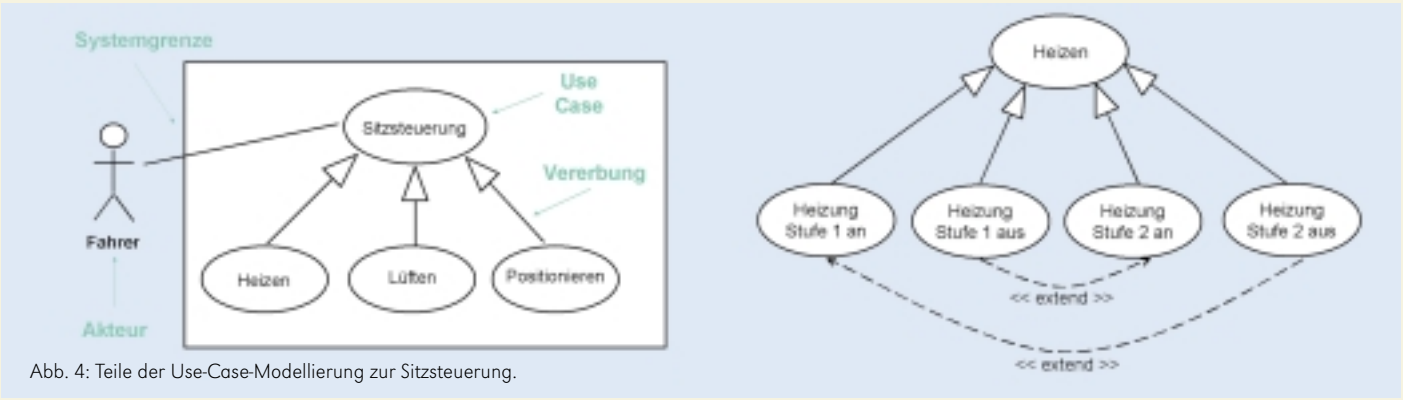


Abb. 4: Teile der Use-Case-Modellierung zur Sitzsteuerung.

gen zwischen Objekten, auch Aggregation genannt, werden im Klassendiagramm durch Linien mit einem Diamanten an der Seite der aggregierenden Klasse dargestellt. Kardinalitätsangaben, wie z.B. drei Motoren in einem Sitz, werden entsprechend im Diagramm ergänzt.

Entstehung des Use-Case-Diagramms

Im Use-Case-Diagramm werden die durch den Benutzer aufrufbaren Funktionen der Sitzsteuerung identifiziert. Dabei ist zu diesem Zeitpunkt nur wichtig, welche Funktionen es gibt. Wie die Funktionen das ihnen zugedachte Verhalten realisieren, wird getrennt modelliert. Grafisch wird eine Funktion, Use-Case genannt, durch eine Ellipse dargestellt. Zu jeder Ellipse existiert ein Text, der den Use-Case genauer beschreibt. Bei diesem Text kann es sich um stichwortartige Beschreibungen oder um ausführliche Darstellungen handeln. Wie beim Klassendiagramm gibt es Beziehungen unter den einzelnen Use-Cases.

Sie können beispielsweise durch Generalisierungsbeziehungen hierarchisch verfeinert werden (siehe Abbildung 4). So sind die Funktionen *Heizung Stufe 1 an, Heizung Stufe 1 aus, Heizung Stufe 2 an, Heizung Stufe 2 aus* Verfeinerungen des Use-Case *Heizen*, was durch eine nicht gefüllte Pfeilspitze an dem allgemeineren Use-Case dargestellt wird. Die Spezialisierungen besitzen eine gegenüber dem Use-Case *Heizen* verfeinerte bzw. erweiterte Funktionalität. Analoges gilt für Lüften und Positionieren. Heizen, Lüften und Positionieren werden ihrerseits zum übergeordneten Use-Case *Sitzsteuerung* generalisiert. Durch die Beziehung zwischen dem Benutzer, der hier die Rolle „Fahrer“ ausübt, dargestellt durch das Strichmännchen (Akteur), und dem Use-Case *Sitzsteuerung* wird modelliert, dass ein Fahrer die Funktion Sitzsteuerung sowie jede ihrer untergeordneten Funktionen aufrufen kann. Entsprechend müssen bei dem Entwurf der Benutzungsschnittstelle die Möglichkeiten zu deren Aufruf z.B. durch Schalter, Schaltflächen, Menüs etc. berücksichtigt werden. Zudem kann modelliert werden, dass Funktionen (optionale) Teile anderer Funktionen sein können. Die <<extend>>-Beziehung zwischen Use-Case *Heizung Stufe 1 aus* und *Heizung Stufe 2 an* besagt beispielsweise, dass der zweite Use-Case unter bestimmten, noch festzulegenden Bedingungen um das Verhalten des ersten erweitert wird.

Was ist ein Statechart?

Im Klassendiagramm wird die Struktur des zu entwickelnden Systems festgelegt; das Use-Case-Diagramm dient dazu, die Funktionalität des Gesamtsystems festzulegen, das der Außenwelt an der Schnittstelle angeboten werden muss. Um die Komponenten und Funktionen mit Leben zu erfüllen, reichen die bisher verwendeten Diagramme allerdings noch nicht aus. Statecharts können diese Aufgabe übernehmen, da sie das in den bisher eingesetzten Diagrammen beschriebene Modell verfeinern und Verhalten hinzufügen. Ein Statechart des Modells ist immer einer Klasse zugeordnet. In Abbildung 5 wird das zur Klasse Heizung gehörende Statechart beschrieben. Ein Statechart besteht aus Zuständen, von denen einer als Startzustand und mehrere als Endzustände ausgezeichnet werden können, und Transitionen zwischen diesen Zuständen, die durch bestimmte Ereignisse (Signale) ausgelöst werden. Zustände können im Sinne einer weiteren Verfeinerung hierarchisch komponiert werden, dann können Start- und Endzustände auf jeder Ebene verwendet werden. Abhängig vom jeweiligen Zustand sind nur bestimmte Zustandsübergänge möglich. So kann vom Zustand Heizung aus durch das Signal Taste 2 über den Zustandsübergang (dargestellt durch einen Pfeil) nur in den Zustand Heizung Stufe 2 gewechselt werden. Dies entspricht dem Einschalten der Heizung auf Stufe 2 durch den Fahrer.

Vom Startzustand (Init) aus wird die Heizung aktiviert (Sitzheizung), wenn das Signal KL15R oder KL15 empfangen wird (bezieht sich auf Klemme 15, siehe Abbildung 2). Die Heizung geht dann in den Zustand Heizung aus über, der Startzustand innerhalb von Sitzheizung ist. Dieser Zustand kann verlassen werden, wenn die Signale Taste 1 oder Taste 2 auftreten. Es wird dann in den Zustand Heizung Stufe 1 bzw. Heizung Stufe 2

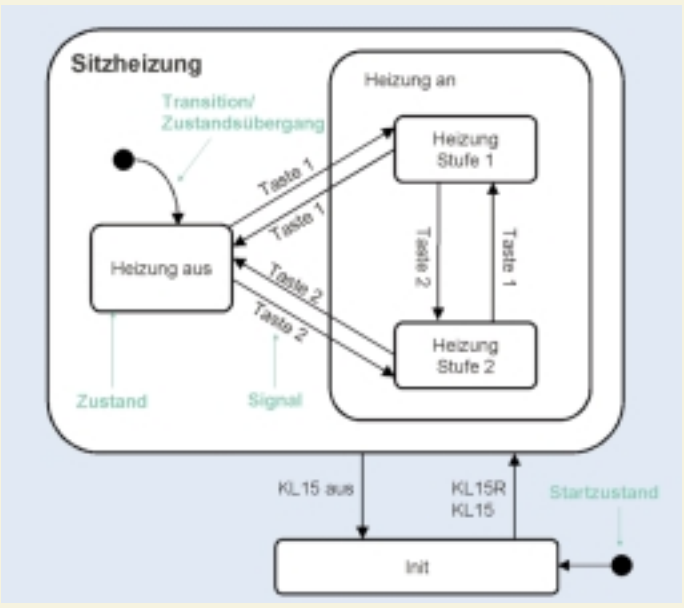


Abb. 5: Statechart-Diagramm für das Verhalten der Heizung.

Abb. 6: Beispiel einer Head-Unit in einem zukünftigen Automobil.

gewechselt. Tritt das Signal KL15 aus auf – egal, in welchem Unterzustand des komponierten Zustands Sitzheizung sich die Heizung gerade befindet – wird wieder in den Initialzustand gewechselt und die Sitzheizung deaktiviert.

Das im Statechart modellierte Verhalten wird nun den im Use-Case-Diagramm spezifizierten Funktionen zugeordnet. Die Transitionen zur Änderung der Heizstufe können auf entsprechende Use-Cases abgebildet werden. So entsprechen die Transitionen von Heizung aus nach Heizung Stufe 2 und von Heizung Stufe 1 nach Heizung Stufe 2 der Ausführung des Use-Cases *Heizung Stufe 2 an* (vgl. Abbildung 4), wobei im zweiten Fall das Verhalten um das Verhalten des erweiternden Use-Cases *Heizung Stufe 1 aus* erweitert wird.

Von UML zu Automotive UML

Diese kleinen Ausschnitte aus einem Modell für ein eingebettetes System in einem Automobil zeigen, dass die UML geeignet ist, um die Kernfunktionalität eines Softwaresystems zu modellieren. Da die Sprache UML jedoch als allgemein einsetzbare Modellierungssprache konzipiert ist, werden ihre Grenzen deutlich, wenn man sie zur Modellierung in einem speziellen Anwendungskontext einsetzt. Dies gilt auch für die Modellierung von Softwaresystemen im Automobil wie etwa die Modellierung der Anzeige- und Steuerungssoftware in einer so genannten Head-Unit eines Autos der Zukunft (vgl. Abbildung 6). Eine derartige Software ist nicht nur eingebettet, da sie mit verschiedenen technischen, insbesondere elektronischen Komponenten interagiert, sondern muss auch Echtzeitanforderungen (z.B. unmittelbare Anzeige der aktuellen Geschwindigkeit) genügen und eine interaktive Schnittstelle zum Fahrer realisieren.

Die Entwickler der UML haben diese Notwendigkeit der Anpassbarkeit an einen bestimmten Anwendungskontext bereits vorausgesehen und Mechanismen für eine Anpassung bereitgestellt. In einem gemeinsamen Forschungsprojekt mit der DaimlerChrysler AG in Stuttgart wird derzeit an einer solchen Entwicklung einer *Automotive UML*, einer Erweiterung der UML für den automobilspezifischen Kontext, gearbeitet. Hierbei steht insbesondere die Modellierung der interaktiven grafischen Bedienoberfläche im Vordergrund. Aufbauend auf früheren Arbeiten zur Entwicklung einer UML-basierten Modellierungssprache für Multimedia-Anwendungen [4, 5] wird UML um eine zusätzliche Diagrammart zur Modellierung der Bedienoberfläche erweitert. Durch dieses *Präsentationsdiagramm* (siehe Abbildung 7) können dann nicht nur die Struktur und das Verhalten einer automobilen Software modelliert werden, sondern auch deren Anzeige und interaktive Bedienung.

Die Kooperation mit der DaimlerCrysler AG und andere

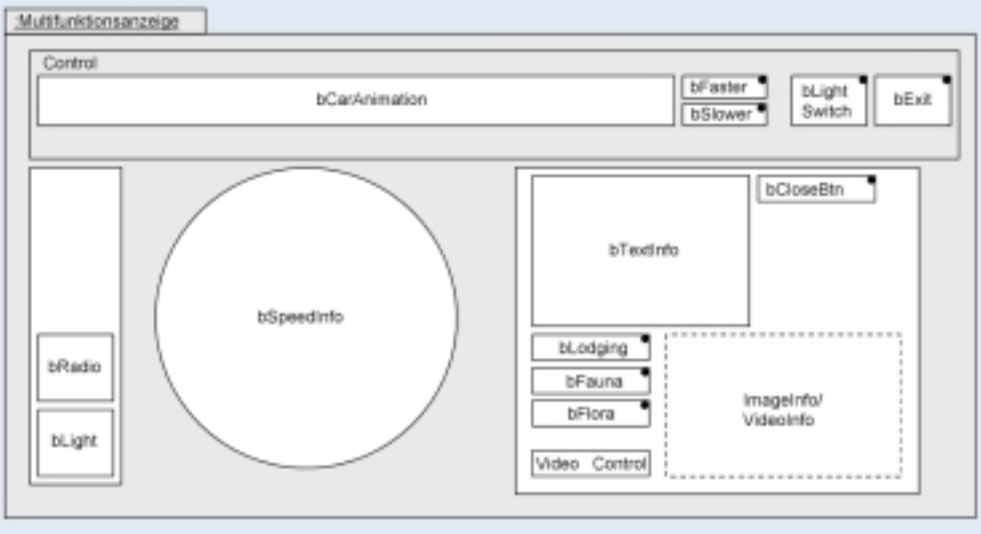


Abb. 7: Präsentationsdiagramm zu der Head-Unit.

Forschungstätigkeiten im Bereich interaktiver multimedialer und eingebetteter Systeme in der Arbeitsgruppe haben belegt, dass für diese Systeme wie für herkömmliche Software das Erstellen eines Modells vor der eigentlichen Realisierung bzw. Implementierung unabdingbarer und zentraler Bestandteil des Entwicklungsprozesses ist. Mit der Unified Modeling Language steht zudem ein Standard zur Verfügung, der es einerseits erlaubt, die Kernaspekte von Systemen aus verschiedenen Anwendungskontexten mit einheitlichen Sprachmitteln zu modellieren, und der andererseits einen eingebauten Erweiterungsmechanismus bereitstellt, der spezifische Erweiterungen der Modellierungssprache für spezielle Anwendungsdomänen gestattet. So kann ein hohes Maß an Flexibilität und Problemangemessenheit bei der Sprachentwicklung und der Modellierung auf Basis einer allgemeinen und weit verbreiteten Modellierungssprache erreicht werden.

Literatur

[1] G. Booch, J. Rumbaugh, I. Jacobsen: The Unified Modeling Language User Guide. Addison-Wesley, Reading MA 1999.

[2] P. Geretschläger, P. Hofmann: Objektorientierte Entwicklung eingebetteter Echtzeitsysteme im Automobil. In P. Hofmann, A. Schürr (Hrsg.): Proc. 1st Workhop on Object-Oriented Modeling of Embedded Realtime Systems (OMER), Mai 1999, Herrsching (Ammersee). Universität der Bw München, Fak. für Informatik, Bericht Nr. 1999-01, Mai 1999, S. 61-66.

[3] Object Management Group (OMG): Unified Modeling Language, Version 1.3, http://www.omg.org.

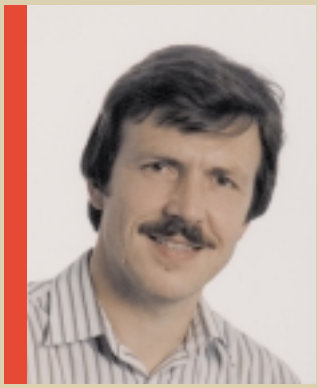
[4] St. Sauer, G. Engels: Extending UML for modeling of multimedia applications. In: Proc. 1999 IEEE Symposium on Visual Languages (VL'99), September 1999, Tokio. IEEE Computer Society, S. 80-87.

[5] St. Sauer, G. Engels: UML-basierte Modellierung von Multimediaanwendungen. In J. Desel, K. Pohl, A. Schürr (Hrsg.): Tagungsband Modellierung'99, März 1999, Karlsruhe. Teubner, Stuttgart 1999, S. 155-170.

Huminstoffe und Trinkwasser

Ein kleiner Ausschnitt aus dem globalen Kohlenstoffkreislauf

Das Element Kohlenstoff kommt auf der Erde nicht fest verteilt vor, sondern es wird ständig zwischen Atmosphäre, Biosphäre, Geosphäre und Hydrosphäre ausgetauscht und befindet sich somit in einem stetigen Kreislauf. In der Hydrosphäre wird Kohlenstoff in Form gelöster anorganischer Verbindungen (Kohlenstoffdioxid bzw. Bikarbonat) sowie in Form partikulärer oder gelöster organischer Komponenten vom Land in die Ozeane transportiert. Die gelösten natürlichen organischen Substanzen sind die so genannten Huminstoffe als Abbauprodukte abgestorbener Biomasse. Obwohl selbst für die menschliche Gesundheit unbedenklich, können sie die Qualität von Trinkwasser über Sekundärwirkungen negativ beeinflussen. Deshalb ist es notwendig, sich mit ihrer Rolle in verschiedenen Wasseraufbereitungsprozessen zu befassen und Verfahren einzusetzen, mit denen sie wirksam entfernt werden können.



Prof. Dr.-Ing. Joachim Fettig ist seit 1990 Professor im Fachbereich 8/Technischer Umweltschutz, Fachgebiet Wassertechnologie, an der Fachhochschulabteilung Höxter der Universität Paderborn. Seine Arbeitsschwerpunkte liegen im Bereich der Aufbereitung von Trink- und Betriebswasser sowie der physikalisch-chemischen Reinigung von industriellem Abwasser.

Der globale Kohlenstoffkreislauf

Vor 24 Jahren, im März 1977, trafen sich in Ratzeburg 66 Wissenschaftler aus 22 Ländern, um während eines einwöchigen Workshops die vorhandenen Daten zur Verteilung des Elements Kohlenstoff auf der Erde und den ständigen Mengenaustausch zwischen den Kompartimenten Geosphäre, Hydrosphäre, Atmosphäre und Biosphäre zu diskutieren und daraus quantitative Erkenntnisse abzuleiten. Obwohl es schon zuvor Ansätze gegeben hatte, Vorkommen und Transport des Kohlenstoffs zu beschreiben, markiert dieser Workshop gewissermaßen den

Abb. 1: Das Hochmoor Mecklenbruch im Solling.

Fotos: privat

Beginn der Modellierung des globalen biogeochemischen Kohlenstoffkreislaufs [1].

Kohlenstoff spielt für die Existenz von pflanzlichem und tierischem Leben eine Schlüsselrolle. Die ständig in der Biosphäre gebundene Masse wird mit mindestens $550x10^{12}$ kg angenommen, wobei über 98 Prozent der Pflanzenwelt zuzurechnen sind. Die Pflanzen nehmen den Kohlenstoff aus der Luft in Form von Kohlenstoffdioxid (CO_2) auf und synthetisieren daraus organische Verbindungen. Nach dem Absterben von Pflanzen oder Pflanzenteilen (Laubfall) wird ein bestimmter Anteil der toten Biomasse durch Mikroorganismen mineralisiert. Dabei entsteht wieder CO_2, das zum Teil direkt zurück in die Atmosphäre gelangt oder, nach Aufnahme durch das Sickerwasser und Umsatz zu Bikarbonat im Grundwasser, als anorganische Kohlenstoffverbindung in die Ozeane transportiert wird.

Der direkte Rücktransport in die Atmosphäre geschieht im Übrigen auch nahezu vollständig bei der Verbrennung von Biomasse (Holz). Solange nur nachwachsende Rohstoffe als Energieträger verbrannt werden, kann daher der CO_2-Gehalt der Atmosphäre, der einer Gesamtkohlenstoffmasse von rund $700x10^{12}$ kg entspricht, auf einem konstanten Niveau verharren. Wird dagegen, wie es seit etwa 150 Jahren in nennenswertem Umfang geschieht, fossiler organischer Kohlenstoff als Energieträger verbrannt, in dem eine Gesamtkohlenstoffmasse von mehr als $5000x10^{12}$ kg gebunden ist, so kann man leicht nachvollziehen, dass damit die Gleichgewichtsverteilung des Kohlenstoffs empfindlich beeinflusst wird, sofern nicht diese zusätzliche CO_2-Menge von jährlich etwa $6x10^{12}$ kg an anderer Stelle gebunden werden kann. Zwar gibt es eine wichtige Senke für atmosphärisches CO_2 in den Ozeanen, doch beobachten wir ja bereits seit mehreren Jahrzehnten einen Anstieg des CO_2-Gehaltes in der Atmosphäre und müssen uns über die daraus zu erwartende Veränderung des Wärmehaushaltes der Erde (Treibhauseffekt) Sorgen machen.

Der mikrobielle Abbau von Biomasse führt jedoch in der Regel nicht zu einer vollständigen Mineralisierung, sondern es entstehen auch organische Restprodukte, die so genannten Huminstoffe. Sie reichern sich in der oberen Bodenschicht an und werden entweder in langsamen Folgeprozessen mineralisiert, oder sie bilden im Boden ein Kohlenstofflager, z.B. in Form von Torf, oder sie werden vom Sickerwasser aus dem Boden herausgelöst. Die ständig in der oberen Bodenschicht vorhandene Masse an organischem Kohlenstoff wird auf etwa $1500x10^{12}$ kg geschätzt, d.h. sie beträgt ein Mehrfaches der in lebenden Pflanzen gebundenen Masse.

Aus dem Humusvorrat des Bodens werden pro Jahr etwa $19x10^{10}$ kg organischer Kohlenstoff in gelöster oder feinpartikulärer Form mit dem Grund- und Oberflächenwasser in die Ozeane transportiert. Der mittlere Gehalt der Flüsse an gelösten organischen Verbindungen beträgt 3,3 mg/l DOC (Dissolved Organic Carbon), hinzu kommen noch etwa 1,8 mg/l organischer Kohlenstoff in partikulärer Form. Das Vorkommen von Huminstoffen auf der Erde ist jedoch nicht abhängig von der jeweils anzutreffenden Biomasse, sondern wird in erster Linie durch das Klima bestimmt. Bei Temperaturen von über 25°C und ausreichender Sauerstoffversorgung findet im Boden keine nennenswerte Akkumulation organischen Kohlenstoffs statt, während bei niedrigeren Temperaturen der Anfall der organischen Reststoffe größer ist als der mikrobielle Abbau, und sich somit auch

Abb. 2: Sickerwasser aus dem Hochmoor Mecklenbruch im Solling.

vermehrt Huminstoffe im Grund- und Oberflächenwasser finden. Dies trifft vor allem auf die borealen Wälder der Nordhalbkugel und in sehr starkem Maß auf die arktische Tundra zu. In unseren Breiten findet man hohe Huminstoffgehalte in erster Linie im Bereich der Hochmoore in den Mittelgebirgen, wie in Abbildung 1 und 2 veranschaulicht wird.

Wer auf Reisen schon einmal das schottische Hochland, Nordskandinavien, Sibirien oder den Norden Kanadas kennen gelernt hat, wird sich vielleicht erinnern, dass es dort nicht nur gelblichbraun gefärbtes Oberflächenwasser, sondern auch Trinkwasser mit einer deutlichen Färbung geben kann, die durch Huminstoffe verursacht wird. Auch bei uns enthält das Trinkwasser Spuren an Huminstoffen, deren Konzentration je nach Herkunft zwischen 0,5 und 2 mg/l DOC schwankt, obwohl es in der Regel keine für das Auge erkennbare Färbung aufweist. Es stellt sich daher generell die Frage, welche Relevanz die Huminstoffe für die Trinkwasserqualität haben und inwieweit sie gegebenenfalls durch eine Aufbereitung zu entfernen sind.

Beeinträchtigen Huminstoffe die Trinkwasserqualität ?

Aquatische Huminstoffe sind chemisch gesehen organische Moleküle uneinheitlicher Struktur mit Molmassen von einigen 100 g/mol bis zu vielen 10 000 g/mol. Sie bestehen aus den Hauptkomponenten Kohlenstoff (ca. 50 Massenprozent), Sauerstoff (ca. 40 Massenprozent) und Wasserstoff (ca. 5 Massenprozent) sowie Stickstoff und Schwefel. Auf Grund ihrer Heteroge-

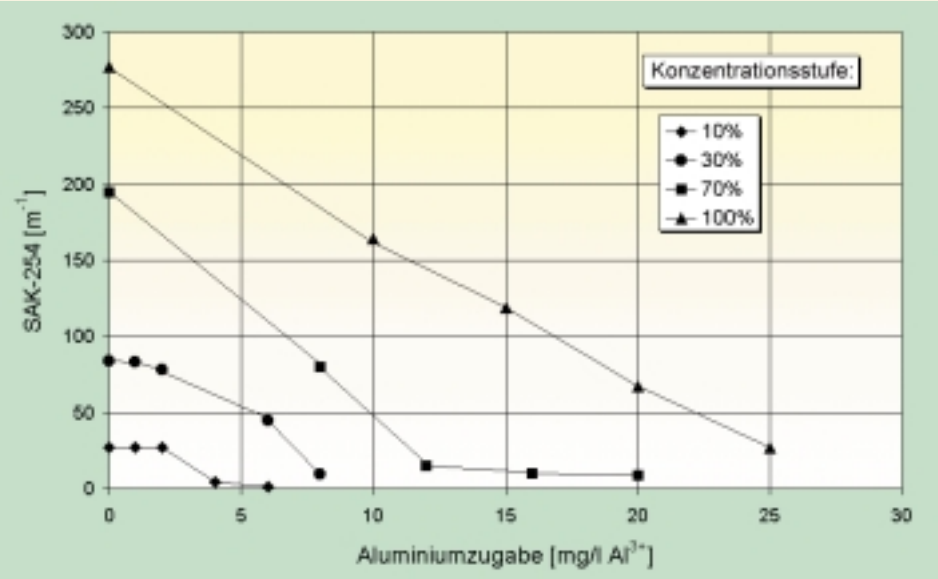


Abb. 3: Entfernung von Huminstoffen durch Flockung mit Aluminiumsulfat.

nität und grundsätzlicher Probleme bei den für Analysen erforderlichen Isolierungsverfahren ist es bis heute nicht gelungen, ihre Strukturen vollständig aufzuklären. Als Endprodukte (schneller) mikrobiologischer Abbauprozesse sind sie im wässrigen Milieu relativ stabil, d.h. sie werden lediglich in langen Zeiträumen von Monaten bis Jahren weiter mineralisiert. In Oberflächengewässern verändern sich ihre Strukturen unter dem Einfluss des Sonnenlichtes. Sie haben das Vermögen, sich an Partikeloberflächen adsorptiv anzulagern und andererseits auch eine Vielzahl gelöster Wasserinhaltsstoffe an sich zu binden.

Huminstoffe an sich sind gesundheitlich nicht bedenklich. Sie können jedoch eine Reihe von Sekundärwirkungen haben, die es erfordern, dass man sich mit ihnen beschäftigt. Nach einem groben Raster lassen sich die durch Huminstoffe bedingten Probleme in solche nutzungsbezogener, technischer oder gesundheitsbezogener Art unterteilen.

Nutzungsbezogene Probleme: Die Huminstoffmoleküle weisen insbesondere bei höheren Molmassen chromophore Strukturen auf, d.h. sie verursachen eine gelb-braune Färbung des Wassers, deren Intensität vom Huminstoffgehalt abhängt. Dadurch wird der ästhethische Eindruck beim Nutzer beeinträchtigt – nicht zufällig fordert die DIN 2000 für Trinkwasser in Deutschland, dass die Attribute farblos, klar und zum Genuss anregend erfüllt sein sollen. Trinkwasser mit einer Eigenfarbe könnte zudem zu der Assoziation führen, dass es auch noch andere „Verunreinigungen" enthält. Im Leitungsnetz können Huminstoffe ausgefällt werden oder sich an den Wandungen anlagern, um dann in unregelmäßigen Abständen als eine Art Schlamm beim Verbraucher aus dem Hahn zu kommen. Die Eigenfarbe kann sich außerdem qualitätsmindernd bei der Wäsche weißer Textilien auswirken.

Technische Probleme: Huminstoffe verhalten sich in zahlreichen Aufbereitungsprozessen nicht inert, sondern beeinflussen diese und werden dabei selbst verändert. Ist also eine Vorbehandlung des Trinkwassers im Wasserwerk erforderlich, z.B. um partikuläre Stoffe zu entfernen oder um es zu desinfizieren, so können die optimalen Verfahrensparameter in starkem Maße von den Huminstoffen abhängen. Zudem sind dann in der Regel deutlich größere Mengen an Behandlungschemikalien erforderlich, um das Aufbereitungsziel zu erreichen. Hohe Huminstoffgehalte können sich bei bestimmten Wasserqualitäten auch ungünstig auf die Korrosion in metallischen Leitungen auswirken.

Gesundheitsbezogene Probleme: Im Jahr 1974 entdeckte der holländische Wasserchemiker Johannes Rook, dass bei der Desinfektion huminstoffhaltigen Wassers mit Chlor nicht nur Mikroorganismen inaktiviert, sondern als so genannte Desinfektionsnebenprodukte auch Chloroform und andere chlororganische Verbindungen im Konzentrationsbereich von einigen 10 µg/l bis über 100 µg/l gebildet werden. Wegen ihrer toxikologischen Relevanz wurde in der Folge für die chemische Gruppe der Trihalogenmethane im aufbereiteten Trinkwasser ein Grenzwert festgelegt, der aber nur eingehalten werden kann, wenn die Huminstoffkonzentration zum Zeitpunkt der Desinfektion relativ niedrig ist. Mit der Weiterentwicklung der instrumentellen Analytik in den letzten 25 Jahren konnte außerdem gezeigt werden, dass Huminstoffe mit zahlreichen Metallionen, und hier vorzugsweise Schwermetallen, Komplexverbindungen bilden können. Darüber hinaus sind sie in der Lage, hydrophobe organische Substanzen, wie z.B. bestimmte Pestizide oder polycyclische Kohlenwasserstoffverbindungen, sorptiv zu binden. Vom norwegischen Chemiker Egil Gjessing sind sie deshalb auch als ein „Staubsauger" im aquatischen Milieu charakterisiert worden. Obwohl selbst eigentlich unbedenklich, können Huminstoffe also durch Anlagerung oder Einbindung von gesundheitlich bedenklichen Substanzen zu problematischen Komponenten im Wasser werden.

Die beschriebenen Wirkungen und Effekte hängen nicht nur von den Konzentrationen der Huminstoffe, sondern auch von ihren Eigenschaften und von weiteren Wasserqualitätsparametern wie pH-Wert, Salzgehalt und Härte ab. Daher wurde es bisher nicht als sinnvoll erachtet, einen generellen Grenzwert für die Huminstoffkonzentration in Trinkwasser festzulegen. Zwar wird derzeit diskutiert, in die neue deutsche Trinkwasserverordnung eine maximale DOC-Konzentration von 5 mg/l aufzunehmen, doch viel stringenter ist das bei uns und in anderen europäi-

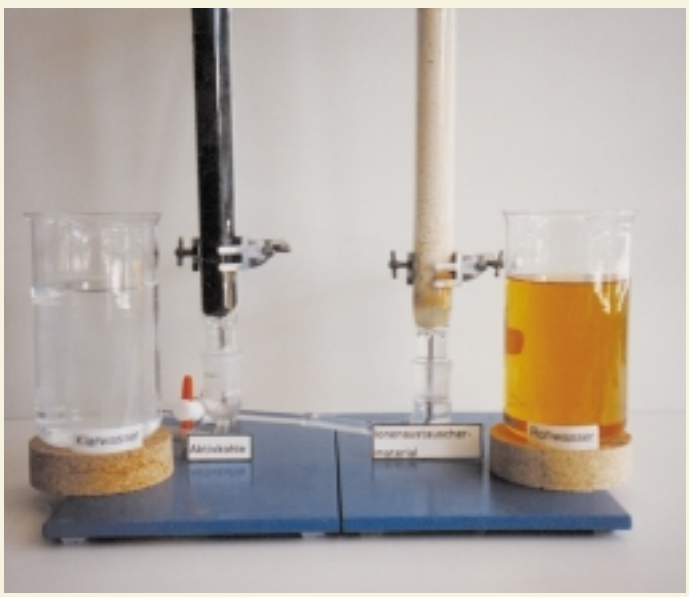


Abb. 4: Körnige Sorbentien zur Entfernung von Huminstoffen.

schen Ländern praktizierte Minimierungskriterium, demzufolge die Bildung von Desinfektionsnebenprodukten bei der Aufbereitung so gering wie möglich zu halten ist.

Aufbereitungsverfahren zur Entfernung von Huminstoffen

Der bereits erwähnte Sachverhalt, dass Huminstoffe bei fast jedem Aufbereitungsverfahren eine aktive Rolle spielen, hat zu zahlreichen Forschungsprojekten geführt, um die jeweiligen Wechselwirkungen aufzuklären. Im Folgenden sollen solche Verfahren kurz dargestellt werden, mit denen Huminstoffe zumindest teilweise aus dem Wasser abgetrennt werden können.

Flockungsverfahren: Unter Flockung versteht man die Überführung von feinpartikulären Trübstoffen im Wasser in Agglomerate (Flocken) durch Zugabe von dreiwertigen Aluminium- oder Eisensalzen. Die Metallionen fallen dabei als Hydroxide aus, die dann einen Teil der Flockenmasse ausmachen. Wenn man die Metallsalze einem trübstofffreien Wasser zusetzt, werden ebenfalls Flocken, und zwar reine Hydroxidflocken, gebildet. Sind in dem Wasser Huminstoffe vorhanden, so werden diese zum Teil in die Hydroxidflocken eingebunden. In Abbildung 3 ist beispielhaft dargestellt, wie der Gehalt an Huminstoffen bei verschiedenen Ausgangskonzentrationen durch Erhöhung der Dosiermengen an Aluminiumsulfat immer weiter verringert werden kann. Neben der Flockenbildung, die im technischen Maßstab in durchströmten Rohrleitungen oder großen Rührbehältern abläuft, umfasst das komplette Behandlungsverfahren noch die Flockenabtrennung durch eine Sedimentations-, Flotations- oder Filtrationsstufe.

Der Wirkungsgrad von Flockungsverfahren für die Huminstoffentfernung liegt üblicherweise im Bereich von 50 bis 80 Prozent, bezogen auf den Parameter DOC. Als maßgeblicher Mechanismus wird eine Sorption der Huminstoffmoleküle an der Oberfläche frisch gefällter Metallhydroxide angesehen, die sich formal über Oberflächenkomplexierungsmodelle beschreiben lässt. Diese Ansätze sind jedoch bislang unbefriedigend, da sie wegen der unvollständigen Kenntnisse der Huminstoffstrukturen zu viele Anpassungsparameter erfordern. Parallel dazu versucht man, den Einfluss bestimmter Randbedingungen wie beispielsweise die Art des Flockungsmittels oder die mittlere Molekülgröße, den Anteil aromatischer Strukturen und die Ladungsdichte der Huminstoffe auf das Flockungsergebnis zu ermitteln. Hierzu sind in den vergangenen Jahren auch eine Reihe von Arbeiten in Höxter, teilweise in Zusammenarbeit mit dem Norwegischen Institut für Wasserforschung in Oslo, durchgeführt worden [2-4].

Weltweit betrachtet kommt den Flockungsverfahren für die Huminstoffentfernung die größte Bedeutung zu. In Deutschland sind sie ein wichtiger Bestandteil der Verfahrenskombinationen zur Aufbereitung von Talsperrenwasser, das vielfach aus Waldgebieten stammt und erhöhte Huminstoffgehalte aufweist. Der anfallende Schlamm wird bislang überwiegend deponiert oder nach einer Trocknung zur Rekultivierung im Landschaftsbau eingesetzt. Wenn es die regionalen Verhältnisse erlauben, kann alternativ nach einer Konditionierung mit Kalk auch eine landwirtschaftliche Verwertung erfolgen. In allen Fällen werden die Huminstoffe also in die Geosphäre zurück verfrachtet, wobei ihre Remobilisierung durch die Einbindung in Flocken jedoch stark eingeschränkt wird.

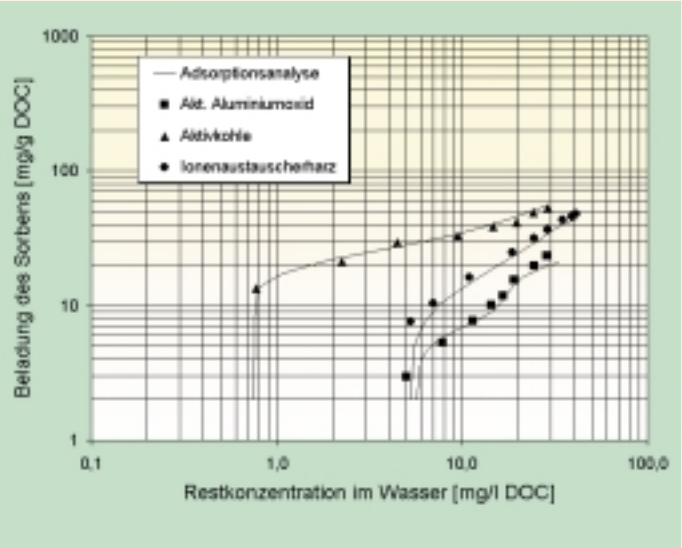


Abb. 5: Sorptionsisothermen für Huminstoffe an drei verschiedenen Sorbentien.

Oxidation/Biofiltration: Huminstoffe werden durch gelösten Sauerstoff nicht angegriffen, wohl aber durch Ozon, dessen entkeimende Wirkung seit über 100 Jahren bekannt ist und das seit vielen Jahrzehnten zur Oxidation gelöster organischer Substanzen bei der Trinkwasseraufbereitung eingesetzt wird. Die zur Mineralisierung der Huminstoffe erforderlichen Ozonmengen sind allerdings unvertretbar hoch, sodass dieses Behandlungsziel unrealistisch ist. Es hat sich jedoch gezeigt, dass bereits bei geringen Ozondosierungen eine Anoxidation der Huminstoffe stattfindet, durch die sie partiell einem schnellen mikrobiologischen Abbau zugänglich gemacht werden. Daher gibt es beispielsweise in Schottland Wasserwerke, in denen das Wasser nach einer Ozonbehandlung über Sandfilterschichten geleitet wird, auf denen sich Mikroorganismen als Biofilm ansiedeln, die dann das organische Substrat umsetzen. Der Gesamtwirkungsgrad liegt hier zwar meist nur bei 20-40 Prozent, bezogen auf den Parameter DOC, doch hat die Voroxidation zwei weitergehende Wirkungen: Zum einen wird die Färbung des Wassers durch die Veränderung der chromophoren Huminstoffstrukturen sehr stark verringert, und zum anderen wird die Reaktivität gegenüber Chlor und damit das Bildungspotenzial für Desinfektionsnebenprodukte deutlich herabgesetzt.

Sorptionsverfahren: Mit Sorption aus der wässrigen Phase wird die Anreicherung gelöster Substanzen an der Oberfläche von Feststoffen (Sorbentien) bezeichnet, die mit dem Wasser in Kontakt stehen. Technische Sorbentien sind hochporös, da nur so die für einen effektiven Einsatz notwendige Oberfläche von einigen 100 m^2/g geschaffen werden kann. Zu ihnen zählen vor allem Aktivkohlen, ferner makroporöse Adsorberharze, Ionenaustauscherharze, aktiviertes Aluminiumoxid und neuerdings auch poröses Eisenoxidhydrat. In den meisten Fällen werden Sorptionsverfahren bei der Trinkwasseraufbereitung zur Entfernung anthropogener Spurenstoffe oder, im Falle der Ionenaustauscher, zur Nitratentfernung und Enthärtung eingesetzt. Alle Sorbentien besitzen aber auch ein gewisses Bindungsvermögen für Huminstoffe, das jedoch nicht vorausberechnet werden kann, sondern experimentell ermittelt werden muss.

Zu den Fragen des Zusammenhangs zwischen Huminstoffeigenschaften und Sorptionskapazitäten sowie der modellmäßigen Beschreibung von Sorptionsprozessen sind auch in Höxter mehrere Untersuchungen durchgeführt worden, über die im

vergangenen Jahr im Rahmen einer internationalen Konferenz in Trondheim (Norwegen) berichtet werden konnte [5, 6]. Abbildung 4 zeigt Laborsäulen, die mit Aktivkohle bzw. Ionenaustauscherharz befüllt sind, um die Entfernung von Huminstoffen aus dem durchfließenden Wasser zu untersuchen. Abbildung 5 zeigt als ein Beispiel für die Datenaufnahme so genannte Sorptionsisothermen für eine Huminstofflösung an den drei Sorbentien Aktivkohle, Anionenaustauscherharz und aktiviertes Aluminiumoxid. Aufgetragen sind die sorbierten Massen als Funktion der Restkonzentrationen im Wasser. Man erkennt, dass sich Sorptionskapazitäten von über 50 mg DOC je g Sorbensmasse erreichen lassen, allerdings nähern sich die Kurven auch der Konzentrationsachse rechts vom Ursprung. Dies ist auf Fraktionen der Huminstoffe zurückzuführen, die durch Sorption nicht entfernt werden können. Ihr Anteil kann zwischen weniger als 5 Prozent bei Aktivkohle und bis zu über 60 Prozent bei Aluminiumoxid liegen. Neben vielen anderen Einflussgrößen spielen dafür auch Molekülausschlusseffekte eine Rolle, da die Sorptionsporen kleiner sein können als große Huminstoffmoleküle.
Die Regeneration erschöpfter Sorbentien erfolgt auf unterschiedlichen Wegen. Aktivkohle wird thermisch reaktiviert, dabei verbrennen die aufgenommenen Substanzen, d.h. der organische Kohlenstoff wird als CO_2 wieder in die Atmosphäre verfrachtet. Die anderen Sorbentien werden mit alkalischen Salzlösungen eluiert, sodass die Huminstoffe dann in konzentrierter Form in entsprechenden Regeneraten vorliegen, deren Entsorgung nicht unproblematisch ist. In Deutschland wurde im Zeitraum von 1978 bis 1995 erfolgreich eine Ionenaustauschanlage zur gezielten Huminstoffentfernung eingesetzt. Sie wurde letztlich deshalb durch ein anderes Verfahren ersetzt, weil die Entsorgung des huminstoffhaltigen Regenerates zu kostenintensiv geworden war. In Skandinavien gibt es eine Reihe kleinerer Wasserwerke, in denen die Huminstoffgehalte im Wasser in weitgehend automatisierten Prozessen durch Ionenaustauscherharze vermindert werden. Dort werden die Regenerate nach einer Neutralisation in Flüsse eingeleitet, d.h. die Huminstoffe verbleiben in der Hydrosphäre.

Membranfiltrationsverfahren: Die Wirkungsweise von Membranverfahren beruht auf dem Durchtritt von Wassermolekülen und gegebenenfalls erwünschten Wasserinhaltsstoffen durch dünne organische Folien oder Hohlfasern, während die abzutrennenden Komponenten zurückgehalten werden. Je nach Trenngrenze unterscheidet man zwischen der Mikrofiltration, Ultrafiltration, Nanofiltration und Umkehrosmose. Für den wirkungsvollen Rückhalt von Huminstoffen sind nominelle Porenweiten von 1 bis 10 nm erforderlich, die dem unteren Bereich der Ultrafiltration bzw. der Nanofiltration entsprechen. In einer technischen Anlage wird der zugeführte Wasserstrom in einen Filtratstrom (Permeat) und einen Konzentratstrom (Retentat) aufgeteilt, bei Systemen zur Huminstoffentfernung typischerweise im Verhältnis von etwa 4:1. Obwohl Membranverfahren seit über 20 Jahren verfügbar sind, werden sie erst seit Mitte der 90er-Jahre gezielt zur Huminstoffabtrennung eingesetzt. Hauptgründe dafür waren neben den anfänglich noch sehr hohen Membrankosten vor allem Probleme mit Belagbildungen auf der Konzentratseite der Membranen im Betrieb, welche die Filtratleistungen stark herabsetzen können. Hierfür mussten geeignete Spül- und Reinigungsmethoden gefunden werden, durch die sich die Leistung der Membranen über Betriebszeiten von mehreren Jahren aufrecht erhalten lassen.
Mit Membranverfahren können Wirkungsgrade für die Huminstoffentfernung von über 80 Prozent, bezogen auf den Parameter DOC, erzielt werden. Da die Konzentrate keine nennenswerten Mengen an Fremdchemikalien enthalten, sondern nur die aus dem Filtrat abgetrennten Stoffe, ist ihre Entsorgung durch Einleitung in einen Fluss meist unproblematisch. Auch hier verbleiben die Huminstoffe somit im Kompartiment Hydrosphäre. Weltweit gibt es inzwischen etwa 200 Membrananlagen, in denen ausschließlich oder unter anderem Huminstoffe abgetrennt werden sollen. Diese Verfahrenstechnik dürfte unter den Aspekten Automatisierbarkeit, Betriebssicherheit und Reststoffentsorgung für kleinere Wasserversorgungen, wie sie in den „Huminstoffregionen“ überwiegend anzutreffen sind, zukünftig die beste Methode sein.

Literatur

[1] Bolin, B., Degens, E.T., Kempe, S., Ketner, P. (Hrg): The Global Carbon Cycle. SCOPE Report No 13, John Wiley & Sons, Chichester, New York, Brisbane, 1979.

[2] Fettig, J., Ratnaweera, H.: Influence of Dissolved Organic Matter on Coagulation/Flocculation of Wastewater by Alum. Wat. Sci. Tech. 27 (1992) No. 11, 103-112.

[3] Ratnaweera, H., Hiller, N., Bunse, U.: Comparison of the Coagulation Behaviour of Different Norwegian Aquatic NOM Sources. Environ. Int. 25 (1999) 2/3, 347-355.

[4] Fettig, J.: Characterisation of NOM by Adsorption Parameters and Effective Diffusivities. Environ. Int. 25 (1999) 2/3, 335-346.

[5] Ødegaard, H. (Hrg): Removal of Humic Substances from Water. Proc. Intern. IAWQ-IWSA Joint Specialist Conference, Trondheim, Norwegen, 24.-26. Juni 1999.

[6] Fettig, J.: Removal of Humic Substances by Adsorption/Ion Exchange. Wat. Sci. Tech. 41 (1999) No. 9, 173-182.

Dipl.-Ing. Claudia Steinert hat an der FH Lippe Biotechnologie studiert und ist seit 1992 Mitarbeiterin für Lehre und Forschung in den beiden Fachgebieten Abwassertechnik und Wassertechnologie im Fachbereich 8. Unter ihrer Leitung sind u.a. die genannten Untersuchungen zur Behandlung huminstoffhaltigen Wassers durchgeführt worden.

Nächstenliebe als Antisemitismus?

Zu einem Problem der christlich-jüdischen Beziehung

Die antisemitische Polemik gegen das Judentum hat sich seit dem 19. Jahrhundert immer wieder auch am Verständnis des biblischen Gebotes der Nächstenliebe festgemacht. Wie sind dabei Nächstenliebe und Judenhass miteinander kompatibel gemacht worden?

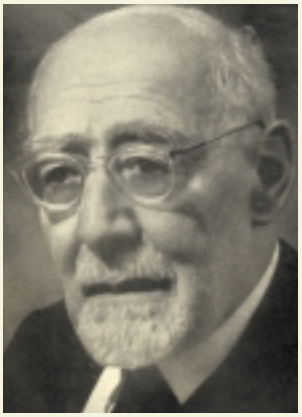
Jüdische Stellungnahme ...

Veranlasst durch die öffentlichkeitswirksame Polemik der Antisemiten, von der Talmudhetze der 1870er-Jahre über den Berliner Antisemitismusstreit ab 1879 und die politischen Bestrebungen, die Emanzipation der Juden rückgängig zu machen, veröffentlichte 1885 der Gemeindevorstand der jüdischen Gemeinde in Berlin „15 Grundsätze der jüdischen Sittenlehre“. Diese beginnen mit den Sätzen:

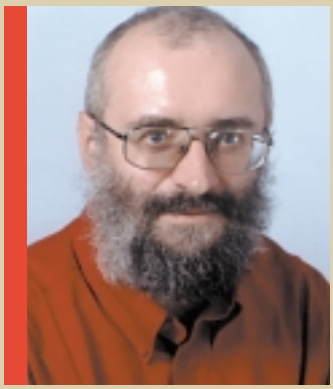
„1. Das Judentum lehrt die Einheit des Menschengeschlechts. Wir haben alle Einen Vater, Ein Gott hat uns alle erschaffen.

2. Das Judentum gebietet: „Liebe deinen Nächsten wie dich selbst“ und erklärt dieses alle Menschen umfassende Gebot der Liebe als Hauptgrundsatz der jüdischen Religion. – Es verbietet daher: gegenüber Jedermann, gleichviel welcher Abstammung er sei, welcher Nation er angehöre und zu welcher Religion er sich bekenne, jede Art von Gehässigkeit, Neid, Missgunst und lieblosses Verhalten; es fordert Recht und Redlichkeit und verbietet Ungerechtigkeit, insbesondre jede Unredlichkeit in Handel und Wandel, jede Uebervorteilung, jede Benutzung (Ausbeutung) der Not, des Leichtsinns oder der Unerfahrenheit eines Andern, sowie jeden Wucher und jede wucherische Ausnutzung der Kräfte Anderer.“

Bis 1889 hatten das Rabbinat der Berliner jüdischen Gemeinde, die Hochschule für Wissenschaft des Judentums und mehr als zweihundert jüdische Theologen Deutschlands diese Sätze approbiert.



Dr. Leo Baeck



Prof. Dr. Martin Leutzsch vertritt an der Universität Paderborn seit 1998 innerhalb des Faches Evangelische Theologie den Schwerpunkt Biblische Theologie und Exegese.

... und christliche Reaktion

Mit wenigen Ausnahmen fand diese Erklärung in der zeitgenössischen christlichen Theologie kein positives Echo. Der evangelische Alttestamentler Bernhard Stade, Mitbegründer der renommierten „Zeitschrift für die alttestamentliche Wissenschaft“ etwa kommentierte:

„Es ist ein durch seine Unverfrorenheit auffallendes Beginnen, wenn versammelte Rabbiner dem christlichen Publikum einzureden versuchen, dass die Juden durch Gebote wie Lev. 19, 18/24, 22 zu gleichem sittlichen Verhalten gegen alle Menschen verpflichtet seien, und das Judentum zur Religion der Menschenliebe stempeln.“

Sofern die Rabbiner nach solchen Grundsätzen handelten, was Stade ihnen zugesteht, täten sie das „unter dem Einflusse christlicher Ethik und gegen die Ethik des talmudischen Judenthums“.

Christliche Vergessenheit

Der englische Philosoph John Stuart Mill schreibt in „Die Nützlichkeit der Religion“ (1853/54), Jesu neues Gebot, einander zu lieben, und weitere wichtige Maximen aus den Reden Jesu stünden „sicherlich zu sehr in Einklang mit dem Denken und Fühlen jedes guten Menschen, um in Gefahr zu geraten, fallen gelassen zu werden, nachdem sie einmal als das Credo des besten und fortgeschrittensten Teils der Menschheit anerkannt“ seien. Mill hebt ihre bleibende Bedeutung für Kultur und Zivilisation hervor. In einer Anmerkung zum Stichwort „neues Gebot“ stellt er fest:

„Das jedoch nicht eigentlich ein neues Gebot ist. Die Gerechtigkeit gegenüber dem großen jüdischen Gesetzgeber gebietet uns, der Tatsache eingedenk zu bleiben, dass die Vorschrift ‚Liebe

deinen Nächsten wie dich selbst' bereits im Pentateuch stand, und es ist in der Tat sehr überraschend, es dort zu finden!"
Mills Verwunderung, das Gebot der Nächstenliebe nicht erst im Neuen, sondern schon im Alten Testament zu finden, ist kein vereinzeltes Phänomen. Eine selektive Bezugnahme auf Sätze des Neuen Testaments, in denen vom Gebot der Liebe als einem neuen Gebot die Rede ist, dürfte zu solchem Vergessen ebenso beigetragen haben wie das von christlicher Theologie immer wieder verbreitete Bild eines Jesus, der mit dem Pathos der Originalität, des Neuen ausgestattet wird. Von Einfluss dürfte auch die bis heute verbreitete Rede von der „christlichen Nächstenliebe" sein, die suggerieren kann, Nächstenliebe sei etwas ursprünglich und exklusiv Christliches.

Die Geburt der christlichen Nächstenliebe aus dem Geist des Antisemitismus

Während der Ausdruck „christliche Liebe" längst in Gebrauch war, findet sich der älteste mir bekannte Beleg für „christliche Nächstenliebe" 1876 in der Schrift „Die ‚goldene' Internationale und die Nothwendigkeit einer socialen Reformpartei" des Juristen C. Wilmanns. Wilmanns kritisiert die dem Geldkapital dienende Gesetzgebung in Kirche und Staat:
„Der herzlose ‚Egoismus', welcher ihr Grundprincip bildet, ist unvereinbar mit christlicher Nächstenliebe. Der Kapitalismus, welcher dadurch blüht, dass er die wirthschaftliche Existenz aller anderen Klassen untergräbt, der Industrialismus, welcher tausende von Menschen körperlich aussaugt und geistig vernichtet, müssen zunächst das christliche Mitgefühl ersticken." „Eine Gesetzgebung, welche die christliche Nächstenliebe ausrottet, erzeugt nothwendig den Klassenhass."
Für die von ihm kritisierte Entwicklung macht Wilmanns die Juden verantwortlich. Deren Wesenszüge entnimmt er dem wenige Jahre zuvor erschienenen, weit verbreiteten antisemitischen Pamphlet „Der Talmudjude" des katholischen Theologen August Rohling.
Wilmanns' Veröffentlichung wird von dem protestantischen Hofprediger Adolf Stoecker im selben Jahr begrüßt, auch im Blick auf die „christliche Nächstenliebe". Stoecker gehörte seit 1879/80 zu den einflussreichen Antisemiten aus christlicher Überzeugung.
Eine Hauptfigur der völkischen Bewegung, der Orientalist Paul de Lagarde, spricht 1881 in dem Pamphlet „Die graue Internationale" von „unserer Nächstenliebe", die indes gegen die Juden nur dann anzuwenden sei, wenn sie durch die Taufe ihrem Judentum völlig entsagt hätten. Ansonsten werde „unsre Aufgabe den Juden Deutschlands gegenüber [...] nicht von der Nächstenliebe, sondern [...] von der Feindesliebe diktiert."
In der Folgezeit wird „christliche Nächstenliebe" immer wieder als antisemitischer Kampfbegriff eingesetzt.

Die antisemitische Polemik gegen die Nächstenliebe der israelitisch-jüdischen Tradition

In der antisemitischen Polemik spielt seit den 1880er-Jahren das Nächstenliebegebot der Hebräischen Bibel eine zentrale Rolle. Viele Antisemiten verstanden sich als christlich und hielten Nächstenliebe für den Inbegriff christlicher Sittlichkeit. Wo Nächstenliebe nicht einfach ohne jede nähere Begründung als unterscheidend christliches Gut in Anspruch genommen wurde, entwickelte man verschiedene Argumentationsstrategien von unterschiedlichem Gewicht und Einfluss. Zusammengenommen zeigen sie den Aufwand an Energie der Bemühungen, dem Judentum einen als zentral angesehenen Wert abzusprechen. Im Einzelnen begegnen folgende Behauptungen:
1. Die Nächstenliebe sei in der Hebräischen Bibel auf Angehörige des Volkes Israel beschränkt und beinhalte einen gänzlich anderen Umgang der Israeliten mit Nichtisraeliten. Diese Behauptung einer jüdischen Binnenethik ist unter all den Argumenten die verbreitetste. Zunächst wird sie nur auf das nachbiblische Verständnis des Gebots im Talmud bezogen. So schreibt der katholische Theologe und spätere Paderborner Bischof Konrad Martin in einem antitalmudischen Pamphlet von 1848, das 1876 erneut aufgelegt wurde:
„Der Nichtjude ist nach Lehre des Talmud weder der Nächste noch der Genosse noch der Bruder des Juden."
Diese Behauptung taucht in der antitalmudischen Hetzliteratur der 1870er- und 1880er-Jahre immer wieder auf. Sie wird im Zuge einer beginnenden Polemik gegen das Alte Testament als einer unter- und unchristlichen Schrift auf das Nächstenliebegebot der Torah selbst angewandt. Der Bonner Theologe und Orientalist Johann Gildemeister behauptete als Gutachter in einem Antisemitenprozess von 1884, sämtliche Professoren hätten längst richtig gestellt, dass in Leviticus 19, 18 das hebräische Wort für „Nächster" nur den Volksgenossen bedeute. Für die Folgezeit sollte diese These in der christlichen Theologie dominieren. Houston Stewart Chamberlain setzt dies in „Die Grundlagen des neunzehnten Jahrhunderts" (1899) bereits ohne Nachweis voraus.
2. Der Vorwurf der Binnenethik konnte mit dem des jüdischen Egoismus verknüpft werden. Der Wagnerianer Arthur Seidl warb für eine Modernisierung, d.h. „Entjudung" des Christentums. Er plädierte dafür, im Nächstenliebegebot auch des Neuen Testaments nicht die höchste Stufe der Ethik zu sehen. Ihm missfiel das „wie dich selbst", „weil dies [...] noch immer seinen fatalen alttestamentlich-selbstischen Beigeschmack mit verrät". Höher stünde das paulinische „Einer trage des andern Last", das zur Identifikation mit dem Nächsten und zum Mitleiden mit ihm bewege und dem christlichen Heiland (wie auch Wagners Parsifal) besser entspreche.
3. Wenige Sätze nach dem Nächstenliebegebot steht in Leviticus 19 das Gebot: „Liebe den Fremden (...) wie dich selbst." (19, 34) Sofern Antisemiten diesen Satz überhaupt berücksichtigten, wurde er im Sinne der schon erwähnten Binnenethik entschärft: Mit den Fremden seien ausschließlich die zum Judentum konvertierten Nichtjuden gemeint.
4. Eine andere Delegitimationsstrategie wurde im Zuge des von dem Orientalisten Friedrich Delitzsch 1902 ausgelösten Babel-Bibel-Streits entwickelt. Das Nächstenliebegebot sei nichts originell Israelitisches, sondern begegne bereits bei den Babyloniern. Dass Delitzsch bei seiner Argumentation zu gefälschten Zitaten griff, fiel indes nur wenigen Fachleuten auf.
5. Hans Blüher, führender Männlichkeitsideologe der Jugendbewegung, akzeptierte die alttestamentliche Herkunft des Nächstenliebegebots in der Jesusüberlieferung. Er bestritt aber, dass dieses Gebot für Jesus zentral sei. Vielmehr sei für Jesus ein aggressives, gegenüber dem Judentum an den Tag gelegtes Heldentum kennzeichnend, das auch für die aktuellen Auseinandersetzungen in der Weimarer Republik vorbildhaft sei.
6. Schließlich konnte auch der im interreligiösen Diskurs nicht

seltene Vergleich der fremden Praxis mit der eigenen Theorie verwendet werden, so etwa von Eugen Dühring:
„Die Heuchelei von Nächstenliebe und Mitleid ist schon sehr früh, lange vor dem Anfang unserer Zeitrechnung, eine Domäne des Judenthums gewesen. Christus nahm das, was Andere heuchelten, gewissermaassen ernst und aufrichtig, wie er denn überhaupt auch die ganze jüdische Religion ernst zu nehmen versuchte."
Unter all diesen Strategien ist das Konzept der jüdischen Binnenethik das dominierende. Noch der junge Georg Wilhelm Friedrich Hegel, aber auch Karl Gutzkow in seinem berüchtigten Roman „Wally, die Zweiflerin" (1835) und David Friedrich Strauß in „Der alte und der neue Glaube" (1872) hatten erklärt, Jesus habe keine neuen Lehren gebracht, und das Nächstenliebegebot als Beispiel dafür angeführt. In der Folgezeit wuchs der Bedarf, das Gebot der „christlichen" Nächstenliebe als etwas Originelles hinzustellen. In einer Zeit, in welcher der antisemitische Diskurs den Juden mangelnde kulturelle Originalität zuschrieb, wurde die Originalität christlicher Nächstenliebe behauptet und als Überwindung jüdischer Doppelmoral und Binnenethik ausgelegt.
Es ist die Behauptung jüdischer Binnenethik, in der politischer Rasseantisemitismus und protestantische Theologie am stärksten konvergierten und kooperierten. Dem sich organisierenden Antisemitismus ging es darum, die Gleichberechtigung der Juden rückgängig zu machen. Die liberale protestantische Theologie brauchte die Abgrenzung vom Judentum, um das Christentum als die überlegene und bessere Religion herauszustellen. In der Frage der Nächstenliebe kam es insbesondere bei den Antisemitenprozessen zu einer Unterstützung der wegen Beschimpfung der jüdischen Religionsgemeinschaft Angeklagten durch protestantische Theologen. Gildemeisters Gutachten im Prozess von 1884 blieb nicht das einzige. 1912/13 etwa, im Gotteslästerungsprozess gegen Theodor Fritsch, einen der einflussreichsten Propagandisten des Rasseantisemitismus, erhielt der Angeklagte durch die protestantischen Bibelwissenschaftler Georg Beer und Johannes Meinhold gutachterliche Unterstützung; Fritschs Freispruch verdankt sich insbesondere dem Obergutachten des angesehenen protestantischen Alttestamentlers Rudolf Kittel, der die jüdische Binnenmoral anhand des Nächstenliebegebots in Leviticus 19, 18 zu erweisen beanspruchte.

Die Behauptung jüdischer Binnenethik in Geschichte und Gegenwart

Im antisemitischen Diskurs der 1870er- und 1880er-Jahre war das Konzept jüdischer Binnenethik vor allem infolge der antitalmudischen Pamphlete Rohlings, Martins und anderer gegenwärtig. Diese griffen zurück auf Johann Andreas Eisenmengers Buch „Entdecktes Judenthum" (1711). Eisenmenger hatte mithilfe entstellter, verfälschter oder aus dem Zusammenhang gerissener Talmudzitate ein Bild des Judentums entworfen, das Solidarität nur unter Juden mit Hass, Verachtung, Erniedrigung und Elimination gegenüber Nichtjuden verband. Vertreter der Aufklärung und des Idealismus und Theologen der Restaurationszeit transportierten dieses Stereotyp weiter. Im Diskurs der Aufklärung und der Restauration um die politische Gleichstellung der Juden diente die jüdische Binnenethik als Argument der Emanzipationsgegner.
Während Eisenmenger und seine hoch- und spätmittelalterlichen Vorgänger die jüdische Binnenmoral aus nachbiblischen Texten herauslasen, hatte eine Generation vor Eisenmenger Baruch Spinoza in seinem „Tractatus theologico-politicus" (1670) eine Beschränkung der Nächstenliebe auf die Volksangehörigen aus der Hebräischen Bibel selbst abgeleitet. Spinoza ist dabei bis in einzelne Formulierungen hinein von Tacitus abhängig. Wenn der römische Geschichtsschreiber sagt, unter den Juden walteten unverbrüchliche Treue und hilfsbereites Mitleid, gegen alle anderen aber feindseliger Hass, greift er ein Stereotyp der antiken Judenfeindschaft auf, das seit dem 3. Jahrhundert v.u.Z. insbesondere bei ägyptischen und römischen Autoren begegnet.
Diese antijüdischen Vorurteile der Antike, des Mittelalters und der frühen Neuzeit werden Ende des 19. Jahrhunderts in der protestantischen Bibelwissenschaft aufgenommen. Die jüdische Binnenethik wird philologisch und historisch-kritisch zu erweisen gesucht. Die Ansicht, Nächstenliebe beschränke sich in der Hebräischen Bibel auf Israeliten und sei erst von Jesus auf alle Menschen ausgedehnt worden, findet Eingang in die maßgeblichen Wörterbücher und hält sich bis heute in theologischen Fachlexika ebenso wie in den Konversationslexika. Die Gegenargumente eines protestantischen Alttestamentlers wie Johannes Hänel, der 1924 ausführlich den volksübergreifenden Geltungsbereich des israelitischen Nächstenliebegebots begründete, fanden ebenso wenig Gehör wie die zahlreichen Veröffentlichungen jüdischer Gelehrter zum Thema.
Bis heute wirksam ist jene Version der Binnenethik-These, die Max Weber in seiner Studie „Das antike Judentum" (1917/19) anhand des israelitischen Zinsrechts und unter Ausblendung des umfassenden Fremdenrechts der Hebräischen Bibel vorlegte. Weber griff die protestantische Exegese und Altertumswissenschaft auf, verarbeitete das antisemitische Pamphlet „Die Juden und das Wirtschaftsleben" (1911) seines Kollegen Werner Sombart und interpretierte das Nächstenliebegebot in Leviticus 19, 18 im Sinn von Nietzsches Nächstenliebekritik der 1880er-Jahre als Ausdruck des racheerfüllten Ressentiments von Zukurzgekommenen. Mit seiner These wirkt Weber auch bis in Bereiche der heutigen christlichen Bibelwissenschaft nach.

Die Nächstenliebe der Antisemiten

Die Mehrheit der Antisemiten im kaiserzeitlichen Deutschland verstand sich als christlich, wenn auch nicht unbedingt als kirchlich. Für ihr christliches Selbstverständnis spielte Nächstenliebe eine wichtige Rolle. Wie vereinbarten sie ihre Betonung der Nächstenliebe mit ihrem Antisemitismus?
Für Julius Langbehn, Verfasser des völkischen Bestsellers „Rembrandt als Erzieher", stand fest:
„Die Deutschen dürfen den Juden so wenig die Liebeshand bieten, wie Christus sie den Pharisäern bot; hier ist das strenge, nicht das sanfte Christenthum am Platze. Otterngezücht!"
Der ariosophische Ordensgründer Georg Lanz, der sich von Liebenfels nannte, forderte: „Liebe deinen rassisch Nächsten und töte alle teuflischen Untermenschen."
Julius Streichers „Der Stürmer" betonte 1927, dass man gegen ein Volk, das den Mord von Golgatha auf dem Gewissen habe, keine Nächstenliebe zu üben brauche.
Der Katholik Josef Alois Kofler schreibt 1928:
„Unsere Religionspolitik gegen die Juden muss sein, sie, die Fremdkörper in unserm Land und in unserer Geisteskultur sind,

auszuschalten aus dem öffentlichen Leben. Das ist die durch die Verhältnisse uns aufgezwungene Form der Nächstenliebe gegen die Juden, ohne allen Hass, aber aus Liebe zu Christentum und Volkstum. Christliche Nächstenliebe in der Judenfrage ist in erster Linie Liebe zu den Nächstgeborenen, zu den Angehörigen des eigenen Volkes, nicht Schutz der Juden vor den Christen, sondern Schutz der Christen vor den Juden. Die Liebe zum eigenen Volk schließt notwendig einen Abwehrkampf gegen das übermächtige Judentum in sich. Dieser Abwehrkampf ist somit kein Verstoß gegen die Nächstenliebe, sondern Betätigung derselben. „Judengegnerschaft ist Notwehr. Notwehr aber ist niemals Verstoß gegen christliche Nächstenliebe."

Wilhelm II. 1929 im Brief an Hans Blüher:

„Mein Nächster, mein Bruder der mir entgegentritt, ist immer und Allewege zuerst der Deutsche!"

Der deutsch-christliche Theologe Friedrich Wieneke 1931:

„Der Herr Jesus hat nie von ‚allgemeiner Menschenliebe' gesprochen, sondern von Nächstenliebe, die in den politischen Angelegenheiten dieser Tage gewiss die Liebe zum deutschen Volkstum wäre."

Joseph Goebbels:

„Das aber ist Christentum: Liebe deinen Nächsten wie dich selbst. Mein Nächster ist mein Volksgenosse. Liebe ich ihn, dann muss ich seine Feinde hassen."

Äußerungen wie diese zeigen, dass die antisemitische Interpretation der Nächstenliebe auf genau jene gewalttätige Binnenethik hinauslief, die man den Juden vorwarf. Liegt hier Projektion vor? Diesen Verdacht äußerte bereits der junge Rabbiner Leo Baeck 1902 in seiner Besprechung von Adolf Harnacks „Das Wesen des Christentums". Der prominente protestantische Theologe hatte dem Judentum zur Zeit Jesu unterstellt, dass der sittliche Quell, der damals in ihm floss, verunreinigt gewesen sei, und dass es die Rabbinen und Theologen seien, die „nachträglich dieses Wasser destillieren". Baeck kommentiert:

„Man muss unwillkürlich an den Satz des Talmuds denken, dass es menschliche Art sei, den eigenen Mangel anderen vorzuwerfen."

Baeck fragt, ob es das Privileg protestantischer Theologieprofessoren sei, das Wesentliche einer Religion hervorzuheben, und fährt dann fort:

„Wenn einer von Zeit zu Zeit einmal schüchtern darauf aufmerksam zu machen sucht, dass im dritten Buche Mosis im neunzehnten Capitel im achtzehnten Verse geschrieben steht: ‚Liebe deinen Nächsten wie dich selbst,' und wenn er dann noch gar hinzuzufügen wagt, dass eben daselbst im vierunddreissigsten Verse gesagt ist: ‚liebe den Fremdling wie dich selbst' und dass unter dem Fremdling der Mensch aus anderem Volke und von anderem Glauben verstanden werden muss, sintemal genannter Vers also begründend fortfährt: ‚denn Fremdlinge waret ihr im Lande Egypten' – alsdann ertönt von verschiedenen Kathedern im deutschen Vaterlande der Vorwurf: Seht den Destillator!"

Projektion der eigenen Binnenethik im antisemitischen Falschzitat?

Der Versuch, die antisemitische Interpretation des Nächstenliebegebots als Projektion gewalttätiger Binnenethik nach außen zu verstehen, kann eine Stütze finden in der Eigenart der Zitate, die in der Debatte auftauchen. Nicht nur Friedrich Delitzsch trug in seinem antisemitischen Alterswerk „Die große Täuschung" 1920 mithilfe freier Erfindung die Nächstenliebe in einen altbabylonischen Text hinein. In einer Zeit, in der Antisemiten das „Wesen des Judentums" als etwas seit alters her Unverändertes und Unveränderliches zu definieren unternahmen, zitierte der Historiker Heinrich von Treitschke 1879 in jenem Artikel, der mit dem Satz „Die Juden sind unser Unglück!" den Berliner Antisemitismusstreit auslöste, den römischen Geschichtsschreiber Tacitus.

„Eine Kluft zwischen abendländischem und semitischem Wesen hat von jeher bestanden, seit Tacitus einst über das odium generis humani klagte."

Die Tacituswendung, die vom Hass gegen das ganze Menschengeschlecht spricht, taucht in der antisemitischen Polemik seit Johann Gottlieb Fichte 1793 immer wieder auf, auch bei dem Anti-Antisemiten Friedrich Nietzsche und selbstverständlich auch bei Theodor Fritsch. Die bedeutenden jüdischen Gelehrten Heinrich Graetz und Manuel Joel wiesen vergeblich darauf hin, dass diese Wendung bei Tacitus auf die von Kaiser Nero verfolgten Christen bezogen sei.

Von diesem Phänomen falscher Zitate aus wäre auch an einen Text der Bergpredigt des Matthäusevangeliums (5, 43) eine Anfrage zu richten. Mit den Worten „Ihr habt gehört, dass gesagt ist: Liebe deinen Nächsten und hasse deinen Feind" wird der Abschnitt über die Feindesliebe eingeleitet. Nur der erste Teil, die Aufforderung, den Nächsten zu lieben, findet sich in der Hebräischen Bibel, nicht jedoch die Aufforderung zum Feindeshass. Ob hier nicht auch eine Projektion eigener Aggression nach außen vorliegen könnte?

Die ungehörte jüdische Stimme

Christian Wiese bestätigt in seinem Buch über „Wissenschaft des Judentums und protestantische Theologie im wilhelminischen Deutschland" (1999) in vielen Hinsichten Gershom Scholems Argumente „Wider den Mythos vom deutsch-jüdischen Gespräch". Jüdische Äußerungen zur Hebräischen Bibel und zum Judentum wurden von protestantischen Theologen oft nicht einmal wahr-, geschweige denn ernst genommen. Die grundlegende Asymmetrie im jüdisch-christlichen Diskurs zeigt sich etwa in Rudolf Kittels Obergutachten im schon erwähnten Gotteslästerungsprozess gegen Theodor Fritsch. Auf jüdischer Seite hatte der bedeutende Neukantianer Hermann Cohen seit 1888 mehrfach in Veröffentlichungen zur Nächstenliebe in der Hebräischen Bibel und im Talmud dargelegt, die Nächstenliebe in Leviticus 19 sei nicht national beschränkt. Kittel ging in seinem Gutachten auf keines von Cohens Argumenten ein. Er begnügte sich mit der Bemerkung, Cohen besitze zwar einen bekannten Namen, doch ein gewesener Professor der Philosophie sei auch dann, wenn er geborener Jude sei, noch nicht ein sachkundiger Erklärer der hebräischen Bibel.

Erneuerung des Verhältnisses der Christen zu den Juden

Eine Änderung der christlichen Beziehung zu den Juden und Jüdinnen wird nur dann möglich werden, wenn die jüdischen Partnerinnen und Partner wahrgenommen und in ihren Äußerungen ernst genommen werden. Für Christen hieße dies zuerst, den Reichtum jüdischer Auslegung des Nächstenliebegebots wahrzunehmen. Dies beginnt mit jener jüdischen Übersetzungs-

tradition, welche die Wendung „wie dich selbst“ als Begründung auffasst und wiedergibt mit: „er ist wie du“. Hermann Cohen, Martin Buber und Franz Rosenzweig, Max Horkheimer und Emmanuel Levinas sind nur einige Befürworter dieser Übersetzung. Entsprechend antiker jüdischer Auslegungstradition wird damit die Gleichheit und Würde der Beteiligten hervorgehoben. Im Hintergrund schwingt die kühne Botschaft der Hebräischen Bibel mit, wonach der Mensch Gottes Ebenbild sei. Und das gilt nach Genesis 1 bekanntlich für alle Menschen.

Flugplanung mit Informatik-Methoden

Optimierungsverfahren erhöhen die Produktivität, steigern die Mitarbeiterzufriedenheit und senken die Kosten bei Fluggesellschaften

»Mein Name ist Ute Schmidt. Ich begrüße Sie auf Ihrem Flug von Frankfurt nach Rom in unserem Airbus A 320 ....«. Bevor diese Ansage ertönen kann, stellen sich der Fluggesellschaft einige komplexe Planungsaufgaben. Wie wichtig es ist, diese Aufgaben möglichst optimal zu lösen, wird klar, wenn man sich vor Augen hält, dass bei großen Fluggesellschaften eine Kostenreduktion von nur 50 Euro pro Abflug eine Ersparnis in zwei- bis dreistelliger Millionenhöhe im Jahr bedeuten kann!
Die Arbeitsgruppe von Prof. Dr. Burkhard Monien aus dem Fachbereich Mathematik-Informatik hat sich in den letzten Jahren intensiv mit einigen dieser Planungsprobleme im Flugbereich beschäftigt. Zusammen mit Industrie- und Universitätspartnern sind im BMBF-Projekt PARALOR, dem EU-ESPRIT-Projekt PARROT und einer direkten Industriekooperation Modelle und Verfahren entwickelt worden, um die komplexen Planungsprobleme der Fluggesellschaften mit Informatik-Methoden anzugehen.

Bereits Jahre vor der Realisierung beschäftigen sich Planer einer Fluggesellschaft mit den Vorstufen zu einem Flugplan. Bis der Passagier im Flugzeug Platz nehmen kann, sind eine ganze Reihe von Planungs- und Entscheidungsproblemen seitens der Fluggesellschaft zu lösen. Angesichts der Komplexität und Größe werden die Probleme in verschiedenen Phasen hintereinander gelöst.

Zwei Jahre vor Abflug beginnt die Auswahl der Städteverbindungen, die von der Fluggesellschaft bedient werden sollen. Passagierströme zwischen den Städten werden prognostiziert und die Häufigkeit der Verbindungen festgelegt. Unter Beachtung vermuteter Angebote der Mitbewerber werden danach die Abflugzeiten definiert und vorläufige Preisklassen bestimmt.

Prof. Dr. Burkhard Monien ist seit 1977 Professor für Theoretische Informatik im Fachbereich 17/Mathematik – Informatik der Universität Paderborn. Sein Arbeitsgebiet ist die effiziente Nutzung von parallelen und verteilten Systemen. Fragestellungen reichen von den Grundlagen (Load Balancing, Mapping) hin zu den Anwendungen (Kombinatorische Optimierung, Computergrafik, Multimedia-Systeme, Spielbaumsuche).

Anderthalb Jahre vor Abflug beginnen die Planungen der Flugzeugtypen für die einzelnen Flüge (Fleet Assignment). Danach werden Einsatzpläne für die Flugzeuge erstellt, die unter anderem vorgeschriebene Wartungen einschließen müssen.

Sechs Wochen vor Abflug werden erste Entscheidungen über den Einsatz der Crews, die aus Piloten und Kabinenpersonal bestehen, getroffen. Dies geschieht in einem ersten Schritt für anonymes Flugpersonal, in einem zweiten für die tatsächlichen Personen mit ihren individuellen Qualifikationen und Präferenzen. Neben Treibstoffkosten ist die Bezahlung der Crews einer der größten Kostenpunkte bei den Fluggesellschaften.

Ein Flugplan wird erfahrungsgemäß **kurz vor Abflug** oft durch unvorhersehbare Umstände gestört, z.B. Ausfall der Maschine, Krankheit eines Piloten oder ungünstige Wetterbedingungen. Um möglichst wenig Umstände für die Passagiere zu verursachen, muss die Fluggesellschaft für Ersatzflugzeuge oder -crews sorgen, Flugzeiten verschieben oder zusätzliche Flüge anbieten.

Zwei Jahre vor Abflug müssen die weltweiten Strecken festgelegt werden, die in den kommenden Flugplänen berücksichtigt werden sollen.

Im Fleet Assignment wird eine günstige Zuweisung der verschiedenen Flottentypen auf die zu fliegenden Strecken bestimmt.

Foto: Werner Krüger/Lufthansa

Manche der beschriebenen Planungsabschnitte werden nochmals in Teilphasen gegliedert. Die Aufteilung des Planungsprozesses in Phasen liefert zwar kleinere, handhabbare Fragestellungen, trennt aber kausale Zusammenhänge des Gesamtproblems. So kann zum Beispiel ein zu dicht gepackter Flugzeugeinsatzplan dazu führen, dass Crews ohne notwendige Pausen im Einsatz wären. Um solche Wechselwirkungen zu dämpfen, werden die einzelnen Phasen in Teilen iteriert. Je näher man dabei dem Abflugtermin kommt, desto eingeschränkter wird der Raum für Änderungen. Um starke Interferenzen zwischen zwei Phasen in der Problem- und Lösungsstruktur zu berücksichtigen, gehen Bestrebungen in der Wissenschaft dahin, einzelne Phasen wieder miteinander zu verschmelzen.

Warum Informatik-Methoden?

Die Komplexität der einzelnen oben beschriebenen Planungsschritte legt es nahe, diese Aufgaben nicht von einem menschlichen Planer allein, sondern mithilfe von Computern und geeigneten Verfahren durchzuführen. Bei einer genaueren Untersuchung der Probleme mit Konzepten der Theoretischen Informatik zeigt sich, dass diese fast alle zur Klasse der so genannten NP-harten Probleme gehören; Probleme dieser Kategorie haben sich bisher erfolgreich einer effizienten Berechnung entzogen und es wird angenommen, dass es für diese Probleme keine effizienten (schnellen) Berechnungsverfahren gibt. Als Beispiel mag die optimale Zusammenstellung von Dienstplänen dienen. Aus den n Einzelflügen müssen Arbeitspakete gruppiert werden, die an einer Heimatbasis starten und enden. Potenziell können hier schon 2^n verschiedene Arbeitspakete gebildet werden. Natürlich ist nicht jedes mögliche Arbeitspaket sinnvoll. Nehmen wir an, 99,99999 Prozent der Arbeitspakete können wir sofort als ungültig erkennen. Dann bleiben uns in einer Planungswoche mit 320 Flügen immer noch $2{,}13 \bullet 10^{90}$ Arbeitspakete – mehr als die vermutete Anzahl Atome im Universum! Informatik-Methoden helfen hier eine Auswahl „guter“ Arbeitspakete zu finden; typischerweise werden mit den geeigneten Verfahren „nur“ mehrere hunderttausend bis einige Millionen Arbeitspakete erzeugt und bewertet und liefern damit schon hinreichend gute und kostengünstige Dienstpläne.

Eine ähnliche kombinatorische Explosion ist bei der nachfolgend beschriebenen Zuordnung von Flugzeugtypen auf die Flüge zu beobachten; auch hier stellt sich die Frage nach möglichst effizienten und guten Lösungsverfahren.

Blick ins Cockpit einer Boeing 747-400 der Lufthansa kurz vor Abflug. Pilot und Copilot dieses Fluges wurden im Crew Scheduling festgelegt.

Foto: Gerd Rebenich/Lufthansa

Foto: Udo Kröner/Lufthansa

Einchecken der Passagiere am Flughafen Frankfurt. Kurz vor dem Abflug werden die vorher geplanten Flugzeuge und Besatzungen noch an die aktuellen Gegebenheiten angepasst.

Fleet Assignment – nehmen wir die Boeing oder den Airbus?

Beim Fleet Assignment (engl. Flottenzuweisung) wird für jeden Flug der Flugzeugtyp bestimmt, der diesen Flug übernehmen soll. Dabei ist die Anzahl der Flugzeuge je Flugzeugtyp eine knappe Ressource, die es einzuhalten gilt. Darüber hinaus soll der Gewinn der Flottenzuweisung maximiert werden. Die Erträge und Kosten eines Fluges werden maßgeblich durch den Flugzeugtyp in Form von Sitzplätzen, Treibstoffverbrauch, Landegebühren und Personalkosten beeinflusst. Die Anzahl der zu erwartenden Passagiere wird vor der Berechnung eines Fleet Assignment durch Marktmodellierung basierend auf Erfahrungswerten (Vorjahr) und Prognosen geschätzt.

Anschaulich geht es beim Fleet-Assignment-Problem darum, ein Flugzeug des richtigen Typs zur richtigen Zeit am richtigen Ort zu haben. Wenn Sie also eines Tages in einem halb leeren Flugzeug sitzen, muss das nicht bedeuten, dass die Fluggesellschaft schlechten Zeiten entgegensieht, sondern es könnte auch sein, dass das Flugzeug an Ihrem Zielflughafen eine bessere Einsatzmöglichkeit hat, die den halb leeren Flug mehr als wett macht.

Wenn sich die Planer einer Fluggesellschaft anderthalb Jahre bis sechs Monate vor Umsetzung des Flugplanes mit dem Problem des Fleet Assignment beschäftigen, geht es neben der eigentlichen Flottenzuweisung auch um deren Folgewirkungen. Das sind Fragen, die die Zusammensetzung der Flotte (welche Flugzeuge sollen verkauft/gekauft werden), Standorte für Wartungsstationen, Wartungspersonal und Ausbildung der Piloten betreffen. In dieser Phase wird von international operierenden Fluggesellschaften ein Fleet Assignment für eine standardisierte Woche berechnet. Sechs Monate bis einen Monat vor Umsetzung des Flugplanes rücken neben der Maximierung des Gewinnes auch andere Kriterien in den Vordergrund, wie z.B. die tatsächlich vorhandenen Piloten für einen Flugzeugtyp. Diese Planungsphase bezieht sich auf einen Zeitraum von zwei bis sechs Wochen, in dem nun – abweichend von der Standardwoche – auch Besonderheiten wie Ferienzeiten und Feiertage einbezogen werden.

Heuristische Lösung des Fleet-Assignment-Problems

Für die Bestimmung eines Fleet Assignments gibt es im Allgemeinen zwei Ansätze: exakte und heuristische Methoden. Bei den exakten Ansätzen wird eine optimale Lösung berechnet, in der Regel mit Linearer Programmierung. Dieser Ansatz stößt bei größeren Probleminstanzen oder bei zusätzlichen Bedingungen oft an die Grenzen des mit heutiger Technologie Berechenbaren. Bei einer heuristischen Herangehensweise verzichtet man auf die Optimalität der Lösung, um in einer akzeptablen Zeit sehr gute Lösungen bzw. überhaupt zulässige Lösungen zu erhalten. Eine *optimale* Lösung ist in der Praxis angesichts langer Berechnungszeiten oft nicht sinnvoll, insbesondere wenn es sich – wie im Fleet Assignment – bei den zu Grunde liegenden Daten nur um Schätzungen handelt.

Im Bereich Fleet Assignment verfolgen wir im Rahmen einer direkten Industriekooperation mit Lufthansa Systems, Frankfurt, einen heuristischen Ansatz, dessen Kernstück eine problemspezifische *lokale Suche* darstellt. Bei einer lokalen Suche werden ausgehend von einer gegebenen (aber i.d.R. nicht optimalen) Lösung einige wenige Veränderungen vorgenommen, sodass sich die Lösung für einen bestimmten Bereich ändert (z.B. durch den Austausch einer Boeing 747 gegen einen Airbus A 320 auf einigen Flügen), aber für den restlichen Bereich (d.h. die restlichen Flüge) gleich bleibt. Solche kleinen Veränderungen (Austausche) werden so lange durchgeführt, bis die Qualität der Lösung zufrieden stellend ist oder sich über eine grössere Anzahl von Austauschen nicht mehr verbessert.

Die Grafik in Abbildung 1 verdeutlicht eine mögliche lokale Änderung an einem kleinen Beispiel. Die Flüge (Pfeile) zwischen den drei Flughäfen werden von zwei Flugzeugtypen (rot und grün) ausgeführt, von denen je ein Flugzeug vorhanden ist. Nehmen wir nun an, dass sich durch einen Wechsel der Flottenzuordnung von grün nach rot von Flug 4 ein beträchtlicher Mehrgewinn erzielen lässt. Aber was passiert, wenn man nur die Flottenzuweisung von Flug 4 ändert? Man benötigt für jeden der beiden Flugzeugtypen zwei Flugzeuge und hat somit keine Lösung mehr! Um weiterhin mit einem Flugzeug je Typ auszukommen, können z.B. die Zuweisungen für die Flugsequenzen (1, 2, 3) und (4, 5) ausgetauscht werden.

Die von uns verwendeten Austausche in der lokalen Suche gleichen den Änderungen, die auch ein menschlicher Experte vornehmen würde. Wieso berechnet unser Ansatz trotzdem viel bessere Lösungen als die Fleet Assignments, die per Hand von

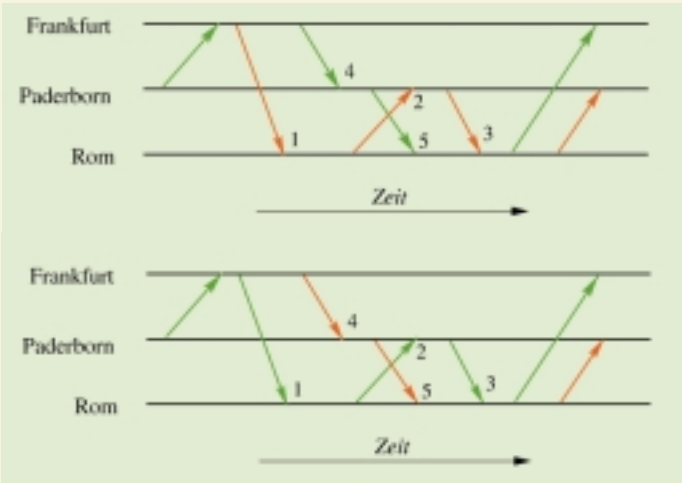


Abb. 1: Um von einem gültigen Fleet Assignment zu einem anderen zu gelangen, müssen die Flottenzuordnungen ganzer Flugsequenzen ausgetauscht werden, wie hier im Fall der Flüge des roten und grünen Flottentyps. Die Grafiken zeigen farblich die Flottenzuweisung vor und nach dem Austausch des Flugzeugtyps für die Flugsequenzen (1, 2, 3) und (4, 5).

Abb. 2: Parallelrechner, wie die Siemens hpcLine des PC² mit 192 Prozessoren, helfen bei der Berechnung großer Fleet Assignments.

den Fluggesellschaften ausgetüftelt wurden? Dies hat mehrere Gründe: Unsere Heuristik betrachtet viele lokale Änderungen (mehrere Millionen), viel mehr als ein Mensch sich je anschauen könnte. Außerdem werden auch verschlechternde Austausche zugelassen. Letzteres ist wichtig, um nicht frühzeitig in Sackgassen zu enden (lokale Optima), bei der durch lokale Änderungen keine Verbesserung mehr möglich ist, der Wert der Lösung aber noch weit vom (globalen) Optimum entfernt ist. Wann man wie viele verschlechternde Übergänge zulässt, wird in unserer Heuristik durch Simulated Annealing, einer globalen Steuerung, entschieden. Simulated Annealing lässt verschlechternde Lösungen mit einer bestimmten Wahrscheinlichkeit zu, wohingegen verbessernde Austausche immer akzeptiert werden. Die Wahrscheinlichkeit für das Akzeptieren von verschlechternden Lösungen ist am Anfang der Berechnung sehr hoch, wird aber im Verlauf der Berechnung schrittweise bis auf null reduziert, sodass am Ende nur noch verbessernde Austausche zugelassen werden. Ein weiterer Grund für die Überlegenheit unseres Ansatzes gegenüber einer Planung per Hand ist die Möglichkeit, längere und komplexer gestaltete Sequenzen von Flügen zu betrachten als es ein menschlicher Experte üblicherweise tut.

Parallelrechner lösen große Fleet-Assignment-Instanzen

Während bei Fluggesellschaften mittlerer Größe (ca. 4 000 Flüge pro Woche, 11 Flugzeugtypen) die Berechnungszeit unseres Ansatzes für eine Planungswoche gering ist (ca. 20 bis 30 Minuten), kann sie für Flugpläne, die bis zu 6 Wochen umfassen (ca. 24 000 Flüge), schon einmal einen ganzen Tag einnehmen. Demnach ist insbesondere im letzteren Fall der Mehrwochenplanung Potenzial und Bedarf für eine parallele Berechnung eines Fleet Assignment gegeben. In einem Parallelrechner arbeiten viele Prozessoren gemeinsam an einer Lösung. Besonders wichtig ist dabei die Verteilung der Berechnungen auf die verschiedenen Prozessoren. Die Kette der Lösungen, die sich durch aufeinander folgende Austausche ergibt, ist nicht im Vorhinein bekannt, sondern ergibt sich erst während der Berechnung. Somit kann diese Kette nicht einfach zu Beginn auf die Prozessoren aufgeteilt werden. Wir haben Methoden entwickelt, welche die Steuerungsheuristik Simulated Annealing parallelisiert (zu finden in der Library „parSA“). Dabei arbeiten die Prozessoren zu Beginn an verschiedenen Lösungsketten. Während der Berechnung schließen sich die Prozessoren zu immer größeren Clustern zusammen, die dann gemeinsam eine viel versprechende Lösungskette bearbeiten. Dies ist sinnvoll, da es im Laufe des Algorithmus immer schwieriger wird, Lösungen zu finden, die akzeptiert werden, und somit sehr viel Zeit mit der Suche nach Austauschen verbracht wird. Zur Implementation von parallelen Programmen stehen im PC² (Paderborn Center for Parallel Computing) einige moderne Parallelrechner zur Verfügung, wie z.B. der Rechner aus der Siemens hpcLine (Abbildung 2).

Regelmäßigkeiten und kurze Rechenzeiten sind in der Praxis wichtig

Für die Umsetzung der berechneten Fleet Assignments in die Praxis ist es notwendig, dass sie zu einem gewissen Maße zusätzliche Eigenschaften besitzen. Eine solche Eigenschaft ist die gleichmäßige Struktur des Fleet Assignment, d.h. Flüge, die mehrmals in einer Woche angeboten werden (z.B. von Frankfurt nach München jeden Werktag um 7.00 Uhr), sollen nach Möglichkeit mit dem gleichen Flugzeugtypen geflogen werden. Eine solche Gleichmäßigkeit ist wünschenswert, damit die Ausführung des Flugplanes für Flug- und Bodenpersonal einfacher wird. Außerdem erhöht sie die Akzeptanz des Flugplanes bei den Kunden. Es hat sich gezeigt, dass sich solche zusätzlichen Bedingungen gut in unseren heuristischen Ansatz integrieren lassen, sodass besser realisierbare Fleet Assignments berechnet werden können. Besonders wichtig für die Praxis ist die Möglichkeit, durch verhältnismäßig kurze Berechnungszeiten, eine Reihe von „Was-wäre-wenn”-Anfragen zu stellen, um z.B. zu evaluieren, inwieweit sich der Gewinn verbessern ließe, wenn ein neues Flugzeug gekauft würde. Bereits ohne diese Anfragen konnte der Gewinn des Fleet Assignment mit unserem Ansatz in Millionenhöhe gesteigert werden (siehe Schlussabschnitt).

Nachdem das Fleet Assignment berechnet wurde, steht fest, dass der Flug von Frankfurt nach Rom mit einem Airbus A 320 geflogen werden soll. Im weiteren Verlauf der Planung stellt sich die Frage nach der Besatzung: »Warum fliegt gerade Ute Schmidt als Pilotin von Frankfurt nach Rom?«

Crew Scheduling – Einsatzplanung der Besatzung

Bei der Zuweisung von Personal an die einzelnen Flüge müssen viele besondere Anforderungen beachtet werden. Dazu gehört eine Vielzahl von gesetzlichen Regelungen zur maximalen Flugzeit oder auch zu Mindestpausen bei Flügen über Zeitzonen. Spezielle Vereinbarungen mit Gewerkschaften oder Vorgaben aus der Firmenphilosophie spielen hier ebenfalls eine Rolle. Individuelle Fähigkeiten, wie Sprachkenntnisse, sind für Kabinenbesatzungen ein wesentliches Auswahlkriterium, wohingegen Qualifikationen und Fluglizenzen für bestimmte Flugzeugtypen im Cockpit-Bereich besonders wichtig sind. Insgesamt können hier durchaus über 100 verschiedene Regeln für die Gestaltung der Dienstpläne (so genannte „Roster“) relevant sein. Neben diesen harten Anforderungen gehen immer mehr Fluggesellschaften dazu über, auch Wünsche der Besatzung stärker in der Planung zu berücksichtigen. Dazu kann das Personal Vorlieben für bestimmte Flugziele benennen, sich für frühe oder späte Abflüge aussprechen oder aber (z.B. im Fall von Ehepaaren) gemeinsame Flüge wählen. Diese Präferenzen sind nicht zwingend für die

Einsatzplanung, werden aber nach Möglichkeit beim Gestalten der Dienstpläne berücksichtigt – schließlich senkt zufriedenes und motiviertes Personal auf lange Sicht die Kosten.
Innerhalb des von der EU geförderten ESPRIT-Projektes PARROT haben wir mit unseren Partnern Carmen Systems (Schweden), ILOG (Frankreich), Lufthansa Systems (Frankfurt), Universität Athen und Olympic Airways (Griechenland) verschiedene Ansätze zur Lösung des Crew-Scheduling-Problems verfolgt.

Lösen des Crew-Scheduling-Problems

Um trotz der hohen inhärenten Komplexität das Crew-Scheduling-Problem hinreichend gut lösen zu können, wird das Problem bei fast allen Fluggesellschaften in Unterprobleme zerlegt:
In einer ersten Phase (Crew Pairing) werden sinnvolle Teilaufgaben zu einem Aufgabenblock zusammengefasst. Typischerweise werden hier Einzelflüge zu einer Kette von Flügen, die an einer Heimatbasis starten und enden, zusammengesetzt. Unterschiedliche Kombinationen der gleichen Aufgaben können dabei durchaus unterschiedliche Kosten für die Fluggesellschaft verursachen. Werden z.B. Teilflüge so ungünstig kombiniert, dass viele gesetzliche Ruhepausen eingehalten werden müssen, kommen unnötig hohe Hotelkosten auf die Fluggesellschaft zu. Aus den erstellten Aufgabenblöcken werden in der nachfolgenden Phase des Crew Rostering für jedes tatsächliche Crew-Mitglied eine Reihe von möglichen Dienstplänen generiert. Diese Dienstpläne genügen allen Regeln und berücksichtigen die Präferenzen des Crew-Mitglieds so weit wie möglich. Schließlich werden nun in der dritten Phase (Crew Assignment) aus der Masse der Dienstpläne für die einzelnen Personen solche gewählt, die kostengünstig sind, gleichzeitig aber auch alle Aufgaben abdecken.
Eine wesentliche Schwierigkeit des Crew Scheduling liegt darin, solche Dienstpläne zu generieren, von denen eine Auswahl die Menge der Flüge genau einmal abdeckt. Probleme entstehen dadurch, dass ein Dienstplan möglichst viele Mitarbeiterwünsche berücksichtigt, und dadurch attraktive Ziele, die von vielen Crew-Mitgliedern präferiert werden, in sehr vielen Dienstplänen enthalten sind. Weniger attraktive Ziele hingegen sind eventuell bei keinem Besatzungsmitglied berücksichtigt worden. Hieraus ergibt sich die Notwendigkeit, die beiden Phasen des Crew Rostering und des Crew Assignment iterativ zu wiederholen und Informationen aus diesen beiden Planungsphasen wechselseitig auszutauschen.
In dem von uns verwendeten Ansatz wird während des Crew Assignment eine neue Bewertung (Dual-Werte) für die Arbeitspakete ermittelt. Vereinfacht ausgedrückt werden durch diese neue Bewertung solche Arbeitspakete verteuert, die in vielen billigen Dienstplänen vorkommen, und solche verbilligt, die kaum oder lediglich in sehr teuren Dienstplänen enthalten sind. Durch diese künstliche Einflussnahme auf die Attraktivität von Arbeitspaketen wird beim Crew Rostering einerseits versucht, solche Einzeldienstpläne zu generieren, die zwar immer noch

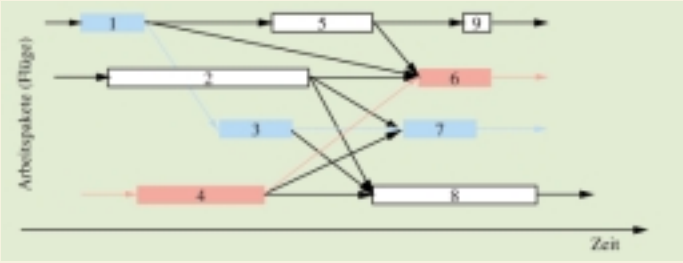


Abb. 4: Der Dienstplan „1, 3, 7" (blau) ist ein Weg durch den Nachfolger-Graphen.

günstig sind, jedoch nicht mehr die Arbeitspakete enthalten, um die mehrere bevorzugte Dienstpläne konkurrieren; und andererseits solche Arbeitspakete, die bisher gemieden wurden, in möglichst günstige Dienstpläne zu integrieren. Mit dieser neuen Bewertung werden in dem nachfolgenden Crew Rostering Einzeldienstpläne generiert, die zu einem kostengünstigeren Gesamtdienstplan führen. Die Idee, anstatt der Menge aller Dienstpläne nur die sukzessiv generierten zu betrachten, basiert auf der Methode des *Column Generation*.
Die Auswahl der Einzeldienstpläne zu einem Gesamtdienstplan erfolgt durch Methoden der *Linearen Programmierung*. Diese liefert auch ein einfaches Kriterium für die Frage, wann die Iteration zwischen Crew Rostering und Crew Assignment beendet werden kann. Sie garantiert nämlich, dass keine weitere Verbesserung mehr erzielt werden kann, wenn die errechnete neue Bewertung der Arbeitspakete bestimmte Eigenschaften nicht mehr erfüllt.

Dienstpläne sind Wege in kostenbewerteten Netzwerken

Die vorgestellte Lösungsmethode erfordert, dass schnell viele billige Einzeldienstpläne bezüglich der aktuellen Bewertung generiert werden können (z.B. die zehn billigsten pro Crew-Mitglied). Jedoch ist dies auf Grund der äußerst komplexen Bedingungen, die an einen gültigen Dienstplan gestellt werden, sehr schwierig.
Um die Generierung von Dienstplänen zu erleichtern, lockert man die Bedingungen an den zu berechnenden Dienstplan. Dafür nimmt man in Kauf, auch ungültige Dienstpläne zu erzeugen. In unserem Ansatz werden alle Einschränkungen für nicht direkt aufeinander folgende Flüge ignoriert und es wird lediglich betrachtet, ob je zwei Flüge legalerweise *direkt* aufeinander folgen können. Dies ermöglicht es, einen sogenannten Nachfolger-Graphen (siehe Abbildung 4) zu konstruieren, in dem jeder Knoten ein Arbeitspaket darstellt. Arbeitspaket *A* wird genau dann mit Arbeitspaket *B* durch eine Kante verbunden, wenn keine Regel das direkte Aufeinanderfolgen dieser Arbeitspakete verbietet. Insbesondere kann jeder legale Dienstplan dann als ein Weg durch das so entstandene Netzwerk angesehen werden. Das Gewicht einer Kante von *A* nach *B* spiegelt die Kosten von Arbeitspakt *A* plus etwaiger Kosten für das Aufeinanderfolgen der beiden Arbeitspakete (z.B. Übernachtungskosten) wider.
Kandidaten für einen kostengünstigen Dienstplan entsprechen kürzesten Wegen durch das Netzwerk, wobei die Kantengewichte

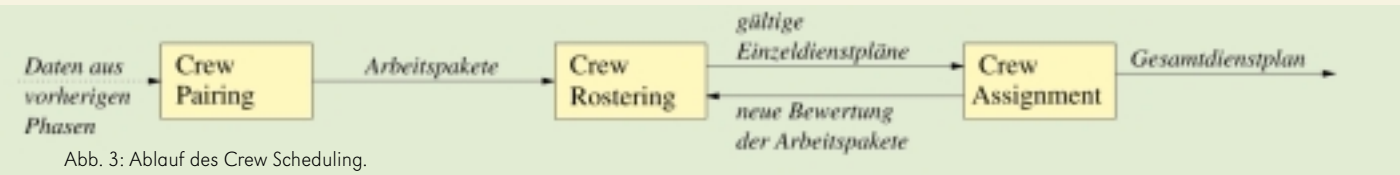


Abb. 3: Ablauf des Crew Scheduling.

die Distanzen zwischen den Knoten des Netzwerkes bestimmen. Für das Kürzeste-Wege-Problem sind sehr effiziente Algorithmen bekannt: Im Wesentlichen brauchen diese Algorithmen eine Laufzeit, die proportional zur Anzahl der Flüge und der Verbindungen zwischen diesen ist. Möglich wurde diese schnelle Methode aber nur dadurch, dass wir einige Regeln gelockert haben und somit auch viele ungültige Dienstpläne erzeugen.
Um möglichst wenig Zeit mit der Generierung ungültiger oder unrentabler Dienstpläne zu verbringen, wird versucht, frühzeitig ungültige und unrentable Lösungsbereiche zu streichen. Das geschieht durch Löschen von Knoten und Kanten im Nachfolger-Graph und durch Festlegen von Knoten, durch die der gesuchte Weg gehen muss. Letzteres entspricht dem Festlegen eines Arbeitspaketes als Bestandteil des Dienstplans. Nehmen wir beispielsweise an, dass Arbeitspaket 1 und 3 zum kürzesten Weg gehören, danach aber eine längere Arbeitspause zwingend notwendig ist, sodass Arbeitspaket 7 nicht nach 1 und 3 ausgeführt werden kann. Dann wird für Wege, die durch 1 und 3 gehen, die Verbindung zu Knoten 7 gelöscht. Um diese Dynamisierung des Nachfolger-Graphen zu formulieren und zur Laufzeit auszurechnen, werden Konzepte des *Constraint Programming* eingesetzt.

Informatik-Methoden als Grundlage für effiziente Planung

Unser Ansatz zum Fleet Assignment wird derzeit mit Erfolg bei einer Fluggesellschaft eingesetzt, die den zu erwartenden Gewinn des Fleet Assignment einer Woche um einige Millionen Dollar steigern und sogar einige Flugzeuge einsparen konnte, während gleichzeitig die Berechnungen einen erheblich geringeren Zeitraum einnahmen als die bisherigen Planungen per Hand. Bei mindestens zwei weiteren Fluggesellschaften wird unser Ansatz in nächster Zeit die Berechnung des Fleet Assignment übernehmen. Die entwickelten Methoden haben nicht nur in der Praxis, sondern auch in der Wissenschaft einen Maßstab gesetzt, konnten damit doch zum ersten Mal Fleet Assignments von sehr guter Qualität in dieser Größe berechnet werden.
Die im Projekt PARROT entwickelten Methoden und Ideen zum Crew Scheduling haben die Integration der ursprünglich getrennten Forschungsgebiete Operations Research und Constraint Programming einen wesentlichen Schritt vorwärts gebracht. Die gewonnenen theoretischen Erkenntnisse konnten mit Erfolg in die Praxis umgesetzt werden. Ihr Einsatz in der Crew-Planungssoftware der schwedischen Firma Carmen Systems konnte z.B. zu einer Reduktion der Rechenzeit von 70 Stunden auf 20 Stunden für einige Referenzprobleme einer grossen europäischen Fluggesellschaft beitragen. Diese Zeitersparnis eröffnet der Fluggesellschaft die Möglichkeit, zwei Tage später mit der endgültigen Crew-Planung zu beginnen, und so noch kurzfristige Änderungen leicht in der Planung zu berücksichtigen. Für andere Gesellschaften ließ sich bei vorgegebener Rechenzeit eine Steigerung der Lösungsqualität (gemessen in Kosten und erfüllten Mitarbeiterpräferenzen) im Bereich von 1 bis 5 Prozent erreichen.
Die beiden Beispiele der Flotten- und Crew-Einsatzplanung zeigen, dass Informatik-Methoden einen wesentlichen Beitrag zur Verbesserung der Flugplanung im Hinblick auf zeitliche, personelle und finanzielle Aspekte leisten.
Die gewonnenen wissenschaftlichen Erkenntnisse sind eine wichtige Grundlage für die weitere Arbeit im Bereich der Flugplanung (Integration von verschiedenen Planungsphasen, Probleme der Flugallianzen). Die Anwendung der entwickelten Lösungsmodelle und -methoden geht weit über die Flugplanung hinaus; sie lassen sich auf andere Gebiete, wie z.B. Planung des Schienenverkehrs oder Einsatzplanung von ambulantem Pflegepersonal, übertragen.

Glossar

BMBF: Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie.

Column Generation: Um Lineare Programme zu verkleinern, werden beim Column-Generation-Ansatz anfangs nur sehr wenige Variablen berücksichtigt und dann nach und nach viel versprechende Variablen ins Lineare Programm hinzugefügt, bis man eine sehr gute oder optimale Lösung gefunden hat. Dabei ist typischerweise die Anzahl der eingefügten Variablen deutlich geringer als die Gesamtzahl der ursprünglichen Variablen. Technisch entspricht die Hinzunahme von Variablen der Generierung und Hinzufügung von Matrixspalten, daher der Name Column Generation.

Constraint Programming: Das zu lösende Problem wird definiert als Menge von Unbekannten und deren Wertebereiche, sowie einer Menge von Bedingungen. Constraint Programming sucht nach einer Belegung der Unbekannten innerhalb der Wertebereiche, welche die gegebenen Bedingungen erfüllt und – bei Optimierungsproblemen – eine zusätzlich gegebene Funktion minimiert bzw. maximiert. Kennzeichnend für das Constraint Programming ist die direkte Weiterleitung von Implikationen von Einschränkungen des Wertebereichs einer Unbekannten auf die Wertebereiche der anderen Unbekannten.

Dual-Werte: In der Linearen Programmierung spiegeln die Dual-Werte den Einfluss einzelner Bedingungen für die Gesamtlösung wider. Sie können z.B. dazu benutzt werden, um systematisch neue verbessernde Lösungen zu finden. Beim Column Generation steuern Dual-Werte in der Regel die Generierung neuer Spalten.

ESPRIT: European Strategic Program for Research in Information Technologies.

Exakte Optimierungsmethoden: Die berechnete Lösung erfüllt alle Bedingungen und hat beweisbar den optimalen Lösungswert unter allen möglichen Lösungen, die ebenfalls alle Bedingungen erfüllen.

Lineare Programmierung: Optimierung linearer Probleme (d.h. die Lösung wird als Vektor dargestellt, wobei die Bedingungen an den Lösungsvektor sowie die Funktion, die es zu minimieren bzw. zu maximieren gilt, linear sind).

NP: Klasse der Probleme, für die man eine Lösung schnell (d.h. in Polynomialzeit, vergleiche Polynomialzeit-Algorithmus) verifizieren kann. Nur für eine Teilklasse von NP weiß man, dass das Finden von Lösungen für diese Probleme ebenfalls in Polynomialzeit gelingt.

NP-hart: Zum Lösen NP-harter Probleme sind keine Polynomialzeit-Algorithmen bekannt. Die Theoretische Informatik klassifiziert diese Probleme als „schwierig"; denn wenn man auch nur für ein NP-hartes Problem einen Polynomialzeit-Algorithmus angeben könnte, dann wären alle Probleme der Klasse NP in Polynomialzeit lösbar.

PARALOR: Parallele Algorithmen zur Lösung großer kombinatorischer Optimierungsprobleme (Forschungsprojekt gefördert vom BMBF). Partner: Lufthansa Systems (Frankfurt), Profi L. (Remscheid), Humboldt Universität Berlin, Universität Köln, Universität Paderborn.

PARROT: Paralleles (Crew) Rostering (EU-ESPRIT-Forschungsprojekt). Partner: ILOG S.A. (Frankreich), Carmen Systems AB (Schweden), Lufthansa Systems (Frankfurt), Olympic Airways (Griechenland), Universität Athen (Griechenland), Universität Paderborn.

Polynomialzeit-Algorithmus: Algorithmus, dessen Laufzeit durch ein Polynom in der Eingabegröße beschränkt werden kann.

Simulated Annealing: Simulated Annealing ist eine Steuerungsheuristik für lokale Suche. Verbessernde Austausche werden immer durchgeführt, verschlechternde nur mit gewisser Wahrscheinlichkeit. Diese Wahrscheinlichkeit nimmt während der Berechnung in Stufen bis zu null ab.

Literatur

Silvia Götz, Sven Grothklags, Georg Kliewer, Stefan Tschöke *Solving the Weekly Fleet Assignment Problem for Large Airlines* Proc. of the Third Metaheuristics International Conference (MIC), S. 241-246, Rio de Janeiro, 1999.

C.A. Hane, C. Barnhart, E.L. Johnson, R.E. Marsten, G.L. Nemhauser, G. Sigismondi *The Fleet Assignment Problem: Solving a Large-Scale Integer Program* Mathematical Programming 70, S. 211-232, 1995.

Torsten Fahle, Ulrich Junker, Stefan E. Karisch, Niklas Kohl, Bo Vaaben, Meinolf Sellmann *Constraint Programming Based Column Generation for Crew Assignment* Journal of Heuristics (to appear).

Meinolf Sellmann, Kyriakos Zervoudakis, Panagiotis Stamatopoulos, Torsten Fahle *Integrating Direct CP Search and CP-based Column Generation for the Airline Crew Assignment Problem.* Proc. of the 2nd intern. Workshop CP-AI-OR, S. 163-171, Paderborn, 2000.

A. Caprara, P. Toth, D. Vigo, M. Fischetti
Modeling and Solving the Crew Rostering Problem
Operations Research, 46(6):820-830, 1998.

G. Yu (editor): *Operations Research in the Airline Industry*
International Series in Operations Research and Management Science, Vol. 9, Kluwer Academic Publishers, 1998.

Ein Teil der Arbeitsgruppe Monien arbeitet im Bereich Optimierung (v.l.): **Dipl.-Inform. Sven Grothklags, Dipl.-Inform. Georg Kliewer, Dipl.-Inform. Meinolf Sellmann, Dipl.-Inform. Torsten Fahle, Prof. Dr. Burkhard Monien, Dipl.-Inform. Silvia Götz und Dipl.-Inform. Stefan Tschöke.**

Fuzzy-Logik

Computergesteuerte Entscheidungstechnik für die Rechtswissenschaft

Das Thema erscheint beim ersten Hinsehen exotisch: Sollte es möglich sein, juristische Entscheidungsfindung an rechnergesteuerte Datenanalyse und damit an eine wenngleich auch hoch technisierte Maschine zu delegieren. Schreckensszenarien wie etwa eine DV-gestützte Strafrechtspflege oder der Mensch als Berechnungsfaktor und Objekt elektronischer Datenverarbeitung drängen sich auf.

Fuzzy-Logik als Entscheidungstechnik?

Bislang ergab sich diese Fragestellung nicht. Zwar hat die elektronische Datenverarbeitung bis zum heutigen Tage riesige Bereiche unseres täglichen Lebens erobert, indem sie mathematische Operationen exakt und mit weniger Zeitaufwand ausführte, als dies dem menschlichen Hirn möglich ist. Zur äußerst komplexen Imitation eines menschlichen Entscheidungsvorganges war die bisherige Technik jedoch nicht in der Lage. Dies liegt an dem Umfang der zu verarbeitenden Daten, die bei der Simulation eines Entscheidungsvorganges mit festen Größen erforderlich sind. Aller Voraussicht nach wird ein solch aufwändiges System – wegen des übermäßigen Aufwandes an Rechnerkapazitäten – niemals realisiert werden können. Auch in der juristischen Methodenlehre haben sich solche Denkansätze nicht durchsetzen können. Jedoch scheinen menschliche Entscheidungsvorgänge eines solchen Aufwandes gar nicht zu bedürfen.
Der Fuzzy-Logik ist es gelungen, menschliche Entscheidungsfindung – vornehmlich in dem Bereich der Technik – überzeugend nachzubilden. Hierin besteht die Chance, Fuzzy-Logik auch auf andere komplexe menschliche Entscheidungsvorgänge anzuwenden.

Die Natur selbst vermeidet – speziell aus zeitökonomischen Gründen – den oben beschriebenen enormen Aufwand der Ermittlung und Verarbeitung unzähliger Einzelwerte bei Entscheidungsvorgängen:

- *Das Max-Planck-Institut für Biologische Kybernetik wies erstmalig nach, dass Stubenfliegen, bei ihrem Landeanflug und zur Kollisionsvermeidung nicht etwa einen optimalen Bremszeitpunkt aus unzähligen „festen" Daten „errechnen". Vielmehr „schließen" Stubenfliegen – in einem fortlaufenden Prozess – aus nur wenigen, sich verändernden Landereizen auf den „geeignetsten" Zeitpunkt für den Beginn und die Durchführung des Landevorganges bzw. des Ausweichens. Ein solches „Schließen" mittels Kombination diverser (und wegen deren Vielzahl und Unterschiedlichkeit) nicht notwendig genauer Einzelinformationen entspricht dem Vorgang der Fuzzy-Logik.*

Prof. Dr. jur. Dieter Krimphove ist Professor im Fachbereich 5/Wirtschaftswissenschaften an der Universität Paderborn. Seine Schwerpunkte sind Wirtschaftsrecht, Europäisches Wirtschaftsrecht, Bankrecht, Arbeitsrecht und Unternehmensrecht.

Auch das menschliche Entscheidungsverhalten – gerade das des Alltages – funktioniert nicht anders. Auch hier kommt man, statt mit einer Vielzahl von kompliziert zu ermittelnden und zu berechnenden „festen" Einzelwerten", mit nur wenigen „unscharfen" Werten aus. Die Reduktion auf nur wenige Werte und die damit verbundene erhebliche Vereinfachung der Entscheidungsfindung gelingt durch die Einteilung der vielen Einzelgrößen in „Entscheidungsmengen".

Beispiele:

- *Ob ein Mensch die Temperatur in einem mit 48,86° oder 47° temperierten Raum als zu heiß empfindet, ist vollständig irrelevant für seine Entscheidung, den Raum zu verlassen bzw. das Fenster zu öffnen. Genaue Messdaten werden hier nicht verlangt.*

- *Gleichgültig für eine Kaufentscheidung ist, ob der Käufer den Preis für ein Kunstwerk der Moderne von 150 Mark oder 1 300 Mark für günstig, da zu niedrig, hält. Auch in diesem Beispiel rechtfertigt allein die – wenn auch unscharfe – Tatsache „Preis ist sehr niedrig" die Kaufentscheidung.*

Gewöhnlich arbeitet der menschliche Intellekt mit sprachlich unscharfen Begriffen bzw. Begriffsmengen, wie „sehr niedrig", „geeignet", „vollständig geeignet", „preiswert", „genügsam" etc. Speziell unter zeitlichem Entscheidungsdruck ermöglicht diese Vereinfachung eine Entscheidungsfindung überhaupt.

Genau an diesem Entscheidungsmuster setzt die sogenannte Fuzzy-Logik an. Es geht der Fuzzy-Logik also nicht um die Bearbeitung „fester Ergebnisse" *(Singletons)*, welche sich auf eine Vielzahl „fester" Daten ableiten lassen: Die Fuzzy-Logik bemüht sich vielmehr – wie ihr Name schon besagt – um Operationen auf „unscharfen linguistischen Begriffen oder (Begriffs-)

Mengen“. Damit ist Fuzzy-Logik in der Lage, gerade die unscharfen Begriffe der Sprache („teuer“, „preiswert“, „geeignet“ etc.) für ein Erkenntnisverfahren nutzbar zu machen; ein Aspekt, der für die Handhabung „unbestimmter Rechtsbegriffe“ von großem Interesse ist.

Vorteile sprachlicher Ungenauigkeit bei der Entscheidungsfindung

Sprachlicher Ungenauigkeit bedient sich der Mensch bei seinen Entscheidungen Zwangsläufig.
Denn:

- Erstens kommt es auf eine zahlenmäßig, rechnerische Differenzierung – wie die obigen Beispiele demonstrieren – überhaupt nicht an.
- Zweitens – und mit dem ersten Gesichtspunkt eng zusammenhängend – ist der notwendige immense Aufwand im Rahmen der mathematischen Überprüfung und Nachbildung menschlicher Entscheidung wegen deren Subjektivität kaum sinnvoll.
- Oder nur unter extrem großen Aufwand darstellbar und somit unpraktikabel.

Die auf den ersten Blick als Paradox anmutende Notwendigkeit (sprachlicher) Unschärfe bei menschlichen Entscheidungsvorgängen erscheint bei genauem Hinsehen vorteilhaft: Denn nur die Unschärfe erzeugt leicht und rasch zu überblickende Urteile.

Traditionelle Einsatzgebiete der Fuzzy-Logik

Bisher fanden Fuzzy-Logik-gesteuerte Entscheidungsvorgänge Anwendung vor allem in der Überwachung und Steuerung chemischer Produktionsprozesse, Müllverbrennungsanlagen, dem Brauen von Reiswein, der Steuerung von Hochhausaufzügen, U-Bahn-Netzen, sowie der Mustererkennung: Das Spektrum reicht hier von dem Erkennen von Umweltverschmutzungen durch bestimmte Chemikalien, über die Zahlen und Schrift(en)erkennung, die Klassifizierung von Objekten, insbesondere Produktionsüberwachung, Werkzeugüberwachung, medizinischen Diagnosesystemen (z.B. EKG-Analyse-Systeme), bis hin zu Techniken der Hindernisumgehung selbstgesteuerter mobiler Industrieroboter.

Trotz der Vielfalt der oben aufgeführten Anwendungsbereiche der Fuzzy-Logik sind deren Einsatzgebiete noch nicht ausgeschöpft:
Die Entwicklung der Einsatzgebiete der Fuzzy-Logik macht nicht im technischen Umfeld halt. Auch bei der Entscheidungsfindung im sozialen Bereich können Fuzzy-Logik gesteuerte Rechnersysteme Anwendung finden: Derzeit befassen sich Systeme mit so komplexen Problemstellungen wie der Auswahlentscheidung zwischen mehreren Kaufobjekten mit einer Anzahl jeweils unterschiedlich realisierter Qualitätsmerkmale (kompensatorische Entscheidungen). Fuzzy-Logik ist ansatzweise ebenfalls im Einsatz bei der Analyse der Kreditwürdigkeit eines Kunden und der Ermittlung, Einschätzung und Prognose von Wertpapieren (Wertpapiermanagement).
Ein weites Feld nimmt aktuell die Erforschung von Fuzzy-Logik-Systemen zur „Risikoprüfung“ von zivilrechtlichen Verträgen, hauptsächlich zum Zweck der Minimierung des finanziellen Risikos der vertragschließenden Parteien ein. Die „Prüfung und Beurteilung“ des Risikos eines Schadenseintrittes steht im besonderen Interesse staatlicher wie privater Versicherungsunternehmen. Schließlich beurteilt sich anhand des „Schadensrisikos“ die Höhe der Versicherungsprämie oder sogar die Frage nach dem Abschluss eines Versicherungsvertrages überhaupt.

Wie funktioniert Fuzzy-Logik

Die Idee der Fuzzy-Logik ist ebenso einfach wie plausibel: Statt einer Unzahl langwierig zu ermittelnden fester Einzelwerte *(Singletons)* und deren aufwändige Berechnung zu einem präzisen Ergebnis, verwendet die Fuzzy-Logik unscharfe linguistische Werte (wie: „sehr hoch“, „hoch“, „mittel“, „niedrig“, „sehr niedrig“) in so genannten Fuzzy-Mengen. Auf diese Weise entstehen „Band- oder Spannbreiten“ (Mengen) in denen der Realisations- oder Gegebenheitsgrad einer bestimmten Qualität – in mehr oder weniger hohem Maße – zutrifft.
Mit der Aneinanderreihung einzelner unscharfer Begriffsmengen (wie sehr hoch, hoch, mittel, niedrig, sehr niedrig) erreicht die Fuzzy-Logik mit nur wenig Aufwand ein Kontinuum der Begrifflichkeit. Dieses Kontinuum macht die Zuordnung eines Objektes zu einer bestimmten Qualität graduell möglich, ohne auf ein Vielfaches an Einzelwerten angewiesen zu sein.

Fuzzy-Logik gewährleistet nicht nur die Handhabung unscharfer linguistischer Begriffe. Mithilfe von Fuzzy-Aussagen lassen sie sich auch zu logischen Sätzen verknüpfen:
Diese logischen Sätze ordnen einen oder mehrere unbestimmte(n) Eingangswert(e) oder eine, bzw. mehrere Eingangsmenge(n) einer bestimmten Konsequenz [Ausgangswert/-menge] zu. WENN die (unbestimmte) Eingangsmenge (z.B.: zu „hohe“ Temperatur) erfüllt ist, DANN ist die Ausgangsmenge (Drosseln der Energiezufuhr) zu aktivieren.

Übertragung der Fuzzy-Logik auf die juristische Entscheidungstechnik

Nicht nur im technischen Bereich findet die Fuzzy-Logik Anwendung. Auf Grund ihrer Entsprechung zu menschlichen Entscheidungsabläufen ist Fuzzy-Logik auch auf die juristische Entscheidungsfindung übertragbar.

Die Übertragung der Fuzzy-Logik auf juristische Sachverhalte rechtfertigen im Wesentlichen zwei Gründe:

- Erstens ist Gegenstand der Fuzzy-Logik die „Aufbereitung“ unscharfer Begriffe für einen Entscheidungsvorgang. Exakt dieselbe Problematik beschäftigt die juristische Methodik der Rechtsfindung. Auch hier geht es um die Verwendung unbestimmter oder/und unscharfer Rechtsbegriffe zum Zweck der Entscheidungsfindung. Der Anwendungsbereich der Fuzzy-Logik beschränkt sich nicht nur allgemein auf linguistisch unscharfe Begriffe. Vorstellbar ist ebenfalls die Anwendung der Fuzzy-Logik speziell auf die „unbestimmten Rechtsbegriffe“ in der Rechtswissenschaft. Die gesamte *juristische Entscheidungsfindung* operiert fast immer mit mehrdeutigen, auslegungsbedürftigen und damit unbestimmten Rechtsbegriffen.

  Der Fuzzy-Logik ist somit im Bereich juristischer Entscheidungsfindung grundsätzlich ein weites Feld eröffnet.

- Der wesentliche Rechtfertigungsgrund der Nutzbarmachung des „Fuzzy-Logik-Schließens“ für den juristischen Bereich besteht – zweitens – in der für die Fuzzy-Logik typischen Entscheidungsstruktur selbst.
  Die logische Verknüpfung der Fuzzy-Logik mit „WENN – DANN“ entspricht exakt der Gliederung gesetzlicher Tatbestände in (diverse) TATBESTANDSMERKMAL(E) und einer RECHTSFOLGE.

Dabei korrespondieren die rechtlichen Tatbestandsmerkmale der Eingangsgröße(n) der Fuzzy-Logik. Die mittels der Auswertung der Tatbestandsmerkmale gefundene Rechtsfolge stimmt mit der Fuzzy-Logik-Ausgangsgröße überein.

Eingangsgrößen	Ausgangsgröße
WENN – das unbestimmte Merkmal A und – das unbestimmte Merkmal B und – das unbestimmte Merkmal C und – das unbestimmte Merkmal N usw. vorliegen	DANN – ist (Rechtsfolge der) Norm gegeben

Speziell die Möglichkeit der Verarbeitung mehrerer Eingangsgrößen durch die Fuzzy-Logik macht ihren Einsatz auf dem juristischen Gebiet interessant:

- Anwendungsbeispiel der Fuzzy-Logik – Die Beurteilung einer „überragenden Marktstellung“ i.S.d. § 19, Abs. 2, Nr. 2, GWB.

Speziell im Wettbewerbs- und Kartellrecht kann der Einsatz der Fuzzy-Logik umfassende schwierige langwierige – und in der Entscheidungspraxis aufwändige Unternehmens- und Marktrecherchen ersparen. Unbestimmte Rechtsbegriffe – wie „wesentlicher Wettbewerb“ z.B.: (§ 19, Abs. 2, Nr. 1, GWB), marktbeherrschende Unternehmensstellung (z.B.: § 19, § 36, GWB), missbräuchliches Ausnutzen einer marktbeherrschenden Stellung (§ 19, Abs. 4, GWB), „unbillige Marktbehinderung“ (§ 19, Abs. 4, GWB), Geeignetheit zur wesentlichen Steigerung der Wirtschaftlichkeit“ (§ 5, Abs. 1, GWB), u.a. – erschweren die Subsumtion wirtschaftlicher Sachverhalte und führen so zu einer ungewollten Verzögerung behördlichen Eingreifens.
Die wünschenswerte Vereinfachung und Beschleunigung wettbewerbsrechtlicher Entscheidungsvorgänge könnte durch die Anwendung der Fuzzy-Logik deren positive Folge sein. Dies sei am Beispiel der Ermittlung des Merkmales der „überragenden Marktstellung“ durch ein Unternehmen i.S.d. § 19, Abs. 2, Nr. 1, GWB dargestellt (Die Fuzzy-Logik-Entscheidungsstruktur zur Feststellung einer „überragenden Marktstellung“ i.S.d. § 19, Abs. 2, Nr. 2, GWB):
Der Gesetzgeber und die Rechtsprechung haben den Faktor „überragende Marktstellung“ durch eine Zahl weiterer „Einzelkriterien“ (wie: Marktanteil, bzw. Marktstruktur, Finanzkraft des Unternehmens, Zugang zu Beschaffungs- und Absatzmärkten, Grad der Unternehmensverflechtung, Forschungsvorsprung, Zugang zu Subventionen, wirtschaftlicher Abstand zum nächstfolgenden Wettbewerber u.a.) umrissen.
Die Fuzzy-Logik-Entscheidungsstruktur zur Feststellung einer „überragenden Marktstellung“ hat folgendes Bild:
In der Praxis der Kartellgerichte und -behörden nehmen die Kriterien des Marktanteiles bzw. der Finanzkraft des Unternehmens vorrangige Stellung bei der Ermittlung einer „überragenden Marktstellung“ des Unternehmens ein. Aus Gründen der Übersichtlichkeit verkürzt die Darstellung zur Analyse des Tatbestandsmerkmales der „marktbeherrschenden Unternehmensstellung“ auf diese beiden Kriterien.

Eingangsgrößen	Ausgangsgröße
WENN – hoher Marktanteil bzw. eine eingeschränkte/oligopolistische Marktstruktur, – große Finanzkraft des Unternehmens, – vorrangiger Zugang zu Beschaffungs- und Absatzmärkten, – hoher Grad seiner Unternehmensverflechtung, – Vorsprung des Unternehmens – z.B. in der Forschung oder – beim Zugang zu Subventionen – großer wirtschaftlicher Abstand zum nächstfolgenden Wettbewerber – u.a.	DANN – überragende Marktstellung

Erste Schlussfolgerungen anhand der „Fuzzy-Logik-Regelbasis“

Es ist unmittelbar einsichtig, dass etwa bei überragender Finanzkraft eines Unternehmens auf einem engen oligopolistischen Markt – also einem Markt mit wenigen relevanten Konkurrenten, welche zudem noch in großer wettbewerblicher Reaktionsverbundenheit stehen – dieses Unternehmen eine absolut bzw. „total“ überragende Marktstellung innehat.
Ebenfalls leuchtet ein, dass auf einem Markt mit „einer Menge“ unterschiedlicher Anbieter und Konkurrenten (etwa einem Polypol) – bei „kaum vorhandener“ Finanzkraft des Unternehmens – keine überragende Marktstellung gegeben sein kann.
Erhöht sich die Finanzkraft des Unternehmens, bei gleicher (polypolistischer) Marktstruktur, auf einen mittleren Wert, so besteht, auch in dieser Situation, „kaum“ eine überragende Marktstellung.
Diese Zusammenhänge gibt Abbildung 1 wieder:

- Nach der Regelbasis in Abbildung 1 ist von dem Vorliegen einer „überragenden Marktstellung“ immer dann auszugehen, wenn der Wert „total“ **(TO)** erreicht ist.
- Bei den Werten „keine“ **(KN)**, „kaum“ **(KM)** und „gering“ **(GR)** „überragende Marktstellung“ ist eine „Überragende Marktstellung i.S.d. § § 19, Abs. 2, Satz 2, GWB zu verneinen.
- In Fällen des „mittleren“ Grades einer „überragenden Marktstellung“ **(MT)** ist daran zu denken, deren Vorliegen zusätzlich mittels *weiterer* Beurteilungskriterien (z.B.: *vorrangiger Zugang zu Beschaffungs- und Absatzmärkten, hoher Grad seiner Unternehmensverflechtung, Vorsprung des Unternehmens z.B. in der Forschung oder beim Zugang zu Subventionen, großer wirtschaftlicher Abstand zum nächst folgenden Wettbewerber, übermäßiger Verhaltensspielraum, Koordination des Marktverhaltens der Anschlussunternehmen, insbesondere Gleichpreisigkeit der Anschlussunternehmen, Umstellungsflexibilität und das Ausweichverhalten der Marktgegenseite auf andere Anbieter u.a.)* zu prüfen.
- Die Feststellung einer „überragenden Marktstellung“ bei Zwischenwerten der Eingangsgrößen „Marktteilnehmer“ und „Finanzkraft“ des Unternehmens:

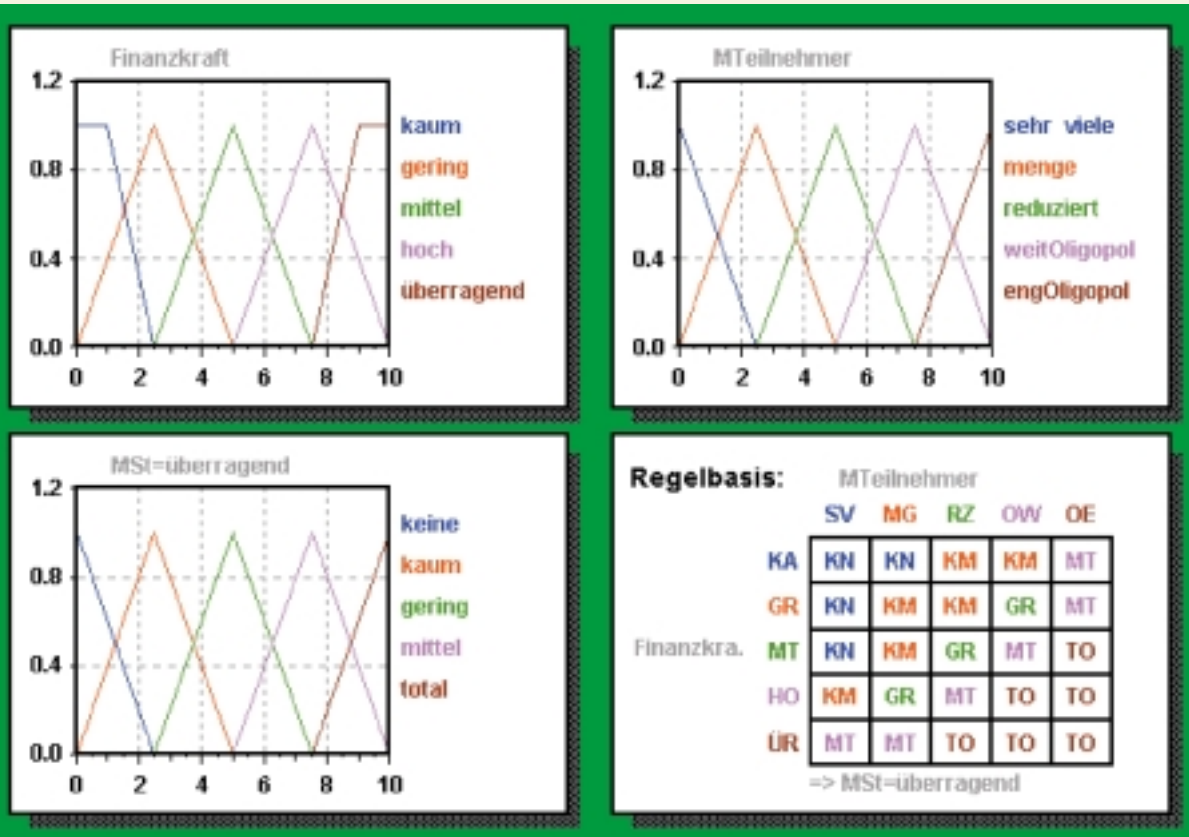
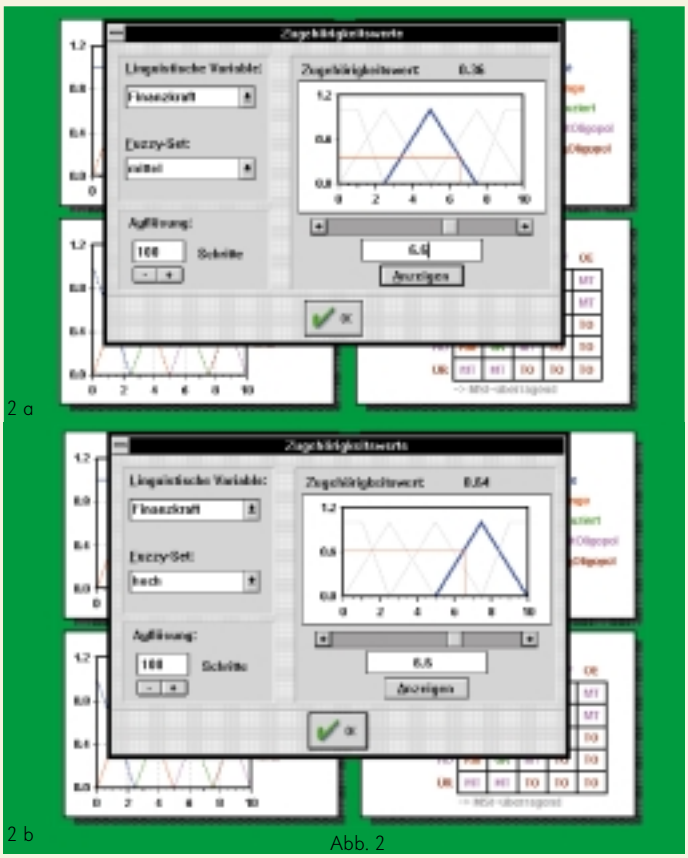


Abb. 1

Häufig sind die jeweiligen Eingangswerte (gering, mittel, hoch usw.) nicht gänzlich, d.h. nicht zu einem Gegebenheitsgrad von 1,0, sondern in unterschiedlichen Maßen erfüllt. Zwischenwerte, die beispielsweise die Finanzkraft des Unternehmens zwischen den Fuzzy-Logik-Mengen **„mittel“** und **„hoch“** einordnen, kommen in der Praxis überwiegend vor. Beispielsweise kann eine bestimmte Finanzkraft (hier 6,6 Einheiten) der Fuzzy-Logik-Menge **„mittel“** (in Abbildung 2a dargestellt durch das Dreieck mit der Basis 2,5 bis 7,5) zu einem Gegebenheitsgrad von **0,36** und gleichzeitig der Fuzzy-Logik-Menge **„hoch“** (in Abbildung 2b dargestellt durch das Dreieck mit der Basis 5 bis 10) zu einem Grad von **0,64** zugeordnet sein.



2 a

2 b

Abb. 2

Zwischenwerte erzeugen, durch jeweilige Übertragung der Höhe ihres jeweiligen Zugehörigkeitsgrades (0,36 „mittel“ und 0,64 „hoch“) in die entsprechende Fuzzy-Logik-Ausgangsmenge eine, i.d.R. unsymmetrische, Fuzzy-Logik-Entscheidungsmenge.

Diese (unsymmetrische) Entscheidungsmenge (hier: „überragende Marktstellung“) ist dargestellt durch die abgetragenen Höhen der sich überschneidenden Eingangsmengendreiecke (Abbildung 3).

Aus dieser unsymmetrischen Entscheidungs-Ausgangsmenge (hier: blaues Feld) ist ein einziger fester Ausgangswert zu ermitteln (bzw. zu defuzzifizieren). Dies geschieht mithilfe der Schwerpunktmethode **(Center of Gravity)**. Der Schwerpunkt der Entscheidungs-Ausgangsmenge (hier: gelbe senkrechte Linie) liegt im vorliegenden Beispiel bei einem Wert von 6,53455. Dieser Wert ist dann wieder in eine „handhabbare Größe“ (z.B.: „kaum“, „geringe“, „mittlere“, „starke“ Marktstellung) zu überführen: Der Wert 6,53455 ergibt, dass bei Vorliegen einer hohen **Finanzkraft (HO)** des Unternehmens – zu einem Gegebenheitsgrad von 0.64 – und bei einer reduzierten Anzahl von Konkurrenten **(RZ)** auf dem Markt – zu einem Gegebenheitsgrad von 1,0 – eine mittel-überragende Marktstellung des Unternehmens **(MT)** – zu einem Gegebenheitsgrad von 0,64 – vorliegt.

Nach den oben dargelegten normativen Parametern unserer Fuzzy-Logik-Untersuchung bedeutet dies, dass § 19, Abs. 1, Satz 2, GWB nicht vorliegt, und somit auch eine „marktbeherrschende Stellung“ dieses Unternehmens nicht festgestellt werden kann.

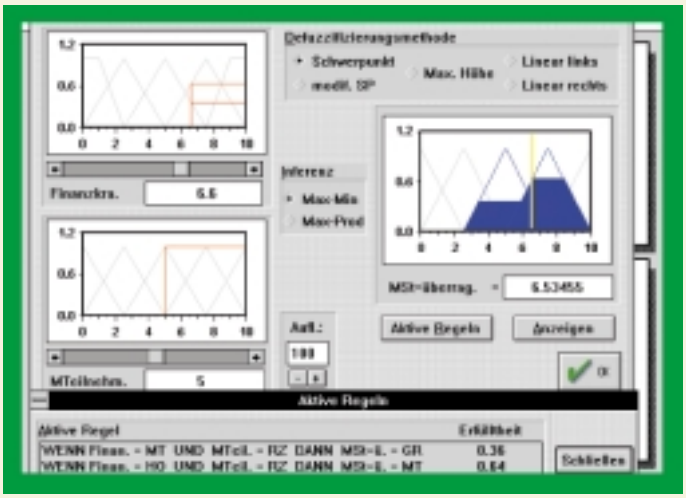


Abb. 3

Mithilfe dieser Vorgehensweise lassen sich zahllose Marktsituationen simulieren:

- Verringert sich – bei gleich bleibender **Finanzkraft (HO)** des Unternehmens – die Zahl der Konkurrenten auf dem betreffenden Markt um ein Viertel (hier ausgedrückt mit 7,5) sodass die **Marktstruktur** eines *weiten Oligopols* **(WO)** – zu einem Gegebenheitsgrad von 1,0 – vorliegt, ist – zu einem Gegebenheitsgrad von 0,64 – von einer „totalen **überragende Marktstellung**“ des Unternehmens **(TO)** auszugehen.

Das Unternehmen besitzt auf diesem Markt folglich eine „überragende Stellung“ i.S.d. § § 19, Abs. 2, Satz 2, GWB.

Auch Zwischenwerte für die Zahl der Marktteilnehmer lassen sich nach obiger Vorgehensweise problemlos analysieren (Abbildung 4):

- Bei einer mittleren **Finanzkraft (MT)** des Unternehmens mit einer großen Anzahl (Menge) an **Marktteilnehmern (MG)** besteht kaum **(KM)** eine überragende Unternehmensstellung. Der Tatbestand des § § 19, Abs. 2, Satz 2 ist daher nicht erfüllt.

Auf diese Weise lässt sich – mithilfe der Fuzzy-Logik – ein Kontinuum sämtlicher Konstellationen das Vorliegen einer „überragenden Marktstellung“ i.S.d. § § 19, Abs. 2, Satz 2, GWB überprüfen (Abbildung 5).

Der Bestimmung einer „überragenden Marktstellung“ durch die Fuzzy-Logik ist auch – in derselben, oben geschilderten Weise –

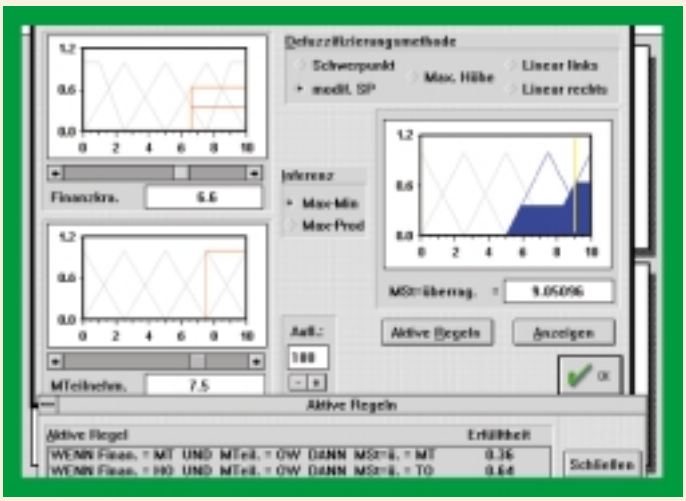


Abb. 4

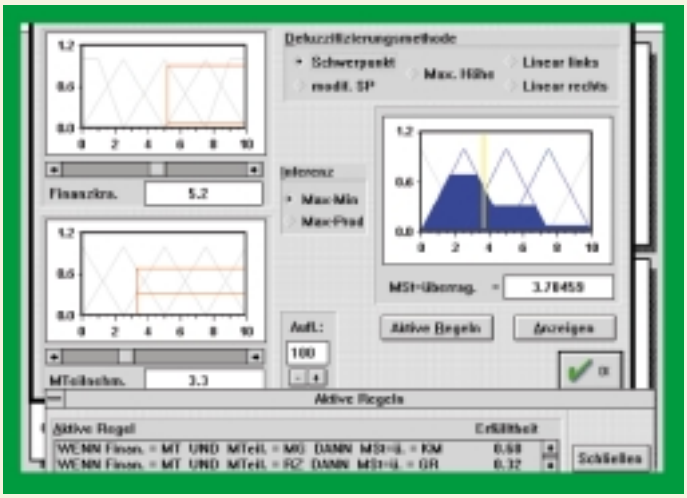


Abb. 5

der Einfluss mehrerer als nur der zwei Eingangsgrößen „Finanzkraft“ und „Marktteilnehmer“ zugänglich.

Die obigen Ausführungen erläutern hinreichend die Einsatzfähigkeit und Anwendungsmöglichkeit der Fuzzy-Logik in der Rechtswissenschaft, ohne dass es weiterer, rein quantitativer Ergänzungen, bedarf. Aus Platzgründen wurde daher auf die entsprechende Darstellung weiterer Konstellationen verzichtet.

Fazit

Trotz bestehender Anwendungshindernisse der Fuzzy-Logik als universale Entscheidungstechnik der rechtswissenschaftlichen Praxis (Einzelheiten: Krimphove: der Einsatz der Fuzzy-Logik in der Rechtswissenschaft, in Rechtstheorie, Heft 3, 1999), kommt dem Einsatz von Fuzzy-Logik in juristischen Entscheidungen – aber neben den „klassischen“ rechtswissenschaftlichen Methoden, als deren Ergänzung – große Bedeutung zu:

- So kann Fuzzy-Logik sinnvoll zur Überprüfung juristischer Entscheidungen herangezogen werden.
- Die Überprüfung juristischer Entscheidungen mithilfe des konstant gleich bleibenden Instrumentariums der Fuzzy-Logik befriedigt nicht nur die Bedürfnisse zur Gleichmäßigkeit juristischer Entscheidungen.
- Ebenfalls eröffnet die – ergänzende – Anwendung von Fuzzy-Logik im rechtswissenschaftlichen Bereich das Auffinden neuer Beurteilungsaspekte und das kritische Hinterfragen der Ergebnisse klassischer juristischer Methodik und Entscheidungsfindung.

Allerdings bedarf es zu einer solchen Nutzbarmachung von Fuzzy-Logik der Aufgeschlossenheit des Juristen gegenüber neuen, innovativen Methoden.

Sofern diese interdisziplinäre Aufgeschlossenheit nicht ohnehin schon als selbstverständlich in der Rechtswissenschaft begriffen wird, ist es Ziel und erklärter Gegenstand vorliegender Untersuchung, hierfür zu werben.

Umsonst durchs All: GENESIS startet 2001

Mathematische Methoden in der Raumfahrt

Auf Grund notwendig gewordener finanzieller Einsparungen wurde die Weltraumbehörde NASA (National Aeronautics and Space Administration) in der jüngsten Vergangenheit immer mehr dazu gezwungen, Raumfahrtmissionen besonders kostengünstig zu gestalten. In diesem Zusammenhang werden jetzt auch aktuelle Ergebnisse der Angewandten Mathematik eingesetzt, die es ermöglichen, energetisch besonders günstige Flugbahnen für Raumfahrzeuge zu bestimmen, um so den Treibstoffverbrauch und damit das Gewicht der Sonden stark zu reduzieren. Im Folgenden wird ein Einblick in die zu Grunde liegenden mathematischen Methoden gegeben, die zum ersten Mal von Mitarbeitern des NASA Jet Propulsion Laboratory (JPL) im Rahmen der Planung der Mission GENESIS – Start ist im Jahr 2001 – eingesetzt wurden. Die Zusammenarbeit zwischen dem JPL und der Angewandten Mathematik an der Universität Paderborn wird zurzeit durch die NSF (National Science Foundation, USA) und den DAAD (Deutscher Akademischer Austauschdienst) unterstützt.

Prof. Dr. Michael Dellnitz ist seit 1998 Professor für Angewandte Mathematik im Fachbereich 17/Mathematik – Informatik der Universität Paderborn. Forschungsschwerpunkte liegen in den Bereichen Numerische Mathematik, Dynamische Systeme und Optimierung. Im Vordergrund stehen interdisziplinäre Fragestellungen aus den Ingenieur- und Naturwissenschaften.

Flugbahnen im Drei-Körper-Problem

Für das Auffinden geeigneter Flugbahnen für Raumfahrzeuge wäre es viel zu aufwändig, die Bewegung sämtlicher Himmelskörper im Sonnensystem mit Hilfe eines Computerprogramms zu simulieren. Daher beschränkt man sich bei den numerischen Explorationen zunächst auf ein für die jeweilige Mission angemessenes Drei-Körper-Problem. Hierbei handelt es sich um ein mechanisches System, welches die zeitliche Bewegung dreier Massen unter dem gegenseitigen Einfluss ihrer Gravitationskräfte beschreibt. Für die Entwicklung von mondnahen Raumfahrtmissionen ist z.B. die Betrachtung der drei Körper Erde, Mond und Raumfahrzeug adäquat. Im Vergleich zu Erde und Mond ist die Masse des Raumfahrzeugs vernachlässigbar und daher wird vereinfachend angenommen, dass das Raumfahrzeug keine Masse besitzt und sich Erde und Mond auf kreisförmigen Bahnen um ihren gemeinsamen Schwerpunkt bewegen.

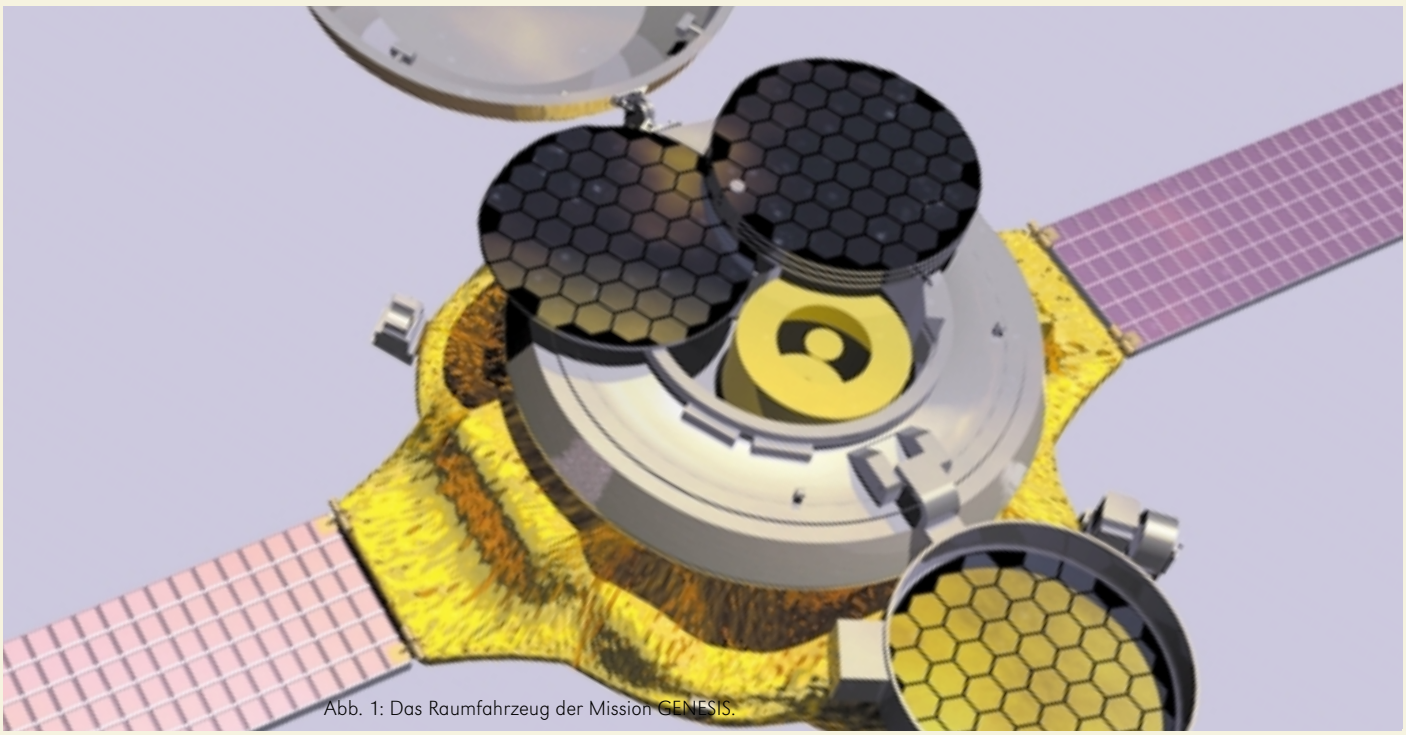


Abb. 1: Das Raumfahrzeug der Mission GENESIS.

Abb. 2: Henri Poincaré, 1854-1912.

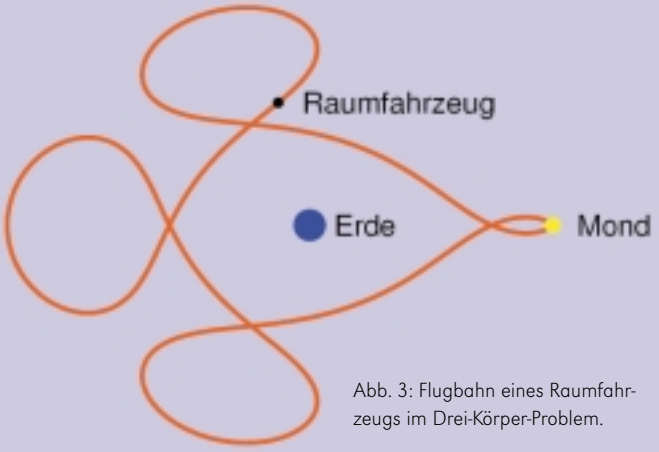


Abb. 3: Flugbahn eines Raumfahrzeugs im Drei-Körper-Problem.

Der französische Mathematiker Henri Poincaré (Abbildung 2) hat bereits am Ende des 19. Jahrhunderts gezeigt, dass es prinzipiell unmöglich ist, das Drei-Körper-Problem analytisch – d.h. mit Papier und Bleistift – zu lösen und sämtliche auftretenden Bewegungen des masselosen Körpers anzugeben. Heutzutage kann man jedoch mit Hilfe numerischer Verfahren Lösungen des Systems sehr genau ermitteln. Eine mögliche Bewegung eines Raumfahrzeugs im Drei-Körper-Problem Erde/Mond/Raumfahrzeug ist in Abbildung 3 dargestellt. Die gezeigte Flugbahn ist geschlossen, so dass sich ein auf dieser Bahn einmal befindendes Raumfahrzeug ohne weitere Energiezufuhr für immer entlang der Bahn bewegen würde.

Bewegungen eines Pendels

Wir gehen nun der Frage nach, wie man energetisch günstige Flugbahnen im Drei-Körper-Problem auffinden kann, die gewisse Orte im Sonnensystem miteinander verbinden. Hierzu betrachten wir zunächst ein anderes, sehr einfaches mechanisches System, an dem die zu Grunde liegende Idee gut illustriert werden kann: wir analysieren die Bewegungen eines mathematischen Pendels, wie es in Abbildung 4 dargestellt ist. Es wird angenommen, dass eine Masse an einem starren, jedoch masselosen Stab befestigt ist und dass das Pendel sich reibungsfrei bewegt. Sieht man einmal von den beiden vertikalen Ruhelagen ab, so kann ein solches Pendel drei prinzipiell verschiedene Bewegungen ausführen:

(i) Ist die Auslenkung des Pendels nicht zu groß und seine Geschwindigkeit klein genug, so wird das Pendel um seine Ruhelage periodisch hin- und herschwingen. Eine derartige Bewegung ist im Phasendiagramm in Abbildung 5 dargestellt. Dort sind auch die stabile und instabile Ruhelage des Pendels eingezeichnet, die dem vertikal nach unten bzw. nach oben gerichteten Pendel entsprechen. Die instabile Ruhelage (senkrecht nach oben) ist in der Abbildung durch zwei Punkte repräsentiert, da bei einer Umdrehung des Pendels der Winkel um 2π anwächst.

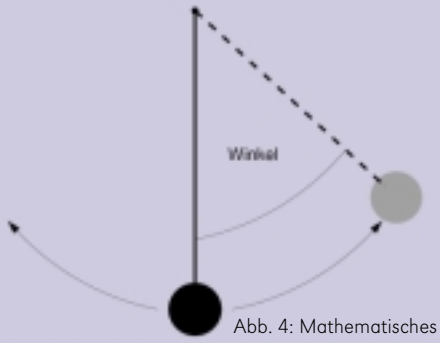


Abb. 4: Mathematisches Pendel.

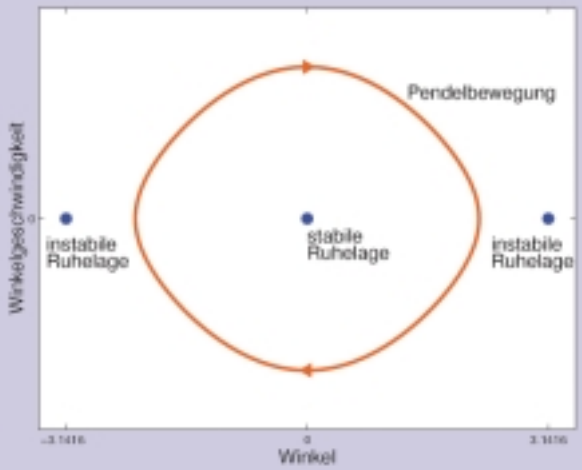


Abb. 5: Pendelbewegung als geschlossene Kurve im Phasendiagramm.

Abb. 6: Das Pendel überschlägt sich.

(ii) Ist die Geschwindigkeit des Pendels sehr hoch, so wird es sich überschlagen. Die entsprechende Bahn im Phasendiagramm ist in Abbildung 6 dargestellt.

(iii) Neben den in (i) und (ii) beschriebenen Fällen existiert nun noch eine weitere Bewegung, die gewissermaßen den „Grenzfall“ zwischen der Pendel- und der Überschlagbewegung darstellt. Stellen wir uns vor, dass das Pendel in seiner instabilen Ruhelage (senkrecht nach oben) verharrt. Geben wir dem Pendel jetzt einen „beliebig kleinen“ Stoß, so fällt es und wird sich am Ende der Bewegung wieder (von der anderen Seite) der instabilen Ruhelage annähern; vergleiche Abbildung 8.

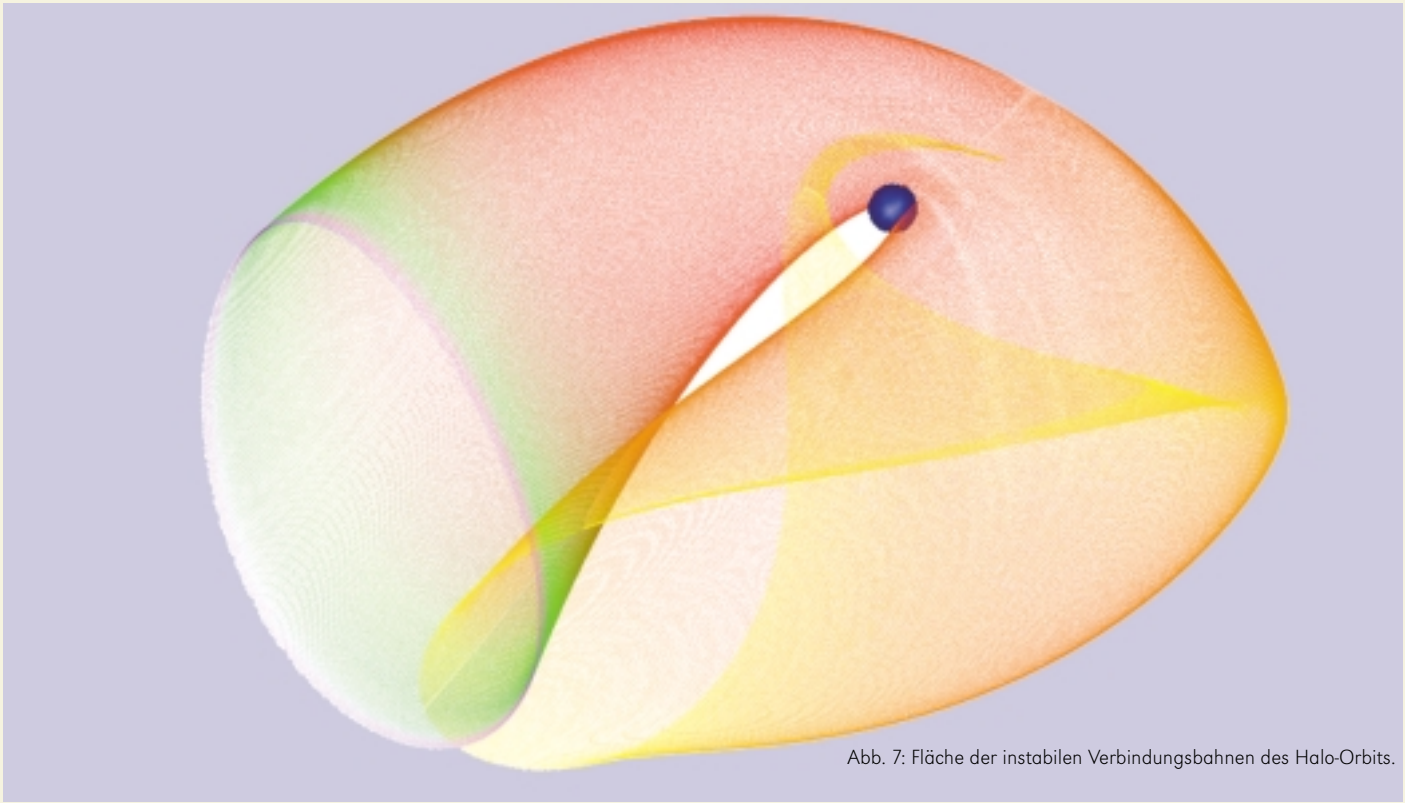


Abb. 7: Fläche der instabilen Verbindungsbahnen des Halo-Orbits.

Verbindungsbahnen im Pendel

Das in (iii) beschriebene dynamische Verhalten des Pendels ist dasjenige, welches für das Auffinden energetisch günstiger Flugbahnen von besonderem Interesse ist. Um dies einzusehen, stellen wir uns vor, ein vertikal aufgerichtetes Pendel mit möglichst wenig energetischem Aufwand aus seiner instabilen Position entfernen und später wieder in diese Position zurückbringen zu wollen. Die Lösung dieses Problems ist genau die in (iii) beschriebene Bewegung, welche die instabile Ruhelage des Pendels durch eine beliebig kleine Energiezufuhr mit sich selbst verbindet. Entsprechende Bahnen in mechanischen Systemen, die vorgegebene Positionen ohne zusätzlichen energetischen Aufwand asymptotisch miteinander verbinden, bezeichnen wir im Folgenden als Verbindungsbahnen. Gäbe es nun auch im Drei-Körper-Problem Verbindungsbahnen zwischen verschiedenen Zuständen des Systems, so könnten wir diese, analog zum Pendel, als energetisch besonders günstige Flugbahnen für das Raumfahrzeug nutzen.

Die Mission GENESIS

Tatsächlich wurde diese Idee erstmals vom JPL zur Entwicklung der Flugbahn eines Raumfahrzeugs im Rahmen der Mission GENESIS umgesetzt. Bei dieser Mission geht es darum, mit einer Sonde (vgl. Abbildung 1) auf einer geschlossenen Bahn im Sonnensystem zwei Jahre lang Sonnenstaub zu sammeln und anschließend das gesammelte Material zur Erde zurückzubringen. Der Start ist für Mitte 2001 geplant.

Naturwissenschaftler versprechen sich von der Analyse des Sonnenstaubs neue Erkenntnisse über die Entstehung unseres Sonnensystems. Detaillierte Informationen über den naturwissenschaftlichen Hintergrund dieser Mission findet man im Internet unter http://genesismission.jpl.nasa.gov.

Die GENESIS-Flugbahn

Betrachten wir nun einmal die in Abbildung 9 dargestellte Flugbahn des Raumfahrzeugs im Detail: zunächst bewegt sich das Raumfahrzeug entlang der lila Linie auf die geplante Umlaufbahn, einen so genannten Halo-Orbit, zu. Dieser ist in der Abbildung als rotes Oval zu erkennen.

Die Sonde verbleibt für etwa zwei Jahre auf diesem Halo-Orbit, um Sonnenstaub zu sammeln, anschließend fliegt sie zurück zur Erde.

Bei einer genaueren Betrachtung des Rückflugs fällt auf, dass das Raumfahrzeug nicht direkt zur Erde zurückkehrt, sondern zunächst entlang der grünen Bahn an der Erde vorbei fliegt, um sich erst anschließend in einem großen Bogen der Erde anzunähern. Der Grund für die Wahl dieser so sonderbar anmutenden Flugbahn liegt darin, dass diese aus Verbindungsbahnen zusammengesetzt und somit energetisch besonders günstig ist.

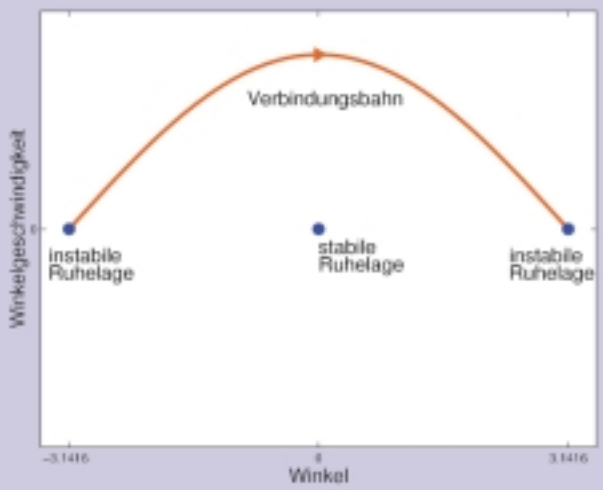


Abb. 8: Eine Bewegung des Pendels, welche die instabile Ruhelage mit sich selbst verbindet.

Abb. 9: Schematische Darstellung der Flugbahn des Raumfahrzeugs zum Halo-Orbit und zurück zur Erde.

Verbindungsbahnen des Halo-Orbits

Das der Mission GENESIS angemessene Drei-Körper-Problem besteht aus Sonne, Erde und Raumfahrzeug. In diesem mechanischen System ist – analog zur instabilen Ruhelage im Pendel – der Halo-Orbit eine instabile geschlossene Bahn. Dementsprechend besitzt auch der Halo-Orbit Verbindungsbahnen, entlang derer sich ein Raumfahrzeug prinzipiell ohne zusätzlichen energetischen Aufwand bewegen kann. Der Grund für die Wahl der eben beschriebenen GENESIS-Flugbahn wird nun klar: die Sonde nähert sich dem Halo-Orbit auf dem Hinflug entlang einer bestimmten Verbindungsbahn an und entfernt sich auf dem Rückflug entlang einer anderen Verbindungsbahn, die sie zunächst an der Erde vorbei führt.

Im Gegensatz zum Pendel besitzt der Halo-Orbit unendlich viele verschiedene Verbindungsbahnen. Diese bilden eine Fläche, wie sie in Abbildung 7 gezeigt ist.

Die Berechnung dieser Fläche wurde mit dem an der Universität Paderborn entwickelten Softwarepaket GAIO („Global Analysis of Invariant Objects“, http://www.upb.de/math/~agdellnitz/gaio) durchgeführt.

Die im Rahmen der Mission GENESIS als Flugbahn gewählte Verbindungsbahn liegt auf dieser Fläche. Ein den Flug illustrierendes Video entlang dieser Fläche kann im Internet unter http://www.upb.de/math/~agdellnitz/gaio/halo.html heruntergeladen werden. Die Flugbahn verbindet den Halo-Orbit mit einem weiteren Orbit, der auf der anderen Seite der Erde liegt (vgl. wiederum Abbildung 9). Auch dieser Orbit – er ist zwar in Abbildung 9 nicht dargestellt, man kann sich jedoch dessen Existenz gut vorstellen – besitzt eine Verbindungsbahn, entlang derer das Raumfahrzeug schließlich zurück in Richtung Erde fliegt. Die im Vergleich zum klassischen mission-design erzielten energetischen Einsparungen sind enorm.

Literatur

M. Dellnitz, G. Froyland und O. Junge: The algorithms behind GAIO – Set oriented numerical methods for dynamical systems. Erscheint 2001.

W.-S. Koon, M. W. Lo, J. E. Marsden, S. D. Ross: Heteroclinic Connections between Periodic Orbits and Resonance Transitions in Celestial Mechanics. Chaos **10**, 2000.

H. Poincaré: Les Méthodes Nouvelles de la Mécanique Céleste, Paris, 1899.

Dr. Oliver Junge, Fachbereich 17/Mathematik – Informatik, Mitarbeiter im Forschungsprojekt.
Dipl.-Math. Bianca Thiere, Fachbereich 17/Mathematik – Informatik, Mitarbeiterin im Forschungsprojekt.

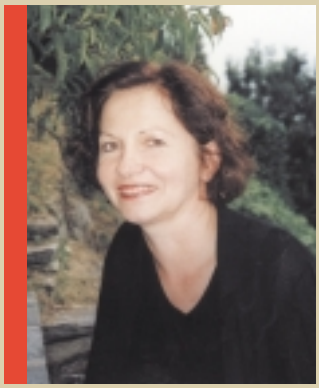
Umordnungen der Dinge

Kulturwissenschaftliche Untersuchungen sozialer und ästhetischer Ordnungen

Dinge sind nicht einfach gegeben – sie werden hergestellt, bewertet, mit Bedeutungen aufgeladen und immer wieder neu gestaltet, arrangiert und „codiert" – im Alltag ebenso wie in den Künsten. Die Untersuchung dieser Prozesse gibt Auskunft über unser Selbstverständnis als Nutzerinnen oder Besitzerinnen von Dingen und über die kulturellen Ordnungen, in denen wir uns bewegen.

Die Auslotung der „Kulturellen Transformationen von Dingen" ist der Gegenstand einer seit 1996 bestehenden Landesforschungsarbeitsgemeinschaft, die durch den Innovationsfonds des nordrhein-westfälischen Wissenschaftsministeriums gefördert wird. Die Arbeitsgemeinschaft, an der – unter der Federführung der Paderborner Professorin Gisela Ecker – nordrhein-westfälische Hochschullehrerinnen aus unterschiedlichen kulturwissenschaftlichen Disziplinen (Kunstwissenschaft, Design, Geschichte, Literaturwissenschaften, Filmwissenschaften, Ethnologie) beteiligt sind, hat sich zunächst in mehreren kleineren Arbeitstagungen und Workshops mit Objekten und ihrer kulturellen Transformation befasst. Der auf diese Weise zustandegekommene interdisziplinäre Dialog – im deutschen Universitätsalltag noch immer eine Seltenheit – erlaubt es, sich den Dingen auf komplexe Weise zu nähern: Die verschiedenen kulturwissenschaftlichen Disziplinen sind auf bestimmte Medien und Strategien des Umgangs mit Dingen spezialisiert, die sich, wie der interdisziplinäre Zugang erweist, allerdings verknüpfen und überlagern. Kürzlich ist im Helmer-Verlag, Königstein eine von Gisela Ecker und Susanne Scholz herausgegebene erste, grundlegende Publikation aus dem Kontext dieser Arbeit erschienen, die den Titel *Umordnungen* der Dinge trägt. Deren Ergebnisse werden im Folgenden auszugsweise, unter besonderer Berücksichtigung des Paderborner Anteils, vorgestellt. In Vorbereitung ist ein weiterer Band, der sich auf zwei spezifische Formen des Umgangs mit Dingen – das Sammeln und das Wegwerfen – konzentrieren wird. In den Räumen des Heinz Nixdorf MuseumsForums, Paderborn, fand zudem im November 2000 eine internationale Tagung statt.

Abb. 1: Vesna Stalljohann: *Ursel P. und ihre Freundinnen beim Tee*, 1998, Tassen als Stellvertreterobjekte, Projekt zur ästhetischen Forschung.



Prof. Dr. Gisela Ecker ist seit 1993 Professorin im Fachbereich 3/Allgemeine Literaturwissenschaft an der Universität Paderborn. Arbeitsschwerpunkte sind u.a. Vergleichende Literaturwissenschaft, *Gender Studies* und *Cultural Studies*. Einen weiteren Schwerpunkt bildet die Untersuchung von Geschlechterordnungen in kulturellen Leitvorstellungen und Praktiken.

Dinge erscheinen uns als objektiv gegeben, ihre Bedeutung scheint auf den ersten Blick erschließbar. Ein zweiter Blick, der sich darauf richtet, wie diese Bedeutung zustandekommt, verkompliziert ihre Wahrnehmung allerdings. Dinge können Gebrauchsgegenstand, Schaustück, Sammlungsobjekt, heimlicher Fetisch oder Teil eines Sammelsuriums sein. Sie werden gebraucht, gehütet, gepflegt oder weggeworfen, betrachtet, bestaunt oder ignoriert, registriert, gezählt oder archiviert. Dinge können von dem aus beschrieben werden, der sie handhabt, oder im Hinblick auf die Kultur, deren Bestandteil sie sind. Sie können im Kontext anderer Dinge oder isoliert wahrgenommen werden, als fremd oder nah, exotisch, künstlerisch, sakral oder trivial. So unmittelbar gegeben uns die Bedeutung von Gegenständen auch vorkommen mag: sie entsteht erst im jeweiligen historischen und kulturellen Kontext – und kann auch nur aus diesem heraus verstanden werden.

Dinge machen Sinn

Die Untersuchung der „Kulturellen Transformation von Dingen" ist daher ein genuin kulturwissenschaftliches Projekt.

Foto: Werner J. Hannappel, Essen.

Abb. 2: Timm Ulrichs: *San Gimignano*, 1983/86, Installation aus 13 Kopf stehenden bemalten Tischen und hochkant gestellten ziegelroten Beton-Pflastersteinen, 80 x 800 x 450 cm.

Abb. 3: Gabriele Schnitzenbaumer: *Lola*, 1997, Ton und Kupferboiler, 37 x 29 x 51 cm.

Die Kulturwissenschaften befassen sich mit Prozessen der Sinnstiftung und der Wertesetzung. Allerdings ist die Thematik der „Dinge" im kulturwissenschaftlichen Zusammenhang bisher eher vernachlässigt worden. Im Mittelpunkt der Forschung der letzten Jahrzehnte standen bevorzugt menschliche Individuen (fachsprachlich: „Subjekte") und die Entstehung ihrer Identitäten (z.B. als Männer und Frauen, Arbeiterinnen oder Adlige). Gerade in diesem Kontext aber gewinnt die Untersuchung der Dinge eine zentrale Bedeutung. Denn Subjekte gewinnen ihre Konturen im Umgang mit Dingen – eben z.B. als Käuferinnen, Sammler und Nutzer. Dinge fungieren als Medien, die zwischen dem Individuum und der sozialen Ordnung vermitteln, in der es sich bewegt. Sie sind ein wichtiger Faktor für Sinnstiftungen und Wertsetzungen, kurz: für kollektive und individuelle Konstruktionen von Welt.

Wichtiger Bestandteil der Art und Weise, wie wir eine „Ordnung der Dinge" erstellen, ist die Vorstellung von Eigentum und Besitz. Die Menschen definieren sich über die Dinge ihrer Umgebung, die sie hüten, verehren, gebrauchen oder begehren, in jedem Fall aber besitzen – oder besitzen wollen (vgl. James Clifford). Die Anordnung der Dinge, die das Subjekt umgeben, ist für ihn oder sie ein wichtiges Mittel, sich selbst darzustellen. Im Rahmen der Arbeitsgruppe sind Tableaus der Selbstdarstellung, beispielsweise auf dem Frisiertisch von Belinda in Alexander Popes *The Rape of the Lock*, auf Schreibtischen von Schriftstellern und in Objektarrangements privater Innenräume untersucht worden. Umgekehrt kann eine Spurensuche von den hinterlassenen Dingen einer verstorbenen Person hin zu dieser Person selbst unternommen werden (Abbildung 1).

Ins Spiel kommt dabei immer wieder die Frage nach Geschlechtszuschreibungen, denn viele Dinge werden durch ihr Auftreten in bestimmten Gebrauchszusammenhängen an die Geschlechter gebunden, oder sie werden von vornherein mit bestimmten Assoziationen umgeben: Küchentische, Kochutensilien, Kosmetika, Schmuckstücke scheinen auf eine ähnlich „natürliche" Art mit Weiblichkeit assoziiert zu sein wie Computer, Werkzeuge und Geldstücke mit Männlichkeit. Auf Grund dieser Zuordnungen – die selbstverständlich gerade nicht natürlich, sondern kulturspezifisch und veränderbar sind – gestaltet sich die Beziehung zwischen Ding und Mensch unterschiedlich.

Die Installation „San Gimignano" (Abbildung 2) von Timm Ulrichs spielt mit solchen Vorstellungen: Dreizehn auf den Kopf gestellte gebrauchte Küchentische und in sie platzierte Ziegelsteine nehmen dort eine Umordnung der Geschlechtszuschreibungen vor, denn nun wird die Tätigkeit der Hausfrau (die wir am Küchentisch vermuten mögen) in den Hintergrund gedrängt zu Gunsten einer Repräsentation von gesellschaftlicher Macht, die in den Geschlechtertürmen der toskanischen Stadt San Gimignano zum Ausdruck kommt. Was bleibt von den ursprünglichen Bedeutungsgebungen? Was wird durch das Kunstwerk umgemünzt, unterschwellig weitergetragen und in Beziehung gesetzt?

Auf den Kopf gestellte Zuschreibungen

Solche Umbesetzungen alltäglicher Gebrauchsgegenstände

Abb. 4: *Mantua mit Petticoat*, England ca. 1740. Victoria und Albert Museum, London.

Abb. 5: Ursula Neugebauer: *tour en l'air*, 1998, Sieben rote Taftkleider, Dekobüsten, computergesteuerte Elektromotoren, 280x700x1100 cm (Ausschnitt).

werden in den Künsten auf besonders auffällige – und deshalb aufschlussreiche – Weise vorgenommen. Ein weiteres Beispiel dafür kann in Gabriele Schnitzenbaumers Skulptur „Lola" gesehen werden, die aus einem ausrangierten Boiler und Ton besteht (Abbildung 3). Der Boiler ist vom Reich der Technik in das der Kunst übergegangen, er mutiert vom leblosen Ding zu einem als belebt vorgestellten Körper. Jenseits seines alten Nützlichkeits- oder Gebrauchswertes dient er nun anderen Zwecken – ohne dass er dabei jede Verbindung zu seinem „vorherigen Leben" verlöre. Während der Boiler in der Fantasie belebt wird, mechanisiert er im Gegenzug den Hundekörper, zu dessen Teil er geworden ist. Die Vorstellung von im Inneren zirkulierender warmer Flüssigkeit trägt zum Effekt der Skulptur bei. Auch die Spannung zwischen den Materialien Ton und Metall, die durch die mattierende Behandlung des Tons im Gegensatz zum glänzenden Kupfer erzeugt wird, trägt zu den Bedeutungseffekten des Kunstwerks bei: der „Zivilisationsmüll Boiler" wird mit dem Reich der (lebendigen?) Natur ebenso verknüpft wie kontrastiert.
Aber auch Dinge, die im alltäglichen Gebrauch verbleiben, sind nicht unumstößlich in ihrer Bedeutung festgelegt. Hier entscheidet ebenfalls der (vom Subjekt möglicherweise bewusst umgemünzte) Funktionszusammenhang, der sich unter anderem auch darüber herstellt, wo ein Objekt positioniert wird. Das Accessoire, das sich *per definitionem* am Rand seiner Trägerin bzw. seines Trägers befindet, erscheint als weniger wesentlich und weniger wertvoll als das im Zentrum platzierte Ding. Nicht zuletzt deshalb und wegen seiner Körpernähe ist es häufig weiblich konnotiert. Die Arbeitsgruppe hat sich innerhalb dieses Themenbereichs Dingen wie Schmuckstücken, Handtaschen, Knöpfen und dem Reifrock als historischem Beispiel zugewandt. Dieser funktionierte im 18. Jahrhundert nicht nur als Spiegelbild des Reichtums seiner Trägerin bzw. ihres Ehemannes, sondern auch als Verteidigungswall gegen befürchtete Angriffe auf ihre Ehre. Seine ausladende Fülle verkörperte weibliche Verschwendungssucht, männliche Finanzkraft und die Fragilität des darunter verborgenen Körpers gleichermaßen (Abbildung 4).

Dinge haben ein Eigenleben/ Dinge und Cyborgs

Ganz gleich aber, was Menschen mit Dingen tun oder wofür sie diese benutzen: die Dinge haben darüber hinaus ein „Eigenleben". In ihrem jeweiligen Umfeld erzeugen sie einen Bedeutungsüberschuss, der dazu führt, dass sie immer auch anders bewertet und geordnet werden könn(t)en. Während Dinge einerseits die Identität eines Subjekts (oder auch sein Alter Ego) repräsentieren mögen, erzeugen sie andererseits einen Überschuss an Materialität oder Dinghaftigkeit, der nicht im Dienst des Subjekts und seiner kulturellen Ordnung aufgeht. Dinge erzeugen Situationen, welche die vermeintliche Herrschaft der Menschen über sie unterlaufen.
Nicht immer sind Ding und Mensch klar voneinander zu trennen – Übergänge werden inszeniert oder stellen sich hinter dem Rücken des Subjekts ein. Szenarios von sich vermenschlichenden Computern und „cyborgisierten" Menschen bevölkern Texte, Filme, Computerspiele und virtuelle Räume gegenwärtig zuhauf. Sie thematisieren Ängste um die intellektuelle Vorherrschaft des Menschen, aber auch Fragen nach den Kriterien menschlicher Identität überhaupt. Ein Teilprojekt untersucht Schaufensterpuppen, die so arrangiert werden, dass sie zwischen Menschlichem und Künstlichem oszillieren, sie lassen die Übergänge zwischen denkendem Bewusstsein und animiertem Ding fließend werden: Grenzverschiebungen stellen sich ein (Abbildung 5).
Die kulturellen Prozesse der Umbesetzung und Umwertung der Dinge, die scheinbare Selbstverständlichkeiten aufheben, sind von besonderem Interesse für die Arbeit der Gruppe. Sie erlauben es, die Prozesse der Herstellung kultureller Ordnungen in

den Blick zu bekommen. So lässt z.B. die Frage, welche Dinge dazu auffordern, sie aufzuheben und vor dem Untergang zu retten, Rückschlüsse über kulturelle Werthierarchien zu, ebenso wie die Auswahl dessen, was als banal, trivial, wertlos und damit entbehrlich gilt. Warum überhaupt heben wir alte Dinge auf? Wer bestimmt, welche alten Möbel Antiquitäten und welche Müll sind? Sind diese Wertzuschreibungen für alle Mitglieder einer Gesellschaft verbindlich? Und sind sie in Geld auszudrücken? Obgleich Dinge in einer Konsumgesellschaft immer auch Waren sind, können sie nicht auf ihren Tauschwert reduziert werden. Denn neben ihrem materiellen haben sie einen symbolischen oder ideellen Wert, der nicht selten sogar bestimmend wird für ihre Bedeutung: Ein Stück Tuch mag an sich wenig wertvoll sein, als Fahne eingefügt in ein Ritual nationaler Sinnstiftung aber verändert sich dessen Wert kategorial.

Im Heimatmuseum wird die Welt sortiert

Sammlungen und künstlerische Installationen arrangieren Objekte in öffentlich wirksamer Form. Die Ansammlung von Ausstellungsstücken im Heimatmuseum, die ein weiteres Paderborner Projekt untersucht, zeigt diese Problematik deutlich. Dort werden Dinge gesammelt und ausgestellt, die vermeintlich ein Bild der „Heimat“ aus vergangenen Tagen zeigen. Doch bereits die Auswahl der Schaustücke ist erklärungsbedürftig: Sind sie Teil der „Heimat“, weil sie in einer geografisch klar umrissenen Gegend gefunden oder hergestellt wurden, oder weil sie jemand aus dieser Gegend besessen hat? Zeigen sie, wie es dort wirklich gewesen ist, oder projizieren sie unter dem Siegel der Authentizität ein Bild der „guten alten Zeit“ bzw. eine „schöne alte Welt“, die sie „Heimat“ nennen? Welche Menschen werden in ihrem Lebensstil und Lebensumfeld dort ausgestellt, welche ausgeblendet? Vom Leben der Bettler, Vaganten und Prostituierten erfährt man im Heimatmuseum in den seltensten Fällen; ihre „Heimatlosigkeit“ bezeichnet nicht nur ein Leben ohne festen Wohnsitz, sondern auch eine Existenz am Rand der Gesellschaft, die das Heimatmuseum reproduziert. Im Heimatmuseum wird eine bestimmte Vorstellung von kultureller und gesellschaftlicher Ordnung konstruiert und in die Vergangenheit projiziert, um als Norm oder Glücksvorstellung für das heutige Leben zu gelten.
Die Ausstellungsstücke im Heimatmuseum, die Schaustücke der frühneuzeitlichen Wunderkammern (die in einem anderen Beitrag des Bandes untersucht werden) und die Anordnung von Gemälden in einer Galerie sind Ordnungsschemata unterworfen, die erst dem kritischen Blick als „gemachte“, d.h. als kulturspezifische und historisch bestimmte zu Tage treten. Im historischen oder kulturellen Vergleich stellt sich beispielsweise heraus, dass uns heute als Sammelsurium erscheinen mag, was früher als unabänderliche und „wahre“ Ordnung der Dinge galt. Gegenwärtige wissenschaftliche Ordnungsschemata mögen späteren Zeiten ebenso absurd vorkommen wie uns heute die Denkweise früher (oder fremder) Enzyklopädien: Am Beispiel einer „gewissen chinesischen Enzyklopädie“ hat Michel Foucault vorgeführt, wie kulturspezifisch unsere Ordnungsmechanismen sind. Hier werden Dinge zusammengruppiert, die in unseren westlichen Kulturen in völlig verschiedene – bzw. als nicht sinnvoll erscheinende – Kategorien fallen, z.B. „Tiere, die dem Kaiser gehören“ und „Tiere, die mit einem ganz feinen Pinsel aus Kamelhaar gezeichnet sind“. Ähnliches trifft auf die Wunderkammern zu: auch hier wurden Dinge ausgestellt, die in den modernen naturwissenschaftlichen Klassifikationsschemata nicht vorkommen, wie etwa Hörner vom Einhorn. Welche Raster werden die „Wunderkammern“ des 21. Jahrhunderts erfinden?

Literatur

Bal, Mieke: Telling Objects. A Narrative Perspective on Collecting. In: John Elsner u. Roger Cardinal (Hg.): *The Cultures of Collecting*. London: Reaktion Books 1994, S. 97-115.

Clifford, James: On Collecting Art and Culture. In: Ders.: *The Predicament of Culture. Twentieth-Century Ethnography, Literature, and Art*. Cambridge, MA, London: Harvard University Press 1988, S. 215-51.

Foucault, Michel: *Die Ordnung der Dinge. Eine Archäologie der Humanwissenschaften*. Frankfurt am Main: Suhrkamp [9]1990.

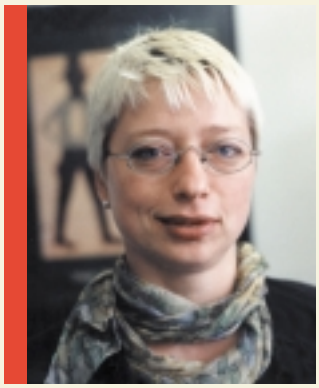
Garber, Marjorie: Fetish Envy. In: *October* 54, 1990, S. 45-56.

Kirkham, Pat (Hg.): *The Gendered Object*. Manchester, New York: Manchester University Press 1996.

Pearce, Susan (Hg.): *Experiencing Material Culture in the Western World*. London, Washington: Leicester University Press 1997.

Stewart, Susan: *On Longing. Narratives of the Miniature, the Gigantic, the Souvenir, the Collection*. Durham: Duke University Press 1993.

Dr. Claudia Breger ist wissenschaftliche Mitarbeiterin im Fach Allgemeine Literaturwissenschaft der Universität Paderborn. Arbeitsschwerpunkte sind u.a. Geschlecht und Ethnizität in Literatur und Kultur des 18.-20. Jhd.



Dr. Susanne Scholz ist wissenschaftliche Assistentin im Fach Allgemeine Literaturwissenschaft der Universität Paderborn (zurzeit DFG-Habilitationsstipendium). Arbeitsschwerpunkte sind Literatur und Kultur der Englischen Renaissance und des 18. Jhd., Cultural Studies und Feministische Theorie.

Auf dem Weg zu einer Theorie der Naturgesetze

Wie Naturgesetze innerhalb eines materialistischen Weltbildes verstanden werden können

Eine Liste von Naturgesetzen aufzustellen ist keine schwierige Aufgabe. Viel schwieriger als zu entscheiden, welche Kandidaten auf einer solchen Liste Platz finden sollten (siehe Kasten unten), ist die Frage zu beantworten, welches eigentlich die entscheidende Qualität ist, die den Kandidaten der Liste gemeinsam ist. Was also sollen wir unter „Naturgesetzen" verstehen? Da Naturgesetze offenbar durch einen bestimmten Typ von Aussagen repräsentiert sind, läuft diese Frage darauf hinaus, zu klären, wie es um die Wahrheitsbedingungen dieses Typs von Aussagen bestellt ist: Welche Sachverhalte müssen in der Natur erfüllt sein, damit wir zurecht davon sprechen können, dass ein Naturgesetz gilt? Die philosophische Tradition des Begriffs „Naturgesetz", dessen Karriere, von sporadischem Vorkommen in Antike und Mittelalter abgesehen, erst im 17. Jahrhundert beginnt, begünstigt zunächst eine supra-naturalistische Deutung von Naturgesetzen.

So sind für René Descartes (1596-1650) Naturgesetze universelle und unveränderliche Grundsätze, die Gott den Gegenständen der Natur vorgegeben hat. Die noch heute gebräuchliche Sprechweise, nach der ein Gegenstand ein Naturgesetz „befolgt", spiegelt diese Vorstellung wider. Die einflussreiche Analyse von David Hume (1711-1776) führt hingegen Naturgesetze auf Regelmäßigkeiten in der Natur zurück. Naturgesetze verlieren ihren supra-naturalen Charakter, d.h. ihre Geltung, unabhängig von den Dingen, die sie befolgen. Die moderne Debatte über Naturgesetze in der analytischen Wissenschaftsphilosophie zeigt eine noch weiter gehende Tendenz zur Naturalisierung von Naturgesetzen. Danach sind Naturgesetze Konsequenzen der inneren Eigenschaften natürlicher Systeme. Der vorliegende Beitrag zeigt, dass die Naturalisierung, die Einbettung von Naturgesetzen in ein materialistisches Weltbild, Grenzen hat. Naturgesetze können nicht völlig naturalisiert werden.

Prof. Dr. Andreas Bartels war seit 1997 Professor für Philosophie im Fachbereich 1/Philosophie der Universität Paderborn. Forschungsschwerpunkte sind die Interpretation moderner physikalischer Theorien, wie Relativitätstheorie und Quantenfeldtheorie, sowie die Analyse der Eigenschaften wissenschaftlicher Modelle. 2000 hat er einen Ruf an die Universität Bonn angenommen.

Naturalisieren durch Regularitäten

Die älteste Form der Naturalisierung von Naturgesetzen ist die Regularitätsauffassung. Nach dieser Auffassung, die auf David Hume zurückgeht und in einer neueren Variante von Norman Swartz vertreten wird, sind Naturgesetze wahr auf Grund der Wahrheit von Aussagen über das Auftreten instantiierender Regelmäßigkeiten. Naturgesetze sind danach keine supra-naturalen Entitäten. Sie tragen keine Wahrheit „sui generis", die unabhängig wäre von Fakten über aktualisierte Instanzen.
Naturgesetze beziehen sich nach Swartz auf tatsächliche Ereignisse oder auf realisierte Sachverhalte. Die Alternative ist also: Entweder, der Aussagegehalt von Naturgesetzen erschöpft sich in

1. Mendelgesetz (Uniformitätsgesetz):	Kreuzt man zwei reinerbige Rassen, die sich in einem Allelpaar unterscheiden, so sind alle F_1-Hybriden unter sich gleich.
2. Mendelgesetz (Spaltungsgesetz):	Werden Monohybride der F_1-Generation unter sich weiter gekreuzt, so sind die Individuen der F_2-Generation untereinander nicht gleich: Die Genotypen treten im Kombinationsverhältnis 1:2:1 auf.
Boltzmanns Entropieformel:	$S = k \cdot \ln W$
Einsteins Feldgleichung:	$R_{ik} - \frac{1}{2} g_{ik} R = -\kappa T_{ik}; \quad i, k = 1, 2, 3, 4.$
Galileis Fallgesetz:	$s = \frac{1}{2} at^2.$
Snellius Brechungsgesetz:	$\frac{\sin\alpha_1}{\sin\alpha_2} = \frac{CB}{CA} = n.$

Abb. 1: Eine Liste von Naturgesetzen zu erstellen ist nicht schwer ...

aktualen Regularitäten und damit sind Sachverhalte jenseits der aktualen Regularitäten als mit den Naturgesetzen unvereinbar ausgeschlossen. Dies ist Swartz' Position. Oder, Naturgesetze lassen Sachverhalte als *möglich* zu, die sich in den beobachteten Regularitäten unserer Welt niemals zeigen. Auch wenn Coca-Cola-Flüsse in unserem Universum niemals aufgetreten sind und (hoffentlich) niemals auftreten werden, so wären sie danach möglich, weil durch kein bekanntes Naturgesetz verboten.

Die zweite Perspektive, also ein als real interpretierter Spielraum physikalischer Möglichkeiten, hat nach Swartz rein fiktiven Charakter und er sei weder mit unserem üblichen Verständnis von Naturgesetzen noch mit der Praxis der Wissenschaft vereinbar. So müsste z.B. die Annahme, Menschen könnten möglicherweise die Meile unter vier Minuten laufen, als fiktiv angesehen werden, wenn es in der bisherigen Menschheitsgeschichte keine Leistungen gegeben hätte, die sich der vier Minuten-Grenze auch nur genähert hätten. Swartz weist jedoch keineswegs alle Potenzialitäten zurück. Die Wahrheit des Satzes „Dieses Stück Salz wird sich auflösen, wenn ich es in Wasser gebe" kann auf eine entsprechende Disposition von Salz gestützt werden, ebenso wie die Wahrheit des Satzes „Wenn Roger Bannister einen guten Tag erwischt, kann er die Meile unter vier Minuten laufen" auf bereits erwiesene athletische Dispositionen von Roger Bannister gestützt werden kann, die sich in einer stetigen Annäherung an die vier Minuten-Grenze manifestieren. In diesen Fällen stützen Dispositionen Wahrheitsansprüche von Konditionalen, aber nur deswegen, weil regelhafte Erfahrungen mit Erfüllungen solcher Dispositionen oder wenigstens Erfahrungen stetiger Annäherung an solche *Erfüllungen* uns dazu berechtigen, die Existenz der Dispositionen anzunehmen. Eine Welt, in der niemals ein Fall, in dem Salz sich in Wasser auflöst, oder athletische Leistungen in Nähe der vier Minuten-Grenze aufgetreten wären, würde uns, so Swartz, nicht die geeigneten empirischen Gründe bieten, an die entsprechenden Dispositionen zu glauben.

Testbare Dispositionen

Swartz' Forderung, die Annahme von Dispositionen müsse *empirisch testbar* sein, halte ich für unstrittig. Allerdings folgt aus der Akzeptanz dieser Forderung nicht, dass eine Erweiterung des Anwendungsbereichs von Naturgesetzen auf das physikalisch Mögliche verhindert wird. Denn es kann tatsächlich, im Widerspruch zu Swartz' Position, „methodologisch einwandfreie" Dispositionen geben (solche also, die empirisch getestet werden können), die sich jedoch in den Regularitäten der Welt niemals widerspiegeln. So könnte es sein, dass die genetisch bedingte physische Konstitution von Menschen in unserer Welt Unter-vier-Minuten-Zeiten über eine Meile zulässt, die Naturgesetze also mit einem Unter-vier-Minuten-Lauf vereinbar sind, die von Anbeginn der Menschheit tradierten Ernährungsgewohnheiten aber die Entfaltung dieser Ressource für alle bisher lebenden Menschen verhindert hat. Die Existenz entsprechender Möglichkeiten könnte z.B. in Laboruntersuchungen der Eigenschaften menschlicher Muskelfasern nachgewiesen werden. In derselben Weise ist die Annahme der Existenz einer Disposition der Sonne, mithilfe der Gravitationskraft die Erde auf eine Kepler-Ellipse zu zwingen, methodologisch unproblematisch, obgleich die Erde – auf Grund der Anwesenheit der anderen Planeten – niemals eine exakte Kepler-Ellipse beschrieben hat. Der empirische Test kann also an einem System erfolgen, in dem die exakte Erfüllung der fraglichen Disposition ausgeschlossen ist. Die Auffassung, nach der Naturgesetze von physikalischen Möglichkeiten handeln, die unabhängig von ihrer Aktualisierung als real interpretiert werden können, ist methodologisch vertretbar und mit der Praxis der Wissenschaft vereinbar. Ein kurzer Blick auf Galileis Argumentation für die Geltung des Fallgesetzes soll den letzten Punkt unterstreichen.

Abb. 2: Galileo Galilei (1564–1642).

Dispositionsauffassung bei Galilei

In den Discorsi versucht Salviati den Simplicius von der Geltung eines Naturgesetzes zu überzeugen: Schwere Körper fallen bei gleichen Anfangsbedingungen gleich schnell – unabhängig von Gewicht und Gestalt. Was Salviati und seine Gesprächspartner vor Augen haben, sind schwere Körper unterschiedlicher Gestalt und Masse, die durchaus unterschiedlich schnell fallen. Die bloßen Beobachtungen bestätigen Salviatis Behauptung nicht. Regelmäßige Abfolgen beobachtbarer Erscheinungen sind auch gar nicht der Gegenstand der Behauptung Salviatis. Dass das von ihm behauptete Naturgesetz gilt, ist für ihn offenbar nicht damit gleichbedeutend, dass alle fallenden schweren Körper, die wir beobachten können, Instanzen des behaupteten Gesetzes sind:

Streng genommen müssen wir gar keine Instanzen des Gesetzes beobachten können, um seine Geltung einzusehen. Was Salviati genügt, ist, plausibel zu machen, dass für den Fall einer vollständigen Ausschaltung aller störenden Einflüsse – also für den Fall eines idealen Vakuums – alle Körper gleich schnell fallen werden. Das Vakuum wird also als Prüfstein für die Geltung des Fallgesetzes angesehen. Aber sollen ausgerechnet hypothetische, nicht beobachtete Instanzen des vermuteten Gesetzes für seine Geltung zu Buch schlagen? Salviati nennt den Grund: Das Vakuum wird zum Prüfstein, weil nur im Vakuum alle Verschiedenheit im Fallverhalten unterschiedlicher Körper auf die Körper

Fred Dretske, Professor an der Universität Stanford, ist einer der Vertreter des top-down-Ansatzes.

selbst zurückzuführen wäre. Könnte man also plausibel machen, dass alle Körper im Vakuum gleich schnell fielen, so wüsste man, welchen Anteil an der Fallbewegung die Körper selbst hätten. Alle Körper haben exakt dieselbe Disposition zur Fallbewegung. Dies, so meint offenbar Salviati, ist es, was angenommen werden muss, wenn die Geltung des Gesetzes behauptet wird.

Das Fallverhalten eines neuen Gegenstandes, der bisher noch nicht im Fall beobachtet wurde, ist dann keine kontingente Tatsache, sondern folgt notwendig aus seiner Disposition zur Fallbewegung, die nach Annahme mit jener der schon beobachteten Körper identisch ist. Dass Körper im Vakuum gleich schnell fallen, ist zwar eine kontingente Tatsache; schließlich argumentiert Salviati dafür mithilfe der empirischen Tatsache, dass die Verschiedenheit im Fallverhalten mit der Reduzierung des Widerstands abnimmt. Salviati hat kein prinzipielles Argument dafür, dass es sich nicht ebenso gut anders verhalten könnte. Aber gegeben die Annahme, dass das Galileische Fallgesetz gilt – und dies ist nichts anderes als die Annahme, dass alle schweren Körper dieselbe Disposition zur Fallbewegung besitzen – zieht die Disposition des neuen getesteten Körpers dieses mit allen anderen fallenden Körpern konkordante Fallverhalten *notwendig* nach sich.

Naturalisieren durch Reduktion auf Eigenschaften: Top-down-Ansatz

Einerseits scheint – wenn wir die in Naturgesetzen implizierte *physische Notwendigkeit* ernst nehmen – kaum ein Weg an einer Dispositionsauffassung vorbeizuführen. Andererseits kann von einer Naturalisierung von Naturgesetzen nicht die Rede sein, solange unerklärbar bleibt, weshalb den Dingen unserer Welt durch Naturgesetze gerade diese Dispositionen zugesprochen werden. Das Auftreten einer Disposition sollte auf Grund der Eigenschaften des Gegenstands erklärbar sein, dem sie zugesprochen wird – so wie das Auftreten der Disposition „Zerbrechlichkeit" bei einer Glasscheibe auf Grund der molekularen Struktur der Glasscheibe erklärbar ist.

Dies ist die Motivation des top-down-Ansatzes von David Armstrong, Fred Dretske u.a. Die Möglichkeit der Zusprechung bestimmter Dispositionen soll auf Ebene der *Eigenschaften* materieller Systeme unserer Welt erklärbar werden, und zwar so, dass bestimmte Paare von Eigenschaften eine natürliche Relation *(„Necessitation")* erfüllen, die zum Inhalt hat, dass das Auftreten des einen Partners das Auftreten des anderen Partners notwendig nach sich zieht. Damit kann die Notwendigkeit, mit der Fälle, welche die Antezedenzbedingung eines Gesetzes erfüllen, zu Instanzen eines Gesetzes werden, von „oben nach unten", d.h. von der Ebene der Eigenschaften für die Ebene der Einzelfälle erklärt werden:

Zu erklären ist: „(Für alle x) dass x ein F ist, *macht es notwendig*, dass x ein G ist". Erklärt wird mithilfe von: „Ein F zu sein, ist eine Eigenschaft, welche die Eigenschaft, ein G zu sein, *notwendig* nach sich zieht".

Abgesehen davon, dass diese angestrebte Erklärung der Einzelfall-Notwendigkeit ihr Ziel nicht erreicht (vgl. Bas van Fraassen, 1989, S. 105 f.), enthält Armstrongs Lösung auch kein Potenzial, um notwendige Beziehungen auf der Eigenschafts-Ebene und strenge Regularitäten auf der Instanzen-Ebene voneinander zu trennen. Notwendigkeit auf der Eigenschafts-Ebene repräsentiert nach Armstrong dieselbe kausale Relation, die wir auch in der Erfahrung singulärer kausaler Relationen wahrnehmen. Wenn aber Eigenschaften nicht anders existieren als in ihren Vorkommnissen, dann besteht die Notwendigkeit auf der Eigenschafts-Ebene in nichts anderem als in dem Auftreten der entsprechenden singulären kausalen Relation für alle Instanzen. Sie kann daher selbst nur in Form strenger Regularitäten in der Abfolge (oder der Koexistenz) von Instanzen der notwendig verknüpften Eigenschaften in Erscheinung treten.

Damit fällt der top-down-Ansatz in eine Regularitätstheorie zurück. Da alle Naturprozesse störenden Einflüssen ausgesetzt sind, sind die Vorkommnisse beobachtbarer Eigenschaften natürlicher Systeme aber in der Regel nicht so stark miteinander verklammert, wie durch diese Auffassung gefordert. Wir stellen in den beobachtbaren Prozessen gerade nicht solche exakt konstanten Beziehungen des Auftretens von Eigenschaften fest. Dennoch würden wir nicht sagen, *dass jede* Abweichung von einer Regelmäßigkeit im Auftreten von Eigenschaften schon die Geltungsgrenze eines Naturgesetzes verletzt. Naturgesetze können auch dort in Geltung sein, wo – auf Grund starker Störungen – sehr „irreguläre" Prozesse stattfinden. Die Geltung eines Naturgesetzes hängt vielmehr davon ab, ob unter gewissen idealisierenden Bedingungen ein Verhaltenstyp durch alle Elemente einer Klasse von Systemen realisiert wird. Der top-down-Ansatz scheitert also, weil er empirisch inadäquat ist.

Naturalisieren durch Reduktion auf Eigenschaften: Der bottom-up-Ansatz

Der top-down-Ansatz identifiziert Naturgesetze mit Relationen, die zwischen den Eigenschaften materieller Dinge in *kontingenter* Weise bestehen. Dieselben materiellen Systeme mit denselben charakteristischen Eigenschaften könnten also auch andere Eigenschafts-Relationen, d.h. andere Naturgesetze erfüllen. Daher kann der Ansatz den grundlegenden Intuitionen einer Naturalisierung von Naturgesetzen nicht gerecht werden. Relationen zwischen natürlichen Dingen können völlig artifiziell und willkürlich sein, wenn nicht die Forderung gestellt wird, dass diese Relationen *auf Grund* der besonderen Eigenschaften dieser natürlichen Dinge bestehen.

Die Idee des bottom-up-Ansatzes (Rom Harré/Edward Madden, Chris Swoyer, John Bigelow u.a.) ist nun, die im top-down-Ansatz postulierten Eigenschafts-Relationen *von unten*, d.h. von der Ebene der instantiierten Eigenschaften selbst aus zu erklären. Eigenschaften, z.B. die Eigenschaft, die Ladung q zu besitzen, sind starr mit Dispositionen (Powers) verbunden; z.B. ein

Coulomb-Potenzial q/r zu produzieren. Solche Dispositionen rufen andere Eigenschaften hervor, z.B. die Eigenschaft der Coulomb-Anziehung, welche die Ladung q auf andere Ladungen ausübt. Deswegen stehen die Eigenschaft, die Ladung q zu besitzen, und die Eigenschaft, die Coulomb-Anziehung auf andere Ladungen auszuüben, in jener charakteristischen Relation, die wir ein Naturgesetz nennen. Ein naturalistisches Verständnis eines Naturgesetzes (z.B. des Coulomb-Gesetzes) ist erreicht, wenn gezeigt werden kann, dass dieses Naturgesetz *auf Grund* von Dispositionen gilt, die mit den Eigenschaften jener natürlichen Gegenstände verknüpft sind, die im Anwendungsbereich des Gesetzes liegen. Relativ zu den Eigenschaften dieser Gegenstände, gelten Naturgesetze nicht mehr, wie nach dem top-down-Ansatz, in kontingenter, sondern in *notwendiger Weise*.

Das Wasserstoff-Spektrum, dessen Erklärung ein spektakulärer Erfolg der Quantenmechanik war, verwenden Harré und Madden als Beispiel eines Naturgesetzes, das *auf Grund* der Struktur des Wasserstoff-Atoms gilt. Zu den Eigenschaften des Wasserstoff-Atoms zählt die Eigenschaft, bestimmte Energieniveaus für das Elektron zuzulassen; diese Eigenschaft ihrerseits ist mit der Disposition verknüpft, beim Übergang von einem höheren zu einem niedrigeren Energieniveau Fotonen einer bestimmten Energie (und der dadurch bestimmten Wellenlänge) auszusenden, und genau diese Disposition ist es, die für die Struktur des Wasserstoff-Spektrums verantwortlich ist. Freilich ist die Eigenschaft des Wasserstoff-Atoms, bestimmte Energieniveaus auszuzeichnen, eine durch die Quantenmechanik *zugesprochene* Eigenschaft. Die grundlegenden Gesetze der Quantenmechanik müssen also schon bekannt sein, um diese Eigenschaft, die mit ihr verbundene Disposition und schließlich daraus das fragliche Gesetz bestimmen zu können. Wenn wir *in diesem, von der Quantenmechanik belehrten Sinne* „wissen was Elektronen sind", lässt sich das Wasserstoff-Spektrum ableiten.

Kritik am bottom-up-Ansatz

Der bottom-up-Ansatz kann aber letztlich nicht erfolgreich sein. Einen Grund dafür liefert eine Beobachtung des englischen Philosophen John Locke (1632-1704).

Im 4. Buch des *„Essay Concering Human Understanding"* (Kap. VI, 11) warnt John Locke einmal davor, *die Powers*, die mit der Natur einer Substanz verbunden sind, allein auf die *primären Eigenschaften* dieser Substanz zurückzuführen. Die Wirkungen, die eine Substanz in unseren Wahrnehmungsorganen erzeugt, hängen eben nicht allein von den Eigenschaften dieser Substanz ab, sondern letztlich vom Resultat der Überlagerung aller Wirkungen, die uns von den Substanzen unserer näheren und weiteren Umgebung erreichen.

Wenn wir die Wirkungen zur Individuierung von Powers verwenden (wie dies sowohl Locke als auch Harré/Madden tun), können wir also nicht ohne weiteres hoffen, die für eine Substanz charakteristischen Dispositionen herauszugreifen. Epistemisch gewendet: Keine Modellierung eines physikalischen oder biologischen Systems kommt daran vorbei, Annahmen über die globale Struktur der Umgebung zu machen und diese Annahmen entscheiden mit über die Aussagen, die im Modell möglich werden.

Die in einem Modell zugesprochenen Dispositionen sind also nicht bereits mit den inneren Eigenschaften der modellierten Systeme gegeben (sie sind *keine* essenziellen Eigenschaften dieser inneren Eigenschaften), sondern hängen ebenso von der gewählten System-Umgebung ab. Lockes *Powers* lehren uns, unser Verständnis von Dispositionen im Blick auf die Bedingungen ihrer Zusprechung zu modifizieren.

Es bleibt das Ziel der theoretischen Naturerklärung, zu den charakteristischen Dispositionen der Substanzen vorzustoßen – z.B. ein Wasserstoff-Atom „monadisch" zu beschreiben, so als sei es das einzige System in der Welt. Dazu müssen wir Idealisierungen vornehmen, die Anteile der Umgebung an den Effekten herausfiltern sollen. Erst unter solchen Bedingungen werden *Naturgesetze* formulierbar: Nur wenn *Powers* von der Umgebung eines Systems isoliert und in dieser „gereinigten" Form dem System selbst zugesprochen werden, können wir hoffen, *notwendige* Zusammenhänge zwischen den Eigenschaften des Systems und den durch sie hervorgerufenen Effekten zu entdecken.

Die ontologische Konzeption des bottom-up-Ansatzes wäre nur dann plausibel, wenn solche Idealisierungsprozesse ein zumindest prinzipiell erreichbares eindeutiges Ziel besäßen. Eine vollständige Idealisierung, die Beiträge der Umgebung für die Effekte, die ein System hervorruft, zum Verschwinden bringt, scheint es, wie die folgenden Beispiele belegen, jedoch nicht zu geben. Die Quantenmechanik verwendet als Hintergrund eine klassische Raumzeit, in der Quantenfeldtheorie wird ein Minkowski-Raum als Einbettung der Systeme vorausgesetzt; jede Beschreibung eines Systems, das der allgemeinen Relativitätstheorie folgt, findet im Rahmen der speziellen Raum-Zeit-Struktur eines Modells statt, und selbst die Beschreibung des Sonnensystems

John Lockes philosophisches Hauptwerk „An Essay Concerning Human Understanding" erschien 1690. Gemälde von Sylvanus Brownover.

mithilfe einer Schwarzschild-Metrik kommt nicht ohne Randbedingungen im Unendlichen aus. Diese raumzeitlichen Hintergrund-Bedingungen sind nicht eliminierbar – es sei denn zu Gunsten eines alternativen raumzeitlichen Hintergrunds. Der raumzeitliche Hintergrund aber ist selbst ein Träger von Energie und beeinflusst daher die Form der Wirkungen, die ein System hervorruft. Wir können die Idealisierung nicht beliebig weit treiben, wir können nur wählen, welche Idealisierung wir vornehmen wollen. Aus diesem Grund kann die Auswahl einer Idealisierung, die zur Zusprechung bestimmter Dispositionen für ein System führt, nicht „von Natur vorgegeben" sein. Zumindest *sie*, die Auswahl der Idealisierung, ist nicht naturalisierbar.

Fazit

Meine Überlegungen haben ergeben: Naturgesetze sprechen Dispositionen zu. Aber die zugesprochenen Dispositionen sind nicht starr mit inneren („primären") Eigenschaften von Systemen verbunden. Wir leben nicht in einer Welt monadischer Einzelwesen, die alle und nur die für sie charakteristischen Dispositionen in sich tragen. Welche Dispositionen zugesprochen werden, hängt stets davon ab, welche Teile der Umgebung eines Systems in die Modellierung eingehen. Das Vorhaben der Naturalisierung von Naturgesetzen, am vollständigsten durchgeführt in der versuchten Rückführung nomologischer Dispositionen auf Eigenschaften, stößt daher an eine Grenze. Was wir vorgefunden haben, war keine externe Grenze der Naturalisierung. Wir sind nicht auf Gegenstände gestoßen, die prinzipiell der Naturalisierung unzugänglich sind. Die Grenze der Naturalisierung war vielmehr eine interne Grenze, ein uneliminierbarer Rest bei der Durchführung des Naturalisierungs-Programms. *Man kann alles naturalisieren – aber nicht alles auf einmal.*

Literatur

Armstrong, D.M. (1983): What is a Law of Nature?, Cambridge: Cambridge University Press.

Dretske, F. (1977): Laws of Nature, *Philosophy of Science* 44 (1977), S. 248-268.

Galileo Galilei (1964): Unterredungen und mathematische Demonstrationen über zwei neue Wissenszweige, die Mechanik und die Fallgesetze betreffend, Darmstadt: Wissenschaftliche Buchgesellschaft.

Harré, R./Madden, E. (1975): Causal Powers. A Theory of Natural Necessity, Oxford: Basil Blackwell.

Hüttemann, A. (1998): Laws and Dispositions, *Philosophy of Science* 65 (March), S. 121-135.

Swartz, N. (1985): The Concept of Physical Law, Cambridge: Cambridge University Press.

Swartz, N. (1995): A Neo-Humean Perspective: Laws as Regularities, in: Friedel Weinert (ed.): Laws of Nature. Essays on the Philosophical, Scientific and Historical Dimensions, Berlin: De Gruyter, S. 67-91.

Van Fraassen, B. (1989): Laws and Symmetry, Oxford: Clarendon Press.

Glossar

Supra-naturale-Entitäten: Entitäten übernatürlicher Art, wie es z.B. Wunder oder Engel wären.

Aktualisierte Instanz: Ein konkreter Fall, in dem sich ein Ereignis gemäß dem in einem Naturgesetz beschriebenen Verhalten ereignet.

Aktuale Regularitäten: Konkrete Ereignisfolgen, die den durch Gesetze beschriebenen Typen von Regulariäten folgen.

Kapazitätssteigerung in heutigen und zukünftigen Mobilfunksystemen

Optimierung mittels Computer-Simulationen

Angesichts ständig wachsender Teilnehmerzahlen und der Forderung nach höheren Datenraten für Multimedia-Anwendungen stellt die Steigerung der Funknetzkapazität eine der großen Herausforderungen für einen Mobilfunkbetreiber dar. Im Fachbereich Elektrotechnik der Abteilung Meschede werden verschiedene Methoden zur Kapazitätssteigerung mit Hilfe von Computer-Simulation auf ihre Effizienz untersucht und optimiert. Einige zentrale Ergebnisse zum Vergleich zwischen dem heutigen GSM-Standard (Global System for Mobile Communications) und dem zukünftigen Universal Mobile Telephone System (UMTS) werden in diesem Beitrag diskutiert.

In Deutschland hat sich die Zahl der Mobilfunkteilnehmer und -teilnehmerinnen in den vergangenen fünf Jahren bei Zuwachsraten von ca. 60 Prozent pro Jahr auf etwa 20 Millionen (Anfang 2000) verzehnfacht. Auch wenn sich der Anstieg in dieser Form nicht so fortsetzen kann, so ist doch zumindest mit einer Verdopplung der Teilnehmerzahl zu rechnen. Ferner wird erwartet, dass die Nutzung von Multimedia- und Internetdiensten über Mobilfunksysteme in den nächsten Jahren stark zunimmt. Daraus resultieren der Bedarf an höheren Datenraten und ein Zuwachs des Datenvolumens pro Mobilfunkteilnehmer und -teilnehmerin.

Prof. Dr. Christian Lüders ist seit 1998 Professor für Physik und Datenverarbeitung im Fachbereich 15/Nachrichtentechnik der Abteilung Meschede. Sein Forschungsgebiet liegt im Bereich der Planung und Kapazitätsanalyse von Mobilfunknetzen.
Dipl.-Ing. (FH) Markus Quente erwarb sein Diplom im Fachbereich 15/Nachrichtentechnik der Fachhochschulabteilung Meschede. Seit Anfang 1999 ist er als Mitarbeiter von Prof. Lüders in dem Drittmittelprojekt „Funknetzsimulationen für zukünftige Mobilfunksysteme" tätig.

Somit besteht eine große Herausforderung bezüglich der Steigerung der Funknetzkapazität sowohl für heutige Mobilfunksysteme, wie das GSM, als auch für zukünftige Systeme, wie das UMTS.
Im Rahmen eines von der Siemens AG geförderten Drittmittelprojekts werden verschiedene Verfahren zur Steigerung der

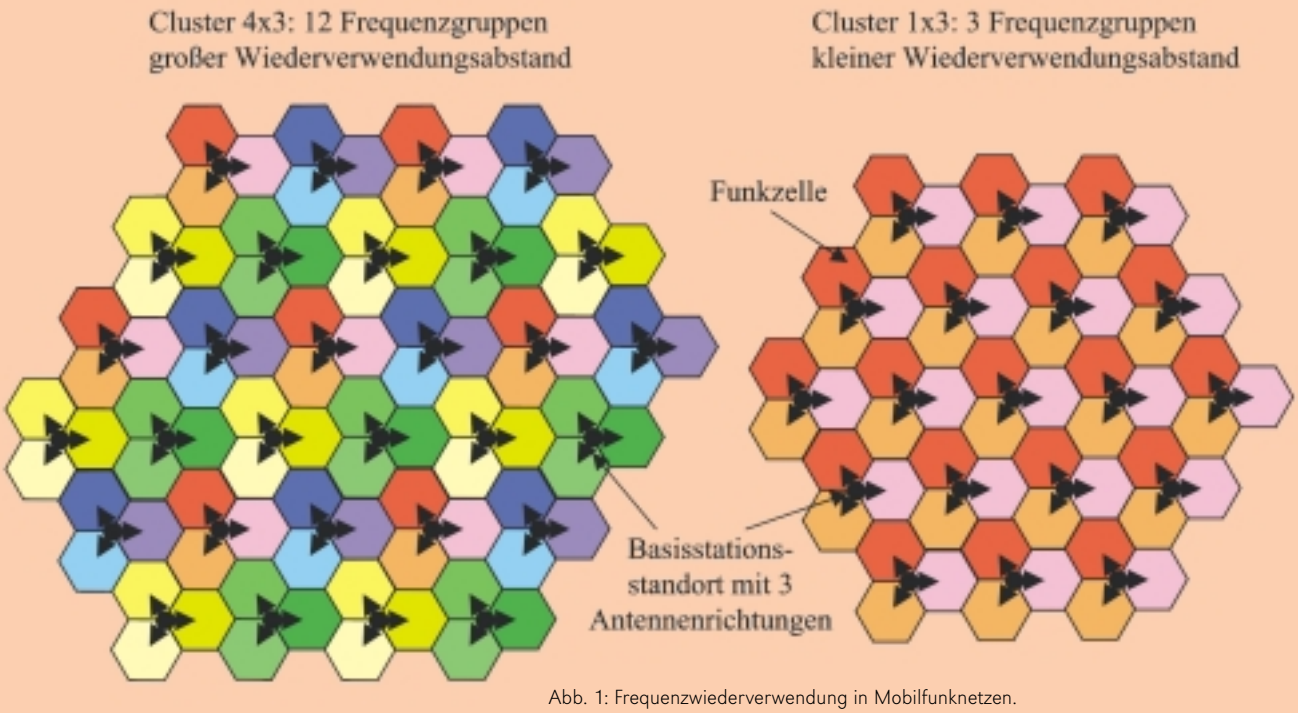


Abb. 1: Frequenzwiederverwendung in Mobilfunknetzen.

Funknetzkapazität mit Hilfe von Computer-Simulationen auf ihre Effizienz hin untersucht und optimiert. Dabei wurden ein von der Siemens AG entwickeltes Simulationstool, an deren Weiterentwicklung die Autoren mitgearbeitet haben, sowie von den Autoren entwickelte Simulationsmodelle verwendet.
In der Simulationssoftware werden die relevanten Strukturen und Abläufe innerhalb eines Mobilfunknetzes nachgebildet. Dazu gehören:
- Die Anordnung der Funkstationen und der zugehörigen Antennenkonfigurationen.
- Die Berechnung von Nutz- und Störleistungspegeln mittels realitätsnaher Funkausbreitungsmodelle.
- Die Verteilung sowie das Bewegungs- und Gesprächsverhalten der Mobilfunkteilnehmer.
- Mobilfunkspezifische Algorithmen, wie z.B. die automatische Regelung der Sendeleistung auf Basis der Empfangsleistung.

Die während der Simulationsläufe gesammelten Statistiken (Bitfehler, Datenrate, zurückgewiesene und abgebrochene Gespräche) geben Aufschluss über die Netzgüte bei den jeweils simulierten Bedingungen. Durch Variation der Teilnehmerzahl und anderer Parameter lässt sich so die maximal erzielbare Kapazität bei einer vorgegebenen Netzgüte bestimmen.
Dieser Beitrag diskutiert einige ausgewählte Ergebnisse zum Vergleich von GSM und UMTS.

Kapazität von Mobilfunknetzen

Bevor die untersuchten Verfahren und deren Leistungsfähigkeit im Detail diskutiert werden, soll zunächst erläutert werden, worin die Herausforderung bei der Steigerung der Funknetzkapazität begründet liegt.
Die entscheidende Randbedingung besteht darin, dass für Mobilfunkanwendungen nur ein eingeschränktes Frequenzspektrum zur Verfügung steht, welches von den Regulierungsbehörden in gewissen Portionen an die Mobilfunkbetreiber zugeteilt wird. Die beiden D-Netzbetreiber haben zur Zeit etwa jeweils 25 MHz Spektrum für das GSM-System im Einsatz. Was dies für die Funknetzkapazität bedeutet, zeigt eine kurze Überschlagsrechnung: Geht man davon aus, dass für einen Sprachkanal – wie derzeit bei GSM – 50 kHz Frequenzspektrum erforderlich sind, so hat jeder der beiden D-Netzbetreiber etwa 500 Sprachkanäle im Einsatz, denen ca. jeweils 10 Millionen Teilnehmer gegenüber stehen. Selbst wenn von diesen nur 2 Prozent gleichzeitig telefonieren, so muss jeder der Kanäle zumindest in 200 000/500 = 400 Funkzellen gleichzeitig verwendet werden. Eine hohe Funknetzkapazität ist also nur durch eine räumliche Wiederverwendung von Frequenzen möglich.
Zwei Muster für eine Frequenzwiederverwendung sind in Abbildung 1 für Funknetze aus hexagonalen Sektorzellen illustriert, welche üblicherweise bei der Untersuchung von Kapazitätsfragen verwendet werden. Sektorzellen heißt, dass von einem Standort einer Basisstation (Funkfeststation) aus 3 Funkzellen versorgt werden, die hier als regelmäßige Sechsecke symbolisiert sind. Sektorzellen lassen sich durch Antennen realisieren, die das Signal im Bereich von etwa 120° bündeln.
In den in Abbildung 1 gezeigten Beispielen ist die Gesamtzahl der Frequenzen N_{ges} in K = 12 bzw. K = 3 Frequenzgruppen eingeteilt, die durch verschiedene Farben gekennzeichnet sind. K nennt man die Cluster-Größe der Anordnung. Die Anzahl der Frequenzen pro Funkzelle ergibt sich also aus $N_{zelle} = N_{ges}/K$. Mit abnehmender Cluster-Größe steigt also die Funknetzkapazität; andererseits nehmen in diesem Fall die Störungen im Funknetz zu, da Zellen, die gleiche Frequenzen verwenden, näher zusammenrücken. Vom Standpunkt der Funknetzkapazität ist also eine möglichst geringe Cluster-Größe anzustreben; der günstigste Fall ist der einer Cluster-Größe K = 1, bei der Frequenzen auch in unmittelbar benachbarten Zellen (Sektoren) wiederverwendet werden können.
Geringe Cluster-Größen lassen sich aber nur mit speziellen Maßnahmen erreichen, nämlich
- durch eine Erhöhung der Störfestigkeit des Übertragungsverfahrens (z.B. durch Erhöhung des Fehlerschutzes) und
- durch Reduktion der Störungen im Funknetz.

Der vorliegende Beitrag konzentriert sich auf Methoden zur Reduktion von Störungen im Netz, von denen einige im Folgenden erläutert werden:
Bei der **Sendeleistungsregelung** passt sich die Sendeleistung einer Mobil- oder Basisstation automatisch an die jeweils herrschenden Empfangsbedingungen an. So reduziert z.B. eine Mobilstation, die sich nahe an der Basisstation befindet, ihre Sendeleistung und verursacht damit weniger Störungen in anderen Zellen. Beim GSM-System erfolgt die Regelung etwa im Sekundentakt. Dadurch ist es möglich, auf den entfernungsabhängigen Anteil des Empfangspegels und gegebenenfalls auf veränderte Abschattungsverhältnisse zu reagieren. Für eine Reaktion auf das sogenannte Kurzzeitfading ist die Regelung bei GSM jedoch zu langsam. Kurzzeitfading entsteht dadurch, dass das Funksignal durch Reflexionen an Bergen oder Gebäuden auf verschiedenen Wegen zum Empfänger gelangt. Die zugehörigen Funkwellen überlagern sich und können sich gegenseitig auslöschen oder verstärken. Die resultierenden starken Pegelschwankungen erfolgen auf einer Längenskala von einigen Zentimetern (im Bereich der halben Wellenlänge) und sind zudem frequenzabhängig. Bei einer Mobilstation im Fahrzeug bei 50 km/h ändert sich das Kurzzeitfading also im Bereich von Millisekunden. Im Vergleich dazu ist demnach eine Regelung im Sekundentakt als langsam einzustufen.
Discontinuous Transmission beruht darauf, dass bei einer Sprachverbindung jeder der beiden Gesprächspartner im Mittel nur zu etwa 50 Prozent spricht. Gesprächspausen und das Wiedereinsetzen der Sprechaktivität können von der Sprachkodiereinheit im Sender automatisch erkannt werden. Dadurch ist es möglich, die Sendeeinheit bei Sprachpausen automatisch abzuschalten, sodass andere Teilnehmer weniger gestört werden.
Beim **Frequenzsprungverfahren** (Frequency Hopping) wechseln Sender und Empfänger in einem bestimmten Rhythmus die Sende- und Empfangsfrequenz. Auf diese Weise lassen sich frequenzabhängige Empfangs- und Störverhältnisse ausgleichen, sodass für alle Verbindungen möglichst gleichmäßige Bedingungen auftreten und eine „worst-case-Planung“ vermieden werden kann.

Kapazität von GSM-Mobilfunknetzen

Das Potenzial der zuvor erläuterten Methoden soll kurz anhand des GSM-Systems illustriert werden, um einen Ausgangs- und Vergleichspunkt für die anschließenden weiterführenden Untersuchungen zu haben. Ohne besondere kapazitätssteigernde

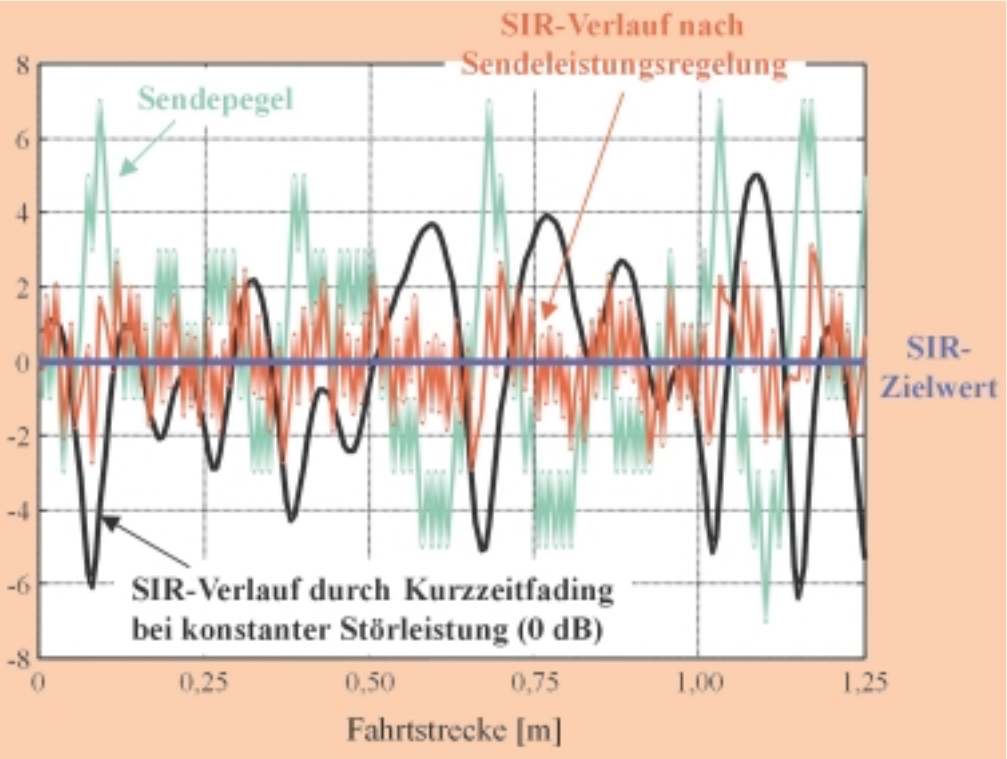


Abb. 2: Illustration der Sendeleistungsregelung bei UMTS für eine Mobilstation im Fahrzeug bei 36 km/h. Falls das gemessene Verhältnis von Signal- zu Störleistung (SIR) unter dem vorgegebenen Zielwert liegt, wird der Sendepegel um einen festen Betrag (hier 2 dB) erhöht, anderenfalls erniedrigt.

Methoden lässt sich für das GSM-System bei dem in Abbildung 1 gezeigten Muster aus 12 Frequenzgruppen eine akzeptable Sprachqualität erzielen.

Für einen Betreiber mit z.B. 19,2 MHz Spektrum ergeben sich daraus N_{ges} = 19,2 MHz/50 kHz = 384 Kanäle insgesamt, d.h. N_{zelle} = 384/12 = 32 Kanäle pro Zelle. Auf Grund des Gesprächsverhaltens der Teilnehmer ist eine vollständige Auslastung dieser 32 Kanäle jedoch nicht erreichbar. Verlangt man, dass nicht mehr als 2 Prozent der Gesprächswünsche wegen voll belegter Funkkanäle zurückgewiesen werden dürfen, so lässt sich im Mittel eine Auslastung von ca. 75 Prozent erzielen; man sagt ein Verkehrsangebot A = 0,75·32 = 24 Erl kann versorgt werden. Die (Pseudo-)Einheit Erlang (Erl) charakterisiert die Last in Telekommunikationsnetzen. 24 Erl bedeuten, dass im Mittel über einen längeren Beobachtungszeitraum gleichzeitig 24 Gespräche in der Funkzelle geführt werden.

Da die Kapazität mit dem zur Verfügung stehenden Gesamtspektrum wächst, führt man zur Charakterisierung der Leistungsfähigkeit eines Mobilfunksystems die spektrale Kapazität kS als versorgbares Verkehrsangebot pro Frequenz und Zelle ein – gewissermaßen als Verhältnis zwischen Gewinn und Kosten für einen Mobilfunkbetreiber. Für das Beispiel beträgt sie:

$$\kappa_s = \frac{24}{19{,}2} \cdot \frac{Erl}{MHz \cdot Zelle} = 1{,}25 \cdot \frac{Erl}{MHz \cdot Zelle}$$

Computer-Simulationen [1], [2], die zum Teil für sehr realitätsnahe Funknetzstrukturen und Ausbreitungmodelle durchgeführt wurden, haben gezeigt, dass sich durch Leistungsregelung, Discontinuous Transmission und Frequency Hopping ein Cluster 1 realisieren lässt, wobei allerdings gleichzeitig nur etwa 30 bis 40 Prozent der Kanäle belegt sein dürfen, um die gegenseitigen Störungen erträglich zu halten. Auf diese Weise kann die spektrale Kapazität auf etwa k_S = 3 Erl/MHz·Zelle mehr als verdoppelt werden.

Übertragungsverfahren bei UMTS

Bei der Auswahl des Übertragungsverfahrens für UMTS (Ende 1997/Anfang 1998) war die effektive Nutzung des verfügbaren Frequenzspektrums, d.h. eine hohe spektrale Effizienz, ein entscheidendes Kriterium. Seitens des „European Telecommunications Standardisation Institute“ (ETSI) wurden zwei von fünf vorgeschlagenen Verfahren für die weitere Standardisierung ausgewählt, welche sich in den beiden für verschiedene Anwendungsbereiche konzipierten Betriebsmodi des UMTS wiederfinden [3].

In beiden Betriebsmodi kommt das CDMA-Verfahren (CDMA: Code Division Multiple Access) zum Einsatz, bei dem auf einem Frequenzträger gleichzeitig mehrere Verbindungen untergebracht werden. Die Trennung der Verbindungen am Empfänger erfolgt anhand verbindungsspezifischer Codesignale, mit denen die zu übertragenden Daten am Sender multipliziert werden. Diese Multiplikation führt zu einer Spreizung des Signals im Frequenzbereich, d.h. der erforderliche Frequenzbereich ist um ein Vielfaches (den sogenannten Spreizungsfaktor) größer als es zur Übertragung des Signals für eine einzelne Verbindung notwendig wäre. Für UMTS sind etwa 5 MHz Spektrum für einen Frequenzträger erforderlich.

Zumindest ausgeglichen wird diese Bandbreitenerhöhung durch die Übertragung von zahlreichen Verbindungen auf einem Frequenzträger. Vom Standpunkt der spektralen Effizienz aus liegt der Hauptvorteil von CDMA darin begründet, dass durch die hohe Trägerbandbreite und die große Anzahl von Verbindungen pro Träger die Empfangs- und Störverhältnisse sehr gleichmäßig, d.h. nur geringen Schwankungen unterworfen sind. Sowohl die frequenzabhängigen Kurzzeit-Fading-Einbrüche als auch Störungen aus Nachbarzellen können weitgehend ausgemittelt werden.

Von entscheidender Bedeutung für die Leistungsfähigkeit eines CDMA-Mobilfunksystems ist eine schnelle und gut funktionierende Sendeleistungsregelung, denn wenn viele Verbindungen über eine Frequenz geführt werden, kann jede nur dann erfolgreich detektiert werden, wenn am Empfänger in etwa gleiche Empfangsleistungen für all diese Verbindungen vorliegen.

Um auf Pegelschwankungen, wie sie sich aus dem nach der Frequenzmittelung verbleibenden Kurzzeitfading ergeben, reagieren zu können, wurde der Regelungstakt gegenüber dem GSM-System deutlich (etwa um den Faktor 1 000) erhöht. In dem von uns untersuchten Betriebsmodus teilt der Empfänger auf Basis der Empfangsverhältnisse dem Sender alle 625 µs mit, ob die Sendeleistung zu erhöhen oder zu senken ist. Die Schrittweite dieser Änderungen kann vom Netzbetreiber eingestellt werden.

Ebenso lässt sich das SIR-Ziel, also das gewünschte Verhältnis aus Nutz- und Störleistung (SIR: Signal-to-interference-ratio) vorgeben.

Simulationsergebnisse zur Leistungsregelung bei UMTS

Eine wesentliche Aufgabe innerhalb des Drittmittelprojekts war es, den Prozess der Leistungsregelung zu optimieren, d.h. die Werte für die Schrittweite, das SIR-Ziel und weitere Parameter zu finden, die zu maximaler Funknetzkapazität bei vorgegebener Dienstqualität führen.

Das prinzipielle Verhalten der Sendeleistungsregelung bei UMTS ist in Abbildung 2 illustriert, und zwar für eine einzelne Mobilstation, bei der sich die Störungen nur durch das Empfängerrauschen, aber nicht durch andere Mobilstationen ergeben.

Die schwarze Kurve zeigt einen typischen Verlauf des Kurzzeitfadings bei UMTS. Bei einer Fahrzeuggeschwindigkeit von v = 36 km/h erfolgt die Sendeleistungsregelung etwa alle 6 mm, sodass bei einer Schrittweite von 2 dB die Sendeleistungsregelung auf den Kurzeitfadingverlauf reagieren kann und dafür sorgt, dass das resultierende Signal-zu-Störleistungsverhältnis (SIR) um den hier eingestellten Zielwert von 0 dB konzentriert ist. Bei Fußgängergeschwindigkeiten sind kleinere Schrittweiten (ca. 0,5 dB) günstig. Bei hohen Geschwindigkeiten (>100 km/h) hat die Sendeleistungsregelung Schwierigkeiten – selbst bei größeren Schrittweiten -, dem Kurzzeitfading zu folgen.

Ähnliche Werte für die Schrittweite erhält man auch für den Fall größerer Funknetze (57 Sektorzellen mit jeweils 10 bis 100 aktiven Mobilstationen pro Zelle in der Simulation), bei denen sich die Mobilstationen gegenseitig stören. Der entscheidende Parameter, der dabei die Kapazität und Qualität im Netz bestimmt, ist das SIR-Ziel. Sein Einfluss und seine Wirkungsweise lassen sich am besten anhand der Paketdatenübertragung illustrieren, die z.B. bei einem mobilen Internetzugang zum Tragen kommt. Bei der Paketdatenübertragung wird der zu übertragende Datenstrom in Pakete zerlegt, die mit einer Kombination aus fehlerkorrigierenden und fehlererkennenden Codes versehen werden. Beim Empfang eines Datenpaketes versucht der Empfänger, eventuelle Übertragungsfehler zunächst selbstständig zu korrigieren. Erkennt der Empfänger, dass selbst nach der Korrektur noch Fehler verblieben sind, so fordert er das Paket vom Sender nochmals an. Auf diese Weise ist es möglich, eine geringe Bitfehlerrate zu garantieren. Allerdings sinkt bei häufigen Wiederholungen der effektive Datendurchsatz, d.h. die Anzahl erfolgreich übertragener Nutzbits pro Zeit, deutlich ab.

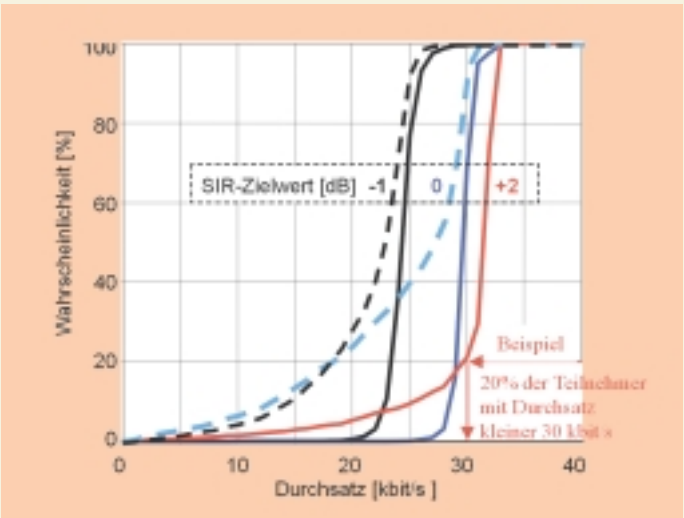


Abb. 3: Wahrscheinlichkeitsfunktionen für den Durchsatz bei einem Paketdatendienst bei verschiedenen Regelungszielwerten und Lasten (durchgezogene Kurven: 20 Erl pro Zelle, gestrichelte Kurven: 30 Erl pro Zelle). Die Ordinate gibt die Wahrscheinlichkeit an, dass der erzielte Datendurchsatz unter dem zugehörigen Wert auf der Absizze liegt.

Die folgenden Diagramme zeigen den Einfluss der Sendeleistungsregelung, sowie den Einfluss des SIR-Ziel-Wertes auf den effektiven Datendurchsatz pro Teilnehmer. Für das diskutierte Beispiel wurde der maximale Durchsatz pro Teilnehmer auf 32 kbit/s begrenzt. Untersuchungen für höhere Datenraten lieferten qualitativ das gleiche Verhalten.

In Abbildung 3 sind Wahrscheinlichkeitskurven für den Durchsatz pro Teilnehmer für verschiedene SIR-Zielwerte und verschiedene Lasten zu sehen. Die Kurven geben die Wahrscheinlichkeit an, dass der Durchsatz durch Wiederholungen von Paketen unter den Wert auf der x-Achse fällt. So beträgt z.B. bei der roten Kurve die Wahrscheinlichkeit 20 Prozent, dass der Durchsatz unter 30 kbit/s fällt.

Betrachtet man zunächst die durchgezogenen Kurven für eine Last von 20 Erl pro Funkzelle, so fällt auf, dass bei einer Senkung des SIR-Ziels von 0 dB auf –1 dB der Durchsatz pro Teilnehmer etwa gleichmäßig um ca. 5 kbit/s sinkt, wie an der Verschiebung der blauen Kurven gegenüber der schwarzen zu erkennen ist. Durch die Absenkung der Sendeleistung erhalten die Teilnehmer ein schlechteres Verhältnis zwischen Nutz- und Störleistung, sodass alle Teilnehmer Datenpakete häufiger wiederholen müssen.

Erhöht man das SIR-Ziel (von 0 dB auf 2 dB), so tritt ein anderer interessanter Effekt auf.

Für gut 80 Prozent der Teilnehmer ist der Durchsatz besser (ca. 2 kbit/s) als bei der schwarzen Kurve, andereseits liegt er für knapp 20 Prozent der Teilnehmer zum Teil deutlich niedriger, wie anhand der roten Kurve zu sehen ist. Bei 2 Prozent der Teilnehmer liegt der Durchsatz sogar unter 10 kbit/s. Bei einer Erhöhung des SIR-Zielwertes profitiert ein Teil der Mobilteilnehmer zu Lasten eines anderen Teils, denn durch die damit verbundene Erhöhung des Sendeleistungspegels steigt die Störleistung im Netz insgesamt. Der optimale SIR-Zielwert von 0 dB (schwarze Kurve) gewährleistet allen Teilnehmern einen nahezu gleichmäßig guten Durchsatz.

Bei einer Erhöhung der Last auf 30 Erl pro Zelle (gestrichelte Kurven) steigen die Störungen im Netz insgesamt so weit an, dass bei allen SIR-Zielwerten ein nicht zu vernachlässigender Anteil der Teilnehmer nur einen geringen Durchsatz erhält. Legt man als Qualitätsmaßstab einen Satz von 98 Prozent zufriedenen Teilnehmern zugrunde, wobei gemäß Bewertungsrichtlinien aus der UMTS-Standardisierung ein zufriedener Teilnehmer für den gezeigten Fall einen Durchsatz von mindestens 6,4 kbit/s erhalten muss, so ist bei etwa 30 Erl pro Zelle die Lastgrenze erreicht. Der Mindestdurchsatz von 6,4 kbit/s wird bei 2,5 Prozent der Teilnehmer unterschritten.

Aus den Wahrscheinlichkeitsfunktionen für den Durchsatz lassen sich mit dem mittleren Durchsatz pro Teilnehmer sowie dem Anteil „unzufriedener Teilnehmer“ mit einem Durchsatz unterhalb von 6,4 kbit/s zwei wichtige Kenngrößen extrahieren.

Diese beiden Kenngrößen sind in den folgenden Diagrammen für verschiedene Lasten und SIR-Zielwerte aufgetragen. Die

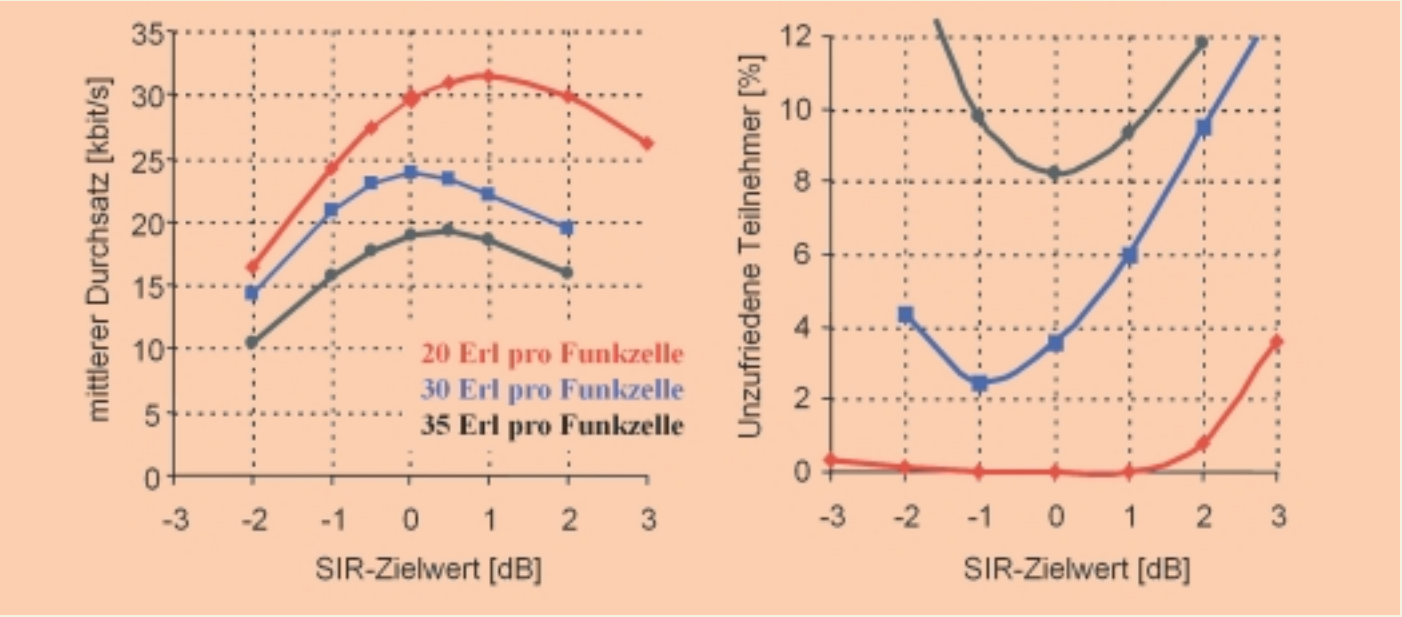


Abb. 4: Mittlerer Durchsatz pro Teilnehmer und Anteil unzufriedener Teilnehmer (Durchsatz kleiner als 6,4 kbit/s) in Abhängigkeit vom Regelungszielwert für verschiedene Lasten.

Kapazitätsgrenze ist bei etwas unter 30 Erl pro Zelle erreicht; beim optimalen SIR-Zielwert von –1 dB sind 2,5 Prozent der Teilnehmer unzufrieden. 35 Erl pro Zelle stellen eine Überlastsituation mit mehr als 9 Prozent unzufriedenen Teilnehmern dar. Bei 20 Erl pro Zelle liegt die Last deutlich unter der kritischen Grenze, dementsprechend ist der Anteil unzufriedener Teilnehmer über einen recht großen Bereich von SIR-Zielwerten nahezu 0 Prozent. Ferner zeigen die Kurven für den mittleren Durchsatz, dass es einen lastabhängigen optimalen Wert für das SIR-Ziel gibt.

An der Lastgrenze von etwa 30 Erl in einer mit einem Träger (Bandbreite 5 MHz) bestückten Funkzelle ergibt sich bei einem mittleren Durchsatz von 24 kbit/s eine spektrale Effizienz von

$$\kappa_{\eta} = \frac{30 \cdot 24\ kbit/s}{5\ MHz \cdot Zelle} = 144 \frac{kbit/s}{MHz \cdot Zelle}$$

Zu beachten ist, dass bei der spektralen Effizienz für Datendienste nicht nur die Verkehrslast pro Zelle (Anzahl aktiver Teilnehmer), sondern auch die Datenrate, die jeder dieser Teilnehmer im Mittel erhält, zu berücksichtigen ist.

Ähnliche Untersuchungen wie für die Paketübertragung wurden auch für die Sprachübertragung durchgeführt. Auch hier zeigte sich eine empfindliche Abhängigkeit der erreichbaren Kapazität vom SIR-Zielwert. Als spektrale Effizienz ergab sich der Wert k_S = 5 Erl/MHz·Zelle. Bei einer für UMTS angenommenen Sprachdatenrate von 8 kbit/s entspricht dies einem Wert von k_S = 40 kbit/s/MHz·Zelle.

Vergleich der Kapazitätszahlen

Bei einem Vergleich der spektralen Effizienzen für die Sprach- und Paketdatenübertragung fällt auf, dass der Wert für die Paketdatenübertragung deutlich höher liegt. Dies lässt sich darauf zurückführen, dass der Wiederholmechanismus eine Fehlerschutzmaßnahme ist, die sich gut an die jeweils herrschenden Empfangsbedingungen anpasst – sie muss nur eingesetzt werden, wenn bei schlechten Empfangsbedingungen tatsächlich Übertragungsfehler auftreten. Dagegen sind die Fehlerschutzmaßnahmen bei der Sprachübertragung, die auf ungünstige Empfangsverhältnisse ausgelegt sind, auch bei guten Verhältnissen aktiv und erfordern somit ein hohes Maß an Redundanzbits. Das sind Bits, die der Sender auf systematische Weise den eigentlich zu übertragenden Nutzbits hinzufügt, um dem Empfänger eine automatische Korrektur etwaiger Übertragungsfehler zu ermöglichen.

Vergleicht man die spektralen Effizienzen für Sprachdienste bei GSM (3Erl/MHz·Zelle) mit UMTS (5 Erl/MHz·Zelle), so ist UMTS knapp doppelt so effizient in der Spektrumsnutzung wie GSM. Zu beachten ist, dass für UMTS höhere Anforderungen an die Versorgung (mehr als 98 Prozent) und die Sprachqualität (Bitfehlerate < 0,1 Prozent), als sie bei GSM üblich sind, gestellt und bei den Simulationen auch berücksichtigt wurden. Bei GSM geht man üblicherweise von einer Versorgungswahrscheinlichkeit von 90 Prozent aus; eine akzeptable Sprachqualität liegt bei einer Bitfehlerrate unterhalb von etwa 0,5 Prozent vor. Würde man für GSM die UMTS-Anforderungen zu Grunde legen, wäre das Verhältnis der spektralen Effizienzen etwa 3:1 zu Gunsten von UMTS. Diese Effizienzsteigerung bei UMTS ist auf die erhöhte Übertragungsbandbreite, die verbesserte Sendeleistungsregelung sowie auf einen stärkeren Fehlerschutz zurückzuführen.

Zusammenfassung und Ausblick

Das Potenzial verschiedener kapazitätssteigernder Maßnahmen wurde für das GSM- und UMTS-Mobilfunksystem untersucht. Insbesondere an Hand der Paketdatenübertragung bei UMTS ließ sich mittels Computer-Simulationen der große Einfluß einer gut arbeitenden Sendeleistungregelung auf die Funknetzkapazität und -qualität illustrieren. Die Simulationen lieferten optimale Systemparameter-Einstellungen für verschiedene Szenarien und Kapazitätswerte für die Sprach- und Datendienste.

Zusätzliche Kapazitätssteigerungen von mehr als 100 Prozent lassen sich sowohl für GSM [4] als auch für UMTS durch intelligente Antennensysteme erzielen, wie weitere Simulationen gezeigt haben.

Während bei den bisherigen Simulationen die Sprach- und Paketdatendienste mit ihren unterschiedlichen Qualitätsanforderungen getrennt betrachtet wurden, konzentrieren sich die weiteren Forschungen auf Fragen der Funknetzplanung, insbesondere auf die Optimierung der Funkkanalzuteilung und der Sendeleistungsregelung bei einem Dienstemix.

• Wir danken den Mitarbeitern der Siemens AG aus den Dienststellen ICN CA MR MA 8, ICN CA MR EI 1 und PSE TN MNR 22 für die gute Zusammenarbeit, viele wertvolle Diskussionen und Anregungen. Insbesondere danken wir Dr. E. Schulz, dass er uns das Siemens-eigene Funknetz-Simulationstool sowie die erforderliche beträchtliche Rechnerleistung für die Untersuchungen zur Verfügung gestellt hat.

Literatur

[1] K. Ivanov, C. Lüders, et al.: „Spectral Capacity of Frequency Hopping GSM“, in „Evolution Towards 3rd Generation Systems“, Editors: P. Jung, K. Kammerlander, P. Zvonar, Kluwer Academic Press, 1999.

[2] U. Rehfueß, K.Ivanov: „Comparing Frequency Planning Against 1x3 and 1x1 Re-Use in Real Frequency Hopping Networks“, Proc. 49th IEEE Veh. Technol. Conf., June 1999.

[3] W. Mohr: „Internationale Standardisierungsaktivitäten zur Definition der dritten Mobilfunkgeneration“, Frequenz, Band 54, März/April 2000.

[4] K. Ivanov, C. Lüders, U. Rehfueß, „Comparing and Estimating the Upper Bounds on the Capacity Gain from Different Smart Antenna Concepts in GSM Mobile Radio Networks”, Proc. 48th IEEE Veh. Technol. Conf., Ottawa, May 1998.

Naturstoffe – Quelle neuer Produkte für Pflanzenschutz und Pharma

Isolierung von biologisch aktiven Substanzen aus Pilzen

Ein Forschungsteam im Arbeitskreis der Organischen Chemie (AK Krohn, Fachbereich Chemie und Chemietechnik) beschäftigt sich mit der Isolierung und Strukturaufklärung von biologisch aktiven Stoffen aus Pilzen. Ziel ist dabei die Auffindung neuer Leitstrukturen (1) für Pharma und Pflanzenschutz. „Am Anfang steht die Leitstruktur" lautet ein Credo der modernen Forschung für Arznei- und Pflanzenschutzmittel [1]. Das gilt auch heute noch, obwohl sich die Strategie zur Auffindung neuer Präparate im Laufe der letzten Jahre mehrfach gewandelt hat.

Trends bei der Leitstrukturfindung (1)

Bis in die Achtzigerjahre hinein war die chemische Synthese einzelner Verbindungen oder die Isolierung bestimmter Naturstoffe und deren aufwändiges Testen an ganzen Organismen die Regel. Sehr oft stand der Zufall Pate bei der Entwicklung neuer bioaktiver Verbindungen. Ende der Achtzigerjahre kam die Vision auf, durch Analyse der Rezeptor-Inhibitor-Modelle endlich zu einem rationalen „Denovo-Design" von Wirkstoffen zu gelangen. Mit der rasanten Entwicklung der Rechenleistung der modernen Computer und ihren Möglichkeiten zur grafischen Darstellung kam das „Molecular Modeling" in Mode. Nach diesen Modellen wurde der Wirkstoff, der meist viele flexible dreidimensionale Gestalten (Konformationen) einnehmen kann, in das aktive Zentrum von Proteinen eingepasst. Abbildung 1 zeigt die Einlagerung von Galanthamin in die Acetylcholinesterase (Arbeitsgruppe Prof. Gregor Fels).

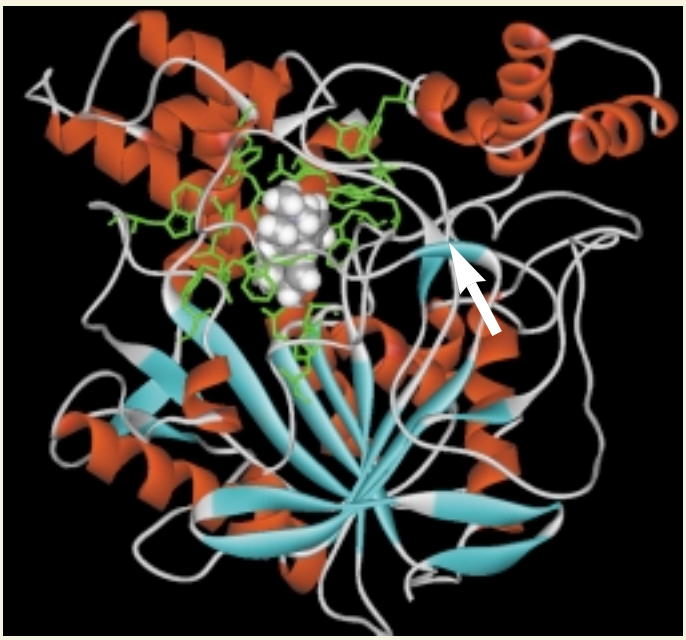


Abb. 1: Einlagerung von Galanthamin in die Acetylcholinesterase (Berechnung durch Prof. Gregor Fels).

Prof. Dr. Karsten Krohn ist seit 1991 als Fachvertreter für Organische Chemie im Fachbereich 13 an der Universität Paderborn tätig. Seine Arbeitsgebiete umfassen die chemische Synthese von Chinon-Antibiotika, die Isolierung von Wirkstoffen aus Pilzen, Nachwachsende Rohstoffe aus Zucker und Methoden zur Reduktion und Oxidation.

Eine grundlegende Rolle bei diesem Ansatz spielte dabei die Anfang des 20. Jahrhunderts von dem Chemiker Emil Fischer aufgestellte Schloss-Schlüssel Hypothese [2]. Emil Fischer war der erste deutsche Wissenschaftler, der für seine überragenden Leistungen 1902 den Nobelpreis für Chemie bekam. Nach dieser Vorstellung passt der Wirkstoff wie ein Schlüssel in das Schloss (das „aktive Zentrum") des körpereigenen „Rezeptors". Die Rezeptorproteine übernehmen im komplexen Zusammenspiel der Körperfunktionen unterschiedliche Regelfunktionen, die durch das Wechselspiel mit niedermolekularen Wirkstoffen beeinflusst werden.

Obwohl in den nächsten Jahrzehnten eine gewisse Ernüchterung in der rationalen Leitstruktur-Auffindung durch Molecular Modeling aufkam, erfüllt diese Methode doch nach wie vor eine überaus wichtige Funktion bei der Leitstruktur-Optimierung! Kennt man die biologischen Aktivitäten einer Reihe ähnlicher Verbindungen, so lassen sich die physikochemischen Parameter einzelner Gruppen in den Molekülen über die QSAR-Methode miteinander korrelieren (Quantitative Structure-Activity Relationship). Neuerdings kann man sogar die dreidimensionalen Feldeigenschaften mit einbeziehen (CoMFA Comparative Molecular Field Analysis).

Was ist der heutige Trend in der Auffindung neuer Leitstrukturen? Interessanterweise hat man dabei bewusst dem Zufall wieder eine größere Rolle zugestanden! Mehrere Entwicklungen haben dazu beigetragen. Durch den Fortschritt in der Biologie und Biotechnologie einerseits und der computergesteuerten Automatisierung andererseits ist man heute in der Lage, eine sehr hohe Zahl von Verbindungen in kurzer Zeit zu testen (High Throughput Screening). Allerdings nicht am kompletten Organismus sondern an ganz speziellen biochemischen „Targets", die im

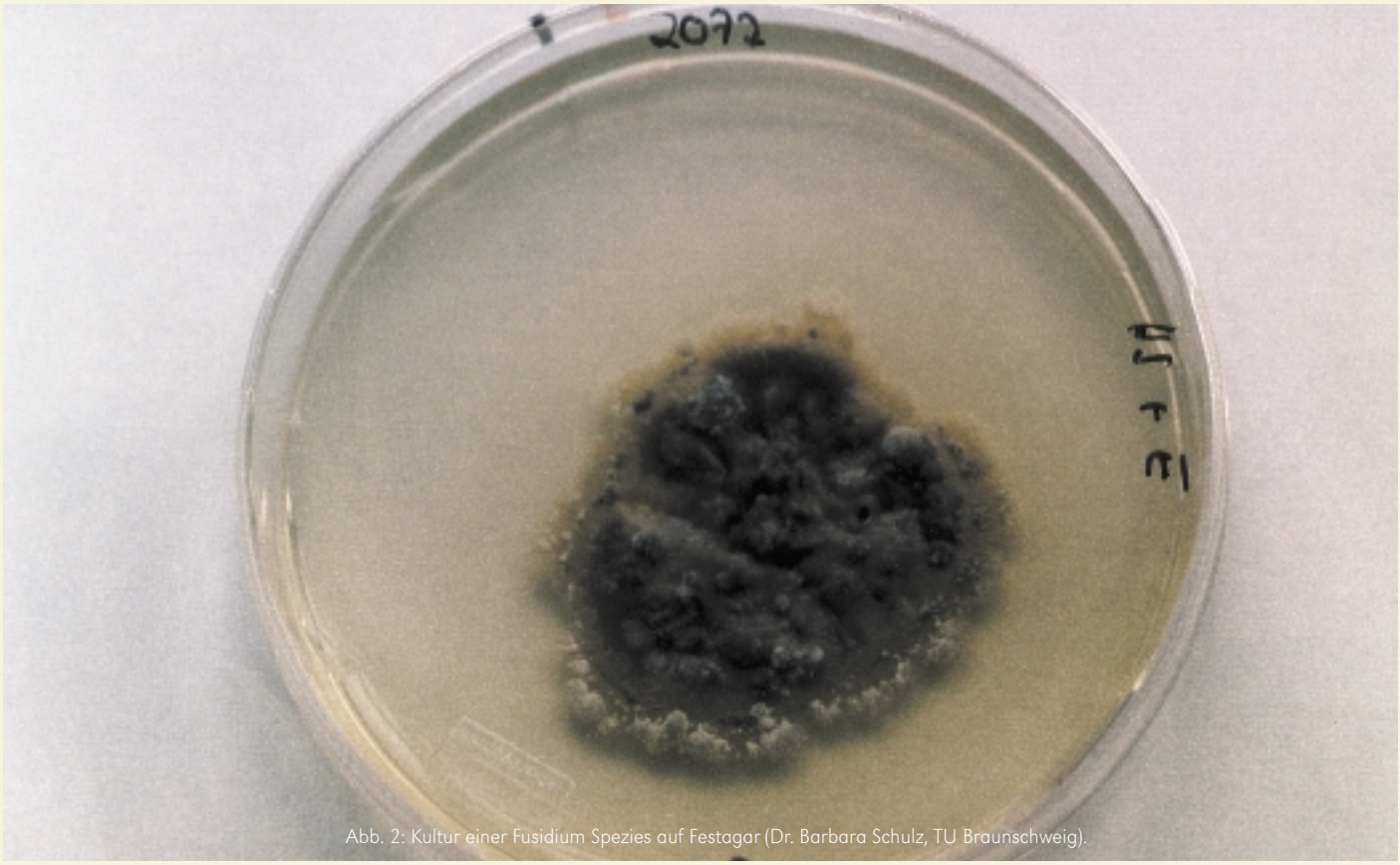


Abb. 2: Kultur einer Fusidium Spezies auf Festagar (Dr. Barbara Schulz, TU Braunschweig).

Krankheitsbild eine Rolle spielen und die eine Analogie zu dem oben beschriebenen Schloss-Schlüssel-Modell zeigen. Findet man in diesen Tests auffällige Verbindungen, die als Leitstrukturen dienen können, so müssen sie natürlich nachfolgend optimiert und gründlich auf ihre Wirkung und Toxikologie geprüft werden. Dennoch trägt dieses Verfahren heute auch dazu bei, die Zahl der notwendigen Tierversuche auf ein unbedingt notwendiges Mindestmaß zu beschränken.

Aber wie kommt man zu so vielen Verbindungen, die in das High Throughput Screening eingeschleust werden können? In der Tat ist die Bereitstellung von interessanten Verbindungen zu einem Nadelöhr geworden. Die Synthesechemiker haben neue effiziente Verfahren zur so genannten „parallelen Synthese“ und zur „kombinatorischen Synthese“ entwickelt, mit denen oft tausende von Verbindungen (jedoch meist in kleinen Mengen oder in unreiner Form) gleichzeitig hergestellt werden können. Ein Merkmal der neuen Testsysteme ist nämlich auch der geringe Substanzbedarf für die einzelnen Tests. Heute können mit wenigen Milligramm Substanz eine Vielzahl von „in vitro“-Tests durchgeführt werden. Die großen pharmazeutischen Firmen haben auch riesige „Substanzbibliotheken“ auf Lager, die aus früheren Synthesen der Firma oder aus Zukäufen von dritter Seite (auch von Hochschulen) stammen. Da die Biochemie immer neue Testsysteme bereitstellt, werden auch bekannte Verbindungen erneut in diese Tests eingeschleust. Man sagt im Übrigen voraus, dass durch die Kenntnis des menschlichen Genoms immer mehr und auch immer bessere Targets für bestimmte pathologische Erscheinungsbilder entwickelt werden.

Die Rolle der Naturstoffe

Welche Rolle spielen nun die Naturstoffe bei der Leitstruktursuche? Von der Zahl her nehmen sie zunächst einen eher bescheidenen Platz ein. Genau 16 827 004 organische und anorganische Verbindungen (ohne Sequenzen) waren am 15. Juni 2000 strukturell genau beschrieben und in der größten Datenbank der Welt, den „Chemical Abstracts“, erfasst. Die umfassende Datenbank für Naturstoffe von Chapman and Hall beinhaltet aber in der neuesten Version gerade einmal 145 000 genau analysierte Naturstoffe!

Und dennoch spielen Naturstoffe in Pharma und im Pflanzenschutz nach wie vor bei der Suche nach Leitstrukturen eine überragende Rolle. Die zufällige Entdeckung des Penicillins 1928 durch Alexander Fleming und die daraus bis heute ständig weiterentwickelten hochwirksamen β-Lactam-Antibiotika zählen zu den größten Erfolgen der Medizingeschichte [3]. Die Besorgnis erregende Entwicklung von Resistenzen z.B. durch den Eitererreger *Staphylococcus aureus* oder durch *Enterococcus faecalis* zwingt aber dazu, bei den Bemühungen um die Entwicklung neuer Antibiotika nicht nachzulassen [4]. Neben den Antibiotika stammt auch ein großer Teil der mit zunehmendem Erfolg eingesetzten Antitumormittel aus natürlichen Quellen (z.B. Anthracycline, Taxol). Auch hier geht die Suche nach noch besseren Medikamenten mit unverminderter Intensität weiter.

Ein weiteres prominentes Beispiel aus der modernen Medizin ist das Cyclosporin, ohne das die gesamte Transplantations-Chirurgie wegen der Abstoßung fremder Organe gar nicht möglich wäre. Rapamycin und KF-506 sind zwei weitere Naturstoffe, die für die Entwicklung neuer Immunsuppressiva bereitstehen [5].

Im Pflanzenschutz werden als Insektizide heute noch „Pyrethroide“ verwendet, Verbindungen, die ursprünglich aus den Blüten von Pyrethrum-Arten gewonnen wurden. Als jüngstes Erfolgsbeispiel seien die Abkömmlinge der natürlichen Pilzmetabolite (die Strobilurine) genannt, die chemisch zur Gruppe der β-Methoxyacrylate gehören. Große Firmen haben daraus Präparate gegen pflanzlichen Pilzbefall von bisher nicht erreichter Umwelt-

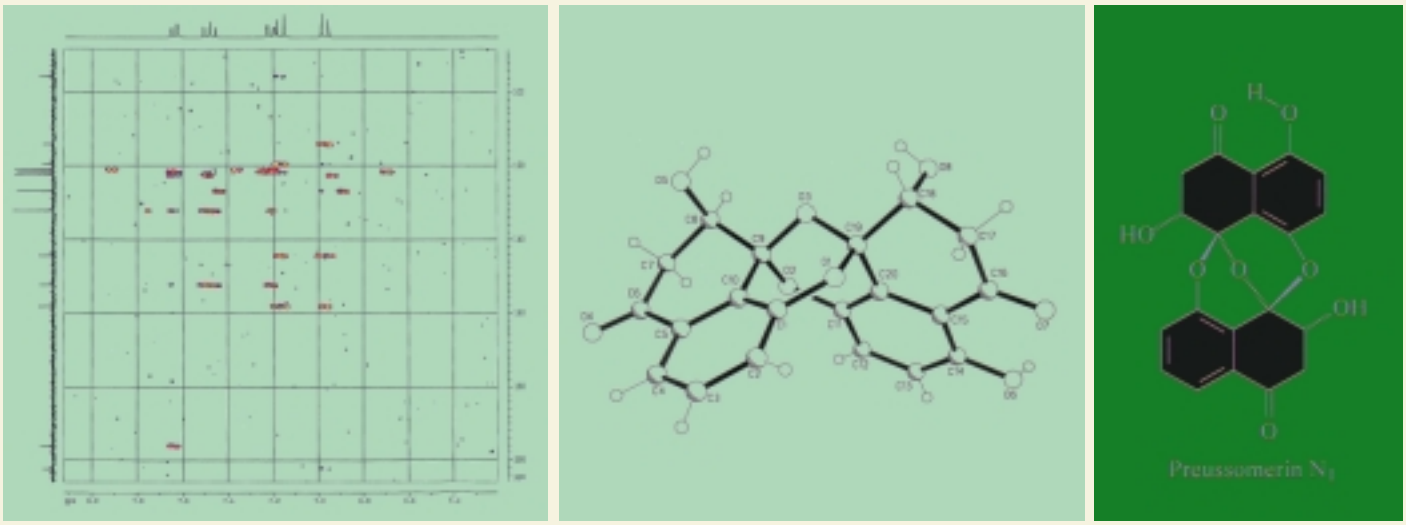


Abb. 3: Zweidimensionales NMR-Spektrum (Dr. Victor Wray, GBF Braunschweig), Darstellung der Röntgenstrukturanalyse (Dr. Ulrich Flörke) und chemische Formel eines neuen Preussomerins (Natalie Root).

verträglichkeit entwickelt (BASF: Kresoxim-methyl; ICN: Azoxystrobin) [6].

Der Erfolg der Naturstoffe hat vor allem zwei Gründe. Einmal sind die vorhandenen Substanzbibliotheken und auch die durch kombinatorische Synthese hergestellten Verbindungen in Ihrer strukturellen Vielfalt sehr viel geringer einzuschätzen als Naturstoffe. Es gibt heute eine Reihe von mathematischen Modellen, welche die „strukturelle Diversität" genau beschreiben. Die Natur ist in der Lage, mit erstaunlich wenigen Biosynthesewegen eine große Strukturvielfalt bereitzustellen. Und erst ein kleiner Teil des vorhandenen Schatzes ist gehoben. Der überwiegende Teil der Pilze, Mikroorganismen, Pflanzen, Insekten (gar nicht zu reden von der Vielfalt der Meeresorganismen) ist auf Inhaltsstoffe überhaupt noch nicht genau untersucht!

Als zweiter Grund wird angenommen, dass die Natur in den Milliarden Jahren ihrer Evolution gerade auch die niedermolekularen Naturstoffe („Sekundärmetabolite") für den Überlebenskampf entwickelt hat. Man kann auch sagen, dass diese Stoffe für Interaktion mit anderen Bestandteilen der Biosphäre im Laufe der Evolution „trainiert" sind. Diese Annahme hat natürlich recht hypothetischen Charakter und bei den meisten Sekundärmetaboliten ist der genaue Zweck auch nicht bekannt. Bei anderen Bestandteilen lässt sich aber durchaus eine plausible Vermutung für ihre Funktion anführen.

Neue Naturstoffe aus Pilzen

Im Fachbereich Chemie und Chemietechnik beschäftigen wir uns in der Organischen Chemie u.a. mit der Auffindung und chemischen Charakterisierung von Verbindungen aus Pilzen und neuerdings auch aus solchen Pflanzen, die in der traditionellen Medizin in Indien oder Indonesien eine Rolle spielen.

Wie gehen wir im Arbeitskreis bei der Suche nach neuen Leitstrukturen aus natürlichen Quellen vor? Unsere Kooperationspartner von der Technischen Universität Braunschweig, die Mikrobiologen und Mykologen Prof. Hans-Jürgen Aust, Dr. Sigfried Dräger und Dr. Barbara Schulz, testen die Extrakte kleiner Pilzkulturen zunächst mit eine Reihe von Testorganismen auf Wirksamkeit gegen Bakterien, Pilze und Keimlinge. Abbildung 2 zeigt die Kultur einer Fusidium Spezies auf einem Festagar.

Werden in diesen ersten biologischen Tests interessante Wirkungen festgestellt, so erhält der zweite Kooperationspartner, die BASF AG (die Knoll AG für die Pharmatests) einen Extrakt, der dann mit den oben beschriebenen Targets getestet wird. Nur wenn in beiden Testsystemen interessante Wirkungen festgestellt werden, machen sich die Braunschweiger Kollegen an die recht aufwändige Kultivierung größerer Kulturen. Die meist nicht in flüssigem Medium sondern auf Agar angezogenen Kulturen brauchen oft mehrere Monate zur Produktion der begehrten Sekundärmetabolite. Durch eine chromatographische Analyse wird der Gehalt an interessanten Stoffen laufend kontrolliert. Schließlich wird ein Extrakt des Kulturmediums nach Homogenisierung in einem organischen Lösungsmittel hergestellt, der dann unserer Arbeitsgruppe übergeben wird. Es folgt die mühsame Aufreinigung des oft sehr komplexen Gemisches in die einzelnen Komponenten. Zur Trennung und Aufreinigung steht ein ganzes Arsenal von so genannten „chromatographischen" Verfahren zur Verfügung. Es gibt verschiedenen Materialien, die meist in eine Säule gepackt zur Auftrennung von Substanzgemischen genutzt werden. Bei geringeren Mengen kann das Trennmaterial (z.B. getrocknetes Kieselgel mit einer enorm großen inneren Oberfläche) auch auf eine Platte aufgegeben werden. Leichter flüchtige Verbindungen können auch über die Gasphase aufgetrennt werden (Gaschromatographie). Die Regel ist, dass die Ermittlung der genauen räumlichen „Struktur" einer Verbindung nur mit hochreinen Verbindungen gelingt. Im Gegensatz zur chemischen Synthese, bei der man in jeder Stufe meist nur eine begrenzte Zahl von Verbindungen erhält, liegt in Extrakten aus natürlichen Quellen meist ein äußerst komplexes Gemisch vor. Außerdem kennt man bei Synthesen auch die Struktur der Vorstufen, von denen man ausgegangen ist. Bei Naturstoffen können die Verbindungen aus den verschiedensten Substanzklassen stammen, über die man zunächst einmal gar nichts weiß.

Vom reinen Stoff zur Struktur

Nehmen wir an, die Auftrennung eines Rohextrakts in mehrere reine Verbindungen ist nach intelligenter Anwendung der Trennverfahren gelungen, was oft einige Monate in Anspruch nehmen kann. Im nächsten Schritt erfolgt dann die Messung einer Reihe wichtiger physikalischer Parameter und der so genannten spektra-

len Daten, aus denen der Experte nach genauer Analyse die Struktur ermitteln kann. Der Fachbereich Chemie ist mit allen dafür wichtigen Methoden ausgestattet. Aber falls nur geringe Mengen (unter 3 mg) einer reinen Verbindung vorliegen, sind wir doch auf die Hilfe von Kollegen an anderen Standorten angewiesen, die bestimmte Messungen vornehmen. Das wichtigste Werkzeug ist dabei die kernmagnetische Resonanzspektroskopie (NMR, Nuclear Magnetic Resonanz) und die Röntgenstrukturanalyse. Abbildung 3 zeigt ein so genanntes zweidimensionales heteronuclear-korreliertes NMR-Spektrum und das Ergebnis der Röntgenstrukturanalyse eines Naturstoffs, dessen Struktur wir kürzlich aufgeklärt haben. Eine Röntgenstrukturanalyse ist leider nur von kristallinen Verbindungen möglich.

In dieser Phase der Untersuchungen bangen Mitarbeiter und Arbeitsgruppenleiter darum, dass der von ihnen untersuchte Stoff auch neu ist. Parallel zu der Erfassung und Auswertung der Daten beginnt eine intensive Suche in Datenbanken. Man kann dabei schon Bruchteile der physikalischen und spektroskopischen Daten eingeben und mit der Suche beginnen. Die Arbeit an schon bekannten Verbindungen gilt heute als größtenteils verlorene Zeit. Neue Leitstrukturen sind gefragt! Deshalb gilt es, möglichst rasch herauszufinden, ob ein Naturstoff schon bekannt ist oder nicht. Dank der guten Ausstattung mit Computern und Programmen im Fachbereich konnte die Zeit zur Identifizierung bekannter Verbindungen drastisch verkürzen werden.

Die Auswahl der Pilze ist bei der Erfolgsquote auf der Jagd nach neuen Strukturen von großer Wichtigkeit. Unsere Kooperationspartner in Braunschweig geben sich die allergrößte Mühe, die Pilze vor der chemischen Analyse genau zu bestimmen. Als Quellen für die erfolgreiche Struktursuche gelten Pilze aus extremen Standorten – wie schwermetallhaltige Abraumlager, ja selbst wüstenartige Gegenden oder gar die antarktischen Räume! Als ergiebige Quelle haben sich auch die so genannten endophytischen Pilze erwiesen, die in den meisten Pflanzen wachsen, ohne dass der Betrachter irgendeinen Pilzbefall erkennen kann. Diese eigenartige Symbiose mit Pflanzen führt oft zur hohen Produktion von Sekundärmetaboliten. Außerdem ist diese Klasse von Pilzen nicht so leicht zu kultivieren (sie wachsen z.B. oft nur auf Festagar) und sind deshalb oft noch nicht auf Inhaltsstoffe untersucht. Dennoch kommt es immer wieder vor, dass schon

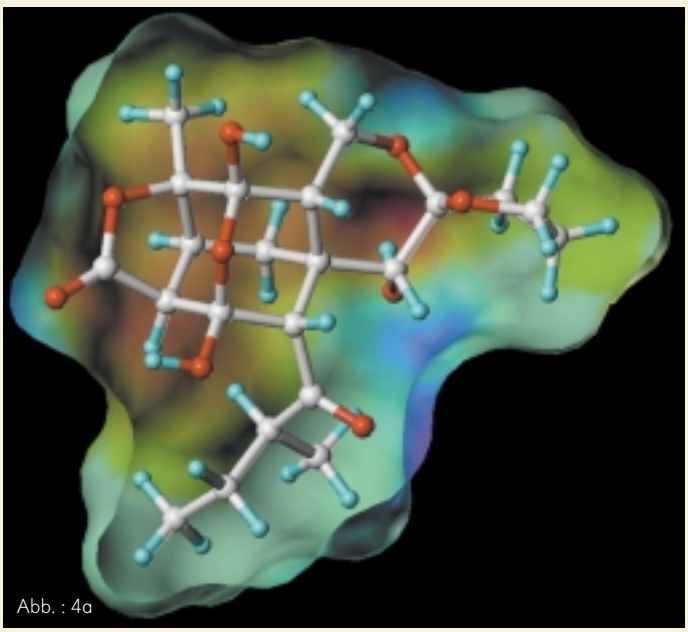


Abb. : 4a

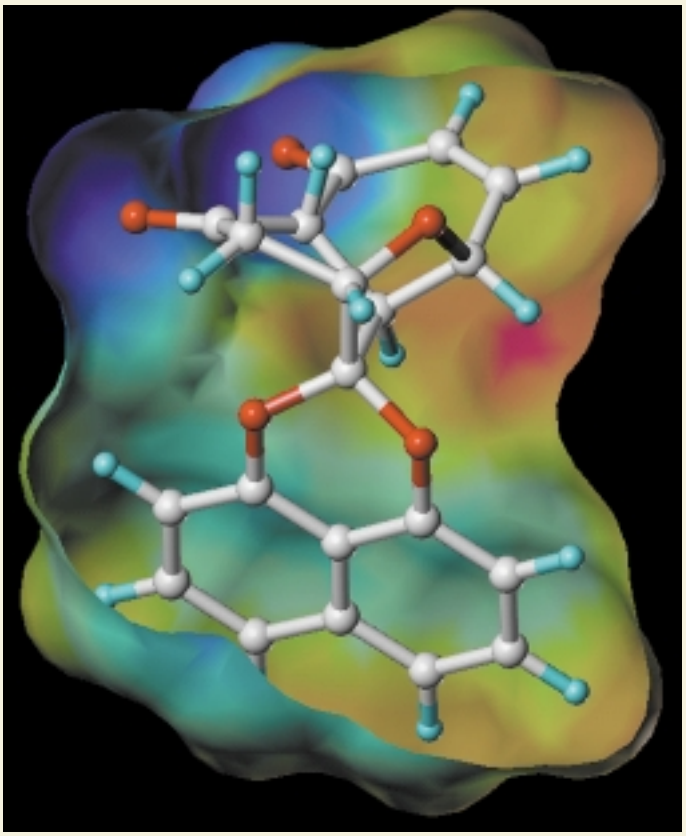


Abb. 4b: Berechnungen (Prof. Gregor Fels) der Konformation eines Palmarumycins und von Fusidilacton III (Abb. 4a: Dr. Karl-Heinz Drogies).

bekannte Stoffe auch aus noch nicht untersuchten Pilzstämmen aufgefunden werden. Bekanntlich verfügen Mikroorganismen ja über eine Reihe von Möglichkeiten, genetische Information rasch auszutauschen. So auch über die Synthese bestimmter Sekundärmetabolite, die deshalb oft in verschiedenen Organismen in gleicher Form auftauchen. Ein ähnliches Phänomen trägt ja auch zu der oben schon erwähnten Resistenzbildung bei Bakterien gegen Antibiotika bei.

Dennoch haben wir im Arbeitskreis auch eine große Zahl bisher in der Literatur noch nicht beschriebener Verbindungen aufgefunden und in ihrer chemischen Struktur aufgeklärt. In vielen Fällen handelt es sich dabei um Variationen von Stoffen, die in ähnlicher Form schon bekannt sind. Das sind dann zwar publikationswürdige neue Naturstoffe, deren Wirkung auch von der BASF getestet wird, aber von einer neuen Leitstruktur kann man in diesen Fällen dann leider nicht sprechen. Es gab aber auch eine Reihe von Glücksfällen, in denen Verbindungen mit ganz neuem Grundgerüst aufgefunden wurden. Dazu gehören z.B. die Verbindungen aus *Coniothyrium palmarum*, die wir „Palmarumycine“ genannt haben. Einige dieser Verbindungen sind Inhibitoren der so genannten „Ras Farnesyltransferase“, die in einem modernen Ansatz der Antitumorforschung eine große Rolle spielen [7]. Eine weitere Verbindungsklasse mit völlig neuartigem Grundgerüst sind die Fusidilactone. Die dreidimensionale Gestalt („Konformation“) und die farbige Darstellung der Ladungsdichte an der Oberfläche eines Palmarumycins und des Fusidilactons III sind in Abbildung 4 gezeigt (Berechnungen von Prof. Gregor Fels).

Danksagung: Meinen Mitarbeitern Natalie Root und Klaus Steingröver, den Kooperationspartnern (Prof. Hans-Jürgen Aust, und Dr. Sigfried Draeger, Dr. Barbara Schulz, TU Braunschweig)

und Dr. Joachim Rheinheimer (BASF) danke ich herzlich für die Unterstützung bei der Forschung und meinem Kollegen Prof. Gregor Fels bei der Berechnung einiger Molekülstrukturen.

(1) Leitstruktur: Bezeichnung für meist niedermolekulare Wirkstoffe aus der Natur oder der Synthesechemie, deren strukturelle Neuartigkeit bezüglich einer bestimmten Indikation die Basis zu vertiefenden Studien bildet. Ziel ist eine Wirkstoffentwicklung zum Pharmakon oder zur Agrarchemikalie. In der Regel werden mit Suchsystemen im Wirkstoff-Screening zunächst so genannte Hits aufgefunden, deren biologische Aktivität durch weitergehende biologische Austestung (Sekundärtestung) und in-vivo-Studien weiter zu verifizieren sind. Hieraus abgeleitete Leitstrukturen werden u.a. zur Optimierung von Struktur-Wirkungsbeziehungen chemisch gezielt variiert, um einen möglichst erfolgreichen Entwicklungskandidaten auszuwählen.

Literatur

[1] J. Stetter, F. Lieb: Innovation im Pflanzenschutz: Trends in der Forschung. Angew. Chem. 2000, 112, 1792-1812.

[2] F. W. Lichtenthaler: Hundert Jahre Schlüssel-Schloss-Prinzip: Was führte Emil Fischer zu dieser Analogie, Bd. 106, Angew. Chem. 1994, S. 2456-2467.

[3] J. Emsley: Sonne, Sex und Schokolade. Chemie im Alltag II, Wiley-VCH, Weinheim 1998, S. 83 ff.

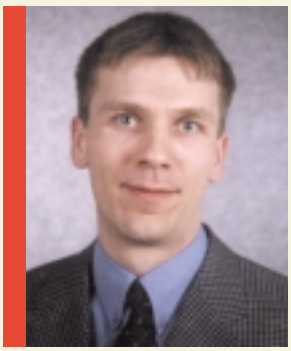
[4] S. B. Levy: Antibiotikaresistenz: eine globale Herausforderung. Spektrum der Wissenschaft 1998, 5, 34-41.

[5] B. Werth: Das Milliarden-Dollar-Molekül, Wiley-VCH, Weiheim 1996.

[6] H. Sauter, W. Steglich, T. Anke: Strobilurine: Evolution einer neuen Wirkstoffklasse. Angew. Chem. 1999, 111, 1416-1438.

[7] D. M. Leonard: Ras Farnesyltransferase: A New Therapeutic Target. J. Med. Chem. 1997, 40, 2971-2990.

Natalia Root arbeitet im Arbeitskreis von Prof. Krohn an ihrer Dissertation über biologisch aktive Sekundärmetabolite aus endophytischen Pilzen.



Klaus Steingröver arbeitet im Arbeitskreis von Prof. Krohn an seiner Dissertation über biologisch aktive Sekundärmetabolite aus endophytischen Pilzen.

Elektronische Kurvenscheiben

Eine klassische Antriebsaufgabe mit neuen Technologien gelöst

Überall in der industriellen Produktion werden Bewegungen unterschiedlichster Komplexität benötigt. Forderungen nach immer hochwertigeren und flexibel beeinflussbaren Bewegungsabläufen machen den Einsatz intelligenter Antriebstechnik erforderlich. Dabei spielt insbesondere die Bewegungsplanung und ihre steuerungsinterne Erzeugung eine entscheidende Rolle.

Immer dort, wo zyklische Arbeitsbewegungen im Maschinenbau benötigt werden, sind überwiegend mechanische Kurvenscheiben im Einsatz. Oftmals werden sie mit den so genannten Gelenkgetrieben kombiniert, um besondere Arbeitsbewegungen zu erzeugen. All diesen Lösungen ist das Ziel gemeinsam, eine ungleichförmig-übersetzte Bewegung zu erzeugen. Diese kann man sich auch so vorstellen, dass ein Zahnradgetriebe nach einer vorgeschriebenen Gesetzmäßigkeit sein ansonsten konstantes Übersetzungsverhältnis permanent ändert. Es entsteht eine permanent beschleunigte und verzögerte Bewegung, die oft auch Stillphasen beinhaltet, wobei die eigentliche Antriebsbewegung von einem mit konstanter Drehzahl betriebenen Antrieb bereitgestellt wird. Bei komplexeren Anlagen und Maschinen treibt dieser „Haupt"-Antrieb gleich mehrere Kurvenscheiben an, sodass alle Arbeitsbewegungen synchron zum Hauptantrieb erfolgen. Wird der Hauptantrieb eingeschaltet, so bewegen sich alle Arbeitsorgane zwangsläufig lagesynchron zueinander.

Prof. Dr.-Ing. Jürgen Bechtloff ist Professor für Mess-, Steuerungs- und Regelungstechnik im Fachbereich 11/Maschinenbau – Datentechnik der Universität Paderborn, Fachhochschulabteilung Meschede. Seine Schwerpunkte sind Messdatenverarbeitung, Anwendung der Fuzzy-Logic in der Regelungstechnik sowie Antriebsregelungen.

Dabei sind Verstellmöglichkeiten bezüglich der Relativlage verschiedener Arbeitsbewegungen zueinander im laufenden Betrieb gar nicht oder nur mit erheblichem mechanischen Aufwand möglich. Solche Eigenschaften sind gerade im Einrichtbetrieb einer Anlage oder beim Optimieren des Prozesses sehr wichtig. Ebenso ist ein schrittweiser Einrichtbetrieb, in dem nur einzelne Antriebsgruppen laufen und andere sich in einer Parkposition befinden, nur bedingt und ebenfalls mit einem hohen mechanischen Aufwand realisierbar.

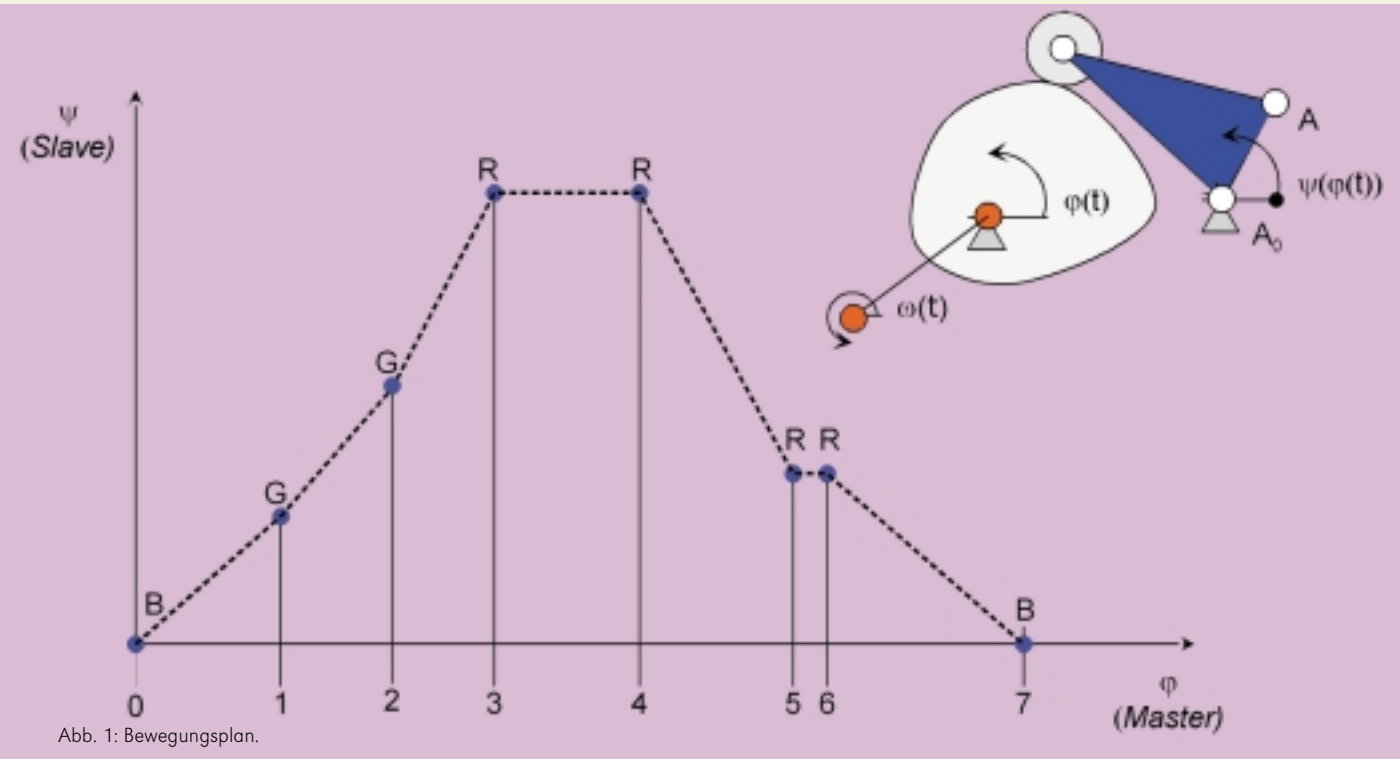


Abb. 1: Bewegungsplan.

Um nun eine höhere Flexibilität hinsichtlich der Planung, Einstellbarkeit bzw. Justage und einer dynamischen Bewegungsmodifikation zu erreichen, drängt der Sondermaschinen- und Verpackungsmaschinenbau verstärkt auf die Substitution mechanischer Gelenk- und Kurvengetriebe durch „elektronische Kurvenscheiben“. Dabei spielt die Integration einer solchen elektronischen Lösung in eine automatisierungstechnische Gesamtlösung eine entscheidende Rolle.

Einsatz geregelter Antriebe in der Automatisierungstechnik

In der Automatisierungstechnik werden zunehmend geregelte Antriebe (Servoantriebe) zur Realisierung hochdynamischer lagegeregelter Bewegungen eingesetzt. Die Bewegungssteuerung wird mit speziellen im Allgemeinen dezentral ausgeführten „Spezial“-Steuerungen realisiert. Daher wird hinsichtlich ihrer Komplexität der einzelnen Steuerungsarten unterschieden in NC-Steuerungen (Numerical Control → Werkzeugmaschinen), Positionier-Steuerungen und Sonder-Regler (z.B. Synchrongleichlaufregler, Fliegende Säge etc., *elektronische Kurvenscheibe*).

Elektronische Kurvenscheibe

Eine elektronische Kurvenscheibe stellt aus Sicht der Antriebssteuerung ein elektronisches Getriebe mit einer variablen Übersetzung dar. In Abhängigkeit des momentanen Lageistwertes des Leitantriebs bzw. Masters wird das Übersetzungsverhältnis permanent verändert. Eine besonders effektive Planung der Bewegung der Folgeachse bzw. des Slaves lässt sich durch Einsatz der sich in der Praxis vielfach bewährten Entwurfsmethoden für mechanische Kurvengetriebe erzielen. Diese Methode beruht auf der Idee, dass eine Bewegung dann kinematisch und dynamisch am günstigsten verläuft, wenn möglichst wenig Zwänge auf die Bewegung ausgeübt werden. Daher beruht die Methode der Bewegungspläne auf der Vorgabe weniger Bewegungsabschnitte und zugeordneten Bewegungszuständen.

In Anlehnung an die VDI-Richtlinie 2143 [1] wird die Abhängigkeit des Slaves vom Master in einem Bewegungsplan dargestellt. Die Bewegung wird in einzelne Abschnitte unterteilt. Jeder Übergang von einem Bewegungsabschnitt in den nächsten wird durch seinen Bewegungszustand charakterisiert. Es wird unterschieden in:

Zustand	V	A	
Rast	v=0	a=0	Slave befindet sich momentan in Ruhe
Geschwindigkeit	v≠0	a=0	Slave bewegt sich mit momentan konstanter Geschwindigkeit
Umkehr Bewegung	v=0 v≠0	a≠0 a≠0	Slave kehrt seine Bewegung um allgemeiner Bewegungszustand

Die Werte von Geschwindigkeit und Beschleunigung charakterisieren also den momentanen Bewegungszustand. Grafisch ergibt sich z.B. der in Abbildung 1 gezeigte Bewegungsplan:

Die eigentliche Abtriebsbewegung ergibt sich dadurch, dass durch Interpolation einer abschnittsweisen Funktion ein resultierender „glatter“ Verlauf ensteht [2]. Mit „glatt“ wird ein Verlauf mit einer Stetigkeit bis mindestens zur zweiten Ableitung der gesamten abschnittsweise definierten Funktion verstanden (Gewährleistung der Stoß- und Ruckfreiheit). Das hier entwickelte Verfahren stammt im Ursprung aus dem Verfahren der lokalen

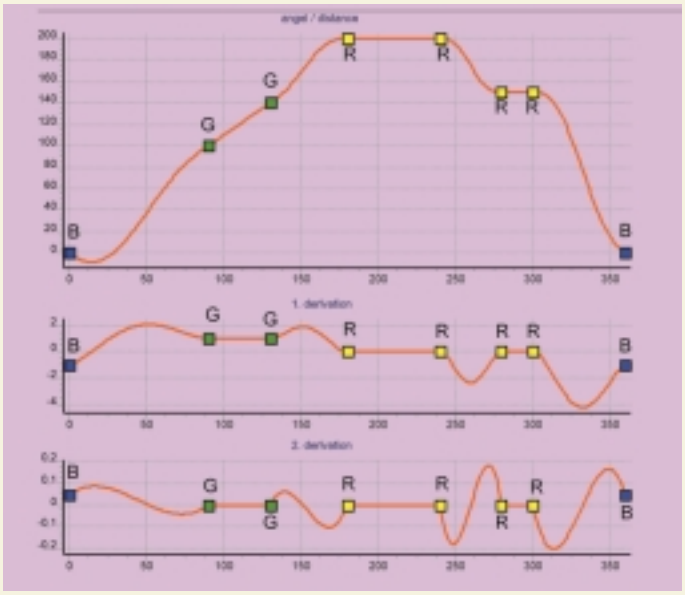


Abb. 2: Übertragungsfunktionen nullter bis zweiter Ordnung.

Spline-Interpolation für die Generierung von Bewegungstrajektorien für Industrieroboter in Echtzeit [3]. Die ursprüngliche Spline-Bedingung wird nun lediglich bei den B-Punkten vollständig zur Berechnung der optimalen ersten und zweiten Ableitung genutzt, bei G-Punkten wird nur die erste und beim U-Punkt die zweite Ableitung nach dem Spline-Kriterium genutzt, die andere Ableitung wird gemäß der oben stehenden Tabelle zu Null gesetzt. Wird der in Abbildung 1 grafisch dargestellte Bewegungsplan berechnet, so ergeben sich die in Abbildung 2 gezeigten Übertragungsfunktionen bis zur 2. Ordnung:

Hier kann festgestellt werden, dass die Beschreibung einer allgemeinen Bewegungsaufgabe mithilfe des Bewegungsplanes folgende Vorteile aufweist:

- Kompakte Beschreibung auch komplexer Bewegungsaufgaben
- Kinematischer Zugang zur Bewegungsdefinition.

Realisierung

Die Umsetzung in der Praxis orientiert sich in einem ersten Schritt an der Steuerungsarchitektur [4]. Der klassische Ansatz in der Automatisierungstechnik beinhaltet die Realisierung in einem zentralen Steuerungsrechner, der je nach Leistungsfähigkeit über mehrere parallel arbeitende Zentraleinheiten für die unterschiedlichen Aufgaben wie Ablaufsteuerung, Kommunikation und Bewegungssteuerung verfügen kann. Vorteil solcher zentraler Lösungen ist die direkte Kopplung von Ablaufsteuerung und Bewegungserzeugung, wie sie oft in CNC-Steuerungen realisiert wird. Daneben besteht auch die Möglichkeit, Multi-Achsen-Bewegungen ideal miteinander z.B. mit einer elektronischen Kurvenscheibenfunktion koppeln zu können.

In der Antriebstechnik geht der Trend eindeutig hin zu digitalen Verstärkern, die mittels Feldbussystemes wie Profibus DP oder CAN-Bus an eine übergeordnete Ablaufsteuerung (SPS – speicherprogrammierbare Steuerung) angekoppelt werden. Über den Feldbus lassen sich Programme bzw. Sollwerte an den Antrieb übertragen. Neben einer Positionierlogik lassen sich in diesen Antrieben weitere Technologiefunktionen, wie z.B. elektronisches Getriebe, fliegende Säge oder auch die elektronische Kurvenscheibe realisieren.

Unabhängig von der Steuerungsarchitektur, in der die Technologiefunktion realisiert wird, gibt es zwei Alternativen, wie die

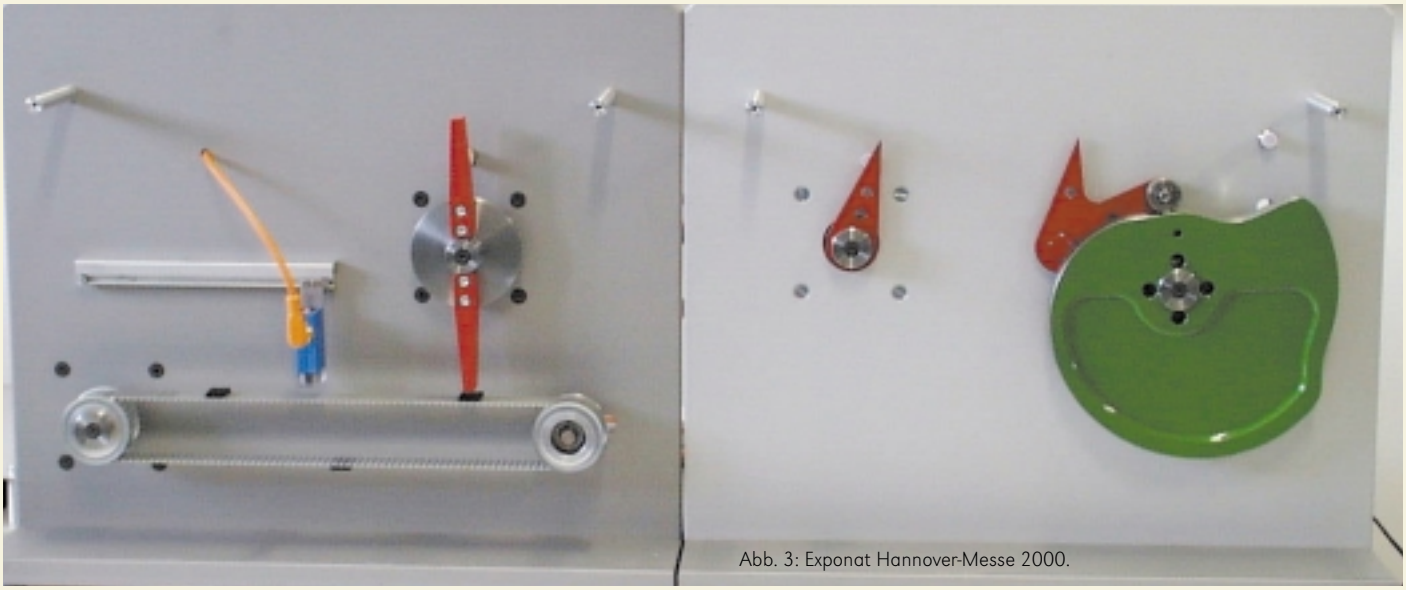


Abb. 3: Exponat Hannover-Messe 2000.

Echtzeitverarbeitung zu realisieren ist:

a) Lineare oder kubische Interpolation einer vorausberechneten Interpolationstabelle, die aus dem vorgegebenen Bewegungsplan generiert wurde.
b) Direkte Interpolation des Bewegungsplanes mit Polynomen fünften Grades zur Wahrung der Stoß- und Ruckfreiheit der resultierenden Bewegung.

Die Variante a) hat den Vorteil, eine von der Bewegungssteuerung unabhängige Schnittstelle für die Bewegungsplanung der elektronischen Kurvenscheibe zur Verfügung zu haben, sodass auch dem geregelten Antrieb nachgeschaltete Mechanismen in ihrem nichtlinearen Übertragungsverhalten berücksichtigt werden können. Nachteilig ist zum einen das hohe Datenaufkommen, um die Fehler der Echtzeit-Tabelleninterpolation zu verringern und zum anderen, dass der zugrunde liegende Bewegungsplan während der Antriebsbewegung nicht mehr z.B. durch Sensorsignale veränderbar ist.
Die Variante b) zeichnet sich durch die Kompaktheit der kinematischen Beschreibungsweise durch den Bewegungsplan aus und interpoliert die Bewegung exakt. Lediglich bei Anwendungen, bei denen Mechanismen mit nichtlinearem Übertragungsverhalten nachgeschaltet sind, muss auf eine Vielpunktdarstellung ausgewichen werden, was aber prinzipiell ohne Erweiterungen des Konzeptes möglich ist.
Um eine vergleichende Analyse verschiedener Realisierungen der Technologiefunktion „Elektronische Kurvenscheibe“ durchführen zu können, wurde ein Laborsteuerungssystem entwickelt. Auf der Basis eines Industrie-PCs auf Hutschiene mit den zur Ansteuerung industrieller Servoantriebe notwendigen Baugruppen wie schnelle D/A-Wandler, Zählerbaugruppen usw. wurde ein kompletter Bewegungskern, eingebettet in ein Echtzeitbetriebssystem, entwickelt. Durch seinen modularen Aufbau und die Unterteilung in Satzvorbereitungs-, Interpolations- und Lageregel-Task lassen sich auch Aspekte der Integration in heute üblichen Steuerungsarchitekturen untersuchen. Mithilfe dieser Entwicklungsstrategie können Konzepte sowohl auf der Planungs- als auch auf der Steuerungsebene erarbeitet werden, die direkt in industriellen Projekten einsetzbar sind.
Die Leistungsfähigkeit der Lösung konnte auf dem Gemeinschaftsstand Forschungsland NRW auf der Hannover-Messe Industrie 2000 in Augenschein genommen werden. Das zweiteilige Exponat zeigt im linken Teil die direkte Umsetzung einer mechanischen Kurvenscheibe.
Im rechten Teil wurde der Quersiegelprozess, ein Teilprozess aus einer Horizontalschlauchbeutelmaschine, dargestellt (Abbildung 3). Dabei ist die Anwendung der Funktion „elektrische Kurvenscheibe“ auf den ersten Blick gar nicht offensichtlich. Wird die Bewegung des Quersieglers kinematisch analysiert, so ist gerade die Realisierung als ereignisgesteuerte elektronische Kurvenscheibe die effizienteste Lösung dieser typischen Bewegungsaufgabe aus dem Verpackungsmaschinenbau.

Literatur

[1] VDI-Richtlinie 2143, Blatt 1: Bewegungsgesetze für Kurvengetriebe – Theoretische Grundlagen. VDI-Verlag. Düsseldorf 1980.
[2] Bechtloff, J.: Einsatz von Spline-Interpolationsmethoden zur Bewegungsgesetzgenerierung. VDI-Berichte Nr. 847, S. 125-138, Düsseldorf: VDI-Verlag 1990.
[3] Bechtloff, J.: Neue Verfahren zur Echtzeitinterpolation für die Bahnsteuerung 6-achsiger Industrieroboter. VDI-Berichte Nr. 1094, S. 511-524, Düsseldorf: VDI-Verlag 1993.
[4] Bechtloff, J.: Reduzierung von Engineering-Kosten durch Software-Standardisierung – Teil 2: Biegeanlage. MSR-Magazin 1-2/97, S. 44-45.
[5] Bewegungserzeugung mit elektronischen Kurvenscheiben. VDI-Bericht Nr. 1533, S. 521-538, Düsseldorf: VDI-Verlag 2000.

Mikroelektronik neuronaler Netze

Chips imitieren Funktionsmechanismen des Gehirns

Es scheint eine Konstante der Wissenschaftsgeschichte zu sein, dass das menschliche Gehirn immer am jeweils kompliziertesten Artefakt aus menschlicher Hand gemessen wird. Waren es im Altertum mechanische Konstrukte wie pneumatische Maschinen oder Uhrwerke in der Renaissance, sind es in jüngerer Zeit elektronische Modelle wie Telegrafensysteme am Ende des 19. Jahrhunderts sowie Computer und Rechnernetze in heutiger Zeit. Basistechnologie heutiger Informationsverarbeitung ist die Mikroelektronik. Wieweit man mit dieser Technologie derzeit an das biologische Vorbild herankommt, wird im Folgenden anhand von konkreten Realisierungen künstlicher neuronaler Netzwerkmodelle aus dem Fachgebiet Schaltungstechnik exemplarisch aufgezeigt.

Gehirn und Computer: zwei grundverschiedene Welten

Längst sind Computer zu einem selbstverständlichen Bestandteil unserer immer stärker auf die Verarbeitung von großen Datenmengen angewiesenen „Informationsgesellschaft“ geworden. So eindrucksvoll die Leistungsfähigkeit heutiger Rechnersysteme in vielen Domänen auch ist, es gibt Bereiche, in denen diese nimmermüden Rechenknechte immer noch eine ausgesprochene Beschränktheit im Vergleich zu den Möglichkeiten der „informationsverarbeitenden Maschine Gehirn“ haben.

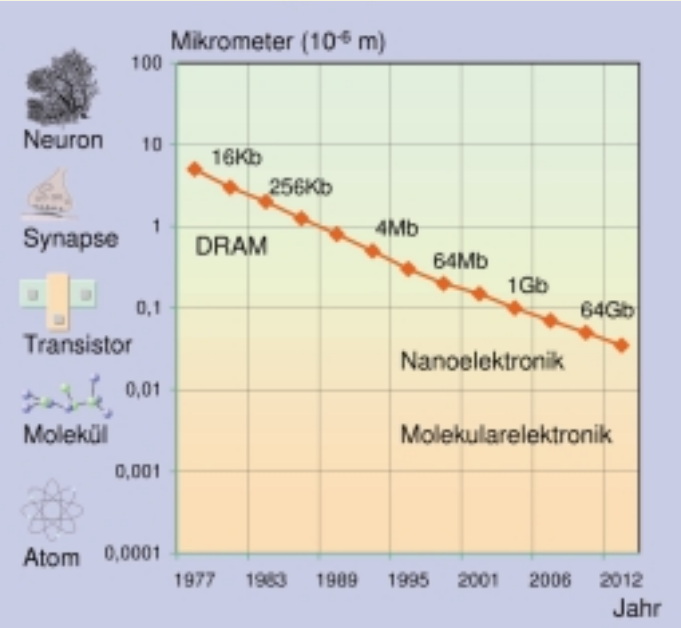
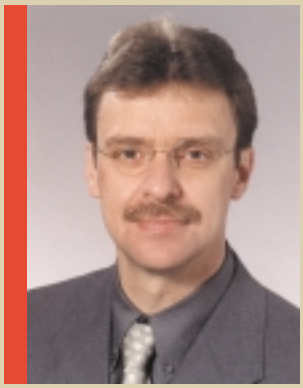


Abb. 1: Entwicklung der Leiterbahnbreiten (in Mikrometer) und der Speicherkapazität von dynamischen Speicherbausteinen (DRAM, K=Kilo 10^3, M=Mega 10^6, G=Giga 10^9 Bit).



Prof. Dr.-Ing. Ulrich Rückert ist seit 1995 Professor für Schaltungstechnik am Heinz Nixdorf Institut und am Fachbereich 14/Elektrotechnik und Informationstechnik der Universität Paderborn. Seine Arbeitsgebiete sind die mikroelektronische Systemintegration, die neuronale Informationstechnik und kognitive Systeme.

Einer dieser Bereiche ist die Kognition, die für jene Prozesse steht, durch die Sinnesinformationen umgewandelt, ausgewertet, gespeichert und wieder verwendet werden. Wenn es beispielsweise darum geht, Objekte zu erkennen und Szenen zu interpretieren, sich an gespeichertes Wissen assoziativ zu erinnern oder komplexe Bewegungsabläufe zu steuern, dann sind im Allgemeinen die Computer und Roboter unserer Tage den biologischen, aus Nervenzellen (Neuronen) aufgebauten „Rechenmaschinen“ noch deutlich unterlegen. Selbst das winzige, aus „nur“ rund 500 000 Neuronen bestehende Gehirn einer Fliege ermöglicht diesem Insekt, Flugbahnen durch unser Wohnzimmer mit großer Geschwindigkeit zu steuern, ohne dabei mit irgendwelchen Objekten oder anderen Fliegen zusammenzustoßen. Eine enorme Leistung, die selbst Hochleistungsparallelrechner nicht in der geforderten Zeit erbringen können – und das, obwohl die Einzeloperationen im Computer mindestens tausendfach schneller ablaufen als im Gehirn. Dies muss offenbar daran liegen, dass im Gehirn ungeheuer viele Operationen gleichzeitig durchgeführt werden, während in dem zentralen Rechenwerk eines modernen Computers nur wenige Operationen pro Zeiteinheit stattfinden können. Heutige Parallelrechner muten noch immer bescheiden gegenüber einer Parallelität von etwa zehn Milliarden asynchron arbeitenden Nervenzellen in der Gehirnrinde des Menschen an.
Aber nicht nur die massive Parallelität sondern auch die Ressourceneffizienz, d.h. die Realisierung der Systemeigenschaften mit geringem Aufwand hinsichtlich Baugröße, Energie und Reaktionszeit, ist für den Ingenieur faszinierend und herausfordernd zugleich. Es stellt sich somit die Frage, ob und wie gut mit heutiger Technik, der Mikroelektronik, diese Systemeigenschaften nachgebildet werden können.

Schlüsseltechnologie Mikroelektronik

Der atemberaubende Fortschritt der Mikroelektronik ist die treibende Kraft für die Entwicklung neuer technischer Produkte mit deutlich besserer Funktionalität und höherer Leistung bei gleichzeitig niedrigeren Kosten. Mit der damit einhergehenden Zunahme der Anwendung dieser Produkte in nahezu allen Lebensbereichen hat sich die Mikroelektronik zur Schlüsseltechnologie der modernen Informationsgesellschaft entwickelt. Mit der Entwicklung mikroelektronischer Systeme zu immer höheren Integrationsgraden – Stand der Technik sind mehrere Milliarden Bauelemente (Transistoren) auf einer Fläche von einem Quadratzentimeter – erhebt sich die Frage nach neuen Schaltungs- und Systemkonzepten, die dieses technologische Potenzial effizient auszunutzen vermögen. Die Grenzwerte der technologischen Möglichkeiten werden derzeit nur von Speicherbausteinen erreicht (Abbildung 1), bei denen auf Grund ihrer ausgesprochen regulären und modularen Architektur im Wesentlichen „nur“ schaltungstechnische und technologische Probleme zu lösen sind. Bei komplexen logischen Bausteinen, wie z.B. Mikroprozessoren, ist das Problem der hohen Entwurfs- und Testkomplexität in den Vordergrund getreten.

Gefordert sind daher Systemkonzepte, die einerseits die Möglichkeiten der Halbleitertechnologie weitestgehend ausschöpfen und andererseits die Entwurfs- und Testkomplexität auf ein beherrschbares Maß reduzieren. Einen interessanten Ansatz bieten in diesem Zusammenhang künstliche neuronale Netze. Die Natur soll hier einmal mehr als Wegweiser für die Erforschung neuer Prinzipien in der Informationsverarbeitung und für die Entwicklung neuer technischer Verarbeitungsstrukturen dienen.

Künstliche neuronale Netze

Mit künstlichen neuronalen Netzen werden biologische neuronale Netze als informationsverarbeitende Systeme nachgeahmt. Die Untersuchungen über die informationsverarbeitenden Fähigkeiten neuronaler Netze haben zur Definition vieler Modellvarianten geführt. Gemeinsam ist diesen Modellen ihr im Allgemeinen modularer und regulärer Aufbau. Sie bestehen aus sehr vielen gleichartigen und relativ einfachen Verarbeitungseinheiten, den Modellneuronen (Abbildung 2), die über gewichtete Verbindungsleitungen hochgradig miteinander vernetzt sind und erst durch ihre Zusammenschaltung zu komplexen Gesamtoperationen fähig werden. Die meisten Modelle gehen von einer regelmäßigen Verbindungsstruktur aus. Sie weisen daher die für den

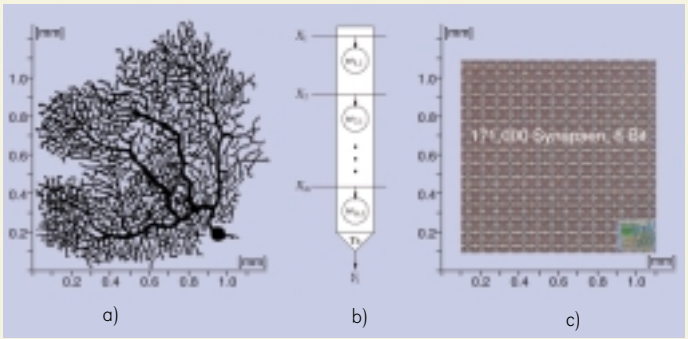


Abb. 2: Prinzipielles Aussehen eines biologischen Neurons (Purkinje-Zelle (a)), eines Modellneurons (b) und einer mikroelektronischen Schaltung eines Modellneurons (c).

X → KNN → Y		Aufgabe	Bewertung	Modell
endliche Menge	endliche Menge	Assoziation/ Speicherung	Speicherkapazität	NAS
unendliche Menge	endliche Menge	Klassifizierung/ Quantisierung	Fehlerwahrscheinlichkeit	SOM
unendliche Menge	unendliche Menge	Approximation/ Interpolation	Abstandsmaß	LCNN

Abb. 3: Anwendungsbereiche künstlicher neuronaler Netze.

integrationsgerechten Entwurf mikroelektronischer Systeme charakteristischen Eigenschaften der Modularität und Regularität auf /1/.

Ein künstliches neuronales Netz (KNN) ist somit ein Netzwerk aus formalen Neuronen und wird durch drei Komponenten spezifiziert: Das Neuronenmodell beschreibt die Eingangs-Ausgangs-Beziehung eines einzelnen Neurons. Ein Neuron hat dabei viele Eingänge und einen Ausgang. Die Verknüpfungsstruktur gibt an, wie die Modellneuronen untereinander verbunden sind. Sie wird meist durch einen Graphen beschrieben und legt ferner fest, welche Neuronen der Netzeingabe und der Netzausgabe dienen. Die Lernregel gibt an, wie sich die Verbindungen (Synapsen oder Gewichte w_{ij}, Abbildung 2b) zwischen Neuronen verändern.

Die ansonsten üblichen Schwierigkeiten bei der Programmierung paralleler Rechnerarchitekturen treten hier insofern nur eingeschränkt auf, als ein KNN nicht programmiert, sondern mithilfe von Beispieldaten trainiert (angelernt) wird. An Stelle der Programmierung erfolgt eine Konfiguration des KNN, durch die das Neuronenmodell, die Netzwerkstruktur und das Lernverfahren möglichst optimal bezüglich einer gewünschten Anwendung ausgewählt werden. Es gibt dementsprechend nicht nur ein KNN, das für alle Aufgaben gleichermaßen gut geeignet ist. Über die Frage, welches KNN für welche Anwendung das beste Modell ist, gibt es noch keinen allgemeinen Konsens.

In der Fachgruppe Schaltungstechnik beschäftigen wir uns mit den drei Anwendungsbereichen Assoziation, Klassifikation und Approximation (Abbildung 3). Für jeden dieser Bereiche haben wir uns jeweils ein Modell ausgewählt, das unserer Meinung nach interessante Aspekte der Informationsverarbeitung beinhaltet sowie eine ressourceneffiziente Realisierung verspricht.

Der Assoziativspeicher

Die grundlegende Aufgabe eines Assoziativspeichers ist es, eine endliche Menge von binären Musterpaaren (x, y) abzuspeichern und zu jedem Eingabemuster das zugehörige Antwortmuster wieder auszulesen. Dabei soll allerdings auch bei Eingabe einer leicht verrauschten Variante eines Musters x das korrekte Muster y ausgegeben werden. Anschaulich kann man diese Muster als Frage und Antwort bzw. Reiz und Reaktion auffassen. Hierbei ist man natürlich daran interessiert, möglichst viele Musterpaare zu speichern und diese schnell fehlerfrei auszulesen.

Die einfachste Form eines KNN, bestehend aus einer Schicht von Modellneuronen mit binären Ein- und Ausgaben sowie Verbindungsgewichten (Abbildung 4a), ermöglicht bereits eine überraschend effiziente Realisierung eines „neuronalen Assoziativspeichers“ (NAS). Die Speicherung der Muster erfolgt in einer einzigen Lernepoche, d.h. jedes Paar (x, y) wird dem

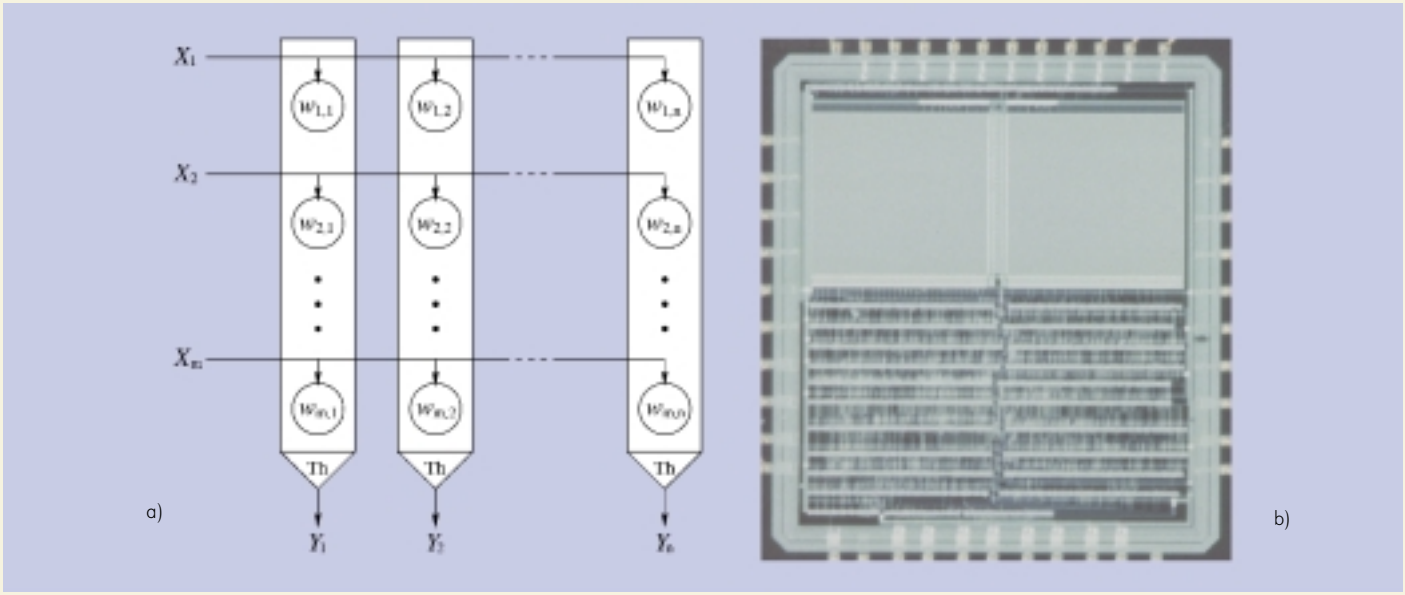


Abb. 4: Architektur (a) und mikroelektronische Realisierung (b) eines neuronalen Assoziativspeichers.

NAS genau einmal präsentiert. Aus theoretischen Arbeiten über das asymptotische Verhalten ist bekannt, dass eine sehr hohe Speicherkapazität und ein schneller Datenzugriff nur für sogenannte spärlich kodierte Musterpaare gilt /2/. Ein Muster ist spärlich kodiert, wenn nur sehr wenige Komponenten des Musters Eins und der überwiegende Anteil Null sind. Diese Kodierungsart ist unüblich in der Informationstechnik, aber dafür vermutlich in biologischen Nervennetzen wesentlich. Das Prinzip der spärlichen Kodierung ist daher aktuelles Forschungsthema in den Neurowissenschaften.

Auf Grund der sehr einfachen und leicht erweiterbaren Architektur eines NAS, kann dieser entsprechend effizient /3/ realisiert werden (Abbildung 4b). Mit heutigen Standardtechnologien lassen sich etwa 16 000 Neurone mit jeweils 16 000 Eingängen auf einem Quadratzentimeter Silizium realisieren. In einem solchen Baustein können mehr als vier Millionen Muster assoziativ gespeichert (z.B. Name und Telefonnummer aller Einwohner Berlins) und jede dieser Assoziationen in weniger als einer Mikrosekunde ausgelesen werden. Die dabei erforderliche elektrische Energie entspricht einem Bruchteil der Energie, die moderne Mikroprozessoren für diese Aufgabe benötigen. Auf Grund dieser Kenndaten kann das NAS-Konzept sehr wohl klassischen Methoden (Hashing, Suchbäume) überlegen sein.

Selbstorganisierende Klassifikatoren

Wird die Grundarchitektur nach Abbildung 4 durch die Anordnung der Modellneuronen in einer Ebene (Abbildung 5a) mit einer Verkopplung benachbarter Elemente erweitert, erhält man die selbstorganisierende Merkmalskarte (SOM, (4)). Die Eingabe des Netzes bilden Vektoren, die allen Neuronen parallel zugeführt werden. Der Lernalgorithmus hat die Eigenschaft, dass ähnliche Eingabevektoren auf Neurone abgebildet werden, die in der Karte benachbart sind. Die Ausbildung dieser topologischen Ordnung auf der Karte vollzieht sich selbstorganisierend durch wiederholtes Anbieten der Eingabevektoren in der Lernphase und kann zu einer Visualisierung von Gruppenbildungen (Cluster) im Eingabedatenraum (Abbildung 5b) verwendet werden. Ausgabe in der Abrufphase ist die Position des Neurons, das der Eingabe bezüglich einer vorgegebenen Metrik am ähnlichsten ist, d.h. die SOM ordnet die Eingabevektoren einer endlichen Menge von Klassen zu (Klassifikation). Sie eignet sich sowohl für zeitkritische Erkennungsaufgaben (z.B. Fehlererkennung an laufenden Motoren) als auch für die schnelle Clusteranalyse sehr großer Datenmengen.

Die SOM stellt höhere Anforderungen an die mikroelektronische Realisierung, da die Modellneuronen, die Verbindungsstruktur und der Ablauf der Lernphase komplexer als beim NAS sind. Durch geringfügige Anpassungen des Verfahrens an die Gegebenheiten einer digitalen Schaltungstechnik lassen sich derzeit etwa 1 000 Neurone (mit z.B. 256 Synapsen) auf einem Quadratzentimeter Silizium realisieren, mit denen mehrere Millionen Vektorklassifikationen pro Sekunde durchgeführt werden können. Fehlerhafte Verarbeitungselemente werden durch den Lernalgorithmus automatisch ausgeblendet. Abbildung 5c zeigt die Realisierung einer Testschaltung von 5x5 Verarbeitungseinheiten in digitaler Schaltungstechnik /5/. Möchte man größere Karten realisieren, dann können SOM-Bausteine einfach zusammengeschaltet und so sehr große Netzwerke realisiert werden (Abbildung 5d).

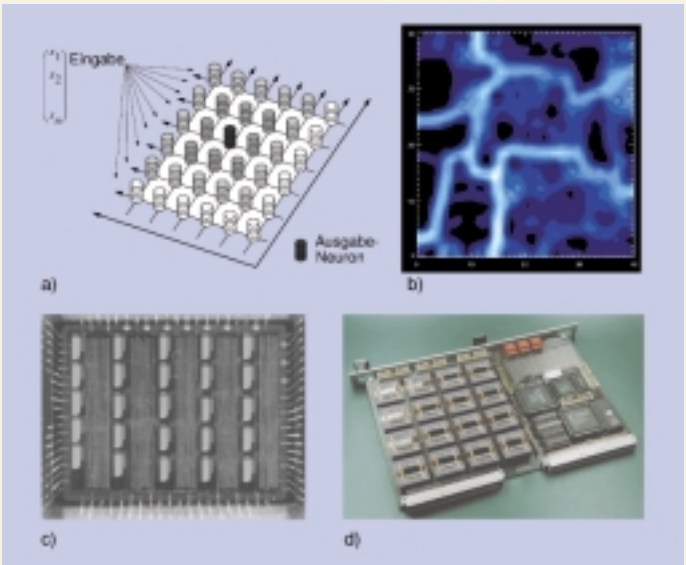


Abb. 5: Selbstorganisierende Merkmalskarten: Architektur (a), Visualisierung des Lernergebnisses (b), mikroelektronische Realisierung (c) und Zusammenschaltung von SOM-Bausteinen zur Realisierung größerer Netze (d).

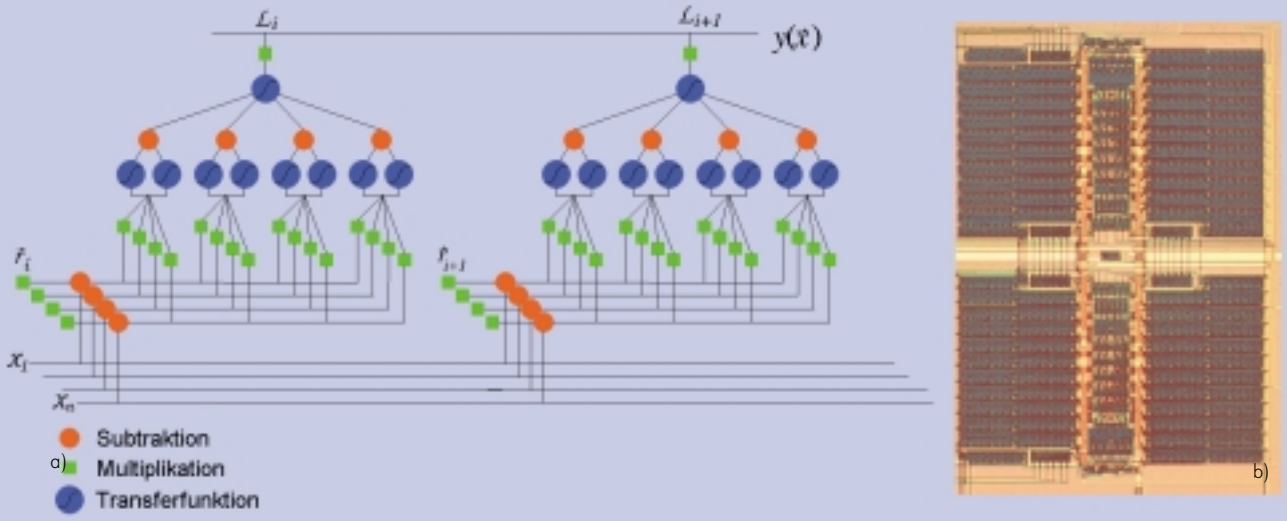


Abb. 6: Prinzipieller Aufbau eines „Local Cluster Neural Networks“ (a) und mikroelektronische Realisierung in analoger Schaltungstechnik (b).

Mehrschichtige Netze

Für den dritten Anwendungsbereich (Funktionsapproximation) haben wir das sogenannte „Local Cluster Neural Network“ (LCNN) ausgewählt. Ein LCNN (Abbildung 6a) ist ein mehrschichtiges Netzwerk mit lokal operierenden Verarbeitungseinheiten und kann für einfache Regelungen eingesetzt werden. Da Sensor- und Aktorsignale in technischen Anwendungen überwiegend analog ausgelegt sind und ein LCNN für einfache Regelungen im Allgemeinen nicht sehr groß wird, haben wir uns für eine analoge Realisierung entschieden /6/. Der realisierte Baustein (Abbildung 6b) enthält zwei Netze mit jeweils sechs Eingaben und einer Ausgabe auf einer Fläche von 5 mm^2 (15 Khz Betriebsfrequenz). Diese Realisierung eignet sich z.B. für die Anwendung in der Mikrosystemtechnik, da hier Sensorik, Aktorik und Informationsverarbeitung auf kleinster Fläche bei minimaler Verlustleistung integriert werden müssen.

Ausblick

Von der Gehirnforschung geht schon immer eine große Faszination, aber auch ein gewisses Unbehagen aus. Für den Mikroelektroniker bietet diese ein riesiges Reservoir an neuen Ansätzen für ressourceneffiziente Architekturen für die Informationstechnik und insbesondere für kognitive Systeme. Mit Strukturgrößen von 0,1 Mikrometer (10^{-7} Meter) hat die Halbleitertechnologie die biologischen Strukturgrößen (Abbildung 1 und 2) erreicht. Allerdings nutzt die Natur alle drei Raumdimensionen im Gehirn effizient aus, während die Mikroelektronik im Wesentlichen die Oberfläche des Siliziums, d.h. nur zwei Raumdimensionen intensiv nutzt.

Die hier vorgeschlagenen Lösungen für künstliche neuronale Netze zeichnen sich durch eine sehr gute Ressourceneffizienz aus, da sie für Realzeitanwendungen bei extrem niedrigem Energiebedarf selbst für kleinste Baugrößen geeignet sind. Ferner ist eine Anpassung der Realisierung an neue Technologien auf Grund der modularen Architektur relativ einfach möglich, ohne dem Problem der enormen Entwurfs- und Testkomplexität ausgesetzt zu sein. Diese Architekturen profitieren damit von den zukünftigen Entwicklungen der Mikroelektronik besser als Programmlösungen auf Mikroprozessoren. Ihre Systemeigenschaften (Speicherkapazität, Klassifikationsgüte, Approximationsgenauigkeit) verbessern sich zudem merklich mit steigender Netzwerkgröße.

Gewiss, wir sind heute noch sehr weit davon entfernt, die Funktionsmechanismen eines Gehirns wirklich zu durchschauen, und die Konstruktion eines „künstlichen Gehirns“ wird wohl auch weiterhin sehr lange – wenn nicht für immer – eine Utopie bleiben. Die Realisierung von neuronalen Netzen in Hardware hat daher nicht die detailgetreue Nachbildung von Nervensystemen zum Ziel, sondern schlicht die effiziente Nutzung heutiger technologischer Möglichkeiten zur Lösung technischer Fragestellungen. Die technologische Basis ist heute die Mikroelektronik. Auf die technologische Basis von morgen dürfen wir gespannt sein.

Literatur

/1/ Ramacher, U., Rückert, U.: VLSI Design of Neural Networks, Boston, Kluwer Academic Publishers, 1991.

/2/ Palm, G.: Neural Assemblies, Berlin, Springer-Verlag 1982.

/3/ Heittmann, A., Rückert, U.: Mixed Mode VLSI Implementation of a Neural Associative Memory, Tagungsband zur MicroNeuro’99, Granada, IEEE Computer Society, Los Alamitos, 1999, S. 299-306.

/4/ Kohonen, T.: Self-Organization and Assoziative Memory, Berlin, Springer-Verlag, 1984.

/5/ Rüping, S., Porrmann, M., Rückert, U.: SOM Accellerator System, Neurocomputing 21, Elsevier, 1998, S. 31-50.

/6/ Sitte, J., Körner, T., Rückert, U.: Local Cluster Neural Net: Analog VLSI Design, Neurocomputing 19, Elsevier, 1998, S. 185-197.

Die Kunst der Kunststoff-Herstellung

Innovative Reaktionsführung zur Herstellung von Polymeren

Kaum ein anderer Bereich der Chemie hat unser Leben in den letzten Jahrzehnten so nachhaltig beeinflusst wie die Chemie der *Polymeren*. Dem Laien besser unter dem Begriff *Kunststoffe* bekannt, haben Polymerprodukte enorme Bedeutung erlangt und sind mittlerweile unverzichtbarer Bestandteil im Alltagsleben unserer hoch technisierten Industriegesellschaft geworden. Dabei wäre es unpassend, Kunststoffe nur als schlichte Ersatzmaterialien klassischer Werkstoffe wie Keramik, Glas, Holz, Metall, Baumwolle usw. abzustempeln. Selbst in heutiger Zeit, gut 90 Jahre nach ihrer ersten Synthese im Labor, bieten sich den Forschern bei der Entwicklung innovativer Polymerwerkstoffe und der Erschließung neuer Anwendungsbereiche noch völlig ungeahnte Möglichkeiten, weshalb Polymere im wahrsten Sinne des Wortes als „Kunst"-Stoffe angesehen werden können.

Derzeit werden allein in Deutschland jährlich rund neun Millionen Tonnen Kunststoffprodukte erzeugt, die sich auf die in Abbildung 1 dargestellten Einsatzbereiche verteilen.
Um auch zukünftig den stetig steigenden Bedarf an Kunststoffen decken zu können, werden beträchtliche Anstrengungen unternommen, die großtechnischen Herstellungsverfahren der Polymere unter verschiedensten Gesichtspunkten wie Prozesstechnik, Betriebsbedingungen, Produktionskapazität oder Wirtschaftlichkeit ständig weiter zu entwickeln bzw. zu optimieren; ein Aufgabenfeld, dem sich das Fachgebiet der *Polymerreaktionstechnik* widmet.

Definition, Aufbau und Einteilung der Kunststoffe

Die Polymerchemie versteht sich selbst als Chemie der synthetischen Riesenmoleküle bzw. *Makromoleküle*. Vergleichbar mit einer Perlenschnur, auf der einzelne Perlen zu einer Kette aufgefädelt werden, entsteht ein Makromolekül durch stufenweises

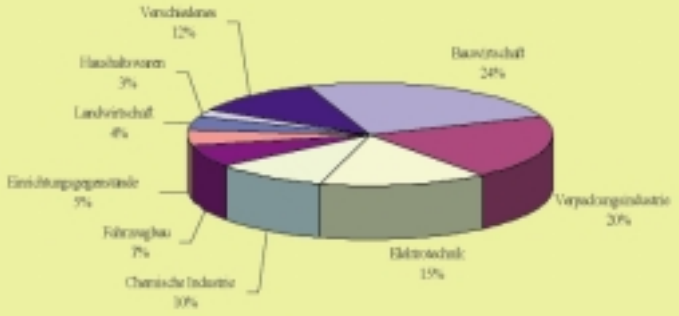


Abb. 1: Einsatzbereiche der jährlich produzierten Kunststoffprodukte.

Prof. Dr.-Ing. Hans Joachim Warnecke, Leiter des Fachgebietes Technische Chemie und Chemische Verfahrenstechnik der Universität Paderborn (links).

Prof. Dr. Hans-Christoph Broecker, Leiter des Fachgebietes Kunststoffe (Mitte).

Dr. rer. nat. Mike Bobert, Geschäftsleitung des „Westfälischen Umwelt Zentrums".

Verknüpfen einer Vielzahl kleiner Grundbausteine, sogenannter Monomere, zu einem ausgedehnten Molekülverbund. Dabei können Makromoleküle vielfältige Strukturen annehmen. So lassen sich Makromoleküle aus lediglich einer Sorte von *Monomeren* aufbauen *(Homopolymere)*, die dann in Form unverzweigter Kettenmoleküle und verzweigter Makromoleküle mit Seitenketten vorliegen können. Weitaus mehr Erscheinungsformen ergeben sich allerdings, wenn zwei oder mehr verschiedene Monomersorten zu Mischpolymeren *(Copolymere)* verknüpft werden. Zusätzlich zu den bereits angedeuteten Makromolekülstrukturen können hier dreidimensionale Molekülnetzwerke erhalten werden. Weitere Variationsmöglichkeiten resultieren zudem aus der rein zufälligen oder auch geordneten Abfolge der unterschiedlichen Monomerkomponenten entlang der Polymerketten.
Durch die Art der Monomerbausteine und vor allem durch die Struktur des aus ihnen aufgebauten Molekülverbundes wird bereits im Wesentlichen festgelegt, ob ein Kunststoff hart,

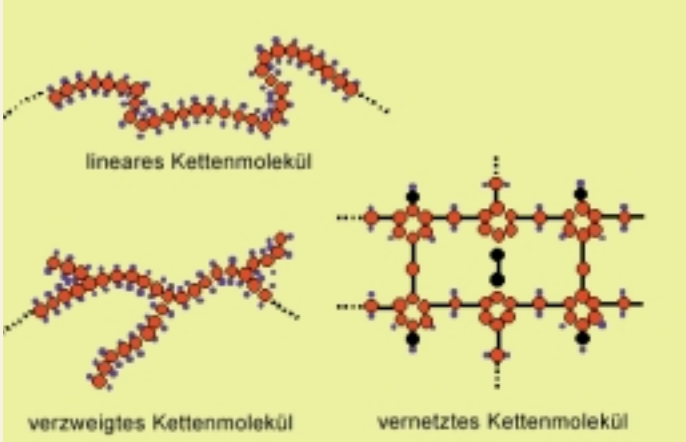


Abb. 2: Makromolekülstrukturen.

Kunststoffname	Eigenschaften	Verwendung
- Duroplaste – Synthetische, härtbare Kunststoffe		
Phenoplaste (Bakelit, Trolon)	Durch Pressen äußerst hart, wasserbeständig, temperaturunempfindlich bis 100°C	Formteile für Elektrotechnik; Edelkunstharze; Maschinenteile; Leimharze
Aminoplaste (Resopal, Pollopas)	Hellfarbig, lichtbeständig, geschmack- und geruchlos, gute Isolatoren	Haushaltsgeräte; Dekorationsplatten; Leimharze; Küchenplatten; Isoliermittel
- Thermoplaste – Synthetische, in der Wärme verformbare Kunststoffe		
Polyethylen (PE), Polypropylen (PP)	fest, korrosionsfest, kältebeständig, leichteste Kunststoffe, Sorten verschiedener Biegsamkeit, Härte und Wärmebeständigkeit	Isoliermaterial; Wasserleitungen; Lebensmittelverpackungen; Spritzgussmassen (Folien, Haushaltsgegenstände)
Polyvinylchlorid (PVC)	Vielseitiger Kunststoff, starr, hornartig, aber mit Weichmacher beliebig leder- oder gummiartig einstellbar	Hart-PVC (Rohrleitungen); Weich-PVC (Dichtungen, Fußbodenbeläge); PVC-Paste, Schuhe
Polystyrol (PS) (aufgeschäumt: Styropor)	hart, glasklar, korrosionsfest, spröde, knickempfindlich	Haushaltsgegenstände; Rohre; Schaumstoff; Kühlraumisolierung; Rettungsringe
Polymethylmethacrylat (Acrylglas, Plexiglas)	Glasklar, durchsichtig, hart, alterungsbeständig	Verglasungen; medizin. Geräte; Lichtbänder; Leuchten; Haushaltsgegenstände
- Weitere wichtige Kunststoffe		
Polyamide (PA) (Perlon, Nylon)	Hornartig, zäh, elastisch, zum Spinnen feiner Fäden geeignet	Maschinenteile (geräuscharme Zahnräder ohne Schmierung); Drähte; Fasern
Polyurethane (PU)	Wasser- und benzinfest, mechanisch vielseitig einstellbar, PU-Schaum isoliert sehr gut	Klebemittel für Glas, Metall, Keramik; Schaumstoff (Isoliermaterial, Polsterung)
Polyester	Härten glasig aus, als Gießharz drucklos zu verarbeiten, große Festigkeit mit Glasfasern	Wellplatten für Vordächer; Sportboote, Balkonbrüstungen; Segelflugzeuge, Airbus
Polycarbonate (PC) (Makrolon)	Niedrige Wasseraufnahme, gute Isolierfähigkeit, sehr dimensionsstabil und schlagfest	Spritzgussartikel, Isolationsfolien, als Fasern in Mischgeweben
Silikone	Hitze- und kältefest von –60°C bis +300°C, Wasser abstoßend, korrosionsbeständig	Silikonöl (Schmiermittel); Silikonkautschuk (nicht alternde Dichtung)

Tabelle 1: Gruppierung von Kunststoffen, Eigenschaften und Verwendung.

verformbar, elastisch, schmelzbar, zäh, schlagfest usw. ist. Um bestimmte Eigenschaften zu verfeinern oder weitere Eigenschaften zu erzielen, werden Kunststoffen zudem noch spezielle chemische Substanzen beigegeben. Als wichtigste Zusätze sind Flammschutzmittel und Weichmacher zu nennen, des Weiteren kommen Ultraviolettstabilisatoren zum Schutz vor Verwitterung, Gleitmittel zur Verringerung von Reibungsproblemen oder auch Pigmente zur Farbgebung des Kunststoffes zum Einsatz.

Die Polymere lassen sich nach den Kriterien Verarbeitung, Eigenschaften und Verwendung in ein übersichtliches Feld weniger Kunststoffgruppen einteilen.

Zu den *Thermoplasten* zählen dabei aus linearen und eventuell verzweigten Makromolekülen aufgebaute Kunststoffe, die sich prinzipiell beliebig oft aufschmelzen und verformen lassen, beim Abkühlen aber ihre Form beibehalten. Die hoch vernetzten *Duroplaste* hingegen werden beim Erhitzen hart und sind nicht mehr weiter verarbeitbar. Bei den gummiartigen *Elastomeren* handelt es sich um schwach vernetzte, bei Krafteinwirkung leicht reversibel verformbare Polymere. *Fasern* finden ihr Anwendungsgebiet hauptsächlich auf dem Textilsektor, *Schaumstoffe* werden als Isoliermaterialien und für Polsterungen eingesetzt und *Polymerlösungen* dienen u.a. als Klebstoffe, Anstrichmittel und für Beschichtungszwecke. In Tabelle 1 ist diese Einteilung ansatzweise wiedergegeben, sie zeigt zudem die Zuordnung der geläufigsten Kunststoffsorten.

Die technische Kunststoffherstellung, eine Aufgabe der Polymerreaktionstechnik

Die technische Herstellung von Kunststoffen oder Kunststoffprodukten setzt sich wie jedes andere Produktionsverfahren in der chemischen Technik aus einzelnen, nacheinander geschalteten Prozessstufen zusammen:

1. Die Vorbereitung der Einsatzstoffe (Monomere, Zusatzstoffe, Hilfsstoffe).
2. Die chemische Umsetzung der Einsatzstoffe zu den Polymeren.
3. Das Aufbereiten und gegebenenfalls Weiterverarbeiten der Endprodukte zum fertigen Produkt.

Die erste und dritte Verfahrensstufe umfassen dabei Prozesse, bei denen auf mechanischem oder thermischem Wege lediglich rein physikalische Änderungen von Komponenten bewirkt werden. Zu diesen Verfahrensschritten gehören Zerkleinerungs-, Misch- oder auch Trennprozesse (z.B. Filtration und Destillation). Sie sind wissenschaftlicher Gegenstand der *Chemischen Verfahrenstechnik*.

Die zweite Stufe der Prozesskette ist für die Kunststoffherstellung der wichtigste Prozess. In einem als *Chemischer Reaktor* bezeichneten Reaktionsapparat vollziehen sich dabei jene zuvor im Labormaßstab wissenschaftlich erforschten chemischen Reaktionen, in denen aus den niedermolekularen Bausteinen die gewünschten Polymerverbindungen aufgebaut werden. Sie werden unter den allgemeinen Oberbegriffen *Polyreaktionen* bzw. *Polymerisationsreaktionen* zusammengefasst.
Der Reaktor ist stets das Kernstück eines chemisch-technischen Prozesses. Im Hinblick auf

- seine Art und Größe,
- seine Betriebsweise (aufgeteilt in separate Chargen oder kontinuierlich betrieben),
- die Betriebsbedingungen (Temperatur, Druck, Reaktionszeit, Konzentration, usw.) und
- die verwendeten Werkstoffe

ist der Reaktor derart zu konstruieren, zu bemessen und zu betreiben, dass die den Laborversuchen zu Grunde liegende chemische Reaktion sicher in den technischen Maßstab übertragen und unter gegebenen Voraussetzungen ein Produkt bestimmter Qualität und Menge bei minimalem Gesamtkostenaufwand hergestellt wird. Mit dieser Aufgabe befasst sich die Fachdisziplin der Chemischen Reaktionstechnik, in diesem speziellen Fall die *Polymerreaktionstechnik*.

Die Schwierigkeit bei der Gestaltung eines Reaktors liegt vor allem darin, dass neben der rein chemischen Umsetzung stets auch physikalische Prozesse, Werkstoff- und Fertigungsfragen sowie Wirtschaftlichkeitsbetrachtungen zu berücksichtigen sind.
Die Reaktionsmasse muss dem Reaktor zugeführt, während der Umsetzung gut durchmischt und nach der Umsetzung aus dem Reaktor abgeführt werden. *Stofftransport* und *strömungstechnische Vorgänge* sind damit eng verbunden. Die meist große Wärmeentwicklung bei Polyreaktionen erfordert Maßnahmen für eine gute Wärmeabfuhr an die Umgebung *(Wärmetransport)*. Jede chemische Umsetzung verläuft bis zu einem definierten Endzustand, durch den ein maximaler Umsatz bedingt ist. Die Geschwindigkeit, mit der sich die Reaktionsmasse diesem Endzustand nähert *(Reaktionskinetik)*, legt die erforderliche Reaktionszeit und damit für eine vorgegebene Produktionsmenge die notwendige Größe des Reaktors fest. Der Reaktor muss derart beschaffen sein, dass er sowohl den einzustellenden Betriebsbedingungen als auch der chemischen Aggressivität der Reaktionsmasse standhält *(Werkstoffkunde)*. Schließlich gilt es, für den Betrieb der Anlage das wirtschaftliche Optimum zu finden *(Ökonomie)*.
Die Konstruktion jedes Reaktionsapparates erfordert die Zusammenarbeit von Chemie und Physik einerseits und den Ingenieur- und Wirtschaftswissenschaften andererseits. Die Polymerreaktionstechnik baut auf dem Wissen dieser wissenschaftlichen Disziplinen auf und wendet sie auf ein jeweiliges Problem an.

Polymerreaktionstechnik an der Universität Paderborn

Für alle chemisch-technischen Produktionsprozesse gilt, dass das technische Know-how eines Unternehmens ausschlaggebend ist für seine Marktsituation, die Markterfolge seiner Produkte und seine zukünftigen Entwicklungschancen. Unternehmen, die mit ihren Produkten im weltweiten Wettbewerb langfristig bestehen wollen, müssen es sich zur steten Aufgabe machen, herstellungstechnische Verbesserungspotenziale für ihre Produkte zu erkennen und in Innovationen umzusetzen.
Zur Bewältigung dieser Aufgabe bietet sich das Fachgebiet Technische Chemie und Chemische Verfahrenstechnik der Universität Paderborn den Industrieunternehmen als kompetenter Kooperationspartner an, um auf Basis eines Know-how-Verbundes wertvolle Impulse für Innovationen zu setzen und die Optimierung von Herstellungsprozessen technisch bedeutsamer Polymere zu erarbeiten.
Ein aktuelles Forschungsprojekt beschäftigt sich mit der Optimierung der Prozessführung bei der Herstellung solcher Festharze, die aus relativ kleinen Makromolekülen mit einer möglichst einheitlichen Zahl von Monomerbausteinen zusammengesetzt sind, oder anders gesagt, die eine niedrige Molekularmasse bei enger Molekularmassenverteilung aufweisen.
Diese Festharze werden hauptsächlich in wasserbasierten Druckfarben- und Lack-Systemen verwendet, welche sich sehr gut als Alternativen für lösemittelhaltige Systeme bewährt haben. Durch den Ersatz der organischen Lösemittel durch Wasser konnten dabei Probleme wie die Belastung der Umwelt mit Lösemit-

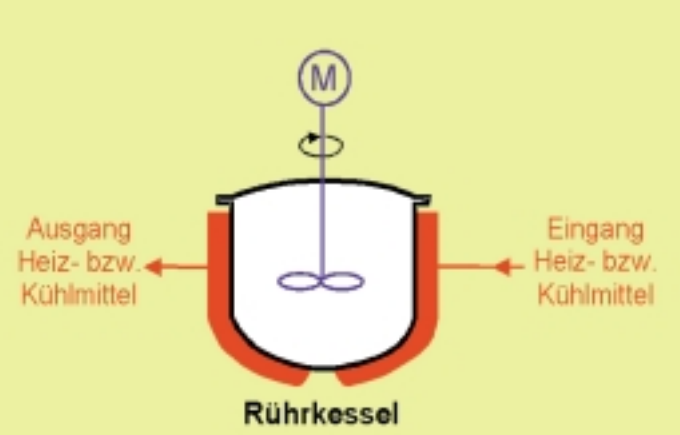


Abb. 3: Absatzweise betriebener Rührkessel.

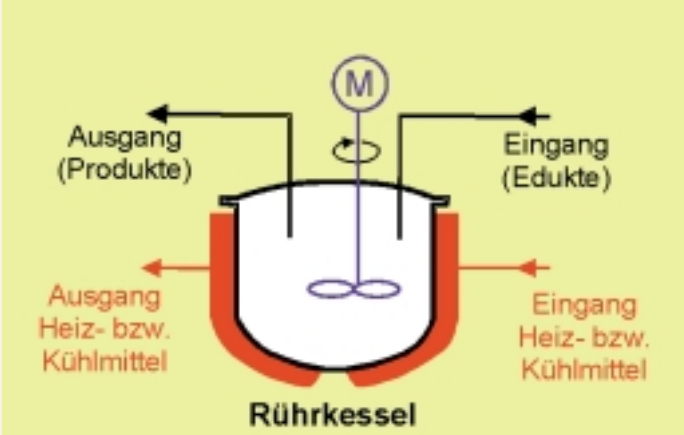


Abb. 4: Durchflussrührkessel.

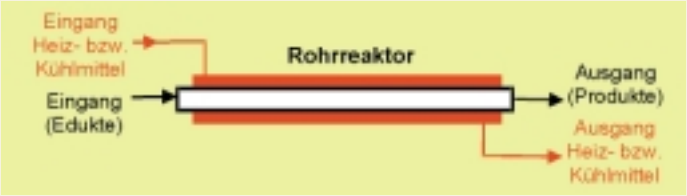


Abb. 5: Kontinuierlich betriebener Rohrreaktor.

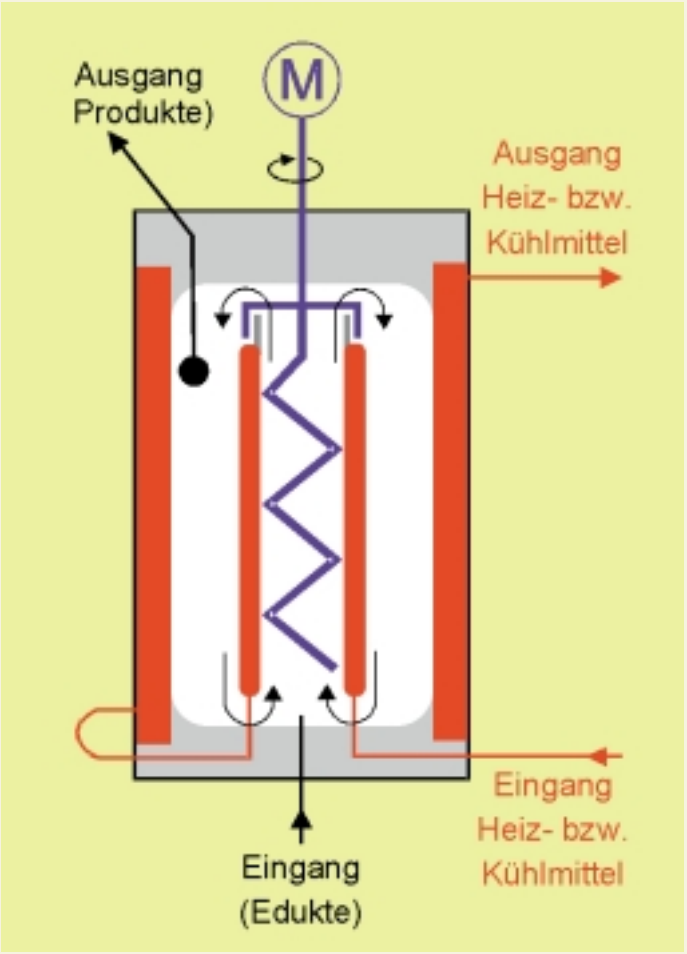


Abb. 6: Schnecken-Schlaufenreaktor.

teldämpfen, die leichte Entzündlichkeit der Druckfarben und Lacke und die Giftigkeit der Systeme beseitigt und die Qualität der Druckfarben und Lacke verbessert werden.
Die wässrigen Flexo- und Tiefdruckfarben zeichnen sich durch eine ausgezeichnete Haftung auf den meisten Untergründen, hohen Glanz, gute Verdruckbarkeit, schnelle Trocknung, gute Wiederauflösbarkeit und hohe Beständigkeit aus. Die wässrigen Überdrucklacke zeigen gute Flexibilität, sehr hohen Glanz, gute Wasser- und Nass-Abriebfestigkeit sowie Hitzebeständigkeit.

Für die technische Herstellung der Festharze kommen derzeit diskontinuierlich oder *absatzweise betriebene Rührkesselreaktoren*, auch als *Batchreaktoren* bezeichnet, zum Einsatz (siehe Abbildung 3).
Dieser diskontinuierliche Produktionsbetrieb *(Chargenbetrieb)* ist dadurch gekennzeichnet, dass zunächst die Ausgangskomponenten und die notwendigen Begleitstoffe in einen Rührkesselreaktor eingefüllt werden, in dem sie nach Start der chemischen Umsetzung für eine definierte Reaktionszeit unter festgelegten Reaktionsbedingungen verweilen. Während durch effektives Rühren dafür gesorgt wird, dass der Reaktorinhalt stets innig durchmischt und homogen ist, ändert sich wegen der Umsetzung der Ausgangsstoffe zu den Produkten die Zusammensetzung des Reaktorinhalts mit fortschreitender Zeit. Zur Einhaltung der vorgesehenen Reaktionstemperatur lässt sich der Rührkessel über die Behälterwand abkühlen oder aufheizen. Nach beendeter Umsetzung wird der Reaktor schließlich entleert, gereinigt und steht für die nächste Produktionscharge zur Verfügung.
Als nachteilige Eigenschaft des Batchreaktors ist vor allem seine hohe Produktionsausfallzeit *(Totzeit)* zum Befüllen, Entleeren, Aufheizen und Abkühlen sowie für die aufwändigen Reinigungsarbeiten zu nennen. Außerdem fallen infolge des Aufheizens und Abkühlens bei jeder produzierten Charge hohe Energiekosten an. Die mit der variierenden Zusammensetzung des Reaktorinhalts gekoppelte Veränderung wichtiger physikalischer Größen und Prozessparameter macht einen erheblichen Steuerungs- und Überwachungsaufwand erforderlich, um stets die vorgesehenen Betriebsbedingungen einzuhalten. Dies betrifft im Besonderen die Steuerung der Reaktortemperatur, da die Reaktorkühlung der sich zeitlich verändernden, freigesetzten Reaktionswärmemenge anzupassen ist. Hinzu kommt, dass ein Rührkessel im Verhältnis zu seinem Volumen nur eine kleine Behälterwandungsfläche aufweist, wodurch die notwendige Wärmeabfuhr an das Kühlmittel stark eingeschränkt wird. Trotz aller Steuerungs- und Regelungstechnik ist die Qualität der erzeugten Festharze immer Schwankungen unterworfen, die im Extremfall zur Unbrauchbarkeit einer kompletten Charge und somit zu erheblichen wirtschaftlichen Einbußen führen können.

Im Rahmen des Forschungsvorhabens werden im Fachgebiet Technische Chemie und Chemische Verfahrenstechnik zunächst im Technikumsmaßstab neue Reaktorkonzepte entwickelt und erprobt, mittels derer die beim Batchreaktor aufgezeigten Nachteile umgangen werden können und die eine deutlich verbesserte Prozessführung bei der Produktion der Festharze gestatten. Hauptansatzpunkt ist dabei die oftmals schwer beherrschbare Umstellung der Festharzproduktion von diskontinuierlicher auf kontinuierliche Betriebsweise des eingesetzten Reaktionsapparates *(Fließbetrieb)*.
Der Fließbetrieb zeichnet sich dadurch aus, dass ein Gemisch aus Ausgangskomponenten und Begleitstoffen in einem gleich bleibenden Mengenstrom in den Reaktionsapparat eingespeist wird. Ebenso kontinuierlich wird ein Endgemisch, bestehend aus Reaktionsprodukt, nicht umgesetzten Ausgangsstoffen und Begleitstoffen, aus dem Reaktor ausgetragen. Werden dabei alle wichtigen Betriebsparameter wie die Reaktortemperatur, die Zulauf- und Ablaufmengen, die Zusammensetzung des Zulaufgemisches usw., zeitlich konstant gehalten, so tendiert ein im Fließbetrieb arbeitendes System dazu, einen zeitunabhängigen, gleich bleibenden Zustand einzunehmen; man spricht dann von einem *stationären Betriebszustand*.
Im Vergleich zum Chargenbetrieb ergeben sich durch die kontinuierliche Betriebsführung entscheidende Vorteile:

- Bei richtiger Auslegung des Fließbetriebs ist für eine bestimmte Produktionsleistung ein deutlich kleinerer Reaktionsapparat erforderlich, da die Totzeiten für das Befüllen, Entleeren, Aufheizen, Abkühlen und Reinigen entfallen.
- Infolge der Einhaltung (annähernd) konstanter Betriebsbedingungen kann dauerhaft eine (nahezu) einheitliche Produktqualität aufrecht erhalten werden.
- Der Fließbetrieb gestattet eine weitgehende Automatisierung

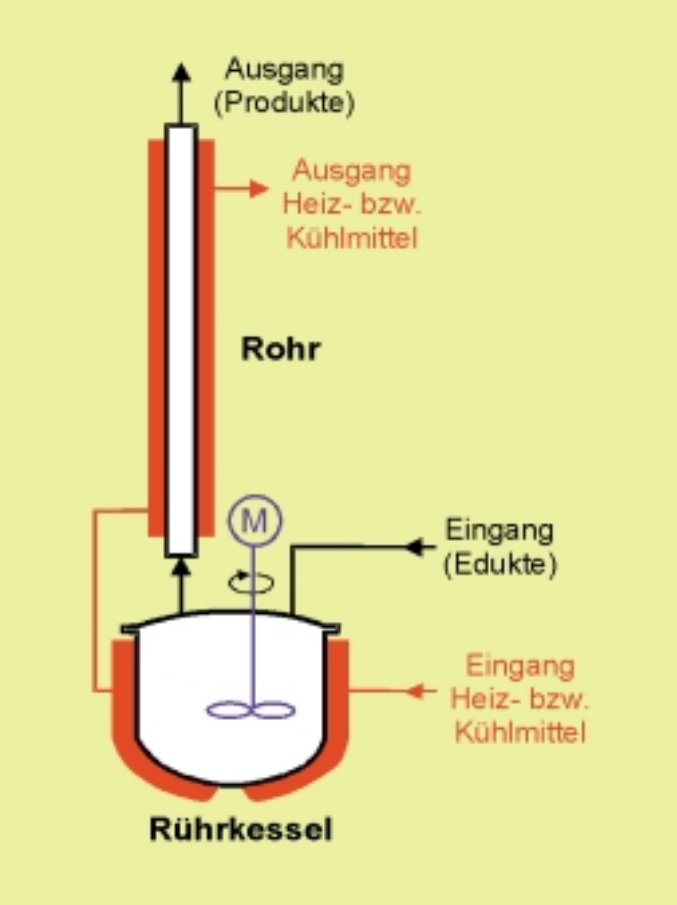


Abb. 7: Rührkessel-Rohr-Reaktor.

und lässt sich im Allgemeinen sehr viel leichter steuern und regeln. Gemeinsam mit dem Wegfall der Totzeiten ist ein erheblich geringerer Personalaufwand notwendig.

Der kontinuierlich betriebene Rührkesselreaktor *(Durchflussrührkessel*, Abbildung 4) ist einer der zwei Grundtypen chemischer Reaktoren für den Fließbetrieb. Im Vergleich zum Chargenreaktor besitzt er lediglich zusätzliche Einlauf- und Ablaufvorrichtungen. Durch intensive Durchmischung wird dafür gesorgt, dass der Eingangsstrom sich sofort mit dem Kesselinhalt innig vermengt. Im Produktausgang liegt daher nahezu die gleiche Zusammensetzung vor wie im Kessel.
Obwohl dieser Reaktortyp für die Herstellung von Festharzen bereits signifikante prozesstechnische Verbesserungen bringen würde, sind, bedingt durch die relativ kleine Behälterwandungsfläche des Rührkessels, noch gravierende Probleme bei der Reaktortemperierung zu lösen.

Der zweite Grundtyp kontinuierlich betriebener *Reaktoren* ist der Rohrreaktor (Abbildung 5). In diesem Reaktor bewegt sich das Reaktionsgemisch vom Reaktoreingang entlang einer Rohrstrecke zum Ausgang und ändert dabei seine Zusammensetzung *örtlich längs des Rohres*. Da die Länge des Rohres im Verhältnis zu seinem Durchmesser sehr groß ist, ergibt sich eine weitaus größere Wärmeaustauschfläche, mit der die Temperierungsprobleme zu überwinden sind. Für die Festharzherstellung wirkt sich jedoch nunmehr die unzureichende Durchmischung des Reaktionsgemisches vor allem im ersten Abschnitt des Strömungsrohres nachteilig auf die Produktqualität aus.

Ausführungsformen chemischer Reaktoren, mittels derer die Anforderungen an die Temperierung und Durchmischung gleichermaßen zu erfüllen sind, können allerdings durch wohl überlegte Verschaltung oder Modifikation von Reaktorgrundtypen realisiert werden. Zwei Ausführungsformen kombinierter bzw. modifizierter Reaktoren sind im Fachgebiet Technische Chemie und Chemische Verfahrenstechnik für diese Zwecke als geeignet erarbeitet worden: der Schnecken-Schlaufenreaktor und der Rührkessel-Rohr-Reaktor (Abbildung 6 und 7).

Beim *Schnecken-Schlaufenreaktor* wird versucht, das Durchmischungsverhalten eines Rohrreaktors dem eines Durchflusskessels dadurch nahe zu bringen, dass man mittels einer zentral eingebauten Förderschnecke einen Zirkulations- oder Kreislaufstrom erzeugt, der das Reaktionsgemisch wieder an den Reaktoreingang zurückführt und mit dem zufließenden Eingangsstrom vermischt. Der Zirkulationsstrom kann dabei dem gewünschten Durchmischungsgrad entsprechend eingestellt werden.
Beim *Rührkessel-Rohr-Reaktor* vergrößert man auf der einen Seite die limitierte Wärmeaustauschfläche eines einzelnen Durchflussrührkessels durch Hintereinanderschaltung eines Durchflussrührkessels und eines Rohrreaktors mit insgesamt gleichem Reaktorvolumen. Auf der anderen Seite kann die unzureichende Durchmischung im Anfangsabschnitt eines einzelnen Rohrreaktors durch den vorgeschalteten Durchflussrührkessel vermieden und die Produktqualität positiv beeinflusst werden.

Sowohl das Schnecken-Schlaufenreaktor- als auch das Rührkessel-Rohrreaktor-Konzept werden derzeit für die Herstellung von Festharzen an der Universität Paderborn im Technikumsmaßstab intensiv beforscht und für eine Optimierung erschlossen.

Strategien in Informations- und Kommunikationsbranchen

Über den Umgang mit Netzwerkeffekten, Lock-in-Situationen und First-Mover-Vorteilen

Wie können Unternehmen aus Informations- und Kommunikationsbranchen Netzwerkeffekte, positives Feedback und Lock-in strategisch nutzen? Weshalb erscheint es notwendig, sich mit Wettbewerbsstrategien in diesen Branchen gesondert zu befassen? Gelten in der Informations- und Kommunikationsindustrie nicht die gleichen strategischen Maximen wie in jeder anderen Branche?

In seiner knapp dreißigjährigen Firmengeschichte hat das Softwarehaus SAP einen kometenhaften Aufstieg erlebt. Zwischen 1972 und 1999 stieg der Umsatz um mehr als das 16 000-fache. Die bestechende Geschäftsidee, eine integrierte Standardsoftware zu entwickeln, war nicht der alleinige Grund für diesen Erfolg. Warum haben sich nicht viele Unternehmen mit ähnlichen Produkten in diesem boomenden Markt behaupten können? Diese Frage lässt sich unter Berücksichtigung von Netzwerkeffekten, positivem Feedback und Lock-in-Situationen beantworten. SAP vertreibt mit seiner Software ein Produkt, dessen Kundennutzen mit wachsender Kundenzahl zunimmt. Je mehr Unternehmen SAP R3 einsetzen, desto leichter wird es für diese Unternehmen, mit Kunden bzw. Lieferanten zu kommunizieren, qualifizierte Arbeitskräfte einzustellen und mit anderen Unternehmen zu kooperieren. Mit zunehmender Kundenzahl steigt aber nicht nur der durchschnittliche Kundennutzen. Gleichzeitig sinken auf Grund des hohen Fixkostenanteils (Softwareentwicklungskosten) die Stückkosten. Kurz gesagt: Das Produkt wird wertvoller und zugleich billiger. Ferner können Kunden, die sich einmal für R3 entschieden haben, nicht kostenlos auf ein Konkurrenzprodukt umsteigen. Sie müssten ihre Kommunikationssysteme anpassen, ihre Mitarbeiter umschulen und Änderungen ihres Kooperationsdesigns vornehmen. Infolge dieser Wechselkosten können sich sogar Unternehmen mit besseren Produkten häufig nicht durchsetzen.

Netzwerkeffekte, positives Feedback und Lock-in sind die Ursache dafür, dass es in Informations- und Kommunikationsbranchen einerseits zu Gewinnexplosionen und schneeballartigem Unternehmenswachstum kommt, andererseits aber Unternehmen trotz überlegener Produkte wieder vom Markt verschwinden. Der vorliegende Beitrag analysiert diese Zusammenhänge und zeigt, welche Wettbewerbsstrategien in der Informations- und Kommunikationsbranche erfolgreich sind.

Positives Feedback: VHS gegen Beta

In der Informations- und Kommunikationsökonomie gibt es zwei Arten von Netzwerken: reale und virtuelle. In realen Netzwerken sind die Nutzer eines Produktes physisch miteinander verbunden. Klassische Beispiele sind Telefonnetze, Börsenhandelssysteme (vgl. Dietl/Pauli/Royer 1999) und das Internet. Bei virtuellen Netzwerken besteht keine physische, sondern nur eine logische Verbindung zwischen den Netzwerkmitgliedern. Beispielsweise bilden die Benutzer von CD-Playern, Videorekordern oder dem Betriebssystem MS Windows jeweils virtuelle Netzwerke.

Sowohl reale als auch virtuelle Netzwerke sind ökonomisch durch so genannte Netzwerkeffekte gekennzeichnet. Hierunter versteht man das Phänomen, dass der Durchschnittsnutzen jedes Netzwerkteilnehmers mit zunehmender Anzahl der Teilnehmer ansteigt. In realen Netzwerken steigert jeder neue Teilnehmer direkt den Nutzen aller Netzwerkteilnehmer. Dieser direkte Netzwerkeffekt lässt sich beispielsweise beim Telefonnetz leicht nachvollziehen. In virtuellen Netzwerken basiert der Netzwerkeffekt auf einer indirekten Nutzensteigerung. Mit zunehmender Netzwerkgröße steigt nämlich der Anreiz, Netzwerkkomplemente bereitzustellen. Nimmt beispielsweise die Anzahl der Benutzer von CD-Playern zu, dann steigen die Anreize, Komplemente wie z.B. CDs herzustellen. Die zunehmende Vielfalt und Verfügbarkeit von Komplementen resultiert wiederum in einem ansteigenden Durchschnittsnutzen für alle Netzwerkteilnehmer.

Netzwerkeffekte sind notwendige aber nicht hinreichende Bedingung positiver Feedback-Effekte. Hierunter versteht man die Kombination aus nachfrage- und angebotsseitigen Größenvorteilen (Economies of Scale). Auf der Nachfrageseite führen direkte

Prof. Dr. Helmut M. Dietl ist seit 1998 Inhaber des Lehrstuhls für Betriebswirtschaftslehre, insbesondere Organisation und Internationales Management an der Universität Paderborn. Seine Forschungsschwerpunkte umfassen u.a. die Gebiete Corporate Governance, Virtuelle Netzwerke und Internationales Management.

Dr. Susanne Royer ist seit 1995 als wissenschaftliche Mitarbeiterin am Lehrstuhl für Organisation und Internationales Management an der Universität Paderborn tätig. Im Frühjahr 2000 promovierte sie mit dem Thema „Strategische Erfolgsfaktoren horizontaler kooperativer Wettbewerbsbeziehungen". Zur Zeit lehrt und forscht Susanne Royer an der Queensland University of Technology in Brisbane, Australien.

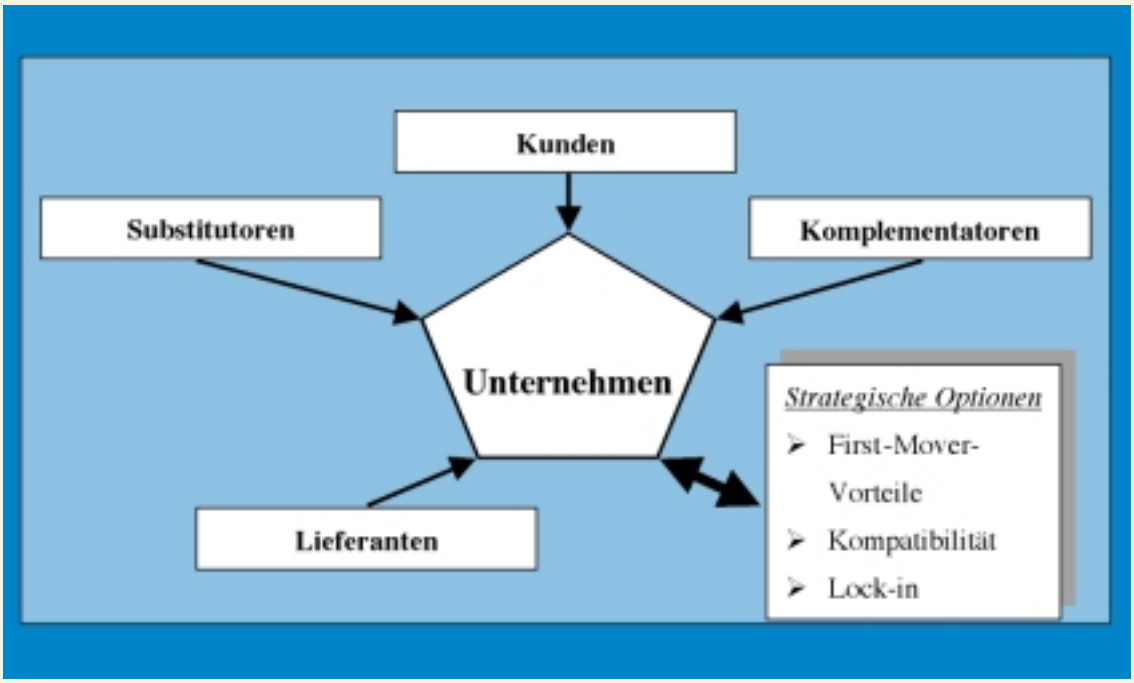


Abb. 1: Analyserahmen für Wettbewerbsstrategien in Informations- und Kommunikationsbranchen.

bzw. indirekte Netzwerkeffekte zu Größenvorteilen. Je größer das Netzwerk, desto größer die Zahlungsbereitschaft jedes Netzwerkmitglieds. Angebotsseitige Größenvorteile führen dazu, dass die Stückkosten eines Produktes mit steigender Produktionsmenge fallen. Dieser Effekt ist in Informations- und Kommunikationsbranchen stark ausgeprägt. Die Produktion von Informations- und Kommunikationsgütern ist häufig mit sehr hohen Fixkosten verbunden. Beispielsweise belaufen sich die Fixkosten bei der Produktion von Kinofilmen auf mehrere 100 Millionen Euro. Zugleich sind die Grenzkosten häufig nahe Null. Die Kopie eines Filmes auf eine Videokassette kostet nur wenige Cents. Die Kombination aus nachfrage- und angebotsseitigen Größenvorteilen resultiert in einer unaufhaltsamen Vorteilsspirale. Nachfrageseitig steigt über die Netzwerkeffekte die Zahlungsbereitschaft der Kunden. Zugleich sinken infolge der angebotsseitigen Größenvorteile die Stückkosten. Diese positiven Feedback-Effekte führen dazu, dass Unternehmen bei wachsenden Marktanteilen immer höhere Gewinnmargen erzielen.

Positives Feedback wirkt nicht im Sinne aller Wettbewerber in einem Markt. Durch positives Feedback werden starke Unternehmen stärker und schwache Unternehmen schwächer. Es kommt häufig zur Dominanz einer Technologie im Markt. Kämpfen zwei oder mehr Unternehmen um einen Markt mit großem positiven Feedback, kann es meist nur einen Gewinner geben. Der Ausgang des Kampfes zwischen dem VHS-System von JVC und dem Beta-System von Sony auf dem Markt für Videorekorder illustriert ein positives Feedback (für JVC). Nachdem JVC eine installierte Basis von Videorekorder-Nutzern erreicht hatte, begann das positive Feedback zu wirken: Je mehr Videorekorder es gab, desto größer wurde das Angebot an Videofilmen. Die große Auswahl an Videofilmen sorgte wiederum dafür, dass immer mehr Menschen einen Videorekorder kauften (vgl. Shapiro/Varian 1999, S. 96). Das führte zu Größenvorteilen bei den Herstellern der Geräte und bei den Herstellern von Komplementen. Die Stückkosten der Videorekorder und der Komplemente (Videokassetten, Videofilme etc.) sanken. Ferner nahm die Vielfalt der Komplemente zu. Beispielsweise wurde zusätzliche Ausrüstung für das Drehen von Heimvideos und für die Programmierung der Rekorder entwickelt. Hierdurch wurde der Kundennutzen noch größer, die Stückkosten sanken noch stärker.

Ein weiteres Charakteristikum der Informations- und Kommunikationsökonomie sind so genannte Lock-in-Effekte. Hierunter versteht man Wechselkosten, die einen Kunden davon abhalten, von einem Produkt auf ein anderes umzusteigen. Ein historisches Beispiel für eine Lock-in-Situation ist die QWERTY-Anordnung der Tastaturen auf Schreibmaschinen und Computern (bzw. im deutschsprachigen Raum die QWERTZ-Anordnung). Diese Anordnung hat sich deshalb durchgesetzt, weil sie bei mechanischen Schreibmaschinen verhinderte, dass sich die einzelnen Buchstaben ineinander verhakten. Dieses Problem besteht heute nicht mehr und trotzdem finden wir die QWERTY-Anordnung der Tasten auf unseren Computertastaturen. Ein typischer Fall von Lock-in: Es erscheint zu teuer und zu unbequem, auf eine andere Tastatur umzusteigen (vgl. David 1985, S. 332ff.).

Strategische Gestaltungspotenziale

Unternehmen müssen Netzwerke etablieren, um von Netzwerkeffekten und positivem Feedback zu profitieren. Dabei ist es wichtig, dass nicht nur Wettbewerber, Abnehmer und Zulieferer in die strategischen Überlegungen einbezogen werden. Wie Abbildung 1 zeigt, müssen darüber hinaus Komplemente, Substitute und Kooperationspartner (vgl. Royer 2000) sowie First-Mover-Vorteile, Kompatibilitätsprobleme, Lock-in- und Signalisierungseffekte berücksichtigt werden.

First-Mover-Vorteile oder „Wer zuerst kommt, mahlt zuerst“

In Informations- und Kommunikationsbranchen ist es häufig entscheidend, als Erster eine Technologie auf den Markt zu bringen, um dauerhafte Wettbewerbsvorteile zu realisieren. Ist erst eine installierte Basis von Nutzern erreicht, tun sich die so genannten Second-Mover schwer, gleichzuziehen. Der Grund dafür liegt in den erläuterten Netzwerkeffekten in Verbindung mit positivem Feedback. Hat der First-Mover mit der eigenen Technologie bei den Kunden Wechselkosten aufgebaut, gehen diese Wechselkosten in die Kaufüberlegungen der potenziellen Kunden des Second-Mover ein. Durch die frühe Marktpräsenz einer Technologie können langfristige Vorteile erlangt werden. Bevor Wettbewerber im Markt sind, ist es bereits möglich, eine kritische Masse an Nutzern aufzubauen und Wechselkosten für

diese zu erzeugen. In einer solchen Situation haben die Second-Mover mit hohen Markteintrittsbarrieren zu kämpfen und können sich oft selbst mit einem besseren Produkt nicht mehr durchsetzen. Diesen Zusammenhang spiegelt auch das oben aufgezeigte QWERTY-Beispiel wider. Es ist entscheidend, der Erste zu sein, um genügend Kunden zu gewinnen, damit das positive Feedback im Sinne des eigenen Unternehmens arbeitet.
Gerade für Informationsanbieter eröffnen sich vor dem Hintergrund der geschilderten Kostenstruktur mit hohen Fixkosten und geringen variablen Kosten vollkommen neue Vertriebsmöglichkeiten (vgl. Dietl 1999). Insbesondere die Gewährung der Möglichkeit, sich Produkte zunächst kostenlos anzusehen, erscheint als viel versprechende Vertriebsmöglichkeit für Unternehmen, die frühzeitig im Markt sind. So wird es möglich, eine Lock-in-Situation möglichst vieler Kunden zu erreichen. Mit den Updates, verbesserten Versionen oder zusätzlichen „Features" des Produkts lassen sich Gewinne realisieren. Ist eine kritische Masse von Nutzern und eine Lock-in-Situation erreicht, wird es möglich, auch für das Basisprodukt profitable Preise zu setzen.

Komplemente oder „Manchmal muss man mit der Wurst nach dem Schinken werfen"

Netzwerke bestehen aus konkurrierenden und komplementären Komponenten. Bei der Entwicklung von Strategien ist die Nutzung von Hebeleffekten einer vorhandenen installierten Basis einzubeziehen. Diese Nutzung kann in dem Verkauf komplementärer Produkte liegen. Unternehmen müssen erkennen, dass der Wert ihres Produktes massiv von der Quantität und Qualität der Komplemente abhängt. Ein Beispiel für Komplemente sind Hardware und Software. Es liegt praktisch ein Ei-Henne-Problem vor: Softwareentwickler haben ohne bestehende Hardware keinen Anreiz, Software zu entwickeln. Kein Kunde wird in Hardware investieren, für die keine ausreichende Menge an Software zur Verfügung steht. Vor diesem Hintergrund kann es für Unternehmen der Informations- und Kommunikationsbranche sinnvoll sein, die Hersteller von Komplementen für den Eintritt in den Markt zu bezahlen, um später selbst größere Gewinne erzielen zu können.
Diese Strategie wählte beispielsweise das Unternehmen 3DO, das Anfang der neunziger Jahre eine 32-Bit-CD-ROM-Hardware-Softwaretechnologie für die nächste Generation der Videospiele entwickelt hatte. 3DO hat für eine Gebühr von drei Dollar Lizenzen an Softwarehäuser zur Entwicklung der 3DO-Spiele vergeben. Um Software verkaufen zu können, musste es Kunden geben, die bereit waren, die Hardware zu erstehen. Diese frühen Nutzer hatten jedoch das Problem, dass es nur wenig Software gab. Um eine Initialzündung zu erreichen, musste die Hardware so billig wie möglich sein. 3DO hat deshalb kostenlose Lizenzen für die Herstellung der Hardwaretechnologie vergeben, wodurch Hardwarehersteller wie Matsushita (Panasonic), AT&T, GoldStar, Sanyo, Samsung und Toshiba in das Geschäft eintraten. Weil jede 3DO-Software auf jeder 3DO-Hardware lief, fand der Wettbewerb zwischen den Hardwareherstellern allein auf der Basis der Kosten statt. 3DO wollte so erreichen, dass die Hardware zu einem Massenprodukt wurde, was wiederum den Preis der Komplemente sinken ließ. Darüber hinaus war 3DO der Ansicht, dass die Hardware sogar unterhalb der Kosten verkauft werden musste, damit Bewegung in den Markt kam. Deshalb erhielten die Hardwarehersteller eine Prämie pro verkauftem Gerät. Gleichzeitig wurden die Lizenzgebühren für die Softwarehersteller verdoppelt. Mit den zusätzlichen Lizenzgebühren wurden die Hardwarehersteller subventioniert (vgl. Brandenburger 1995). 3DO zahlte also dafür, dass es entsprechende Komplemente zu den 3DO-Produkten auf dem Markt gab. Für 3DO war es entscheidend, den Wert der eigenen Technologie zu steigern, nicht die eigene Kontrolle über diese Technologie. Es kann sinnvoll sein, anderen Unternehmen bestimmte Geschäfte zu überlassen. Auch bei der SAP AG ist eine solche Strategie zu erkennen. SAP hat das Beratungsgeschäft zu einem großen Teil Beratungs- und Wirtschaftsprüfungsgesellschaften überlassen. Diese haben zum einen Unternehmen bei der Softwareauswahl beraten und zum anderen weitere Komplemente entwickelt, die die SAP-Software noch attraktiver werden ließen.

Kompatibilität oder „Revolution vs. Evolution"

Einheitliche Gesamtsysteme bestehen aus kompatiblen Produkten. Führt ein Unternehmen ein neues Produkt ein, sieht es sich mit einem Trade-off zwischen Kompatibilität und revolutionären Produkten konfrontiert. Es muss sich entscheiden, ob eine so genannte Evolutions- oder Revolutionsstrategie verfolgt werden soll (vgl. Shapiro/Varian 1999, S. 192 ff.). Bei der Evolutionsstrategie entstehen keine Schwierigkeiten hinsichtlich der Rückwärtskompatibilität der Produkte. Allerdings muss meist eine eingeschränkte Leistungsverbesserung in Kauf genommen werden. Die Programmierung der dBASE-Datenbank war z.B. beschränkt, weil jede neue dBASE-Version in der Lage sein musste, alle früher geschriebenen dBASE-Programme zu unterstützen. Das hat sich negativ auf die Leistung ausgewirkt. Programme wie Access konnten dBASE aus diesem Grunde verdrängen. Die Umsetzung einer Revolutionsstrategie führt zu eindeutigen Leistungsverbesserungen, es kommt jedoch zu Kompatibilitätsproblemen mit bereits bestehenden Produkten. Ein Beispiel für eine Revolutionsstrategie ist die Einführung der CD-Technologie. Die sehr viel bessere Klangqualität der CD-Player überzeugte viele Nutzer, die hohen Wechselkosten von einem Plattenspieler auf einen CD-Player in Kauf zu nehmen. Eine Revolutionsstrategie ist risikoreicher als eine Strategie der Evolution. Für die Durchführung einer Revolution ist eine große Menge von Abnehmern notwendig. Außerdem ist diese Strategie ohne Kooperationspartner meist nicht zu bewältigen.

Lock-in-Management oder „Mit gehangen, mit gefangen"

Die Erschaffung einer Kunden-Lock-in-Situation stellt hinsichtlich der Strategieentwicklung in Informations- und Kommunikationsbranchen eine wichtige Erfolgskomponente dar. Dabei existiert neben der Möglichkeit, eine natürliche Lock-in-Situation zu erschaffen (beispielsweise durch Software, die spezifische Lernkosten impliziert), auch die Möglichkeit, eine Situation künstlichen Lock-ins zu etablieren. So hat beispielsweise das Unternehmen Webmiles diesen Gedanken für den Bereich des E-Commerce entdeckt. Für Ware, die über das Netz gekauft wird, werden jeweils so genannte Webmeilen gutgeschrieben, wenn die Einkäufe bei Partnern von Webmiles getätigt werden und sich der Nutzer als „Webmiles-User" identifiziert. Die Webmeilen können gesammelt und später gegen Ware eingetauscht werden. Für 5 000 Webmeilen gibt es beispielsweise eine Minolta Riva

Zoom-Kamera. Dem Kunden entstehen durch diese Loyalitätsprogramme künstlich erzeugte Wechselkosten. Hat er z.B. bereits 4 800 Webmeilen gesammelt, wird er die 200 fehlenden Meilen haben wollen, um die Kamera zu bekommen. Sowohl natürliche Wechselkosten (z.B. Wechsel von einem Plattenspieler auf einen CD-Player) als auch künstliche Wechselkosten (z.B. Mengenrabatte) führen zu Kunden-Lock-in-Situationen.

Unternehmen können Wettbewerbsvorteile generieren, wenn ein Lock-in ihrer Kunden vorliegt. Deshalb haben Unternehmen große Anreize, eine solche Situation herzustellen. Lock-in muss für den Kunden nicht nur negative Folgen haben, sondern die Kunden können auch von einer langfristigen Beziehung zu einem einzigen Zulieferer profitieren. Insbesondere Großkunden haben in einer solchen Situation häufig beträchtlichen Einfluss auf die weitere Entwicklung des Produktes. Potenzielle Kunden, die eine Lock-in-Situation antizipieren, d.h. sich darüber bewusst sind, dass sie hohe Wechselkosten haben werden, wenn sie sich für ein Produkt entschieden haben, könnten sich gegen ein Produkt entscheiden, das eine Lock-in-Situation zur Folge hat. Eine andere Möglichkeit für die Kunden liegt darin, aus der Antizipierung der zukünftigen Lock-in-Situation Kapital zu schlagen, indem sie sich ihren Einstieg in das Produkt von dem entsprechenden Anbieter „bezahlen" lassen. Unternehmen müssen in ihren Strategien also auch berücksichtigen, dass Kunden zukünftige Lock-in-Situationen vorhersehen und versuchen werden, sich nicht ohne Absicherung in derartige Situationen zu begeben.

Signalisierung oder „Die Spatzen müssen es von den Dächern pfeifen"

Es bedarf von Unternehmensseite her bestimmter Signale, die den Kunden veranlassen, in das jeweilige Produkt zu investieren. Hinsichtlich einer erfolgreichen Signalisierung ist die Glaubwürdigkeit des signalisierenden Unternehmens entscheidend (vgl. Picot/Dietl/Franck 1999, S. 90 ff.). Das entsprechende Unternehmen muss den (potenziellen) Kunden glaubwürdig versichern, dass das eigene Produkt sich gegen die Produkte der Wettbewerber durchsetzen wird. Eine Möglichkeit, das gegenüber Kunden und auch Wettbewerbern zu tun, besteht in der Ausweitung der Produktionskapazitäten, wodurch eine große Nachfrage signalisiert wird. Beispielsweise hat Philips signalisiert, dass sich die CD-Technologie durchsetzen wird, indem das Unternehmen frühzeitig große Produktionsstandorte für die CD-Produktion errichtet hat. Auf diese Weise wurde ein glaubwürdiges Signal an potenzielle Käufer ausgesandt, dass die damals neue Technologie das Potenzial hatte, sich am Markt durchzusetzen.

Konsumenten profitieren im Hinblick auf Informationstechnologien häufig, wenn sie ein gängiges Format oder System nutzen. Auch das Bemühen, einen Standard zu entwickeln, kann das Signal an die Kunden senden, dass es sich lohnt, in ein Produkt zu investieren. Ein Beispiel hierfür liefert der Wettbewerb zwischen Beta- und VHS-System im Bereich der Videorekorder. JVC gelang es durch die Bildung verschiedener Allianzen und das Bemühen um die Etablierung des VHS-Systems als Standard zu signalisieren, dass es sich lohnt, in die VHS-Technologie zu investieren. Das Gleiche gelang Sony für das konkurrierende Beta-System nicht. Zwar war das Beta-System dem VHS-System technisch ebenbürtig, aber JVC hat die glaubwürdigeren Signale gesetzt als Sony.

Eine Strategie des Preiskampfes kann ebenfalls Signalwirkung haben. Drohen neue Wettbewerber in einen Markt einzutreten, kann eine zeitweise Preissenkung für ein bereits etabliertes Unternehmen von Vorteil sein. Auf diese Weise kann es die eigene installierte Basis vergrößern und eine unverwundbare Marktposition erreichen. Ist eine solche Position erreicht, können die Preise wieder erhöht werden. Ein glaubwürdiges Signal, das Kunden zur Investition in ein Produkt bringt, kann auch darin liegen, dass ein Produkt verliehen und nicht verkauft wird. Wird diese Alternative gewählt, geht allerdings die Möglichkeit verloren, einen (materiellen) Kunden-Lock-in zu erschaffen. Bei dieser Strategie nimmt der Verkäufer das Risiko der frühen Konsumenten auf sich. Sind mit dem Produkt jedoch immaterielle Wechselkosten wie z.B. Lernkosten verbunden, kann auch in dieser Situation ein (immaterieller) Lock-in erschaffen werden.

Große Kundenbasis entscheidend

In Informations- und Kommunikationsbranchen sind Netzwerkeffekte, positives Feedback und Lock-in von herausragender strategischer Bedeutung. Unternehmen dieser Branchen müssen versuchen, möglichst rasch eine große Kundenbasis aufzubauen. Mittels Lock-in-Effekten kann diese Kundenbasis langfristig auch gegenüber Konkurrenten mit besseren Produkten verteidigt werden. Da viele Informations- und Kommunikationsprodukte nur in Verbindung mit entsprechenden Komplementen Werte schaffen, ist das Komplementmanagement einer der wichtigsten strategischen Erfolgsfaktoren. Die Bedeutung des koordinierten Angebots von Komplementen wird heute schon den Kleinsten bewusst: Die 150 verschiedenen Pokémons von Nintendo tummeln sich nicht nur auf Tauschkarten, sondern sind gleichzeitig in Videospielen, Kinofilmen, Comic-Heftchen und Büchern präsent.

Literatur

Brandenburger, A. M. (1995): The Right Game: Power Play (C): 3DO in 32-bit Video Games, Harvard Business Case Study (9-795-104), 1995.

David, P. A. (1985): Clio and the Economics of QWERTY, in: Economic History, Vol. 75, No.2, 1985, S. 332-337.

Dietl, H. M. (1999): Selling Information on the Internet, in: Hitotsubashi Journal of Commerce and Management, 34, 1999, S. 1-14.

Dietl, H. M./Pauli, M./Royer, S. (1999): Internationaler Finanzplatzwettbewerb: Ein ressourcenorientierter Vergleich, Gabler Verlag 1999.

Picot, A./Dietl, H. M./Franck, E. (1999): Organisation – Eine ökonomische Perspektive, Schäffer-Poeschel, 2. Auflage 1999.

Royer, S. (2000): Strategische Erfolgsfaktoren horizontaler kooperativer Wettbewerbsbeziehungen, Rainer Hampp Verlag 2000.

Shapiro, C./Varian, H. R. (1999): Information Rules: A Strategic Guide to the Network Economy, Harvard Business School Press, 1999.

Erdtangente und Stille Post

Eine Retrospektive mit sechs künstlerischen Arbeiten und drei kleinen Texten

Kunst ist neben der Wissenschaft ein wichtiges System zur Entwicklung und Bereitstellung von Kenntnissen und Wissen. Die Wissenschaft bemüht sich darum, die Erkenntnisse und Aussagen so zu formulieren, dass sie für alle gelten, dass sie allen einleuchten. Sie ist auf der Suche nach der allgemeinen, intersubjektiven oder objektiven Erkenntnis. Diese Erkenntnis muss nachprüfbar und wiederholbar sein, eben wissenschaftlich. Wissenschaftlichkeit ergibt sich dadurch, dass bestimmte Regeln und Vorgehensweisen eingehalten werden. Das System Wissenschaft wacht (hoffentlich) sehr streng darüber, dass die Regeln eingehalten werden. Künstler bemühen sich um individuelle, subjektive Sichtweisen. Sowohl auf der Seite des Künstlers als auch auf der Seite des Betrachters geht es darum, immer wieder neue Zugangs- und Sichtweisen zu erfinden. Die Kunst ermöglicht auf ihre ganz spezifische Art und Weise, neue Erfahrungen und Erkenntnisse.

Meine Arbeiten

entwickele ich so, dass sie auch ein anderer verwirklichen kann. Ich suche nach der „eleganten" ökonomischen Lösung. Die Ausführung soll leicht, einfach und schnell sein; mit dem richtigen Verhältnis von Aufwand und Ergebnis. Ich suche nach Konzepten und Ideen, die vor mir noch keiner formuliert und verwirklicht hat. Manchmal stelle ich mir die Möglichkeiten der Kunst wie eine Landkarte am Anfang des 19. Jahrhunderts vor, mit vielen weißen, unerforschten Flecken. Das Problem ist nur, dass auf der Karte der Kunst die weißen Flecken nicht zu sehen sind, man muss sie erst finden. Von Beginn der 80er- bis in die Mitte der 90er-Jahre habe ich mich hauptsächlich mit der Erde als großem geometrischen Körper befasst. So bin ich 1982 mit dem Auto so schnell wie möglich über drei Längengrade von Ost nach West gefahren, ein vergeblicher Versuch, die Erddrehgeschwindigkeit auszugleichen und die Sonne einzuholen. 1993 habe ich eine Tangente an die Erde angelegt. Seit den 90er Jahren arbeite ich an Skulpturen und Zeichnungen, die ich „Stille Post" nenne. Die Formen ergeben sich aus der Ungenauigkeit bei der Übertragung. Ich erfinde die Formen nicht, ich erforsche, was bei einer einfachen Vorgehensweise – der Kopie von der Kopie – entsteht. Grundlage ist keine Formvorstellung, kein gestalterischer Entwurf, kein Bild, sondern eine bestimmte Methode. Entscheidend für den Erfolg ist, dass ich mich möglichst ehrlich und „dumm" wie eine Maschine verhalte ... Meine Arbeiten sind in erster Linie die Konzepte, die ich entwickle. Viele meiner Arbeiten lassen sich deshalb ohne Bilder erzählen, ohne dass ein entscheidender Verlust auftreten würde. Ideen sind immer „sauberer" als die Ausführung. Weil ich neugierig bin, führe ich sie trotzdem selbst aus.

Prof. Franz Billmayer
Ist Künstler und Kunsterzieher. Seit 1998 lehrt er im Fachbereich Kunst, Musik und Gestaltung Kunst und ihre Didaktik mit Schwerpunkt Bildhauerei.

Kunst

Der Kunstdiskurs handelt im Wesentlichen davon, was Kunst ist. Spätestens seit der Romantik geht man davon aus, dass Kunst nicht völlig erschöpfend definiert werden kann. Das hängt zum einen damit zusammen, dass die Avantgarde von innen heraus dauernd versucht, die Grenzen des Systems anzugreifen. Diese Grenzzwischenfälle führen dazu, dass sich der Bereich der Kunst thematisch und medial dauernd ausdehnt. Zum anderen versuchen die „klassischen" Definitionen, den Kunstbegriff an bestimmten Eigenschaften festzumachen. So wie man das Rot eines Bildes an bestimmten Eigenschaften festmachen kann, meint man, müsse man das auch mit seinem Kunstcharakter leisten. Das geht nicht; denn Kunst ist keine Eigenschaft, sondern die Art und Weise, wie wir mit bestimmten Gegenständen umgehen. So lässt sich Kunst vom Gebrauch, den wir von Kunstwerken machen, definieren. Kunst ist das, was wir als Kunst verwenden. Kunst ist eine Gebrauchsanweisung.
Seit der Ausstellung eines Urinoirs mit dem Titel „fontaine" im Jahre 1917 – von Marcel Duchamp als Dada-Gag in New York inszeniert – ist der Kunstwelt immer klarer geworden, dass jeder Gegenstand oder Sachverhalt grundsätzlich als Kunstwerk betrachtet werden kann. Seither müssen Kunstwerke nicht mehr unbedingt hergestellt werden, man kann sie auch bloß hinstellen. Kunstwerke sind dazu da, interpretiert zu werden. Dazu muss man ihnen einen tieferen Sinn unterstellen. Andere Äußerungen, die Menschen hervorbringen, damit sie gedeutet werden, Zeichen also, haben im Idealfall eine genaue, allgemein akzeptierte Interpretation. Wenn wir sie verwenden, wollen wir möglichst genau herausbekommen, was der Zeichenproduzent gemeint hat. Kunstwerke verwenden wir anders, ihre Qualität besteht gerade darin, dass sie beim Betrachter ganz unterschiedliche Deutungen hervorrufen. Selbst wenn eine Deutung der anderen widerspricht, muss sie jeweils nicht falsch sein. In der Kunst geht es darum, neue Sichtweisen zu finden und sie in die Welt zu bringen. Kunst dient dazu, den Raum der Möglichkeiten zu erweitern.

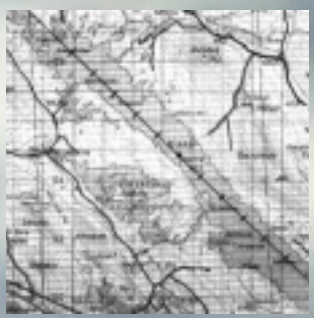
A TANGENT TOUCHING EARTH

Im März 1993 habe ich auf dem Eis des Västra Kikkejaure in Lappland einen Tangentialpunkt bestimmt. Von dort aus haben wir in zwei Richtungen jeweils eine fünf Kilometer lange Linie ausgemessen und jeweils in einen Kilometer lange Teile unterteilt. An den einzelnen Punkten habe ich Dreiecke aufgestellt, deren Spitzen den Verlauf der Tangente markierten.

Entfernung vom Tangentialpunkt 1 000 Meter – Abstand der Tangente von der Erdoberfläche ca. 8 Zentimeter.

Entfernung vom Tangentialpunkt 5 000 Meter - Abstand der Tangente von der Erdoberfläche ca. 186 Zentimeter.

Vermessung auf dem Västra Kikkejaure im März 1993.



Lage der Tangente auf dem Västra Kikkejaure.

Teleaufnahme der Markierungen, zu sehen sind die Dreiecke ab 2 000 Meter vom Tangentialpunkt.

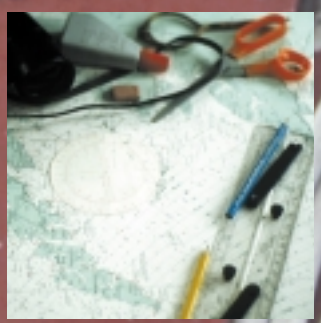
24 HOURS TOWARDS THE SUN 1987

Zur Zei der Mitternachtssonne nördlich des Polarkreises 24 Stunden mit dem Schiff auf die Sonne zufahren

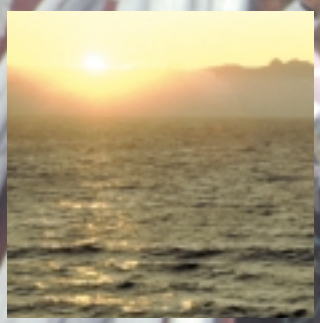
Diese Arbeit habe ich am 11. und 12. Juni 1987 mit einem Fischerboot auf dem Vestfjord zwischen den Lofoten und dem Festland realisiert. Es entstand eine riesige Zeichnung der Erddrehung auf der Erdoberfläche. Das Schiff war das Zeichenwerkzeug, das Meer der Bildträger. Die Spur der Zeichnung, das Kielwasser, löste sich sehr schnell wieder auf, wurde aber in einem 24-stündigen Video festgehalten. Jede Stunde wurde ein Foto in Richtung Sonne gemacht. Außerdem wurden jeweils stündlich zwei Flaschen ins Meer geworfen. Beide enthielten eine kurze Beschreibung des Projektes und die Positionsangabe des Schiffes; die eine versank als Markierung (der Zeichnung) auf den Meeresgrund, die andere trieb als Flaschenpost fort. Die beiden Flaschen stehen für den beweglichen und den festen Teil der Zeichnung. Auf vier der 25 Flaschpostbriefe habe ich eine Nachricht bekommen. Die Fundorte lagen z.T. mehrere hundert Kilometer von der Aussetzstelle entfernt.

Seekarte des Vestfjords, südwestlich der Lofoten. In der Mitte sieht man den „idealen" Kreis. Aufgrund von Meeresströmungen fuhren wir im Endeffekt eine Eiform.

24 Stunden auf einem Fischerboot bei langsamer Fahrt.



Kartentisch auf der Trollfjord.



Die Mitternachtssonne

Schnelle Fahrt zur Startposition.

THE WORLD IS STILL GOING AROUND

Bei 44,360079° geografischer Breite dreht sich ein Punkt auf der Erdoberfläche im Verhältnis zur Rotationsachse der Erde mit 331 Meter pro Sekunde, mit der durchschnittlichen Schallgeschwindigkeit. Dieser Breitengrad verläuft etwas südlich von Ravenna. Im Oktober 1989 haben wir dort von Ost nach West neun Punkte eingemessen, jeweils mit einem Abstand von 331 Metern. Die Strecke war insgesamt fast drei Kilometer lang. Der Schall braucht von einem Ende zum anderen neun Sekunden. Zur Zeit des Sonnenuntergangs haben sich an den Punkten jeweils zwei Personen mit Fotoblitzgeräten aufgestellt, eine in Richtung Westen, eine in Richtung Osten. Mit einer Schiffshupe habe ich gehupt und sobald die Posten den Schall hörten, lösten sie ihre Blitzlichter aus. Eine menschliche Laufblitzanlage, die die Geschwindigkeit der Erdrotation anzeigte.

Wenn die Posten den Schall hören, lösen sie ihre Blitzgeräte aus.

Auf 44,360079° nördlicher Breite werden neun Punkte eingemessen. Die Abstände betragen jeweils 331Meter – die Strecke, die der Schall durchschnittlich in einer Sekunde zurücklegt.

Die Ost-West-Richtung des Breitengrades wird bestimmt.

Signalhupe (Foto vom Dokumentarvideo).

44,360079° nördliche Breite (bei Ravenna).

STILLE POST

Ich zeichne eine einfache Form, einen Kreis oder eine Ellipse, auf eine Holzplatte und säge sie mit der Bandsäge so genau wie möglich aus. Dann lege ich sie auf die nächste Holzplatte und zeichne sie darauf ab. Diese Zeichnung säge ich aus, zeichne sie ab und säge wieder ... die Kopie von der Kopie von der Kopie. Die Holzplatten staple ich aufeinander zu baumgleichen Skulpturen, die Umrisse zeichne ich auf Papier übereinander.

Stille-Post-Stamm 200 Sperrholzplatten ausgesägt und verleimt.

Stille-Post-Ellipse 800 Spanplatten (6mm) – Ausgangsform war eine Ellipse.

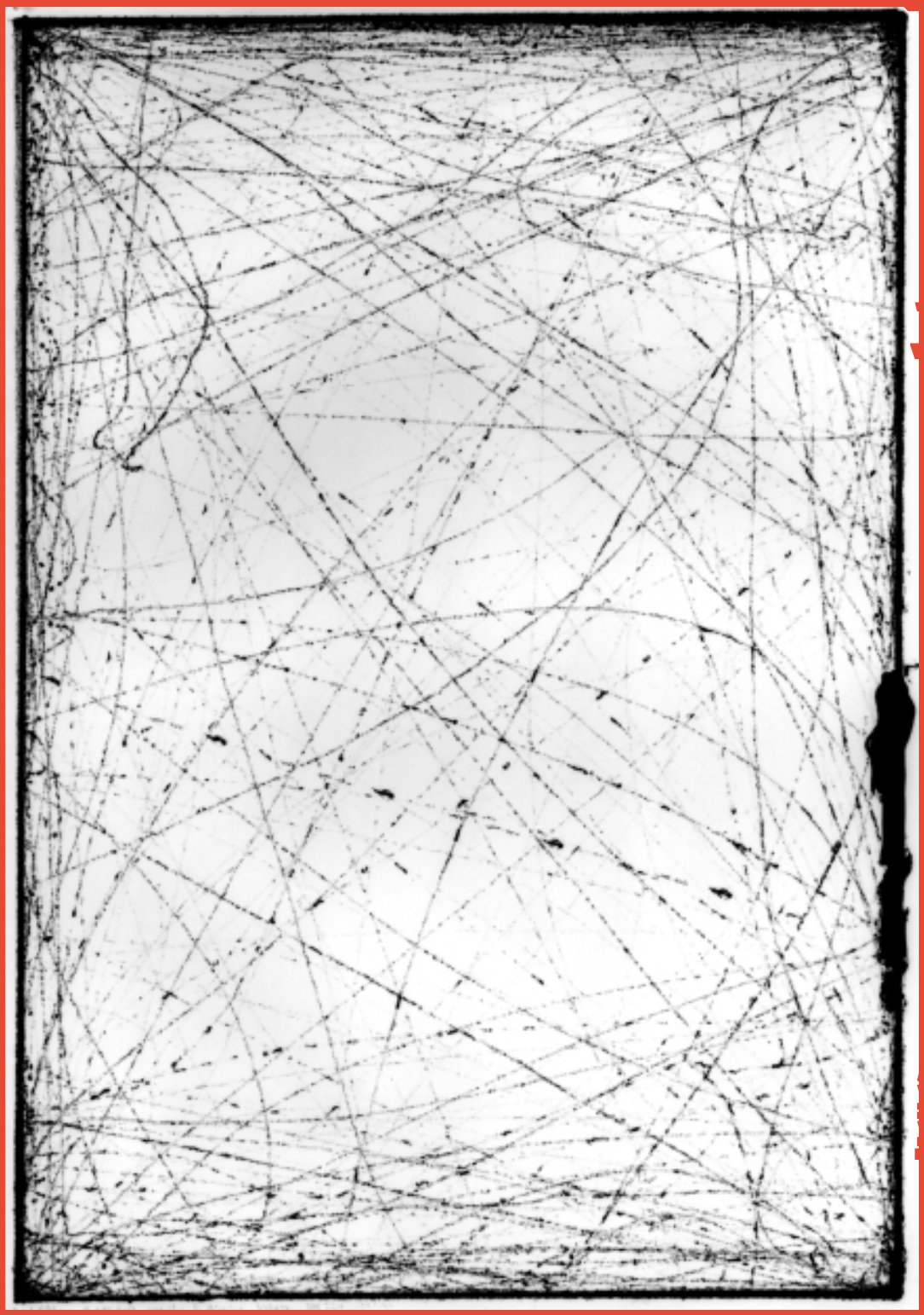
Die Umrisse aller Holzplatten der linken Skulptur übereinander gezeichnet. (Bleistift auf Papier).

FISHSTICKS RETURN TO SEA 1988

Am 16. Mai 1988 habe ich in Dorfen in Oberbayern eine Schachtel mit tiefgefrorenen Fischstäbchen gekauft. Diese habe ich mit Hilfe einer Kühltasche und Zwischenlagerungen in Kühlschränken und Tiefkühltruhen auf einer kombinierten Bahn-, Schiffs- und Flugreise auf die Lofoten transportiert. Dort habe ich sie dann am 25. Mai 1988 in Svolvær aus der Packung ins Meer geworfen.

AUTOBILDER

Im Gepäckraum meines Autos hat in den Jahren 1990 und 1991 ein Rahmen gelegen. In diesen Rahmen habe ich täglich ein Blatt Papier gelegt und darauf etwas Druckerschwärze geschmiert. Während meiner Autofahrten lief eine Stahlkugel im Rahmen frei über das Papier. Aus dem „Depot“ hat sie Farbe mitgenommen und die Bilder gezeichnet. Jeden Tag ein Bild.



21.1.1991 Wörth – Bad Aibling – Wörth – 124km.

Energiebilanzen landwirtschaftlicher Betriebe

Energetischer Bewertungsansatz landwirtschaftlicher Produktionssysteme an ausgewählten Beispielen

In der öffentlichen Diskussion nimmt die Umweltproblematik (Ozonloch, Treibhauseffekt, Waldsterben, Bodenerosion, Nitratverlagerung) einen immer höheren Stellenwert ein. Konzentrationen in der Viehhaltung, Spezialisierung, zunehmende Größe von Betrieben und Einzelflächen, die Flurbereinigung und der Einsatz moderner Landtechnik mit immer größeren und schlagkräftigeren Maschinen können als Quellen zunehmender Umweltbelastungen durch die moderne Landbewirtschaftung in Betracht gezogen werden. Seit den Ölkrisen in den 70er-Jahren werden Energiebilanzen in allen Bereichen unserer Gesellschaft immer häufiger aufgestellt, analysiert und diskutiert. Dies trifft auch in zunehmendem Maße für die Landwirtschaft zu.

Ursprünglich dienten Energiebilanzierungen im landwirtschaftlichen Bereich dem Vergleich der Effizienz der Bereitstellung von Energie aus der in der landwirtschaftlichen Produktion erzeugten Biomasse und anderen regenerativen und fossilen Energieträgern. Die Abschätzung der Klima- und Umweltrelevanz in der landwirtschaftlichen Erzeugung wird auf der Grundlage von Energiebilanzen und hier in erster Linie in der pflanzlichen Produktion diskutiert. Ihre Anwendung zur Beurteilung nachhaltiger Wirtschaftsweisen ist insbesondere im Rahmen von Öko-Audits und Öko-Bilanzen zu finden.

Prof. Dr. Norbert Lütke Entrup ist Leiter des Forschungsprojektes „Energiebilanzen landwirtschaftlicher Betriebe".

Der Ansatz

Bilanzierungen lassen sich für landwirtschaftliche Betriebe auf verschiedenen Ebenen durchführen. Bei Hoftorbilanzen wird der gesamte Betrieb in die Betrachtungen aufgenommen, Stallbilanzen erfassen nur den Teilbereich der Tierproduktion – des Stalles. Über Flächen- und Schlagbilanzen wird die pflanzliche Produktion erfasst. Die Ergebnisse der einzelnen Flächen- und Schlagbilanzen können zu Bilanzen für Kulturarten, über ganze

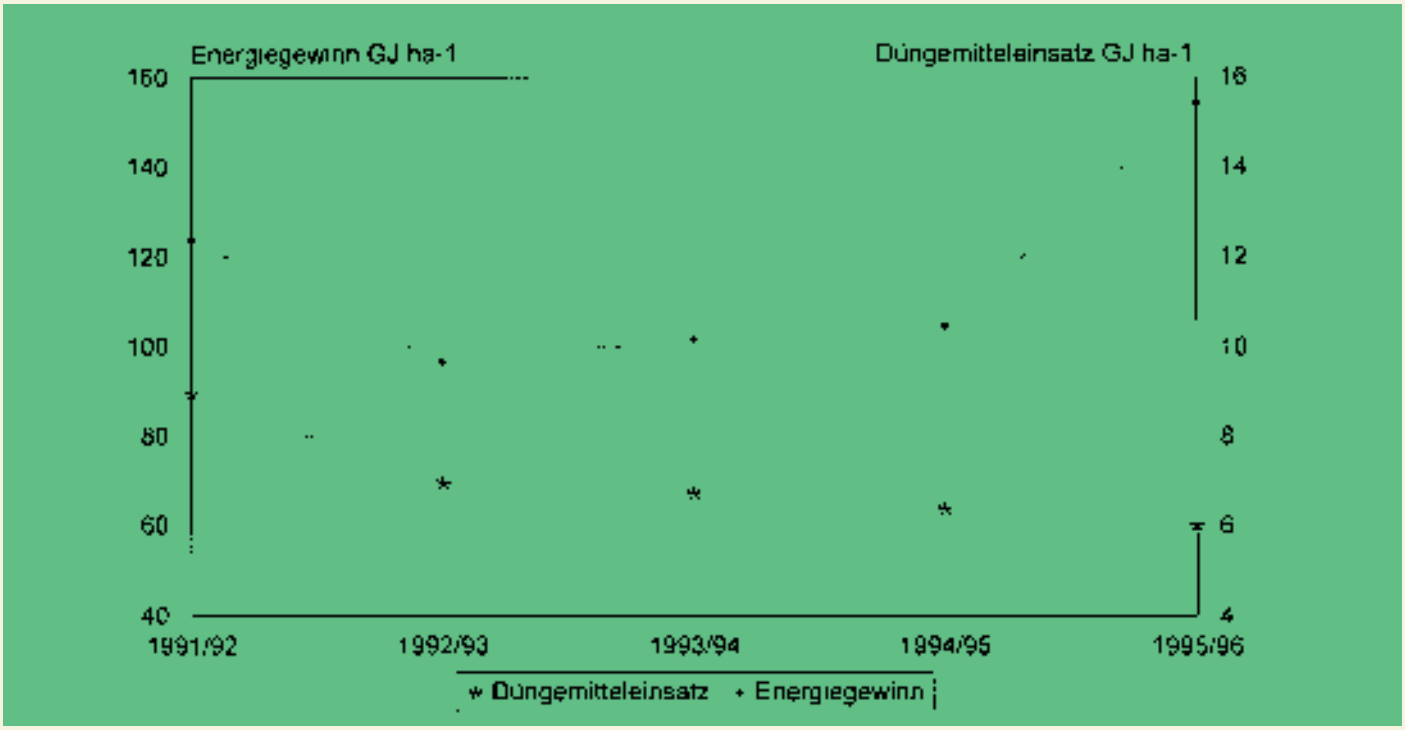


Abb. 1: Entwicklung des Düngemitteleinsatzes und des Energiegewinns I eines Betriebes von 1991/92 bis 1995/96 für den gesamten Ackerbau.

Fruchtfolgen oder die gesamte pflanzliche Produktion aufsummiert werden. Bisherigen Ergebnissen von Energiebilanzierungen verschiedener Autoren liegen unterschiedliche Zielsetzungen zu Grunde. Viele der veröffentlichten Arbeiten beziehen sich auf die pflanzliche Produktion oder erfassen den gesamten Betrieb. Energetische Bilanzierungen werden zum Vergleich verschiedener Landbewirtschaftungssysteme oder unterschiedlicher Anbau- und Bestelltechniken durchgeführt. Die Ergebnisse werden je nach Zielsetzung der Bilanzierung hinsichtlich des Energiegewinnes, des Output-Inputverhältnisses oder des fossilen Inputs analysiert und diskutiert.

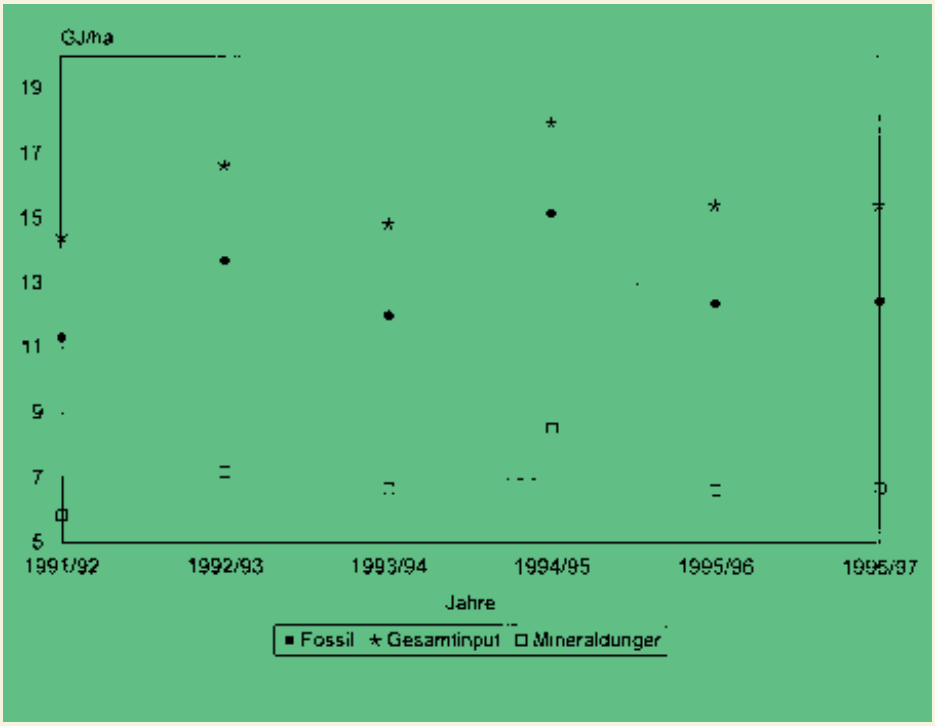


Abb. 2: Entwicklung des Inputs für Winterweizen in einem Ackerbaubetrieb am Beispiel des Gesamtinputs, des fossilen Inputs und des Mineraldüngereinsatzes.

Ein Beispiel – Düngemitteleinsatz und Energiegewinn

Anhand mehrerer Jahresergebnisse können mithilfe energetischer Bilanzen Entwicklungen innerhalb eines Betriebes dokumentiert werden. Die Abbildung 1 zeigt die Entwicklung eines Betriebes über den Zeitraum von fünf Bilanzjahren am Beispiel des Düngemitteleinsatzes und des Energiegewinns für den gesamten Ackerbau, wobei nur die Haupternteprodukte in die Betrachtungen einbezogen werden. Das heißt, die Ergebnisse beziehen sich auf den Energiegewinn I ohne Berücksichtigung von Nebenernteprodukten.

Nach dem ersten Bilanzjahr 1991/92 wurde der Einsatz an Düngemitteln sehr deutlich gesenkt. Dies zog einen Abfall des Energiegewinns I nach sich. In den folgenden Jahren von 1992/93 bis 1995/96 konnte der Energiegewinn I wieder gesteigert werden und übertraf 1995/96 den Ausgangswert von 1991/92 um rund 25 GJ/ha. Bei steigenden Energieerträgen war eine weitere relativ konstante Senkung des Einsatzes von Düngemitteln möglich und erreichte 1995/96 einen vorläufigen Tiefststand mit knapp 6 GJ/ha.

In der Ausgangssituation des Betriebes W1 zu Beginn des Projektes „Leitbetriebe Integrierter Landbau in NRW“ konnte die Nährstoffversorgung der Böden als mittel bis hoch beschrieben werden. Durch die regelmäßige Anwendung eines speziellen Analysenprogramms und die Einflussnahme der Beratung auf den Betriebsleiter wurde der Düngemitteleinsatz im Bilanzjahr 1992/93 stark gesenkt, was eine ebenfalls deutliche Senkung der Erträge bzw. Energiegewinne nach sich zog.

Mit der Beteiligung des Betriebes an einem Markenfleischprogramm seit 1993/94 wurde der Maisanteil an der Anbaufläche deutlich gesenkt und durch Getreide sowie Erbsen ersetzt. Gleichzeitig konnte der Anteil des Zwischenfruchtanbaues von 32,5 Prozent der Anbaufläche 1992 auf 53,8 Prozent im Jahr 1997 gesteigert werden. Gleiches gilt für die Steigerung des Anteils der Mulchsaat nach Zwischenfrüchten. Weiterhin wird auf dem Betrieb seit Beginn des Projektes „Leitbetriebe“ konsequent zu 100 Prozent pfluglos gewirtschaftet. Die Entwicklung des Betriebes hinsichtlich der Bearbeitungssysteme und des Einsatzes von Betriebsmitteln resultiert vorwiegend aus dem Wissensgewinn des Betriebsleiters. Dies schlägt sich letztlich auch in kontinuierlich steigenden durchschnittlichen Erträgen bzw. Energiegewinnen und gleichzeitig sinkendem Einsatz von Düngemitteln nieder.

Verschiedene Inputgrößen und ihre Aussagekraft

Jährliche Schwankungen lassen sich nicht nur für Erträge (Outputs), sondern auch für den Input abbilden (Abbildung 2). Die Punkte des Gesamtinputs und des fossilen Inputs liegen in nahezu gleich bleibendem Abstand zueinander. Der regenerative Input für Winterweizen ist in diesem Betrieb somit relativ konstant. Der geringe Abstand zwischen dem Gesamtinput und dem fossilen Input zeigt die Höhe des regenerativen Inputs (hier lediglich der Heizwert des Saatgutes) am Gesamtinput an. Der fossile Input gliedert sich in die fünf Inputgrößen Saatgut (Prozessenergie), Pflanzenschutzmittel, Maschinen, Kraftstoffe und Mineraldüngereinsatz. Da der überwiegende Anteil am fossilen Input in diesem Fall durch den Einsatz mineralischer Düngemittel bestritten wird, ist der Mineraldüngereinsatz zusätzlich abgebildet worden. Er beträgt im Verlauf der Bilanzjahre nahezu 50 Prozent am fossilen Input.

Haupteinfluss auf die jährlichen Inputschwankungen üben die Schwankungen im Mineraldüngereinsatz, der dem größten Anteil sowohl am fossilen Input als auch am Gesamtinput entspricht, aus. Die Höhe des Einsatzes der mineralischen Düngemittel wird im Wesentlichen vom Witterungsverlauf innerhalb eines Vegetationsjahres bestimmt. So ist das Jahr 1992/93 durch eine erschwerte Aussaat des Winterweizens bedingt durch anhaltende Niederschläge im Oktober/November, starke Auswinterungen durch Nachtkahlfröste bis minus 10°C im Februar und einer Hitze-/Trockenperiode im Frühsommer, die zu einer verfrühten ertragsmindernden Abreife bei Winterweizen führte, gekennzeichnet. Das Jahr 1994/95 wird durch gute Saatbedingungen im Herbst, eine eher zu milde Witterung mit anhaltenden Niederschlägen ausgangs des Winters (eher niedrige

Reststickstoffbestände im Boden) und einen „Jahrhundert-Hochsommer“ (die Ernte war Ende der ersten Augustwoche bereits abgeschlossen) charakterisiert.
In beiden Jahren wurde der Mineraldüngereinsatz auf dem Betrieb erhöht, um den sich abzeichnenden Ertragseinbußen entgegenzuwirken. Hinzu kamen Mehraufwendungen für Pflanzenschutzmittel, Maschinen und (mit dem Maschineneinsatz gekoppelt) Kraftstoffe, die jedoch nur einen geringen Einfluss auf die fossilen Inputschwankungen ausüben. Die durch die Witterung bedingten Ertragsrückgänge konnten durch den erhöhten Einsatz pflanzenbaulicher Maßnahmen nicht vollständig kompensiert werden. Eine detaillierte Darstellung der Inputgrößen in einer energetischen Bilanz kann somit schon relativ konkrete Aussagen über Bewirtschaftungsmaßnahmen zulassen, im Gegensatz zu einer Darstellung auf höherer Ebene, wie beispielsweise des Gesamtinputs oder des Energiegewinns.

Einfluss der Fruchtfolgegestaltung

In einem Betrieb werden drei unterschiedliche Fruchtfolgen praktiziert. Die tragenden Blattfrüchte sind Mais, Zuckerrüben und Raps. Für einen Vergleich der drei gesamten Fruchtfolgen wurden die Bilanzen jeder Fruchtfolge über den gesamten Bilanzzeitraum von fünf Jahren zu Mittelwerten in der Einheit GJ/ha zusammengefasst. Unterschiede der Fruchtfolgen innerhalb der einzelnen Bilanzglieder können durch den direkten Vergleich veranschaulicht werden (Tabelle 1).
Die Unterschiede im Input sind in ihrer Gesamtheit betrachtet zwischen den drei Fruchtfolgen eher gering. Relativ deutlich sind sie im Bereich des Düngemitteleinsatzes. Die Maisfruchtfolge mit einem Input an Düngemitteln (mineralisch und organisch) von 11,86 GJ/ha gegenüber 11,05 GJ/ha bei der Zuckerrübenfruchtfolge und lediglich 8,67 GJ/ha in der Rapsfruchtfolge erzielt allerdings auch den mit Abstand höchsten energetischen Output mit 522,00 GJ/ha. Leichte Unterschiede sind im Maschineneinsatz bzw. Kraftstoffverbrauch festzustellen. Den 6,26 GJ/ha in der Mais- bzw. 5,83 GJ/ha in der Zuckerrübenfruchtfolge stehen 4,36 GJ/ha der Rapsfruchtfolge gegenüber. Der durchschnittlich geringere Maschineneinsatz (und daran gekoppelt der geringere Kraftstoffverbrauch) sind in der Rapsfruchtfolge vor allem auf den geringeren Einsatz technischer Hilfsmittel beim Anbau des Feldgrases in dieser Fruchtfolge zurückzuführen. Da die Böden in ihrer Beschaffenheit auf dem Betrieb relativ homogen sind, ist auch in der Ausgestaltung des Maschineneinsatzes innerhalb der drei vorgestellten Fruchtfolgen kein gravierender Unterschied festzustellen.
Eventuelle Einsparungspotenziale liegen in erster Linie in den Bereichen der mineralischen Düngung und des Maschineneinsatzes. Untersuchungen haben gezeigt, dass Energie in Form von Treibstoff vor allem beim Pflügen erforderlich ist. Im praktischen Vergleich konnte der Treibstoffeinsatz durch Mulchsaat statt Pflug etwa halbiert werden. Sofern die Witterungsbedingungen und die Bodenverhältnisse es zulassen, könnte man auch im Betrieb W4 zumindest gelegentlich auf den Pflugeinsatz verzichten.
Der Einsatz mineralischer Stickstoffdünger birgt noch immer Einsparungsmöglichkeiten im konventionellen Landbau. Dies gilt insbesondere für den viehhaltenden Betrieb. Durch die regelmäßige Analyse von Bodenproben kann anhand der erhaltenen Daten sehr zielgerichtet gedüngt werden. Stehen neben den mineralischen Düngemitteln auch organische zur Verfügung, so können beispielsweise gezielte Güllegaben zur Zwischenfrucht, zur Strohdüngung oder zur Frühjahrsfurche den mineralischen Düngebedarf zusätzlich senken.
Die Unterschiede im Output hängen vor allem von dem Energiegehalt der geernteten Haupternteprodukte und den Erträgen je Flächeneinheit ab. So sind beispielsweise die hohen Energiegewinne der Maisfruchtfolge vor allem auf den Mais als Hauptern-

GJ/ha	Mais-fruchtfolge	Zuckerrüben-fruchtfolge	Raps-fruchtfolge
INPUT			
Input-Fossil			
Saatgut (Prozess)	0,61	0,53	0,44
Mineraldünger	6,74	6,73	5,16
PSM	0,60	0,89	0,51
Maschinen	2,78	2,71	1,97
Kraftstoffe	3,48	3,12	2,89
∑ Input-Fossil	14,23	13,62	10,47
Input-Regenerativ			
Saatgut (Heizwert)	2,10	1,90	1,58
Organ. Dünger (N, P, K)	5,12	4,32	3,51
∑ Input-Regenerativ	7,23	6,22	5,09
∑ INPUT	21,45	19,84	15,56
Output I	522,00	272,44	231,62
Energiegewinn I	500,55	252,60	216,06
Output: Input I	24,34:1	13,73:1	14,89:1

Tab. 1: Vergleich von drei Fruchtfolgen eines Betriebes in der Soester Börde im Mittel von fünf Bilanzjahren (GJ/ha).

	91/92	92/93	93/94	94/95	95/96	∅
Input						
Mineraldünger	4,75	5,48	6,25	6,43	5,25	5,63
Summe Dünger	6,90	8,05	8,35	8,31	7,07	7,74
Maschinen	2,50	2,29	2,22	2,09	2,14	2,25
Input-Fossil	10,38	10,86	11,82	11,54	10,42	11,00
Input-Regenerativ	5,56	6,03	5,73	5,52	5,09	5,59
∑ Input	15,94	16,89	17,35	17,06	15,51	16,55
Output I	115,14	114,25	120,01	130,23	133,91	122,71
Energiegewinn I	99,21	97,34	102,46	113,17	118,40	106,12
Output:Input I	7,32:1	6,84:1	6,89:1	7,63:1	8,58:1	7,45:1

Tab. 2: Entwicklung der Energiebilanzen für Winterweizen in Betrieben der Gruppe 1 mit geringerer Bodengüte (GJ/ha).

	91/92	92/93	93/94	94/95	95/96	∅
Input						
Mineraldünger	7,12	6,59	7,18	7,57	6,31	6,95
Summe Dünger	7,44	7,21	7,67	8,21	6,87	7,48
Maschinen	2,36	2,31	2,14	2,31	2,12	2,25
Input-Fossil	12,99	12,15	12,38	13,23	11,73	12,50
Input-Regenerativ	3,48	3,62	3,48	3,66	3,99	3,64
∑ Input	16,47	15,77	15,86	16,89	15,72	16,14
Output I	138,48	147,85	138,96	135,66	161,52	144,49
Energiegewinn I	122,02	132,08	123,10	126,98	148,88	130,61
Output:Input I	8,43:1	9,48:1	8,86:1	8,52:1	10,43:1	9,14:1

Tab. 3: Entwicklung der Energiebilanzen für Winterweizen in Betrieben der Gruppe 2 auf besseren Standorten (GJ/ha).

teprodukt mit sehr hohen Energiewerten je dt Erntegut und mit hohen Flächenerträgen begründet. Von den gewählten Kulturarten innerhalb einer Fruchtfolge hängt somit ganz entscheidend die Höhe des Energiegewinns der Fruchtfolge ab. Die Höhe des Energieinputs beeinflusst dagegen vor allem die Relation von Output zu Input, sodass einem guten Energiegewinn (Zuckerrübenfruchtfolge) ein relativ schlechtes Output-Input-Verhältnis gegenüber stehen kann. Es erscheint sinnvoll, sowohl den Energiegewinn als auch die Relation von Output zu Input im Sinne einer nachhaltigen und umweltverträglichen Landwirtschaft in Effizienzbetrachtungen einzubeziehen, damit ein überproportionaler Input zur Erzielung eines hohen Outputs vermieden wird.

Einfluss der Bodengüte auf Energiebilanzen am Beispiel des Winterweizens

Für diese Untersuchung wurden neun Betriebe in zwei Gruppen unterteilt. Die Gruppe 1 beinhaltet alle Betriebe mit durchschnittlich geringerer Bodengüte und schwierigeren Bearbeitungsbedingungen. Dieser Gruppe stehen die von der Natur etwas begünstigten Betriebe als Gruppe 2 gegenüber. Die Tabellen 2 und 3 stellen die zeitliche Entwicklung der beiden Gruppen in den für die Energiebilanzen wichtigsten Bilanzgrößen dar. Der Vergleich der beiden Betriebsgruppen mit unterschiedlichen Bodengüten ergibt einen nahezu gleichen Input (im Durchschnitt der Bilanzjahre) bei einem deutlich höheren Output der Gruppe 2, das heißt der Gruppe, die durchschnittlich über die bessere Bodenqualität verfügt. Da der Bilanzzeitraum für alle Betriebe fünf Bilanzjahre beinhaltet, können sowohl größere Wirkungen der Fruchtfolgen als auch Effekte der Witterungsverhältnisse weitestgehend ausgeschlossen werden. Wenn das zu Beginn der Datenerhebungen unterschiedliche Know-how der Betriebsleiter vernachlässigt wird, können die Unterschiede der beiden Gruppen im Output und daraus resultierend im Energiegewinn sowie im Verhältnis Output zu Input zu einem überwiegenden Teil auf die unterschiedlichen Bodengüten zurückgeführt werden. Bei einem Vergleich von landwirtschaftlichen Betrieben in der pflanzlichen Produktion muss neben anderen Aspekten somit auch die eventuell unterschiedliche Beschaffenheit der Bodenqualität Eingang in die Schlussfolgerungen der vergleichenden Betrachtungen finden, um Fehleinschätzungen der Wirtschaftsweisen auf den Betrieben zu vermeiden.

Gegenwart und Zukunft

Aufstellung und Interpretation energetischer Bilanzen landwirtschaftlicher Produktionssysteme erscheinen nur über einen längeren Zeitraum betrachtet sinnvoll. Eine einmalig aufgestellte Bilanz besitzt nahezu keine Aussagekraft. Je detaillierter die

Bilanz in sinnvoller Weise aufgeschlüsselt ist, desto eher können Schwachstellen und Optima in Bewirtschaftungssystemen erkannt werden. Bei der Beurteilung landwirtschaftlicher Energiebilanzen sind immer die Einflussfaktoren Witterung, Klima, Bodenqualität und geografische Lage in die Schlussfolgerungen einzubeziehen, da sie in nicht unerheblichem Maß die Ergebnisse der Bilanzen beeinflussen können.
Generell sind anhand des vorhandenen Datenmaterials auf den landwirtschaftlichen Betrieben Bilanzen für den gesamten Betrieb am schnellsten und einfachsten zu erstellen. Mithilfe der entwickelten Eingabemasken unter dem Tabellenkalkulationsprogramm EXCEL lassen sich Hoftorbilanzen eines Betriebes anhand von Buchführungsunterlagen für mindestens zwei Jahre innerhalb eines Tages aufstellen. Sind die notwendigen Unterlagen für die Berechnung der Schlag- und Flächenbilanz vorhanden, so ist die Aufstellung der Energiebilanz für alle Schläge mit Zusammenfassung der Schläge für die jeweilige Kulturart für ein Bilanzjahr innerhalb eines Tages möglich.
Forschungsbedarf besteht vor allem im Bereich der tierischen Erzeugung. Durch die unterschiedlichsten Produktionssysteme in den tierhaltenden Betrieben dieses Projektes sind Vergleiche nur schwer möglich, sodass Anhaltspunkte für eine Bewertung der aufgestellten Stallbilanzen fehlen. Weiterführende Arbeiten mit der Zielsetzung, Betriebe mit gleichen Produktionssystemen unter gleichen Produktionsbedingungen zu vergleichen, könnten einen wichtigen Beitrag zur Erarbeitung von Anhaltspunkten und Bewertungsgrundlagen in energetischen Stallbilanzen der landwirtschaftlichen Tierhaltung liefern.

• Dank gebührt der Deutschen Bundesstiftung Umwelt für die Förderung dieses Projektes.

Literatur

DIEPENBROCK,W.; PELZER, B.; RADTKE, J. (1995): Energiebilanz im Ackerbaubetrieb, KTBL-Arbeitspapier 211.

HAAS, G. (1996): Maßzahlen der Energieeffizienz: Brennwerte oder Lebensmittel erzeugen? Mitt. Ges. f. Pflanzenbauwiss. 9, S. 101-102.

KALK, W.-D.; HÜLSBERGEN, K.-J. (1996): Methodik zur Einbeziehung des indirekten Energieverbrauchs mit Investitionsgütern in Energiebilanzen von Landwirtschaftsbetrieben. Kühn.-Arch. 90 (1996) 1, S. 41-56, Landwirtschaftsverlag, Münster-Hiltrup.

KALTSCHMITT, M.; REINHARDT, G. A. (1997): Nachwachsende Energieträger. Grundlagen, Verfahren, ökologische Bilanzierung. Vieweg-Verlag, 1997.

LÜTKE ENTRUP, N.; ONNEN, O.; TEICHGRÄBER, B. (1998): Integrierter Landbau in Deutschland und Europa – Studie zur Entwicklung und den Perspektiven. In: Fördergemeinschaft Integrierter Pflanzenbau (Hrsg.): Zukunftsfähige Landwirtschaft, Heft 14/1998.

MOERSCHNER, J.; ECKERT, H.; HOFFMANN, M.; ISERMANN, K.; KÜSTERS, J. (2000): 5. Entwurf VDLUFA – Standpunkt „Energiebilanzen im Landwirtschaftsbetrieb".

MOERSCHNER, J.; GEROWITT, B. (1998): Energiebilanzen von Raps bei unterschiedlichen Anbauintensitäten. In: Landtechnik 6/98, S. 384-385.

ZERHUSEN-BLECHER, P. (1996): Bericht über das Forschungsvorhaben „Leitbetriebe Integrierter Landbau in Nordrhein-Westfalen", 1. Zwischenbericht der 2. Projektphase, Versuchsjahr 1994/1995.

Dipl.-Ing. agr. Britta Hackenschmidt betreut das Projekt „Energie- und Kohlendioxidbilanzen landwirtschaftlicher Betriebe – Energetischer Bewertungsansatz landwirtschaftlicher Produktionssysteme am Beispiel der Leitbetriebe Integrierter Landbau in NRW".