

Hans-Joachim Fritz und Robert Schaback (Hg.)  
Die Natur der Information

erschienen im Universitätsverlag Göttingen 2019

---

Hans-Joachim Fritz  
Robert Schaback (Hg.)

# Die Natur der Information

Beiträge aus der Arbeit  
der gleichnamigen  
Forschungskommission  
der Akademie der Wissenschaften  
zu Göttingen



Universitätsverlag Göttingen  
2019

## Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

### *Anschrift der Herausgeber*

Akademie der Wissenschaften zu Göttingen  
Theaterstraße 7  
37073 Göttingen  
Tel.: +49 (0)551 39-5362  
Fax.: +49 (0)551 39-5365  
E-Mail: [adw@gwdg.de](mailto:adw@gwdg.de)  
[www.adw-goe.de](http://www.adw-goe.de)

Dieses Buch ist auch als freie Onlineversion über die Homepage des Verlags, über den Dokumentenserver der Akademie der Wissenschaften zu Göttingen [res.doctae \(https://rep.adw-goe.de/\)](https://rep.adw-goe.de/) sowie über den Göttinger Universitätskatalog (GUK) bei der Niedersächsischen Staats- und Universitätsbibliothek Göttingen (<https://www.sub.uni-goettingen.de>) erreichbar. Es gelten die Lizenzbestimmungen der Onlineversion.

Satz und Layout: Robert Schaback  
Umschlaggestaltung: Jutta Pabst  
Titelabbildung: Paul Klee (1879-1940), „Grenzen des Verstandes“  
Bayerische Staatsgemäldesammlungen – Sammlung Moderne Kunst in der Pinakothek der Moderne München (Inv.-Nr. 14234)

© 2019 Universitätsverlag Göttingen  
<https://univerlag.uni-goettingen.de>  
ISBN: 978-3-86395-409-3  
DOI: <https://doi.org/10.17875/gup.2019-1157>

## Inhaltsverzeichnis

Präambel.....	3
<i>Robert Schaback, im Auftrag der Kommission „Die Natur der Information“</i>	
Definitionen von Information .....	11
<i>Robert Schaback, Fakultät für Mathematik und Informatik, Universität Göttingen</i>	
Wahrscheinlichkeit und Information – wie Henne und Ei?.....	27
<i>Kurt Schönhammer, Fakultät für Physik, Universität Göttingen</i>	
Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information .....	45
<i>Kurt Schönhammer, Fakultät für Physik, Universität Göttingen</i>	
Quanteninformation: Versuch einer Einführung für Nicht-Physiker .....	73
<i>Kurt Schönhammer, Fakultät für Physik, Universität Göttingen</i>	
Informationsverarbeitung in der Kognitionspsychologie und in den kognitiven Neurowissenschaften .....	89
<i>Gerd Lüer, Fakultät für Biologie, Universität Göttingen</i>	
<i>Uta Lass, Courant Forschungszentrum „Textstrukturen“, Universität Göttingen</i>	
Informationsverarbeitung in der Hörbahn von Insekten .....	119
<i>Bernhard Ronacher, Institut für Biologie, Humboldt-Universität zu Berlin</i>	

Informationstheoretische Metaphern im Vokabular der Molekularen Genetik .....	143
<i>Hans-Joachim Fritz, Institut für Mikrobiologie und Genetik der Universität Göttingen</i>	
Informationsgesteuerte Synthese – eine Blockbuster-„Erfindung“ der Evolution .....	167
<i>Hans-Joachim Fritz, Institut für Mikrobiologie und Genetik der Universität Göttingen</i>	
Daten, Prozesse und Information .....	183
<i>Robert Schaback, Fakultät für Mathematik und Informatik, Universität Göttingen</i>	
Vortragsreihe der Akademiekommission .....	195

## Präambel

*R. Schaback, im Namen der Akademiekommision „Die Natur der Information“  
Institut für Numerische und Angewandte Mathematik, Lotzestraße 16-18, 37083  
Göttingen*

In einem Brief vom 17.10.2005 stellte Prof. Dr. Hans-Joachim Fritz an die Professoren Norbert Elsner, Werner Lehfeldt, Gerd Lüer, Robert Schaback und Kurt Schönhammer die Frage, „*ob es fruchtbar sein könnte, in ein Gespräch einzutreten, das den Informationsbegriff von verschiedenen Seiten beleuchtet*“, und um den Versuch zu machen, „*Sie als Vertreter von Physik, Informatik, Psychologie, Neurobiologie und Sprachwissenschaften zu einem ersten Gedankenaustausch über das skizzierte Thema zusammenzubringen*“. So begann die Arbeit, deren Abschlussbericht hier vorgelegt wird.

Die Gruppe verständigte sich auf ihrem ersten Treffen am 11.11.2005 darauf, unter Federführung von Prof. Dr. Hans-Joachim Fritz „*die Diskussion zunächst informell im Sinne einer gegenseitigen Nachhilfe zu führen*“, mit den Vorträgen

- Zum Begriff der Information in Mathematik und Informatik  
(Schaback, 16.12.2005 und 08.12.2006)
- Der Entropiebegriff in der Thermodynamik und der Statistischen Mechanik  
(Schönhammer, 06.01.2006 und 23.06.2006)
- Anmerkungen zum Informationsbegriff in der Linguistik  
(Lehfeldt, 20.01.2006)
- Informationsgesteuerte Synthese-  
eine Blockbuster-Erfindung der Evolution  
(Fritz, 10.03.2006)
- Datenträger und Informationsfluss in der Molekularen Biologie  
(Fritz, 30.03.2006)
- Kognitionspsychologie - eine Psychologie der Informationsverarbeitung  
(Lüer, 28.04.2006)
- Bau und Funktionsweise natürlicher Neuronen  
(Elsner, 15.12.2006)

Schon am Ende des ersten Semesters kam die Gruppe überein, „*in eine explorative Phase*“ übergehen zu wollen, nämlich durch Erweiterung des in Göttingen vorhandenen Fachwissens durch Einladung externer Fachleute zu Vorträgen. Es zeichnete

sich früh ab, dass der *semantische Informationsbegriff*, der Information als semantisches Ergebnis der Interpretation einer im Allgemeinen sprachlich fixierten Nachricht sieht (vgl. Aufsatz ab S. 11), von zentraler Bedeutung ist. Dabei ist *Sprache* im weitesten Sinn zu verstehen: von der DNA der Molekularbiologie über tierische und menschliche Kommunikation bis hin zu den Programmiersprachen der Informatik. Weil nach wichtigen Arbeiten von Manfred Eigen<sup>1 2</sup> die biologische Evolution auch als Entwicklung der genetischen Sprache und ihres Informationsgehalts aufgefasst werden kann, wurde als eine erste thematische Fokussierung am 10.06.2006 „*Ko-Evolution von Sender und Empfänger*“ vorgeschlagen, und auf die „*Suche nach Gesetzmäßigkeiten für die Entwicklung von Sprache auf verschiedensten Ebenen*“ erweitert.

Am 26. Januar 2007 wurde aus der informellen Arbeitsgruppe eine Forschungskommission der Akademie unter Vorsitz von H.-J. Fritz, und im weiteren Verlauf kamen

- Klaus-Peter Lieb (Physik)
- Gert Webelhuth (Linguistik)

als Mitglieder und

- Konrad Bachmann (IPK Gatersleben)
- Konrad Cramer (Philosophie)
- Jens Frahm (Biomedizinische NMR Forschungs GmbH am MPI)

als regelmäßige und

- Wolfgang Künne (Philosophie, Hamburg)
- Bernd Ronacher (Neurophysiologie, Humboldt-Universität, Berlin) und
- Manfred R. Schroeder (Physik)

als sporadische Gäste hinzu.

Das Thema *Information* wurde auch in reguläre Sitzungen der Akademie eingebracht:

- Prof. Dr. M. Eigen (Göttingen): Was ist Information? (28.06.2007)
- Prof. Dr. Luciano Floridi (Oxford): The future development of the information society (09.11.2007),

---

<sup>1</sup>Selforganization of matter and the evolution of biological macromolecules, Die Naturwissenschaften 58, 1971

<sup>2</sup>Der Ursprung der genetischen Information, in: *Evolution – Zufall und Notwendigkeit der Schöpfung*, Elsner, N., Fritz, H.-J., Gradstein, R. und Reitner, J. (Hrsg.), Wallstein-Verlag Göttingen 2009



und insbesondere war Luciano Floridi im Jahre 2007 auf Initiative der Kommission als Gauß-Professor in Göttingen.

Die Erweiterung der Kommission in Richtung auf Philosophie, Linguistik, Psychologie und Kognitionswissenschaften ergab sich aus der Fokussierung auf den semantischen Informationsbegriff. *Nachrichten* sind als Strukturen auf physikalischen Medien sicher dem materiellen Bereich zuzuordnen, die Gewinnung von Information durch Interpretation der Nachrichten verweist aber auf die Ebene des *Sinns* oder des *Geistes*. Der semantische Informationsbegriff berührt deshalb in der Philosophie das Leib/Seele- bzw. Materie/Geist-Problem, und betrifft in der Linguistik die Semantik von Sprachen. Das *Formalisieren* von Sprachen, in der Informatik gang und gäbe, bleibt dort zunächst auf der syntaktischen Ebene, wird aber durch *formalisierte Semantik* auf die semantische ausgeweitet. In der Linguistik entspricht dem der Übergang von der üblichen Grammatik zu formalen Semantiksystemen, z. B. der *Diskursrepräsentationstheorie*. Die Verarbeitung von Information im menschlichen Denken und Entscheiden und deren Modellierung ist dann Gegenstand der Kognitionspsychologie und der kognitiven Neurowissenschaften.

In den Folgesemestern wurde ein vielfältiges Vortragsprogramm durchgeführt (im Anhang ab Seite 195), das die Rolle von „*Information*“ in verschiedenen Wissenschaftsbereichen darstellte:

- *Physik*
- *Neuronen und Neuronale Netze*
- *Biologie und Evolution*
- *Linguistik*
- *Philosophie*
- *Psychologie, Kommunikation und Kognition*
- *Mathematik und Informatik*
- *Recht und Gesellschaft*

Am 9. Oktober 2009 ging der Vorsitz der Kommission von H.-J. Fritz an R. Schaback über. Das Programm wurde unverändert beibehalten.

Die Kommission bestritt die öffentliche Akademiesitzung in der Paulinerkirche am 21.1.2011 mit den Vorträgen

- Zum Begriff der Information in Mathematik und Informatik (Robert Schaback),
- Nachricht und Information in der belebten Welt (Hans-Joachim Fritz),
- Die Verarbeitung von visueller Bewegungsinformation in der Großhirnrinde (Stefan Treue, Primatenzentrum).

Im Sommersemester 2011 wurde die Beschäftigung mit der Evolution biologischer Makromoleküle in die unter Vorsitz von Hans-Joachim Fritz neu gegründete Forschungskommission *Origin of Life* ausgegliedert und dort mit Untersuchungen zu früher Erdgeschichte und präbiotischer Chemie zusammengeführt. Unter letzterer ist geo-organische Chemie im Zeitabschnitt vor dem Auftreten erster potentiell nachrichtentragender Makromoleküle zu verstehen, der mit „genetischer Information“ *prima vista* nichts zu tun hat, weil die Moleküle noch zu klein sind. Ob komplexe chemische Systeme, die sich fernab des Gleichgewichts in einem quasi-stationären Zustand befinden und sehr langsam in Richtung auf Leben hin evolvieren, in irgendeinem Sinne Informationen enthalten oder verarbeiten, ist zu klären, wenn solche Systeme identifiziert sind.

Durch den Tod des Akademie-Vizepräsidenten und Gründungsmitglieds Norbert Elsner nach kurzer schwerer Krankheit verlor die Kommission am 16. 06. 2011 einen ihrer wichtigsten Initiatoren. Auch der Tod von Konrad Cramer am 12. 02. 2013 hinterließ eine nicht auffüllbare Lücke.

Im Rückblick auf die vergangenen Jahre stellte die Kommission im WS 2013/14 fest, dass das bisherige Vortragsprogramm einen guten Überblick über die „*Natur der Information*“ geliefert und damit seinen Zweck erfüllt hat. Sie beschloss deshalb, das Vortragsprogramm einzustellen und vorrangig ein publizierbares Manuskript zu erarbeiten, das die bisherige Kommissionsarbeit zusammenfasst und auswertet. Ferner war im WS 2014/15 die Amtszeit des bisherigen Vorsitzenden R. Schaback abgelaufen und wurde auf eigenen Wunsch nicht verlängert. Die Kommission setzte mit Herrn Fritz als kommissarischem Koordinator ihre Arbeit fort bis zur Anfertigung dieses Berichtes. Er besteht aus folgenden Einzelbeiträgen:

- *Definitionen von Information* von Robert Schaback (S. 11) stellt
  1. den *mathematischen Informationsbegriff* nach Shannon dar, der den Informationsgehalt eines zufälligen Ereignisses als den negativen Zweierlogarithmus von dessen Wahrscheinlichkeit definiert,
  2. den *bayesianischen* als *change of beliefs*, der bei Formalisierung als Übergang eines Systems von Wahrscheinlichkeiten zu bedingten Wahrscheinlichkeiten beschrieben werden kann,
  3. nach Solomonoff/Kolmogoroff/Chaitin den *algorithmischen Informationsgehalt* einer Struktur als die Komplexität des einfachsten Verfahrens, das die Struktur erzeugt und
  4. den *semantischen*, der Information als semantischen Gehalt von strukturierten Daten definiert,

und grenzt diese Begriffe gegeneinander ab.

- *Wahrscheinlichkeit und Information - wie Henne und Ei?* von Kurt Schönhammer (S. 27) ergänzt die im ersten Aufsatz dargestellte Methode, aus der Wahrscheinlichkeit  $p$  eines Ereignisses die Information  $-\log_2(p)$  abzuleiten, durch ihren Gegenpol, nämlich die bayesianische Sichtweise, bei der

sich durch Hinzutreten von Informationen Wahrscheinlichkeiten ändern. Der schon im ersten Artikel kurz erwähnte *bayesianische Informationsbegriff* als *change of rational beliefs* wird detailliert dargestellt und durch verschiedene Beispiele illustriert, mit Anwendungen in medizinischen Tests und in der Wissenschaftstheorie.

- *Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information* von Kurt Schönhammer (S. 45) beginnt mit der Entropie als Zustandsfunktion der Thermodynamik und schildert ihre Rolle beim zweiten Hauptsatz in der Form

*Die Entropie eines abgeschlossenen Systems kann niemals abnehmen. Sie nimmt zu bei allen natürlichen, mit endlicher Geschwindigkeit verlaufenden Prozessen.*

Die Statistische Mechanik liefert die mikroskopische Begründung der Thermodynamik. Dabei spielt der Entropiebegriff eine zentrale Rolle. Der von Boltzmann vorgeschlagene Ausdruck für die Entropie wurde von Gibbs und von Neumann verallgemeinert. Auf Vorschlag von Letzterem gab Shannon seinem Ausdruck für den Informationsgehalt einer Nachricht den Namen *Informationsentropie*, da er dieselbe Form wie die Entropie in der Statistischen Mechanik hat.

- *Quanteninformation: Versuch einer Einführung für Nicht-Physiker* von Kurt Schönhammer (S. 73) zeigt, dass sich zum Problemkreis *Kommunikation und Information* neue Gesichtspunkte ergeben, wenn die Quantennatur des Systems der Nachrichtenübertragung eine Rolle spielt. Weil man Quantenzustände nicht „klonieren“ kann und das „Abhören“ eines quantenmechanischen Zustandes diesen verändert, wenn der Abhörende nicht zufällig das Abzuhörende bereits kennt, kann das Abhören unbekannter Nachrichten auf einem quantenmechanischen Kommunikationskanal nicht unbemerkt geschehen. Und bei quantenmechanischer *Verschränkung* zweier Teilsysteme beeinflusst die Messung in einem Teilsystem den Zustand im anderen Teilsystem, auch bei räumlicher Trennung.
- *Informationsverarbeitung in der Kognitionspsychologie und in den kognitiven Neurowissenschaften* von Gerd Lüer und Uta Lass (S. 89) schildert den Paradigmenwechsel vom Behaviorismus zur Psychologie der Informationsverarbeitung. Weil die Beschränkung des Behaviorismus auf Reiz-Reaktions-Modelle wesentliche kognitive Prozesse ignorierte, die zwischen Reiz und Reaktion liegen, und weil die quantitative Formalisierung des Informationsgehalts von Nachrichten nach Shannon nicht mit Kapazitätsberechnungen des Gedächtnisses übereinstimmte, rückt die heutige Psychologie der Informationsverarbeitung kognitive Prozesse in den Fokus ihrer Forschung. Dazu werden zwei wichtige Arbeitsrichtungen vorgestellt: die *Modellierung kognitiver Architekturen* und die *Entwicklung konnektionistischer Netzwerkmodelle*. In beiden Fällen handelt es sich um Realisierungen von Theorien, die auf

Computern als lauffähige Programme implementiert werden können, um die zugrundeliegenden Theorien über kognitive Prozesse und Strukturen und deren Eigenschaften sowie auch deren dynamischen Veränderungen detailliert kennenzulernen und auch zu testen. Die kognitiven Neurowissenschaften prüfen diese Modelle an experimentellen Daten, und dazu werden Beispiele angegeben.

- *Informationsverarbeitung in der Hörbahn von Insekten* von Bernhard Ronacher (S. 119) behandelt am Beispiel der akustischen Kommunikation von Heuschrecken einige grundlegende Probleme für die Verarbeitung von Umweltreizen im Nervensystem. Warum wird trotz der Uneindeutigkeit der Sinnesreize ein bestimmter Gegenstand immer als derselbe erkannt? Wie gravierend sind die Auswirkungen der intrinsischen Variabilität neuronaler Signale auf die Verarbeitung und Repräsentation sensorischer Information?
- *Informationstheoretische Metaphern im Vokabular der Molekularen Genetik* von Hans-Joachim Fritz (S. 143) schildert, wie weit der Informationsbegriff in die Wissenschaftssprache der Molekulargenetik eingedrungen ist, wenn vom „metabolischen Netzwerk“, „Informationsgehalt“, „Codierung“, „Auslesen“, „Codon“, „Message“ die Rede ist. Das verleitet aber auch, überall computerähnlich funktionierende Mechanismen mit totalem Determinismus und einer zentralen Steuerinstanz zu sehen, wo eher selbstorganisierende netzwerkartige Prozesse und keine strikt linearen Kausalketten vorliegen.
- *Informationsgesteuerte Synthese - eine Blockbuster-Erfindung der Evolution* von Hans-Joachim Fritz (S. 167) untermauert die These, dass die Entstehung von Leben auf unserem Planeten, so wie wir es kennen, mit einem in jeder Zelle ablaufenden, komplexen Stoffwechsel, die Koevolution von informationsgesteuerter Synthese katalytisch wirksamer Makromoleküle zwingend voraussetzt.
- *Daten, Prozesse und Information* von Robert Schaback (S. 183) versucht, die Arbeit der Akademiekommission begrifflich zusammenzufassen, indem Information als Zustandsänderung eines Prozesses definiert wird, die durch Daten bewirkt ist. Wie sich dieser Informationsbegriff in verschiedenen Disziplinen anwenden lässt und wie er die anderen Begriffe von Information umfasst, zeigt sich in den vorgenannten Beiträgen und wird im genannten Artikel detailliert erläutert.

Die Mitglieder der Kommission haben durch Vorträge und Diskussionen eine Menge über die Natur der Information gelernt, mindestens ebenso viel aber über die Natur interdisziplinären Arbeitens. Es ist nicht damit getan, dass mehrere Personen ein Problem aus verschiedenen Blickwinkeln betrachten, sich ihre jeweilige Ansicht gegenseitig vortragen und ansonsten geduldig zuhören. Entgegen landläufiger Meinung wird das interdisziplinäre Arbeiten erst dann wirklich erfolgreich, wenn

in möglichst vielen einzelnen Köpfen einer Gruppe gedanklich mit den verschiedenen Aspekten der verschiedenen Disziplinen experimentiert wird und man sich das Ergebnis solcher Experimente wieder gegenseitig zur kritischen Prüfung vorlegt. Dazu braucht es den Mut, sich aus dem Fenster der eigenen Spezialkompetenz zu lehnen und die Bereitschaft aller, sich der Kritik aus den anderen Disziplinen zu stellen. Funktionieren kann so etwas nur in der vertrauensvollen Atmosphäre einer kleinen Gruppe von Spezialisten, die sich über Fachgrenzen hinweg kollegial bis freundschaftlich verbunden sind.

Göttingen, den 26. Juni 2019

R. Schaback

unter Mitwirkung der Mitglieder der Akademiekommision „Die Natur der Information“.

Die Herausgeber danken Frau Jutta Pabst vom Universitätsverlag Göttingen für die engagierte und kundige Betreuung dieses Publikationsprojekts und der Pinakothek der Moderne, München, für die großzügige Erlaubnis der Verwendung des Umschlagsbilds.



# Definitionen von *Information*

*R. Schaback*

*Institut für Numerische und Angewandte Mathematik,  
Lotzestraße 16-18, 37083 Göttingen*

## Zusammenfassung

In den Wissenschaften finden sich vier Definitionen von *Information*, die gegeneinander einigermaßen klar abgrenzbar sind:

1. Der *mathematische* Informationsbegriff nach *Shannon*. Er basiert auf den *Kolmogoroff*'schen Axiomen der Wahrscheinlichkeitstheorie, definiert Information als negativen Logarithmus einer Wahrscheinlichkeit, führt zur *Entropie* und ist deshalb auch in der Physik von zentraler Bedeutung.
2. Verwendet man stattdessen einen bayesianistischen Zugang zu Wahrscheinlichkeiten als *degree of belief*, so ist *Information* das, was einen *change of beliefs* bewirkt. Dieser Informationsbegriff sei hier *bayesianistisch* genannt.
3. Die *algorithmische* Information, die in einem Text steckt, ist nach *Solomonoff/Kolmogoroff/Chaitin* die minimale Länge eines Programms, das diesen Text produziert. Dieser Begriff ist an ein Maschinenmodell gebunden und liefert ein Maß für *Komplexität*.
4. Der *semantische* Informationsbegriff definiert Information als den *semantischen Gehalt von strukturierten Daten*.

Das Ziel dieses Beitrags ist, diese Informationsbegriffe etwas genauer darzustellen, gegeneinander abzugrenzen und eventuelle Gemeinsamkeiten aufzuzeigen. Eine besondere Tiefe wird dabei nicht angestrebt. Sie bleibt anderen Beiträgen in diesem Band vorbehalten. Und der Initiator der Informationstheorie, C. E. Shannon, stellte schon 1953 fest:

*It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field.*  
[19, p. 180]

# 1 Der mathematische Informationsbegriff

Der Informationsbegriff in der Mathematik geht auf die von *Claude Shannon* begründete *Informationstheorie* zurück. Man kann die These vertreten, letztere sei eher eine Theorie der *Kommunikation*, und dies wird sogar durch die Titel von Shannons Originalarbeiten [16, 17, 18] gestützt. Hier kann kein kompletter Überblick über die Informationstheorie [11] gegeben werden, insbesondere nicht mit allen mathematischen Details. Stattdessen wird auf diejenigen Gesichtspunkte fokussiert, die für die Querverbindungen zu anderen Informationsbegriffen wichtig sind.

## 1.1 Ereignisse

Man betrachtet eine Menge  $M$  im Sinne der mathematischen Mengenlehre, deren Elemente man Ereignisse nennt und die in der Regel Ereignisse aus der „realen“ Welt modellieren. Ein typischer Fall ist die Menge  $M := \{1, 2, 3, 4, 5, 6\}$  der möglichen Ergebnisereignisse des herkömmlichen Würfels. Das *Eintreten eines Ereignisses* ist dann mathematisch nichts anderes als die Auswahl eines Elementes der Ereignismenge. Im Unterschied zu unstrukturierten Mengen haben Mengen von Ereignissen also die Zusatzeigenschaft, dass die Art der Auswahl von Elementen ebenfalls modelliert wird. Dies geschieht in der Regel dadurch, dass man den Ereignissen  $E$  Wahrscheinlichkeiten  $p(E)$  zwischen Null und Eins zuordnet, die den Auswahlprozess sauber modellieren, auch wenn er „zufällig“ ist.

Man definiert dann die *Information* eines Ereignisses  $E$  als

$$I(E) := -\log_2 p(E),$$

d. h. die Information eines Ereignisses ist der negative Zweierlogarithmus der Wahrscheinlichkeit dieses Ereignisses.

Dieser Informationsbegriff hat dann die folgenden Eigenschaften:

1. Ein Ereignis, das mit Wahrscheinlichkeit Eins eintritt, hat Information Null.
2. Ein Ereignis, das mit Wahrscheinlichkeit Null eintritt, hat positiv unendliche Information.
3. Wenn zwei Ereignisse  $E_1$  und  $E_2$  unabhängig voneinander eintreten können, hat nach den Regeln der Wahrscheinlichkeitstheorie das Und-Ereignis ( $E_1$  und  $E_2$ ) die Wahrscheinlichkeit  $p(E_1)p(E_2)$ , und deshalb gilt für die Information

$$I(E_1 \text{ und } E_2) = I(E_1) + I(E_2).$$

4. Ein Ereignis liefert umso mehr Information, je unwahrscheinlicher es ist.



## 1.2 Vergleich mit anderen Informationsdefinitionen

Dieser mathematische Informationsbegriff modelliert also ein Szenario, das aus nichts als zufälligen Ereignissen besteht, deren Wahrscheinlichkeiten man oben-drein kennen muss. Weil der Logarithmus eine umkehrbare Funktion ist, kann man alles, was man durch mathematische Information ausdrücken kann, auch durch Wahrscheinlichkeiten ausdrücken und umgekehrt. Insofern geht der mathematische Informationsbegriff nicht über den von Ereigniswahrscheinlichkeiten hinaus. Man könnte die herkömmliche Wahrscheinlichkeitstheorie durch eine „Informationstheorie“ ersetzen, bei der man immer statt mit Wahrscheinlichkeiten  $p$  mit der Information  $-\log_2(p)$  rechnet. Es ist daher nicht zu erwarten, dass sich aus dem mathematischen Informationsbegriff mehr herausholen lässt als aus der gut etablierten Wahrscheinlichkeitstheorie.

Für die Abgrenzung zum semantischen Informationsbegriff ist festzustellen, dass es keine „Bedeutung“ von Ereignissen im Sinne einer Semantik gibt. Und in Bezug auf den algorithmischen Informationsbegriff spielt der Entstehungsprozess der Ereignisse keine Rolle. Das schränkt die Anwendbarkeit des mathematischen Informationsbegriffs erheblich ein, denn wenn man natürlichsprachigen Texten oder biologischen Signalen oder Abschnitten aus dem Genom *mathematische* Information beismessen will, müsste man alle diese als zufällig ansehen.

## 1.3 Mathematische Entropie

Der mathematische *Entropiebegriff* bezieht sich nicht auf Einzelereignisse und ihren mathematischen Informationsgehalt, d. h. den negativen Logarithmus ihrer Wahrscheinlichkeit, sondern immer auf *Mengen* möglicher Ereignisse. Es sei  $\mathcal{E}$  eine Menge von Ereignissen  $E$ , die jeweils eine Wahrscheinlichkeit  $p(E)$  und eine Information  $I(E) := -\log_2 p(E)$  haben. Jetzt bildet man den Erwartungswert der Information, d. h.

$$\mathbb{E}_{\mathcal{E}}(I) := \sum_{E \in \mathcal{E}} p(E)I(E) = - \sum_{E \in \mathcal{E}} p(E) \log_2 p(E).$$

Auf Anraten von John von Neumann nennt Shannon diesen Ausdruck *Entropie*. Er gibt den *mittleren Informationsgehalt* aller Ereignisse aus  $\mathcal{E}$  wieder, wenn diese mit ihren Wahrscheinlichkeiten gewichtet werden, und bezieht sich auf ein konkretes Zufallsexperiment:

Wenn man immer wieder zufällig ein Ereignis  $E$  aus  $\mathcal{E}$  auswählt (natürlich mit Wahrscheinlichkeit  $p(E)$ ) und dann dessen Information  $-\log_2 p(E)$  betrachtet, so kann man bei genügend vielen Versuchswiederholungen im Mittel die Information  $\mathbb{E}_{\mathcal{E}}(I)$  pro Ereignis erwarten.

Genauereres bringt der Artikel *Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information* von Kurt Schönhammer, ab S. 45 in diesem Band.

## 1.4 Spezial- und Grenzfälle

Beim Würfeln hat man also die Entropie

$$\mathbb{E}_{\mathcal{E}}(I) = -6 \cdot \frac{1}{6} \log_2 \left( \frac{1}{6} \right) = \log_2(6),$$

und man sieht daran auch, dass die Entropie einer Menge von  $n$  gleichwahrscheinlichen Ereignissen immer  $\log_2 n$  ist, also mit dem mathematischen Informationsgehalt jedes Einzelereignisses übereinstimmt.

Die Entropie eines Systems ist Null, wenn nur ein einziges Ereignis möglich ist. Fragt man bei fester Ereignismenge nach der Wahrscheinlichkeitsverteilung, die zu maximaler Entropie führt, so bekommt man die Gleichverteilung heraus. Mit anderen Worten: das Entropiemaximum ist gegeben, wenn sich das System am wenigsten gut vorhersagen lässt, weil alle Ereignisse gleichwahrscheinlich sind. Kurz, lax und ungenau formuliert:

Ein System hat umso mehr Entropie, je weniger leicht es vorhersagbar ist, d. h. je „chaotischer“ es ist.

## 1.5 Entropie und Codierung

Um den Zusammenhang zwischen Entropie und Codierung von Nachrichten herzustellen, betrachten wir das Beispiel eines Lokals, das acht Speisen anbietet, die von den Gästen mit gleicher Wahrscheinlichkeit  $1/8$  gewählt werden. Um eine fehlerfreie primitive Kommunikation mit der Küche zu realisieren, kann man die 8 Gerichte mit den 3 Binärziffern kodieren, die man für die Zahlen  $0, 1, \dots, 7$  braucht, nämlich

$$000, 001, 010, 011, 100, 101, 110, 111.$$

Man kann in der Küche drei Lampen anbringen, die das Bitmuster simultan zeigen (bitparallele Übertragung) oder eine Pulsfolge vereinbaren, die einen hohen Ton für die Null und einen tiefen für die Eins vereinbart (bitserielle Übertragung). Man braucht  $3 = \log_2 8 = \mathbb{E}_{\mathcal{E}}(I)$  Lampen, d. h. die Entropie des Systems gibt die Anzahl der Lampen an.

Die obige Codierung ist aber unpraktisch, wenn einzelne Gerichte sehr viel häufiger bestellt werden als andere, d. h. wenn das System eine niedrigere Entropie als 3 hat. Nehmen wir an, das Gericht 0 würde mit einer großen Wahrscheinlichkeit  $p$  bestellt, während die sieben übrigen Gerichte sich die Restwahrscheinlichkeit  $1 - p$  gleichmäßig teilen. Dann codieren wir

Gericht 0	durch	0
Gericht 1	durch	1001
Gericht 2	durch	1010
Gericht 3	durch	1011
Gericht 4	durch	1100
Gericht 5	durch	1101
Gericht 6	durch	1110
Gericht 7	durch	1111

und berechnen den Erwartungswert der Länge der zu übertragenden Bitfolge:

$$p \cdot 1 + (1 - p) \cdot 4 = 4 - 3 \cdot p$$

und das lohnt sich gegenüber der Drei-Lampen-Lösung schon ab  $p > \frac{1}{3}$ . Die Entropiedefinition liefert direkt das Prinzip der entropieoptimalen Codierung:

Man sollte die Länge der Binärcodierung eines Ereignisses mit Wahrscheinlichkeit  $p$  etwa wie  $-\log_2 p$  wählen, um zu erreichen, dass der Erwartungswert der Codierungslänge etwa gleich der Entropie wird.

Besser geht es nicht, wie man beweisen kann. Insbesondere ist also die mathematische Information eines Ereignisses ein Maß für die Länge einer optimalen Binärcodierung dieses Ereignisses.

Würde die mathematische Informationsdefinition auf Sprachen uneingeschränkt zutreffen, und würden die Sprachen alles optimal kodieren, so müssten die am häufigsten eintretenden Sachverhalte durch die kürzesten Wörter beschrieben werden. Inwieweit dies auf natürliche Sprachen oder das Genom zutrifft, ist zu klären.

## 1.6 Konsequenzen

Der mathematische Entropiebegriff modelliert den Informationsgehalt einer Menge von Ereignissen mit bekannten Einzelwahrscheinlichkeiten, und gibt die zu erwartende mathematische Information bei zufälliger Auswahl eines Ereignisses an. Die mathematische Entropie ist nur für vollständig bekannte Systeme sauber definiert (das hat Folgen für den Entropiebegriff in der statistischen Physik, siehe *Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information* von Kurt Schönhammer, ab S. 45). Wie beim mathematischen Informationsbegriff gibt es keinen Bezug zu irgendeiner Form von Semantik. Aber es gibt einen Bezug zu Sprachen: die Entropie gibt den Erwartungswert der Länge einer optimalen Binärcodierung der Ereignisse des Systems an.

## 2 Bayesianischer Informationsbegriff

Dieser Abschnitt ist bewusst sehr kurz gehalten, weil eine detailliertere und gründlichere Darstellung im Artikel *Wahrscheinlichkeit und Information – wie Henne und Ei?* von Kurt Schönhammer ab S. 27 in diesem Band vorliegt.

### 2.1 Information als *change of beliefs*

Der Bayes'sche Wahrscheinlichkeitsbegriff interpretiert Wahrscheinlichkeit als *Grad persönlicher Überzeugung (degree of belief)*. Neue Informationen ändern diesen Grad, und dann kommt man zu „*Information is whatever forces a change of rational beliefs*“ [3, S. 4]. Im Sinne des Bayesianismus ist *belief* eine a-priori-Annahme gewisser Wahrscheinlichkeitsaussagen im Sinne des Abschnitts 1, und

durch zusätzliche Informationen, z. B. Messungen, können sich diese Wahrscheinlichkeitsaussagen verändern, wodurch man einen *change of beliefs* hat. Caticha erklärt das in [3] so:

S. 4: *It may be worthwhile to point out an analogy with Newtonian dynamics. The state of motion of a system is described in terms of momentum – the „quantity“ of motion – while the change from one state to another is explained in terms of an applied force. Similarly, in Bayesian inference a state of belief is described in terms of probabilities – the „quantity“ of belief – and the change from one state to another is due to information. Just as a force is defined as that which induces a change in motion, so information is that which induces a change of beliefs.*

S. 7: *When there is no new information there is no reason to change one's mind.*

Man sollte letzteres umdrehen: Wenn sich die *beliefs* nicht ändern, war auch keine Information im Bayes'schen Sinne da.

## 2.2 Bedingte Wahrscheinlichkeiten

Der einfachste Fall einer Änderung von Wahrscheinlichkeiten durch Hinzunahme einer neuen Information als Resultat einer Beobachtung wird beschrieben durch *bedingte Wahrscheinlichkeiten*. Genaueres findet sich im Artikel *Wahrscheinlichkeit und Information – wie Henne und Ei?* von Kurt Schönhammer ab S. 27.

Falls man „*change of beliefs*“ nicht wahrscheinlichkeitstheoretisch, sondern kognitiv interpretiert, bewirkt Information eine Änderung eines mentalen oder kognitiven Zustands. Das entspricht der in Abschnitt 4.6 genauer dargestellten Rolle der Information als Zustandsänderung eines interpretierenden Systems. Die *Informationsverarbeitung in der Kognitionspsychologie und in den kognitiven Neurowissenschaften* wird im Artikel von Gerd Lüer und Uta Lass ab S. 89 in diesem Band behandelt, während eine allgemeine Theorie der Information als Zustandsänderung von Prozessen im Beitrag *Daten, Prozesse und Information* von Robert Schaback ab S. 183 dargestellt wird.

## 3 Algorithmischer Informationsbegriff

Nach Gregory Chaitin [4] ist die algorithmische Informationstheorie „*the result of putting Shannon's information theory and Turing's computability theory into a cocktail shaker and shaking vigorously*“.

### 3.1 Texte als Zeichenketten

Gegenüber dem mathematischen Informationsbegriff aus Abschnitt 1 geht es hier nicht um Wahrscheinlichkeiten, sondern wie beim semantischen Informationsbegriff

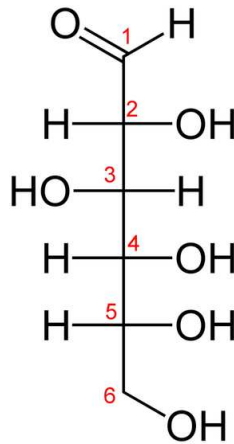


Abbildung 1. Strukturierte Daten aus der Chemie

des Abschnitts 4 um den Informationsgehalt von *Texten* oder *Nachrichten*. Diese sind *Zeichenketten*, d. h. Folgen oder mehrdimensional strukturierte Anordnungen von Zeichen eines Alphabetes  $A$ . Weil man Satzzeichen und Steuerzeichen (z. B. Zeilenvorschub, Seitenanfang) mit zum Alphabet rechnet, fallen alle natürlich-sprachlichen Texte unter diese Definition, ebenso wie einige Spezialsprachen:

- $a^2 + b^2 = c^2$  in der Mathematik,
- `for (i=0; i<j; i++)` in der Informatik,
- CTTATTCATCTGGTGATTTGGCTACTTCTTAA im Genom,
- LCLYTHIGRNIYYGSYLYSETWNTGIMLLLITMATAFMG...  
im FASTA-Format zur Beschreibung der Primärstruktur von Proteinen,
- *An jenem Tag im blauen Mond September...* in der Literatur,
- Partituren in der Musik, siehe Abb. 3,
- Strukturformeln in der Chemie, siehe Abb. 1.

### 3.2 Syntax und Semantik von Texten

Texte im obigen Sinne haben eine Struktur, die gewissen Regeln genügen muss, die die *Syntax* des Textes formal beschreiben. Sie haben auch eine *Semantik*, die unten im Abschnitt 4 eine zentrale Rolle spielen wird und im weitesten Sinne die „Bedeutung“ des Textes ausmacht.

### 3.3 Komplexitätsdefinition

Der Informationsgehalt eines Textes wird aber beim algorithmischen Informationsbegriff nicht über Syntax oder Semantik eines Textes definiert, sondern nur durch dessen *Komplexität*. Letztere wird exakt definiert durch die *minimale Länge eines Algorithmus, der in einem vorgegebenen Maschinen- oder Sprachbeschreibungsmodell diesen Text produziert*, unter Einhaltung der Syntaxregeln. In der Informatik gibt es diverse Maschinenmodelle, und für jedes Modell ist klar definiert, was ein Algorithmus ist und welche Länge er hat. Die allgemeine *Komplexitätstheorie* der Informatik ist sehr viel weiter gefasst und wird hier ignoriert.

Zum Standardlehrstoff der Informatik gehört auch der Zusammenhang zwischen Maschinenmodellen und Klassen formaler Sprachen, wobei man allerdings nicht direkt die Komplexität einzelner Texte behandelt, sondern danach fragt, welches Maschinenmodell erforderlich ist, um beliebige Texte aus einer durch eine Klasse von generativen Grammatiken definierte Klasse von Sprachen syntaktisch zu analysieren. Das stiftet einen Zusammenhang zwischen Maschinenmodellen und der Komplexität von generativen Grammatiken.

### 3.4 Konsequenzen

Der obige Komplexitätsbegriff ist quantitativ nicht praktikabel, weil man ziemlich leicht beweisen kann, dass es kein Programm auf einer Turingmaschine gibt, die zu einem gegebenen Text dessen Komplexität berechnet. Er liefert dennoch einige nützliche Einsichten zum Thema Information:

- Wenn die biologische Evolution als Evolution der Genome als Zeichenketten gesehen wird, ist der Zusammenhang zwischen der Komplexität des Textes mit der Struktur der generierenden „Maschine“ interessant.
- Auch die Evolution natürlicher Sprachen kann unter dem Gesichtspunkt der Komplexitätssteigerung gesehen werden, ebenso wie der Verlauf des Spracherwerbs bei Kindern.
- Effiziente technische Übertragungstechniken für Audio- und Videodaten versenden nicht die Daten selber, sondern Anweisungen zu deren Synthetisierung. Man versendet nicht den Kuchen, sondern das Rezept. Der algorithmische Informationsgehalt des Kuchens ist gleich der Länge des kürzestmöglichen Rezepts für diesen Kuchen in einer festgelegten „Rezeptsprache“.
- Bei der Notenschrift, der Sprache des Genoms und den Syntheseanweisungen für Proteine ist ebenfalls die Nachricht gleich dem Syntheserezept.
- Das sind jeweils Texte, die durch *Ausführung* interpretiert werden, siehe Abschnitt 4.5.

## 4 Semantischer Informationsbegriff

### 4.1 Information durch Interpretation von Nachrichten

Nach dem klassischen Informatik-Lehrbuch [2] von F. L. Bauer und G. Goos entsteht *Information* durch *Interpretation* einer *Nachricht*. Dabei sind Nachrichten als *strukturierte Daten* oder *Signale* zu sehen, die einem physikalischen *Medium*, z. B. Papier, Schall, elektromagnetischen Wellen als *Struktur* aufgeprägt sind. *Information* entsteht erst, wenn diese Daten durch irgendeinen Prozess, der z. B. in Menschen, Tieren, Zellen oder Maschinen ablaufen kann, *interpretiert* werden. Die entstehende Information ist von diesem Interpretationsprozess abhängig. Die Frage nach der „Natur der Information“ wird dadurch reduziert auf die Analyse der verschiedenen Interpretationsprozesse. Man wird deshalb sehr verschiedene Antworten bekommen, je nachdem welchen Interpretationsprozess man ins Auge fasst.

### 4.2 Information als „*semantic content*“

Diese der Philosophie zuzurechnende Definition ist auf den ersten Blick nicht wesentlich anders. Sie hat eine interessante Geschichte [15] und wird von Floridi [8] formuliert als

*$\sigma$  is an instance of information, understood as semantic content, if and only if:*

- *$\sigma$  consists of one or more data;*
- *the data in  $\sigma$  are well-formed;*
- *the well-formed data in  $\sigma$  are meaningful.*

### 4.3 Vergleich

Der Informationsbegriff nach Floridi setzt *semantic content* und *meaning* voraus, was immer das auch sei, während der Informationsbegriff nach 4.1 nur von einem *Interpretationsprozess* spricht, dessen Ergebnis *Information* ist. Bei Floridi wird nicht spezifiziert, was *Semantik* oder *meaning* ist und durch welchen Vorgang sie entsteht, während in 4.1 ein unspezifizierter *Interpretationsprozess* postuliert wird, dessen Ergebnis *Information* ist. Begriffe wie *semantic content* oder *meaning* sind nachgeordnet und werden nicht gebraucht.

Bei Floridi gehören die *well-formed data* mit zum Informationsbegriff, während *Daten* oder *Nachrichten* bei der Definition in Abschnitt 4.1 als potentiell uninterpretierte Objekte nichts mit Information zu tun haben, solange sie nicht interpretiert werden. Diese Diskrepanz sieht Mingers [15] so, dass der Interpretationsprozess die Information in 4.1 *subjektiv* macht, während sie bei vollständiger Bindung an die Daten *objektiv* genannt werden kann und auch ohne Interpretation existiert.

Wenn die Information an die Daten gebunden ist, und nicht erst durch einen Interpretationsprozess entsteht, muss man einen wie auch immer gearteten, jedenfalls aber sehr allgemeinen und objektiven Semantikbegriff voraussetzen, der seinerseits der Klärung bedarf, ebenso wie der Begriff des Interpretationsprozesses, wenn man der subjektiven Definition den Vorzug gibt.

Wie man sich auch immer entscheidet: es ist in beiden Fällen klärungsbedürftig, was *Semantik* ist. Entweder als Zugangsvoraussetzung zum Informationsbegriff oder als Ergebnis von Interpretationsprozessen. Der zweite Weg bietet die Chance, gleichzeitig *Information* und *Semantik* zu untersuchen, zumal nicht klar ist, wie man *Semantik*, ohne den Begriff der *Information* zu haben, überhaupt verstehen und an den Anfang einer Begriffsbildung von *Information* stellen kann.

Es erweist sich deshalb für das Folgende als pragmatisch vorteilhaft, Interpretationsprozesse genauer unter die Lupe zu nehmen, um dem semantischen Informationsbegriff und damit auch der Semantik näher zu kommen.

Im Gegensatz zu den anderen Informationsbegriffen geht es beim semantischen Informationsbegriff weder um irgendwelche Zufallsprozesse noch um die Komplexität eines datenerzeugenden Verfahrens. Insofern ist keine Überschneidung mit dem mathematischen und dem algorithmischen Informationsbegriff aus den Kapiteln 1 bzw. 3 erkennbar. Ein Zusammenhang mit dem bayesianischen Informationsbegriff aus dem Kapitel 2 findet sich in Abschnitt 4.6.

## 4.4 Beispiele

Ein typischer Fall sind die strukturierten Daten, die uns der Anblick des Sternenhimmels bietet. Bei Floridi's Definition sind die Daten und die Information eine Einheit. Bei der Sichtweise aus Abschnitt 4.1 gewinnen wir diesen Daten erst durch Interpretation Informationen ab, beginnend bei den Tierkreiszeichen bis hin zu Spektraldaten von Galaxien. Die Informationen hängen vom Interpretationsprozess ab und können sehr verschiedenartig sein, obwohl die Daten dieselben sind und unabhängig vom Interpretationsprozess existieren. Es wird kein „Sender“ angenommen, der „Information“ in die Nachricht packt und diese dann absichtsvoll verschickt. Information lässt sich nur aus der Perspektive des Interpretierenden oder des „Empfängers“ sehen [6]. Der Satz *Semantic information is the propositional content of data* [15, S. 391] aus der „objektiven“ Variante der semantischen Informationsdefinition ist beim Anblick des Sternenhimmels problematisch.

Strukturierte Texte wie in Abschnitt 3.1 auf Seite 16 und in den Abbildungen 1, 2 und 3 sind die gängigsten Beispiele für Nachrichten bzw. strukturierte Daten. Sie haben eine syntaktische Struktur und einen semantischen Gehalt, der sich bei geeigneter Interpretation erschließt. Syntax und Semantik werden üblicherweise weder den strukturierten Nachrichten noch der eventuell daraus ableitbaren Information zugeordnet, sondern den formalen oder natürlichen *Sprachen*, in denen die Texte abgefasst sind.

Die Programmiersprachen der Informatik werden syntaktisch exakt definiert [10] und ihre Syntaxanalyse wird vom Computer nach den Syntaxregeln ausgeführt. Bei modernen Sprachen wie JAVA findet allerdings lediglich eine Übersetzung



in eine andere Sprache statt, die von einer „virtuellen Maschine“ durch Aktion interpretiert wird. Deshalb befindet sich die Spezifikation der Semantik der Sprache in der Definition der zugehörigen virtuellen Maschine [12]. Die Interpretation durch Aktion wird in Abschnitt 4.5 genauer behandelt.

Als ein weiteres Beispiel strukturierter Daten betrachten wir Abb. 2. Auch wenn

道	可	道
非	常	道
名	可	名
非	常	名

Abbildung 2. Strukturierte Daten

keinerlei semantisches Verständnis vorliegt, ist die syntaktische Strukturierung durch eine Art Reimschema und gewisse Symmetrien klar erkennbar. Das ist allerdings nur eine Syntaxanalyse und keine Interpretation. Sie könnte auch von einem Computer durchgeführt werden. Eine Interpretation<sup>1</sup> setzt sehr viel mehr voraus und ist offenbar unter Sinologen und Philosophen bis heute heftig umstritten.

An dieser Stelle kann man einwenden, dass Eigenschaften der *Struktur* der Daten, z. B. Symmetrien, bereits Information sind und objektiv den Daten angehören. Eigenschaften von Strukturen sind aber nicht gleich den Strukturen, sondern Aussagen über Strukturen, während Strukturen nur dadurch Strukturen sind, dass sie bewirken, dass sich strukturierte Daten überhaupt als etwas Gegebenes vom Nicht-Gegebenen unterscheiden lassen. Insofern gehen Eigenschaften von Strukturen über strukturierte Daten hinaus. Sie sind selbst nicht unmittelbar gegeben, sondern erfordern einen Interpretationsprozess. Wenn dieser auf die Struktureigenschaften fokussiert, kommt er mit Notwendigkeit zu „objektiven“ Ergebnissen. Das kann der subjektive semantische Informationsbegriff durchaus konzederieren.

In gewissen Fällen kann eine starke und dem üblichen Interpretationskontext entsprechende Strukturierung eines Textes dazu verleiten, eine nicht existente Semantik vorzugaukeln. Zum Beispiel ist es relativ leicht, für wissenschaftliche Spezialsprachen Texte zu generieren, die syntaktisch korrekt und semantisch absolut sinnlos sind. Mit MathGen [5] kann man sich eine hochtrabend klingende com-

<sup>1</sup>Láo-zi: *dào dé jīng* (Lao-tse: *Tao-Te-King*)

In der strukturtreuen Übersetzung von Günter Debon:

*Könnten wir weisen den Weg, es wäre kein ewiger Weg;*

*Könnten wir nennen den Namen, es wäre kein ewiger Name*

putergenerierte mathematische Arbeit in Sekundenbruchteilen schreiben lassen. Mehrfach sind solche Produkte zur Publikation angenommen worden, wenn auch nur in dubiosen Journalen [9].

The image shows a musical score for five voices: Soprano (S), Alto (A), Tenor (T), and Bass (B). The music is in 4/4 time and G major. The lyrics are: "auf - - ge - rafft, und nie - mand ach - - tet" and "auf - - ge - rafft, und nie - mand ach - - tet drauf". The Alto part has red letters 'b', 'a', 'c', 'h' under the notes for "und nie - mand ach - - tet drauf".

Abbildung 3. Strukturierte Daten

Abbildung 3 zeigt ein Beispiel [1] für versteckte Information in strukturierten Daten. Musiker interpretieren diese strukturierten Daten durch Aktion und stellen ein Tonsignal her, das ebenfalls aus strukturierten Daten besteht und das von Hörenden neu interpretiert werden muss. Es ist in diesem Falle allerdings fraglich, ob alle Hörer die versteckte B-A-C-H- Information herausinterpretieren können, denn „... niemand achtet drauf“.

## 4.5 Interpretation durch Aktion

In vielen Fällen erfolgt die Interpretation strukturierter Daten durch eine konkrete Aktion des Interpretierenden:

- auf dem Kasernenhof durch unmittelbares und unreflektiertes Befolgen sprachlicher Befehle,
- in der Informatik ebenfalls durch Befehlsausführung, z. B.  $z=3*x-y$ ;
- durch Ausführen von biochemischen Synthesen bei Interpretation von Genomdaten,
- durch unmittelbare Reaktionen von Organismen auf Reize jedweder Art, z. B. Warn- oder Locksignale.

Wenn die Ausführung durch die Nachricht eindeutig definiert ist, kann man mit gutem Recht die Information, die aus der Nachricht durch Interpretation entsteht, mit der wohldefinierten und im Idealfall auch stets eintretenden Aktion des Interpretierenden gleichsetzen. In diesem Falle liegt die „Natur der Information“ auf der Handlungsebene, und die Information ist eine Aktion. Sofern diese objektiv beobachtbar ist, wie etwa in der Verhaltensforschung, kann man Information dieser Art exakt beschreiben.

## 4.6 Information als Zustandsänderung

Liegen keine Interpretationsprozesse vor, die als Aktionen äußerlich und objektivierbar manifest werden, kann man zumindest eine potentielle Zustandsänderung betrachten, die im Interpretierenden durch die Interpretation erfolgt, und diese als Interpretationsergebnis und damit als Information ansehen. Das trifft zu bei der Lektüre von Büchern oder dem Hören von Sprache oder Musik.

Wenn man eine Aktion eines Organismus, eines Bewusstseins, einer Maschine oder eines biochemischen Reaktionssystems als Konsequenz einer Zustandsänderung des Systems ansieht, ist die in Abschnitt 4.5 beschriebene Interpretation durch Aktion eine Konsequenz einer Interpretation durch Zustandsänderung, also ein Spezialfall dieses Abschnitts.

Diese Argumentationslinie berührt den sehr allgemeinen Begriff von Kognition

*If a living system enters into cognitive interaction, its internal state is changed in a manner relevant to its maintenance, and it enters into a new interaction without loss of its identity [14, S. 13]*

bei Maturana [13]. Dort wird jede Interaktion eines lebenden Organismus mit der Umwelt als kognitiver Prozess gesehen. In diesem Sinne kann man das Umfeld eines Organismus als eine permanent vorhandene und sich zeitlich und räumlich ändernde Nachricht sehen, die ständig vom Organismus interpretiert wird und dann zu lebenserhaltenden Aktionen des Organismus führt. Der kognitive Prozess ist in ständigem Wandel durch *Lernen* aus der Beobachtung der Umwelt, und er führt dann zur Interpretation von Signalen durch Aktion:

*In animal communication, „meaning“ is generated when the recipient has learned that the occurrence of signal A reliably predicts event B [7, S. 133].*

Dieses Beispiel stützt die These, eher Semantik aus Interpretationsprozessen abzuleiten als umgekehrt Semantik vorauszusetzen, um Information zu definieren.

Auch den Informationsbegriff als „*change of beliefs*“ aus Abschnitt 2.1 kann man hier einordnen, denn es geht um eine Zustandsänderung, in diesem Falle von „*beliefs*“. Diese können einerseits innerhalb des kognitiven Bereichs liegen, andererseits aber auch quantitativ durch Veränderungen wahrscheinlichkeitstheoretischer „*priors*“ erfasst werden.

## 4.7 Konsequenzen

Der semantische Informationsbegriff hat zur Folge, dass Information aus einer sehr allgemein zu fassenden *Zustandsänderung des eine Nachricht interpretierenden Systems* besteht. Die Frage nach der „Natur der Information“ erfordert dann eine spezifische Betrachtung solcher Systeme und ihrer typischen Zustandsänderungen. Dies wird im Beitrag *Daten, Prozesse und Information* ab S. 183 ausgearbeitet.

## Literatur

- [1] <http://de.wikipedia.org/wiki/B-A-C-H> <2019-03-14>.
- [2] F. L. Bauer und G. Goos. *Informatik: Eine einführende Übersicht, Teil 1, Heidelberger Taschenbücher*, Band 80. Springer, 1973.
- [3] A. Caticha. Information and entropy. *AIP Conference Proceedings*, 954(1):11–22, 2007.
- [4] G. J. Chaitin. Zahlen und Zufall. Algorithmische Informationstheorie. Neueste Resultate über die Grundlagen der Mathematik. In *Naturwissenschaft und Weltbild*, Seiten 30–44. Hölder-Pichler-Tempsky, Wien, 1992.
- [5] N. Eldredge. <http://thatsmathematics.com/mathgen> <2019-03-14>.
- [6] J. Fischer. Where is the Information in Animal Communication? In R. Menzel, R. and J. Fischer, editors, *Animal Thinking: Contemporary Issues in Comparative Cognition*, 151–161. MIT Press, Cambridge, 2011.
- [7] J. Fischer and K. Hammerschmidt. Information and Influence in Animal Communication. In A. D. M. Smith, editor, *The Evolution of Language*, pages 129–136. World Scientific, 2010.
- [8] L. Floridi. Semantic Conceptions of Information. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2014. <http://plato.stanford.edu/archives/spr2014/entries/information-semantic> <2019-03-14>.
- [9] S. Friedl. <http://thatsmathematics.com/blog/archives/185> <2019-03-14>.
- [10] J. Gosling, G. Steele, G. Bracha, and A. Buckley. *The Java Language Specification, Java SE 8 edition, 2014*. <http://docs.oracle.com/javase/specs/jls/se8/jls8.pdf> <2019-03-14>.
- [11] G. Hotz. *Algorithmische Informationstheorie*. Teubner-Texte zur Informatik, 25. B. G. Teubner Verlagsgesellschaft mbH, Stuttgart, 1997.
- [12] T. Lindholm, F. Yellin, G. Bracha, and A. Buckley. *The Java Virtual Machine Specification, Java SE 8 edition, 2014*. <http://docs.oracle.com/javase/specs/jvms/se8/jvms8.pdf> <2019-03-14>.

- 
- [13] H. Maturana. Cognition. In P. M. Hejl, W. K. Köck, and G. Roth, Herausgeber: *Wahrnehmung und Kommunikation*, pages 29–49. Frankfurt: Peter Lang, 1978.
- [14] H. R. Maturana and F. J. Varela. *Autopoiesis and Cognition: The Realization of the Living*, volume 42 of *Boston Studies in the Philosophy of Science*. Reidel, 1980.
- [15] J. Mingers. Prefiguring Floridi’s Theory of Semantic Information. *TripleC*, 11:388–401, 2013.
- [16] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.
- [17] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana, Ill., 1949.
- [18] C. E. Shannon und W. Weaver. *Mathematische Grundlagen der Informationstheorie*. R. Oldenbourg Verlag, 1976.
- [19] N. J. A. Sloane and A. D. Wyner. *Claude E. Shannon: Collected Papers*. Wiley-IEEE Press, 1993.



# Wahrscheinlichkeit und Information – wie Henne und Ei?

*Kurt Schönhammer*

*Institut für Theoretische Physik, Friedrich-Hund-Platz 1, D-37077 Göttingen*

## 1 Einleitung

Gezielte Information ist hilfreich, um besser einschätzen zu können wie es weiter geht. Oft bewirkt eine Information direkt eine Handlung. Wegen der häufigen Verspätungen bei der Bahn, rufe ich meine Frau aus dem Zug an, wenn er in Kassel abgefahren ist. Fünf Minuten später fährt sie mit dem Auto los, um mich am Göttinger Bahnhof abzuholen. Wie kurz kann das Gespräch sein, um die Aktion in Gang zu setzen? Reicht ein „jetzt“ oder gar nur ein bit, falls man das mit dem Handy könnte? Das hängt natürlich von den getroffenen Vereinbarungen ab. Claude Shannon hat 1948 ein *quantitatives Maß* für den Informationsgehalt einer Nachricht vorgeschlagen. Es ist eine Funktion der *Wahrscheinlichkeit* der Nachricht.

Auch der Begriff der Wahrscheinlichkeit scheint intuitiv verständlich. Der Versuch ihn zu definieren, führt aber leicht zu Zirkelschlüssen. In meinem Beitrag zum Entropiekonzept in der Physik ab S. 45 habe ich dargelegt, dass der Wahrscheinlichkeitsbegriff in der Statistischen Mechanik eine zentrale Rolle spielt. Die von Neumannsche Entropiedefinition verwendet Wahrscheinlichkeiten. Daher ist es wichtig, sich zu verdeutlichen, wie man Wahrscheinlichkeiten angeben, bzw. interpretieren kann. Im Folgenden wird dargelegt, dass es zwei sehr unterschiedliche Interpretationen gibt. Das Wechselspiel zwischen Wahrscheinlichkeit und Information kehrt sich beim Übergang von der einen Interpretation zur anderen gerade um, was mich zum Titel dieses Beitrags veranlasst hat.

Der Hauptteil des Folgenden ist eine einfache Einführung in die Wahrscheinlichkeitstheorie unter dem Gesichtspunkt, die beiden verschiedenen Interpretationen zu verdeutlichen. Es werden mehrere Beispiele ausführlich behandelt. Diese Einführung ist für Nicht-Experten gedacht und erfordert nur geringe mathematische Vorkenntnisse.

## 2 Wahrscheinlichkeitstheorie

Den Begriff „Wahrscheinlichkeit“ findet man nicht nur in der Mathematik oder den Natur- oder anderen Wissenschaften, sondern auch in der Alltagssprache. Bereits hier zeigt sich, dass der Begriff mit unterschiedlichen Bedeutungen belegt ist. Hier zwei Beispiele:

- Wie groß ist die Wahrscheinlichkeit, dass ich bei einmaligem Würfeln eine Sechs erhalte?
- Die Wahrscheinlichkeit, dass es morgen regnen wird, ist achtzig Prozent.

Auch ohne eine präzise Definition des Begriffs *Wahrscheinlichkeit*, die sich als problematisch herausstellt, weiß man in beiden Fällen, was gemeint ist. Das erste Beispiel steht in Beziehung zur Entstehung der Wahrscheinlichkeitstheorie, als praktisches Hilfsmittel zur Strategie in *Glücksspielen*. Wichtige Beiträge dazu aus dem siebzehnten Jahrhundert stammen von Blaise Pascal, Pierre de Fermat, Jakob Bernoulli und Pierre Simon Laplace [1]. Bereits bei diesen Anfängen stellte es sich als einfacher heraus, mit Wahrscheinlichkeiten zu hantieren, als sie zu definieren. Diese Regeln wurden 1933 von A. Kolmogorov axiomatisch formuliert [2]. Dazu benötigt man zwei Typen von Angaben:

1. die Menge  $\Omega$  der möglichen Ereignisse  $E_i$ ,
2. die Wahrscheinlichkeiten  $P(E_i)$  für das Auftreten der Ereignisse  $E_i$ .

Zu 1.: Die Menge kann *diskret* sein, wie bei der Augenzahl beim einmaligen Würfeln:  $\Omega = \{E_1, E_2, E_3, E_4, E_5, E_6\} \equiv \{1, 2, 3, 4, 5, 6\}$ , d. h.  $E_i$  ist das Ereignis mit der Augenzahl  $i$ .

Neben den „elementaren Ereignissen“ betrachtet man häufig *zusammengesetzte* Ereignisse. Beim einmaligen Würfeln z. B. das Ereignis  $A$ , dass eine gerade Augenzahl gewürfelt wird.

Bei zweimaligem Würfeln besteht  $\Omega$  aus den 36 Paaren  $(ij)$ , z. B.  $(61)$ , wobei es auf die Reihenfolge von  $i$  und  $j$  ankommt, d. h.  $(61)$  und  $(16)$  sind verschiedene Ereignisse.

Die Geschwindigkeit eines Teilchens oder die Länge eines Stabes sind Beispiele, in denen die Ereignisse *kontinuierlichen* Charakter haben. Da dieser Fall mathematisch anspruchsvoller ist, werden wir uns hier auf den diskreten Fall beschränken.

Zu 2.: Die Wahrscheinlichkeit  $P(A)$  für ein zusammengesetztes Ereignis  $A$  ist die *Summe* aus den Wahrscheinlichkeiten  $P(E_i)$ , aus denen sich das Ereignis zusammensetzt, wenn die Einzelereignisse voneinander unabhängig eintreten können. Für das obige Beispiel einer geraden Augenzahl also  $P(A) = P(E_2) + P(E_4) + P(E_6)$ .



Allgemeiner lauten die Axiome

- Für jede Teilmenge  $A \in \Omega$  gilt:  $P(A) \geq 0$ ,
- Für alle elementfremden Teilmengen  $A$  und  $B$  gilt für die Vereinigungsmenge  $A \cup B$ :

$$P(A \cup B) = P(A) + P(B),$$

- Für die gesamte Menge  $\Omega$  gilt:  $P(\Omega) = 1$ .

Der nützliche Begriff der *bedingten Wahrscheinlichkeit* wird später eingeführt.

Dies ist zunächst eine rein mathematische Definition des Wahrscheinlichkeitsbegriffes, ähnlich wie die Peano-Axiome definieren, was eine natürliche Zahl ist. Objekte, welche die Kolmogoroff-Axiome erfüllen, kann man „Wahrscheinlichkeiten“ nennen. Das hilft einem für die Praxis allerdings wenig, da noch nicht klar ist, wo die „Wahrscheinlichkeiten“ herkommen.

Dazu betrachten wir unser erstes Beispiel mit dem Würfeln einer Sechs. Von einem nichtpräparierten Würfel erwartet man aus *Symmetriegründen*, dass alle Augenzahlen *gleichwahrscheinlich* sind. Da die Summe aller sechs Wahrscheinlichkeiten nach den Axiomen gleich eins ist, setzt man also

$$P(E_i) \equiv P(i) = \frac{1}{6}$$

für alle  $i$  von 1 bis 6. Dabei haben wir unterdrückt, dass es sich um Wahrscheinlichkeiten *unter der Annahme eines nichtpräparierten Würfels* handelt.

Ein etwas komplexeres Beispiel, das in der Geschichte der Wahrscheinlichkeitsrechnung eine Rolle gespielt hat, ist die Betrachtung der *Summe*  $S$  der Augenzahlen von *drei* (nichtpräparierten) Würfeln, die Werte von 3 bis 18 annehmen kann. Für den ersten (weißen), zweiten (schwarzen) und den dritten (roten) Würfel gibt es bei jedem Wurf jeweils sechs gleichwahrscheinliche Möglichkeiten. Also ergeben sich  $6 \cdot 6 \cdot 6 = 216$  *gleichwahrscheinliche Ereignisse* für einen Wurf der drei Würfel. Die Wahrscheinlichkeit für einen speziellen Wert der Summe bei einem Wurf erhält man daher, indem man *die Anzahl der verschiedenen Arten diese Summe zu erhalten durch die Gesamtzahl der möglichen Ereignisse teilt*. Für das Beispiel  $S = 17$  muss einer der drei Würfel eine fünf zeigen, die anderen beiden eine sechs, d. h. die Wahrscheinlichkeit für  $S = 17$  ist gleich  $3/216$ .

Wenn man das „Experiment“ *einmal Würfeln* mit einem nichtpräparierten Würfel sehr häufig wiederholt, so merkt man, dass *im Mittel* jede Augenzahl von 1 bis 6 *gleichhäufig* vorkommt. Hat man  $N$ -mal gewürfelt, so gilt für die Zahl  $N_6$  der Ereignisse mit einer Sechs ungefähr  $N_6/N \approx 1/6$ , was  $P(6) = 1/6$  entspricht.

Wenn man sich die Wahrscheinlichkeiten  $P(E_i)$  *nicht* aus einem Symmetrieargument verschaffen kann, so kann man die bei  $N$ -facher Wiederholung des entsprechenden „Experiments“ gefundenen *relativen Häufigkeiten*  $N_i/N$  im Grenzfall („lim“) unendlicher Wiederholung als die gesuchten Wahrscheinlichkeiten betrachten

$$P(E_i) = \lim_{N \rightarrow \infty} \frac{N_i}{N}.$$

Für unsere Würfelbeispiele kommen wir so zu denselben Wahrscheinlichkeiten wie mit dem Symmetrieargument.

Hier also die erste „Interpretation“ des mathematischen Wahrscheinlichkeitsbegriffs, unter der die mathematischen Begriffsbildungen als empirisch zutreffende Behauptungen interpretiert werden können:

*I1: Wahrscheinlichkeiten als relative Häufigkeiten von Ereignissen*

*I1* ist die Interpretation von Wahrscheinlichkeiten, die z. B. Studierende der Physik in Vorlesungen wie „Statistische Mechanik“ oder „Quantenmechanik“ üblicherweise präsentiert bekommen. In der statistischen Mechanik wird ein „Ensemble“ von identisch präparierten (makroskopischen) Systemen betrachtet und man berechnet physikalische Größen als Mittel über die Werte an den einzelnen Systemen. Das Ensemble existiert natürlich nur im Kopf des Theoretikers. Es wäre viel zu teuer all die Kopien anzuschaffen. Man wiederholt lieber das Experiment mit der *einen* Anordnung mehrfach.

Das mit Hilfe der Born'schen Wahrscheinlichkeitsinterpretation des Absolutquadrats der Wellenfunktion vorhergesagte Interferenzmuster bei der Streuung an einem Doppelspalt mit stark verdünnten Strahlen erscheint erst allmählich.

## 2.1 Zusammenhang mit dem Informationsbegriff

Bei der Entwicklung des Morse-Alphabets war es eine wichtige Einsicht, auf die *relative Häufigkeit des Auftretens der Buchstaben* in einem üblichen (englischen) Text zu achten, und den *häufig* auftretenden Buchstaben *kurze* Symbole zuzuordnen. So wird der Buchstabe *e* durch *einen* Punkt, *t* durch *einen* Strich, die seltenen Buchstaben *x*, *y* und *z* aber durch *vier* Symbole dargestellt. Dadurch konnte man die Übertragungslänge eines Textes wesentlich verkürzen.

Nummeriert man die Buchstaben von *a* bis *z* mit  $i = 1, 2, \dots, 26$  durch und nimmt als Ereignis  $E_{27}$  das Leerzeichen hinzu, so entspricht  $E_5$  dem Buchstaben *e*, der in einem deutschen Text mit der größten relativen Häufigkeit (als Wahrscheinlichkeit interpretiert)  $p_5 \approx 0.174$  auftritt, während  $E_{17}$ , das dem Buchstaben *q* entspricht mit  $p_{17} \approx 0.02$  am wenigsten häufig auftritt [4].

Claude Shannon hat 1948 eine quantitative Analyse des *Informationsgehalts* einer Nachricht vorgelegt [3,4,5]. Er arbeitete hauptsächlich mit einem *binären* Alphabet mit den „Buchstaben“ 0 und 1. Ist die Nachricht ein „string“

011...101

der Länge  $N$ , so gibt es  $2^N$  verschiedene strings dieser Länge, d. h. die Ereignismenge besteht aus  $2^N$  Elementen. Kommen alle strings mit *gleicher relativer Häufigkeit* vor, so gilt für alle strings  $p(E_i) = 1/2^N$ . Ein offensichtliches Maß für den Informationsgehalt der Nachricht ist die Zahl  $N$  der übertragenen 0-1-Entscheidungen „bits“, d. h. die *Länge* des strings. Es gilt

$$h(E_i) := \log_2 \left( \frac{1}{p(E_i)} \right) = \log_2 2^N = N \log_2 2 = N,$$

wobei  $\log_2$  der Logarithmus zur Basis 2 ist [6]. Also liefert die Funktion  $h$  das als „offensichtlich“ angesehene Ergebnis. Shannons wichtige Einsicht war, dass der erste Teil der obigen Formel auch für den Fall nicht gleichhäufiger strings *das nützliche Maß für den Informationsgehalt* darstellt:

$$h(E_i) := \log_2 \left( \frac{1}{p(E_i)} \right) = -\log_2(p(E_i)).$$

Der Mittelwert über das gesamte Ensemble wird *Informationsentropie*  $H$  genannt

$$H := - \sum_i p(E_i) \log_2(p(E_i)).$$

Zur formal analogen Entropiedefinition in der statistischen Mechanik verweise ich auf mein Manuskript auf S. 45 zum Entropiebegriff in der Physik. Es sei angemerkt, dass Shannon bei der Namensgebung dem Rat John von Neumanns gefolgt ist, der 1927 die allgemeine Entropiedefinition in der Quantenstatistik formuliert hat: „weil sowieso niemand wisse, was Entropie wirklich ist, sei Shannon in einer Debatte immer im Vorteil“ [7]. Peter Janich schreibt in seinem Buch „Was ist Information?“ dazu, dass hier bestimmte Strukturgleichheiten vorliegen, die aber „nicht in der Gleichheit der Gegenstandsgebiete Nachrichtenübertragung und Wärmeübertragung“ liegen, sondern in der Gleichheit der Betrachtung, mehr noch, „in der Gleichheit der philosophischen Vorurteile“, unter denen der Vergleich der beiden Gebiete steht [8].

Für den betrachteten Fall eines deutschen Textes gilt für die zwei Buchstaben  $e$  und  $q$

$$h(e) = -\log_2(0.174) \approx 2.52, \quad h(q) = \log_2(0.02) \approx 5.64,$$

und die Informationsentropie, d. h. der *mittlere Informationsgehalt* eines Buchstabens ist  $H \approx 4.1$ .

Bei Shannon stand das *praktische* Problem im Vordergrund, eine „Nachricht“ so genau wie möglich von einem Ort an einen anderen zu transferieren. Ob sie tatsächlich eine *Bedeutung* hat oder nicht, spielte für ihn keine Rolle („*These semantic aspects of communication are irrelevant for the engineering problem*“ [3]).

Shannons Definitionen, die auf der Interpretation von Wahrscheinlichkeiten als relativen Häufigkeiten beruhen, besagen zum „Henne und Ei“-Problem

$Wahrscheinlichkeit \rightarrow Information(smaß).$
---

Nun zu unserem zweiten Beispiel der Verwendung von Wahrscheinlichkeiten, der Frage, ob es morgen wohl regnen wird. Offensichtlich macht die Interpretation des Wahrscheinlichkeitsbegriffes als relative Häufigkeit hier keinen Sinn. Man kann die heutige Großwetterlage nicht (mehrfach!) wiederholen. Hier ist die Angabe einer Wahrscheinlichkeit als *Maß für die Überzeugung (Erwartung)*, englisch „*degree of belief*“, anzusehen, dass es morgen regnen wird. Es klingt vielleicht überraschend, dass man mit solch einem vagen Begriff eine quantitative Theorie formulieren

kann. Cox hat 1948 gezeigt, dass jeder Versuch, einer *konsistenten quantitativen Beschreibung von rationalem Schließen* bei unvollständiger Information notwendigerweise zu Regeln führt, die mit den Kolmogorov-Axiomen identisch sind [9]. Die Interpretation

I2: Wahrscheinlichkeiten als „*degree of belief*“

wird *Bayes'scher Standpunkt* genannt, nach Thomas Bayes (1702–1761), dessen Arbeit dazu posthum veröffentlicht wurde. Wie an mehreren Beispielen im Folgenden gezeigt wird, gilt hier

*Information*  $\rightarrow$  *Wahrscheinlichkeit(s-update)*.

Zentrale Bedeutung hat dabei der Begriff der *bedingten Wahrscheinlichkeiten* und das *Bayes'sche Theorem*. Da wir mit der Häufigkeitsinterpretation begonnen haben, wird dieser Begriff im Folgenden auf diese Art eingeführt [9]. Die so motivierte Definition ist aber auch ohne diese Interpretation möglich.

## 2.2 Bedingte Wahrscheinlichkeiten

Wir betrachten das Beispiel einer (bezüglich des folgenden Tests „zufälligen“) Ansammlung aus  $N$  Individuen, in der sich  $N_F$  Frauen befinden. Es wird getestet, wieviele der Anwesenden farbenblind sind. Die Gesamtzahl der Farbenblinden stellt sich als  $N_{fb}$  heraus, davon sind  $N_{F,fb}$  Frauen. Das entspricht mit der Häufigkeitsinterpretation folgenden Wahrscheinlichkeiten:

$$P(F) = \frac{N_F}{N}, \quad P(fb) = \frac{N_{fb}}{N}, \quad P(F, fb) = \frac{N_{F,fb}}{N}.$$

Die „Verbundwahrscheinlichkeit“  $P(F, fb)$  „Frau und farbenblind“ hat in Kolmogoroffs mengentheoretischer Axiomatik die Bedeutung der Wahrscheinlichkeit der Durchschnittsmenge der Menge  $F$  der Frauen und der Menge  $fb$  der Farbenblinden in der Ansammlung, d. h.

$$P(F, fb) = P(F \cap fb).$$

Man kann  $P(F, fb)$  elementar durch „Erweitern“ umformen

$$P(F, fb) = \frac{N_{F,fb}}{N} = \frac{N_{F,fb}}{N_F} \frac{N_F}{N} = \frac{N_{F,fb}}{N_F} P(F) =: P(fb|F)P(F).$$

Dabei haben wir die *bedingte Wahrscheinlichkeit*  $P(fb|F) = N_{F,fb}/N_F$ , „wenn Frau, dann farbenblind“ (im Symbol von rechts nach links gelesen) eingeführt. Diese Größe wird mit dem Wert in medizinischen Statistiken ungefähr übereinstimmen. Alternativ kann man *formal* auch die bedingte Wahrscheinlichkeit  $P(F|fb)$ , „wenn farbenblind, dann Frau“ einführen

$$P(F, fb) = \frac{N_{F,fb}}{N} = \frac{N_{F,fb}}{N_{fb}} \frac{N_{fb}}{N} = \frac{N_{F,fb}}{N_{fb}} P(fb) =: P(F|fb)P(fb).$$

Diese bedingte Wahrscheinlichkeit  $P(F|fb)$  kann man *nicht* irgendwelchen Statistiken entnehmen, da sie offensichtlich von der Zusammensetzung der Ansammlung abhängt. Besteht sie z. B. nur aus Frauen, so gilt  $P(F|fb) = 1$ . Allgemein ergibt sich  $P(F|fb)$  durch Gleichsetzung der beiden Ausdrücke für  $P(F, fb)$  und Division durch  $P(fb)$  als

$$P(F|fb) = \frac{P(fb|F)P(F)}{P(fb)}.$$

Das ist *Bayes' Theorem* (für unseren Fall). Für andere Beispiele (s. u.) muss man nur andere Bezeichnungen verwenden.

Für die elementaren Umformungen, die nötig waren, klingt der Ausdruck „Theorem“ sehr hochgestochen. Bei der Herleitung kann man, wie bereits erwähnt, auf die Frequenzinterpretation verzichten und in beiden Darstellungen für  $P(F, fb)$  jeweils die drei mittleren Identitäten weglassen. Das hat den Vorteil, dass Anwendungen auf Bereiche möglich sind, bei denen die beliebige häufige Wiederholung (wie z. B. beim Würfeln) nicht möglich ist. Das sind Fälle, bei denen fehlende Information die zentrale Rolle spielt und nicht der Zufall. Dabei kann man die Wahrscheinlichkeitstheorie als Erweiterung der Logik auffassen. Wie später diskutiert wird, ist es dann die *Interpretation* obiger Formel, die zu teilweise heftigen Auseinandersetzungen geführt hat [11].

## 3 Beispiele

### 3.1 Urnen

Ein Standardbeispiel zur Veranschaulichung der Wahrscheinlichkeitsrechnung sind Urnen, die verschiedene Kugeln enthalten. An diesem Beispiel sollen beide Interpretationen *I1* und *I2* diskutiert werden. Wir beginnen mit einfachen „direkten“ Beispielen zur Bestimmung von Wahrscheinlichkeiten. Dazu wird Bayes' Theorem nicht benötigt. Es wird aber angegeben, in welchem Sinn die bestimmten Wahrscheinlichkeiten auch als bedingte Wahrscheinlichkeiten aufgefasst werden können, da dies bei dem folgenden „inversen“ Problem Verwendung findet.

Die hier verwendeten „Urnen“ sollen so funktionieren, wie die im Fernsehen zu sehende *Lottomaschine*. Sie besitzt einen „Umrührmechanismus“, der dafür sorgt, dass die anschließend herausfallende Kugel als *völlig zufällig* gewählt betrachtet werden kann. Unsere Urnen sind aber, im Gegensatz zur Lottomaschine, in einen *undurchsichtigen Kasten* eingebaut. Nach Knopfdruck wird erst umgerührt und dann *eine* Kugel ausgespuckt. Wir nehmen an, *zwei* solcher Urnen vorliegen zu haben, mit Aufkleber A und B, sonst aber in ihrem Äußeren völlig identisch.

In Urne A befinden sich 47 weiße ( $w$ ) und 2 rote ( $r$ ) Kugeln. Dann ist die Wahrscheinlichkeit nach Knopfdruck eine weiße oder eine rote Kugeln zu erhalten, leicht durch „Abzählen“ zu finden,

$$p_A^{(0)}(w) = \frac{47}{49}, \quad p_A^{(0)}(r) = \frac{2}{49}.$$

Dabei steht der Index A für die Bezeichnung der Urne und das hochgestellte (0) für die Tatsache, dass noch keine Kugel entnommen wurde. Diese Wahrscheinlichkeiten können auch als die *bedingten Wahrscheinlichkeiten*, wenn Urne A, dann  $w$  oder  $r$  aufgefasst werden

$$p^{(0)}(w|A) = p_A^{(0)}(w) = \frac{47}{49}, \quad p^{(0)}(r|A) = p_A^{(0)}(r) = \frac{2}{49}.$$

Sei nun die (erste) entnommene Kugel eine rote Kugel. In der Urne befinden sich dann noch 48 Kugeln, 47 weiße und eine rote. Nun werde eine zweite Kugel entnommen. Die neuen Wahrscheinlichkeiten sind

$$p^{(r)}(w|A) = p_A^{(r)}(w) = \frac{47}{48}, \quad p^{(r)}(r|A) = p_A^{(r)}(r) = \frac{1}{48}.$$

Dabei deutet das hochgestellte ( $r$ ) an, dass beim ersten Knopfdruck eine rote Kugel erschienen ist. Jetzt kann man z. B. nach der Wahrscheinlichkeit fragen, dass *beide* gezogenen Kugeln rot sind, d. h. die Kugelfolge  $\{r, r\}$  vorliegt. Dazu muss man einfach die beiden oben angegebenen Wahrscheinlichkeiten *multiplizieren*:

$$p_A^{(0)}(\{r, r\}) = p_A^{(r)}(r)p_A^{(0)}(r) = \frac{1}{48} \cdot \frac{2}{49} \approx 0.00085.$$

Die Wahrscheinlichkeit, dass eine dritte gezogene Kugel auch rot ist, ist offensichtlich gleich Null.

In Urne B befinden sich vor der ersten Ziehung 49 rote Kugeln, d. h. *alle* Kugeln sind rot. Für die entsprechenden Wahrscheinlichkeiten gilt also

$$p^{(0)}(w|B) = p_B^{(0)}(w) = 0, \quad p^{(0)}(r|B) = p_B^{(0)}(r) = 1,$$

sowie

$$p^{(r)}(w|B) = p_B^{(r)}(w) = 0, \quad p^{(r)}(r|B) = p_B^{(r)}(r) = 1.$$

Hier ist also alles deterministisch. Bei zweifachem Ziehen aus Urne B wird man z. B. immer zwei rote Kugeln finden

$$p^{(0)}(\{rr\}|B) = 1$$

usw.

Nun werden die Aufkleber A und B an den voll besetzten Urnen entfernt und letztere in unserer Abwesenheit „zufällig“ nebeneinander auf Position 1 und 2 auf einen Tisch gestellt. Dann kann man z. B. fragen, mit welcher Wahrscheinlichkeit

man bei einmaligem Knopfdruck an der beliebig gewählten Urne, z. B. der an Position 1, eine rote Kugel findet. Da für die Wahrscheinlichkeit, Urne A oder B gewählt zu haben,

$$p_1^{(0)}(A) = p_1^{(0)}(B) = \frac{1}{2}$$

gilt, folgt für die Wahrscheinlichkeit, eine rote Kugel zu finden

$$p_1^{(0)}(r) = p^{(0)}(r|A)p_1^{(0)}(A) + p^{(0)}(r|B)p_1^{(0)}(B) = \frac{2}{49} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{51}{98}$$

und  $p_1^{(0)}(w) = 47/98$ . Völlig analog erhält man die Wahrscheinlichkeit, dass zweifaches Ziehen aus der Urne an Platz 1 zwei rote Kugeln liefert als

$$p_1^{(0)}(\{r, r\}) = p^{(0)}(\{r, r\}|A)p_1^{(0)}(A) + p^{(0)}(\{r, r\}|B)p_1^{(0)}(B) = \left(\frac{1}{48} \cdot \frac{2}{49} + 1\right) \cdot \frac{1}{2}.$$

Also ist  $p_1^{(0)}(\{r, r\}) \approx 0.5008$ , d. h. die Wahrscheinlichkeit für diese Kugelfolge nur wenig größer als 50 Prozent. Soweit einfaches, „*direktes*“ Hantieren mit Wahrscheinlichkeiten.

Nun wollen wir ein „*inverses*“ Problem betrachten und zur Lösung *Bayes' Theorem* verwenden. Dabei gehen die gerade berechneten Wahrscheinlichkeiten wesentlich ein.

Die beiden noch voll besetzten Urnen stehen ohne die Aufkleber auf dem Tisch. Sie werden aufgefordert, Urne B möglichst schnell zu identifizieren. Was sollte man tun?

Man überlegt sich zuerst die sogenannten a-priori-Wahrscheinlichkeiten „priors“ [9], wobei hier nur eine Zahlenangabe nötig ist:

$$p_1^{(pr)}(B) = p_2^{(pr)}(A) = 1 - p_2^{(pr)}(B) = 1 - p_1^{(pr)}(A).$$

Ein quantitatives Maß für das Unwissen ist die zugehörige Informationsentropie

$$H^{(pr)} = -p_1^{(pr)}(B) \log_2(p_1^{(pr)}(B)) - p_1^{(pr)}(A) \log_2(p_1^{(pr)}(A)).$$

Wenn man gar kein Vorwissen hat, wird man die beiden Anordnungen der Urnen als gleichwahrscheinlich annehmen, d. h. obige Wahrscheinlichkeiten gleich  $1/2$  setzen, was  $H^{(pr)} = \log_2(2) = 1$  liefert. Um den Informationsstand zu verbessern, wird man anfangen, den Knopf an einer der Urnen, z. B. der ersten, zu drücken.

### Einmal drücken:

Erscheint eine weiße Kugel, was mit Wahrscheinlichkeit  $47/98$  passiert, so ist klar, dass man Urne A gewählt hat. Erscheint dagegen eine rote Kugel, so muss man bei Wahrscheinlichkeiten bleiben.

Bayes' Theorem für die Wahrscheinlichkeit  $p_1^{(0)}(B|r)$ , dass die Urne auf Position 1 die Urne B ist, wenn dort eine rote Kugel erscheint, lautet (die benötigte Wahrscheinlichkeit  $p_1^{(0)}(r)$  haben wir bereits berechnet)

$$p_1^{(0)}(B|r) = \frac{p^{(0)}(r|B)p_1^{(0)}(B)}{p_1^{(0)}(r)} = \frac{1 \cdot \frac{1}{2}}{\frac{51}{98}} = \frac{49}{51} \approx 0.960 \equiv p_1^{(post)}(B).$$

und  $p_1^{(0)}(A|r) = 2/51 \approx 0.040 \equiv p_1^{(post)}(A)$ . Dabei sind  $p_1^{(post)}(B)$  und  $p_1^{(post)}(A)$  die sogenannten *posteriori*-Wahrscheinlichkeiten, dass die Urne auf Position 1 die Urne B oder A ist, wenn dort eine rote Kugel gefunden wurde. Die Wahrscheinlichkeit, dass die Urne auf Position 1 die Urne B ist, ist also durch das Finden der roten Kugel von  $1/2$  auf  $49/51 \approx 0.960$  angewachsen. Die zu diesem Verlauf gehörige Informationsentropie ist

$$H_r^{(post)} = -(49/51) \log_2(49/51) - (2/51) \log_2(2/51) \approx 0.238,$$

also eine deutliche Absenkung verglichen mit  $H^{(pr)} = 1$ . Man kann auch den gewichteten *Mittelwert* der jeweiligen neuen Informationsentropien für die Fälle, eine rote *oder* eine weiße Kugel gefunden zu haben, betrachten

$$\langle H^{(post)} \rangle = p_1^{(0)}(r)H_r^{(post)} + p_1^{(0)}(w)H_w^{(post)} = \frac{51}{98} \cdot H_r^{(post)} \approx 0.123,$$

da beim Ziehen einer weißen Kugel  $p_1^{(post)}(A) = 1$  gilt, und für die resultierende Informationsentropie  $H_w^{(post)} = 0$  folgt.

Falls eine rote Kugel erschienen war, wird man nochmal drücken:

### Zweimal drücken:

Man kann nun entweder wieder bei der Urne auf Position 1 drücken oder aber die andere wählen. Erscheint eine weiße Kugel, so ist alles klar. Ist die zweite Kugel dagegen wieder rot, so muss man nochmals das Bayes-Theorem verwenden. Dazu verwendet man die posteriori-Wahrscheinlichkeiten vom ersten Drücken als *neue priors* ( $pr'$ ) für die zweite Runde

$$p_1^{post} \rightarrow p_1^{pr'}.$$

Für den Fall, dass wieder bei der Urne auf Position 1 gedrückt wird, erhält man

$$p_1^{(r)}(B|r) = \frac{p^{(r)}(r|B)p_1^{(pr')}(B)}{p_1^{(r)}(r)} = \frac{1 \cdot p_1^{(pr')}(B)}{p_1^{(r)}(r)} = \frac{49/51}{p_1^{(r)}(r)}.$$

Also benötigen wir noch die Wahrscheinlichkeit  $p_1^{(r)}(r)$  bei der Urne auf Platz 1 wieder eine rote Kugel zu finden:

$$p_1^{(r)}(r) = p^{(r)}(r|A)p_1^{(pr')}(A) + p^{(r)}(r|B)p_1^{(pr')}(B) = \frac{1}{48} \cdot \frac{2}{51} + 1 \cdot \frac{49}{51} = \frac{49}{51} \cdot \left(1 + \frac{1}{48} \cdot \frac{2}{49}\right).$$



Eingesetzt ergibt sich

$$p_1^{(r)}(B|r) = \frac{1}{1 + \frac{1}{48} \cdot \frac{2}{49}} \approx 0.99915 \equiv p_1^{(post')}(B).$$

Man kann also „fast sicher“ sein, dass die Urne auf Platz 1 tatsächlich Urne B mit den ausschließlich roten Kugeln ist.

Als Test machen wir das „updating“ der Wahrscheinlichkeiten nicht in zwei Schritten, sondern in *einem Schritt* und verwenden Bayes' Theorem in der Form

$$p_1^{(0)}(B|\{r, r\}) = \frac{p^{(0)}(\{r, r\}|B)p_1^{(0)}(B)}{p_1^{(0)}(\{r, r\})} = \frac{1 \cdot \frac{1}{2}}{(\frac{1}{48} \cdot \frac{2}{49} + 1) \cdot \frac{1}{2}} = \frac{1}{1 + \frac{1}{48} \cdot \frac{2}{49}}.$$

Also ergibt sich, wie erwartet, dasselbe Ergebnis.

Es ist interessant, sich auch den Fall anzuschauen, bei dem die zweite Kugel aus der Urne auf Platz 2 entnommen wird:

### Urnenwechsel:

Nachdem aus der Urne auf Platz 1 eine rote Kugel gezogen wurde, ist die Urne auf Position 2 mit hoher Wahrscheinlichkeit  $p_2^{(post)}(A) = p_1^{(post)}(B) = 49/51$  die Urne A. Da die Urne auf Platz 2 noch voll ist, gilt für die Wahrscheinlichkeit dort eine rote Kugel zu erhalten

$$p_2(r) = p^{(0)}(r|A)p_2^{(post)}(A) + p^{(0)}(r|B)p_2^{(post)}(B) = \frac{2}{49} \cdot \frac{49}{51} + 1 \cdot \frac{2}{51} = \frac{4}{51} \approx 0.078.$$

Also findet man mit der komplementären Wahrscheinlichkeit  $47/51 \approx 0.92$  eine weiße Kugel und damit ist in diesem Fall klar, dass die Urne auf Platz 1 die Urne B ist, d. h.  $H_{1r,2w}^{(post')} = 0$ , wobei der Index  $1r, 2w$  angibt, dass erst in Urne 1 eine rote und dann in Urne 2 eine weiße Kugel gefunden wurde.

Für den Fall, dass die Kugel aus der Urne auf Platz 2 auch rot ist, ist man „so klug als wie zuvor“, d. h. vor dem ersten Ziehen. Das ist auch ohne Verwendung des Bayes'schen Theorems klar. Die *Zusatzinformation* hat die Situation relativ zur ersten Ziehung also *verschlechtert* und die Informationsentropie bezüglich der Position von Urne B ist wieder auf 1 *angewachsen*, d. h.  $H_{1r,2r}^{(post')} = 1$ . Zum Glück tritt dies nur in ca. acht Prozent der Zweitziehungen nach dem Auffinden einer roten Kugel bei der ersten Ziehung auf. Die über die beiden Fälle, im zweiten Zug eine weiße *oder* rote Kugel gefunden zu haben, gemittelte Informationsentropie ist gleich

$$\langle H_{12}^{(post')} \rangle = \frac{4}{51} \cdot H_{1r,2r}^{(post')} + \frac{47}{51} \cdot H_{1r,2w}^{(post')} = \frac{4}{51} \approx 0.078$$

und damit *kleiner* als die Entropie der Ausgangslage  $H_r^{(post)} \approx 0.123$  vor dem zweiten Ziehen, nach roter Kugel beim ersten Ziehen. Die Absenkung dieser mittleren

Informationsentropie ist aber deutlich kleiner als im Fall ohne Urnenwechsel (mit  $p_1^{(post')} (B) \approx 0.99915$  folgt  $\langle H_{11}^{(post')} \rangle \approx 0.0098$ ).

Ganz allgemein kann man (unabhängig von unseren speziellen Annahmen über die beiden Urnen) zeigen, dass im Gegensatz zur Entropiezunahme in Einzelfällen (wie bei obigem Beispiel „1r,2r“) die *gemittelte* Informationsentropie bei Zusatzinformation *nicht* zunehmen kann, sondern i. A. *abnimmt*.

### Dreimal drücken:

Hat man zweimal auf Platz 1 gedrückt und zwei rote Kugeln erhalten, so liefert die dritte Kugel aus dieser Urne die endgültige Antwort. Ist sie rot, so haben wir es mit Urne B zu tun, ist sie weiß, mit Urne A.

Gab es rote Kugeln aus *beiden* Urnen, so ist man beim Erscheinen einer weiteren roten Kugel immer noch nicht fertig. Als Fazit dieser Betrachtungen sieht man, dass der Urnenwechsel keine gute Entscheidung ist.

Dieses Beispiel zeigt klar, wie die jeweils *zusätzliche Information* über die gezogene Kugel die Wahrscheinlichkeiten modifiziert:

$$p^{(pr)} + \text{Information} \rightarrow p^{(post)},$$

wobei der Vorgang mehrfach durchlaufen werden kann. Dabei werden die  $p^{(post)}$  des  $m$ -ten Durchlaufs die  $p^{(pr)}$  des  $m + 1$ -ten Durchlaufs.

Wahrscheinlichkeiten als „degree of belief“ hängen vom *Informationsstand* ab.

Ähnlich in der Durchführung, aber näher an der Praxis ist das zweite Beispiel:

## 3.2 Medizinische Tests

Dieses Beispiel macht die Nützlichkeit des Bayes'schen Theorems *für jedermann* deutlich. Mit den überraschenden Ergebnissen sollten sowohl Ärzt(inn)e(n) als auch Patient(inn)en vertraut sein.

Eine Patientin, von der die Ärztin nicht weiß, ob sie eine spezielle Krankheit hat („*k*“ für krank) oder nicht („*g*“ für gesund), lässt einen *Test* über sich ergehen. Solche Tests sind nie fehlerfrei. Bei kranken Patientinnen wird der Test mit einer geringen Wahrscheinlichkeit negativ (*n*) sein, d. h. die Erkrankung wird übersehen. Die entsprechende bedingte Wahrscheinlichkeit sei

$$P(n|k) = \epsilon, \quad P(p|k) = 1 - \epsilon,$$

wobei das positive  $\epsilon$  bei einem guten Test klein verglichen mit eins ist, z. B.  $\epsilon = 0.02$  (zwei Prozent der Erkrankungen werden übersehen). Das Symbol „*p*“ steht hier für positiven Test und *nicht* für Wahrscheinlichkeit. Ebenso passieren „falsch positive“ Fälle

$$P(p|g) = \epsilon', \quad P(n|g) = 1 - \epsilon',$$

wobei das positive  $\epsilon'$  bei einem guten Test ebenfalls klein gegen eins ist, aber durchaus von  $\epsilon$  verschieden sein kann, z. B.  $\epsilon' = 0.01$ .

Die Patientin habe nun ein *positives Testergebnis*. Mit welcher Wahrscheinlichkeit ist sie dann wirklich erkrankt?

Aus dem Urnenbeispiel sollte klar sein, dass man zur Verwendung der Bayes'schen Regel einen *prior* benötigt. Hier sollte man zwei Fälle unterscheiden:

- a) die „Reihenuntersuchung“,
- b) die „gezielte“ Untersuchung,

bei der Resultate von einem *Vortest*, Symptome, die auf das Vorliegen der Krankheit schließen lassen könnten, oder die familiäre Vorgeschichte Grund zur Besorgnis geben.

Bayes' Theorem für die Wahrscheinlichkeit, bei einem positiven Test erkrankt zu sein, lautet

$$P^{(post)}(k) = P(k|p) = \frac{P(p|k)P^{(pr)}(k)}{P(p)} = \frac{(1 - \epsilon)P^{(pr)}(k)}{(1 - \epsilon)P^{(pr)}(k) + \epsilon'(1 - P^{(pr)}(k))}.$$

Für das Folgende spielt es eine zentrale Rolle, welcher der beiden Summanden im Nenner der größere ist. Wenn der (hintere) *falsch positive* Beitrag überwiegt, muss sich unsere Patientin bei einem positivem Befund noch nicht allzu große Sorgen machen. Dies kann besonders dann eintreten, wenn die Krankheit nur *sehr selten* auftritt. Wir betrachten nun die beiden Fälle a) und b):

a) Falls keine Zusatzinformation vorliegt wird man als  $P^{(pr)}(k)$  die Häufigkeitsrate der Krankheit für die entsprechende Altersgruppe verwenden. Leiden z. B. durchschnittlich fünf von tausend Frauen der Altersgruppe an der Krankheit, so folgt  $P^{(pr)}(k) = 5/1000 = 0.005$ . Dann erhält man für die angegebenen Werte von  $\epsilon$  und  $\epsilon'$

$$P^{(post)}(k) = \frac{0.98 \cdot 0.005}{0.98 \cdot 0.005 + 0.01 \cdot 0.995} = \frac{4.9}{4.9 + 9.95} = 0.3299 \dots$$

Bei tausend untersuchten Frauen gibt es in unserem Beispiel im Schnitt fast zehn falsch positive Ergebnisse, ungefähr doppelt soviel wie echt positive. Obwohl  $P^{(post)}(k)$  deutlich größer als  $P^{(pr)}(k)$  ist, stellt sich die Wahrscheinlichkeit wirklich erkrankt zu sein, als „nur“ ungefähr 33 Prozent heraus, also noch kein Grund zu verzweifeln, da es sich eher um einen *falsch positiven* Befund handelt.

Um ein klareres Ergebnis zu erhalten, kann man z. B. den Test wiederholen, wobei das im ersten Test erhaltene Ergebnis von  $P^{(post)}(k)$  nun als  $P^{(pr)}(k)$  dient. Ist der Test wieder positiv, so zeigt die analoge Rechnung nun eine Wahrscheinlichkeit von etwa 0.979, dass man wirklich erkrankt ist. Wem das noch nicht klar genug ist, erhält nach dem dritten Testlauf mit positivem Ausgang eine Wahrscheinlichkeit von etwa 0.9997 erkrankt zu sein, also sehr nah an der Gewissheit.

b) Bei einer „gezielten Untersuchung“ ist es in der Regel nicht einfach, aus der Vorgeschichte einen genauen Wert für  $P^{(pr)}(k)$  anzugeben. Um die Rechnung zu

vereinfachen, nehmen wir als prior auf Grund der Vorinformation  $P^{(pr)}(k) = 1/2$  an. Dann gilt  $P(p) = (1 - \epsilon + \epsilon')/2$  und im Spezialfall  $\epsilon = \epsilon'$  folgt  $P(p) = 1/2$ . Die Wahrscheinlichkeit bei einem positiven Test wirklich krank zu sein, ist dann direkt durch die bedingte Wahrscheinlichkeit gegeben

$$P^{(post)}(k) = P(k|p) = 1 - \epsilon = 0.98.$$

Jetzt muss man sich wirklich Sorgen machen. Ist der Test dagegen *negativ*, so ist die Wahrscheinlichkeit tatsächlich *gesund* zu sein, ebenfalls durch  $1 - \epsilon = 0.98$  gegeben und Freude ist angesagt.

Bei dieser Diskussion zur Anwendung der Bayes'schen Formel in der Medizin wurde zur Vereinfachung angenommen, dass die Tests eine klare Unterscheidung in „positiv“ oder „negativ“ liefern. Dies ist sicher nicht immer der Fall. Bei einer Routineuntersuchung zur Früherkennung von Brustkrebs, wird die Ärztin das Mammographiebild genau analysieren. Dabei gibt es häufig Fälle, bei der eine klare Aussage in „positiv“ oder „negativ“ nicht angebracht ist. Erst das Heranziehen von Zusatzinformation, wie der Vergleich mit dem Mammographiebild des Vorjahrs (falls vorhanden) oder einer weiteren medizinischen Untersuchungsmethode wird die Ärztin zu einer klaren Aussage veranlassen. Dieser Aspekt wird in der Diskussion, wie sinnvoll eine Reihenuntersuchung zur Früherkennung von Brustkrebs ist, oft unterschlagen.

Fazit der dargestellten Überlegungen ist, dass man bei Untersuchungen auf eine seltene Erkrankung, wegen der Häufigkeit von falsch positiven Tests, beim Vorliegen eines positiven Tests auf jeden Fall eine „zweite Meinung“ einholen sollte.

### 3.3 Wissenschaftstheorie

Oft werden konkurrierende Hypothesen zur Erklärung empirischer Fakten vorgeschlagen. Diese Hypothesen sollten natürlich in der Lage sein, diese Fakten zu erklären, und keine Widersprüche in sich enthalten. Wünschenswert ist eine möglichste einfache Hypothese mit großer Vorhersagekraft für einen breiten Anwendungsbereich. Die vorgelegten Hypothesen werden bezüglich dieser Zusatzanforderungen in der Regel nicht äquivalent sein. Die Entscheidung, welche Hypothese sich durchsetzt, fällt oft nach dem Vorliegen neuer empirischer Befunde. Eine Möglichkeit die Wahl zwischen verschiedenen Erklärungsversuchen zu treffen, ist wieder die Verwendung von Bayes' Theorem [12].

An Stelle der Ergebnisse „positiv“ oder „negativ“ im letzten Abschnitt über medizinische Tests, treten hier die empirischen Fakten im folgenden kurz  $e$  genannt. Hier wird nun nicht nach der Wahrscheinlichkeit „gesund“ oder „krank“ zu sein, sondern nach der Glaubwürdigkeit einer speziellen Hypothese  $h$  gefragt. Dann lautet Bayes' Theorem

$$P(h|e) = \frac{P(e|h)P(h)}{P(e)} = \frac{P(e|h)P(h)}{P(e|h)P(h) + P(e|\bar{h})P(\bar{h})},$$

wobei  $\bar{h}$  bedeutet, dass die Hypothese  $h$  falsch ist und  $P(h)$  ist die ursprüngliche Wahrscheinlichkeit, dass die Hypothese  $h$  stimmt. Die Mitglieder der betrachteten

Wissenschaftsgemeinschaft werden sich bezüglich des „priors“ für  $P(h)$  in der Regel nicht einig sein.

Bezüglich des Wechselspiels der „Zusatzaspekte“ gibt es eine Vielzahl von Szenarien. Versucht man z. B. den Anwendungsbereich einer Hypothese zu erweitern, so wird es oft nötig sein, auch die Hypothese zu modifizieren. Liefert die erweiterte Hypothese  $h$  Vorhersagen für neue Phänomene, die dann tatsächlich gefunden werden, oder ist sie in der Lage bis dahin unverstandene Effekte  $e$  zu erklären, so ist  $P(e|h)$  nahe eins, was  $P^{(post)}(h) = P(h|e)$  auch in die Nähe von eins bringt, unabhängig vom schwer einschätzbaren  $P^{(pr)}(h) = P(h)$  in obiger Gleichung. Ein Beispiel dazu ist Maxwells Erweiterung der Gleichungen zur Beschreibung elektrischer und magnetischer Phänomene auf beliebige *zeitabhängige* Ladungs- und Stromverteilungen. Seine „Maxwellschen Gleichungen“ der *Elektrodynamik* sagten die später experimentell gefundene elektromagnetische Strahlung beschleunigt bewegter Ladungen voraus. Dies führte zur schnellen Akzeptanz seiner Gleichungen (auch ohne Verwendung der Bayes'schen Regel). In weniger klaren Fällen kann man sie aber verwenden, um das Vertrauen in eine Hypothese zu quantifizieren [11]. Ein aktuelles Beispiel, für das solch eine Analyse in Zukunft denkbar wäre, ist die Herausforderung einer erfolgreichen Theorie der „Quantengravitation“. Dazu gibt es eine Vielzahl von Versuchen, wobei derzeit zwei sehr verschiedene Zugänge als möglicherweise erfolgsversprechend gelten, die „string theory“ und „loop quantum theory“ [13].

Es gibt allerdings auch heftige Kritiker der Verwendung der Bayes'schen Regel im Kontext der Wissenschaftstheorie [14] („*The reason scientists have to avoid Bayesianism . . .*“).

## 4 Zum Streit der Schulen

Im letzten Jahrhundert gab es einen heftigen Streit zwischen „Frequentisten“ (F) und „Bayesianern“ (B), wobei erstere das stärkere Lager bildeten. Exzellente Lehrbücher zu den beiden Standpunkten stammen von Feller (F) [10] und Jaynes (B), sowie (nach meinem Vortrag erschienen) von der Linden, Dose und Toussaint (B) [10]. Unter meinen Kollegen sind die Fs in der Mehrheit, und wie angedeutet, ist das auch der „übliche“ Standpunkt in den Vorlesungen zur theoretischen Physik. Es gibt aber z. B. auch im Rahmen der Quantenmechanik, bei der Wahrscheinlichkeiten eine zentrale Rolle spielen, eine bayesianische Interpretation der auftretenden Wahrscheinlichkeiten [15]. In der Frage der Interpretation veranschaulicht experimenteller Daten gibt es glühende Befürworter der Bayes'schen Methode unter den Experimentalphysikern [16, 17, 18] und auch so bekannte Theoretiker wie der Nobelpreisträger P. W. Anderson wundern sich, dass es nicht mehr Bayesianer unter den Physikern gibt [19]. Zum Testen von Hypothesen stellen sie die *richtige Frage*, „wie wahrscheinlich eine Hypothese beim Vorliegen der Daten ist“, und nicht die Frage des anderen Lagers „angenommen, die Hypothese stimmt, was ist die Wahrscheinlichkeit der Daten?“.

Dass sich eines der Lager langfristig durchsetzen wird, ist eher unwahrscheinlich.

## Danksagung

Für hilfreiche Anmerkungen und Anregungen zu meinem Manuskript danke ich meinen Kollegen S. Friedrich und R. Schaback.

## Literaturangaben:

- [1] zur Entwicklung des Wahrscheinlichkeitsbegriffs:  
L. Daston, *Classical Probability in The Enlightenment*, Princeton University Press, 1988
- [2] A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, 1933
- [3] C. E. Shannon, *A mathematical theory of communication*, Bell Sys. Tech. J.27 379–423, 623–656, 1948
- [4] ein modernes Lehrbuch dazu ist z. B.: D. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2003
- [5] ein allgemeinverständliche Diskussion dazu findet man z. B. in: H. C. von Baeyer, *Information – The New Language of Science*, Harvard University Press, 2004
- [6] zu meiner Schulzeit spielte die Funktion  $\log_{10}(x)$ , der Logarithmus zur Basis 10, eine wichtige Rolle und man besaß dazu eine *Logarithmentafel*. Für beliebige Basis  $b > 1$  z. B.  $b = 2$  gilt für die mit dem Argument  $x$  monoton anwachsende Logarithmusfunktion
- $$\log_b(xy) = \log_b(x) + \log_b(y), \quad \log_b(b) = 1, \quad x \geq 1 : \log_b(x) \geq 0.$$
- [7] M. Tribus and E. McIrvine, *Energy and Information*, Scientific American 225(3), p. 179, 1971
- [8] P. Janich, *Was ist Information?*, Suhrkamp Verlag, 2006  
Janich argumentiert z. B. überzeugend, dass es eine extreme Einschränkung darstellt, Fragen nur dann als sinnvoll anzusehen, wenn die Menge der möglichen Antworten bereits bekannt ist [7].
- [9] R. T. Cox, *Probability, Frequency and Reasonable Expectation*, Am. J. Physics 14, 1, 1946
- [10] W. Feller, *An Introduction to Probability Theory and its Application*, Vol. I, II, John Wiley and Sons, 1957
- [11] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, 2003 W. von der Linden, V. Dose und U. von Toussaint, *Bayesian Probability Theory: Applications in the Physical Sciences*, Cambridge University Press, 2014
- [12] M. Carrier, *Wissenschaftstheorie – zur Einführung*, Junius Verlag, 2006

- 
- [13] L. Smolin verwendet dieses Beispiel in seiner kritischen Rezension des Buches von R. Dawid: *String Theory and the Scientific Method*, Am. J. Phys. 82, p. 1105, 2014
- [14] M. Bunge, *Evaluating Philosophies*, Springer Science, 2012
- [15] C. A. Fuchs and R. Schack, *Quantum-Bayesian coherence*, Rev. Mod. Phys. 85, p. 1693, 2013
- [16] V. Dose, *Bayesian inference in physics: case studies*, Rep. Prog. Phys. 66, 1421–1461, 2003
- [17] R. Cousins, *Why isn't every physicist a Bayesian?*, Am. J. Phys. 63, 398–410, 1994
- [18] P. Coles, *From Cosmos to Chaos – The Science of Unpredictability*, Oxford University Press, 2006
- [19] P. W. Anderson, *The Reverend Thomas Bayes, Needles in Haystacks, and the Fifth Force*, Physics Today 9–10, January 1992





# Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information

Kurt Schönhammer

Institut für Theoretische Physik, Friedrich-Hund-Platz 1, D-37077 Göttingen

## 1 Einleitung

Die Thermodynamik liefert eine phänomenologische Beschreibung physikalischer Eigenschaften von *makroskopischen* Systemen im *thermischen Gleichgewicht*. In ihr spielen die Begriffe *Energie* und *Entropie* eine zentrale Rolle. Ersterer ist speziell durch die „Energiewende“ Tagesthema, der Entropiebegriff ist weniger bekannt und konzeptionell komplexer. Die Diskussion der Interpretation seiner Erweiterungen wird zum Teil kontrovers geführt.

Ein einfaches Beispiel, an dem die Konzepte der Thermodynamik verdeutlicht werden können, ist ein Gas bei vorgegebener Temperatur, das in ein Gefäß eingeschlossen ist und dessen Volumen mit Hilfe eines beweglichen Stempels verändert werden kann. Solche Systeme spielen bei Berechnung des Wirkungsgrades von Dampfmaschinen, der erstmals von Sadi Carnot (1824) angegeben wurde, eine wichtige Rolle. Das Grundgerüst der Thermodynamik liefern die sog. *Hauptsätze*. Der 1. Hauptsatz ist das zuerst von Julius Robert Mayer (1842) und (unabhängig davon) von Hermann Helmholtz (1847) formulierte Gesetz von der Erhaltung der Energie. Der von Rudolf Clausius 1850 formulierte 2. Hauptsatz verbietet ein *perpetuum mobile 2. Art* (s. Kap.2). Clausius hat diesem Hauptsatz 1865 mit Hilfe des von ihm eingeführten Entropiebegriffs eine abstraktere Form gegeben.

Clausius war auch einer der Begründer der *kinetischen Gastheorie*, in der eine *mikroskopische* Beschreibung im Rahmen der *klassischen Mechanik* versucht wird. Auf Grund der extrem großen Zahl von Atomen (Molekülen) in einem makroskopischen Gasbehälter realisierten insbesondere James Clerk Maxwell und Ludwig Boltzmann, dass statistische Betrachtungen Ergebnisse liefern, die sehr gut mit den Messungen an einem einzelnen System übereinstimmen. Diese Art der Beschreibung bezeichnet man als (klassische) *statistische Mechanik*. Im Rahmen dieses Zugangs fand Boltzmann einen mikroskopischen Ausdruck für die Entropie eines Systems im thermischen Gleichgewicht. Nach der Erkenntnis, dass die Quantenmechanik den korrekten Zugang zum Verständnis des Aufbaus der Mate-

rie liefert, präsentierte John von Neumann 1927 die allgemeine Entropiedefinition im Rahmen der *Quantenstatistik*. Die im Jahr 1948 von Claude Shannon in seiner quantitativen Analyse des Informationsgehalts einer Nachricht vorgeschlagene „Informationsentropie“ hat dieselbe Form wie die von Neumannsche Entropie für physikalische Systeme.

Während bei der Entropiedefinition im Rahmen der Thermodynamik der Bezug zum Begriff „Information“ zunächst schwer erkennbar ist, zeigt die Definition im Rahmen der statistischen Mechanik die enge Beziehung zwischen den Konzepten Entropie und Information, worauf in Kapitel 3 ausführlich eingegangen wird. Da der Entropiebegriff zuerst in der Thermodynamik auftauchte, wird er hier auch in der historischen Reihenfolge vorgestellt.

Da nicht davon ausgegangen wird, dass die in dieser Einleitung vorgestellten Begriffe genauer bekannt sind, werden sie in den folgenden Kapitel erläutert, wobei versucht wird, die mathematischen Ausführungen so einfach wie möglich zu halten. Mathematisch anspruchsvollere Betrachtungen zur statistischen Mechanik werden im Anhang B gegeben. Leider lassen sich solche Betrachtungen auch im Haupttext gelegentlich nicht vermeiden.

## 2 Thermodynamik

In der klassischen Mechanik untersucht man mit Hilfe der Newtonschen Bewegungsgleichungen Systeme von Massenpunkten. Dabei spielen neben Ort, Geschwindigkeit und der Beschleunigung der Teilchen die Begriffe Kraft, Arbeit und Energie eine zentrale Rolle. In mechanischen Systemen treten zwei Formen von Energie auf, die *kinetische* und die *potentielle* Energie. Die kinetische Energie (Bewegungsenergie) eines Massenpunktes der Masse  $m$ , der sich mit einer Geschwindigkeit vom Betrag  $v$  bewegt, ist in der klassischen (nichtrelativistischen) Mechanik durch  $mv^2/2$  gegeben. Die sog. potentielle Energie hängt von der Lage der Massenpunkte im Raum und den Kräften zwischen ihnen ab. Ein einfaches Beispiel ist eine Masse im Schwerfeld der Erde, auf die die Schwerkraft in Richtung zum Erdmittelpunkt wirkt. In der Nähe der Erdoberfläche ist der Stärke der Kraft durch  $mg$  gegeben, wobei  $g$  die *Erdbeschleunigung* ist. Hebt man die Masse um den Betrag  $h$  (Höhe) an, so vergrößert sich die potentielle Energie um die gegen die Schwerkraft geleistete *Arbeit*  $A$ , die durch das Produkt aus *Kraft mal Höhe (Weg)* gegeben ist, d. h.  $A = mgh$ .

Für (vom Rest der Welt) *abgeschlossene* Systeme, auf die keine zeitabhängigen äußeren Kräfte wirken, gilt der Satz von der Erhaltung der (Gesamt-)Energie. Lässt man z. B. die angehobene Masse fallen, so wird die abnehmende potentielle Energie in kinetische Energie umgewandelt. Die Geschwindigkeit der Masse nimmt entsprechend zu.

Die zwei im folgenden wichtigen Begriffe *Temperatur* und *Wärme(menge)* treten in der klassischen Mechanik *nicht* auf.

## 2.1 Thermisches Gleichgewicht, Temperatur

Die Thermodynamik ist eine phänomenologische (makroskopische) Theorie zur Beschreibung der Erscheinungen, in denen der Begriff *Wärme* eine Rolle spielt. Einige der zentralen Konzepte und experimentellen Tatsachen sind jedem aus dem Alltagsleben vertraut:

Bringt man ein warmes und ein kaltes Objekt in Kontakt, so wird das warme Objekt kälter und das kalte wärmer. Nach einer gewissen Zeit ändert sich die Wärmeempfindung nicht mehr, es stellt sich *thermisches Gleichgewicht* ein.

Mit ein bisschen Experimentieren lässt sich auch folgender Sachverhalt zeigen:

Ist Körper A mit Körper B im thermischen Gleichgewicht, und ist ebenso B mit C im thermischen Gleichgewicht, so sind auch A und C im thermischen Gleichgewicht. Diese Erfahrungstatsache bezeichnet man als den *0. Hauptsatz der Thermodynamik*.

Ein wichtiger Schritt der *Wärmelehre* zu einer Wissenschaft bestand darin, ein *quantitatives Maß* für die Begriffe warm und kalt zu finden. Die Grundidee ist, dass es *mechanische Eigenschaften* von Körpern gibt, die sich mit Erwärmen (bzw. Abkühlen) *monoton ändern*, wie die Ausdehnung vieler Flüssigkeiten. Mit einem *Thermometer* kann man dann Körpern, die sich untereinander im thermischen Gleichgewicht befinden, eine Maßzahl zuordnen, ihre *Temperatur*. Um die Temperaturskala zu eichen, benützt man wohldefinierte Systeme um *Fixpunkte* zu setzen. Auf der Celsiusskala verwendet man Wasser bei „Normaldruck“ und setzt die Temperatur des Gefrierpunktes auf 0 Grad ( $T_{Celsius}^G = 0$ ) und die des Siedepunktes auf 100 Grad ( $T_{Celsius}^S = 100$ ). In der Thermodynamik verwendet man die auf Kelvin zurückgehende *absolute Temperatur*  $T$

$$T \equiv T_{Celsius} + 273.16. \quad (1)$$

Der Eckpfeiler der Thermodynamik ist die *Annahme*, dass makroskopische Systeme im thermischen Gleichgewicht durch Angabe von *wenigen Variablen* beschrieben werden können. Für ein Gas aus gleichen Atomen hängt z. B. der Druck  $p$ , den es auf die Wände des Behälters ausübt, nur von der Temperatur  $T$  und der Teilchendichte  $n = N/V$  ab, wobei  $N$  die Zahl der Atome und  $V$  das Volumen des Gases ist

$$p = p(T, n). \quad (2)$$

Die Funktion  $p(T, n)$  lässt sich für jede vorgegebene Gassorte *experimentell* bestimmen. Als besonders einfach erweist sich diese Funktion für stark verdünnte Gase, sog. *ideale Gase*

$$p_{id} = nk_B T = \frac{N}{V} k_B T. \quad (3)$$

Dabei ist die experimentell gemessene „Boltzmannkonstante“  $k_B$  unabhängig von der Gassorte. Eine Vorschrift zu *Berechnung* von  $p(T, n)$  für beliebige Gase liefert die Thermodynamik als phänomenologische Theorie nicht.

## 2.2 Wärme(menge)

Zum Begriff der Wärmemenge zunächst einige historische Anmerkungen. Noch Carnot glaubte an einen „Wärmestoff“ (Caloricum), der beim Temperatenausgleich vom wärmeren zum kälteren Körper übergeht. Das Caloricum stellte man sich als masselose Flüssigkeit vor. Erst Mitte des 19. Jahrhunderts setzte sich die Einsicht durch, dass die übertragenen Wärmemengen eine spezielle Form von *Energie* sind, und die Gesamtenergie bei allen Prozessen eine Erhaltungsgröße ist.

Zur quantitativen Definition der auf ein System übertragenen Wärmemenge betrachten wir ein thermisch isoliertes Gas, auf dessen beweglichen, aber festgeklemmten Stempel sich ein Körper der Masse  $M$  befindet (s. Abb. 1).

Da das System „abgeschlossen“ ist, hat es in diesem Anfangszustand, wie aus der Mechanik bekannt, eine Gesamtenergie  $E_{Anfang}$ . Lässt man nun das System für eine gewisse Zeit mit anderen Systemen in Wechselwirkung treten (z. B. Lösen des Stempels und Wärmekontakt mit einem System anderer Temperatur), isoliert es anschließend wieder und klemmt den Stempel fest, dann hat das System im Endzustand die Energie  $E_{Ende}$ . Die Energiedifferenz  $\Delta E = E_{Ende} - E_{Anfang}$  wird aufgeteilt in die *am* System geleistete *Arbeit*  $A$  und die *auf das System übertragene Wärmemenge*  $Q$

$$\Delta E = Q + A. \quad (4)$$

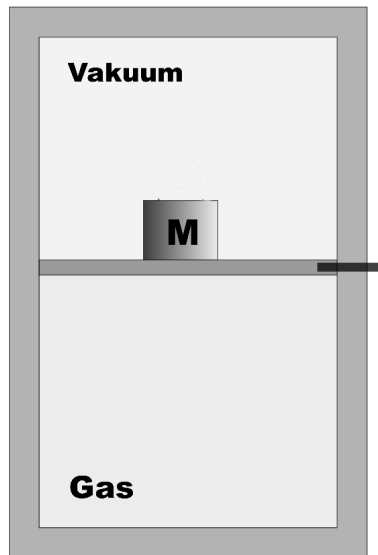


Abbildung 1. Thermisch isoliertes Gas mit festgeklemmtem Stempel

Dabei ist die am System geleistete Arbeit *makroskopisch* z. B. „mechanisch“ definiert. Zur jeweiligen Charakterisierung von  $A$  muss man den abgelaufenen Prozess betrachten. Ist z. B. der Stempel im Endzustand um die Distanz  $a$  nach unten

gedrückt, so gilt  $A = Mga$ . Eine Arbeit kann auch „elektrischer“ Form sein, z. B. durch das Aufladen eines Kondensators, oder „chemischer“ Form, z. B. durch das Aufladen einer Batterie.

Der *1. Hauptsatz der Thermodynamik* ist der Satz von der Erhaltung der Gesamtenergie für makroskopische Systeme. Die Aufspaltung der Energieänderung  $\Delta E = Q + A$ , d. h. in die auf das System übertragene Wärmemenge und die am System geleistete Arbeit, ist der Ausgangspunkt für die weiteren Begriffsbildungen der Thermodynamik. Dazu ist es nützlich, Prozesse mit sehr kleinen („infinitesimalen“) Änderungen zu betrachten:  $\delta E = \delta Q + \delta A$ . Zur Kennzeichnung verwendet man ein *kleines* griechisches Delta  $\delta$ .

War das System im Anfangszustand im thermodynamischen Gleichgewicht und wartet man nach dem Prozessende bis sich im System wieder thermisches Gleichgewicht eingestellt hat, so kann man Anfangs- und Endzustand im Rahmen der Thermodynamik beschreiben. Die Gesamtenergie des einkomponentigen Systems (alle Gasatome sind von der gleichen Sorte) hängt bei fester Teilchenzahl nur von der Temperatur  $T$  und dem Volumen  $V$  ab, d. h.  $E = E(T, V)$ .

Für den *allgemeinen Prozess*, den wir beschrieben haben, lassen sich über die Zwischenzustände im Rahmen der (Gleichgewichts-) Thermodynamik in der Regel keine weiteren Aussagen machen. Es ist daher nützlich, sogenannte *reversible* Prozesse zu betrachten. Das sind Prozesse, die so hinreichend langsam ablaufen („quasistatisch“), dass das System eine *Folge von Gleichgewichtszuständen* durchläuft, wobei der Prozess auch in *umgekehrter* Reihenfolge ablaufen kann. Solche Prozesse stellen eine nützliche Idealisierung dar, da man den gesamten Prozess (mit all seinen Zwischenzuständen) im Rahmen der Thermodynamik beschreiben kann. Insbesondere kann die am (vom) System geleistete Arbeit durch Eigenschaften des Systems ausgedrückt werden. Das sieht man leicht am gewählten Beispiel. Das System ist im mechanischen Gleichgewicht mit dem Gewicht auf dem Stempel, wenn Druck und Gegendruck gleich sind:  $p = Mg/F$ , wobei  $F$  die Fläche des Stempels ist. Bei einer infinitesimalen Verschiebung des Stempels um  $dh$  nach oben, erhöht sich die potentielle Energie des Gewichts um  $Mgdh = pFdh = pdV$ , wobei  $dV$  die Volumenänderung ist. Die *am* System geleistete Arbeit  $\delta A$  ist also

$$\delta A = -pdV. \quad (5)$$

Bei Volumenabnahme, also negativem  $dV$ , verliert die Masse  $M$  an potentieller Energie. Dieser Verlust erscheint als Energiegewinn im System. Für die „differenzielle“ Energieänderung  $dE = E(T, V + dV) - E(T, V)$  bei dem reversiblen Prozess zwischen „benachbarten“ Gleichgewichtszuständen gilt dann

$$dE = \delta Q_{rev} - pdV. \quad (6)$$

In diese Relation gehen bei fester Teilchenzahl die thermodynamischen *Zustandsfunktionen*  $E(T, V)$  und  $p(T, V)$  ein. Die Größe  $\delta Q$  ist dagegen (noch) nicht durch eine Zustandsfunktion des Systems ausgedrückt. Das geschieht mit Hilfe des Entropiekonzepts.

## 2.3 Entropie und der 2. Hauptsatz der Thermodynamik

Die Zustandsfunktion *Entropie*  $S$  wird für abgeschlossene Systeme im thermischen Gleichgewicht eingeführt. Das als Beispiel dienende einkomponentige System wird im thermischen Gleichgewicht durch die Temperatur  $T$ , das Volumen  $V$  und die Teilchenzahl  $N$  charakterisiert. Man *definiert* nach R. Clausius die Zustandsfunktion Entropie  $S(T, V, N)$ , indem man angibt, wie sie sich bei reversibler Wärmezuführung *ändert*:

Führt man dem System, das sich bei der Temperatur  $T$  im thermischen Gleichgewicht befindet, *reversibel* die (kleine) Wärmemenge  $\delta Q_{rev}$  zu, so erhält man für die Änderung („Differential“)  $dS$  der Entropie

$$dS = \frac{\delta Q_{rev}}{T} . \quad (7)$$

(Für mathematisch Vorgebildete: Die Temperatur wirkt als „integrierender Nenner“.)

Mit der Entropiedefinition nimmt der 1. Hauptsatz für infinitesimale reversible Prozesse bei fester Teilchenzahl (Gl. (6)) die für die Thermodynamik zentrale Form

$$dE = TdS - pdV \quad (8)$$

an. Die Änderung der Zustandsfunktion  $S$  bei Änderung von Energie und Volumen erhält man durch einfache Umstellung

$$dS = \frac{1}{T}dE + \frac{p}{T}dV. \quad (9)$$

Weiterhin definiert man die Entropie eines *Gesamtsystems* aus isolierten Teilsystemen, die für sich im thermischen Gleichgewicht sind (möglicherweise bei *verschiedenen* Temperaturen), als *Summe* der Teilentropien. Diese Definition ist wesentliche Voraussetzung für die „abstrakte“ Formulierung des 2. Hauptsatzes der Thermodynamik mit Hilfe des Entropiekonzepts (s. u.). Auch das in Kapitel 3 erläuterte Auffinden eines mikroskopischen Ausdrucks für die thermodynamische Entropie hat wesentlich mit ihrer Additivität zu tun.

Nun zum 2. Hauptsatz der Thermodynamik, für den es verschiedene Formulierungen gibt. Hier drei Versionen mit zunehmendem Abstraktionsgrad [1]:

- A) Bringt man einen warmen und einen kalten Körper in thermischen Kontakt, so wird der warme kälter und der kalte wärmer. Der umgekehrte Vorgang, bei dem der wärmere noch wärmer und der kältere noch kälter wird, findet *spontan* (d. h. ohne anderweitige Änderungen) nicht statt.
- B) Es ist nicht möglich, eine periodisch arbeitende Maschine zu konstruieren, bei der nach einem Umlauf die *einzigsten* Änderungen in der umgebenden Welt darin bestehen, dass Arbeit geleistet wurde und nur *ein* Wärmereservoir abgekühlt wurde.

Vor der dritten Version einige Anmerkungen:

Die Form A) kennt jeder („der Kaffee wird kalt“). In Version B) verbietet der 2. Hauptsatz zum Beispiel, dass ein Ozeandampfer die zum Betrieb seiner Maschine erforderliche Energie aus dem Wärmehalt des Ozeans entnimmt, auf dem er fährt („*Perpetuum mobile 2. Art*“). Man benötigt, wie bei einer Dampfmaschine, mindestens *zwei* Wärmereservoirs mit *verschiedenen* Temperaturen, um eine Maschine zu bauen, die *zyklisch* Arbeit leistet.

- C) Die *Entropie*  $S$  eines abgeschlossenen Systems kann niemals abnehmen. Sie *nimmt zu*,

$$S_{\text{End}} > S_{\text{Anf}} \quad (10)$$

bei allen natürlichen, mit *endlicher* Geschwindigkeit verlaufenden Prozessen.

Ein einfaches Beispiel für die Äquivalenz der Formulierungen A) und C) ist ein abgeschlossenes System aus zwei großen thermisch isolierten Systemen mit *verschiedenen* Temperaturen, die kurz in thermischen Kontakt gebracht und dann wieder isoliert werden. Beim Kontakt wird nach Version A) eine (kleine) *positive* Wärmemenge  $\Delta Q$  auf das kältere System mit der Temperatur  $T_k$  übertragen. Der Wärmeaustausch soll dabei so gering sein, dass man die damit verbundene kleine Temperaturänderung der Systeme vernachlässigen kann. Nach diesem *irreversiblen* Prozess stellt sich in den getrennten Teilsystemen wieder thermisches Gleichgewicht ein. Obwohl der gesamte Prozess irreversibel ist, kann man Gl. (9) verwenden, um die Entropiedifferenz der „benachbarten“ Gleichgewichtszustände, vor dem Wärmeaustausch und nachdem sich nach dem Austausch wieder Gleichgewicht eingestellt hat, zu berechnen. Da sich das Volumen nicht ändert, erhält man die positive Entropieänderung des kälteren Systems  $\Delta S_k = \Delta E_k/T_k = \Delta Q/T_k$ . Da das Gesamtsystem abgeschlossen ist, *verliert* das wärmere System mit der Temperatur  $T_w$  wegen der Erhaltung der Gesamtenergie die Wärmemenge  $\Delta Q$ . Es gilt also  $\Delta S_w = \Delta E_w/T = -\Delta Q/T_w$ . Für die Änderung der Gesamtentropie  $S_G$  gilt damit

$$\Delta S_G = \Delta S_k + \Delta S_w = \frac{\Delta Q}{T_k} - \frac{\Delta Q}{T_w} = \frac{\Delta Q}{T_k T_w} (T_w - T_k) > 0. \quad (11)$$

Nimmt man umgekehrt C), d. h.  $\Delta S_G > 0$  an, so folgt aus Gl. (11), dass  $\Delta Q$  für  $T_w > T_k$  positiv ist, d. h. der Wärmeübertrag erfolgt vom warmen zum kalten System, also Version A).

Die Äquivalenz der drei Formulierungen des zweiten Hauptsatzes wird im Anhang A gezeigt. Wir betrachten im Folgenden die „abstrakte“ Version C).

Ein weiteres einfaches Beispiel für das Anwachsen der Entropie liefert die *plötzliche Vergrößerung* des einem Gas zur Verfügung stehenden Volumens. Besonders einfach wird die Beschreibung für *ideale Gase*. Experimentell findet man (Gay-Lussac, Joule), dass sich nach hinreichend langer Zeit wieder thermisches Gleichgewicht einstellt und sich die Temperatur bei diesem Prozess *nicht* ändert. Also hängt die Energie eines idealen Gases nicht vom Volumen ab:  $E_{id}(T, V) = E_{id}(T)$ , wobei die Abhängigkeit von der festen Teilchenzahl unterdrückt ist. Damit lautet

Gl. (9) unter Verwendung des Resultates für den Druck eines idealen Gases Gl. (3)

$$dS_{id} = \frac{dE_{id}(T)}{T} + Nk_B \frac{dV}{V} = \frac{dE_{id}(T)}{T} + Nk_B d \log V. \quad (12)$$

Dabei ist „log“ der natürliche Logarithmus und es wurde verwendet, dass für die Ableitung des Logarithmus  $d \log V/dV = 1/V$  gilt. Die Funktion  $E_{id}(T)$  ist für *atomare* Gase besonders einfach. Die Messung der Wärmekapazität zeigt, dass  $E_{id}(T)$  proportional zur Temperatur  $T$  ist. Es gilt  $E_{id,at}(T) = (3/2)Nk_B T$ . In diesem Fall liefert die Integration von Gl.(12) für die Entropie als Funktion von  $T$  und  $V$

$$S_{id,at}(T, V) = S_{id,at}(T_0, V_0) + Nk_B \left[ \frac{3}{2} \log(T/T_0) + \log(V/V_0) \right]. \quad (13)$$

Für molekulare ideale Gase ist der rein temperaturabhängige Term komplizierter.

Da sich beim beschriebenen Expansionsversuch die Temperatur nicht ändert, trägt nur der zweite Term in der Klammer zur Entropieänderung bei und man erhält

$$\Delta S = S_{\text{End}} - S_{\text{Anf}} = Nk_B \log(V_{\text{End}}/V_{\text{Anf}}) > 0 \quad (14)$$

für  $V_{\text{End}} > V_{\text{Anf}}$ .

Bis hierher tauchte im Kontext der Thermodynamik der Begriff „Information“ nicht auf. Zur Zeit der Entstehung der Thermodynamik im 19. Jahrhundert, die unabhängig vom Verständnis des mikroskopischen Aufbaus der Materie erfolgte, wurde der Informationsbegriff nicht verwendet. Betrachtet man die Definition der Entropie über ihre Änderung in Gleichung (7), so stellt das Auftreten der Wärmemenge  $\delta Q$  eine Andeutung dazu dar, da Wärme *ungeordnete* Bewegungsenergie ist, über die wir im Detail nichts wissen. Die Erkenntnis, dass das Anwachsen der Entropie mit zunehmender Unordnung verbunden ist, gipfelte in der von R. Clausius und anderen formulierten Hypothese vom „Wärmetod“ des Weltalls.

Bei der Entropieänderung im besprochenen Expansionsversuch kann man sehen, dass „Entropie“ etwas mit „Information“ zu tun hat, da man vor der Expansion genauere Information darüber hatte, *wo* sich die Teilchen befinden. Der Zusammenhang mit der Information über das System wird bei einer kleinen Modifikation des Experiments noch klarer:

Wir betrachten nun zwei *verschiedene* Sorten von idealen Gasen. Im Anfangszustand (vgl. Abb. 2) befinden sich die „*roten*“ Atome in der linken Hälfte des Gefäßes, die „*grünen*“ in der rechten Hälfte, wobei die Anzahl  $N_R$  der roten Atome gleich der Anzahl  $N_G$  der grünen Atome ist.

Nun werde die Zwischenwand herausgezogen (vgl. Abb. 3), wodurch die Durchmischung der Gase beginnt. Nach hinreichend langer Zeit ist die mittlere Dichte beider Gassorten im gesamten Gefäß gleich.



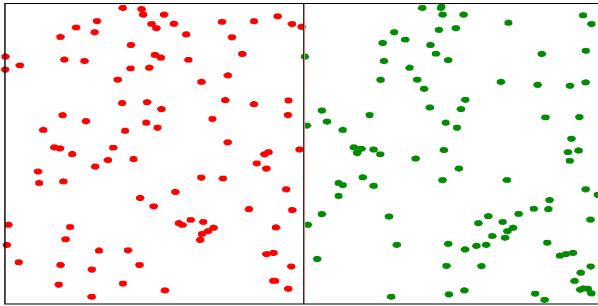


Abbildung 2. Zwei durch eine Trennwand getrennte Gase

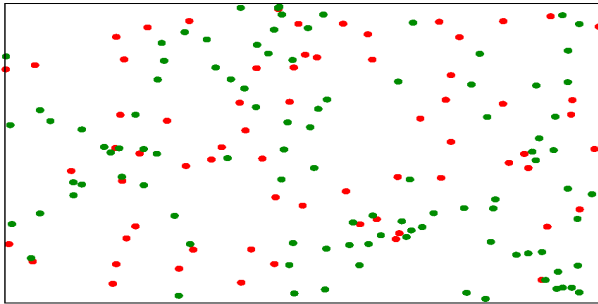


Abbildung 3. Zwei Gase nach Entfernen der Trennwand

Für diesen neuen Gleichgewichtszustand kann man jetzt wieder mit Hilfe von Gl. (14) die Änderung der *Gesamtentropie* berechnen. Man erhält für die *Mischungsentropie* für  $N_R = N_G = N$

$$S_{\text{End}} - S_{\text{Anf}} = 2k_B N \log 2,$$

da sich das Volumen für beide Gassorten verdoppelt hat.

Abgesehen von der fehlenden Trennwand unterscheidet sich für einen *Farbenblinden* der Endzustand *nicht* vom Anfangszustand und er würde behaupten, dass sich die Entropie *nicht* geändert hat.

Für den Experimentator kommt es darauf an, ob er eine Methode hat, die Atome zu unterscheiden, und er z. B. durch eine *semipermeable Wand* in der Lage ist, sie wieder zu trennen [1]. Das Gedankenspiel, die beiden Atomsorten immer ähnlicher und schließlich gleich zu machen, hat zu vielen Diskussionen geführt und wird als „*Gibbs'sches Paradoxon*“ bezeichnet.

Hier taucht zum ersten Mal die *Beobachtungsebene* des Experimentators in den Überlegungen auf, die im Kapitel „Statistische Mechanik“ eine wichtige Rolle spielt.

Wie bereits erwähnt, ist der Name *Thermodynamik* irreführend, da sie ohne Zusatzannahmen *keine* Aussagen über den *zeitlichen Verlauf* von Prozessen liefert. Die Bezeichnung *Thermostatik* wäre also angebrachter. Es gibt aber *Erweiterungen* des Rahmens der „Thermostatik“, bei denen solche zeitlichen Aussagen möglich sind, die sog. *irreversible Thermodynamik* [1]. Dabei beschreibt man das System als bestehend aus „kleinen“, aber doch noch „makroskopischen“ Teilsystemen, in denen *lokales* thermisches Gleichgewicht herrscht, so dass man die Teilsysteme im Rahmen der Thermodynamik beschreiben kann.

In Kap. 3 wird erläutert, wie man allgemein Entropien definieren kann, deren zeitlicher Verlauf unter gewissen Annahmen berechnet werden kann. Es sollte angemerkt werden, dass in der Literatur häufig über das Verhalten „der Entropie“ außerhalb des Bereiches der Gleichgewichtsthermodynamik argumentiert wird, ohne überhaupt eine präzise Definition dieser „Entropie“ angegeben zu haben.

Dem Erfolg der phänomenologischen Thermodynamik liegt offensichtlich folgender Sachverhalt zu Grunde:

Obwohl makroskopisch identisch präparierte Systeme in ihrem mikroskopischen Zustand (Orte und Geschwindigkeiten der Teilchen in einer Beschreibung im Rahmen der klassischen Mechanik) verschieden sind, spielt dieser Unterschied bei der Messung *makroskopischer* Eigenschaften innerhalb der Messgenauigkeit keine Rolle. Also ist es möglich mit makroskopischen Systemen reproduzierbare Experimente zu machen, d. h.

*(makroskopische) Physik ist möglich.*

Die Thermodynamik liefert *Beziehungen zwischen messbaren Größen*. Die *Berechnung* solcher Größen für mikroskopische Modelle der Materie ist Aufgabe der *Statistischen Mechanik*.

## 3 Statistische Mechanik

Die Statistische Mechanik (SM) liefert die *mikroskopische Begründung* der Thermodynamik und deren Erweiterungen. Eine zentrale Rolle spielt wieder der *Entropiebegriff*. Der Zusammenhang mit dem Begriff der *Information* kann nun genauer untersucht werden. Eine kurze mathematischere Darstellung der Grundideen der statistischen Mechanik findet sich in Anhang B.

### 3.1 Ludwig Boltzmann und der Entropiebegriff

Das historisch erste Beispiel, die Gesetze der Thermodynamik über den atomaren Aufbau der Materie zu verstehen, ist die hauptsächlich von J. C. Maxwell und L. Boltzmann entwickelte *kinetische Gastheorie*. In der zweiten Hälfte des 19. Jahrhunderts stellten die Newtonschen Gleichungen die Basis der mikroskopischen Beschreibung dar (klassische Mechanik). Im Sinne von „*Information ist das, was man benötigt, um besser einschätzen zu können wie es weitergeht*“ hat man die

*volle Information* über den Zustand des Systems, wenn man die Orte und Geschwindigkeiten aller Teilchen zu einem gegebenen Zeitpunkt kennt. Die zukünftige Zeitentwicklung ist dann *vollständig* festgelegt. Der sich ergebende Determinismus, auf den insbesondere Laplace hingewiesen hat, hat aber bereits in der klassischen Mechanik seine praktischen Grenzen, da die Dynamik von Vielteilchensystemen üblicherweise *sehr sensitiv* von den Anfangsbedingungen abhängt. Beliebig kleine Änderungen der Anfangsorte oder Anfangsgeschwindigkeiten führen nach einer davon abhängigen späteren Zeit zu völlig unterschiedlichem Systemverhalten („Chaostheorie“).

Weder experimentell noch theoretisch kann man  $\sim 10^{23}$  Teilchen *individuell* handhaben. Der Ausweg, den Maxwell und Boltzmann gewählt haben, ist eine statistische Beschreibung. Im Gegensatz zur Thermodynamik betrachtet man in der statistischen Mechanik die *mikroskopischen* Details des untersuchten Systems.

Wir beginnen mit einer einfachen Überlegung. Der Kasten mit Trennwand von Abbildung 2 sei zunächst leer und vor ihm liegen  $N$  kleine identische Kugeln, die von 1 bis  $N$  nummeriert sind. Mit Hilfe einer Münze wird der Kasten nun wie folgt gefüllt: Erhält man beim Münzwurf „Kopf“, so wird Kugel 1 durch ein Loch in die linke Hälfte des Kastens gelegt, bei „Zahl“ in die rechte. Entsprechend behandelt man alle anderen Kugeln. Falls  $N$  sehr groß ist, werden am Ende beide Hälften ungefähr gleich voll sein, da Kopf und Zahl ungefähr gleich oft gefallen sind. Da es bei jedem Münzwurf zwei Möglichkeiten gibt, können sich insgesamt  $2^N$  verschiedene Konfigurationen für die Füllung des Kastens ergeben. Für „mikroskopische“ Kugeln werden sie im folgenden *Mikrokonfigurationen* genannt. Man kann nun z. B. die Frage stellen, wieviele Mikrokonfigurationen es gibt, bei denen sich in der linken Hälfte des Kastens  $M$  Kugeln befinden (und  $N - M$  Kugeln in der rechten Hälfte). Für  $M = 6$  und  $N = 49$  ist das die Frage, wieviel verschiedene Ergebnisse bei einer Lottoziehung möglich sind. Dieses kombinatorische Problem wird durch die sog. *Binomialkoeffizienten* [2] gelöst. Für gerade  $N$  ergibt sich die größte Zahl der Mikrokonfigurationen für  $M = N/2$ , d. h. gleicher Füllung beider Hälften. Da bei einer fairen Münze alle Mikrokonfigurationen gleich wahrscheinlich sind, hat der Zustand gleicher Füllung die größte Wahrscheinlichkeit.

Will man ein ideales Gas bei endlicher Temperatur im Rahmen der klassischen Physik beschreiben, so kann man die kleinen Kugeln als Modell für die Atome benutzen. Boltzmann schreibt dann zum weiteren Vorgehen, „dass man den Zustand des Wärmegleichgewichts dadurch berechnen kann, dass man die Wahrscheinlichkeit der verschiedenen möglichen Zustände des Systems aufsucht“ [3]. Man muss auch die Geschwindigkeiten der Atome betrachten, die *kontinuierliche* Werte annehmen, so dass ein „Abzählen“ der Mikrokonfigurationen der Geschwindigkeiten ein Problem darstellt. Boltzmann umgeht in der zitierten Arbeit diese Schwierigkeit indem er als (näherungsweise) mathematischen Trick annimmt, dass jedes Molekül nur „eine bestimmte endliche Anzahl von Geschwindigkeiten“ annehmen kann [3]. Er arbeitet dann aber mit der Energie („lebendige Kraft“) und schreibt:

„Zu noch größerer Vereinfachung nehmen wir an, dass die Reihe der lebendigen Kräfte, welche jedes Molekül anzunehmen imstande ist, eine arithmetische

*Progression bildet, z. B. folgende“*

$$0, \epsilon, 2\epsilon, 3\epsilon, \dots, n\epsilon.$$

Was Boltzmann hier aus Gründen der mathematischen Vereinfachung einführt, sieht (bis auf den Übergang  $n \rightarrow \infty$ ) wie Plancks „Akt der Verzweiflung“ zur Beschreibung der Strahlung eines „Schwarzen Körpers“ aus, der dreiundzwanzig Jahre später die Geburtsstunde der Quantentheorie war [1]. Es ist daher nicht verwunderlich, dass die statistische Beschreibung von Systemen im thermischen Gleichgewicht im Rahmen der Quantenmechanik konzeptionell einfacher ist als im Rahmen der klassischen Physik. Dann sind die Gesamtenergien  $E_i$  für Systeme in endlichen Kästen *quantisiert* und man kann die entsprechenden *Zustände*  $|E_i\rangle$  (im Hilbertraum) mit einer natürlichen Zahl  $i$  durchnummerieren.

Für ein einkomponentiges System, wie das Gas im Kasten, benötigt man, wie in Kapitel 2 diskutiert, neben der Teilchendichte nur die *Gesamtenergie*  $E$  als mechanische Größe, um die Gesetze der Thermodynamik zu begründen. Also versucht man in der statistischen Beschreibung solcher Systeme im thermischen Gleichgewicht einen analogen Ansatz. Man nimmt an, dass die Gesamtenergie  $E$  (für festes Volumen und Teilchenzahl) bis auf eine Ungenauigkeit  $\Delta E$  bekannt ist. Die Anzahl  $W$  der für vorgegebene Werte von  $E$  und  $\Delta E$  möglichen Mikrokonfigurationen (Energieeigenzustände) ist für makroskopische System „gigantisch“. Je größer  $W$  ist, umso größer ist die *fehlende Information* über den mikroskopischen Zustand des Systems.

Die Gesamtzahl  $W_{\text{tot}}$  der Mikrokonfigurationen für ein Gesamtsystem aus *zwei* unabhängigen Teilsystemen 1 und 2 mit  $W_1$  und  $W_2$  Mikrokonfigurationen ist durch das *Produkt* dieser Zahlen

$$W_{\text{tot}} = W_1 \cdot W_2 \quad (15)$$

gegeben. Also ist der *Logarithmus* der Zahl der Mikrozustände *additiv*

$$\log W_{\text{tot}} = \log W_1 + \log W_2. \quad (16)$$

In Kapitel 2 wurde darauf hingewiesen, dass die Additivität der Entropie eines Gesamtsystems aus isolierten Systemen im thermischen Gleichgewicht ein wesentlicher Aspekt ihrer Definition ist. Boltzmann erkannte als Erster, dass die Entropie, bis auf einen Faktor, durch den Logarithmus der Zahl der Mikrokonfigurationen gegeben ist [4]. Diese bahnbrechende Einsicht ist auf Boltzmanns Grabstein verewigt ( $\ln \equiv \log, k \equiv k_B$ )

$$S = k \ln W \quad (17)$$

In der statistischen Beschreibung des Gleichgewichtszustands des Systems im Rahmen der Interpretation von Wahrscheinlichkeiten als „relative Häufigkeiten“ [5] stellt man sich (als Gedankenexperiment) ein *Ensemble* von gleichartigen Systemen vor, die alle auf dieselbe Weise präpariert wurden [6]. Wegen der experimentellen Energieunschärfe kann man keinen wohldefinierten Energiezustand präparieren. Man nimmt an, dass der  $i$ -te Zustand mit der *Wahrscheinlichkeit*  $p_i$  auftritt. Gibt

man nun  $E$  und die Unschärfe  $\Delta E$  vor, so ist die Frage, wie man die  $p_i$  in dieser „Energieschale“ mit  $W$  möglichen Zuständen wählen soll. Da wir nichts zusätzlich wissen, sollte die Wahl *so unvoreingenommen wie möglich* sein. Also nimmt man die Zustände als *gleichwahrscheinlich* an, d. h.  $p_i = 1/W$ . Man nennt das entsprechende Ensemble das *mikrokanonische Ensemble*.

### 3.2 Der allgemeine Entropiebegriff

Insbesondere durch die Arbeiten von J. Willard Gibbs (klassische SM) [6] und John von Neumann (Quanten-SM) [7] erkannte man, dass Boltzmanns Ausdruck für die Entropie ein Spezialfall des folgenden *allgemeinen* Ausdrucks ist

$$S = -k_B \sum_i p_i \log p_i \quad (= -k_B S p(\rho \log \rho)). \quad (18)$$

Dabei sind für den Fall des thermischen Gleichgewichts die  $p_i \geq 0$  die Wahrscheinlichkeiten, dass der Zustand mit Energie  $E_i$  vorliegt, wobei die Summe aller Wahrscheinlichkeiten gleich eins ist, d. h.  $\sum_i p_i = 1$  gilt. Da obige Definition der Entropie mit der Summe dieselbe Form hat, wie die später von C. Shannon im Zusammenhang mit dem Informationsgehalt einer Nachricht vorgeschlagene Größe [8], nennt man die in Gleichung (18) definierte Entropie heutzutage auch *Informationsentropie*. Dazu sollte erwähnt werden, dass ein Zusammenhang zwischen „Entropie“ und „Information“ schon früh erkannt wurde, ohne dass der Begriff „Information“ explizit auftauchte. Ein berühmtes Beispiel dazu ist der sog. *Maxwell'sche Dämon* [9]. In einem Gedankenexperiment kann dieser Dämon schnelle und langsame Teilchen in den zwei Hälften eines Gasbehälters unterscheiden. Durch geschicktes Öffnen und Schließen einer kleinen Öffnung in der Trennwand gelingt es dem Dämon, die schnellen Moleküle in der einen und die langsamen in der anderen Hälfte des Containers zu sammeln. Da auf diese Weise spontan ein Temperaturunterschied zwischen beiden Hälften entsteht, kann der Dämon scheinbar den 2. Hauptsatz aushebeln. Im Jahr 1929 lieferte Leo Szilard mit Hilfe eines vereinfachten Modells eine quantitative Untersuchung solch einer „Entropieminderung bei Eingriffen intelligenter Wesen“ und zeigte, dass die Entropie des Gesamtsystems „Gas + Dämon“ nicht abnimmt [10]. Eine ausführliche Diskussion dazu findet sich in Leon Brillouins Buch „Science and Information Theory“ [11].

Bevor wir die „abstraktere Version“ nach dem zweiten Gleichheitszeichen in Gl. (18) mit dem *statistischen Operator*  $\rho$  (s. Anhang B) diskutieren, betrachten wir zunächst den Spezialfall, bei dem die Wahrscheinlichkeiten aller  $W$  möglichen Zustände gleich sind, d. h.  $p_i = 1/W$  gilt. Dann folgt mit  $\log p_i = \log(1/W) = -\log W$  wieder Boltzmanns Resultat Gl. (17), da alle  $W$  Terme in der Summe in Gl. (18) gleich sind.

Für den Spezialfall, dass sich nach der *Entfernung einer Einschränkung*, wie beim erwähnten Herausziehen der Trennwand beim idealen Gas, wieder ein Gleichgewichtszustand gleicher Energie einstellt, versteht man nun leicht, dass sich dabei die Entropie erhöht. In der Zählung der Mikrozustände tragen *zusätzlich* auch die Zustände bei, die mit der Einschränkung nicht auftreten können:  $S_{\text{End}} > S_{\text{Anf}}$ .

In makroskopischen Systemen ist im Allgemeinen nur eine *bestimmte Klasse von Observablen* (messbaren physikalischen Größen)  $\{A_i\}$  von Interesse. Diese Observablen stellen i. A. nur einen *sehr kleinen Teil* aller möglichen Observablen des Systems dar, da meist nur wenige Eigenschaften eines Systems experimentell zugänglich sind. Wir nennen sie die *relevanten* Observablen [12]. Sie legen die *Beobachtungsebene* fest. In einem besser ausgestatteten Labor könnte es möglich sein, eine umfassendere Wahl der Beobachtungsebene zu verwenden  $\{A\} \rightarrow \{\tilde{A}\}$ , da man mehr Observable experimentell kontrollieren kann, und so zu einer detaillierteren Beschreibung gelangt. Beispiele zur Wahl der relevanten Observablen werden später gegeben.

Auch für Anfangszustände, die nicht Systemen im thermodynamischen Gleichgewicht entsprechen, kann man das Konzept eines Ensembles experimentell identisch präparierter Systeme verwenden. Solch ein Ensemble kann man immer durch einen *speziellen* vollständigen Satz von „orthonormierten“ Zuständen  $|n\rangle$  im Hilbertraum beschreiben, wobei der Zustand  $|n\rangle$  im Ensemble mit der Wahrscheinlichkeit  $p_n$  auftritt und  $\sum_n p_n = 1$  gilt (s. Anhang B). Das Ensemble sei zum Zeitpunkt der Präparation durch die *Ensemblemittelwerte*

$$\langle A_i \rangle_{t_0} = \sum_n p_n \langle A_i \rangle_{t_0}^{(n)} \quad (= Sp(\rho(t_0)A_i)) \quad (19)$$

charakterisiert, wobei  $\langle A_i \rangle_{t_0}^{(n)}$  der quantenmechanische *Erwartungswert* des Operators  $A_i$  im Zustand  $|n\rangle$  ist. In der abstrakteren zweiten Form auf der rechten Seite taucht der statistische Operator  $\rho(t_0)$  auf

$$\rho(t_0) = \sum_n p_n |n\rangle \langle n|, \quad (20)$$

der den Anfangszustand des Systems beschreibt und „*Sp*“ bedeutet die Spur über den Hilbertraum (s. Anhang B). Wie die Observablen  $A_i$  wirkt  $\rho(t_0)$  als linearer Operator auf dem Hilbertraum und  $\sum_n p_n = 1$  ist gleichbedeutend zur Gleichung  $Sp(\rho(t_0)) = 1$ .

Für den Fall, dass sich die relevanten Observablen *additiv* aus vielen Operatoren des gleichen Typs zusammensetzen (Gesamtenergie, Teilchenzahl in einem Teil des Systems, ...), werden die *Schwankungen um die Mittelwerte sehr klein* sein. In solchen Fällen stimmen die Werte für ein einzelnes System mit den Mittelwerten sehr gut überein.

Offensichtlich ist durch Gl. (19), d. h. die *alleinige* Vorgabe der Mittelwerte  $\langle A_i \rangle_{t_0}$  der *relevanten* Variablen, der statistische Operator  $\rho(t_0)$  *nicht* vollständig festgelegt. Wie bereits für den Spezialfall von Gleichgewichts-Ensembles besprochen, sollte die Festlegung so *unvoreingenommen* wie möglich geschehen. Darauf wurde insbesondere von Jaynes [13] in Verallgemeinerung der Vorgehensweise von Gibbs [6] und von Neumann [7] hingewiesen:

Jaynes'sches Prinzip des maximalen Unschärfemaßes:

Der statistische Operator  $\rho(t_0)$  sollte die *maximale Informationsentropie*

$$S_{\rho(t_0)} = -k_B Sp(\rho(t_0) \log \rho(t_0))$$

unter den *Nebenbedingungen*  $Sp(\rho(t_0)) = 1$  und  $\langle A_i \rangle_{t_0} = Sp(\rho(t_0) A_i)$  besitzen.

Man kann das als Verallgemeinerung des von Laplace propagierten „*principle of insufficient reason*“ ansehen.

Auch ohne die explizite Form des sich aus diesen Forderungen ergebenden statistischen Operators maximaler Entropie  $R_{\{A\}}(t_0)$  anzugeben, lässt sich die Bedeutung dieses Prinzips für eine präzise Formulierung des zweiten Hauptsatzes im Rahmen der statistischen Mechanik angeben.

Wir betrachten die Dynamik eines vom Rest der Welt isolierten Systems. Dem System wird also von außen *keine* Wärme zugeführt, was als *adiabatischer* Prozess bezeichnet wird. Zeitabhängige äußere Störungen (z. B. durch Anlegen von äußeren Feldern) werden aber zugelassen.

In der *exakten* Beschreibung der Dynamik löst man im „Schrödinger-Bild“ die Schrödingergleichung mit der Anfangsbedingung  $|n(t_0)\rangle = |n\rangle$  (s. Anhang B). Den statistischen Operator  $\rho(t)$  für  $t > t_0$  erhält man, indem man in Gl. (20) die Zustände  $|n\rangle$  durch die Lösungen  $|n(t)\rangle$  der Schrödingergleichung ersetzt, die ebenfalls ein Orthonormalsystem bilden. Die Wahrscheinlichkeiten  $p_n$  bleiben unverändert. Führt man die Spur in Gl. (20) nun mit Hilfe der Zustände  $\{|n(t)\rangle\}$  aus, so sieht man, dass *keine* zeitliche Änderung der *Informationsentropie* auftritt:

$$S_{\rho(t)} = -k_B Sp(\rho(t) \log \rho(t)) = -k_B \sum p_n \log p_n = S_{\rho(t_0)}. \quad (21)$$

Wie passt das mit dem 2. Hauptsatz zusammen?

Nach der Festlegung des Anfangszustandes, über den wir nur beschränkte Information haben, wurde bei der exakten Lösung der Dynamik *keine Information verschenkt*.

Wenn man nur an der Dynamik der relevanten Observablen interessiert ist, kann man wieder den „verkürzten“ statistischen Operator maximaler Entropie  $R_{\{A\}}(t)$  zur späteren Zeit  $t$  betrachten, der die Forderungen  $Sp(R_{\{A\}}(t)) = 1$  und

$$Sp(R_{\{A\}}(t) A_i) = \langle A_i \rangle_t = Sp(\rho(t) A_i) \quad (22)$$

erfüllt. Jetzt tritt ein wichtiger Unterschied zwischen  $R_{\{A\}}(t)$  und  $\rho(t)$  auf. Berechnet man Erwartungswerte von *nicht-relevanten* Observablen  $B_\alpha$ , so erhält man durch  $Sp(\rho(t) B_\alpha)$  das *exakte* Resultat  $\langle B_\alpha \rangle_t$ , wenn der Anfangszustand durch den statistischen Operator  $R_{\{A\}}(t_0)$  gegeben ist. Ersetzt man dagegen  $\rho(t)$  durch den verkürzten statistischen Operator  $R_{\{A\}}(t)$ , so kann man nicht erwarten, dass  $Sp(R_{\{A\}}(t) B_\alpha)$  eine gute Näherung für alle  $\langle B_\alpha \rangle_t$  liefert. Mit der Ersetzung

$$\rho(t) \rightarrow R_{\{A\}}(t) \quad (23)$$

hat man wichtige *Information verschenkt*. Dies hat Auswirkung auf die Entropiebilanz. Auf Grund der Maximaleigenschaft, über die  $R_{\{A\}}(t)$  definiert ist, gilt

$$S_{\{A\}}(t) = -k_B Sp(R_{\{A\}}(t) \log R_{\{A\}}(t)) \geq S_{\rho(t)}. \quad (24)$$

Fasst man nun die Resultate zusammen, so sieht man, dass die *begleitende* (oder „thermodynamische“) Entropie  $S_{\{A\}}(t)$  i. A. zunimmt [10]

$$S_{\{A\}}(t) \geq S_{\rho(t)} = S_{\rho(t_0)} = S_{\{A\}}(t_0).$$

Der Verzicht auf die vollständige Information in  $\rho(t)$ , d. h. die Verkürzung  $\rho(t) \rightarrow R_{\{A\}}(t)$ , ist der Grund für die Zunahme der begleitenden Entropie:  $S_{\{A\}}(t) \geq S_{\{A\}}(t_0)$ .

Im folgenden Unterkapitel wird die explizite Form des statistischen Operators  $R_{\{A\}}$  maximaler Entropie angegeben und Beispiele für die Wahl der relevanten Observablen präsentiert.

### 3.3 Verallgemeinertes Gibbs-Ensemble

Eine recht einfache Rechnung zeigt, dass  $\rho(t_0)$  unter der Forderung maximaler Informationsentropie unter den Nebenbedingungen  $Sp(\rho(t_0)) = 1$  und  $\langle A_i \rangle_{t_0} = Sp(\rho(t_0) A_i)$  die Form einer Exponentialfunktion annimmt [12,13,14]

$$\rho(t_0)_{maxEnt} \equiv R_{\{A\}}(t_0) = \frac{e^{-\sum_i \lambda_i(t_0) A_i}}{Sp e^{-\sum_i \lambda_i(t_0) A_i}}. \quad (25)$$

Dabei sind die  $\lambda_i(t_0)$  sogenannte *Lagrangeparameter*, die so zu bestimmen sind, dass  $Sp(R_{\{A\}}(t_0) A_i) = \langle A_i \rangle_{t_0}$  gilt, und damit Funktionen aller  $\langle A_j \rangle_{t_0}$  sind. Wie in Anhang B gezeigt, bestimmen sie direkt die infinitesimalen Änderungen  $dS_{\{A\}}$  der Entropie  $S_{\{A\}}(t_0) = -k_B Sp(R_{\{A\}}(t_0) \log(R_{\{A\}}(t_0)))$  bei infinitesimalen Änderungen  $d\langle A_i \rangle_{t_0}$  der Mittelwerte  $\langle A_i \rangle_{t_0}$

$$dS_{\{A\}}(t_0) = k_B \sum_i \lambda_i(t_0) d\langle A_i \rangle_{t_0}. \quad (26)$$

Man nennt die  $\lambda_i(t_0)$  auch *thermodynamische Kräfte*.

Der einfachste Fall tritt bei Systemen auf, die sich zum Anfangszeitpunkt  $t_0$  im *thermischen Gleichgewicht* befinden. Die relevanten Observablen sind dann die *additiven Erhaltungsgrößen*. Für ein abgeschlossenes (ruhend) System fester Teilchenzahl ist die Gesamtenergie  $E = \langle H \rangle$  Erhaltungsgröße, wobei  $H$  der Hamiltonoperator (im klassischen Fall die Hamiltonfunktion) ist. Also tritt in Gleichung (26) nur der Term mit  $i = 0$  auf  $dS_H = k_B \lambda_0 d\langle H \rangle = k_B \lambda_0 dE$ . Verwendet man zum Vergleich Gl. (9) aus der Thermodynamik, so erhält man bei festem Volumen  $dS = (1/T)dE$ . Also liefert der Zugang mit dem Hamiltonoperator als einzigem



relevanten Operator eine mikroskopische Begründung der Thermodynamik, wenn man  $\lambda_0$  invers zur Temperatur  $T$  wählt

$$\lambda_0 = 1/(k_B T) \equiv \beta. \quad (27)$$

Man nennt den entsprechenden statistischen Operator „kanonisch“

$$\rho_{kan} \equiv R_H(t_0) = e^{-\beta H} / Sp(e^{-\beta H}) \quad (28)$$

und das entsprechende Ensemble das *kanonische Ensemble*. Die allgemeine Form von  $R_{\{A\}}$  in Gl. (25) bezeichnet man auch als generalisierten kanonischen Operator [12] und das zugehörige Ensemble als *generalisiertes kanonisches Ensemble* oder *verallgemeinertes Gibbs-Ensemble*.

Ein weiteres wichtiges Beispiel ist der Ansatz zur mikroskopischen Begründung der *irreversiblen Thermodynamik* im Rahmen der statistischen Mechanik. Dazu unterteilt man das Gesamtsystem in der theoretischen Beschreibung in schwach wechselwirkende, noch makroskopische, Untersysteme. Der Index  $i$  bei den relevanten Observablen  $A_i$  wird dann zum Doppelindex  $n, i$ , wobei  $n$  das Untersystem und  $i$  den Charakter der Observablen (Energie, Teilchenzahl, ...) bezeichnet. Die Wahl der Energien der Teilsysteme als relevante Variable ist der Ausgangspunkt zur Beschreibung der *Wärmeleitung* [12].

Als drittes Beispiel betrachten wir eine Flüssigkeit, in der sich ein schweres Teilchen mit Masse  $M$  („Geschoss“) befindet, das sich mit einer Anfangsgeschwindigkeit bewegt, die *sehr* viel größer ist, als die mittlere thermische Geschwindigkeit, die das Teilchen im thermischen Gleichgewicht mit der Flüssigkeit hätte. Hier wählt man z. B. die Gesamtenergie, d. h.  $H$ , sowie die drei Komponenten  $P_i$  des Impulses des schweren Teilchens als die relevanten Observablen. Eine einfache Überlegung (s. Anhang B) liefert für die Lagrangeparameter der Impulskomponenten

$$\lambda_i(t) = -\lambda_0(t) \langle P_i \rangle_t / M \equiv -\beta(t) \langle P_i \rangle_t / M, \quad i = 1, 2, 3. \quad (29)$$

Für die zeitliche Änderung der begleitenden Entropie erhält man mit Gl. (26)

$$\frac{dS_{\{A\}}(t)}{dt} = k_B \sum_i \lambda_i(t) \frac{d\langle A_i \rangle_t}{dt}. \quad (30)$$

Kennt man die Zeitableitungen der  $\langle A_i \rangle_t$ , so erhält durch Integration die die begleitende Entropie  $S_{\{A\}}(t)$ .

Als Beispiel dazu betrachten wir das „schwere schnelle Teilchen“ in der Flüssigkeit. Mit Gl. (29) lautet Gl. (30)

$$\begin{aligned} \frac{dS_{\{A\}}(t)}{dt} &= k_B \beta(t) \left[ \frac{d\langle H \rangle_t}{dt} - \frac{1}{M} \sum_{i=1}^3 \langle P_i \rangle_t \frac{d\langle P_i \rangle_t}{dt} \right] \\ &= -k_B \beta(t) \frac{d}{dt} \frac{\langle \vec{P} \rangle_t^2}{2M} \approx -\frac{1}{T} \frac{d}{dt} \frac{\langle \vec{P} \rangle_t^2}{2M}. \end{aligned} \quad (31)$$

Wegen der Gesamtenergieerhaltung gilt  $d\langle H \rangle_t / dt = 0$ . In der zweiten Gleichung wurde zur Abkürzung die Vektorschreibweise verwendet und angenommen, dass das System der Flüssigkeit sehr groß ist, so dass die Temperaturerhöhung beim Abbremsen des „Geschosses“ vernachlässigt werden kann. Damit ist  $S_{\{A\}}(t) + T_{kin}^{(M)}(t)/T$  zeitunabhängig, wobei  $T_{kin}^{(M)}$  die kinetische Energie des schweren Teilchens bezeichnet. Mit  $\langle \vec{P} \rangle_t = M\langle \vec{V} \rangle_t$  gilt also

$$S_{\{A\}}(t) = S_{\{A\}}(t_0) + \frac{1}{T} \left[ \frac{M}{2} \langle \vec{V} \rangle_{t_0}^2 - \frac{M}{2} \langle \vec{V} \rangle_t^2 \right]. \quad (32)$$

Da die Mittelwerte  $\langle V_i \rangle_t$  der Geschwindigkeitskomponenten beim Abbremsen des Teilchens monoton abnehmen, nimmt die begleitende Entropie *monoton* zu. Die Entropiezunahme  $M\langle \vec{V} \rangle_{t_0}^2 / (2T)$  nach vollständigem Abbremsen ist leicht zu verstehen. Die kinetische Energie des schweren Teilchens wird vollständig in Wärmeenergie der Flüssigkeit „dissipiert“.

Der genaue zeitliche Verlauf von  $S_{\{A\}}(t)$  kann erst angegeben werden, wenn die  $\langle V_i \rangle_t$  bekannt sind. Dazu verwendet man näherungsweise eine phänomenologische Bewegungsgleichung für das schwere Teilchen mit einer mittleren Reibungskraft  $\vec{F}_{Reibung} = -\gamma(|\langle \vec{V} \rangle_t|)\langle \vec{V} \rangle_t$ , d. h.

$$M \frac{d}{dt} \langle \vec{V} \rangle_t = -\gamma(|\langle \vec{V} \rangle_t|) \langle \vec{V} \rangle_t. \quad (33)$$

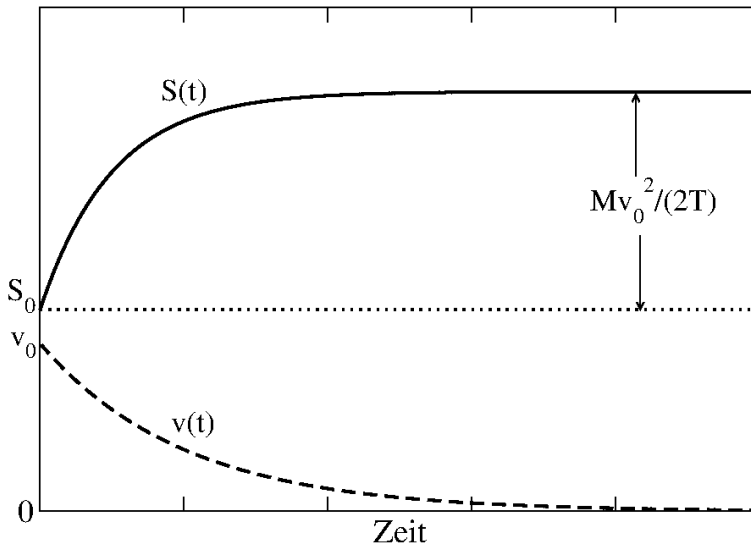
Bei nicht zu großen Anfangsgeschwindigkeiten kann man  $\gamma(|\langle \vec{V} \rangle_t|)$  durch die sog. *Reibungskonstante*  $\gamma$  ersetzen, die proportional zur *Viskosität* der Flüssigkeit ist. Mit dieser Vereinfachung fallen die  $\langle V_i \rangle_t$  zeitlich *exponentiell* ab und der zeitliche Verlauf von  $S_{\{A\}}(t)$  folgt mit Hilfe von Gl. (32):

Die Abbildung zeigt den zeitlichen Verlauf des Betrags der Geschwindigkeit und der begleitenden Entropie.

Da die exakte Dynamik zur Bestimmung der  $\langle A_i \rangle_t$  i. A. nicht durchgeführt werden kann, besteht wie beim Beispiel der Abbremsung des schweren Teilchens der zweite wichtige Schritt der verkürzten Beschreibung darin, dynamische Gleichungen vom Typ

$$\frac{d}{dt} \langle A_i \rangle_t = F_i(\{\langle A \rangle\}). \quad (34)$$

aufzustellen, wobei auf der rechten Seite in die aufzufindende Funktion  $F_i$  nur die Mittelwerte  $\langle A_j \rangle_{\tilde{t}}$  für  $\tilde{t} \leq t$  eingehen. Damit versucht man eine gute *Näherung* für die exakte zeitliche Dynamik der  $\langle A_i \rangle_t$  zu erhalten. In Systemen, in denen eine klare *Trennung* von mikroskopischen und makroskopischen *Zeitskalen* auftritt, spielt die „Vergangenheit“ oft keine Rolle und auf der rechten Seite von Gl. (34) treten nur die  $\langle A_j \rangle_t$  zur *selben* Zeit auf, wie auf der linken Seite. Ein Beispiel für solch stark verschiedene Zeitskalen sind bei verdünnten Gasen die Zeitdauer eines Stoßes zwischen zwei Gasatomen und die sehr viel längere Zeit bis einer der Stoßpartner den nächsten Stoß ausführt. In solchen Fällen vereinfachen sich Gl. (34) zu einem System von *Differentialgleichungen*, in Verallgemeinerung von Gl.(33) für



den Fall der Abbremsung des Teilchens. In diesen Differentialgleichungen treten „Transportkoeffizienten“ oder „Reaktionsraten“ auf, die (im Prinzip) durch den Rückgriff auf die mikroskopische Beschreibung berechnet werden können. Oft werden sie aber als *phänomenologische* Parameter eingeführt und als *Anpassparameter* behandelt. Nach Lösung der genäherten dynamischen Gleichungen für die Ensemblemittelwerte der relevanten Observablen erhält man mit Hilfe von Gl. (26) den zeitlichen Verlauf von  $S_{\{A\}}(t)$ . Er hängt von der speziellen Wahl der relevanten Observablen, sowie der genauen Form der dynamischen Gleichungen ab. Der Zeitverlauf von verschiedenen begleitenden Entropien kann deutlich unterschiedlich sein [14]. Findet man in der Literatur Resultate zum Zeitverlauf „der Entropie“ eines Systems, ohne Angabe ihrer genauen Definition, so ist Skepsis angesagt.

### 3.4 Irreversibilität

In Kapitel 2 wurde für makroskopische Systeme die Idealisierung *reversibler* Prozesse diskutiert, die auch in zeitlich umgekehrter Reihenfolge ablaufen können. Als notwendige Voraussetzung für solche Prozesse wurde ihr „quasistatischer“ zeitlicher Ablauf genannt. Die Mehrzahl der in makroskopischen Systemen ablaufenden Prozesse ist *irreversibel*, d. h. unumkehrbar. Ein drastisches Beispiel stellt das Herunterfallen eines rohen Eies dar. Laufen in makroskopischen Systemen Prozess mit endlicher Geschwindigkeit ab, so lassen sie sich in der Regel nicht umkehren. Das ist anders bei Systemen aus wenigen Teilchen. Lässt man einen gut aufgepumpten Ball senkrecht fallen, so fällt er mit zunehmender Geschwindigkeit. Ist der Fußboden hinreichend hart, so wird er nach der Reflexion (fast) wieder zur Ausgangssituation zurückkehren. Durch die elastische Deformation des Balles und

des Fußbodens und der Übergang der damit verbundenen Energie in ungeordnete Bewegung ist auch das Fallenlassen des Balles genau genommen ein irreversibler Prozess.

Können nicht-quasistatische Prozesse, die in makroskopischen Systemen ablaufen, reversibel sein?

Wir betrachten dazu den in Kap. 2 beschriebenen Expansionsversuch, bei dem einem Gas plötzlich das doppelte Volumen zu Verfügung steht. Wir beginnen mit der Beschreibung im Rahmen der klassischen Mechanik und nehmen an, dass die Orte und Geschwindigkeiten aller Atome vor dem Herausziehen der Trennwand bekannt sind. Nach hinreichend langem Warten ist, außer für extrem unwahrscheinliche Anfangsbedingungen (z. B. *alle* Teilchengeschwindigkeiten senkrecht zur Trennwand), die mittlere Teilchendichte im gesamten Kasten ungefähr gleich. Stünde ein Mechanismus zur Verfügung, alle Teilchengeschwindigkeiten zum selben Zeitpunkt umzukehren, so würde das System zu seinem Anfangszustand zurückkehren, wenn wir das System nochmals dieselbe Zeitdauer durchlaufen lassen und dann wiederum alle Geschwindigkeiten umkehren. Diese exakte Beschreibung der Dynamik des Systems passt zur Verwendung der Informationsentropie, die zeitlich konstant bleibt. Auch im Rahmen der quantenmechanischen Beschreibung kann man, wenn das Gas das gesamte Volumen erfüllt, in der theoretischen Beschreibung eine „Bewegungsumkehr“ (auch „Zeitspiegelung“ genannt), durchführen. Lässt man dann die exakte Dynamik nochmals genau so lange ablaufen und führt anschließend wieder eine Bewegungsumkehr durch, so gelangt man zum Ausgangszustand (mit allen Atomen in der linken Hälfte). In der Praxis lässt sich solch eine Bewegungsumkehr für das betrachtete Beispiel aber nicht durchführen. Das System erweckt nach hinreichend langer Zeit den *Eindruck*, als ob sich der thermische Gleichgewichtszustand eingestellt hat. Man erhält dann die in Kap. 2 berechnete Änderung der „thermostatischen“ Entropie. Das zeitliche Anwachsen einer begleitenden Entropie kann man erhalten, wenn man die Dynamik der Expansion näherungsweise mit Hilfe der *Boltzmann-Gleichung* beschreibt [14].

Wie diskutiert, ist die Bewegungsumkehr bei makroskopischen Systemen in der Regel nicht möglich. Eine Ausnahme bilden makroskopische magnetische Systeme, bei denen in sog. *Spinechoexperimenten* solch eine „Zeitumkehrung“ gelungen ist [14]. Nachdem die makroskopische Magnetisierung von einem endlichen Nichtgleichgewichtswert auf (praktisch) Null abgefallen war, konnte man durch geschickte zeitliche Magnetpulse das System zu seinem Anfangszustand mit endlicher Magnetisierung zurückführen.

Hier ein leichter zu verstehendes Beispiel aus dem täglichen Leben. Wir blicken in einem Leichtathletikstadion auf die Vierhundertmeterbahn und sehen ca. hundert Läufer, die fast *gleichmäßig über die Bahn verteilt* ihre Runden drehen. Es sieht aus, wie ein (ungeordneter) Gleichgewichtszustand, mit gleicher mittlerer Läuferdichte entlang der Bahn. Jetzt kommt die Ansage, dass jeder Läufer sich umdrehen („*Magnetpuls*“) und dann weiterlaufen soll. Nach einigen Minuten wundern wir uns, dass die Verteilung der Läufer über die Bahn immer ungleichmäßiger wird, und plötzlich sind alle Läufer fast an der gleichen Stelle.

Wie kann man es verstehen, dass aus scheinbarer Unordnung auf einmal Ordnung entstanden ist? Hätten wir die *Information* über die Vorgeschichte gehabt, so müssten wir uns nicht wundern: Alle Läufer sind einige Minuten bevor Sie sie erstmals beobachtet haben, gleichzeitig an der Startlinie losgelaufen. Auf Grund der Tatsache, dass sehr unterschiedliche Läufer gestartet sind, hat sich das Anfangsfeld fast gleichmäßig über die Bahn verteilt. Die entstandenen Abstände der Läufer wurden durch die Umkehransage wieder geringer.

## 4 Fazit

Zu den vorgestellten Überlegungen zum Wesen des Entropiebegriffs und seiner engen Beziehung zum Konzept der Information wurde in Kap. 3 nur zum einfachen Beispiel des schnellen schweren Teilchens in einer Flüssigkeit eine detaillierte Rechnung präsentiert. Ein Vielzahl weiterer wichtiger Beispiele findet man in den Referenzen [12] und [14].

Eine Vielzahl von Dingen konnte hier natürlich nicht angesprochen werden. Zur grundsätzlichen Frage, was Information eigentlich ist, gibt es z. B. durchaus verschiedene Ansichten. Im Zusammenhang mit der Frage des minimalen Energieverbrauchs in den Schritten, die ein Computer ausführt, hat Rolf Landauer die Aussage „Information is physical“ propagiert [15]. Er weist darauf hin, dass jede Form von Information notwendigerweise an eine physikalische Darstellung gebunden ist. Dagegen empfiehlt der Philosoph Peter Janich in seinem Buch „Was ist Information?“ die Überwindung des Irrtums, in der Information einen Naturgegenstand zu sehen (und gleichwohl erfolgreich mit dem Informationsbegriff in Technik- und Naturwissenschaften zu arbeiten) [16].

Eine weit über die Physik hinausgehende Diskussion zum Themenkreis „Entropie und Information“ findet man in Manfred Eigens kürzlich erschienener Monographie „From Strange Simplicity to Complex Familiarity“ [17].

Abschließend hier noch einmal die dargestellte Einsicht zum Anwachsen der Entropie:

Ist eine exakte mikroskopische Beschreibung der zeitlichen Entwicklung eines physikalischen Systems möglich, so ist die Informationsentropie *zeitunabhängig*.

In einer verkürzten Beschreibung, die nur „relevante“ Observable beinhaltet, erhält man das *Anwachsen* der zugehörigen *begleitenden Entropie* zum Wert der thermodynamischen Entropie des Gleichgewichtszustandes, der im Langzeitlimit erreicht wird. Der detaillierte zeitliche Verlauf hängt von der speziellen Wahl der relevanten Observablen ab, d. h. auf welche Information über die Dynamik des Systems man bei der Beschreibung verzichtet.

## 5 Anhang A

Im Text wurde die Äquivalenz der Formulierungen A) und C) des zweiten Hauptsatzes der Thermodynamik gezeigt. Hier soll auch die Äquivalenz von A) und B), und damit auch die von B) und C) gezeigt werden.

Die Äquivalenz von A) und B) zeigt man am einfachsten, indem darlegt, dass falls die Behauptung in A) falsch ist, auch die in B) falsch ist, und umgekehrt:

Wir nehmen also an, dass es einen spontanen Prozess gibt, in dem eine Wärmemenge  $Q$  vom kalten System auf das wärmere System übergeht. Anschließend lassen wir zwischen den beiden Wärmereservoirs eine Dampfmaschine laufen, die genau diese Wärmemenge  $Q$  dem wärmeren System wieder entzieht. Dabei wird ein *Teil* von  $Q$  in Arbeit verwandelt, der Rest wird als kleinere Wärmemenge auf das kältere System übertragen. Nach beiden Schritten ist also dem *einen* kälteren System Wärme entzogen worden und Arbeit geleistet worden. Also ist auch B) falsch.

Sei nun die Annahme in B) falsch, d. h. ein Wärmereservoir liefert spontan die Arbeit  $W$ . Dann können wir diese Arbeit dazu verwenden, um eine wärmere System durch „Umrühren“ weiter zu erwärmen, d. h. auch die Aussage in A) ist falsch.

Wegen der im Text gezeigten Äquivalenz von A) und C) und der hier gezeigten Äquivalenz der Formulierungen A) und B), folgt schließlich auch die Äquivalenz von B) und C).

## 6 Anhang B

Die statistische Mechanik geht von der *mikroskopischen* Beschreibung physikalischer Systems aus. Ein wichtiges Anliegen dieses Bereichs der theoretischen Physik ist es, Methoden anzugeben, mit denen man die in der Thermodynamik auftretenden Zustandsfunktionen *berechnen* kann. Die wesentlichen Konzepte der statistischen Mechanik entstanden vor der Geburt der Quantenmechanik (1925), basierend auf den Gesetzen der *klassischen Mechanik*.

Anmerkung: In diesem Anhang steht der Buchstabe  $p$  (groß oder klein, mit Indizes oder nicht) für den *Impuls* der Teilchen (s. u.) und *nicht* für den Druck, wie im Kapitel zur Thermodynamik.

Das makroskopische (klassische) Vielteilchensystem bestehe aus  $N$  „Atomen“, die als „Punktteilchen“ beschrieben werden. Das  $i$ -te Teilchen ( $i = 1, 2, \dots, N$ ) habe die Masse  $m_i$ , und zur Zeit  $t$  den Ort  $\vec{x}_i(t)$ , die Geschwindigkeit  $\vec{v}_i(t)$  und den *Impuls*  $\vec{p}_i(t) = m_i \vec{v}_i(t)$ . Ein beliebiger Anfangszustand des Systems ist durch die Angabe der Orte und Impulse aller Teilchen

$$(\vec{x}_1, \dots, \vec{x}_N, \vec{p}_1, \dots, \vec{p}_N) \equiv \vec{X}$$

zur Anfangszeit  $t_0$  festgelegt. Der Zustand des Systems entspricht damit einem *Punkt* im  $6N$ -dimensionalen *Phasenraum*, der auch „ $\Gamma$ -Raum“ genannt wird. Die

Zeitentwicklung des Systems wird durch eine Trajektorie  $\vec{X}(t)$  in diesem hochdimensionalen Phasenraum beschrieben [1].

Die Grundidee der klassischen statistischen Mechanik ist folgende:

Man kann unmöglich alle  $N \approx 10^{23}$  Teilchenorte und Impulse kontrollieren. Von J. W. Gibbs stammt die Idee, sich ein *Ensemble* von  $I$  makroskopisch identisch präparierten Systemen vorzustellen. Bei dieser Art der Präparation unterscheiden sich die Phasenraumpunkte  $\vec{X}^{(i)}(t)$  für  $(i = 1, \dots, I)$  der verschiedenen Systeme im Ensemble dennoch. Der *Ensemblemittelwert* einer beliebigen physikalischen Größe  $A(\vec{X})$  ist durch

$$\langle A \rangle_t^{(I)} = \frac{1}{I} \sum_{i=1}^I A(\vec{X}^{(i)}(t)) \quad (35)$$

gegeben. Jedes individuelle Ergebnis  $A(\vec{X}^{(i)}(t))$  an einem System des Ensembles geht bei dieser Form der Mittelung mit gleichem Gewicht ein. Diese Art der Mittelung ist so unvoreingenommen wie möglich. Die allgemeine Form der Mittelung erhält man mit einer positiven *Wahrscheinlichkeitsdichte*  $\rho_t(\vec{X})$

$$\langle A \rangle_t^{(\rho)} = \int \rho_t(\vec{X}) A(\vec{X}) d\Gamma, \quad (36)$$

wobei  $d\Gamma$  die Integration über den Phasenraum bedeutet und die Normierungsbedingung  $\int \rho_t(\vec{X}) d\Gamma = 1$  lautet. Ein wichtiges Beispiel für die physikalische Größe  $A$  ist beim Versuch, die Verbindung zur Thermodynamik herzustellen, die sog. *Hamiltonfunktion*  $H(\vec{X})$ , die der Gesamtenergie entspricht.

Die zu einer vorgegebenen Wahrscheinlichkeitsdichte  $\rho(\vec{X})$  gehörige Entropie definiert man (bis auf eine Konstante) durch

$$S = -k_B \int \rho \log(\rho/c) d\Gamma. \quad (37)$$

Dabei ist  $c$  eine Konstante, um das Argument des Logarithmus dimensionslos zu machen.

Soll ein Ensemble zur Beschreibung von Systemen im thermischen Gleichgewicht dienen, so muss  $\rho_t(\vec{X})$  zeitunabhängig sein. Die Abhängigkeit von  $\vec{X}$  sollte nur über die additiven Erhaltungsgrößen stattfinden. Für das betrachtete System im ruhenden Kasten ist das die Gesamtenergie, d. h.  $\rho_{GI} = \rho(H(\vec{X}))$ . Im „kanonischen“ Ensemble zur Beschreibung von Gleichgewichtssystemen mit Temperatur  $T$  wählt man  $\rho_{GI} \sim e^{-H(\vec{X})/(k_B T)}$ . Als wichtiges Resultat der *klassischen* statistischen Mechanik folgt daraus, dass die mittlere kinetische Energie jedes Teilchens proportional zur Temperatur ist, d. h. hohe Temperaturen entsprechen hohen mittleren (ungerichteten) Teilchengeschwindigkeiten.

Vor analogen Überlegungen zur *Quantenstatistik* einige Bemerkungen zur Quantenmechanik [18]. Zum Einstieg betrachten wir *ein* Teilchen der Masse  $m$  in *einer* Raumdimension in einem zeitabhängigen äußeren Potential  $V_t(x)$ . Dann werden in

der klassischen Hamiltonfunktion  $H_t = p^2/(2m) + V_t(x)$  der Ort  $x$  und der Impuls  $p$  zu *linearen Operatoren* (auf einem Hilbertraum (s. u.)), die *nicht vertauschen*. Definiert man den *Kommutator*  $[x, p] \equiv xp - px$ , so lautet die Born-Heisenberg-Vertauschungsregel

$$[x, p] = i\hbar 1, \quad (38)$$

wobei  $\hbar = h/(2\pi)$  und  $h$  das Plancksche Wirkungsquantum ist. Sie impliziert die Heisenbergsche Unschärferelation für den Ort und Impuls des Teilchens. Die beiden Unbestimmtheiten hängen vom Zustand  $|\psi\rangle$  des Teilchens ab, dem im Ortsraum eine komplexe *Wellenfunktion*  $\psi(x)$  entspricht.

Zur Beschreibung der Dynamik gibt es in der Quantenmechanik verschiedene Möglichkeiten. Im „Heisenberg-Bild“ betrachtet man die Heisenbergschen Bewegungsgleichungen für Operatoren  $A_{Heis}(t)$ . Falls die Operatoren, wie Ort und Impuls keine „äußere Zeitabhängigkeit“ haben, lautet die Bewegungsgleichung

$$\dot{A}_{Heis}(t) \equiv \frac{d}{dt}A_{Heis}(t) = \frac{1}{i\hbar}[A, H_t]_{Heis}(t). \quad (39)$$

Für den Ortsoperator ergibt sich mit Gl. (38)

$$\dot{x}_{Heis}(t) = \frac{1}{i\hbar}[x, H_t]_{Heis}(t) = \frac{1}{i\hbar}\left[x, \frac{p^2}{2m}\right]_{Heis}(t) = p_{Heis}(t)/m. \quad (40)$$

Diese Relation zwischen den Operatoren im Heisenbergbild hat dieselbe Form wie für die klassischen Variablen. Der Zustand mit dem Erwartungswerte gebildet werden, ist der Zustand  $|\psi(t_0)\rangle$  zur Anfangszeit  $t_0$ , d. h. im Heisenbergbild liegt die Zeitabhängigkeit ganz bei den Operatoren.

Im Haupttext wurde das „Schrödinger-Bild“ verwendet, in dem die Zeitabhängigkeit über den *Zustand*  $|\Phi(t)\rangle$  des (Vielteilchen-)Systems geschieht. Er gehorcht der *Schrödingergleichung*

$$i\hbar \frac{d}{dt}|\Phi(t)\rangle = H_t|\Phi(t)\rangle, \quad (41)$$

wobei  $H_t$  der Hamiltonoperator des Systems ist. Für zeitunabhängigen Hamiltonoperator erhält man den Zustand des Systems zu einem späteren Zeitpunkt durch Lösung der Schrödingergleichung als  $|\Phi(t)\rangle = e^{-iH(t-t_0)/\hbar}|\Phi(t_0)\rangle$ . Für den Fall eines, etwa durch zeitabhängige äußere Felder, zeitabhängigen Hamiltonoperators gilt dagegen  $|\Phi(t)\rangle = U(t, t_0)|\Phi(t_0)\rangle$ , mit einem komplizierteren unitären ( $UU^\dagger = U^\dagger U = 1$ ) Zeitentwicklungsoperator  $U(t, t_0)$ , der nicht nur von der Zeitdifferenz  $t - t_0$  abhängt.

Verallgemeinert man das Ensemblekonzept auf Quantensysteme, so befinde sich das  $i$ -te System des Ensembles im (normierten) Zustand  $|\phi_i\rangle$ . Die Abhängigkeit von der Zeit ist hier unterdrückt. Der quantenmechanische Erwartungswert der Observablen  $A$  im  $i$ -ten System ist dann gleich  $\langle\phi_i|A|\phi_i\rangle$ . Der Ensemblemittelwert ist analog zu Gl. (35) definiert

$$\langle A \rangle^{(I)} = \frac{1}{I} \sum_{i=1}^I \langle\phi_i|A|\phi_i\rangle. \quad (42)$$



Die allgemeine Form der statistischen Mittelung erhält man mit einem selbstadjungierten, positiven *statistischen Operator*  $\rho$

$$\langle A \rangle^{(\rho)} = Sp(\rho A), \quad (43)$$

wobei  $Sp(C) \equiv \sum_{\alpha} \langle \alpha | C | \alpha \rangle$  die Spurbildung im Vektorraum der physikalischen Zustände bedeutet, der oft unendlichdimensional ist („Hilbertraum“). Dabei bilden die Zustände  $\{|\alpha\rangle\}$  ein beliebiges vollständiges Orthonormalsystem. Außerdem ist der statistische Operator auf eins normiert:  $Sp(\rho) = 1$ . Ausgedrückt durch seine positiven Eigenwerte  $p_n$  und orthonormierten ( $\langle n | m \rangle = \delta_{nm}$ ) Eigenzustände  $|n\rangle$  lautet der statistische Operator  $\rho$

$$\rho = \sum p_n |n\rangle \langle n|. \quad (44)$$

Für die (i. A. unendliche) Summe über die Wahrscheinlichkeiten  $p_n$  gilt  $\sum_n p_n = 1$ . Dieser statistische Operator beinhaltet die Aussage, dass sich das System mit Wahrscheinlichkeit  $p_n$  im Zustand  $|n\rangle$  befindet. Ist der statistische Operator des Gleichgewichts eine Funktion des Hamiltonoperators  $H$ , d. h.  $\rho_{Gl} = \rho(H)$ , dann sind die  $|n\rangle$  die Eigenzustände  $|E_n\rangle$  von  $H$ .

Im Schrödinger-Bild ist der statistische Operator zeitabhängig und  $\rho(t)$  ergibt sich mit Hilfe der Lösung der Schrödingergleichung  $|n(t)\rangle = U(t, t_0)|n\rangle$  für die Anfangsbedingung  $|n(t_0)\rangle = |n\rangle$  als

$$\rho(t) = \sum_n p_n |n(t)\rangle \langle n(t)|. \quad (45)$$

Wegen  $\langle n(t) | m(t) \rangle = \langle n | U^\dagger U | m \rangle = \langle n | m \rangle = \delta_{nm}$  bilden auch die  $\{|n(t)\rangle\}$  ein vollständiges Orthonormalsystem. Die zum statistischen Operator  $\rho(t)$  gehörige Entropie  $S(t)$  ist definiert als

$$S(t) = -k_B Sp(\rho(t) \log \rho(t)) = -k_B \sum_n p_n \log p_n. \quad (46)$$

Zur zweiten Form gelangt man, wenn man die Spur in den Eigenzuständen  $|n(t)\rangle$  von  $\rho(t)$  ausführt. Bei Verwendung der exakten mikroskopischen Dynamik mit Hilfe des unitären Operators  $U(t, t_0)$  ist diese Informationsentropie *zeitunabhängig*. Sie ist ein Maß für die fehlende Information über den quantenmechanischen Zustand eines Systems des Ensembles. Sie verschwindet für den Fall, dass  $p_m = 1$  für  $n = m$  gilt und daher alle anderen Wahrscheinlichkeiten gleich Null sind. Jedes System des Ensembles wird in diesem Fall durch einen wohldefinierten Zustand im Hilbertraum beschrieben. Die Entropie  $S$  ist ein Maß, wie weit der statistische Operator  $\rho$  von solch einem „reinen“ Zustand entfernt ist. Selbst im reinen Fall  $p_m = 1$  und  $p_n = 0$  für  $n \neq m$  verbleibt aber der für die Quantenmechanik charakteristische Wahrscheinlichkeitsaspekt bei der Messung einer Observablen. Für den Fall einer Observablen mit einem *diskreten* Eigenwertspektrum ergibt sich folgendes Bild:

Die möglichen Messwerte der Observablen  $A$  sind deren Eigenwerte  $a_i$  zum Eigenzustand  $|a_i\rangle$ . Da sich jedes System des Ensembles im Zustand  $|n(t)\rangle$  befindet, ist die Wahrscheinlichkeit  $w_i$  den Eigenwert  $a_i$  zu messen, nach Born durch das Absolutquadrat des Skalarproduktes von  $|a_i\rangle$  mit  $|n(t)\rangle$ , d. h.  $|\langle a_i|n(t)\rangle|^2$  gegeben. Im „gemischten“ Fall, in dem mehrere Wahrscheinlichkeiten  $p_n$  in der Entropiedefinition Gl. (46) von Null verschieden sind, hat man also einen „doppelten Wahrscheinlichkeitsaspekt“ bei der Messung der Observablen  $A$ .

Ein direkter Bezug zu Shannons, formal zur zweiten Relation in Gl. (46) identischen Entropiedefinition, lässt sich dann herstellen, wenn der statistische Operator  $\rho$  mit der Observablen  $A$  „vertauscht“, d. h.  $A\rho = \rho A$  gilt. Ein Beispiel ist der Gleichgewichtsfall  $\rho_{Gl} = \rho(H)$ , in dem der Hamiltonoperator  $H$  mit  $\rho$  vertauscht. Dann sind die  $p_n$  in Gl. (46) die Wahrscheinlichkeiten für das „Ereignis“ den Energieeigenwert  $E_n$  zu messen.

Wie im Haupttext besprochen, liefert die Entropiemaximierung unter den Nebenbedingungen der Normierung und der Mittelwerte der relevanten Observablen für den entsprechenden statistischen Operator die Form einer Exponentialfunktion

$$\rho(t_0)_{maxEnt} \equiv R_{\{A\}}(t_0) = \frac{e^{-\sum_i \lambda_i(t_0)A_i}}{Sp e^{-\sum_i \lambda_i(t_0)A_i}}, \quad (47)$$

wobei die Lagrangeparameter  $\lambda_i(t_0)$  so zu bestimmen sind, dass für die relevanten Observablen  $\langle A_i \rangle_{t_0} = Sp(R_{\{A\}}(t_0)A_i)$  gilt. Damit sind die  $\lambda_i(t_0)$  Funktionen der vorgegebenen Mittelwerte  $\langle A_j \rangle_{t_0}$ . Nützliche Relationen zwischen den verschiedenen  $\lambda_i(t_0)$  kann man oft mit Hilfe der Relation  $Sp([B, C]e^{-C}) = 0$  erhalten, die aus der zyklischen Invarianz der Spur folgt. Also gilt für beliebige Operatoren  $B$  für mit  $R_{\{A\}}(t_0)$  gebildete Mittelwerte

$$\langle [B, \sum_i \lambda_i A_i] \rangle = 0. \quad (48)$$

Für das Beispiel des schweren Teilchens mit der Wahl  $A_0 = H$  und  $A_i = P_i$  für  $i = 1, 2, 3$ , erhält man für  $B_i = X_i$  mit Hilfe von  $[X_i, H]/(i\hbar) = P_i/M$  (s. Gl. (40)) und der Born-Heisenbergschen Vertauschungsrelation  $\lambda_i = -\lambda_0 \langle P_i \rangle / M$ .

Die mit  $R_{\{A\}}(t_0)$  berechnete Entropie  $S_{\{A\}}$  lautet

$$\begin{aligned} S_{\{A\}}(t_0) &\equiv -k_B Sp(R_{\{A\}}(t_0) \log R_{\{A\}}(t_0)) \\ &= k_B \sum_i \lambda_i(t_0) \langle A_i \rangle_{t_0} + k_B \log(Sp e^{-\sum_i \lambda_i(t_0)A_i}). \end{aligned} \quad (49)$$

Nach Konstruktion ist sie *größer* (oder gleich) als jede Informationsentropie, die mit einem anderen statistischen Operator gebildet wird, der *dieselben Erwartungswerte*  $\langle A_i \rangle_{t_0}$  liefert.

Differenziert man den zweiten Term auf der rechten Seite von Gl. (49) partiell nach  $\lambda_j(t_0)$ , so erhält man bis auf einen Faktor den Mittelwert  $\langle A_j \rangle_{t_0}$

$$-\frac{\partial}{\partial \lambda_j(t_0)} \log(Sp e^{-\sum_i \lambda_i(t_0)A_i}) = \frac{Sp A_j e^{-\sum_i \lambda_i(t_0)A_i}}{Sp e^{-\sum_i \lambda_i(t_0)A_i}} = \langle A_j \rangle_{t_0}. \quad (50)$$

Tritt nur eine relevante Observable auf, wie beim Beispiel des thermischen Gleichgewichts in Kap. 3, oder vertauschen alle Operatoren  $A_i$  untereinander, so folgt das nach elementarer Rechnung. Selbst wenn die Operatoren  $A_i$  *nicht* vertauschen, spielt das bei der Bildung der *ersten* partiellen Ableitungen innerhalb der Spur keine Rolle (Hellmann-Feynman-Theorem).

Für eine kleine (differentielle) Änderung der Entropie  $S_{\{A\}}$  bei kleiner (differentieller) Änderung der  $\langle A_i \rangle$  erhält man nach partiellem Differenzieren nach diesen Mittelwerten unter Benutzung von Gl. (50) und der Kettenregel

$$dS_{\{A\}} = \sum_i \frac{\partial S_{\{\langle A_i \rangle_t\}}}{\partial \langle A_i \rangle_t} d\langle A_i \rangle = k_B \sum_i \lambda_i d\langle A_i \rangle. \quad (51)$$

Diese Relation wird in Abschnitt 3.3 verwendet.

## Referenzen und Anmerkungen

[1] Richard Becker, *Theorie der Wärme*, Heidelberg Taschenbücher Bd. 10, Springer-Verlag Berlin, Heidelberg, New York 1966

[2] Mit Hilfe des Begriffs der *Fakultät*  $N!$  der Zahl  $N$  als  $N! = N \cdot (N-1) \cdot \dots \cdot 2 \cdot 1$  lautet der Binomialkoeffizient für die Zahl der Mikrokonfigurationen von  $M$  Kugeln in der linken und  $N-M$  Kugeln in der rechten Hälfte

$$\binom{N}{M} = \frac{N!}{M!(N-M)!}.$$

[3] Ludwig Boltzmann, *Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärme Gleichgewicht*, Wiener Berichte 76, 373–435, 1877

Anmerkung: Boltzmann verwendet den Buchstaben  $p$  an Stelle von  $n$ . Da  $p$  aber bereits für den Druck und  $p_i$  für die Komponenten des Impulses verwendet werden, kann diese Änderung mögliche Verwirrung vermeiden.

[4] Ludwig Boltzmann, *Über die Mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie*, Wiener Berichte 53, 195–220, 1866

[5] Zur Interpretation der auftretenden Wahrscheinlichkeiten gibt es zwei Hauptansichten: Wahrscheinlichkeiten als *relative Häufigkeiten* oder als *Maß für die Überzeugung* (degree of belief). Die Kontroverse zwischen beiden „Schulen“ wird in meinem Artikel „*Wahrscheinlichkeit und Information - wie Henne und Ei*“ (in diesem Abschlussbericht, S. 27) dargelegt. Dort wird auch Shannons Analyse zum Informationsgehalt einer Nachricht ausführlich besprochen.

[6] J. Willard Gibbs, *Elementary Principles in Statistical Mechanics, developed with especial reference to the rational foundation of thermodynamics*, 1902; Nachdruck: Dover Publications, New York 1960

- [7] John von Neumann, *Thermodynamik quantenmechanischer Gesamtheiten*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, S. 273–291, 1927; eine ausführlichere Darstellung findet man in John von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin 1932
- [8] Claude E. Shannon, *A mathematical theory of communication*, Bell. Syst. Tech. J. 27, 379–423, 623–656, 1948
- [9] James C. Maxwell, *Theory of Heat*, Longmans Green, London, 1872
- [10] Leo Szilard, *Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen*, Z. Phys. 53, 840–856, 1929
- [11] Leon Brillouin, *Science and Information Theory*, Academic Press, New York 1956; einen Überblick über die Literatur zum Thema findet man in *Maxwell's Demon: Entropy, Information, Computing*, edited by H. S. Leff and A. F. Rex, Princeton Univ. Press, Princeton, 1990
- [12] Eugen Fick und Günter Saueremann, *Quantenstatistik dynamischer Prozesse, Bd. 1 und 2*, Verlag Harry Deutsch, Thun, Frankfurt 1983
- [13] Edward T. Jaynes, *Information Theory and Statistical Mechanics*, Phys. Rev. 106, 620–630, 1957 und 108, 171–190, 1957
- [14] Roger Balian, *Information and statistical physics*, Studies in History and Philosophy of Modern Physics 36, 323–353, 2005; eine ausführlichere Diskussion findet man in Roger Balian, *From Microphysics to Macrophysics*, Vol. 1, Springer-Verlag, Berlin, 1991
- [15] Rolf Landauer, *Information is Physical*, Physics Today 44, 23–29, 1991; *The physical nature of information*, Physics Letters A 217, 188–193, 1996
- [16] Peter Janich, *Was ist Information?*, Suhrkamp Verlag, 2006
- [17] Manfred Eigen, *From Strange Simplicity to Complex Familiarity: A Treatise on Matter, Information, Life and Thought*, Oxford University Press, 2013
- [18] Die im Folgenden verwendeten Grundtatsachen der Quantenmechanik findet man in allen Lehrbüchern dazu, z. B. im historisch einflussreichen *The Principles of Quantum Mechanics* von Paul A. M. Dirac (Oxford University Press, 1930), das ein Vielzahl von Auflagen erlebt hat. Ein empfehlenswertes modernes Lehrbuch ist z. B. *Quantum Physics* von M. Le Bellac, Cambridge University Press, 2006

# Quanteninformation: Versuch einer Einführung für Nicht-Physiker

*Kurt Schönhammer*

*Institut für Theoretische Physik, Friedrich-Hund-Platz 1, D-37077 Göttingen*

## 1 Einleitung

In meinem Vortrag „Wahrscheinlichkeit und Information - wie Henne und Ei?“ (in diesem Band, ab S. 27) bin ich auf Claude Shannons fundamentale Arbeit zum Informationsgehalt einer Nachricht eingegangen. In dieser Arbeit stand das *praktische* Problem im Vordergrund, eine Nachricht so genau wie möglich von einem Ort an einen anderen zu transferieren. Ob sie tatsächlich eine *Bedeutung* hat oder nicht, spielte für Shannons Untersuchungen keine Rolle. Die spezielle physikalische Realisierung des Übertragungskanals für die Nachricht war für die Überlegungen ebenfalls unwesentlich. Als selbstverständlich wurde eine Beschreibung des Vorgangs im Rahmen der klassischen Physik angenommen. Es ist aber seit den zwanziger Jahren des letzten Jahrhunderts klar, dass (zumindest) auf atomarer Ebene die klassische Mechanik durch die Quantenmechanik ersetzt werden muss. Daher kann man die Frage stellen, ob sich zum Problemkreis Kommunikation und „Information“ neue Gesichtspunkte ergeben, wenn die Quantennatur des Systems der Informationsübertragung eine Rolle spielt.

Aus dieser Fragestellung ist in den letzten dreißig Jahren das neue Forschungsfeld „Quanteninformation“ entstanden. Im Vordergrund steht dabei die Idee, die „Merkwürdigkeiten“ der Quantentheorie für praktische Anwendungen zu nutzen. Dabei wird bereits von *Quantentechnologie* gesprochen. Neben der Quantenkryptographie und Quantenteleportation hat insbesondere die Idee der Quantencomputer breite Aufmerksamkeit auch außerhalb eines engeren Forscherkreises gefunden. Obwohl die praktische Realisierung von Quantencomputern erst in der Anfangsphase ist, sind die theoretischen Betrachtungen dazu bereits sehr weit entwickelt. Für *spezielle* Aufgaben, wie der Faktorisierung großer Zahlen oder dem Suchen in einer großen Datenmenge wurden Algorithmen entwickelt, deren Ausführung auf Quantencomputern *sehr* viel (exponentiell) schneller wäre, als die heutigen (für „klassische“ Computern entwickelten) Algorithmen für dasselbe Problem. Als Rechner für den täglichen Gebrauch werden Quantencomputer dagegen (noch?) nicht diskutiert.

Der Versuch, die neuen Ideen nur „mit Worten“ zu vermitteln, ohne die mathematischen Grundideen der Quantenmechanik anzudeuten, kann nicht wirklich tief gehen. Daher werde ich im Folgenden an einem möglichst einfachen Beispiel auch die mathematische Beschreibung darlegen.

Eine Vorbemerkung: Wir wissen, wie schwer das Konzept „Information“ zu fassen ist. Es ist daher nicht verwunderlich, dass eine präzise Definition des Begriffs „Quanteninformation“ ebenfalls nicht einfach ist. Man kann als ersten Versuch Quanteninformation als das definieren, womit sich die Wissenschaftler(innen) in diesem neuen Forschungsfeld beschäftigen. Klar ist zumindest, dass, wie bei Shannon, die Frage der *Bedeutung* von Quanteninformation nicht gestellt wird.

## 2 Mathematische Grundkonzepte der Quantentheorie

Wie in der Einleitung angedeutet, kann man die neuen Ideen des Forschungsfeldes Quanteninformation nicht angemessen vermitteln, wenn man es nur mit Worten versucht. Hier soll zunächst die mathematische Formulierung der Quantentheorie in einfachster Form dargestellt werden. Die wichtigen Konzepte der Überlagerung und Verschränkung von Zuständen verlangen ein zumindest rudimentäres Verständnis der Addition von Vektoren, wie sie z. B. im Physikunterricht in Schulen bei der Überlagerung von Kräften auftritt. Die (*Zustands*)-Vektoren der Quantentheorie sind meist Elemente von (*Hilbert*-)Räumen unendlicher Dimension, was eine Veranschaulichung erschwert. Das zentrale Grundkonzept der Überlagerung in der Beschreibung von Quantensystemen lässt sich aber bereits mit Hilfe der linearen Polarisation einzelner Photonen verdeutlichen, wobei die Begriffe „Polarisation“ und „Photon“ im folgenden Kapitel erläutert werden. Dazu benötigt man nur zweidimensionale Vektoren, wie bei der Addition zweier Kräfte in der Ebene, die sie aufspannen. Das Konzept der Verschränkung tritt erst bei „zusammengesetzten“ Quantensystemen auf. Es erfordert also die vorherige Beschreibung einzelner Quantensysteme. Bevor wir dies angehen, beginnen wir mit einem kurzen Rückblick auf den Physikunterricht an Schulen.

### Vektoren und ihre Überlagerung

Vektoren tauchen im Physikunterricht bei der Überlagerung von Kräften auf. Kräfte haben eine Stärke und eine Richtung. Die Richtung der Kraft stellt man durch einen Pfeil dar, dessen Länge die Stärke der Kraft angibt. Dieses geometrische Objekt nennt man einen Vektor  $\vec{F}$  und wir deuten das durch einen Pfeil über dem Buchstaben „F“ (für „Force“) an. Um die „Gesamtkraft“  $\vec{F}_G$  zu erhalten, muss man die Vektoren „addieren“. Für zwei Kräfte führt das auf das Konzept des Kräfteparallelogramms, das in Abbildung 1 dargestellt ist.

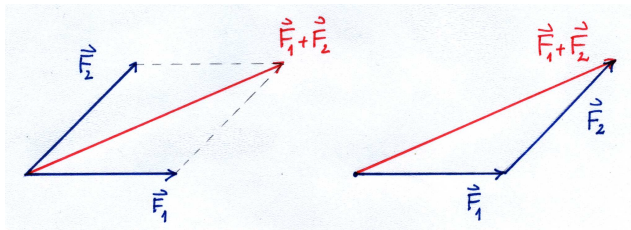


Abbildung 1. Addition zweier Vektoren

Ein wichtiges Beispiel für Kräfte sind diejenigen, die elektrisch geladene Teilchen aufeinander ausüben. Die Kraft  $\vec{F}$  auf eine „Testladung“  $q$  um ein geladenes Objekt wird beschrieben durch das elektrische Feld  $\vec{E}$ , das das Objekt am Ort der Testladung erzeugt:  $\vec{F} = q\vec{E}$ .

Ein zweites Beispiel aus der Mechanik, in der Vektoren als physikalische Größen mit Richtung und Betrag auftreten, ist die Drehung starrer Körper. Dort entspricht die Richtung der Orientierung der Drehachse und die Länge des „Vektors“  $\vec{\omega}$  der Winkelgeschwindigkeit“ ist durch die Drehgeschwindigkeit um die Drehachse gegeben.

Für den praktischen Umgang mit einem beliebigen (reellen) zweidimensionalen Vektor  $\vec{a}$  ist es nützlich, ihn als Summe von Vektoren in zueinander orthogonalen Richtungen, hier in Richtung der  $x$ - und  $y$ -Achsen darzustellen. Dabei wird aus dem Parallelogramm von Abb. 1 der Spezialfall eines Rechtecks:

$$\vec{a} = \vec{a}_x + \vec{a}_y = a_x \vec{e}_x + a_y \vec{e}_y.$$

Nach dem zweiten Gleichheitszeichen wird die Richtung der Vektoren  $\vec{a}_x$  und  $\vec{a}_y$  durch die sog. Einheitsvektoren  $\vec{e}_x$  und  $\vec{e}_y$  der Länge eins ausgedrückt. Die Komponenten  $a_x$  und  $a_y$  können dann beliebige (reelle) Zahlenwerte annehmen. Für positive  $a_x$  zeigt  $\vec{a}_x$  nach rechts, für negative  $a_x$  zeigt  $\vec{a}_x$  nach links. Solche Vektorzerlegungen werden in diesem Vortrag eine zentrale Rolle spielen.

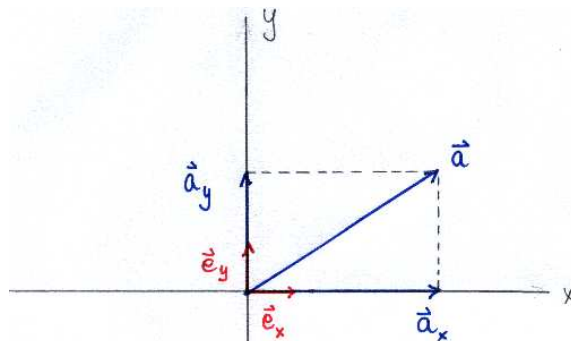


Abbildung 2. Zerlegung eines Vektors nach orthogonalen Einheitsvektoren

### 3 Welle-Teilchen Dualismus bei Licht

Gegen Ende des 19. Jahrhunderts hatte sich im Rahmen der *klassischen Elektrodynamik* die Beschreibung von Licht als Wellenphänomen durchgesetzt. Die mathematische Beschreibung erfolgt im Rahmen der „Maxwellschen Gleichungen“. Der alte Streit zur Wellen- oder Teilchennatur des Lichts schien aber nur für kurze Zeit als entschieden. Zu Beginn des 20. Jahrhunderts konnten neue Experimente nur im Teilchenbild verstanden werden, so dass man sich heute mit dem Dualismus in der Überschrift abgefunden hat. Die Beschreibung von Polarisierungseffekten in beiden Bildern liefert einen einfachen Zugang zu den Besonderheiten der Quantenphysik.

#### 3.1 Elektromagnetische Wellen und ihre Polarisation

Elektrische Felder sind ein zentraler Bestandteil elektromagnetischer Wellen, wie einer Lichtwelle. In einer (ebenen) Welle mit linearer Polarisation schwingt das elektrische Feld  $\vec{E}$  periodisch in Raum und Zeit senkrecht zur Ausbreitungsrichtung  $z$  der Welle (s. Abbildung 3). Senkrecht zu  $\vec{E}$  und der Ausbreitungsrichtung schwingt auch ein Magnetfeld, das wir im Folgenden aber außer Acht lassen können.

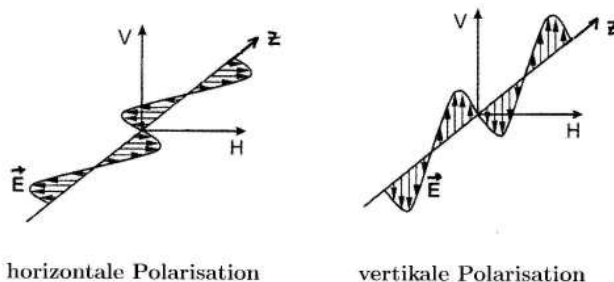


Abbildung 3. Elektrisches Feld  $\vec{E}$  einer horizontal ( $H$ ) bzw. vertikal ( $V$ ) polarisierten elektromagnetischen Welle, senkrecht zur Ausbreitungsrichtung  $z$ .

Wir zerlegen das elektrische Feld  $\vec{E}$  in seine Komponenten in den zur Ausbreitungsrichtung senkrechten Richtungen  $\vec{e}_V$  und  $\vec{e}_H$

$$\vec{E} = E_V \vec{e}_V + E_H \vec{e}_H,$$

wobei wir den eigentlich auftretenden, in Raum und Zeit periodischen Faktor weglassen haben, da er für die folgenden Betrachtungen keine Rolle spielt. Für den Fall, dass  $\vec{E}$  nur eine Komponente in  $V$ -Richtung hat ( $E_H = 0$ ), heißt die Welle „vertikal polarisiert“, für den Fall, dass  $\vec{E}$  nur eine Komponente in  $H$ -Richtung hat ( $E_V = 0$ ), heißt die Welle „horizontal polarisiert“. Eine linear polarisierte Welle



mit beliebigen (reellen) Werten von  $E_V$  und  $E_H$  kann mit Hilfe eines Polarisators in eine Welle mit Feldvektor in Richtung des Polarisators übergeführt werden.

Für einen Analysator in  $V$ -Richtung geht nur die  $V$ -Komponente des Feldes durch den Polarisator, d. h.

$$E_V \vec{e}_V + E_H \vec{e}_H \rightarrow E_V \vec{e}_V.$$

Dabei wird die Intensität der Welle abgeschwächt. Nach der klassischen Elektrodynamik ist der Bruchteil der durchgelassenen Energie durch  $E_V^2/(E_V^2 + E_H^2)$  gegeben.

Dies ist die Beschreibung der Ausbreitung von Licht im „Wellenbild“ der klassischen Elektrodynamik.

### 3.2 Photonen und ihr Polarisationszustand

Im Jahr 1905 erklärte Einstein den „photoelektrischen Effekt“ (Herausschlagen von Elektronen aus einer Metalloberfläche durch Bestrahlung mit „Licht“) im Teilchenbild. Der Lichtstrahl wird beschrieben durch sich mit Lichtgeschwindigkeit  $c$  ausbreitende „Teilchen“, den Photonen. Für ebene Wellen der (Kreis-) Frequenz  $\omega$  ist die Energie des Photons durch  $\hbar\omega$  gegeben. Dabei ist  $\hbar = h/(2\pi)$  mit  $h$  dem Planckschen Wirkungsquantum.

Die erfolgreiche Beschreibung des photoelektrischen Effekts im Teilchenbild legt es nahe, auch die (lineare) Polarisation von elektromagnetischen Wellen auf Photonen zu übertragen.

Alle Photonen des Strahls haben in diesem Bild dieselbe Polarisation, die durch  $\psi_{V(H)} = E_{V(H)}/\sqrt{E_V^2 + E_H^2}$  festgelegt ist. Nach Dirac ersetzen wir die Vektoren und Einheitsvektoren durch neue Symbole

$$\vec{e}_V \rightarrow |V\rangle, \quad \vec{e}_H \rightarrow |H\rangle.$$

Dann lautet der Vektor, der den „Polarisationszustand“ eines Photons beschreibt

$$|\psi\rangle = \psi_V |V\rangle + \psi_H |H\rangle,$$

mit  $\psi_V^2 + \psi_H^2 = 1$ . Im Folgenden werden wir insbesondere die Zustände  $|V'\rangle$  und  $|H'\rangle$  verwenden, in denen  $\psi_V$  und  $\psi_H$  dem Betrage nach gleich sind:

$$|V'\rangle = \frac{1}{\sqrt{2}}|V\rangle + \frac{1}{\sqrt{2}}|H\rangle, \quad |H'\rangle = -\frac{1}{\sqrt{2}}|V\rangle + \frac{1}{\sqrt{2}}|H\rangle.$$

Man erhält sie aus den Zuständen  $|V\rangle$  und  $|H\rangle$  durch Drehung um 45 Grad.

Seien nun  $\psi_V$  und  $\psi_H$  von Null verschieden, d. h. die (lineare) Polarisation sei bezüglich  $\vec{e}_V$  und  $\vec{e}_H$  verdreht. Anschließend trifft der Strahl von Photonen auf einen  $V$ -Analysator. Aus der Interpretation im Wellenbild wissen wir, dass nur der Bruchteil  $\psi_V^2$  der Energie durch den Analysator geht. Dies muss daher auch der Bruchteil der durchgehenden Photonen sein. Da aber alle Photonen dieselbe Polarisation haben, kann man für das einzelne Photon nicht vorhersagen, ob es im

Analysator durchgelassen wird. Es ist nur noch eine Wahrscheinlichkeitsaussage möglich:

$$w_V(|\psi\rangle) = \psi_V^2,$$

wobei  $w_V(|\psi\rangle)$  die Wahrscheinlichkeit (im Sinne von relativer Häufigkeit) für den Durchgang des Photons durch den  $V$ -Analysator ist. Damit sind wir bei einem wesentlichen Aspekt der Quantentheorie angelangt:

Nach der Präparation eines physikalischen Systems kann man über den Ausgang späterer Messungen im Allgemeinen nur Wahrscheinlichkeitsaussagen machen.

Zunächst zum Zusatz „im Allgemeinen“ in den Wahrscheinlichkeitsaussagen. In unserem Beispiel der Photonen-Polarisation gibt es zwei Spezialfälle, in denen definitive Aussagen möglich sind. Wir verdeutlichen das sowohl mit dem  $V$ -Analysator, als auch an Hand der in Abbildung 4 gezeigten Messapparatur („ $V$ - $H$ -Detektor“), in der durch Doppelbrechung in einem Kalzitkristall vertikal und horizontal polarisierte Photonen *räumlich getrennt* werden.

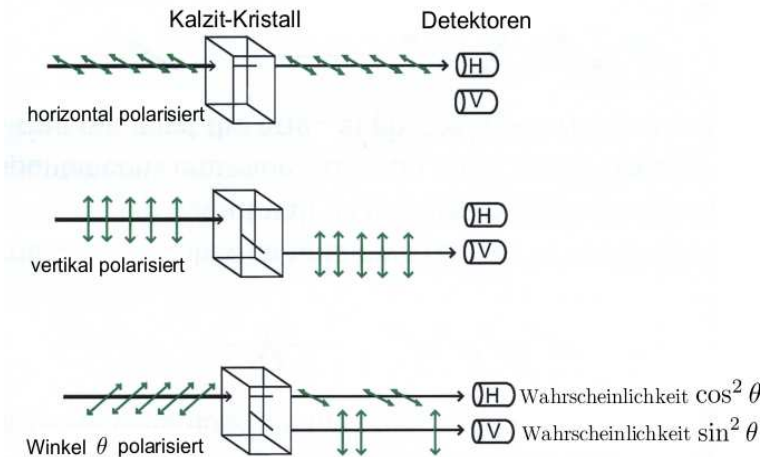


Abbildung 4. Detektor der  $V$ - oder  $H$ -Polarisation der Photonen abfragt. Im unteren Teil der Abbildung sind die einlaufenden Photonen um einen Winkel  $\theta$  relativ zur horizontalen Richtung verdreht (s. auch Anmerkung dazu in Kap. 5.1).

1) Der einlaufende Strahl hat horizontale Polarisation, d. h.  $\psi_V = 0, \psi_H = \pm 1$ . Dann ist  $w_V(|\psi\rangle) = 0$ , d. h. kein Photon geht durch den  $V$ -Analysator. Im  $V$ - $H$ -Detektor wird das Photon mit Sicherheit im  $H$ -Detektor registriert.

2) Der einlaufende Strahl hat vertikale Polarisation, d. h.  $\psi_V = \pm 1, \psi_H = 0$ . Dann ist  $w_V(|\psi\rangle) = 1$ , d. h. jedes Photon geht durch den  $V$ -Analysator. Im  $V$ - $H$ -Detektor wird das Photon mit Sicherheit im  $V$ -Detektor registriert.

Diese beiden Spezialfälle entsprechen den ersten beiden Beispielen in Abb. 4. Den (eigentlich vergeblichen) Versuch, den allgemeinen Fall darzustellen, in dem nur Wahrscheinlichkeitsaussagen möglich sind, zeigt der untere Teil der Abbildung.

## 4 Quanteninformation: einfache Systeme

Bis in die neunziger Jahre des letzten Jahrhunderts war die im letzten Abschnitt verwendete Notation die übliche bei einer Einführung in die Quantenmechanik. In diesem Kapitel wird zunächst für die Polarisation von Photonen eine neue Notation und Begriffsbildung eingeführt. Sie hat sich schnell im Forschungsgebiet der Quanteninformation durchgesetzt. Als einfachste Anwendung der entwickelten Ideen wird die Quantenkryptographie beschrieben. Anschließend werden als wichtige andere mögliche Realisierung von „Qubits“ Spin-1/2 Teilchen vorgestellt.

### 4.1 Quanten-Bits (Qubits, oder Qbits)

Wir nehmen eine weitere Umbenennung unserer Basisvektoren vor

$$|V\rangle \rightarrow |0\rangle, \quad |H\rangle \rightarrow |1\rangle,$$

und analog  $|V'\rangle \rightarrow |0'\rangle$ ,  $|H'\rangle \rightarrow |1'\rangle$ . Dann lautet ein allgemeiner Zustandsvektor für die Polarisation des Photons

$$|\psi\rangle = \psi_V|0\rangle + \psi_H|1\rangle.$$

Man nennt solch eine quantenmechanische Überlagerung, in der „die Null“ (präziser der Zustand  $|0\rangle$ ) mit der Wahrscheinlichkeit  $\psi_V^2$  und „die Eins“ (präziser der Zustand  $|1\rangle$ ) mit der Wahrscheinlichkeit  $\psi_H^2$  vorliegt, seit ca. fünfzehn Jahren ein *Qubit* [1]. Der Begriff hat schnell Eingang in moderne Lehrbücher der Quantenmechanik gefunden [2,3,4].

Der Übergang „Präparation  $\rightarrow$  Messung“ eines Quantensystems erinnert an die Shannonsche Kommunikation „Sender  $\rightarrow$  Empfänger“, in der eine (klassische) Bit-Sequenz übertragen wird. Vor der Messung des einlaufenden Photons steckt die Information über dessen Polarisation im quantenmechanischen Zustand  $|\psi\rangle$ . Diese „Quanteninformation“ wird in Abbildung 5 durch die von der Präparation zur Messung laufende gezackte Linie dargestellt. Das Ablesen an der Messapparatur erzeugt „klassische Information“.

Unter Verwendung der beiden erwähnten Spezialfälle können wir die Polarisation von Photonen nun zunächst verwenden, um im Shannonschen Sinne eine beliebige Bit-Folge: 01001011... zu übertragen. In der Literatur zur Quanteninformation heißt die Absenderin der Nachricht Alice, und der Empfänger Bob. Alice und Bob werden meist als räumlich weit getrennt angenommen.

Eine einfache Methode der Übertragung, die sich eng an eine „klassische“ Art anlehnt, besteht darin, dass Alice ein vertikal polarisiertes Photon für jede Null,

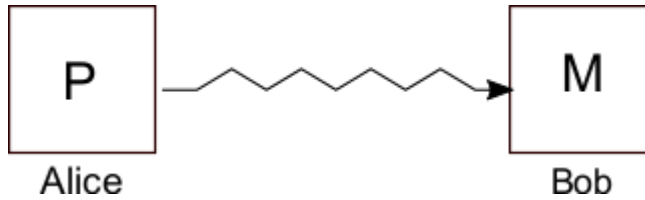


Abbildung 5. Nach der Präparation schickt Alice „Quanteninformation“ (gezackte Linie) an Bob.

und ein horizontal polarisiertes Photon für jede Eins an Bob sendet. Also für obiges Beispiel

$$\begin{array}{cccccccc}
 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
 V & H & V & V & H & V & H & H \\
 |0\rangle & |1\rangle & |0\rangle & |0\rangle & |1\rangle & |0\rangle & |1\rangle & |1\rangle
 \end{array}$$

In der dritten Zeile ist der Zustand des übertragenen Photons angezeigt.

Bob kennt diese Abmachung und kann die gesendete Bit-Folge mit Hilfe seines „V-H-Detektors“ (Abb.4) perfekt rekonstruieren. Er schreibt eine 0, wenn sein V-Zähler anspricht und eine 1, wenn der H-Zähler „klickt“.

Völlig anders sieht es aus, wenn Bob nichts über die (lineare) Polarisation des einlaufenden Photons weiß. Ist dessen Polarisation um den Winkel  $\theta$  bezüglich der horizontalen Richtung verdreht (Abb. 4), so gilt  $w_V(|\psi\rangle) = \psi_V^2 = \sin^2 \theta$ . Für den Drehwinkel 45 Grad liefert das für (V und (H))

$$w_{V(H)}(|V'\rangle) = 1/2 = w_{V(H)}(|H'\rangle).$$

Durch Messung an *einem* Photon ist der Winkel *nicht* bestimmbar. Außerdem gilt:

Der Messprozess hat den ursprünglichen Zustand des Photons zerstört.

Es muss ein ganzer Strahl identischer Photonen präpariert werden, um aus der relativen Häufigkeit des Durchgangs auf den Winkel  $\theta$  zu schließen. Wenn aber nur ein Photon eintrifft, müsste man seinen Zustand mit dem in Abbildung 6 skizzierten Apparat (mehrfach) kopieren („klonen“), um dann mit den geklonten Photonen eine „Reihenuntersuchung“ durchzuführen.

Man kann aber allgemein zeigen [2,3,4]:

Das (perfekte) Klonen eines unbekanntenen Quantenzustands ist unmöglich.

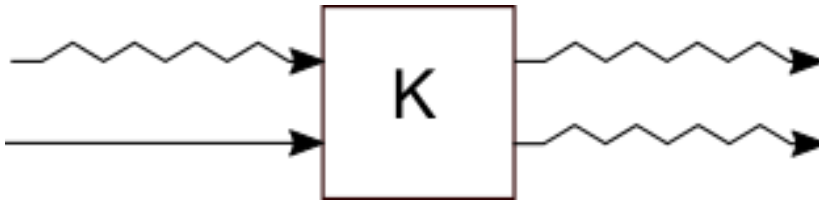


Abbildung 6. Klonen eines Quantenzustandes mit unbekannter Polarisation (gezackte Linie) mit Hilfe eines Zustands beliebiger Polarisation? (gerade Linie)

## 4.2 Quantenkryptographie

Die Tatsache, dass der Versuch, einen Zustand  $|\psi\rangle$  mit unbekannter Richtung der linearen Polarisation durch die (einmalige) Messung am  $V$ -Analysator oder mit dem  $V$ - $H$ -Detektor zu charakterisieren, gescheitert ist, stellt sich für die Quantenkryptographie als nützlich heraus. Geschickt eingesetzt, erlaubt diese Tatsache eine Methode um festzustellen, ob jemand („Eve“) bei der Übertragung der Nachricht „mitgehört“ hat (*eavesdropping*).

Bei der bisher diskutierten „klassischen“ Übertragungsmethode kann Eve unbemerkt mithören, wenn sie die beschriebene Spielregel zwischen Alice und Bob kennt. Sie schiebt ihren  $V$ - $H$ -Detektor (s. Abb. 4) in die Übertragungslinie und notiert eine 0, wenn ein vertikales Photon registriert wurde. Anschließend sendet sie ein Photon im Zustand  $|0\rangle$  an Bob weiter. Spricht ihr  $H$ -Detektor an, so notiert sie eine 1 und ein Photon im Zustand  $|1\rangle$  wird an Bob gesandt. Bob erhält so dieselbe Sequenz, die Alice gesendet hat und Eve hat unbemerkt die Bit-Sequenz registriert.

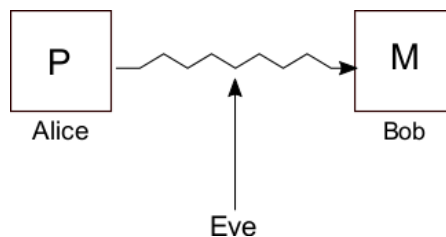


Abbildung 7. Eve versucht die Quanteninformation „anzuzapfen“.

Also muss die Informationsübertragung von Alice zu Bob raffinierter gestaltet werden. Eve muss im Unklaren darüber gelassen werden, welche Polarisation Alice verwendet. Ein einfaches Verfahren [5], das „BB84-Protokoll“ genannt wird, besteht darin, zur Übertragung einer 0 einen  $V$ -Polarisator oder einen um 45 Grad gedrehten  $V'$ -Polarisator zu verwenden, wobei die Auswahl zufällig erfolgt. Analog verwendet Alice für die Übertragung einer 1 zufällig einen  $H$ -Polarisator oder

einen  $H'$ -Polarisator. Die Übertragung der vorher verwendeten Sequenz sieht dann z. B. so aus (die „Striche“ in der zweiten Zeile sind zufällig)

$$\begin{array}{cccccccc}
 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
 V & H & V' & V & H' & V & H' & H \\
 |0\rangle & |1\rangle & |0'\rangle & |0\rangle & |1'\rangle & |0\rangle & |1'\rangle & |1\rangle
 \end{array}$$

Bob verwendet zur Analyse der gesendeten Photonen ebenfalls völlig zufällig einen  $V$ - $H$ - oder einen um 45 Grad gedrehten  $V'$ - $H'$ -Detektor. Immer wenn er (zufällig) das „richtige“ Messgerät verwendet hat, erhält er mit Sicherheit das von Alice gesendete Bit, falls sich Eve nicht eingeschaltet hat. Hier unter den Einstellungen von Alice die von Bob gewählte Folge von Detektoren und deren Wirkung auf die von Alice geschickten Zustände:

$$\begin{array}{cccccccc}
 V & H & V' & V & H' & V & H' & H \\
 VH & V'H' & V'H' & VH & VH & V'H' & VH & VH \\
 |0\rangle & |1\rangle & |0'\rangle & |0\rangle & |1'\rangle & |0\rangle & |1'\rangle & |1\rangle \\
 0 & - & 0 & 0 & - & - & - & 1
 \end{array}$$

wobei „-“ bedeutet, dass (gleichwahrscheinlich) eine 0 oder eine 1 gemessen wurde. Da diese „unsicheren Fälle“ eliminiert werden (s. u.), sind die von Bob tatsächlich gemessenen Werte nicht angegeben.

Da Bob seine Analysatorwahl zufällig getroffen hat, stimmt sie mit der von Alice im Mittel nur in der Hälfte aller Fälle überein. Hat Bob nicht dieselbe Richtung wie Alice gewählt, so ist sein Messergebnis zufällig. Die mögliche Übereinstimmung des registrierten Bits wäre ebenfalls zufällig. Um diese Fälle zu eliminieren, gibt Bob die Sequenz der von ihm verwendeten Analysatorrichtungen nach abgeschlossener Messung öffentlich bekannt. Daraufhin teilt Alice ihm mit, in welchen Fällen sie für ihren Polarisator dieselbe Wahl getroffen hat. In all diesen Fällen stimmen beide Resultate überein und der Rest der Sequenz wird eliminiert (-). Die Bekanntgabe von Bobs Analysatorsequenz schadet nicht, da die gemessenen Werte nicht öffentlich bekannt gemacht werden.

Bei der geschilderten Übertragungsmethode mit der zufälligen Wahl der Präparation der Apparate bei der Erzeugung und Messung kommt die Besonderheit der Quantennatur der Photonen derart zum Tragen, dass nun der Versuch von Eve, die Bit-Sequenz unbemerkt zu registrieren, scheitert.

Wir nehmen an, dass Eve die neuen Spielregeln kennt. Daher rüstet sie sich selbst mit einem  $V$ - $H$ - und einen  $V'$ - $H'$ -Detektor aus, kann aber selbst nur hoffen, jeweils die „richtige“ Wahl (d. h. die von Alice) zu treffen. Jedes Mal, wenn ihr das zufällig gelingt, registriert sie das richtige Bit und schickt dasselbe Qubit an Bob weiter, das Alice geschickt hat. In der anderen Hälfte der Fälle schickt sie nur mit Wahrscheinlichkeit  $1/2$  das richtige Photon an Bob. Also wird verglichen mit der richtigen Sequenz, die ohne das Eindringen von Eve übertragen wird, eine Fehlerrate von 25 Prozent vorliegen. Dies können Alice und Bob feststellen, wenn Sie einen zufälligen Teilabschnitt ihrer (in Wirklichkeit sehr viel längeren) Sequenz „öffentlich“ vergleichen und dann nicht mehr berücksichtigen. Stellen Alice und

Bob fest, dass die verglichene Teilsequenz identisch ist, so schließen sie daraus, dass niemand mitgehört hat. Die verbleibende Bit-Folge verwenden sie anschließend als „Schlüssel“ zur Übertragung einer anderen Nachricht.

Bei dieser Diskussion wurde idealisierend angenommen, dass keine Übertragungsfehler durch klassisches Rauschen auftreten.

Es sind bereits kommerzielle „Quantenkryptographen“ der Firmen ID-Quantique und MagiQ Technologies auf dem Markt, wobei das verwendete Protokoll aber geheim ist.

### 4.3 Spin-1/2-Teilchen als Qubits

Eine weitere wichtige Realisierung von Qubits, d. h. quantenmechanischer „Zweizustandssysteme“ erhält man durch den Spin 1/2 eines Elektrons oder (spezieller) Atomkerne. Der Begriff „Spin“ ist auch Nichtphysikern aus der Kernspin-Tomographie bekannt, die jetzt meist als „MRT“ (Magnet-Resonanz-Tomographie) bezeichnet wird. In einem (unzureichenden) klassischen Bild stellt man sich das Elektron als kleine Kugel vor, die um eine Achse in Richtung des Einheitsvektors  $\vec{n}$  rotiert.

Wie bei der Einführung von Vektoren erwähnt, liefert solch eine Drehbewegung einen Drehimpuls, der in Richtung von  $\vec{n}$  zeigt und für Spin-1/2 Teilchen den Wert  $\hbar/2$  hat. Misst man an diesem „Spinzustand“ die Drehimpulskomponente bezüglich irgendeiner anderen Raumrichtung, so erhält man entweder den Wert  $\hbar/2$  oder  $-\hbar/2$ , wobei die Wahrscheinlichkeit vom Winkel zwischen den beiden Richtungen abhängt [3,4,5]. Im zweidimensionalen Raum der Spinzustände kann man aus Überlagerung von

$$|\vec{e}_z\rangle \equiv |\uparrow\rangle \rightarrow |0\rangle, \quad |-\vec{e}_z\rangle \equiv |\downarrow\rangle \rightarrow |1\rangle$$

wieder Qubits formen [6].

## 5 Quanteninformation: komplexe Systeme

Während das Phänomen der Quantenkryptographie, wie gezeigt, mit Hilfe *einzelner* Photonen erklärt werden kann, benötigt man für kompliziertere Formen der „Quantentechnologie“ zusammengesetzte Systeme. Aber bereits bei der Beschreibung der Quantenkryptographie haben wir – stillschweigend – ein zusammengesetztes System verwendet, das Photon und den Messapparat.

### 5.1 Zusammengesetzte Systeme – Verschränkung

Um die (immer noch etwas kontroverse) Diskussion des quantenmechanischen Messprozesses zu vermeiden, betrachten wir als Teilsysteme jeweils den Spin eines Elektrons. Ein Beispiel sind zwei räumlich getrennte Wasserstoffatome in ihrem Grundzustand. Dabei können die Spins der Elektronen (im räumlichen 1s-Zustand) in den beiden Atomen in beliebige Richtungen  $\vec{n}_1$  und  $\vec{n}_2$  zeigen.

Wie beschreibt man nun ein aus zwei Teilsystemen zusammengesetztes „Gesamt-system“(G)? Besteht zwischen den Systemen keinerlei Korrelation (wie zwischen den Spins in den räumlich auch zu allen früheren Zeiten getrennten Atomen), so sollten die Wahrscheinlichkeit am System 1 den Wert  $a_1$  einer physikalischen Größe und am System 2 den Wert  $b_2$  einer physikalischen Größe dieses Systems zu messen, durch das Produkt der einzelnen Wahrscheinlichkeiten gegeben sein:  $p_{a_1, b_2} = p_{a_1} p_{b_2}$ . Dies wird durch eine Produktform des Gesamtzustandes erreicht. Man schreibt die Zustände einfach „nebeneinander“

$$|\vec{n}_1, \vec{n}_2\rangle_G = |\vec{n}_1\rangle_1 |\vec{n}_2\rangle_2.$$

Etwas mathematischer schreibt man noch ein  $\otimes$  zwischen die Zustände, was man als „Tensorprodukt“ bezeichnet.

Obigen Produktzustand fasst man auf als abstrakten Vektor in einem Raum, dessen Dimension durch das Produkt der Dimensionen der Zustandsräume von System 1 und System 2 (hier jeweils 2) gegeben ist. Im Beispiel der zwei Spins ist der Gesamt-Zustandsraum also vierdimensional. Aus diesen abstrakten Zustandsvektoren des Gesamtsystems kann man wieder lineare Überlagerungen bilden. Sie lassen sich aber im Gegensatz zur Vektoraddition bei den klassischen Kräften nicht bildlich darstellen.

Stehen die beiden Teilsysteme in Wechselwirkung, so treten dadurch üblicherweise Überlagerungen auf. Im Grundzustand eines Wasserstoff-Moleküls  $H_2$  befinden sich die beiden Spins im (sog. Spin-Singulett-) Zustand

$$|\text{Singulett}\rangle_G = |\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2,$$

wobei wir den Normierungsfaktor  $(1/\sqrt{2})$  weggelassen haben. Dieser Zustand des Gesamtsystems lässt sich *nicht* als Produkt aus Zuständen von System 1 und System 2 darstellen. Nach Schrödinger [7] nennt man solche Zustände *verschränkt*.

Es sollte angemerkt werden, dass das Phänomen der Verschränkung bereits im Zusammenhang mit dem  $V$ - $H$ -Detektor auftritt. Durch die Doppelbrechung im Kalzitkristall wird der „Ortszustand“ des Photons mit seinem Polarisationszustand verschränkt. In Abb. 4, unten, ist der eigentlich vergebliche Versuch gemacht, diese Verschränkung durch ein klassisches Bild darzustellen.

## 5.2 Das EPR-Paradoxon und Schrödingers Katze

Der Zustand  $|\text{Singulett}\rangle_G$  hat Eigenschaften, die unserem klassischen Vorstellungsvermögen völlig fremd sind. Wäre nur der erste Term  $|\uparrow\rangle_1 |\downarrow\rangle_2$  vorhanden, so besagt dieser Produktzustand, dass Spin 1 in Richtung  $\vec{e}_z$  zeigt, und Spin 2 in der umgekehrten Richtung  $-\vec{e}_z$ . Für den Produktzustand  $|\downarrow\rangle_1 |\uparrow\rangle_2$  sind die Richtungen der beiden Spins umgekehrt. Für beide Produktzustände sind die Eigenschaften der beiden Teilsysteme festgelegt. Für den obigen verschränkten Zustand gilt das nicht mehr.



Mit Hilfe der Eigenschaften der einzelnen Zustände  $|\vec{n}\rangle_{1,2}$  lässt sich (als theoretischer Physiker) leicht zeigen, dass

$$|\uparrow\rangle_1|\downarrow\rangle_2 - |\downarrow\rangle_1|\uparrow\rangle_2 = |\vec{m}\rangle_1|-\vec{m}\rangle_2 - |-\vec{m}\rangle_1|\vec{m}\rangle_2$$

gilt, wobei  $\vec{m}$  in eine *beliebige* andere Raumrichtung zeigen kann. Das bedeutet:

Während man über das Gesamtsystem eine vollständige Aussage hat, weiß man über die Spinrichtung der Teilsysteme nichts.

Misst man den Zustand von Spin 1 bezüglich der beliebigen Richtung  $\vec{m}$ , so zeigt er mit Wahrscheinlichkeit  $1/2$  in dieser Richtung und mit Wahrscheinlichkeit  $1/2$  in umgekehrter Richtung. Hat man z. B. festgestellt, dass Spin 1 in Richtung  $\vec{m}$  zeigt, so weiß man, dass Spin 2 in Richtung  $-\vec{m}$  zeigt. Diese perfekte Korrelation bleibt auch erhalten, wenn die beiden Spins nach der Präparation des Zustandes  $|\text{Singulett}\rangle_G$  räumlich getrennt werden. Dies hielten Einstein, Podolski und Rosen (1935) für ein „Problem“ der Quantenmechanik, das unter dem Namen EPR-Paradoxon [8] in die Literatur eingegangen ist. Seit der Entwicklung des Forschungsgebietes Quanteninformation hat man realisiert, dass man diese nicht-klassischen Korrelationen „quantentechnologisch“ (z. B. in der *Quantenteleportation* [9]) ausnutzen kann.

In seinem Artikel, in dem erstmals der Begriff der Verschränkung auftritt, gibt Schrödinger ein sehr bekannt gewordenes Beispiel dafür, das heutzutage *Schrödingers Katze* genannt wird:

„Man kann auch ganz burleske Fälle konstruieren. Eine Katze wird in eine Stahlkammer gesperrt, zusammen mit folgender Höllenmaschine (die man gegen den direkten Zugriff der Katze schützen muss): in einem GEIGERschen Zählrohr befindet sich eine winzige Menge radioaktiver Substanz, *so* wenig, dass im Verlauf einer Stunde *vielleicht* eines von den Atomen zerfällt, ebenso wahrscheinlich aber auch keines; geschieht es, so spricht das Zählrohr an und betätigt über ein Relais ein Hämmerchen, das ein Kölbchen mit Blausäure zertrümmert. Hat man dieses ganze System eine Stunde lang sich selbst überlassen, so wird man sich sagen, dass die Katze noch lebt, *wenn* inzwischen kein Atom zerfallen ist. Der erste Atomzerfall würde sie vergiftet haben. Die  $\psi$ -Funktion des ganzen Systems würde das so zum Ausdruck bringen, dass die in ihr lebende und die tote Katze zu gleichen Teilen gemischt oder verschmiert sind.“ [7]

In analoger Formulierung zu den zwei Spins würde man also für den zeitabhängigen Zustand des Gesamtsystems schreiben

$$|\psi(t)\rangle_G = c_0(t)|\text{nicht zerfallen}\rangle_A|\text{Katze lebt}\rangle_K + c_1(t)|\text{zerfallen}\rangle_A|\text{Katze tot}\rangle_K.$$

Dabei gilt für die Koeffizienten  $c_0(t=0) = 1$  und  $c_1(t=0) = 0$ , d. h. die Katze lebt im Anfangszustand. Wenn ca. eine Stunde vergangen ist, sind  $c_0$  und  $c_1$  ungefähr gleich, was Schrödingers „gemischt oder verschmiert“ entspricht. Nach einigen Tagen ist das „Atom“ (fast) sicher zerfallen und es gilt  $c_0 \approx 0$  und  $c_1 \approx 1$ , d. h. die Katze ist (fast) sicher tot. Aus dem Produktzustand zur Zeit  $t = 0$

entwickelt sich ein stark verschränkter Zustand, der nach langer Zeit in den anderen Produktzustand übergeht.

Die gezielte Präparation von verschränkten Zuständen ist ein schwieriges experimentelles Problem. Und wenn sie gelingt, muss sichergestellt sein, dass das präparierte System gut gegenüber der „Umwelt“ isoliert ist. Sonst tritt das Phänomen der *Dekohärenz* auf, das den verschränkten Zustand im Lauf der Zeit unwiederbringlich zerstört [1, 2, 3].

### 5.3 Auf dem Weg zum Quantencomputer

Die nur vom Standpunkt der Theorie bereits weit entwickelten *Quantencomputer* (die experimentelle Realisierung hinkt weit hinterher (s. o.)) sind Systeme aus  $M \gg 1$  verschränkten Qubit-Zuständen. Also muss man mehrfache Tensorprodukte aus den einzelnen Qubit-Zuständen bilden. Die Dimension des Raumes der Zustände des Gesamtsystems, also der möglichen verschiedenen Produktzustände ist  $2^M$ . Der Produktzustand, der für  $M = 8$  der im Text häufig verwendeten Bit-Sequenz entspricht, lautet

$$|0, 1, 0, 0, 1, 0, 1, 1\rangle = |0\rangle|1\rangle|0\rangle|0\rangle|1\rangle|0\rangle|1\rangle|1\rangle.$$

Es gibt  $255 = 2^8 - 1$  weitere solcher Produktzustände. Allgemeine Überlagerungen dieser Produktzustände sind wieder verschränkt.

Der Rechengang eines Quantencomputers besteht aus drei Teilschritten:

#### *Präparation, Manipulation und Messung*

des Quantensystems „Computer“. Die Besonderheit eines Quantencomputers ist, dass bei der *Manipulation* (Wirkung unitärer Operatoren) auf alle Komponenten des Zustandsvektors gleichzeitig eingewirkt wird. Darauf genauer einzugehen, wäre der Inhalt eines weiteren Vortrages.

## Literaturangaben und Anmerkungen:

- [1] B. W. Schumacher, *Quantum coding*, Phys. Rev. A 51, 2738–2747, 1995
- [2] M. Le Bellac, *Quantum Physics*, Cambridge Univ. Press, 2006
- [3] S. Haroche, JM. Raimond, *Exploring the Quantum*, Oxford Univ. Press, 2006
- [4] B. W. Schumacher, M. Westmoreland, *Quantum Processes, Systems, and Information*, Cambridge Univ. Press, 2010
- [5] C. H. Bennett and G. Brassard, *Quantum Cryptography: Public Key Distribution and Coin Tossing*, Proceedings of the IEEE International Conference on Computer Systems and Signal Processing, Bangalore, 1984

- [6] Allgemeine Zustände im zweidimensionalen (unitären) „Spinraum“ erhält man, wenn man komplexe Zahlen als Entwicklungskoeffizienten zulässt. Dasselbe gilt auch für allgemeine Polarisationszustände von Photonen („elliptische Polarisati-on), z. B. für *zirkulare* Polarisati-on).
- [7] E. Schrödinger, *Die gegenwärtige Situation in der Quantenmechanik*, Naturwissenschaften 23, 807–812, 1935
- [8] A. Einstein, B. Podolski, and N. Rosen, *Can Quantum Mechanical Description of Physical Reality be Considered Complete?*, Physical Review 47, 777–780, 1935
- [9] C. H. Bennett, et al. *Teleporting an unknown quantum state via dual classical and Einstein-Podolski-Rosen channels*, Physical Review Letters 70, 1895–1899, 1993



# Informationsverarbeitung in der Kognitionspsychologie und in den kognitiven Neurowissenschaften

Gerd Lüer<sup>1</sup>, und Uta Lass<sup>2</sup>

<sup>1</sup>*Georg-Elias-Müller-Institut für Psychologie, Fakultät für Biologie und  
Psychologie, Goßlerstraße 14, 37073 Göttingen*

<sup>2</sup>*Courant Forschungszentrum „Textstrukturen“, Nikolausberger Weg 23, 37073  
Göttingen*

## 1 Paradigmenwechsel vom Behaviorismus zur Psychologie der Informationsverarbeitung

Seit Mitte des 20. Jahrhunderts ist in der Psychologie allmählich die Überzeugung gewachsen, dass psychische Prozesse am besten durch theoretische Konstrukte beschrieben und erklärt werden können, die auf der Grundlage von Informationsverarbeitungsprozessen konzipiert werden. Eine solche Vorstellung ist das Ergebnis eines radikalen Paradigmenwechsels, der sich vor mehr als einem halben Jahrhundert vollzogen hat und auf den zunächst des Verständnisses wegen eingegangen werden muss.

Mit einem Paradigma kristallisiert sich eine unter Wissenschaftlern allgemein anerkannte Art und Weise heraus, wie eine zu untersuchende wissenschaftliche Problemstellung am besten und am erfolgreichsten anzugehen ist. In der Kognitionspsychologie wird diese Frage mit dem Verweis auf das Informationsverarbeitungsparadigma beantwortet. Danach sind alle höheren geistigen Leistungen, über die Menschen verfügen, wie z. B. das Entscheiden, Urteilen und Problemlösen, die Aufmerksamkeit und der Wissenserwerb sowie die Gedächtnisnutzung am besten durch Prozesse der Informationsverarbeitung zu beschreiben und zu erklären. Solche Informationsverarbeitungsprozesse beinhalten die Aufnahme und das Sammeln von Daten aus der Umwelt, das Speichern, Verändern und Interpretieren dieser Daten, um nur einige dieser Prozesse zu nennen.

### 1.1 *S-R*-Paradigma des Behaviorismus

Das beschriebene Verständnis von kognitiven Leistungen als basierend auf Informationsverarbeitungsprozessen steht im krassen Gegensatz zum Behaviorismus, der mit dem Paradigma von Reiz-Reaktionsverbindungen seine Blütezeit Mit-

te des 20. Jahrhunderts erlebte und die damalige Psychologie beherrschte. Die Vorstellung, dass allein über die Beschäftigung mit auslösenden Reizen und den dadurch geradezu mechanisch ausgelösten Reaktionen psychologische Forschung betrieben werden könne, war durch den experimentellen Nachweis sogenannter bedingter Reaktionen 1927 von Pawlow [35] genährt worden, dem für diese Entdeckung im Jahre 1904 der Nobelpreis verliehen wurde. Zusätzlich gewann dieses Reiz-Reaktions-Paradigma seine Attraktivität durch jeglichen Verzicht auf mehr oder weniger mystische Zusatzannahmen. Eine derartige Psychologie ließ sich allein mit naturwissenschaftlichen Methoden betreiben wie dem Experiment und den daran anschließenden statistischen Datenanalysen. Die von den damaligen wissenschaftlich arbeitenden Psychologen erwünschte Trennung von Philosophie und Psychologie konnte so noch konsequenter und endgültiger erreicht werden.

Gegründet wurde der Behaviorismus in den USA. Sein rigorosester Vertreter war John Broadus Watson (1878-1958). Er veröffentlichte im Jahre 1913 ein sogenanntes behavioristisches Manifest [52], in dem Psychologie als eine Wissenschaft des Verhaltens definiert wurde. Verhalten wird durch Reizung des Organismus produziert. Aufgabe psychologischer Forschung ist es nach dieser Sichtweise allein, beobachtbares und messbares Verhalten zu beschreiben, zu prognostizieren und durch die Variation von auslösenden Reizen zu verändern und zu erklären. Dieses geschieht unter vollkommener Auslassung mentalistischer Begriffe wie Empfindungen, Vorstellungen, Bewusstsein u. a. Man spricht deshalb in diesem Falle auch von einer reinen *S-R*-Psychologie oder auch Black-Box-Psychologie, wobei *S* den Stimulus oder Reiz und *R* die Reaktion symbolisieren und der Organismus als Black Box und mechanischer Reaktionsgeber nicht weiter untersucht wird.

Der Behaviorismus hat eine enorm große Flut experimenteller Forschungsansätze in der Psychologie in Gang gesetzt. Hauptthema war das Lernen, das man vorzugsweise an Tieren wie Ratten und Tauben studierte, ganz im Einklang mit der Auffassung, dass geistige Prozesse keine relevante Rolle beim Zustandekommen einer Leistung spielen.

Wie der Philosoph und Wissenschaftstheoretiker Thomas Kuhn [18] es 1967 beschrieben hat, erfahren Paradigmen Veränderungen, weil neue Forschungsergebnisse sie nicht mehr in Gänze unterstützen. Dieses Schicksal erfuhr auch der Behaviorismus. Besonders kritisch wurde die Annahme diskutiert, dass zwischen Reiz und Reaktion nur eine Mechanik angenommen wurde, die für den Organismus stand. Deshalb wurde im Neo-Behaviorismus die *S-R*-Formel zur Formel *S-O-R* erweitert, wobei das *O* für Organismus steht. In diesem Zusammenhang sprach man zunächst noch vorsichtig von *intraorganic reactions*, dann folgten aber Begriffe wie *intervening variables* als Bezeichnung für die Tätigkeit des mechanischen Organismus. Besonders bemerkenswert – weil nicht vereinbar mit den behavioristischen Grundannahmen – ist der Begriff *cognitive map*, den Tolman 1948 [49] prägte und aus Rattenexperimenten entwickelte. Er konnte zeigen, dass sich diese Tiere die Wege durch ein Labyrinth allein durch ihre Erkundungsversuche sehr gut eingeprägt hatten. In ihrem Verhalten wurde das Gelernte und im Gedächtnis Behaltene allerdings erst sichtbar, wenn es einen entsprechenden Anreiz gab, also Futter beim Erreichen der Zielbox zu finden war. Damit war die Annahme

von komplexen Repräsentationen im Organismus in der Form eines Gedächtnisses geboren – eine in der Tat ganz nichtbehavioristische Sichtweise. So bewahrheitete sich Kuhns Vorhersage, dass Paradigmen – eben auch dem Behaviorismus – eine begrenzte Lebenszeit beschert ist, weil immer neue Forschungsergebnisse die Grundannahmen Stück für Stück invalidieren und damit die Erklärungs- und Vorhersagekraft des Paradigmas schwächen.

Mit dem Konzept von *cognitive maps* war ein wichtiger Schritt in Richtung Informationsverarbeitungsprozesse für höhere geistige Leistungen getan. Es hat jedoch schon in der Blütezeit des Behaviorismus außerhalb der USA ebenfalls solche Forschungsansätze gegeben, die als Vorläufer des Informationsverarbeitungsparadigmas verstanden werden können. Dazu gehört das von dem Engländer Frederic Bartlett (1932) [5] konzipierte Schemakzept, das er als Begriff von dem Neurologen Henry Head (nach Baddeley [4], 1997) übernommen hat. Ein Schema ist eine Wissensstruktur, die zusätzlich zu einem einzelnen Faktum abrufbar ist. Wenn ich z. B. das Wort *Hochzeit* erfahre, kann ich sofort viele Dinge dazu aufrufen, die zu diesem Begriff gehören: Brautkleid, Brautstrauß, Hochzeitgesellschaft u. v. a. m. Eine derartige Wissensrepräsentation geht weit über eine einzige S-R-Verbindung hinaus und erfordert eine gedächtnismäßige Konstruktion und damit einen Informationsverarbeitungsprozess. Interessanterweise haben die US-Forscher Minsky [26], Rumelhart [39] und Schank [42] im Jahre 1975 das Schemakzept 40 Jahre später „wiederentdeckt“.

Als Fazit aus den vorangegangenen Darstellungen ist festzuhalten, dass das S-R-Paradigma des Behaviorismus, das wie kein anderes rigoros ein- und auch durchgeführt wurde, sich vollkommen überlebt hat. Der Grund für das Scheitern ist in der immer wichtiger werdenden Beachtung höherer geistiger Prozesse zu suchen, wie sie in behavioristisch geplanten Experimenten unerwartet zu Tage traten und neue Formen der Untersuchungen notwendig werden ließen. In diese Lücke trat die Psychologie der Informationsverarbeitung.

## 1.2 Wege zum Paradigma der Informationsverarbeitung

Die Entwicklung eines neuen Paradigmas der Informationsverarbeitung wurde nach Lachman, Lachman und Butterfield [19] 1979 durch die aufkommende Computerwissenschaft stark beeinflusst. Ein wichtiger Anstoß kam aus der Mathematik und geht auf den Göttinger Mathematiker David Hilbert zurück. Hilbert listete eine lange Reihe mathematischer Probleme auf, die ungelöst waren. Gäbe es eine allgemeine Beweistheorie für mathematische Probleme, würde diese das Lösen bisher ungelöster Probleme erleichtern. Tatsächlich fand Hilbert die gesuchte Beweistheorie nicht. Eine neue Einsicht in die Beweisbarkeit mathematischer Probleme publizierte Gödel [13] 1931. Er konnte zeigen, dass die Berechenbarkeit und damit Lösbarkeit eines mathematischen Problems nicht in demselben Axiomensystem bewiesen werden kann, in dem dieses Problem formuliert worden ist. Diese Einsicht stellte einen erneuten Rückschlag für das Auffinden einer allgemeinen Beweistheorie in der Mathematik dar.

### 1.3 Turing-Maschine

Einen ganz anderen Weg beschritt der damals noch junge Mathematiker Alan Mathinson Turing. Er publizierte im Jahre 1936 eine Arbeit [50], in der er das Problem der Vollständigkeit und der Berechenbarkeit mathematischer Probleme über sogenannte *effective procedures* zu lösen versuchte, indem er diese Vorgehensweisen analysierte. Derartige Prozeduren sind Algorithmen zur Lösung von Problemen vergleichbar. Solche Algorithmen bestehen aus einer nur kleinen Anzahl von Schritten, die bei lösbaren Problemen zum erwünschten Ergebnis führen. Wie können wir nun erfahren, so fragte sich Turing weiter, ob ein angewandter Lösungsalgorithmus für ein Problem vollständig ist? Wenn man die zur Lösung wenigen Operationen in eine Maschine implementiert, die diese Schritte ohne jede zusätzliche Intelligenz blind ausführt, so war Turings Argument, dann handelt es sich um eine vollständige Problemlösung. Die sogenannte Turing-Maschine war erfunden. Dabei ist zu beachten, dass diese Maschine als abstrakte Maschine von Turing gedacht wurde. Sie ist zum Zeitpunkt der Entstehung nicht mit Computern in Verbindung gebracht worden, die es damals noch gar nicht gab.

Weiterführend an Turings Ansatz war, dass seine universelle Maschine nicht nur mit Zahlen operieren konnte, sondern auch mit Symbolen. Hierzu bot sich die formale Logik an, weil man mit ihrem Formalismus logische Schlüsse in Symbolen abbilden konnte. Damit gelang es, logische Schlussfolgerungen durch Übersetzung in aussagenlogische Kalküle auf der Grundlage von Symbolmanipulationen formalisiert darzustellen. Als dann zum Beginn der 50er Jahre des letzten Jahrhunderts Computer einsetzbar wurden, konnte man diese logischen Operationen auch auf den Rechner übertragen, der die logischen Schritte durchführen konnte und zur Lösung gelangte. Das war eine Konkretisierung der Symbolmanipulation, so wie sie schon vor der Existenz von Computern vorgedacht worden war. Als Beispiel für eine solche Konkretisierung kann die Arbeit von Newell und Shaw [32] 1957 gelten. Mit ihrem Programm *Logic Theorist* rechneten sie auf diese Weise alle aussagenlogischen Beweise durch, die in dem Standardwerk [53] der Logik von Whitehead und Russell (1935) vorkamen. Das war nicht nur ausnahmslos möglich, sondern brachte auch Fehler ans Tageslicht, die den Buchautoren unterlaufen waren. Es sollte sich zeigen, dass dieser Ansatz fruchtbar weitergeführt werden konnte und zu einer der Grundlagen der Kognitionspsychologie wurde.

### 1.4 Informationstheorie – eine Sackgasse für die Kognitionspsychologie

Als weiteren wichtigen Anstoß zur Entstehung des Paradigmas der Informationsverarbeitung nennen Lachman, Lachman und Butterfield [19] 1979 die Informationstheorie. Für die Nachrichtentechnik hat Claude Elwood Shannon [45] 1948 die *Entropie  $H$*  zur Quantifizierung des mittleren Informationsgehalts der über einen Kanal übertragenen Nachrichten eingeführt. Sie wird in Bit gemessen, und berücksichtigt die Anzahl von Alternativen, die in einer Informationsmenge auftreten, und deren Auftretenswahrscheinlichkeiten. Genaueres findet sich im Beitrag



*Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information* von K. Schönhammer in diesem Band, ab S. 45). Auf der Grundlage dieser Größe  $H$  sowie der Redundanz, mit der ein „Informationsüberschuss“ beschrieben werden kann, hat Shannon eine Kommunikationstheorie entwickelt, die anfänglich auch in der Psychologie Interesse fand. Mit dieser Theorie hat G. A. Miller [23] 1951 die Leistung des Sprechens und Verstehens der englischen Sprache untersucht. Danach verwendet ein normaler Sprecher im Englischen 39 verschiedene Sprachlaute (Phoneme), und zwar in einer Durchschnittsgeschwindigkeit von 12,5 Phonemen pro Sekunde. Der Sprecher erzeugt damit  $H = 66$  bit pro Sekunde, die sein Zuhörer übernimmt. Diese Informationsmenge reduziert sich allerdings durch zusätzliche Einflussgrößen: Die ungleiche Auftretenshäufigkeit der Phoneme reduziert den  $H$ -Wert auf 60 bit. Außerdem vollzieht sich das geübte Sprechen nicht in Lauten, sondern in Silben, wovon es im Englischen ca. 4500 gibt und von denen man durchschnittlich fünf pro Sekunde ausspricht. Das senkt den  $H$ -Wert weiter auf 46 bit. Zusätzlich senken die von einem durchschnittlichen Sprecher verwendete Vokabelmenge von ca. 22 000 Wörtern und die Sprechgeschwindigkeit von drei Wörtern pro Sekunde den  $H$ -Wert auf elf bit. In geordneten und sinnvollen Sätzen schränken vorangehende Wörter die Alternativen für nachfolgende Wörter weiter ein, so dass sich schließlich in einem durchschnittlichen Sprechverkehr zwischen Menschen ein Wert von acht bit pro Sekunde zur Kennzeichnung der übertragenen Informationsmenge errechnet. Aus diesem Wert ergibt sich, dass sowohl ein Sprecher als auch ein Zuhörer durchschnittlich  $2^8 = 256$  theoretisch gleichwahrscheinliche Alternativen pro Sekunde zu berücksichtigen hat.

Derartigen Berechnungen haben Psychologen auch bei weiteren höheren geistigen Leistungen vorgenommen und z. B. durchgeleitete Informationsmengen, Kanalkapazität und hemmende Filter bestimmt, so wie man sie aus der Nachrichtentechnik kannte. Diesen Ansatz hat der englische Psychologe Broadbent verfolgt, der mit seinem Buch [8] aus dem Jahre 1958 mit dem Titel *Perception and Communication* (man beachte die Verwendung des Wortes *Communication*, das schon im Titel von Shannons Buch [46] vorkommt) als ein Begründer der Kognitionspsychologie gilt. Broadbent war es auch, der mit den Begriffen der Informationstheorie eine Filtertheorie der Aufmerksamkeit formulierte und so die wissenschaftliche Untersuchung nicht beobachtbarer Prozesse noch zur Hochzeit des Behaviorismus salonfähig machte.

Aber der Ansatz von Shannon erwies sich als Sackgasse für die Kognitionspsychologie. G. A. Miller (1956), der den Informationsfluss beim Sprechen mit Hilfe der Informationstheorie so präzise berechnen konnte [24], fand in seinen Experimenten zur Kapazität des menschlichen Gedächtnisses (eine typische Fragestellung nach der Informationstheorie) heraus, dass die Kapazitätsberechnung des Gedächtnisses nicht mit der Menge zu behaltener Zahlen, Buchstaben oder Wörtern übereinstimmte. Menschen fassen vielmehr entsprechendes Lernmaterial zu Chunks (Klumpen) zusammen, die sie anstelle von abzählbaren Einheiten im Gedächtnis speichern. Soll ich mir z. B. die Zahlenfolge 1771149219181941 merken, wird mir das nur schwer gelingen. Bilde ich daraus jedoch mehrere Chunks in der Form von Jahreszahlen, nämlich 1771 – 1492 – 1918 – 1941, werde ich die Zahl

wohl schneller im Gedächtnis behalten können. Dieser Effekt hatte sich, wie oben dargestellt, schon bei Millers Untersuchung des Informationsflusses beim Sprechen gezeigt: Die Anzahl zu berücksichtigender Alternativen beim Sprechen und auch beim Zuhören wurde durch einen Sinnzusammenhang mit vorangehenden Wörtern deutlich gesenkt. Mit der Chunk-Bildung hatte Miller eine kognitive Leistung entdeckt, die für die Untersuchung höherer geistiger Leistungen beim Menschen eine sehr große Rolle zu spielen begann. In Extremfällen wie z. B. bei Rechenkünstlern (Bredenkamp, Klein, von Hayn & Vaterrodt [7] 1988) erklärt die Fähigkeit, Chunks zu bilden, geradezu die außergewöhnlichen Leistungen, zu denen solche Menschen in der Lage sind. Genau aber die Entstehung von Chunks kann mit dem Bit-Maß nicht erfasst werden. Unter anderem aus diesem Grund wurde die Informationstheorie von Shannon nach einer ca. zehnjährigen Zeit der Anwendung als Forschungsmethode in der Kognitionspsychologie nicht weiter verfolgt.

## 1.5 Fazit

Der Weg zu einer Psychologie der Informationsverarbeitung kann aus verschiedenen wissenschaftlichen Einzelentwicklungen heraus erklärt werden. Auf einer sehr generellen wissenschaftshistorischen Ebene muss dieser Weg als ein Paradigmenwechsel verstanden werden. Die geradezu planmäßige Vernachlässigung des Blickes auf Prozesse, die zwischen dem Reiz und Reaktion liegen wie die höheren kognitiven Prozesse, war in der vom Behaviorismus geforderten Strenge nicht durchzuhalten. Vielmehr zeigte sich auch in behavioristisch angelegten Experimenten, dass zwischen einem gesetzten Reiz und der beobachteten Reaktion Prozesse abliefen, die der psychologischen Untersuchung und der Aufklärung bedurften. Diese Einsicht forderte immer stärkere Beachtung und durchlöcherte damit den theoretischen Ansatz des Behaviorismus. Mit der Überzeugung, dass gerade die vom Behaviorismus negierten bzw. vernachlässigten kognitiven Prozesse von der Psychologie untersucht und in ihrer Gesetzmäßigkeit aufgeklärt werden müssten, vollzog sich ein radikaler Paradigmenwechsel zu einer Psychologie der Informationsverarbeitung.

Hilfreiche Anregungen zu einem neuen Weg in eine Informationsverarbeitungspsychologie konnten vor allem aus den Computerwissenschaften und der Nachrichtentechnik bezogen werden. Besondere Bedeutung kam dabei Turings [50] Entdeckung aus dem Jahre 1936 zu, Lösungen von Problemen durch eine Aufeinanderfolge von nur wenigen grundlegenden Operationen zu finden. Mit dieser noch vor der Zeit von Computern entworfenen Turing-Maschine war die Einsicht in eine schrittweise zu vollziehende Informationsverarbeitung verbunden, von der schon Turing annahm, dass sie der Informationsverarbeitung im menschlichen Gehirn sehr ähnlich sei. Damit war ein erster Schritt getan, dass artifizielle Systeme intelligente Leistungen erbringen können.

Einen anderen Weg, Informationsverarbeitung zu untersuchen, bot die Nachrichtentechnik. Mit Shannons Formalisierung der Informationsmenge in einem quantitativen Maß als Herzstück seiner Kommunikationstheorie ergaben sich neue Möglichkeiten, Kapazitätsgrößen bei der Durchleitung und Speicherung von Infor-

mationen beim Menschen zu bestimmen. Bei entsprechenden Experimenten stellte sich jedoch heraus, dass die Informationsmaße aus der Nachrichtentechnik dazu wenig geeignet erschienen. Damit verblieb der von Turing gewiesene Weg zur Konzeptualisierung der menschlichen Informationsverarbeitung die erfolgversprechendere Möglichkeit.

## 2 Psychologie der Informationsverarbeitung

In diesem Abschnitt werden zwei sehr unterschiedliche Ansätze der Informationsverarbeitung beschrieben, derer sich die kognitive Psychologie bedient hat. Die symbolische Informationsverarbeitung hat sich direkt aus den Computerwissenschaften heraus entwickelt und profitierte insbesondere von der Informatik. Die subsymbolische Informationsverarbeitung berief sich auf Vorstellungen der theoretischen Neurologie und bezeichnete ihre wissenschaftliche Vorgehensweise mit dem Begriff *brain-style modeling*. Diese Bezeichnung ist allerdings nicht identisch und auch nicht verwandt mit der heutigen Hirnforschung in den Neurowissenschaften, auf die später eingegangen wird.

### 2.1 Symbolische Informationsverarbeitung

Mit dem Verlassen des S-R-Paradigmas zugunsten einer Informationsverarbeitungspsychologie wurde auch die Vorstellung einer Black Box aufgegeben, die der Behaviorismus als Zwischenstück zwischen dem auf das Individuum einwirkenden Reiz und seiner Reaktion als nicht weiter aufklärungsbedürftig deklariert hatte. Das Interesse der Forschung konzentrierte sich jetzt explizit darauf, den Prozess zu beschreiben, wie ein dargebotener Reiz vom Individuum aufgenommen und verändert wird und am Ende dieses Weges zur Grundlage einer Reaktion wird. Damit wurde aus der Black Box eine White Box.

#### 2.1.1 Beispiel eines Informationsverarbeitungsprozesses

Ein frühzeitiges Beispiel hat Sternberg 1966 [47] publiziert, das unter der Bezeichnung Sternberg-Paradigma bis heute in der kognitionspsychologischen Forschung angewendet wird. Im Sternberg-Paradigma wird einem Probanden kurzzeitig eine kleine Anzahl von einstelligen Zahlen – in der Regel nicht mehr als sechs Ziffern – gezeigt, die er sich merken soll. Danach erscheint eine einzige Zahl in Verbindung mit der Frage, ob sie in der vorher gezeigten Zahlensequenz vorkam. Der Proband wird aufgefordert, seine Antwort so schnell wie möglich durch das Drücken einer Ja- oder einer Nein-Taste zu geben. Dabei wird die Zeit gemessen, die zwischen dem Erscheinen der Testziffer und der Antwortabgabe verstrichen ist. Diese Prozedur wird mit unterschiedlichen Längen der zu merkenden Ziffernfolgen vielfach wiederholt. Getrennt ausgewertet werden dann die benötigten Antwortzeiten für „Treffer“ und für „Nieten“. Um Nieten handelt es sich, wenn die Testziffer nicht in der ursprünglichen Sequenz auftrat. Sternberg fand heraus, dass die Antworten von den Probanden ziemlich schnell abgegeben werden konnten. Die Zeiten

varyierten, je nach Länge der zu merkenden Zahlensequenz, zwischen 380 und 600 Millisekunden. Weiterhin ergab sich, dass zwischen der zu behaltenden Zahlenmenge und der Beurteilungszeit eine annähernd lineare Beziehung besteht. Für jede weitere Ziffer der dargebotenen Zahlensequenz wurde die Reaktionszeit um 38 Millisekunden verlängert. Aus diesem Befund entwickelte Sternberg eine Abfolge der zugrunde liegenden Informationsverarbeitungsschritte bei der Bewältigung der gestellten Aufgabe. Sie visualisierte er in einem Flussdiagramm, das die einzelnen Informationsverarbeitungsschritte von der Darbietung der Testziffer bis zur gegebenen Antwort beschreibt, in Abbildung 1 dargestellt am Beispiel eines Probanden, der die Ziffernfolge 2 – 4 – 7 gesehen hatte und nun als Testziffer die Zahl 4 bekommt.

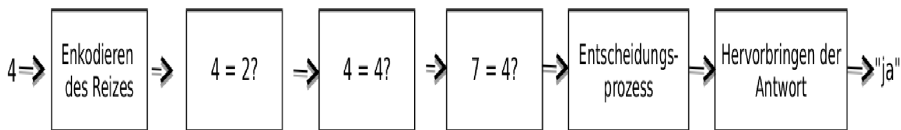


Abbildung 1. Sternberg-Vergleichskette

Nach der Wahrnehmung der Testziffer 4 findet ein dreifacher Vergleich mit den im Kurzzeitgedächtnis gespeicherten Ziffern der Sequenz statt. Danach erst folgt die Entscheidung „ja“ oder „nein“. Im letzten Schritt wird die Antwort generiert und ausgeführt. Dieses vollständig sequentielle Vorgehen erscheint auf den ersten Blick unplausibel, da eigentlich schon nach Prüfung der zweiten Ziffer aus der Sequenz, der 4, die Antwort „ja“ gegeben werden könnte. Sternberg erklärte jedoch die vollständige Prüfung damit, dass bei diesem Vorgehen nur ein einziger Entscheidungsschritt benötigt wird, um festzustellen, ob an irgendeiner Stelle der Zahlensequenz eine Ja-Markierung angebracht wurde oder nicht. Entscheidungen erfordern einen vergleichsweise hohen kognitiven Aufwand und benötigen dafür viel Zeit. Deshalb ist es ökonomischer, nur einen einzigen Entscheidungsschritt auszuführen als – wie im Beispielfall – drei davon. Es lässt sich unschwer vorstellen, dass die hier im Experiment geforderte Entscheidungsaufgabe sich auf der Grundlage von Symbolen so abbilden lässt, dass eine Turing-Maschine das Experiment erfolgreich nachvollziehen kann.

### 2.1.2 Simulation des menschlichen Denkens

Kehren wir nun zurück zu den Wurzeln der Informationsverarbeitung in der kognitiven Psychologie. Nach erfolgreichen Versuchen, aussagenlogische Beweise in Symbolen abzubilden und sie in einem Beweisprozess durch schrittweise Veränderungen zu einem logischen Schluss bringen zu können, so wie das mit dem *Logic Theorist* gelungen war, stellten sich Psychologen die Frage, wie denn Menschen solche aussagenlogischen Aufgaben lösen. Dazu gaben sie Probanden nach einer

entsprechenden Einführung derartige Beweisaufgaben vor und baten sie, alles, was ihnen auf dem Wege der Lösung durch den Kopf ging, zu berichten. Alle Äußerungen wurden aufgenommen, und daraus wurde ein Protokoll angefertigt. Man nennt dieses Verfahren „Methode des lauten Denkens“. Ein solches empirisches Vorgehen führten Newell, Shaw und Simon [33] 1957 in einer sehr frühzeitigen Phase aus (s. a. Lüer [21] 1973). Mit der Protokollauswertung von den Probandenäußerungen fanden sie heraus, dass Menschen beim Lösen derartiger Probleme sogenannte Heuristiken anwandten, mit denen, anders als in strengen formalen mathematische Beweisführungen, eher im Ungefähren operiert werden konnte, womit sich die Problemlöser die gestellten Aufgaben vereinfachten. Eine solche Heuristik ist die *Means-Ends-Analysis*, mit der Problemlöser durch den Vergleich des Zielzustands des gestellten Problems mit dem inzwischen erreichten Zustand, der aber noch nicht die Lösung ist, feststellen, ob sie sich der Lösung schon genähert haben oder nicht. Diese Einsicht, dass Probanden Heuristiken anwenden, führte dazu, ein Computersystem wie den *Logic Theorist* zusätzlich mit Heuristiken auszustatten. Damit entstand das System *General Problem Solver (GPS)*, das in einer Publikation der oben genannten Autoren beschrieben wird als „a system that simulates human thought“. Schon im folgenden Jahr 1958 kam es zu dem berühmt gewordenen RAND-Symposium in Santa Monica unter der Leitung von Newell und Simon. Hier wurden erste Ergebnisse mit dem noch nicht fertigen GPS-System vorgestellt und diskutiert. Der Gedanke, nun ein wirkungsvolles Instrument zur Simulation menschlichen Denkens und Problemlösens in der Hand zu haben, wurde ausführlich diskutiert und als Forschungsprogramm aus der Taufe gehoben. Höhepunkte dieser Entwicklung einer kognitiven Psychologie auf der Grundlage der Informationsverarbeitung waren das Erscheinen von grundlegenden Werken über die Simulation kognitiver Prozesse beim Menschen. Dazu gehören die Standardwerke *Human Problem Solving* [34] von Newell und Simon (1972) und *Information und Verhalten* [17] von Klix (1971).

Die schon ursprünglich von Turing geäußerte Idee einer Analogie zwischen der Arbeitsweise des Computers und der kognitiven Informationsverarbeitung beim Menschen hatte durch die Beobachtung menschlichen Problemlösens eine wesentliche Weiterentwicklung erfahren. Nach dieser Analogie nehmen Computer symbolischen Input auf, formen ihn zu neuen Ausdrücken um, treffen auf dieser Grundlage Entscheidungen, speichern Symbole ganz oder teilweise und geben wieder Symbole als Output aus. Und im Sinne dieser Analogie sind das genau jene Prozesse, für die sich kognitive Psychologen interessieren, nämlich wie Menschen Informationen aufnehmen, sich merken oder auch vergessen, wie sie Entscheidungen treffen, wie sie ihr Wissen verändern und wie sie dieses Wissen in Verhalten übersetzen. Demnach kann ein Computerprogramm als eine Art Theorie über die Arbeitsweise des zu simulierenden Systems und über die relevanten Variablen, die die Leistung beeinflussen, aufgefasst werden. Ein solches Simulationsprogramm erklärt kognitive Leistungen, indem es über diskrete Zeitintervalle hinweg die Veränderungen im kognitiven System beschreibt. Das Programm charakterisiert jeden neuen Zustand des kognitiven Prozesses als eine Funktion des unmittelbar vorangegangenen Zustandes.

Die diesem Verständnis von intelligenter kognitiver Informationsverarbeitung zugrundeliegende Hypothese wurde von Newell und Simon (1976) als *Physical Symbol System Hypothesis* eingeführt:

*There is no 'intelligence principle', just as there is no 'vital principle' that conveys by its very nature the essence of life. But the lack of a simple deus ex machina does not imply that there are no structural requirements for intelligence. One such requirement is the ability to store and manipulate symbols ... A physical symbol system consists of a set of entities, called symbols, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure) ... A physical symbol system is a machine that produces through time a series of evolving symbol structures. Such a system exists in a world of objects wider than just these symbolic expressions themselves ... We now can state a general scientific hypothesis – a law of qualitative structure for symbol systems: The Physical Symbol System Hypothesis. A physical symbol system has the necessary and sufficient means for intelligent action.*

(Zitiert nach Strube [48], 1990, S. 132)

Computerprogramme zur Simulation kognitiver Prozesse haben einen neuen und sehr fruchtbaren Weg der Theorienbildung eröffnet. Mit ihrer Hilfe können beispielsweise Lücken in unserem Verständnis von kognitiven Phänomenen aufgefunden werden. Denn das Programm muss vollständig sein, sonst würde sein Ablauf unterbrochen werden und stoppen. Werden Lücken erkannt, kann gezielt nach Hypothesen gesucht werden, sie zu schließen. Schließlich kann der Output mit tatsächlichem Verhalten verglichen werden und dann dazu benutzt werden, die dem Programm zugrunde liegenden theoretischen Annahmen zu modifizieren. In diesem Sinne wird jeder Programmdurchlauf zu einem Experiment. Computersimulationsprogramme können somit als eine Art der Operationalisierung des O zwischen dem S und dem R verstanden werden.

### **2.1.3 Kognitive Architekturen – eine Weiterentwicklung kognitiver Simulationsprogramme**

Das oben beschriebene Flussdiagramm zur Darstellung der Informationsverarbeitungsschritte bei der Bewältigung der Sternberg-Aufgabe (s. Abb. 1) beschreibt inhaltlich nur einen winzig kleinen Ausschnitt kognitiver Leistungen des Entscheidens, wenn auch in präziser und experimentell jederzeit nachprüfbarer Art und Weise. Einen auch nur einigermaßen repräsentativen Überblick über intelligente Leistungen des Menschen kann es jedoch nicht liefern. Um diesem in der Kognitionswissenschaft immer wieder beklagten Mangel (Neisser [28, 29, 30] 1967, 1976 und 1982 sowie Newell [31] 1973) abzuhelfen, haben einige Psychologen sogenannte *kognitive Architekturen* entwickelt. Dabei handelt es sich um lauffähige Computerprogramme, die sehr viele und auch vielfältige Leistungen erbringen können, die menschlichen kognitiven Leistungen sehr ähnlich sind oder sogar gleichen. Kognitive Architekturen verfügen über Wissensbasen vom eigenen System

sowie über Repräsentationen von der Umwelt. So können sie z. B. solches Wissen mit Zielen vergleichen, die sie in einer zu erfüllenden Aufgabe erreichen möchten. Dabei werden diese Programme so eingestellt, dass sie menschlichen Vorgehensweisen möglichst ähnlich sind. Das betrifft u. a. sowohl die Geschwindigkeit, mit der Probleme gelöst werden können, als auch Kapazitätsbegrenzungen, wie sie im Gedächtnis vorkommen. In kognitiven Architekturen werden z. B. Verstehensprozesse so nachgebildet, dass sie bei der Aufnahme von Informationen aus der Umwelt dafür ein Verständnis entwickeln können, um das nun Verstandene im weiteren Verarbeitungsprozess wie z. B. beim Problemlösen erfolgreich einsetzen zu können.

Mit der Programmierung einer kognitiven Architektur wird also – wie oben bereits bei den einfachen Simulationsprogrammen beschrieben – eine Theorie über diejenigen kognitiven Prozesse erstellt, die mit dem System simuliert werden sollen. Da in solchen Architekturen sehr viele kognitive Leistungen abgebildet werden können, deren Zusammenwirken auch bei sehr komplexen geistigen Anforderungen verfolgt, überprüft und immer wieder nachjustiert werden kann, ist es möglich, auf diesem Wege eine umfassende Theorie über kognitive Prozesse zu realisieren. Dabei hat es sich als besonders wirkungsvoll erwiesen, die Konsistenz und Widerspruchsfreiheit der Theorie zu überprüfen, die die Voraussetzung für die Lauffähigkeit des Programms darstellen. Kognitive Architekturen als laufende Programme werden häufig in sogenannten tutoriellen Systemen verwendet. Das sind Lernprogramme, die in einem interaktiven Prozess mit Lernenden den Stand des Wissens diagnostizieren, gezielt Trainings ausführen, um Wissenslücken zu schließen, und so zu einem gefestigten Lernfortschritt beim Probanden führen. Besondere Einsatzfelder sind das Erlernen von Programmiersprachen und von mathematischen Kenntnissen.

Sehr bekannt geworden ist die von John R. Anderson [1] 1983 programmierte kognitive Architektur ACT-R. Neben der symbolischen Repräsentation von Wissen und auch Zugriffsmöglichkeiten auf subsymbolische Wissens Ebenen verfügt diese Architektur über sogenannte Produktionssysteme. Sie umfassen Regeln, die zur Bewältigung von Anforderungen geeignet sind. Es handelt sich um formale Darstellungen von Operationen, die auf den Wissensbasen operieren können. Produktionen bestehen aus einem Wenn-Teil und aus einem Dann-Teil.

*Wenn* ich einen Spaziergang machen möchte  
und wenn es draußen regnet  
und wenn ich nicht nass werden will  
*Dann* nehme ich einen Regenschirm mit.

Im Wenn-Teil ist eine Bedingung enthalten, der Dann-Teil enthält eine Handlung oder auch kognitive Aktion. In der Bedingung ist ebenfalls ein Ziel angegeben verbunden mit der Prüfmöglichkeit für die Anwendbarkeit einer Regel. Fällt die Prüfung positiv aus, wird die Regel angewendet, d. h. eine Handlung oder kognitive Aktion wird ausgeführt.



Für die Bewältigung von Anforderungen hat Anderson (2001,[2] S. 252f.) folgende Merkmale der Produktionen angegeben:

1. *Bedingtheit:*  
Jede Produktionsregel besteht aus einer Bedingung, die beschreibt, wann sie angewendet werden soll, und einer Aktion, die beschreibt, was zu tun ist.
2. *Modularität:*  
Die Gesamtfähigkeit des Problemlösens wird in viele Produktionen zergliedert, jeweils eine für jeden Operator (Handlungsmöglichkeit).
3. *Zielzerlegung:*  
Jede Produktion ist einem bestimmten Ziel zugeordnet ...
4. *Abstraktheit:*  
Jede Regel wird auf eine Klasse von Situationen angewendet ...

Mit dem skizzierten Verfahren konnte Anderson z. B. die Sprachentwicklung seines Kindes in den ersten Lebensjahren erfolgreich simulieren und damit neue entwicklungspsychologische Erkenntnisse generieren.

Noch umfassendere kognitive Architekturen hat Dörner [9, 10] erarbeitet. Die Buchtitel *Bauplan für eine Seele* (1999) und *Die Mechanik des Seelenwagens* (2002) benennen gleichzeitig den Anspruch dieser Systeme. Prozesse für die Wahrnehmung, für die Vorstellung, für Emotionen, für die Selbstreflexion, für soziale Bedürfnisse und für die Sprache werden detailliert als Informationsverarbeitungsprozesse beschrieben und zu einer Architektur zusammengefügt, die als Programm zum Laufen gebracht werden kann und damit als überaus komplexes System in seiner Arbeitsweise zu beobachten ist. Vergleiche mit Daten von Probanden werden ebenfalls möglich.

Kognitiven Architekturen wird der Vorteil zugeordnet, der Beliebigkeit im grundlegenden theoretischen Konzept leichter zu entgegen. Wenn, wie in den Anfängen der Simulationsforschung geschehen, für jede Denksportaufgabe oder für jedes Brettspiel ein eigenes Programm entworfen wird, tragen diese Realisierungen sehr wenig an Allgemeinverbindlichkeit in sich. Darüber hinaus konnte gezeigt werden, dass für die Simulation einer jeden speziellen kognitiven Leistung in der Regel mehrere verschiedene Programmierlösungen gefunden werden können. Deshalb kennzeichnete man diese Forschungsrichtung ironischerweise auch als „Modellschreinerei“. Das ist bei kognitiven Architekturen anders. Sie sind in der Lage, sehr viele – sogar möglichst alle – kognitiven Leistungen mit ein und demselben Programm erfüllen zu können. Dadurch, so wird angenommen, wird der mit solchen Systemen zu erzielende Erkenntnisgewinn über die Funktionsweise von Informationsverarbeitungsprozessen deutlich in seiner Fruchtbarkeit erhöht.



### 2.1.4 Kritik an Simulationsprogrammen für kognitive Prozesse

Die Versuche, das Problemlösen von Menschen in Computern zu simulieren, führte auch zu sehr kritischen Überlegungen. Zuerst führte Turing [51] 1950 einen Versuch durch, bei dem eine Versuchsperson über eine Tastatur bei der Bewältigung einer Aufgabe einmal mit einem Menschen als Versuchsleiter und zusätzlich auch mit einer Turing-Maschine kommunizierte. Konnte die Versuchsperson nicht unterscheiden, welches der Versuchsleiter und welches der Computer war, der Antworten gab, hatte das künstliche System den sogenannten Turing-Test bestanden. Das System, so Turings Anspruch, hatte das Vorgehen des Versuchsleiters simuliert. An dieser Stelle sei noch angemerkt, dass Turings Idee von der Arbeitsweise des menschlichen Gehirns in Analogie zur Informationsverarbeitung beim Computer nicht auf der Ebene von Neuronen und ihren Vernetzungen angesiedelt ist, wie es in den heutigen Neurowissenschaften üblich ist.

Gegen Turings Interpretation argumentierte Searle ([43], 1981). Sein Gedankenexperiment ist als *Chinesisches Zimmer* bekannt geworden. In einem Zimmer befindet sich eine Person, die die chinesische Sprache nicht beherrscht. Von außen werden der Person Zettel mit chinesischen Zeichen zugesteckt sowie eine Geschichte in chinesischer Schrift, zu der die einzelnen Schriftzeichen passen. Die Person im Zimmer hat ein rein syntaktisches Regelwerk zur Verfügung, mit der sie die Zeichen kombinieren kann, ohne den Sinn zu verstehen. Auf diese Weise kann die Person im Zimmer in chinesischer Sprache formulierte Fragen beantworten, indem sie Zettel mit passenden Zeichenkombinationen wieder aus ihrem Zimmer hinausreicht, ohne dass sie die Bedeutung der Geschichte und der Zeichen sinngemäß nachvollziehen kann. Ein Chinese kann jedoch die gegebenen Antworten von der Person im Zimmer verstehen. An diesem Beispiel zeigt Searle, dass die Simulation durch eine Turing-Maschine sich grundlegend von menschlichem Denken unterscheidet. Der Turing-Maschine fehlt das semantische Verständnis für das, was sie tut. Dies gilt aber als ein Kennzeichen menschlichen Denkens und Handelns. Eine vollständige Äquivalenz zwischen menschlichem Denken und einer Simulation durch eine Turing-Maschine oder einen Computer besteht nicht, deshalb ist der Turing-Test auch keine adäquate Prüfung für eine Äquivalenzentscheidung zwischen Simulation und menschlichem Denken.

Als abgeschlossen ist diese Diskussion wohl immer noch nicht zu sehen, wie sich im Wettkampf zwischen einem Schachweltmeister und einem Schachcomputer gezeigt hat. Der IBM-Computer Deep Blue konnte den damals amtierenden Schachweltmeister Garry Kasparow im Jahre 1997 mit 3,5 zu 2,5 besiegen. Da der Besiegte die Vermutung geäußert haben soll, dass auf Seiten des Computers zusätzlich ein Mensch eingegriffen haben müsste, was nicht der Fall gewesen ist, hat dieses System wohl auch den Turing-Test bestanden.

## 2.2 Subsymbolische Informationsverarbeitung

Im vorangegangenen Abschnitt wurde die symbolische Informationsverarbeitung dargestellt, die der Computermetapher folgt und jede Verbindung zu neurophy-

siologischen Grundlagen kognitiver Prozesse auslöst. Im nun folgenden Abschnitt wird es primär um den Konnektionismus gehen, einem Forschungsansatz, der sich sehr kritisch mit der Computermetapher der symbolischen Informationsverarbeitung auseinandersetzt und der durch Anleihen aus der Neurophysiologie zu geradezu konträren Vorstellungen über die Verarbeitung von Informationen im kognitiven Geschehen gelangt. Hauptauslöser der Kritik an der Computermetapher war die Annahme einer seriellen Informationsverarbeitung, wie sie von den Vertretern dieser Forschungsrichtung propagiert wird. Da die als Simulationswerkzeug verwendeten Computer in der von John von Neumann erdachten Architektur ausschließlich eine serielle, d. h. schrittweise Verarbeitung zuließen, versuchten die Anhänger des Konnektionismus den Nachweis zu führen, dass diese Vorstellung mit neurophysiologischen Grundlagen der Arbeitsweise des Gehirns nicht in Einklang zu bringen ist. Durch ihre Orientierung an bekannten Fakten der Neurologie und der Hirnforschung wiesen sie Merkmale wie seriell und symbolisch als generelle Charakteristika menschlicher Informationsverarbeitung zurück.

### 2.2.1 Marksteine auf dem Weg zum Konnektionismus

Der spanische Neuroanatom Ramón y Cajal ([37], 1911) erforschte durch Anwendung spezieller histologischer Einfärbungsmethoden die Nervenbahnen der grauen Substanz im Gehirn. Aufgrund seiner Ergebnisse entwickelte er eine Theorie, nach der das Nervensystem aus ein Vielzahl von Neuronen und ihren Fortsätzen besteht, die zusammen ein eng verschaltetes Netzwerk bilden. Eine bahnbrechende Weiterentwicklung dieser anatomisch beschriebenen Netzwerkverschaltungen gelang in den 40er Jahren des 20. Jahrhunderts dem amerikanischen Psychiater McCulloch (1943 [22]; einem Schüler des Behavioristen C. L. Hull) mit seinem Studenten Pitts. Sie verfolgten die Hypothese, dass geistige Leistungen als Gehirntätigkeiten verstanden werden können. Dabei konzentrierten sie sich auf die Neuronen und interpretierten sie als *logische Geräte*, die auf der Grundlage ihrer Verschaltungen miteinander kommunizieren und deren Austausch mit der Booleschen Logik präzise erfasst und beschrieben werden können. Wichtig dabei ist die Erkenntnis, dass die Kommunikation der Neurone untereinander auf elektrochemischem Wege erfolgt. Eine Nervenzelle feuert nur – gibt also nur Informationen weiter – wenn ein Schwellenwert erreicht wird, über den jede einzelne Zelle verfügt. Von großer Bedeutung ist es, dass eine Nervenzelle sowohl aktivierende als auch hemmende Signale von allen Neuronen aufnimmt und aufsummiert, mit denen sie verbunden ist. Es ist dann aber vom ihr eigenen Schwellenwert abhängig, ob sie feuert oder nicht. Durch die Zusammenschaltung von Tausenden von Neuronen in einem Netzwerk ergibt sich eine ungeheuer große Verarbeitungskapazität des gesamten Zellverbandes sowie auch eine Verarbeitungsgeschwindigkeit von imposantem Ausmaß. Die Konzeption und Funktion von einem neuronalen Netz mit der Eigenschaft der parallelen Verarbeitung war damit als Alternative zur seriellen Informationsverarbeitung in den Blickpunkt geraten. Noch John von Neumann erkannte zu Lebzeiten, dass man einen Computer bauen könne, der die beschriebenen Prinzipien der parallelen neuronalen Verschaltungen umsetzt. Zu jener Zeit gab es allerdings noch

nicht einmal Transistoren, wodurch eine technische Realisierung allein mit Röhren und Relais sehr schwierig geworden wäre.

Die bisher beschriebenen Fakten haben sich noch alle aus Forschungen außerhalb der Psychologie ergeben. Zu den frühen für den Konnektionismus relevanten Arbeiten ist die Dissertation [25] von Minsky aus dem Jahre 1951 (publiziert 1954) zu zählen, der auf der Grundlage eines neuronalen Netzes eine lernende Maschine aufbaute. Sie bestand aus 400 Röhren, die untereinander verbunden werden mussten. Allein schon dieser technische Aufwand, der nur durch fleißiges Löten erbracht werden konnte, bremste die Konstruktion solcher Netzwerke aus. Später führten Minsky und Papert ([27], 1969) dann aus, dass sich vernetzten Neuronenverbänden, wollte man sie maschinell herstellen und theoretisch beschreiben, auch prinzipiell unüberwindliche mathematische Schwierigkeiten entgegenstellen, so dass sich ein Weiterverfolgen dieses Weges nicht lohnen würde. Minsky ging deshalb der Idee artifizierlicher neuronaler Netzwerke nicht weiter nach und wandte sich dem Ansatz der symbolischen Informationsverarbeitung wieder zu.

Aber es gab dennoch einen konsequenten Versuch, die von John von Neumann explizierte Idee eines parallel verarbeitenden Computers zu realisieren. Dieses unternahm Rosenblatt ([38], 1958), der einen Computer konstruierte, der Formen erkennen und klassifizieren konnte, also Wahrnehmungsleistungen von unterschiedlichem Niveau zeigte. Rosenblatts Maschine, die er *Perceptron* nannte, bestand aus Input-Units für sensorischen Input und Output-Units für den motorischen Output. Input- und Output-Units waren assoziativ verbunden. Damit sollte in einer ersten Annäherung den Gegebenheiten des menschlichen Gehirns Rechnung getragen werden. Der Erfolg dieses Versuches fiel eher bescheiden aus: Die Verarbeitung der Reize gelang nur in einer geringen Streubreite und die Reaktionen fielen sehr simpel aus. Zudem stellte sich die Lernfähigkeit des *Perceptron* als äußerst begrenzt heraus.

Um den bisher wenig überzeugenden Erfolgen bei der maschinellen Realisierung von künstlichen neuronalen Netzwerken zu einem wirklichen Durchbruch zu verhelfen, bedurfte es weiterer zusätzlicher Ideen, die man wiederum aus den Neurowissenschaften bezog. Der kanadische Neuropsychologe Donald O. Hebb hat in einer bahnbrechenden Arbeit [14] schon im Jahre 1949 zeigen können, dass gespeichertes Wissen nicht in einzelnen Nervenzellen sondern in Zellverbänden abgelegt wird. Als wichtiges Faktum stellte sich dabei heraus, dass Verbindungen zwischen Neuronen, die gleichzeitig aktiv sind, sich mit jeder neuen Inanspruchnahme verstärken und umgekehrt, bei Nichtverwendung, sich abschwächen. Damit war ein wichtiges Ergebnis gefunden worden, das mit Lernvorgängen unmittelbar in Beziehung gesetzt werden konnte. Dieses Resultat zeigte darüber hinaus, dass Lernen offensichtlich – verteilt – in neuronalen Netzen stattfindet und dass dieser Vorgang mit einer Parallelverarbeitung einhergeht. Damit stellten sich die Verbindungen zwischen Neuronen als durch Lernvorgänge veränderliche Strukturen heraus. Wir werden später auf diesen von Hebb nachgewiesenen Befund bei der Beschreibung künstlicher neuronaler Netzwerke zurückkommen. Erwähnenswert in diesem Zusammenhang ist, dass sich die Idee der verteilten Repräsentation von Wissen schon bei dem Tierpsychologen Karl Lashley ([20], 1929) findet. Die Rat-

ten in seinen Experimenten lernten, sich in einem Labyrinth zurechtzufinden. Nach einer anschließenden operativen Entfernung von Gewebeteilen an sehr unterschiedlichen Stellen des Gehirns waren die Ratten immer noch in der Lage, ihr zuvor gelerntes Ortswissen wenigstens teilweise abzurufen. Übrigens war Donald Hebb ein Doktorand von Karl Lashley.

Die beschriebenen Bestrebungen, zusätzlich zu der Computermetapher eine alternative Sichtweise für die Erforschung der Informationsverarbeitung beim Menschen zu entwickeln, kamen in den Jahren von ca. 1950 bis 1980 nur in kleinen Schritten voran und waren auch von Rückschlägen nicht verschont worden. Während die Entwicklung von Computersimulationen nach dem Prinzip der symbolischen Informationsverarbeitung große Fortschritte in dieser Zeit vermelden konnte, verlief die wissenschaftliche Entwicklung in der subsymbolischen Forschung eher schleppend. Dabei spielte das Konkurrenzverhalten der beiden Entwicklungslinien sicher auch eine Rolle. Da die auf die Computerwissenschaften zurückgreifenden Forscher nicht nur sehr erfolgreich waren und zudem mit geringeren finanziellen Mitteln arbeiten konnten als die Neurowissenschaftler, kam es zunächst zu einer deutlich erkennbaren Asymmetrie in der Durchsetzungsfähigkeit der beiden konkurrierenden Forschungsansätze.

Der Durchbruch der Netzwerkmodelle begann mit den 80er Jahren des letzten Jahrhunderts. Der theoretische Physiker John Hopfield veröffentlichte im Jahre 1982 einen Artikel [15], in dem die Eigenschaften neuronaler Netzwerke soweit vereinfacht werden konnten, dass sie exakten mathematischen Beschreibungen zugänglich wurden. Außerdem war seine Argumentationslinie eng an neurowissenschaftliche Sichtweisen angelehnt. Zusätzlich wies er nach, dass in einem Netzwerk, das aus einfachen homogenen Elementen besteht, relativ stabile und dauerhafte Aktivitätsmuster als Zielzustand entstehen können, die als physikalisches Substrat für die Informationsspeicherung verstanden werden können. Wesentliches Merkmal von Hopfield-Netzen ist, dass sie inhaltsadressierbar arbeiten: Das Angebot eines Reizmusters aktiviert ein entsprechend gespeichertes Neuronenmuster ohne Zwischenschaltung eines Indexes und ohne das Durchforsten einer Adressenliste. Das führt zu einer Verarbeitung, wie sie auch für menschliche Wahrnehmungs- und Lernleistungen charakteristisch ist:

- Das Angebot eines Teilmusters aktiviert dennoch das gesamte Muster.
- Für mehrdeutige Muster wird der statistisch wahrscheinlichste Anschluss gesucht.
- Muster, die große Übereinstimmungen zeigen, verschmelzen zu einem gemeinsamen Muster. Damit kann das System generalisieren und kategorisieren.
- Der Ausfall einzelner Elemente hat lediglich eine teilweise abträgliche Wirkung (*fail soft*) auf das Ergebnis.

Durch die Einbeziehung von Rückkoppelungsprozeduren und der damit einhergehenden Veränderung der Gewichte von Verbindungen zwischen den Netzelemen-

ten kam Hopfield auch der von Hebb 1949 aufgestellten Regel nahe, nach der die gemeinsame Aktivität zweier Neurone zur Verstärkung der Verbindung zwischen diesen Neuronen führt.

### 2.2.2 Konnektionismus

Mit den beschriebenen wissenschaftlichen Entwicklungen konnte zu Beginn der 80er Jahre des letzten Jahrhunderts der Konnektionismus als subsymbolische, parallele Informationsverarbeitung und den Neurowissenschaften verbundene Alternative in Konkurrenz treten zur symbolischen, seriellen Informationsverarbeitung, die aus den Computerwissenschaften heraus sich entwickelt hatte. Man gab dieser neuen Richtung auch den Namen *brain-style modeling*, was zu mindestens für den Zeitraum der 80er Jahre des 20. Jahrhunderts als euphemistisch bezeichnet werden muss.

Von sehr großer Bedeutung für die Psychologie der subsymbolischen Informationsverarbeitung war das 1986 von Rumelhart und McClelland herausgegebene zweibändige Werk *Parallel distributed processing* [40, 41]. Sie konnten die Bewährung von Netzwerkmodellen für die Erforschung der Mustererkennung, der Wahrnehmung, des Lernens und des Gedächtnisses mit vielen empirisch gewonnenen Daten nachweisen. Nicht zuletzt ist der in diesem Werk dokumentierte Wissensstand auch den sehr erfolgreichen Entwicklungen sowohl in der Hardware als auch in der Software zu verdanken. Beide Fortschritte machten es möglich, die noch von Rosenblatt (1958) und Minsky und Papert (1969) erfahrenen Schwierigkeiten zu überwinden.

Die Eigenschaften von künstlichen neuronalen Netzwerken dieser „zweiten“ Generation, die auf Computern realisiert werden können, lassen sich nach Eysenck und Keane ([11], 1990) in acht Punkten zusammenfassen, die hier in deutscher Übersetzung mit Hervorhebungen wie im Original wiedergegeben werden.

- Ein Netzwerk besteht aus Elementen, die auch als *units* (künstliche Neurone) oder *nods* (Knoten) bezeichnet werden. Sie sind untereinander verbunden, so dass jedes einzelne Neuron mit vielen anderen Neuronen in Kontakt steht.
- Neuronen beeinflussen andere Neuronen, in dem sie sich entsprechend ihrer gewichteten Verbindungen stimulieren oder hemmen.
- Das Neuron verwendet üblicherweise eine gewichtete Summe über alle Inputverbindungen und produziert einen einzigen Output zu einem anderen Neuron, wenn die gewichtete Summe den Wert der in der Zelle vorhandenen Schwelle übersteigt.
- Das Netzwerk als Ganzes wird von den Eigenschaften der Neuronen bestimmt, aus denen es besteht und der Art und Weise, in denen die Neuronen miteinander verbunden sind. Zusätzlich beeinflussen auch die Algorithmen, die die Verbindungen zwischen den Neuronen verändern können, die Eigenschaften des Netzwerkes.

- Netzwerke können verschiedene Strukturen oder Schichten haben und bestehen in der Regel aus Input-Schichten, Zwischenschichten (hidden units) und Output-Schichten.
- Die Wissensrepräsentation, d. h. die Repräsentation eines Konzeptes, wird in einem über etliche Neurone verteilten Muster gespeichert, und zwar als ein Muster von Aktivation im Netz.
- Ein einziges Netzwerk kann in dieser Weise viele verschiedene Muster speichern ohne dass sie sich gegenseitig stören, wenn sie hinreichend unterschiedlich sind.
- Ein Algorithmus zur Auswertung von falschen Antworten, *backward propagation of errors*, ermöglicht das Lernen in Netzwerken.

Die genannten Prinzipien sollen beispielhaft an einem einfachen und wenig komplexen neuronalen Netzwerk demonstriert werden, das von Rumelhart und McClelland (1986, [40], Vol. 1) publiziert worden ist. – Was passiert, wenn auf ein Neuron des Netzwerkes ein Aktivierungsimpuls trifft? Von allen Neuronen, mit denen dieses Neuron verbunden ist, kann es exzitatorische oder inhibitorische Impulse erhalten. Von all diesen erhaltenen Impulsen bildet das Neuron eine gewichtete Summe. Nur wenn diese gewichtete Summe die im Neuron enthaltene Schwelle übersteigt, feuert das Neuron und gibt damit Aktivation an ein anderes Neuron weiter oder auch an mehrere andere. Diese nun getroffenen Neuronen leiten die Aktivierung entsprechend ihrer eigenen Schwellen und gewichteten Verbindungen weiter. Damit breitet sich über das Netz ein spezifisches Aktivationsmuster aus. Wird als letztes Element eine *output unit* getroffen, verhält sich auch dieses Neuron wie alle Vorgänger: Wird der Schwellenwert übertroffen, kommt es zu einem Output und damit zu einer Ergebnisausgabe. Wenn der Output den Erwartungen entspricht, wird das Netz in einem erneuten Durchgang ebenso verfahren und dieselbe Antwort ausgeben, weil es nun über ein Netz von Aktivierungs- und Hemmungsimpulsen verfügt, das als Muster die passende Antwort produziert. Voraussetzung für die Funktionsfähigkeit eines Netzes ist das Vorhandensein mehrerer, häufig sogar vieler Schichten mit Neuronen, wobei den *hidden units* zur Feinabstimmung und Modulation des Outputs eine besondere Rolle zukommt. Stabilisierte Muster, die zu einer richtigen Antwort führen, erstrecken sich meistens über sehr viele oder auch alle Schichten von Neuronen. Dieser Prozess der Antwortgenerierung, der durch eine parallele Verarbeitung gekennzeichnet ist und auf keiner Stufe irgendeine Übersetzung auf eine symbolische Ebene erfordert, also subsymbolisch abläuft, kann zu der falschen Annahme verführen, dass das Netz über eine Regel verfüge, mit der es einen bestimmten Input in einen erwarteten Output überführt. Tatsächlich ist es aber so, dass es in derartigen neuronalen Netzwerken keine Regeln zur Generierung eines erwarteten Outputs gibt. Herbeigeführt wird das Ergebnis allein durch eine Aktivationsausbreitung im Netz gemäß bestehender Neuronenverbindungen und definierter Schwellen. Abbildung 2 veranschaulicht ein neuronales Netzwerk.

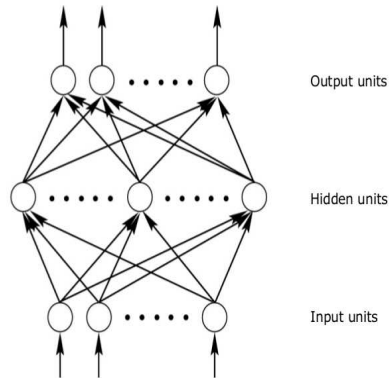


Abbildung 2. Neuronales Netz

Aus dem beschriebenen Prozess der Antwortgenerierung in einem neuronalen Netz wird deutlich, dass es nur sehr selten einmal und nur ganz zufällig auf Anhieb zu einer richtigen Antwort bzw. zu einer zutreffenden Aktivierung in der Output-Schicht des Netzes kommen kann. Das Netz muss erst lernen, ein Muster aufzubauen, das tatsächlich zu einer richtigen Antwort führt. Erfahrungen mit neuronalen Netzen haben gezeigt, dass häufig viele hundert oder auch tausende Durchläufe nötig sind, bis eine Netzanordnung sich gebildet und stabilisiert hat, die zu einer erwarteten Antwort führt. Das hat in früheren Zeiten wegen des damit verbundenen Rechenaufwandes bei begrenzter Computerkapazität zu Problemen geführt, muss allerdings heute mit der Verfügung über leistungsstarke Rechner kein Nachteil mehr sein. Entscheidender ist die Frage, wie kann das neuronale Netz lernen, das „richtige“ Muster zu bilden?

Eine sehr häufig verwendete Methode, das neuronale Netz lernen zu lassen, ist der Algorithmus *backward propagation of errors*. Diese Vorgehensweise vergleicht eine durch das Netz erhaltene Antwort aus der Output-Schicht mit der erwarteten Antwort und stellt Differenzen oder eine Übereinstimmung fest. Werden Differenzen festgestellt, erfolgt eine Rückmeldung an das Netz, das sein Aktivierungsmuster ändert. Durch immer wieder neue Durchläufe erfolgt auf diesem Wege eine Justierung des Aktivitätsmusters, bis keine Differenz mehr festgestellt wird und eine richtige Antwort erzeugt wurde. Wie oben schon erwähnt, kann so ein Prozess eine sehr große Zahl von Korrekturdurchläufen erfordern.

Die beschriebene Funktionsweise von neuronalen Netzen sowie der durch *backward propagation of errors* herbeigeführte Lernvorgang zum Aufbau eines optimal passenden Aktivitätsmusters zeigt deutliche Parallelen zu der von Hebb (1949) aufgestellten Regel zum Lernen in Neuronenstrukturen im Gehirn auf, die sowohl durch wiederholte Aktivierung eine Verstärkung der Verbindungen zwischen den Elementen konstatiert als auch die Erhöhung der Komplexität der Neuronenstruk-



turen mit zunehmenden Anforderungen vorhersagt. Diese Übereinstimmungen sind auf der Ebene von Analogien bemerkenswert. Ob damit allerdings eine Bezeichnung wie *brain-style modeling* gerechtfertigt werden kann, erscheint zweifelhaft.

Künstliche neuronale Netze haben eine sehr große Verbreitung und auch vielfältige Anwendung erfahren. Seit dem Werk von Rumelhart und McClelland (1986), in dem schon sehr viele Beispiele für die subsymbolische Modellierung kognitiver Prozesse enthalten sind, scheint diese Modellierungsart insbesondere seit der gigantischen Zunahme der Hirnforschung immer mehr an Bedeutung zu gewinnen und die konkurrierenden symbolischen Informationsverarbeitungsmodelle in Anzahl und Bedeutung zu überholen. – Ein frühes Beispiel sei hier zur Illustration angeführt. Sejnowski und Rosenberg ([44], 1987) haben ein konnektionistisch aufgebautes System publiziert, das sie *NETalk* genannt haben. Dieses System ist in der Lage, bei Eingabe eines englischsprachigen Textes als Ausgabe englisch gesprochene Sprache zu produzieren. *NETalk* ist also in der Lage, auf einem akzeptablen Niveau vorzulesen, wobei allerdings die Anzahl der ausgesprochenen und trainierten Wörter noch begrenzt ist. Die Angemessenheit der produzierten Sprache wird mit 90% angegeben.

### 2.2.3 Kritik an der subsymbolischen Informationsverarbeitung

Die mit der Entstehung der Forschung über künstliche neuronale Netze geäußerte Kritik eines unangemessenen Rechenaufwandes, bis sich die Aktivationsmuster zu einer erwarteten Antwort stabilisiert haben, kann heute mit der Verfügung über leistungsfähige Rechner beiseite gelegt werden. Ernster zu nehmen ist die Kritik von Fodor und Pylyshyn ([12], 1988) an der subsymbolischen Simulation kognitiver Prozesse. Sie haben diese Forschungstradition als einen Rückfall in das Black-Box-Paradigma bezeichnet, weil mit den Elementen der Netzwerke sowie mit den untereinander vorhandenen Verbindungen keinerlei semantische Bedeutungszuweisungen verbunden sind. Weder die künstlichen Neuronen noch die zwischen ihnen bestehenden Verbindungen und auch nicht die durch das Netz fließende Aktivierung steht für irgend etwas, ihnen kommt allein der Rang von technischen Größen zu.

Schließlich werden der subsymbolischen Simulation auch Schwierigkeiten bei der Aufklärung höherer kognitiver Prozesse wie z. B. der Selbstreflexion und des Verstehens zugeschrieben. Diese Kritik wird bei der Darstellung einer neurowissenschaftlich orientierten Informationsverarbeitungspsychologie wieder aufzunehmen sein.



### 3 Kognitive Neurowissenschaften

In den vorangegangenen Abschnitten wurden die in der Kognitionspsychologie verwendeten Forschungsparadigmen dargestellt und diskutiert, die, wenngleich zu abstrakten, so doch zu inhaltlich ausformulierten Modellen für kognitive Informationsverarbeitung führen. Es handelt sich dabei also um Realisierungen von Theorien, die auf Computern als lauffähige Programme implementiert werden können, um die zugrundeliegenden Theorien über kognitive Prozesse und Strukturen und deren Eigenschaften sowie auch deren dynamischen Veränderungen detailliert kennenzulernen und auch testen zu können. In weiteren Schritten ist es danach jedoch zwingend notwendig, die Validität der Modelle an von Probanden experimentell gewonnenen Daten zu überprüfen. Dabei ist das Unterfangen, komplexe Simulationsmodelle auf ihre Validität hin exakt zu prüfen, in der Regel sehr schwierig. Hierzu können z. B. Daten herangezogen werden, die mit der Methode des lauten Denkens erhoben worden sind und eher als „weiche“ Daten gelten, wenn man dazu den Vergleich mit Messdaten wie Reaktionszeiten heranzieht. In solchen Forschungsansätzen, die der symbolischen und auch der subsymbolischen Informationsverarbeitung zuzuordnen sind, spielen Rückgriffe auf neurowissenschaftliche Ergebnisse wie z. B. aus der Hirnforschung kaum eine Rolle. Zwar werden hier trotzdem Interpretationen nicht selten in neurowissenschaftlichen Begriffen vorgenommen, dabei handelt es sich aber in der Regel nur um Analogien, sowie sie z. B. bei der subsymbolischen Informationsverarbeitung mit dem hervorgehobenen Anspruch der Parallelverarbeitung zu finden sind. Auch die dort postulierten Netzwerke bestehen nicht etwa aus „echten“ Neuronen, sondern sind Konstrukte, die einige wenige Eigenschaften von lebenden Neuronen abbilden. Es bleibt also als wichtig festzuhalten, dass wir es bei solchen Computersimulationen mit abstrakten Realisierungen von Theorien zu tun haben und die Eigenschaften einzelner simulierter Komponenten allenfalls als analog zu neurophysiologischen lebenden Systemen zu bezeichnen sind.

Nun gibt es eine lange Tradition in der Wissenschaft, auch Fragen von der Art nachzugehen „Welche Hirnstrukturen ermöglichen kognitive Leistungen wie z. B. das Denken oder das Lernen?“ Dass das Gehirn der Ort und das Organ ist, an bzw. in dem kognitive Prozesse zustande kommen und ablaufen, ist eine alte und unbestrittene Erkenntnis. Als schwierig umzusetzen erwies sich jedoch der zusätzliche Wunsch, über die Arbeit des Gehirns selber verlässliche Daten und Einsichten zu gewinnen. Zwar ergab sich insbesondere nach Kriegen die Möglichkeit, Hirnverletzte mit den ihnen noch verbliebenen kognitiven Fähigkeiten zu untersuchen bzw. Ausfälle in deren Leistungen mit verletzungsbedingten hirnphysiologischen Defiziten in Verbindung zu bringen. Auch ergänzten Beobachtungen bei neurodegenerativen Erkrankungen Einsichten in die Funktionen und Arbeitsweisen des Gehirns. Schließlich ermöglichten auch experimentell gesetzte Läsionen in Tierexperimenten – wie dies z. B. von Lashley (1929) unternommen worden war – weitere Erkenntnisse über die Hirntätigkeit. Große Durchbrüche erzielte man jedoch erst im 20. Jahrhundert, als Hans Berger in Jena im Jahre 1924 [6] die Methode der Messung von Hirnströmen, das Elektroenzephalogramm (EEG), entwickelte (pu-

bliert erst im Jahre 1929). Eine wichtige Weiterentwicklung stellten Verfahren dar, mit denen elektrische Aktivitäten von einzelnen Zellen im Gehirn abgeleitet werden konnten, wie das z. B. Hubel und Wiesel ([16], 1959) in Tierexperimenten bei der Untersuchung des visuellen Kortex gelang.

Im letzten Jahrzehnt des 20. Jahrhundert gelangen mit der Erfindung von bildgebenden Verfahren, der Positronen-Emissions-Tomographie (PET) und der Technik des Functional magnetic resonance imaging (fMRI), bis dahin niemals gekannte Einblicke in die Arbeitsweise des menschlichen Gehirns. Zudem sind diese neuen Methoden nichtinvasiv, sie können also die untersuchten Gehirne nicht verletzen. Mit PET und fMRI hat sich die Chance ergeben, dem Gehirn quasi bei der Arbeit zuzusehen, wenn es mit der Vorbereitung und Ausführung kognitiver Leistungen beschäftigt ist. Diese nun neu gewonnenen methodischen Beobachtungsmöglichkeiten lassen es auch zu, Informationsverarbeitungsprozesse im Gehirn mit bisher nicht gekanntem Auflösungsvermögen zu untersuchen. Diese Fortschritte haben sich die kognitiven Neurowissenschaften zunutze gemacht und damit eine substantielle neue Forschungsdisziplin aus der Taufe gehoben. Bevor wir uns zur Illustration einem Beispiel einer derartig untersuchten und aufgeklärten kognitiven Leistung mit den daran beteiligten Informationsverarbeitungsprozessen zuwenden, müssen einige Überlegungen von grundsätzlicher Art angestellt werden, welchen prinzipiellen Erkenntnisgewinn dieser neue Forschungszugang erbringen kann.

Mit bildgebenden Verfahren erhält der Wissenschaftler bildliche Darstellungen vom Gehirn, in die Markierungen eingetragen sind, welche Regionen dieses Organs neuronal aktiviert sind. Erfasst werden mit diesen Methoden Veränderungen im Stoffwechsel des Gehirns sowie Eigenschaften des Blutflusses, wobei insbesondere steigende Sauerstoffgehalte im Blut registriert werden. Dieser Vorgehensweise liegt die Annahme zugrunde, dass Gebiete im Gehirn, die aktuell mit der Verarbeitung von Informationen beschäftigt sind, einen erhöhten Stoffwechselumsatz zeigen und eine gesteigerte Sauerstoffzufuhr erhalten müssen. Es handelt sich dabei also um indirekte Messungen der neuronalen Aktivitäten, die von der Annahme ihren Ausgang nehmen, dass neuronale Aktivitäten mit den genannten Veränderungen einhergehen. Auf diese Weise lässt sich ein Gehirn z. B. bei der Lösung einer gestellten Aufgabe über umschriebene Zeitdistanzen „bei der Arbeit“ beobachten. Als Ergebnis erhält man Markierungen für die neuronalen Aktivitäten in unterschiedlichen Gehirnregionen, die während des Lösungsprozesses auch in der tatsächlichen zeitlichen Abfolge erfasst werden konnten (wenn auch mit einer gewissen zeitlichen Verzögerung, wie weiter unten ausgeführt wird).

Welche Erkenntnisprozesse sind nun neu, die durch bildgebende Verfahren ermöglicht werden? Aus Messungen z. B. der Hirnströme aber auch durch Verletzungsausfall bestimmter Hirnteile war schon lange vorher bekannt, dass das Gehirn modular aufgebaut ist, das heißt, dass verschiedene Hirnregionen unterschiedliche Aufgaben erfüllen. So konnte dem präfrontalen Kortex die Verarbeitung semantischer Inhalte zugeschrieben werden, der linken Hemisphäre des Gehirns die sprachliche Verarbeitung und der rechten Hemisphäre die bildliche Verarbeitung zugeordnet werden, und die visuelle Wahrnehmung wurde im Okzipitallappen lokalisiert. Diese beispielhaft genannten bekannten Verortungen von Informationsver-

arbeiten im Gehirn können natürlich auch ebenso mit bildgebenden Verfahren erhoben und verfolgt werden, sowie sie auch schon vorher aus der Neuro- und Experimentalpsychologie zugeordnet werden konnten. Neu hinzugekommen ist neben genaueren Lokalisationen der neuronal aktiven Regionen auch ein über längere Zeiträume verfolgbarer Ablauf von Gehirnaktivitäten unter genauer Dokumentation der beteiligten Regionen. Und schließlich bietet diese Aufzeichnungsmöglichkeit auch den Blick in tiefer gelegene Hirnregionen, die also unterhalb des Kortex liegen, um damit ein vollständigeres Gesamtbild der Arbeitsweise des Gehirns zu erhalten. Demgegenüber liefern die Messungen der Hirnströme lediglich Aktivitätsmuster vom Kortex.

Kognitive Neurowissenschaften nutzen diese beschriebenen methodischen Zugänge, indem sie die in einem Informationsverarbeitungsmodell angenommenen und miteinander verknüpften Prozesse in Bezug setzen zu den beobachtbaren Hirnaktivitäten und mit bildgebenden Verfahren erheben, ob dieselben Prozesse auch in der angenommenen Reihenfolge, Ausprägung und Interaktion im Gehirn beobachtbar sind. Dazu müssen die im Modell angenommenen Komponenten der Informationsverarbeitung auch als lokalisierbare Regionen im Gehirn bekannt sein. Es geht also im Prinzip darum, für die im Modell postulierten Informationsverarbeitungsprozesse neuronale Entsprechungen im Gehirn zu finden, die dann ihrerseits in ihrer inhaltlichen Funktion genauer spezifiziert werden können. Damit können Aussagen getroffen werden, welche Informationsverarbeitungsprozesse an bestimmten Orten des Gehirns ablaufen, wenn in den bildgebenden Verfahren entsprechende Aktivitätsmuster auftauchen. Umgekehrt können Einsichten darüber erhalten werden, an welchen Orten des Gehirns bestimmte Arten der Informationsverarbeitung geleistet werden können und in welcher Weise Interaktionen zwischen den verschiedenen Verarbeitungszentren ablaufen. Denn Untersuchungsbefunde mit bildgebenden Verfahren klären zwar auf, dass in bestimmten Regionen des Gehirns Aktivierungen stattfinden, sie können aber nicht nachweisen, welcher Art die Informationen sind, die dort verarbeitet werden. Informationsverarbeitungsmodelle postulieren demgegenüber detaillierte Hypothesen darüber, welche Art von Informationen in bestimmten Prozessen verarbeitet werden ohne zu wissen, ob und wo sie im Gehirn des Menschen aufzufinden sind. Erst aus der Zusammenschau beider Ergebnisquellen lassen sich weiterführende Erkenntnisse ableiten, die unser Verständnis über kognitive Prozesse beim Menschen vervollständigen können. Und dieser Anspruch wird in den kognitiven Neurowissenschaften verfolgt und bearbeitet. Man kann diese Forschungsrichtung auch als den Versuch bezeichnen, Synergien zwischen einer rein theoretischen Informationsverarbeitungspsychologie einerseits und einer neurowissenschaftlichen Hirnforschung andererseits herzustellen und zu nutzen. Im Folgenden soll nun an einem Beispiel das synergetische Zusammenspiel der unterschiedlichen kognitiven Wissenschaften verdeutlicht werden.

Anderson (2007) beklagt in seinem Lehrbuch [3], dass die meisten neurowissenschaftlichen Untersuchungen von kognitiven Prozessen mit zu einfachen mentalen Anforderungen arbeiten wie z. B. das Erkennen von senkrechten und waagerechten Strichen bei der Aufdeckung rezeptiver Felder von Neuronen im visuellen Kortex

durch Hubel und Wiesel (1959). Als weiterführend zitiert er eine Untersuchung von Qin et al. ([36], 2004), die diesen Mangel überwindet und als Untersuchungsgegenstand eine komplexere mentale Anforderung heranzieht. Es handelt sich dabei um das Lösen von mathematischen Gleichungen, die Kindern im Alter von 11 bis 14 Jahren zur Bearbeitung vorgelegt wurden.

Qin et al. (2004) legten den Schülern Gleichungen von unterschiedlicher Komplexität vor. Die Unterschiede wurden durch die Anzahl der notwendigen Umformungen definiert, die zur Lösungsfindung nötig waren. Es wurden drei Aufgabentypen verwendet.

- (1) Null Umformungsschritte:  $1x + 0 = 4$   
Diese Gleichung erfordert keine Umformung, um die Lösung  $x = 4$  zu finden.
- (2) Ein Umformungsschritt:  $3x + 0 = 12$   
Die Umformung beinhaltet die Division als einen Schritt:  $12/3 = 4$  sowie den Abruf des entsprechenden mathematischen Wissens.
- (3) Zwei Umformungsschritte:  $7x + 1 = 29$   
Der erste Schritt der Umformung besteht in der Umkehr des Pluszeichens:  $29 - 1 = 28$  mit dem Abruf des relevanten Fachwissens. Der zweite Schritt beinhaltet die Division:  $28/7 = 4$ , ebenfalls verbunden mit dem Abruf des einschlägigen Wissens.

Die ungewöhnliche Schreibweise musste eingeführt werden, damit der visuelle Komplexitätsgrad aller drei Aufgabentypen untereinander vergleichbar blieb, um unterschiedliche Verarbeitungsprozesse zwischen den Aufgaben im visuellen Kortex zu vermeiden. Als Antwortmöglichkeit erhielten die Kinder einen speziellen Handschuh übergezogen, so dass sie ihre Antwort durch Drücken mit den verschiedenen Fingern signalisieren konnten: 1 = Daumen; 2 = Zeigefinger; 3 = Mittelfinger; 4 = Ringfinger; 5 = kleiner Finger.

Das von Qin et al. (2004) für diese Gleichungsaufgaben entwickelte Informationsverarbeitungsmodell unterscheidet drei Klassen von kognitiven Prozessen:

- (1) vorgestellte Umformungen der Gleichung (visuelle Verarbeitung);
- (2) Abruf arithmetischer und algebraischer Informationen;
- (3) Programmierung der motorischen Antwort für den Fingerdruck.

Während die Probanden die drei unterschiedlichen Aufgabentypen nacheinander lösten, wurden mit der fMRI-Methode ihre neuronalen Aktivitäten im Gehirn gemessen und aufgezeichnet. Dabei wurde darauf geachtet, zwischen den Aufgaben genügend Zeit zu lassen, um das Aktivierungsniveau im Gehirn vor jeder neuen Messung wieder auf den Ausgangszustand zurückkehren zu lassen.

Aufgrund bekannter Ergebnisse aus der Bildgebungsforschung wurden ausgehend von dem aufgestellten Informationsverarbeitungsmodell folgende Erwartungen bezüglich Gehirnaktivitäten formuliert:

- die Vorbereitung der Handbewegung zur Abgabe einer Antwort sollte sich im linken motorischen und somatosensorischen Kortex zeigen, da die rechte Hand verwendet wurde;
- die vorstellungsmäßig durchgeführte Umformung jeder Gleichung müsste im linken parietalen Kortex zu Aktivierungen führen, da diese Region auch an bildlichen Vorstellungen beteiligt ist;
- der Abruf arithmetischer und algebraischer Informationen sollte im linken präfrontalen Kortex zu Aktivierungen führen.

Die gefundenen Lösungszeiten für die drei Aufgabentypen betragen 2,0 Sekunden für null Umformungsschritte, 3,6 Sekunden für einen Schritt und 4,8 Sekunden für zwei Schritte. Alle 1,2 Sekunden wurden sog. BOLD (blood oxygen level dependent)-Antworten erhoben, und zwar für jede der benannten Gehirnregionen. Diese BOLD-Antworten wurden in prozentualen Änderungen des Aktivierungsniveaus über dem Grundniveau gemessen. Je höher ein BOLD-Wert ausfällt, je größer ist die Aktivierung. Dabei ist weiterhin zu beachten, dass die hämodynamische Antwort, die den Sauerstoffanteil im Blut anzeigt, Zeit benötigt, ehe sie an der neuronal aktiven Stelle im Gehirn gemessen werden kann. Üblicherweise erreicht diese hämodynamische Antwort erst vier bis fünf Sekunden nach dem auslösenden Ereignis ihr Maximum, also – je nach Schwierigkeit der Aufgabe – auch erst nachdem der Proband schon seine Antwort abgegeben hat.

In diesem Experiment wurden für die motorischen Hirnregionen die Maxima der BOLD-Werte bei Null-Schritt-Aufgaben erst nach 6,6 Sekunden, bei Ein-Schritt-Aufgaben nach 7,9 Sekunden und bei Zwei-Schritt-Aufgaben nach 9,9 Sekunden erreicht. Für alle drei Aufgabentypen zeigte sich hier dieselbe Ausprägung der Aktivierung. Das ist plausibel, da die Vorbereitung des Fingerdrucks zur Antwortabgabe bei allen Aufgabentypen dieselbe ist.

Die BOLD-Antworten in der parietalen Region des Gehirns erreichten, wie auch schon in der motorischen Region, zu unterschiedlichen Zeiten ihr Maximum, was – wie erwähnt – den unterschiedlichen Zeiten geschuldet, die zur Lösung der Gleichungen benötigt wurden. Außerdem erreichten die BOLD-Werte in der parietalen Region ihre Maxima früher als in der motorischen Region. Damit wird angezeigt, dass die mental geforderten Umformungen der Gleichungen vor der Antwortvorbereitung stattgefunden haben, was wiederum plausibel ist. Schließlich fielen die maximalen Höhen der BOLD-Messungen unterschiedlich aus, worin sich die Abstufungen der Schritte zur Umformung der Gleichungen widerspiegeln, die in den drei Aufgabentypen erforderlich sind.

Die BOLD-Werte in der präfrontalen Region ähnelten denjenigen aus den parietalen Gehirnregionen. Allerdings bestand eine Ausnahme darin, dass hier bei der Null-Schritt-Aufgabe kein Ansteigen der Messwerte gefunden wurde. Dieses

Ergebnis zeigt an, dass es für die Lösung der leichtesten Aufgabe unnötig war, arithmetisches Faktenwissen zur Lösung abzurufen. Die Antwort folgte direkt aus der vorhandenen Repräsentation dieser Aufgabe. Für die Ein-Schritt- und Zwei-Schritt-Aufgaben mussten hingegen zusätzlich mathematisches Wissen abgerufen werden, das erst eine Lösung ermöglichte. Das spiegelt sich in den jeweils erhöhten Messwerten in der präfrontalen Region wider.

Das beschriebene Experiment zeigt, dass die fMRI-Methode aufgrund ihrer Möglichkeit zur feinkörnigen Auflösung räumlicher Gegebenheit im Gehirn sehr gut in der Lage ist, hypothetisch angenommene Informationsverarbeitungsprozesse in ihrer neuronalen Aktivität zu unterscheiden sowie den zeitlichen Ablauf der kortikalen Verarbeitung zu dokumentieren. Das vorangegangene Beispiel eines Experimentes zeigt zudem, wie durch Synergieeffekte, die aus der Kombination von rein theoretischen, jedoch inhaltlich fundierten Modellannahmen und neurowissenschaftlichen Methoden Erkenntnisgewinne erzielt werden können, die keiner der beiden Forschungsansätze allein erbringen kann. Dieser Weg wird in den kognitiven Neurowissenschaften verfolgt, wenngleich – noch – Grenzen einzelner Methoden dieser Forschungsrichtung offensichtlich sind, wie das geringe zeitliche Auflösungsvermögen der fMRI-Methode.

## Literatur

- [1] J. R. Anderson. *The architecture of cognition*. Harvard University Press, Cambridge MA, 1983.
- [2] J. R. Anderson. *Kognitive Psychologie. 3. Auflage*. Spektrum Verlag, Heidelberg, 2001.
- [3] J. R. Anderson. *Kognitive Psychologie. 6. Auflage*. Spektrum Verlag, Heidelberg, 2007.
- [4] A. Baddeley. *Human memory. Theory and practice*. Psychology Press Ltd., Hove, 1997.
- [5] F. Bartlett. *Remembering*. Cambridge University Press, Cambridge, 1932.
- [6] H. Berger. Über das Elektroencephalogramm des Menschen. *Archiv für Psychiatrie*, 87:527–570, 1929.
- [7] J. Bredenkamp, K. M. Klein, S. von Hayn, und B. Vaterrodt. Gedächtnispsychologische Untersuchungen eines Rechenkünstlers. *Sprache & Kognition*, 7:69–83, 1988.
- [8] D. E. Broadbent. *Perception and communication*. Pergamon Press, London, 1958.
- [9] D. Dörner. *Bauplan für eine Seele*. Rowohlt, Reinbek, 1999.

- 
- [10] D. Dörner, C. Bartl, F. Detje, J. Gerdes, D. Halcour, H. Schaub, und U. Starker. *Die Mechanik des Seelenwagens*. Huber, Bern, 2002.
- [11] M. W. Eysenck and M. T. Keane. *Cognitive psychology. A student's handbook*. Erlbaum, Hove, 1990.
- [12] J. A. Fodor and Z. W. Pylyshyn. Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28:3–71, 1988.
- [13] K. Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. *Monatshefte der Mathematik und Physik*, 38:173–198, 1931.
- [14] D. O. Hebb. *The organization of behavior*. Wiley, New York, 1949.
- [15] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, 79:2554–2558, 1982.
- [16] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*, 148:574–591, 1959.
- [17] F. Klix. *Information und Verhalten*. Huber, Bern, 1971.
- [18] T. Kuhn. *Die Struktur wissenschaftlicher Revolutionen*. Suhrkamp, Frankfurt, 1967.
- [19] R. Lachman, J. L. Lachman, and E. C. Butterfield. *Cognitive psychology and information processing: An introduction*. Lawrence Erlbaum, Hillsdale, NJ, 1979.
- [20] K. S. Lashley. *Brain mechanisms and intelligence: A study of injuries to the brain*. University of Chicago Press, Chicago, 1929.
- [21] G. Lüer. *Gesetzmäßige Denkabläufe beim Problemlösen*. Beltz, Weinheim, 1973.
- [22] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:135–137, 1943.
- [23] G. A. Miller. *Language and communication*. McGraw–Hill, New York, 1951.
- [24] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63:81–97, 1956.
- [25] M. L. Minsky. *Neural nets and the brain model problem*. PhD thesis, Princeton University, 1954.



- 
- [26] M. L. Minsky. A framework for representing knowledge. In P. H. Winston, editor, *The psychology of computer vision*, pages 211–277. McGraw Hill, New York, 1975.
- [27] M. L. Minsky and S. Papert. *Perceptrons*. M. I. T. Press, Cambridge, MA, 1969.
- [28] U. Neisser. *Cognitive psychology*. Appleton-Century-Crofts, New York, 1967.
- [29] U. Neisser. *Cognition and reality*. Freeman, San Francisco, 1976.
- [30] U. Neisser. Memory: What are the important questions? In U. Neisser, editor, *Memory observed*. W. H. Freeman, New York, 1982.
- [31] A. Newell. You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase, editor, *Visual information processing*, pages 283–310. Academic Press, New York, 1973.
- [32] A. Newell and J. C. Shaw. Programming the logical theory machine. *Proceedings of the Western Joint Computer Conference*, pages 230–240, 1957.
- [33] A. Newell, J. C. Shaw, and H. A. Simon. Problem solving in humans and computers. *Carnegie Technical*, 21:34–38, 1957.
- [34] A. Newell and H. A. Simon. *Human problem solving*. Prentice-Hall, Englewood Cliffs, NJ, 1972.
- [35] I. P. Pavlov. *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Oxford University Press, Oxford, 1927.
- [36] Y. Qin, J. R. Anderson, E. Silk, V. A. Stenger, and C. Carter. The change of the brain activation patterns along with the children's practice in algebra equation solving. *Proceedings of the National Academy of Sciences USA*, 101:5686–5691, 2004.
- [37] S. Ramón y Cajal. *Histologie du système nerveux de l'homme et des vertébrates*. Maloine, Paris, 1911.
- [38] F. Rosenblatt. The Perceptron: A probabilistic mode for the information storage and organization in the brain. *Psychological Review*, 65:386–407, 1958.
- [39] D. E. Rumelhart. Notes on a schema for stories. In D. G. Bobrow and A. Collins, editors, *Representation and understanding*, pages 211–236. Academic Press, New York, 1975.
- [40] D. E. Rumelhart and J. L. McClelland. *Parallel distributed processing Vol. 1: Explorations in the microstructure of cognition*. M. I. T. Press, Cambridge, MA, 1986.



- 
- [41] D. E. Rumelhart and J. L. McClelland. *Parallel distributed processing Vol. 2: Explorations in the microstructure of cognition*. M. I. T. Press, Cambridge, MA, 1986.
- [42] R. C. Schank. *Conceptual information processing*. North Holland, Amsterdam, 1975.
- [43] J. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3:417–424, 1981.
- [44] T. J. Sejnowski and C. R. Rosenberg. Parallel networks that learn to pronounce English text. *Complex systems*, 1:145–168, 1987.
- [45] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.
- [46] C. E. Shannon und W. Weaver. *Mathematische Grundlagen der Informationstheorie*. R. Oldenbourg Verlag, 1976.
- [47] S. Sternberg. High-speed scanning in human memory. *Science*, 153:652–654, 1966.
- [48] G. Strube. Neokonnektionismus: Eine neue Basis für Theorie und Modellierung menschlicher Kognition? *Psychologische Rundschau*, 41:129–143, 1990.
- [49] E. C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55:189–208, 1948.
- [50] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematics Society (Series 2)*, 42:230–265, 1936.
- [51] A. M. Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950.
- [52] J. B. Watson. Psychology as a behaviorist views it. *Psychological Review*, 20:158–177, 1913.
- [53] A. N. Whitehead and B. Russell. *Principia Mathematica*. The University Press, Cambridge, 1935.



# Informationsverarbeitung in der Hörbahn von Insekten

*Bernhard Ronacher*

*Institut für Biologie, Humboldt-Universität zu Berlin*

*„Vision is the process of discovering from images what is present in the world, and where it is. Vision is therefore, first and foremost, an information processing task, but we cannot think of it just as a process. For if we are capable of knowing what is where in the world, our brains must somehow be capable of representing this information. . . . The study of vision must therefore include . . . also an inquiry into the nature of internal representations by which we capture this information and thus make it available as a basis for decisions about our thoughts and actions. This duality – the representation and the processing of information – lies at the heart of most information processing tasks and will profoundly shape our investigation of the particular problem posed by vision.“ (aus David Marr 1982, p. 3)*

## Zusammenfassung

Um ihre Umwelt sinnvoll interpretieren und in ihr agieren zu können, um Objekte zu identifizieren und zu lokalisieren, stehen den Nervensystemen von Tieren nur die Meldungen diverser Sinnesorgane zur Verfügung, die im zentralen Nervensystem als Folgen von Aktionspotentialen ankommen. Diese Meldungen müssen in mehreren Verarbeitungsschritten richtig interpretiert werden, um eine passende Repräsentation der Umwelt zu konstruieren. Zusätzlich zu den akut eintreffenden Informationen sind jedoch auch gespeicherte Informationen essentiell für diese Aufgabe – ohne „Vorwissen“ wäre sie wegen der Nicht-Eindeutigkeit von Sinnesmeldungen über die Umwelt wohl unlösbar. Solches Vorwissen kann im Laufe der Individualentwicklung durch Lernvorgänge erworben werden, oder auch im Laufe der Evolution bereitgestellt werden, in Form angeborener Informationen und Strukturen.

Ein erstes schwieriges Problem bei der Objektrepräsentation und Objekterkennung durch Nervensysteme rührt daher, dass auch ein und dasselbe Objekt bei verschiedenen Gelegenheiten nicht die gleichen Meldungen in den Sinnesorganen hervorrufen wird. Mein Schreibtisch ruft in den Sehzellen meiner Augen ganz verschiedene Meldungen hervor, je nach Entfernung, Beleuchtung, oder Verdeckung

durch andere Gegenstände. Dennoch wird er dank vielfältiger Invarianzleistungen meines Gehirns stets als derselbe Gegenstand erkannt.

Ein zweites, grundsätzliches Problem für die Verarbeitung von Umweltreizen im Nervensystem liegt in der inhärenten Variabilität neuronaler Signale, die durch Zufallsprozesse beim Öffnen und Schließen von Ionenkanälen bedingt ist. Infolge dieses „intrinsic Rauschens“ unterscheiden sich die Antworten eines sensorischen Neurons bei mehrfacher Präsentation auch eines identischen externen Signals. Diese intrinsische Variabilität muss sich auf die Verarbeitung und Repräsentation sensorischer Information auswirken. Es ist jedoch nicht von vorneherein klar, wie gravierend die Auswirkungen dieses Rauschproblems sind, in Relation zu der zuvor erwähnten Vielfalt von Erscheinungsformen desselben Gegenstandes.

Diese Fragen werden am Beispiel der akustischen Kommunikation von Heuschrecken behandelt. Akustische Kommunikationssysteme eignen sich für die Untersuchung derartiger Fragen besonders gut, da die Tiere auf die Kommunikationssignale von Paarungspartnern zuverlässig antworten. Aus diesen Antworten können wir die Bedeutung verschiedener Signale, ihren Informationsgehalt für den Empfänger, erschließen. Ein weiterer Vorteil ist das relativ übersichtliche Nervensystem dieser Tiere, was es uns erlaubt, die Schritte der Codierung und Repräsentation von Umweltsignalen im Nervensystem in aufeinander folgenden Verarbeitungsstufen zu verfolgen und in detail zu analysieren.

Dieser Aufsatz ist dem Andenken an den Neuroethologen Prof. Dr. Norbert Elsner, Mitglied und langjähriger Vizepräsident der Göttinger Akademie und Förderer der Göttinger Neurobiologentagung, gewidmet.

## 1 Einleitung

Auch scheinbar sehr einfache Tiere wie Quallen oder Muscheln müssen bestimmte Umweltgegebenheiten wie Lichtverhältnisse, Temperatur, Objekte, erfassen – wenn auch letztere vielleicht nur grob – um sich adäquat verhalten zu können. Sobald vielzellige Organismen eine gewisse Größe erreichten, wurde auch die Koordination verschiedener Körperregionen zum Problem und ein schnelles System zur Informationsübertragung wurde wichtig. Für diese Aufgaben haben sich in der Evolution Nervensysteme entwickelt, welche die Koordination leisten und die Meldungen von Sensoren – Sinneszellen und Sinnesorganen – verarbeiten, interpretieren und über Effektoren – Muskeln und Drüsen – Verhaltensäußerungen ermöglichen.

Wie werden im zentralen Nervensystem (ZNS) die Informationen über relevante Umweltreize verarbeitet und repräsentiert? Dazu muss man sich in Erinnerung rufen, dass einem Gehirn an Information über die Umwelt nur die Meldungen von verschiedenen Sinnesorganen zur Verfügung stehen, die verarbeitet und interpretiert werden müssen. Diese Meldungen erreichen das Gehirn über verschiedene Nervenbahnen in Form von Folgen von Aktionspotentialen („Spiketrains“). Aktionspotentiale („Spikes“) bestehen in einer sehr kurzen lokalen Spannungsänderung an der Nervenzellmembran, die über das Axon der Nervenzelle weitergeleitet wird. Aktionspotentiale sind Alles-oder-Nichts-Ereignisse, die in verschiedensten

Nervenzellen sehr einheitliche Form haben. Sie stellen die grundlegenden Informationsträger in Nervensystemen dar, mittels derer Information schnell über größere Entfernungen übertragen werden kann. [Wie Aktionspotentiale zustande kommen und weitergeleitet werden, ist in einem Anhang kurz dargestellt.]

Welche Information ist in solchen „Spiketrains“ enthalten? Den neuronalen Code – oder die neuronalen Codes – des Nervensystems zu entziffern, ist eine der großen spannenden Fragen der Neurobiologie. Da man die im ZNS verwendeten Codierungsprinzipien nicht a priori versteht, lautet die Grundfrage, wie diese neuronalen Signale zu interpretieren sind. Anders formuliert: welche Bedeutung hat eine Folge von Aktionspotentialen, die von einem bestimmten Neuron weitergeleitet wird? Welche Information ist in den Spikeantworten eines sensorischen Neurons enthalten und wie kann sie von nachgeschalteten Neuronen „ausgelesen“ werden? So wird man fragen müssen, ob in den Meldungen von Sinnesorganen die wesentliche Information – etwa über die Intensität eines Licht- oder Schallreizes – durch die Spikerate, d. h. die über längere Zeit gemittelte Frequenz der Aktionspotentiale, übermittelt wird, oder ob es vielmehr auf die Muster von Aktionspotentialen beziehungsweise deren genauen Zeitpunkte ankommt?

Diese Probleme sollen am Beispiel der Verarbeitung akustischer Signale näher behandelt werden. Auch wenn die vorangestellten Zitate von David Marr sich auf den Sehvorgang bezogen, können sie ebenso auch als Ausgangspunkt für die hier behandelten Themen dienen. Auch beim Hören geht es letztlich darum, Objekte (Schallsignale) zu identifizieren und zu lokalisieren, wobei als aktuelle Informationen für diese Aufgaben nur die Meldungen der verschiedenen Sinnesrezeptoren in den Ohren zur Verfügung stehen. Diese müssen in mehreren Stufen der Verarbeitung richtig interpretiert werden, um eine adäquate Repräsentation der Umwelt zu generieren. Zusätzlich zu den akut eintreffenden Meldungen der Sinnesorgane sind jedoch auch gespeicherte Informationen essentiell; ohne „Vorwissen“ über die Umwelt wäre die Aufgabe wegen der Nicht-Eindeutigkeit und Variabilität der Sinnesmeldungen wahrscheinlich hoffnungslos. Solches „Vorwissen“ kann im Laufe der Individualentwicklung durch Lernvorgänge erworben werden, aber auch im Laufe der Evolution, in Form angeborener Verarbeitungsprinzipien und -strukturen, angelegt worden sein.

Einige Prinzipien neuronaler Verarbeitung werden hier exemplarisch, am Beispiel der Verarbeitung akustischer Kommunikationssignale im Hörsystem von Feldheuschrecken, behandelt. Diese Tiere bieten den großen Vorteil, dass es sich bei ihrer Kommunikation um ein völlig angeborenes Verhalten handelt, wir haben es also mit einem „festverdrahteten“ neuronalen Erkennungssystem zu tun. Außerdem haben die Heuschrecken ein – verglichen mit Wirbeltieren – recht übersichtliches Nervensystem, wobei viele der auditorischen Neurone anhand ihrer charakteristischen Struktur als Individuen identifiziert werden können (siehe Abb. 1). Man kann ein bestimmtes Neuron daher in jedem untersuchten Tier wiederfinden und synaptische Verbindungen zwischen Neuronen identifizieren, was die Analyse der Verarbeitungsschritte und des Informationsgehalts neuronaler Signale wesentlich erleichtert. So können wir schrittweise verfolgen, wie die Meldungen über die Umwelt auf verschiedenen Verarbeitungsstufen codiert und repräsentiert werden.

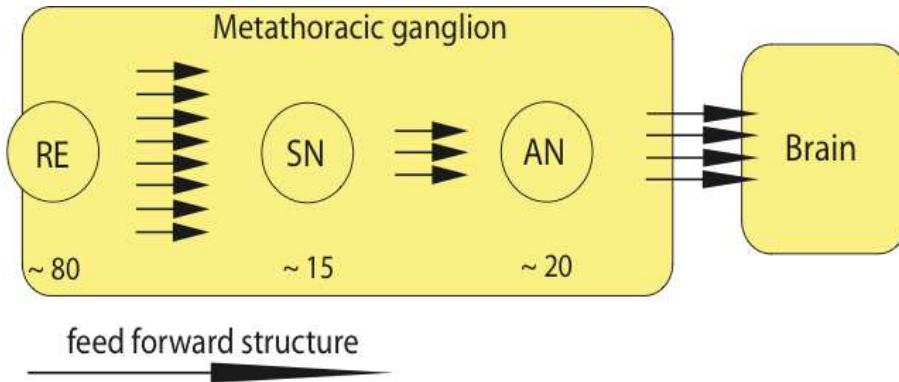


Abbildung 1. Die Hörbahn von Feldheuschrecken

Die beiden Ohren sitzen bei Feldheuschrecken seitlich am ersten Hinterleibs-Segment. Die Rezeptorzellen (RE) sitzen innen am Trommelfell und transformieren dessen Schwingungen in eine Folge von Aktionspotentialen. Die Axone der 60-80 Rezeptorzellen ziehen von jedem Ohr ins dritte Thorakalganglion (Metathorakalganglion) und kontaktieren dort eine Gruppe von etwa 10 – 15 lokalen Neuronen (SN), in denen erste wichtige Verarbeitungsschritte stattfinden. Wir können etwa 10 – 15 verschiedene lokale Neuronen als Individuen identifizieren und charakterisieren. Die lokalen Neurone ihrerseits geben die vorverarbeitete Information weiter an sogenannte aufsteigende Neurone (AN); nur deren Axone reichen bis ins Gehirn, wo die endgültige Entscheidung fällt, ob der Gesang eines potentiellen Partners als arteigen erkannt und akzeptiert wird. Die Gruppe der etwa 20 aufsteigenden Neurone stellt also einen Flaschenhals für die Information dar, die dem Gehirn für seine Entscheidung – Gesang akzeptieren oder nicht – zur Verfügung steht. Näheres in Stumpner und Ronacher (1991) und Ronacher und Stange (2013); Abbildung modifiziert nach Ronacher (2014).

Auch aus einem weiteren Grund sind solche Kommunikationssysteme für die Untersuchung neuronaler Informationsverarbeitung besonders geeignet: weil der Gehalt der übermittelten Nachricht genau bestimmt werden kann. Die Heuschrecken reagieren recht zuverlässig auf ihre artspezifischen Kommunikationssignale, die „Gesänge“ ihrer Partner; so haben wir durch ihre Verhaltensantwort einen klaren Indikator, welche Bedeutung verschiedene akustische Signale für den Empfänger haben (Elsner und Popov 1978; Gerhardt und Huber 2002).

Die Kommunikation der heimischen Heuschrecken dient der Identifizierung und Anlockung von arteigenen Sexualpartnern. Die Gesänge verschiedener Arten unterscheiden sich stark und ihre korrekte Erkennung ist eine wichtige Barriere gegen Hybridisierung. Bei der Verarbeitung der artspezifischen Signale steht also viel auf dem Spiel: eine Verpaarung mit einem artfremden Partner würde zu Hybriden führen, die stark reduzierte Fortpflanzungs- und Überlebenschancen haben. Ein Missgriff bei der Gesangserkennung hätte also massive Auswirkungen auf den Fortpflanzungserfolg, die entscheidende „Währung“ der Evolution. Man wird daher erwarten, dass das Nervensystem der Tiere auf diese so wichtige Aufgabe speziali-

siert ist, was die Untersuchung erleichtert – verglichen mit einem System, welches ein breites Spektrum an Aufgaben abdecken muss und vielleicht zusätzlich durch Lernvorgänge modifiziert wird.

An dieser Stelle möchte ich gleich vorweg betonen, dass man trotz der Größen- und Organisationsunterschiede der Nervensysteme bei diesen Insekten sehr ähnliche Verarbeitungs- und Codierungsprinzipien findet wie bei den viel größeren Nervensystemen von Wirbeltieren. Dieser vergleichende Aspekt wird weiter unten noch zur Sprache kommen.

Am Kommunikationsgeschehen ist noch ein weiterer Aspekt interessant: Weibliche Feldheuschrecken investieren sehr viel mehr in die Nachkommen als die Männchen. Während die Männchen nur Spermia liefern und keine weitere Investition in die Nachkommen leisten, erfordert die Produktion der großen Ei-Pakete viel Energie. Wegen dieser großen Investition sollten die Weibchen wählerisch sein und sich nur mit attraktiven Männchen verpaaren, von denen sie wenigstens „gute Gene“ erwarten können (Andersson 1994). Tatsächlich beobachtet man das Phänomen der „Damenwahl“ oder „sexuellen Selektion“ gegenüber arteigenen Männchen: die Weibchen lassen bestimmte Männchen bevorzugt zur Paarung zu, wobei der Gesang eines Männchen eine wichtige Rolle spielt. Das impliziert, dass die Tiere sehr subtile Gesangsmerkmale unterscheiden können, was uns im Zusammenhang mit der Zuverlässigkeit neuronaler Verarbeitung weiter unten noch beschäftigen wird.

Im folgenden Abschnitt (2) werden zunächst die intrinsische Variabilität neuronaler Signale und die damit verbundene Unsicherheit neuronaler Codierung im Zentrum stehen. Verschiedene Möglichkeiten werden diskutiert, wie die Folgen dieser Unzuverlässigkeit neuronaler Signale erfasst und minimiert werden können. Dazu wird ein Maß für die Variabilität von Spiketrains, eine Metrik, eingeführt. Anhand dieses Maßes können wir prüfen, wie stark sich die intrinsische Variabilität in Relation zu externer Signalvariabilität auswirkt. Im dritten Abschnitt wird auf die Unterscheidbarkeit ähnlicher Signale im Kontext sexueller Selektion eingegangen werden. Sehr präzise Verhaltensantworten scheinen dabei im Widerspruch zu variablen neuronalen Antworten zu stehen. Im Abschnitt 4 erfolgt die Auflösung dieses scheinbaren Widerspruchs: es erfolgt offenbar eine Änderung des Codierungsprinzips, ein Übergang zu einem Populationscode mit „sparsamer“ Codierung. Ähnliche Codierungsprinzipien findet man auch bei den deutlich größeren Nervensystemen von Wirbeltieren, dort allerdings meist erst auf höheren Verarbeitungsstufen. Die bisherige Argumentation folgte einer bottom-up Strategie, in der die Verarbeitungsprinzipien einzelner oder weniger Neurone betrachtet wurden. Im fünften Abschnitt wird die Blickrichtung umgekehrt und mittels eines Modells – in einem top-down Ansatz – versucht, von den Verhaltensantworten auf neuronale Verarbeitungsschritte zurückzuschließen. Das vorgestellte Modell vermag mehrere bisher unverstandene Befunde zu erklären und liefert eine Reihe von interessanten Vorhersagen. Auf einen Ausblick folgt als Anhang noch eine kurze Einführung in die Generierung und Weiterleitung von Aktionspotentialen.

## 2 Variabilität neuronaler Signale: ein grundsätzliches Problem

Die elektrischen Signale des Nervensystems – Ruhepotential, synaptische Potentiale und Aktionspotentiale – beruhen auf Bewegungen von Ionen, die durch spezialisierte Kanäle in der Zellmembran fließen. Es gibt für die verschiedenen Ionsorten – besonders wichtig:  $K^+$ ,  $Na^+$ ,  $Cl^-$ ,  $Ca^{++}$  – jeweils eigene Kanäle, die unter bestimmten Bedingungen geschlossen oder geöffnet sind (siehe Anhang). Aus einer Vielzahl verschiedener Ionenströme resultiert das jeweilige Membranpotential, wobei zwei Zustände besonders hervorzuheben sind: das Ruhepotential – cum grano salis der Ruhezustand der Nervenzelle, in dem die Innenseite der Zellmembran negativ geladen ist – und das Aktionspotential, das als eine sehr schnelle, stereotype Spannungsänderung auftritt, bei der die Zellinnenseite nun kurzzeitig positiv geladen ist. Diese Umpolung des Zellmembran-Kondensators wird durch das schnelle Öffnen von spezialisierten Ionenkanälen verursacht, durch die  $Na^+$ -Ionen in die Zelle und  $K^+$ -Ionen aus der Zelle wandern können (**siehe Anhang**). Wird ein Aktionspotential am Beginn eines Axons ausgelöst, so wandert es das Axon entlang, bis es schließlich die Kontaktstellen mit anderen Neuronen, die Synapsen, erreicht, wo die Meldungen an nachgeschaltete Neurone übertragen werden können und entscheidende Verarbeitungsschritte stattfinden.

Das Öffnen/Schließen eines Ionenkanals ist nun kein völlig determinierter Prozess, sondern erfolgt nur mit bestimmten Wahrscheinlichkeiten. Insbesondere unterliegen die Öffnungsdauern der Kanäle erheblichen Schwankungen, was zu unterschiedlich starken Ionenströmen führt. Zwar mitteln sich solche Schwankungen teilweise wieder aus, wenn sehr viele Ionenkanäle beteiligt sind, dennoch führt diese Stochastizität der Membranströme zu einer intrinsischen Variabilität der Spikeantworten. Das wird deutlich, wenn man im Experiment mehrere Male einen identischen externen Stimulus präsentiert: die Antworten eines sensorischen Neurons unterscheiden sich von Mal zu Mal, oft sogar erheblich (siehe Abb. 2 unten).

Diese durch intrinsisches Rauschen verursachte „trial-to-trial“-Variabilität stellt ein ganz grundsätzliches Problem für neuronale Informationsverarbeitung dar: Wie soll ein Gehirn entscheiden, ob es sich um den gleichen Gegenstand, oder ob es sich um einen anderen aber ähnlichen Gegenstand in der Umwelt gehandelt hatte, wenn die Spiketrains der Sensoren schon aus stochastischen Ursachen recht unterschiedlich sein können? Diese Frage wird besonders relevant, wenn es um eher subtile Unterschiede in den Signalen der Außenwelt geht (siehe Abschnitt sexuelle Selektion).

### **Kann Mittelung eine Lösung für das Problem unzuverlässiger Spiketrains sein?**

Die Statistik lehrt, dass bei zufälligen Messfehlern eine Mittelwertbildung hilft, den „wahren“ Messwert genauer zu bestimmen. Ein probates Rezept schiene demnach, die Spikerate eines Neurons über längere Zeit zu mitteln. Die zeitliche Mittelung hat aber zwei wesentliche Nachteile: Wenn man die relevante Information erst



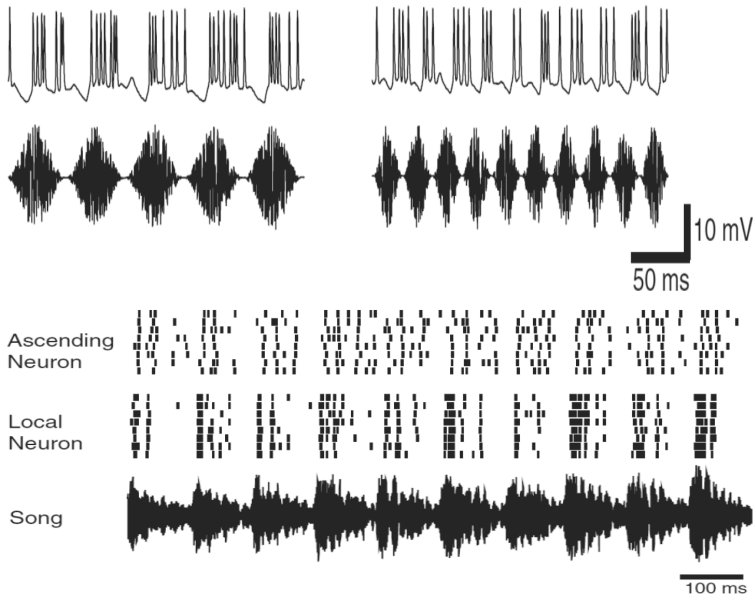


Abbildung 2.

Oben: Antwort eines auditorischen Neurons der Heuschrecke auf amplitudenmodulierte Schallreize (Modulationsfrequenzen 20 und 40 Hz). Obere Spur: Ableitung mit Aktionspotentialen (auch sichtbar: synaptische Potentiale, auf denen die Aktionspotentiale „aufsitzen“), darunter der jeweilige Reiz. Die mittleren Spikeraten waren 144 Hz bzw. 132 Hz. Unten: Variabilität der Antwort zweier Nervenzellen, bei wiederholter Präsentation eines identischen akustischen Signals (jeweils 8 Wiederholungen). Unten: Amplitudenverlauf des Gesangs eines Heuschreckenmännchens. Darüber dargestellt Spiketrains von zwei auditorischen Neuronen der Heuschrecke, einem lokalen und einem aufsteigenden Neuron (vergl. Abb. 1). Jeder kleine senkrechte Strich gibt den Zeitpunkt eines Aktionspotentials wieder. Eine Zeile entspricht der Reaktion auf eine Reizdarbietung. Man beachte die hohe „trial-to-trial“ Variabilität, vor allem beim aufsteigenden Neuron (ascending neuron). Aus Ronacher et al. (2008).

nach längerer Mittelung erhält, kann man vielleicht nicht schnell genug reagieren. Außerdem würde man, bei schnellen Änderungen, wie sie besonders in der Akustik relevant sind, auch die Information über die schnellen Amplitudenänderungen wegmitteln. Das wird in Abb. 2 oben deutlich. Hier ist die Reaktion eines Neurons auf zwei Rauschreize mit sinusförmiger Amplitudenmodulation (20 und 40 Hz) dargestellt. Die mittleren Spikeraten für die 20 Hz und 40 Hz Amplitudenmodulationen unterschieden sich nur geringfügig (144 vs. 132 Spikes/s), die Eigenschaften der zwei akustischen Signale werden erst in den unterschiedlichen Spikemustern sichtbar. Die Information über den Sinnesreiz ist in diesem Beispiel offensichtlich im Zeitmuster der Aktionspotentiale enthalten, nicht in der mittleren Spikerate.

Eine bessere Alternative scheint da eine räumliche Mittelung über viele parallele Kanäle zu sein (für die das intrinsische Rauschen jeweils unabhängig ist). Diese Lösung für das Variabilitätsproblem ist sicher an vielen Stellen in Nervensystemen verwirklicht. Allerdings gibt es auch bei diesem Verfahren Nachteile, die sehr deutlich beim Sehsystem von Wirbeltieren zutage treten. Unsere Augen enthalten zwei Rezeptorsysteme, die für verschiedene Helligkeitsbereiche zuständig sind. Das Zapfensystem für hohe Lichtintensitäten, die Stäbchen für das Sehen bei sehr geringen Intensitäten, wo es darauf ankommt, jedes Lichtquant zu nutzen. Viele Stäbchen sind auf eine retinale Ganglienzelle zusammengeschaltet, die für den Informationstransport ins Gehirn zuständig ist. Dadurch geschieht eine Form der räumlichen Mittelung, die einen genaueren Helligkeitswert zu bestimmen erlaubt. Gleichzeitig ist damit aber der exakte Ort, wo die Quanten eintrafen, nicht mehr rekonstruierbar. Daher nimmt die Sehschärfe notwendig ab – bei Dämmerlicht eine Erfahrung, die jedem geläufig ist. Eine Alternative wäre es, Lichtquanten über längere Zeit zu sammeln, also die Photorezeptoren mit längerer Integrationszeit auszustatten. Manche Kröten haben diesen Kompromiss gewählt. Bei sehr niedrigen Lichtintensitäten haben die Stäbchen in ihrer Retina Integrationszeiten von mehr als einer Sekunde. Kröten können damit noch bei deutlich geringeren Lichtintensitäten sehen als wir Menschen, dann allerdings keine schnellen Bewegungen erfassen. Eine schnell laufende Schabe wäre für die Kröte dann unsichtbar, aber eine langsame Schnecke oder ein Regenwurm kann trotz geringster Lichtintensität wahrgenommen und geschnappt werden.

Besonders für Hörsysteme, die sich rasch ändernde Signale erfassen müssen (Abb. 2), dürfte eine zeitliche Mittelung problematisch sein. Die Alternative, sich viele parallele Nervenzellen mit ähnlichen Aufgaben zu leisten, über deren Reaktionen dann gemittelt wird, steht sicher nicht allen Tieren zur Verfügung, und insbesondere nicht kleinen Tieren mit begrenzten neuronalen Ressourcen. Da gibt es möglicherweise Platzprobleme und vor allem Probleme mit der Energieversorgung. Der „Betrieb“ von Nervenzellen ist energetisch sehr kostspielig, da die Ionenflüsse beim Aktionspotential durch Pumpen, die unter Energieverbrauch gegen ein Konzentrationsgefälle arbeiten, wieder kompensiert werden müssen (siehe Anhang). Mehr zum Thema Energiebedarf von Nervensystemen findet sich im Kapitel über Populationscodes und sparsame Codierung.

Zurück zu Abbildung 2 unten, in der die intrinsische Variabilität von Spiketrains bei wiederholter Reizdarbietung zum Ausdruck kommt. Man hat den klaren Eindruck, dass die oben dargestellte Zelle wesentlich variabler reagiert als das darunter dargestellte lokale Neuron. Um jedoch abschätzen zu können, wie kritisch solche Variabilität für die weiteren Verarbeitungsprozesse im ZNS ist, benötigt man präzisere Angaben. Wie kann man die Abweichungen zwischen zwei Spiketrains quantitativ erfassen? Hier helfen sogenannte Spiketrain-Metriken weiter, die ab Mitte der 1990er Jahre eingeführt wurden. Sie erlauben es, die Verschiedenheit zweier Spiketrains durch einen Zahlenwert zu beschreiben, der ein völlig intuitives Abstandsmaß darstellt: kleine Werte zeigen hohe Ähnlichkeit, große Werte starke Verschiedenheit von zwei Spikefolgen an (siehe Abb. 3).

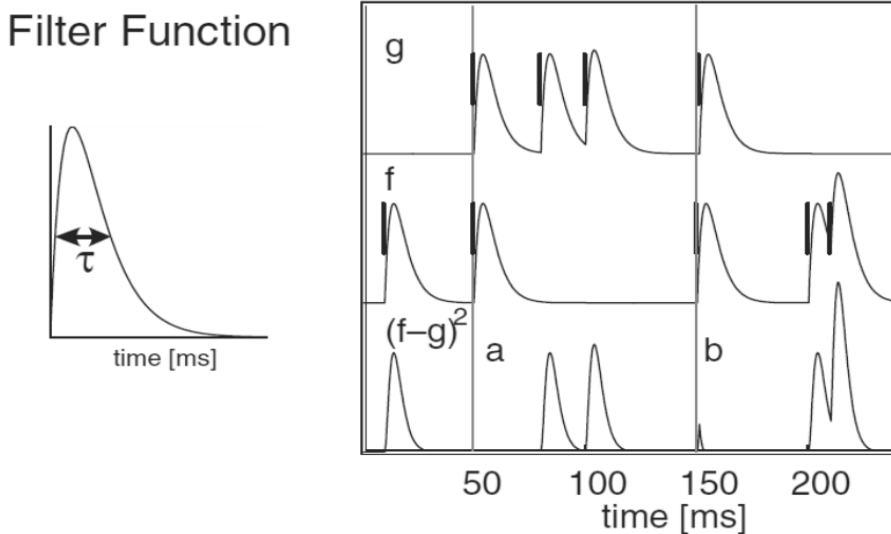


Abbildung 3. Prinzip der Spiketrain-Metrik nach van Rossum (2001)

Dargestellt sind zwei Spiketrain-Folgen (oberste und mittlere Spur), jeder Spike ist durch eine senkrechte Linie symbolisiert. Als erster Schritt wird jeder Spike durch eine Kurvenfunktion (links) ersetzt. Der entscheidende Schritt besteht darin, dass die resultierenden Funktionen  $f$  und  $g$  voneinander subtrahiert werden (das Resultat wird außerdem quadriert, um negative Werte zu vermeiden), und dann wird das Integral über die Funktion  $(f - g)^2$  gebildet. An den Stellen  $a$  und  $b$  wird der Clou des Verfahrens deutlich. Wenn die Spikes in den beiden Folgen gleichzeitig ( $a$ ) oder fast gleichzeitig ( $b$ ) auftreten, dann hebt sich durch die Subtraktion ihr Beitrag zu dem Integral auf. Wenn es für einen Spike keinen „Gegenpart“ im anderen Spiketrain gibt, dann erhöht sich das Verschiedenheitsmaß. Entsprechend ergibt diese Metrik ein kleines Resultat, wenn in zwei Spiketrains viele Spikes zur gleichen Zeit (koinzident) vorkommen – was genau unserem Verständnis von großer Ähnlichkeit entspricht. Näheres in Ronacher und Stange (2013) und Ronacher (2014). (Aus Ronacher et al. 2008).

Mit dieser Metrik kann man die Verschiedenheit von Spiketrains quantitativ erfassen und nun fragen, wie stark sich die Spiketrains allein aufgrund des intrinsischen Rauschens unterscheiden, oder wie verschieden sie sein müssen, damit ein Gehirn folgern kann, sie entsprechen zwei verschiedenen Objekten der Außenwelt.

### Ist „intrinsisches Rauschen“ kritisch für neuronale Informationsverarbeitung?

Auch wenn die neuronale Variabilität in Abb. 2 unten beträchtlich erscheint, so ist dennoch unklar, wie bedeutsam sie ist, wenn man sie mit den in der Außenwelt einwirkenden Störungen vergleicht. Gerade in der Akustik tritt ja häufig noch ein

zweites, „externes“ Rauschproblem auf: in der Regel werden akustische Signale auf dem Weg vom Sender zu einem Empfänger verändert, z. B. durch Reflexionen und Streuung an Gegenständen, durch Hintergrundgeräusche wie Blätterrascheln, oder auch durch Überlagerung mit störenden Signalen anderer Sender – denken Sie etwa an den Lärmpegel bei Unterhaltungen in den Kaffeepausen von Tagungen. Im Allgemeinen kommen akustische Signale also bereits in abgeschwächter und verzerrter Form beim Empfänger an, der damit vor ein zusätzliches Problem bei der Informationsverarbeitung gestellt wird. Dies trifft in besonderem Maß auf die Kommunikationssignale von kleinen Tieren zu, die in der Natur von vielfältigen Geräuschen überlagert werden – wer je das Glück hatte, einen tropischen Regenwald zu erleben, kann das bestätigen.

Man könnte also vermuten, dass diese äußeren Signalverzerrungen so gravierend sind, dass das intrinsische neuronale Rauschen für die Tiere einen vernachlässigbaren Effekt darstellen könnte? Diese Annahme trifft jedoch nicht zu: das intrinsische neuronale Rauschen ist sehr relevant für die Interpretation der Umwelt durch Gehirne. Wir konnten das für unsere Heuschrecken zeigen, indem wir ihnen Gesänge vorspielten, die stufenweise immer stärker verrauscht wurden, und dabei die Aktionspotentiale von identifizierten auditorischen Neuronen ableiteten (Neuhof et al. 2011). Mit Hilfe der Metrik konnten wir berechnen, welchen Anteil das intrinsische neuronale Rauschen an der Gesamt-Verschiedenheit von Spiketraits hatte und wieviel die externen akustischen Signalverzerrungen beitrugen. Selbst bei der maximalen Verzerrung der Gesänge trug das externe Rauschen nur einen etwa gleich großen Anteil zur Verschiedenheit der Spiketraits bei wie die intrinsische „trial-to-trial“-Variabilität. Dieses unerwartete Ergebnis zeigte, dass das intrinsische neuronale Rauschen für die Verarbeitung von Sinnesinformation und damit für die Interpretation der Umwelt sicher nicht ignoriert werden kann, weder vom ZNS des Tieres noch vom Untersucher. Die oben gestellte Frage – ist intrinsisches Rauschen kritisch? – ist also mit einem klaren Ja zu beantworten (Ronacher 2014).

### 3 Unterscheidung ähnlicher Signale im Kontext sexueller Selektion

Nun soll das in der Einleitung erwähnte Problem der Unterscheidbarkeit ähnlicher Signale erneut aufgegriffen werden. Das Erkennen von Qualitätsmerkmalen oder etwaiger Defizite von potentiellen Paarungspartnern erfordert eine sehr genaue Unterscheidung von im Prinzip ähnlichen Signalen. Sicherlich ist die Aufgabe, aus mehreren arteigenen Männchen den qualitativsten anhand seines Gesangs auszuwählen, um Größenordnungen schwieriger, als ein artfremdes Männchen abzulehnen, da die Gesänge verschiedener Arten sich in ihren Zeitmustern stark unterscheiden. Es geht nun also um die Unterscheidbarkeit von ähnlichen akustischen Signalen und damit direkt um das oben angeschnittene Problem: Wie gut kann das ZNS entscheiden, ob es sich bei zweimaligem Hören um das gleiche Objekt, den Gesang des Männchens A, handelte, oder um ein anderes Objekt, den Gesang

eines Männchens B, wenn die Spiketrains der Sensoren bereits aus stochastischen Ursachen so unterschiedlich sein können?

Um dies zu untersuchen, haben wir einer Heuschrecke die Gesänge von acht verschiedenen Männchen der gleichen Art vorgespielt und dabei die Spiketrains der Rezeptoren und nachgeschalteten Neurone registriert. Anschließend wurde mit Hilfe des Metrik-Verfahrens die Ähnlichkeit bzw. Unähnlichkeit der Spiketrains bestimmt und untersucht, wie gut sich anhand dieser Ähnlichkeiten die vorgespielten Gesänge unterscheiden und zuordnen lassen. Dieses Verfahren haben wir auf die ersten drei Verarbeitungsebenen der Hörbahn (s. Abb. 1) angewandt (Machens et al. 2003; Wohlgemuth und Ronacher 2007; Wohlgemuth 2008).

Wie in Abb. 1 dargestellt, bilden die aufsteigenden Neurone ein Nadelöhr für den Informationstransport ins Gehirn, wo die Entscheidung darüber fällt, ob ein Gesang akzeptiert wird oder nicht. Man könnte daher vermuten, dass die Spiketrains der aufsteigenden Neuronen besonders präzise Information tragen müssten. Dagegen spricht jedoch die Aussage von Abb. 2 unten, wo die „trial-to-trial“-Variabilität des aufsteigenden Neurons besonders hoch ist. Die dort dargestellte Reaktion ist kein Einzelfall, sondern bestätigte sich ganz generell: die Spiketrains der aufsteigenden Neurone sind in ihren Zeitmustern besonders variabel: von den Rezeptoren, über die lokalen Neurone, hin zu den aufsteigenden Neuronen, nimmt die intrinsische Variabilität der Spikeantworten deutlich zu. Dementsprechend nimmt die Klassifikationsgüte ab, mit der ähnliche Gesänge anhand der Spiketrains unterschieden werden können: von erstaunlichen 80-90% korrekter Zuordnung bei Rezeptoren und einigen lokalen Neuronen, auf etwa 40%-korrekt bei den aufsteigenden Neuronen (Machens et al. 2003; Wohlgemuth 2008).

Dieser Befund scheint in klarem Widerspruch zur beobachteten Leistung des Hörsystems zu stehen: aus eher subtilen Unterschieden zwischen den Gesängen verschiedener Männchen der gleichen Art Informationen über mögliche Qualitätsunterschiede eben dieser Männchen zu gewinnen. Die hohe Variabilität passt auch nicht zu Verhaltenstests, in denen eine hohe Präzision der Tiere bei der Unterscheidung von akustischen Stimuli zutage tritt (Ronacher und Stumpner 1988).

## 4 Änderung des Codierungsprinzips: hin zu einem Populationscode und „sparsamer Codierung“

Die wahrscheinliche Auflösung dieses scheinbaren Widerspruchs ist vor kurzem publiziert worden: offenbar ändert sich in der Hörbahn von Heuschrecken schon sehr früh das Codierungsprinzip (Clemens et al. 2011). Bereits zwischen der zweiten und dritten Verarbeitungsstufe findet ein Übergang zu einem Populationscode statt. Die etwa 60-80 Rezeptoren stellen eine homogen reagierende Gruppe von Nicht-Spezialisten dar, sie sind sozusagen Allrounder. Von verschiedenen Rezeptoren wird die im Prinzip gleiche Information übertragen, sie bilden das Zeitmuster des akustischen Reizes, die Amplitudenmodulation, sehr präzise in ihren Spikemus-

tern ab (s. Abb. 2 oben). Ähnliches gilt auch für die zweite Verarbeitungsstufe, die lokalen Neurone. Dagegen ändert sich das Codierungsprinzip auf der nächsten Verarbeitungsebene, bei den aufsteigenden Neuronen zu einem Populationscode mit verteilten Rollen. Auf dieser Verarbeitungsstufe findet bereits eine Spezialisierung statt. Verschiedene Neuronen-Typen – die wir als Individuen unterscheiden können – reagieren recht unterschiedlich. Einzelne Neurone sind auf bestimmte Merkmale akustischer Reize spezialisiert, z. B. auf das Vorhandensein von Pausen einer gewissen Mindestdauer, oder auf schnelle Amplitudenanstiege, etc. Ein einzelnes Neuron reagiert also besonders stark oder ausschließlich dann, wenn „sein“ spezifisches Merkmal im akustischen Reiz präsent war. So ist es nicht weiter verwunderlich, dass sich unsere Klassifikation der Gesangsmuster verschlechtert, wenn man die Reaktionen eines einzelnen aufsteigenden Neurons für sich alleine betrachtet: die für die Unterscheidung ähnlicher Gesänge notwendige Information verteilt sich eben auf verschiedene Kanäle, ein einzelner Kanal repräsentiert mit seiner Aktivität nur Teilinformationen (Clemens et al. 2011). Bei den Meldungen der aufsteigenden Neurone geht es auch nicht mehr um die genauen Zeitpunkte der Spikes, sondern um die Stärke der Antwort; nun signalisiert die Spikerate, ob ein bestimmtes Merkmal im Gesang vorhanden war oder nicht. Ein derartiger Raten-code hat den Vorteil, dass er von nachgeschalteten Neuronen leicht „ausgelesen“ werden kann.

## Ähnliche Verarbeitungsprinzipien wie bei Wirbeltieren

Das gleiche Prinzip eines Populationscodes ist auch in den viel größeren Nervensystemen von Wirbeltieren auf vielfältige Weise verwirklicht, dort allerdings meist nicht in der Peripherie, sondern erst auf wesentlich höheren Verarbeitungsstufen. Dass die aufsteigenden Neurone der Heuschrecken ihre Codierungsaufgaben auch mit deutlich geringeren Spikeraten erfüllen als die peripheren Neuronen, also eine sogenannte „sparsame“ Codierung („sparse coding“) stattfindet, ist eine weitere Parallele zu Wirbeltier-Nervensystemen. Bei auditorischen Neuronen im Gehirn von Grillen hat man sogar eine um 90% geringere Spikerate gefunden als bei den aufsteigenden Neuronen (Kostarakos und Hedwig 2015). Wenn die Information mit Hilfe von wenigen Spikes repräsentiert werden kann, senkt das die „Betriebskosten“ des Zentralnervensystems – offenbar ein wichtiger Selektionsdruck (Review dazu in Niven und Laughlin 2008).

## 5 Ein einfaches Modell zur Verarbeitung

Lag in der bisherigen Betrachtung der Schwerpunkt auf den Antworten von einzelnen oder einer Population von Neuronen, in Art eines bottom-up Zugangs, so soll nun die Blickrichtung umgekehrt werden: wir gehen nun – top down – von der Leistung des Gesamtsystems aus – d. h. von den Verhaltensantworten der Tiere, die sie als Reaktion auf eine Vielzahl akustischer Stimuli zeigen – und versuchen, aus diesen Antworten Vorstellungen über die zugrundeliegenden neuronalen Ver-

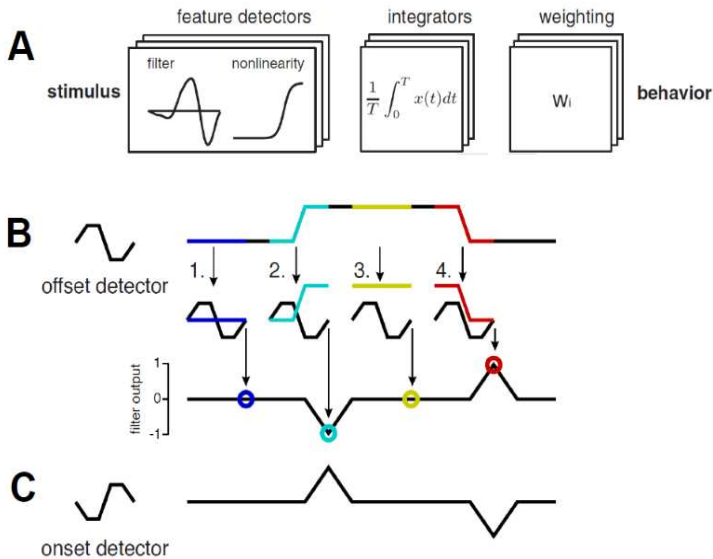


Abbildung 4. Grundstruktur des LN-Modells zur Beschreibung der Verhaltensantworten (A) und Wirkungsweise von zwei linearen Filtern (B, C). Nähere Beschreibung im Text. Modifiziert aus Clemens und Ronacher (2013) und Ronacher et al. (2015)

arbeitungsschritte abzuleiten. Dazu hat Jan Clemens, als Doktorand in meiner Arbeitsgruppe, ein bemerkenswert einfaches Modell entwickelt, welches drei Verarbeitungsschritte postuliert (Abb. 4A):

- Eine Filterbank aus parallel angeordneten Merkmalsdetektoren, welche bestimmte Stimulus-Merkmale mittels LN-Filtern extrahieren; jeder Detektor besteht aus einem Linearen Filter – also einer Art „Vorbild“, das mit dem aktuellen akustischen Stimulus verglichen wird und auf das Vorhandensein spezifischer Merkmale im Stimulus reagiert – gefolgt von einer Nichtlinearität (daher der Name LN-Modell).
- Der nächste Verarbeitungsschritt besteht in einer zeitlichen Integration, also in der Bildung eines Mittelwerts über längere Zeit für den Ausgang jedes dieser Merkmalsdetektoren. Dies ist eigentlich der erstaunlichste Teil dieses Modells, da er dem, was oben über zeitliche Mittelung gesagt wurde, zu widersprechen scheint (mehr dazu unten).
- Schließlich werden die Ausgangswerte der parallel arbeitenden Merkmalsdetektoren mit bestimmten Gewichtungen linear zu einer Vorhersage für die Verhaltensantwort kombiniert (Clemens und Ronacher 2013; Ronacher et al. 2015).

Die Wirkungsweise eines linearen Filters ist schematisch in Abb. 4 B dargestellt. Der Filter wird von links nach rechts über den Stimulus bewegt. Zu jedem Zeitpunkt wird der Filter mit dem aktuellen Stimulus-Verlauf (farbig markiert) multipliziert, das punktweise Produkt wird summiert und ergibt den Ausgangswert des Filters (farbige Kreise in der mittleren Spur). Der Ausgangswert gibt die Ähnlichkeit zwischen dem Filter und dem aktuellen Amplitudenverlauf des Stimulus wieder. Eine Filterkombination mit positiver und symmetrischer negativer Ausbuchtung, wie in Abb. 4B dargestellt, reagiert besonders stark auf eine rasche Verringerung der Amplitude (als „offset-Detektor“). Bleibt die Stimulus-Amplitude konstant (an den Stellen 1., 3.) so heben sich die positiven und negativen Ergebnisse der Multiplikation in der Summe auf und der Ausgangswert ergibt Null. Man beachte, dass dabei nicht zwischen zwei verschiedenen hohen Amplitudenwerten unterschieden wird (vergeiche die Filterwerte an den Stellen 1 und 3). Dieser Detektor reagiert also nur auf Änderungen, nicht auf Gleichbleibendes! Ein Amplitudenanstieg (2.) führt zu einem negativen Wert, weil die negative Flanke des Filters mit dem höheren Amplitudenwert multipliziert wird, der negative Anteil in der Summe also überwiegt. Ein Amplitudenabfall führt dagegen zu einem positiven Ausgangswert der Filterung (Stelle 4, daher die Bezeichnung offset-Detektor). Werden die nach unten und oben gerichteten Ausbuchtungen des Filters vertauscht, so erhält man einen „onset“-Detektor, der spezifisch auf eine schnelle Amplitudenerhöhung reagiert (Abb. 4 C).

Die hier dargestellte Auswertung repräsentiert ein ganz allgemeines, häufig zu findendes Prinzip bei der neuronalen Verarbeitung von Sinnesreizen: nur Änderungen sind interessant, Gleichbleibendes wird weitgehend ignoriert!

In Abbildung 5 ist das Wirkungsprinzip des Modells an einem Beispiel-Stimulus dargestellt. In A sind die Zeitverläufe der beiden optimalen Filter farbig dargestellt. In B sieht man oben einen Ausschnitt aus einem akustischen Stimulus, der zuerst einen Amplitudenabfall ausweist, nach kurzer Pause einen Amplitudenanstieg und danach wieder einen kleineren Amplitudenabfall. Solche Amplitudenmodulationen wurden als Testreize verwendet, quasi als Nachbildung eines Heuschreckengesangs, wie er in Abb. 2B zu sehen ist. Darunter ist in B die Antwort der beiden Filter auf diesen Stimulus zu sehen. Beide reagierten, etwas zeitversetzt, mit positiver, negativer und wieder positiver Antwort auf diesen Stimulus. In C ist die Wirkung der Nichtlinearität zu sehen. Die Nichtlinearität weist einen sehr steilen Anstieg auf und wirkt daher wie ein Schwellenwert. Nur Werte über der Schwelle werden „durchgelassen“, alle niedrigeren Filterausgänge führen zum Wert Null. Die Funktion der dem Filter nachgeschalteten Nichtlinearität könnte man physiologisch so interpretieren, dass eine durch synaptische Potentiale verursachte Depolarisation der Nervenzellmembran dadurch in eine Spikerate übersetzt wird. In D sind die Meldungen der beiden Detektoren zu sehen, zunächst zeitlich aufgelöst, anschließend nach der zeitlichen Integration als zwei Zahlenwerte (0.26 und 0.17). In F ist das Endresultat zu sehen, als gewichtete Summe der beiden Detektorausgänge.

Überraschenderweise lieferte eine sehr einfache Version dieses Modells, mit nur zwei Merkmalsdetektoren, bereits eine ausgezeichnete Vorhersage für die Verhaltensantworten der Heuschreckenweibchen (Abb. 5 F; die Korrelation zwischen Vor-



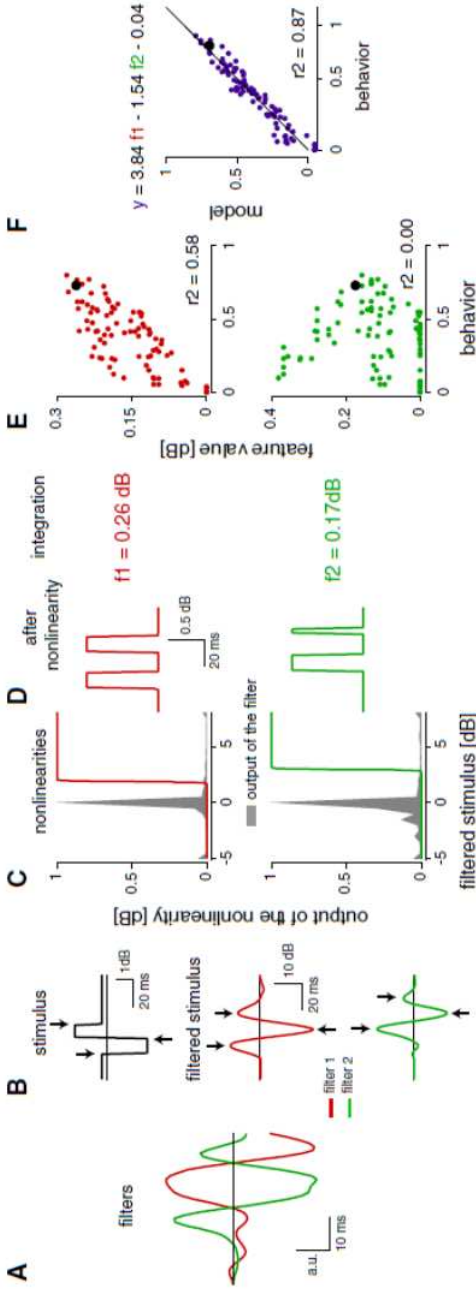


Abbildung 5. Das LN-Modell von Abb. 4 mit nur zwei parallelen Detektor-Systemen. In A sind die beiden Filterfunktionen, die von dem Lernalgorithmus gefunden wurden, dargestellt. In B ist die Antwort der beiden Filter auf einen Reiz mit Amplitudenreduktion, gefolgt von einem starken Amplitudenanstieg zu sehen (obere Spur). C zeigt die zugehörigen Nichtlinearitäten, die nur positive Werte oberhalb einer Schwelle durchlassen und praktisch binäre Antworten (D) erzeugen. Diese Antworten werden dann über die Zeit integriert, so dass jedes Detektorsystem für diesen Stimulus einen Zahlenwert liefert (hier 0.26 und 0.17). Diese beiden Werte werden dann zu einer Vorhersage für die Verhaltensantwort kombiniert, und zwar so, dass f1 mit dem Gewicht 3.84 und f2 mit negativem Gewicht (-1.54) in die Kombination eingeht. In F sind die Vorhersagen des Modells für eine große Zahl verschiedener Schallmuster mit den gemessenen Verhaltensantworten verglichen. Das Modell mit nur zwei Detektoren gibt eine ausgezeichnete Vorhersage für die gemessenen Verhaltensantworten. In E ist dieser Vergleich für die beiden Detektoren getrennt durchgeführt (siehe Text). Aus Clemens und Ronacher (2013), Näheres dort).

hersage und Verhalten ist ausgezeichnet,  $r^2 = 0.87$ ; die Einziehung eines zusätzlichen Detektors verbessert die Vorhersagequalität nur marginal). Dabei muss betont werden, dass abgesehen von der in Abb. 4A dargestellten Grundstruktur des Modells keine weiteren Vorwegannahmen gemacht wurden. Alle spezifischen Eigenschaften der Filter, die Form der Nichtlinearitäten und die Gewichtungen für die Linearkombination wurden durch einen genetischen Lernalgorithmus gefunden, der an einer großen Zahl von verschiedenen Gesangstypen und den Reaktionen der Weibchen auf diese Gesänge trainiert wurde, um die optimalen Filter, Nichtlinearitäten und Gewichtungsfaktoren zu finden (es wurde nur jeweils ein Teil der Daten für das Training verwendet, der Vorhersageerfolg wurde dann an den restlichen Daten validiert).

Auf Details des Modells soll hier nicht weiter eingegangen werden, interessierte Leser seien auf Clemens und Ronacher (2013), Hennig et al. (2014), Ronacher et al. (2015) verwiesen. Es sollen nur einige erstaunliche Eigenschaften und Vorhersagen des LN-Modells hervorgehoben werden.

- (i) Der zeitliche Integrationsschritt des Modells scheint ja insofern problematisch, dass die Zeitstruktur des akustischen Reizes dadurch weggemittelt und verloren gehen würde. Das ist jedoch nicht der Fall: Die linearen Filter arbeiten vor der zeitlichen Mittelung und reagieren da hoch spezifisch auf bestimmte Stimulus-Merkmale, z. B. auf bestimmte Amplitudenmodulations-Muster in dem Heuschreckengesang. Es kommt also sehr wohl darauf an, dass bestimmte Merkmale im Gesang enthalten sind, nur ist nach dem Modell der genaue Zeitpunkt, wann genau im Gesang diese Merkmale auftreten, später nicht mehr wichtig. Zu dieser Modelleigenschaft passen Ergebnisse von früheren Verhaltenstests, in denen die Tiere gut auf Gesangsattrappen reagierten, deren Zeitmuster verändert waren. Ja, die Heuschreckenweibchen reagierten sogar unerwartet gut auf einen künstlichen Gesang in dem zwei ansonsten völlig unwirksame Muster kombiniert waren (von Helversen und von Helversen 1998). Solche Ergebnisse waren lange Zeit etwas rätselhaft, können mit dem neuen Modell jedoch zwanglos erklärt werden. Die gute Beantwortung solcher „shuffled songs“ im Verhaltenstest zeigt, dass die Tiere zwar das Vorhandensein bestimmter Merkmale bewerten, aber nicht so sehr, wann genau im Gesang diese Merkmale auftreten.

Der zeitliche Integrationsschritt des Modells gibt außerdem eine plausible Erklärung für die große Variabilität der Spikeantworten, die man speziell bei den aufsteigenden Neuronen beobachtet. Die beobachtete hohe Variabilität betrifft vor allem die genauen Zeitpunkte der Aktionspotentiale; dagegen variieren die über längere Zeit gemittelten Spikeraten viel weniger. Wenn aber – wie vom Modell vorhergesagt – für nachgeschaltete Neuronen im Gehirn der Tiere nur der Mittelwert der Spikerate über längere Zeit wichtig ist, so ist die Variation der Spike-Zeitpunkte eben nicht nachteilig und konnte offenbar in Kauf genommen werden.

- (ii) Dieses Modell mit LN-Filtern bietet einen großen Vorteil: anders als bei künstlichen Neuronennetzwerken (s. Beitrag von Lürer und Lass ab S. 89 in diesem Band) sind in diesem Modell die Zwischenschritte konkret fassbar und ermöglichen prüfbare Vorhersagen über Eigenschaften von realen Neuronenverbänden. Tatsächlich hat man im Gehirn von Grillen Neuronen identifizieren können, deren Antworten den nach dem LN-Modell postulierten Vorhersagen sehr gut entsprechen (Kostarakos und Hedwig 2012; Clemens und Hennig 2013; Hennig et al. 2014).
- (iii) Die Vorhersagekraft des Modells soll an einem weiteren Beispiel deutlich werden: Die Kombination von zwei Detektorsystemen ermöglichte eine ausgezeichnete Beschreibung der Verhaltensantworten (Abb. 5F). Vergleicht man jedoch die Ausgänge jedes Merkmal-Detektors separat mit den Verhaltensdaten (Abb. 5E), so ergibt sich für den ersten Detektor immer noch eine recht ordentliche Vorhersage ( $r^2 = 0.58$ ); die Meldungen des zweiten Detektors korrelieren jedoch, für sich betrachtet, überhaupt nicht mit der Verhaltensantwort ( $r^2 = 0.00$ ) – obwohl dieser Detektor ganz wesentlich zur ausgezeichneten Leistung des Gesamtsystems beiträgt. Das gibt einem Elektrophysiologen, der mit Ableitungen von Einzelzellen arbeitet, ausreichend Stoff zum Grübeln. Es reicht offenbar nicht aus, Neurone gesondert und einzeln zu betrachten, man muss ihre Funktion im Kontext der Antworten mehrerer Neurone mit berücksichtigen. Sonst läuft man Gefahr, den Beitrag von wichtigen Neuronen nicht zu erkennen, wenn ihre Antworten, für sich genommen, keine klare Korrelation mit den getesteten Parametern aufzeigen (s. a. Ronacher 2014, Ronacher et al. 2015).
- (iv) Es gibt eine einfache Möglichkeit, die Filter, wie sie in Abb. 4 und 5a beschrieben wurden, physiologisch zu realisieren. Sie können als Kombination von erregenden und hemmenden Eingängen dargestellt werden (Clemens und Hennig 2013; Hennig et al. 2014; Ronacher et al. 2015). So liefert dieses Modell eine faszinierende Erkenntnis: bereits durch kleine Änderungen in den Zeitbeziehungen der erregenden und hemmenden Komponenten lassen sich unterschiedliche Verhaltenspräferenzen für bestimmte Gesangsmuster einer Vielzahl verschiedener Grillen- und Laubheuschreckenarten vorhersagen (Clemens und Hennig 2013, Hennig et al. 2014). So kann man sich vorstellen, dass bei der Evolution solcher Kommunikationssysteme bereits kleine Änderungen in synaptischen Gewichtungen zu geänderten Präferenzen für Gesangsmuster geführt haben, was die Aufspaltung in getrennte Arten begünstigt haben dürfte, die sich primär in ihren akustischen Paarungssignalen unterscheiden.

## 6 Ausblick: ähnliche Verarbeitungsprinzipien im ZNS der Heuschrecke wie bei Wirbeltieren

Wenn zwei Gegenstände oder Ereignisse der Umwelt unterschieden werden sollen, müssen die von ihnen hervorgerufenen neuronalen Signale ausreichend verschieden sein, um dem ZNS die Entscheidung: verschieden oder Abwandlung des gleichen Gegenstandes zu ermöglichen. Ein grundsätzliches Problem für die Verarbeitung von Umweltreizen im Nervensystem liegt daher in der intrinsischen Variabilität neuronaler Signale begründet. Dieses Problem tritt bei den Insekten besonders deutlich zutage, da die Größenbeschränkung ihrer Nervensysteme es nicht erlaubt, sehr viele Neuronen für die gleichen Aufgaben einzusetzen und dadurch die intrinsische Variabilität ihrer Spiketrains zu kompensieren. Insekten sind gezwungen, ihre neuronalen Ressourcen effizient zu verwenden. Die Effizienz ist unter anderem daran zu erkennen, dass komplexe Verarbeitungsschritte, an denen bei Wirbeltieren Tausende Neurone beteiligt sind, bei Insekten teilweise von einzelnen Neuronen geleistet werden. Eine häufig zu findende Strategie besteht darin, Schritte der Informationsfilterung schon sehr peripher, bereits in den Rezeptorzellen oder kurz danach durchzuführen. Dadurch wird die weitere Verarbeitung auf höheren Verarbeitungsstufen entlastet, allerdings geht die Filterung zu Lasten der Informationsvielfalt. Die Hörbahn von Insekten kann daher kein Vielzweck-Organ sein, um die ganze akustische Vielfalt der Umwelt aufzunehmen, sondern ist auf wenige klar definierte, verhaltensrelevante Aufgaben fokussiert: vor allem Signale von Artgenossen und Geräusche sich nähernder Fressfeinde. Diese Aufgaben werden allerdings, trotz der Variabilität der Spikeantworten, mit hoher Präzision vollbracht.

Abschließend soll noch einmal betont werden, dass man trotz der Größen- und Organisationsunterschiede ihres Nervensystems bei diesen Insekten ähnliche Verarbeitungs- und Codierungsprinzipien findet wie bei den viel größeren Nervensystemen von Wirbeltieren. Z. B. wird Gleichbleibendes weitgehend ignoriert, die Verarbeitung ist auf Änderungen in der Umwelt fokussiert. Auch Änderungen des Codierungsschemas hin zu einem Populationscode mit verteilten Rollen der einzelnen Elemente, und sparsame Codierung scheinen ganz allgemein verwirklichte Prinzipien zu sein. Eine vergleichende Betrachtung ist wohl der beste Weg, allgemeine Verarbeitungsprinzipien und grundlegende Algorithmen zu identifizieren, die im Laufe der Evolution von Nervensystemen entwickelt wurden. In der kompletten Vernachlässigung evolutiver Prozesse scheint mir im Übrigen ein weiterer wichtiger Grund für das im Beitrag von Lüer und Lass hervorgehobene Scheitern des strikten Behaviorismus zu liegen.

## Anhang: Zustandekommen von Ruhepotential und Aktionspotential einer Nervenzelle

Für eine ausführlichere Darstellung der neurobiologischen Grundlagen sei z. B. auf Reichert 2000 verwiesen.

In Abb. 6A ist schematisch die Verteilung der wichtigsten Ionensorten innerhalb und außerhalb der Zelle dargestellt. Die innen negativ geladene Zellmembran (Ruhepotential hier: -80 Millivolt, mV) ist durch die doppelt-gestrichelte senkrechte Linie angedeutet. (A-) symbolisiert negativ geladene Proteine im Zellinneren, die die Zelle wegen ihrer Größe nicht verlassen können. Natrium-Ionen ( $\text{Na}^+$ ) finden sich außerhalb der Zelle in hoher, innen in niedriger Konzentration, für Kalium-Ionen ist das Konzentrationsverhältnis umgekehrt; Chlorid-Ionen finden sich überwiegend außerhalb der Zelle. Die unterschiedliche Verteilung der positiv geladenen  $\text{Na}^+$ - und  $\text{K}^+$ -Ionen wird durch Ionenpumpen aufrechterhalten, die unter Energieaufwand  $\text{Na}^+$  aus der Zelle und  $\text{K}^+$  in die Zelle transportieren. Die Ionen können die Zellmembran nur durch spezifische Ionenkanäle passieren, von denen es eine große Anzahl verschiedener Typen mit unterschiedlichen Eigenschaften gibt. Für eine grundlegende Betrachtung des Zustandekommens von Aktionspotentialen sind vor allem die spannungsabhängigen Natrium- und Kalium-Kanäle relevant.

Kalium-Ionen tendieren dazu, nach ihrem Konzentrationsgefälle aus der Zelle zu wandern; da sie aber eine positive Ladung nach außen tragen, würde der Transport jedes Ions das Ladungsungleichgewicht vergrößern. Es können daher nur so viele positive Ladungen nach außen wandern, bis ein Gleichgewicht zwischen elektrischer Arbeit und der entgegengesetzt gerichteten Konzentrationsarbeit erreicht ist. Das Gleichgewichtspotential, bei dem kein Nettotransport von Ionen mehr stattfindet, kann für jede Ionensorte mit Hilfe der Nernstschen Gleichung berechnet werden, es ist für  $\text{K}^+$  und  $\text{Cl}^-$  normalerweise noch negativer als das Ruhepotential, während es für  $\text{Na}^+$  weit im positiven Bereich liegt, bei etwa + 70mV.

Man beachte, dass innerhalb und außerhalb der Zelle gleich viele positive und negative Ladungen vorkommen; die innen negative Ladung der Zellmembran wird durch eine verglichen mit den Gesamt-Ionenkonzentrationen extrem kleine Anzahl an Ionen an der Zellmembran getragen, die Zellmembran verhält sich wie ein geladener Kondensator mit geringer Kapazität.

Ein Aktionspotential wird ausgelöst durch eine gewisse Depolarisation der Membran, d. h. durch eine Verringerung des Membranpotentials bis zu einem Schwellenwert von im Beispiel etwa -65 mV (Abb. 6C). Wird dieser Schwellenwert überschritten, so öffnen sich spannungsabhängige  $\text{Na}^+$ -Kanäle, die beim Ruhepotential geschlossen sind, was zu einem „explosiven“ Einstrom von  $\text{Na}^+$  in die Zelle führt. Der schnelle Einstrom ist deshalb möglich, weil für  $\text{Na}^+$  sowohl das Konzentrationsgefälle als auch der elektrische Gradient in die gleiche Richtung (zell-einwärts) wirken. Der  $\text{Na}^+$ -Einstrom führt nun zu einer schnellen Umpolung des Membrankondensators, an der Spitze des Aktionspotentials ist jetzt das Zellinnere positiv geladen – siehe Abb. 6B für eine Momentaufnahme der Ionenverteilung an der Spitze des Aktionspotentials.

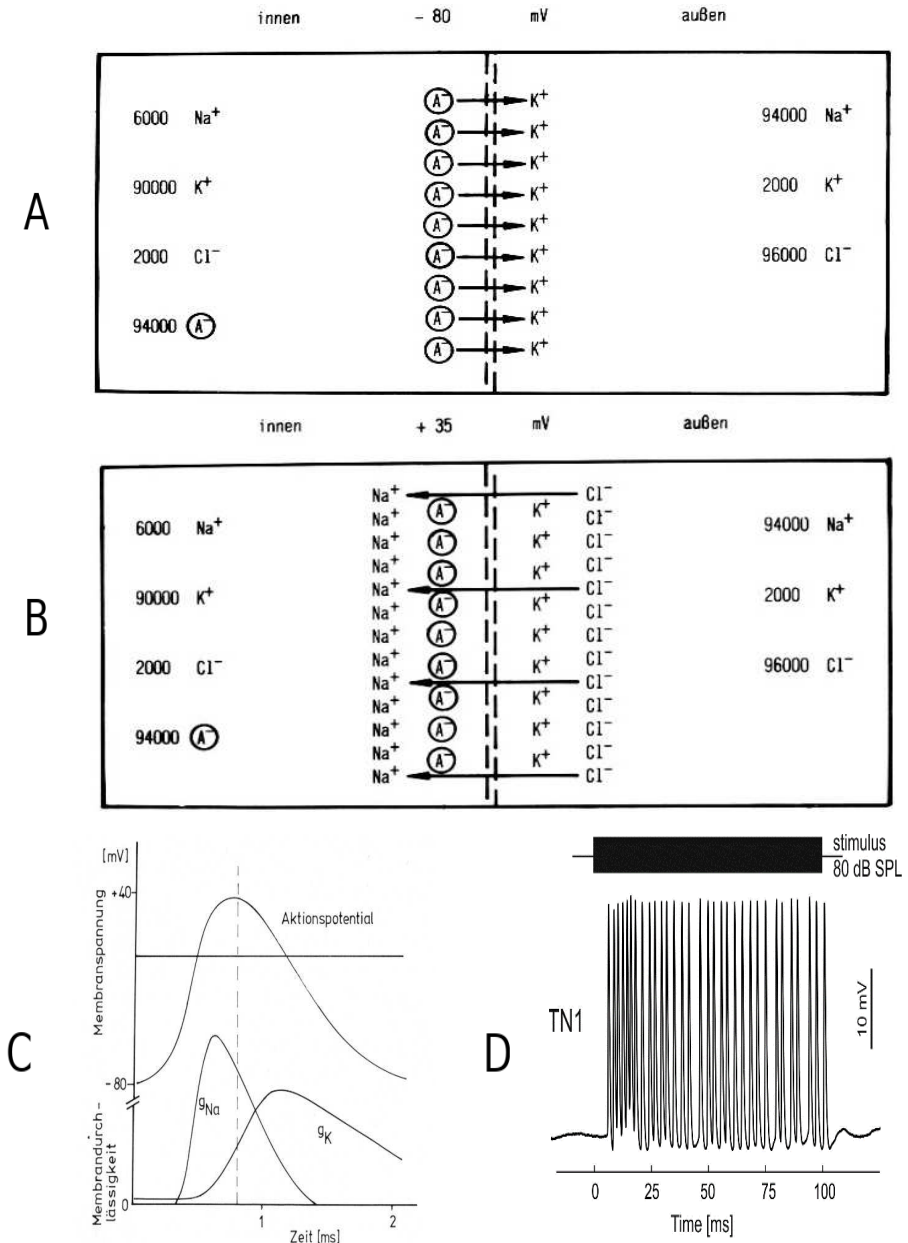


Abbildung 6. A) Ionenverteilung innerhalb und außerhalb einer Nervenzelle im Ruhezustand. B) Ionenverteilung an der Spitze eines Aktionspotentials. C: oben: zeitlicher Verlauf der Membranbranspannung während eines Aktionspotentials, darunter: die Membrandurchlässigkeiten für Natrium- und Kalium-Ionen. Erläuterungen im Text. D: Antwort eines lokalen auditorischen Neurons der Wanderheuschrecke auf einen lauten Schallreiz; Ableitung von Dr. Monika Eberhard (Greifswald). A-C aus Ronacher und Hemminger (1984)

In Abb. 6C ist oben der Verlauf des Membranpotentials beim Aktionspotential dargestellt, darunter die Membrandurchlässigkeit (Leitfähigkeit) für  $\text{Na}^+$  und  $\text{K}^+$ . Der geniale Trick beim Aktionspotential besteht darin, dass die Natriumkanäle (Kurve  $g_{\text{Na}}$ ) nach sehr kurzer Zeit automatisch wieder schließen und, etwas zeitverzögert, Kaliumkanäle ( $g_{\text{K}}$ ) öffnen. Da die Zelle jetzt innen positiv geladen ist, besteht nun für  $\text{K}^+$ -Ionen die gleiche Situation wie vorher für die Natriumionen: elektrischer Gradient und Konzentrationsgradient wirken nun für sie in die gleiche Richtung, d. h. nach außen, und  $\text{K}^+$ -Ionen verlassen die Zelle, wodurch die Zellinnenseite rasch wieder negativer wird und die  $\text{K}^+$ -Kanäle schließen. Binnen 1-2 ms ist das Ruhepotential wieder hergestellt – es müssen später nur noch die in die Zelle eingedrungenen  $\text{Na}^+$  und die aus der Zelle gewanderten  $\text{K}^+$  durch die  $\text{Na}^+/\text{K}^+$ -Pumpe wieder zurückgepumpt werden.

Diese Pumparbeit ist energetisch kostspielig. Deshalb ist es für die Energiebilanz von Neuronen wichtig, dass bei einem Aktionspotential nur relativ wenige Ionen die Ionenkanäle passieren müssen (Im Beispiel etwa ein Zehntausendstel, s. Abb. 6B). Vermutlich ist der Energieverbrauch der  $\text{Na}^+/\text{K}^+$ -Pumpen auch ein wichtiger Grund dafür, dass in zentralen Nervensystemen häufig ein Prinzip des ‚sparse codings‘ zu finden ist, d. h. dass die Neuronen mit sehr geringen Spikeraten arbeiten (vergl. Niven und Laughlin 2008).

In Abb. 6D ist eine Folge von Aktionspotentialen zu sehen, die in einem auditorischen Neuron einer Heuschrecke durch einen lauten Schallreiz ausgelöst wurde. Dieser Reiz wurde mit einer hohen Rate von 320 Aktionspotentialen pro Sekunde beantwortet.

## Fortleitung des APs:

In einer typischen Nervenzelle werden Aktionspotentiale gebildet, wenn synaptische Eingänge an einer bestimmten Stelle (der „spike-initiiierenden Zone“) eine ausreichende Depolarisation der Membran bewirken, d. h. wenn dort die Schwelle zur Öffnung der spannungsabhängigen  $\text{Na}^+$ -Kanäle überschritten wird. Wie wird das Aktionspotential nun über das ganze Axon, das in großen Tieren meterlang sein kann, fortgeleitet? Dazu ist ein Verstärkungsmechanismus notwendig, da sich ohne Verstärkung ein elektrisches Signal an einer Nervenzelle nur höchstens wenige Millimeter weit ausbreiten würde. Die Verstärkung passiert dadurch, dass ein AP an einer bestimmten Stelle des Axons (1 in Abb. 7 a) mit der positiv geladenen Innenseite einen Stromfluss zu benachbarten Stellen bewirkt, wodurch die Membran an Stelle 2 depolarisiert wird. Sobald dort die Schwelle von -65 mV überschritten wird, wird dort ebenfalls ein AP ausgelöst und so wandert ein AP durch ständige Neubildung über das ganze Axon (Abb. 7 b). An der Stelle, wo ein AP eben stattgefunden hat, sind die  $\text{Na}^+$ -Kanäle für kurze Zeit in einem nicht erregbaren, „refraktären“ Zustand. Dadurch wird gewährleistet, dass die Zelle nicht dauererregt wird und sich das AP nur in eine Richtung ausbreitet. Für weitere Informationen sei auf Lehrbücher verwiesen, z. B. Reichert 2008)



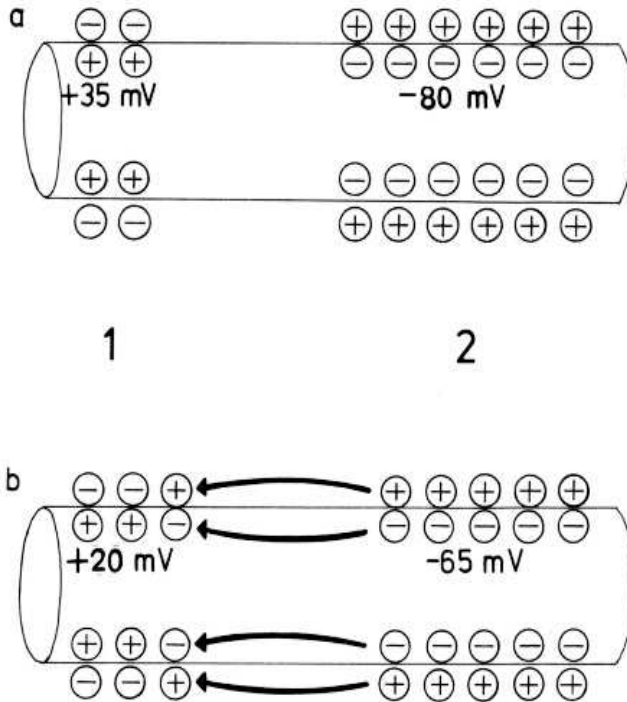


Abbildung 7. Vereinfachtes Schema zur Fortleitung des Aktionspotentials durch ständige Neubildung (aus Ronacher und Hemminger 1984)

## Literatur

Andersson, M. (1994): *Sexual selection*. Princeton University Press.

Clemens, J., Hennig, RM. (2013): Computational principles underlying the recognition of acoustic signals in insects. *J. Comput. Neurosci.* 35: 75–85.

Clemens, J., Kutzki, O., Ronacher, B., Schreiber, S., Wohlgemuth, S. (2011): Efficient transformation of an auditory population code in a small sensory system. *Proc. Natl. Acad. Sci. USA* 108: 13812–13817.

Clemens, J., Ronacher, B. (2013): Feature extraction and integration underlying perceptual decision making during courtship in grasshoppers. *J. Neurosci.* 33:12136–12145.

Elsner, N., Popov, A. V. (1978): Neuroethology of acoustic communication. *Adv. Insect. Physiol.* 13, 229–355.



Gerhardt, H. C., Huber, F. (2002). *Acoustic communication in insects and anurans*. Chicago University Press, Chicago IL.

Hennig, R.M., Heller K.-G., Clemens, J. (2014): Time and timing in the acoustic recognition system of crickets. *Frontiers in Physiology* 5, p. 286, doi: fphys.2014.00286.

Kostarakos, K., Hedwig, B. (2012): Calling song recognition in female crickets: temporal tuning of identified brain neurons matches behavior. *J. Neurosci.* 32(28): 9601–9612.

Kostarakos, K., Hedwig, B. (2015): Pattern recognition in field crickets: concepts and neural evidence. *J. Comp. Physiol. A* 201: 73–85.

Machens, C. K., Schütze, H., Franz, A., Stemmler, M. B., Ronacher, B., Herz, A. V. M. (2003): Auditory receptor neurons preserve characteristic differences between conspecific communication signals. *Nature Neurosci.* 6, 341–342.

Marr D. (1982): *Vision*. Freeman, New York.

Neuhofer, D., Stemmler, M., Ronacher, B. (2011): Neuronal precision and the limits for acoustic signal recognition in a small neuronal network. *J. Comp. Physiol. A* 197, 251–265.

Niven, J. E., Laughlin, S. B. (2008): Energy limitations as a selective pressure on the evolution of sensory systems. *J. Exp. Biol.* 211, 1792–1804.

Reichert, H. (2000); *Neurobiologie*. Thieme, Stuttgart.

Ronacher, B. (2014): Processing of species-specific signals in the auditory pathway of grasshoppers. In B. Hedwig (ed.): *Insect hearing and acoustic communication*. Springer, Berlin p. 185–204.

Ronacher, B., Hemminger, H. (1984): *Einführung in die Nerven- und Sinnesphysiologie*. 3. Auflage. Quelle und Meyer, Darmstadt.

Ronacher, B., Hennig, R.M., Clemens, J. (2015): Computational principles underlying recognition of acoustic signals in grasshoppers and crickets. *J. Comp. Physiol. A* 201: 61–71.

Ronacher, B., Stange N. (2013): Processing of acoustic signals in grasshoppers – a neuroethological approach towards female choice. *J Physiol (Paris)* 107: 41-50 (Review)

Ronacher, B., Stumpner, A. (1988): Filtering of behaviourally relevant temporal parameters of a grasshopper's song by an auditory interneuron. *J. Comp. Physiol. A* 163: 517–523.

Ronacher, B., Wohlgemuth, S., Vogel, A., Krahe, R. (2008): Discrimination of acoustic communication signals by grasshoppers: temporal resolution, temporal integration, and the impact of intrinsic noise. *J. Comp. Psychol.* 22, 252–263.

Stumpner, A., Ronacher, B. (1991): Auditory interneurons in the metathoracic ganglion of the grasshopper *Chorthippus biguttulus*: I. Morphological and physiological characterization. *J. Exp. Biol.* 158, 391–410.

van Rossum, M. C. W. (2001): A Novel Spike Distance. *Neural Comput.* 13, 751–763.

von Helversen, D., von Helversen, O. (1998): Acoustic pattern recognition in a grasshopper: processing in the time or frequency domain? *Biol. Cybern.* 79, 467–476.

Wohlgemuth, S. (2008): *Repräsentation und Unterscheidbarkeit amplitudenmodulierter akustischer Signale im Nervensystem von Feldheuschrecken*. Dissertation, Humboldt Universität zu Berlin.

Wohlgemuth, S., Ronacher, B. (2007): Auditory discrimination of amplitude modulations based on metric distances of spike trains. *J. Neurophysiol.* 97, 3082–3092.

# Informationstheoretische Metaphern im Vokabular der Molekularen Genetik

*Hans-Joachim Fritz*

*Institut für Mikrobiologie und Genetik der Universität Göttingen  
Derzeitige Adresse: Institut für Organische Chemie und Biomolekulare Chemie  
der Universität Göttingen, Tammannstraße 2, D-37077 Göttingen*

Wie reden wir über die Inhalte unserer Wissenschaft? Fange ich mit meiner Sprache wirklich *genau* das ein, was ich mitteilen will – und welche Konsequenzen hat es, wenn ich dieses Ziel verfehle? Typische Alltagsorgen eines Naturwissenschaftlers sehen anders aus – bis ein spezieller Anlass einen dann doch darüber nachdenken lässt. Für den Verfasser dieses Aufsatzes bestand ein solcher Anlass im Zusammenfinden von Vertretern verschiedener Disziplinen zu einer Arbeitsgruppe – mit dem selbst gestellten Auftrag, den Informationsbegriff aus verschiedenen Blickwinkeln zu betrachten. Das wurde schon in der Präambel dieses Bandes (ab S. 3) zum Ausdruck gebracht.

Erst einmal eine gemeinsame Sprache finden zu müssen, ist eine durchaus erwartete Startschwierigkeit eines solchen Unterfangens. Beim Nachdenken darüber jedoch zu bemerken, dass es mit der Kommunikation auch *innerhalb* der eigenen Disziplin nicht ausschließlich zum Besten und Klarsten steht, war schon überraschender und hat beim Verfasser neues Nachdenken ausgelöst – mit dem im Folgenden ausgebreiteten Ergebnis.

Für den engeren Zweck dieses Aufsatzes soll unter einer *Metapher* das Entleihen eines Worts aus einem allgemein gut verstandenen Sinnzusammenhang durch die Molekulare Genetik verstanden werden, mit dem Ziel, das Verständnis eines genetischen oder molekularbiologischen Sachverhalts zu fördern – und zwar durch Unterstellung von Verwandtschaft zu der Bedeutung, die das entlehene Wort in seinem Ursprungsumfeld besitzt. In diesem Sinn spricht man zum Beispiel in der Biochemie von einem metabolischen „Netzwerk“, um anzuzeigen, dass der Stoffwechsel der Zelle nicht aus einer einzigen, linearen Kette chemischer Reaktionen, sondern einer größeren Anzahl von Strängen und Zyklen besteht, zwischen die viele Querverbindungen eingezogen sind.

Mit der Angabe der Informationstheorie als Ursprung der zu betrachtenden Metaphern soll keine Untersuchung angekündigt werden, wie passgenau sich Begriffsinhalte der informationstheoretischen Fachterminologie auf molekularbiologische Sachverhalte projizieren lassen. Vielmehr geht es ganz praktisch darum, ob sich bei Gebrauch solcher Metaphern im Kopf des Adressaten auch tatsächlich die

gewünschte Erhellung einstellt – und da die meisten Menschen keine Informatio-  
nsthoretiker sind, ist im Normalfall damit zu rechnen, dass die hervorgerufenen  
Assoziationen im eher umgangssprachlichen Bedeutungsfeld des jeweiligen Worts  
landen.

Das aber liegt in aller Regel nahe am fachterminologischen Gebrauch – zum  
einen weil die Fachterminologie ihrerseits das Wort ursprünglich mit einer gewissen  
umgangssprachlichen Vorbelastung adoptiert hat (man denke zum Beispiel an das  
Wort „Speicherkapazität“), zum anderen weil in dem Maße, wie die Informations-  
und Kommunikationswissenschaften von der Öffentlichkeit wahrgenommen wer-  
den, die fachspezifischen Begriffsinhalte wieder zurück auf den umgangssprachli-  
chen Gebrauch abfärben. In anderen Worten: Fachterminologische und umgangs-  
sprachliche Belegung der Begriffe sind zwar zu unterscheiden, aber keineswegs  
unabhängig voneinander. In einer „informationstheoretischen Metapher“, wie im  
Titel angekündigt, soll also der genaue fachterminologische Begriffsinhalt „nur“  
über seine sekundäre Brechung im Umgangsdeutsch mitschwingen.

Informationstheoretische und kryptographische Metaphern sind im Vokabular  
der Molekularen Genetik keine exotischen Raritäten, es ist mit solchen geradezu  
gespickt – Beispiele sind Wörter wie „Informationsgehalt“, „Codierung“, „Ausle-  
sen“, „Message“ und andere mehr. Dieser Sachverhalt, so die zu untersuchende  
These, ist der Verständigung innerhalb der Disziplin, aber auch ihrer Darstellung  
nach außen, in Teilen dienlich, in anderen Teilen jedoch ebenso abträglich.

Was folgt, wird sich im wesentlichen auf zwei Begriffe konzentrieren, einmal  
auf den der *Information* selbst, zu dessen theoretischer Grundlegung sich Robert  
Schaback in seinem Beitrag *Definitionen von Information* ab S. 11 in diesem Band  
äußert, und auf den des *Code*. Wo nötig, wird an ein paar Basisfakten der Moleku-  
laren Genetik erinnert – auf frühere Untersuchungen zum Thema wird aus Platz-  
gründen nur punktuell eingegangen. Auf jeden Versuch einer auch nur annähernd  
erschöpfenden Zitation wurde verzichtet. Für möglicherweise als schmerzlich emp-  
fundene Lücken bittet der Autor um Nachsicht.

## 1 Information

Die Rede von der „Genetischen Information“ hat drei Wurzeln:

1. Die Struktur des genetischen Materials (der Gene) selbst.
2. Die Tatsache, dass dieses Material in seiner spezifischen Struktur von einer  
Zellgeneration zur nächsten weitergegeben wird.
3. Das dem Material innewohnende Potential, den Ablauf einer Reihe biologi-  
scher Prozesse zu steuern.

Gene bestehen aus Desoxyribonukleinsäure (DNA) und diese ist bekanntlich ein  
aus vier verschiedenen Grundbausteinen aufgebautes, lineares Makromolekül: Tau-  
sende bis Millionen davon in aperiodischer Reihung. Dies alleine schon drängt den  
Vergleich mit einem Text auf (vergleiche oben: erste Wurzel), der in einem aus vier

Symbolen bestehenden Alphabet geschrieben ist und in der Molekularbiologie ist man es tatsächlich gewohnt, die Abfolge der vier Bausteine entlang des Molekülfadens – die „Sequenz“ der DNA – in vier Buchstaben des lateinischen Alphabets (A für Adenin, C für Cytosin, G für Guanin und T für Thymin) zu notieren.

Zellteilung ist der universelle Vermehrungsprozess in der belebten Welt. Wenn eine Hefezelle eine Tochterzelle knospt und schließlich als eigenständige Zelle abschnürt, dann verfügen beide Zellen über dieselbe genetische Ausstattung; offenbar wird die in der Mutterzelle vorhandene DNA zunächst exakt kopiert und eine Kopie vor der Abschnürung in der Tochterzelle deponiert. In anderen Worten: Eine Nachricht („Text“) wird von A nach B gesendet, was einen Akt von Kommunikation konstituiert – vergleiche oben: zweite Wurzel. (Für die in diesem Zusammenhang geltende, spezielle Bedeutung des Worts „Nachricht“ – insbesondere seiner Abgrenzung vom Begriff „Information“ – vgl. den Artikel von R. Schaback ab S. 11 in diesem Heft).

Die Grundvoraussetzung des Kopierens liegt in der besonderen Struktur der DNA: Der „Text“ ist in zwei miteinander verdrehten Makromolekülen doppelt angelegt – allerdings nicht als „Original und Durchschrift“ sondern gewissermaßen als „Positiv und Negativ“ (plus/minus). In der DNA-Doppelhelix stehen sich nämlich an jeder Kettenposition je ein Element vom einen und vom anderen Strang gegenüber, wobei spezielle Regeln der Nachbarschaftsbeziehung eingehalten werden: es bilden immer A und T ein Pärchen „von Strang zu Strang“ und ebenso G und C. Die strukturchemischen Grundlagen dieser Paarbildung sind sehr schlicht, für die Zwecke des gegenwärtigen Arguments aber unerheblich.

Der Kopiervorgang selbst ist wie folgt zu beschreiben: Zunächst werden die beiden Stränge voneinander getrennt und jeder bildet die Vorlage, um aus den monomeren (heißt: nicht zu einem Faden verknüpften) Grundbausteinen einen neuen Strang zu synthetisieren – immer nach der Regel „A gegenüber T und G gegenüber C“. Der „plus-Strang“ dient also als Vorlage zur Synthese eines neuen „minus-Strangs“ und vice versa. Das Ergebnis sind zwei Doppelstränge, die untereinander und mit dem Muttermolekül strukturell identisch sind, wobei in beiden Fällen jeweils ein Strang durchgehend von der ursprünglichen Doppelhelix stammt und der andere ebenso durchgehend neu aus den Monomeren synthetisiert wurde. Dies ist eine sehr spezielle Art einen Text zu kopieren; tritt man aber einen Schritt zurück und schließt die Augen gegenüber den molekularen Details, entspricht das Ergebnis völlig dem des Abschreibens eines Buchs.

Soviel zu den beiden ersten der oben erwähnten drei Wurzeln, zu Struktur und Kopieren („Replikation“) des genetischen Materials. Mit deren Verständnis ist aber wenig mehr erreicht, als die Oberfläche anzukratzen: Die Weitergabe des genetischen Textes von einer Generation zur nächsten ist kein Selbstzweck, vielmehr erfüllt der Text eine Funktion – und die besteht darin, Entstehen und fortlaufenden Betrieb der komplexen molekularen Maschinerie zu ermöglichen, die ein Organismus (auch) ist.

Die tiefere Frage ist also die, wie sich diese Funktion aus der bloßen Nachricht gleichsam erhebt und ein erster Ansatz, dem Problem näherzukommen, kann darin bestehen, es unter dem Blickwinkel des semantischen Konzepts der Information

(vgl. wieder den Beitrag *Definitionen von Information*, S. 11 in diesem Band) zu betrachten. Dann rückt der Prozess der Belegung der genetischen Nachricht mit semantischem Gehalt (versuchsweise gleichzusetzen mit der von ihr auszulösenden Funktion) in die Nähe des Lesens eines technischen Textes durch einen sachkundigen Ingenieur, der zum einen die Sprache versteht, in der das Buch geschrieben ist und zum anderen über die notwendigen Fachkenntnisse verfügt, die vorgefundenen Sätze in sinnvolle Beziehung zueinander zu setzen. Man fragt sich dann allerdings, welche Instanz einer unscheinbare Hefezelle dies leisten soll und man scheint in Gefahr zu geraten, diese mit anthropomorphem Ballast zu überfrachten: Die Fähigkeit der Interpretation einer Nachricht schreiben wir Personen wie selbstverständlich zu; eine Hefezelle aber scheint mit den dazu vorauszusetzenden kognitiven Fertigkeiten bei weitem überfordert.

Im Folgenden werden Verfassen, Weitergabe und Interpretation von Nachrichten zunächst in einen allgemeineren Zusammenhang gestellt. Erst hernach wird die Diskussion der genetischen Nachricht wieder aufgegriffen und deren weitere Verarbeitung in der Zelle verfolgt.

Was die *Entstehung* von Nachrichten angeht, ist die Sache unproblematisch: Es gibt viele Beispiele dafür, dass wir ganz unbefangen – und mit intuitiv guter Berechtigung – auch dann von Nachricht (bzw. unschärfer von Information) reden, wenn der Ursprung ein automatischer, möglicherweise sogar in der unbelebten Welt sich abspielender Prozess ist. In diesem Zusammenhang sei nur an die sogenannten Jahresringe von Bäumen und an die streifenförmig alternierende Magnetisierungsrichtung des Meeresbodenbasalts erinnert – beides Phänomene, die ohne bewusstes Zutun entstehen und dennoch eine Nachricht darstellen, die der Mensch mit semantischem Gehalt über die Geschichte der Erde füllen kann – bezüglich verschiedener oszillierender Vorgänge, die sich auf sehr unterschiedlichen Zeitskalen ereignen: warme und kalte Episoden von Wetter und Klima im einen Fall und Umpolungen des Erdmagnetfelds im anderen.

Reziprok zur Situation eines nicht personifizierten *Senders* schickt ein Mensch, der eine numerisch gesteuerte Werkzeugmaschine programmiert, eine Nachricht an einen seelenlosen Automaten. Die Flut von Nullen und Einsen, aus der die Nachricht besteht, ist gemäß einer festgelegten Syntax in Pakete gebündelt und diese stellen Anweisungen dar, die der Automat in geordnete Bewegungen von Maschinenteilen umsetzt, womit er – über den bloßen Nachrichtentransfer hinaus – sogar so etwas zustande zu bringen scheint wie ein Belegen der Nachricht mit semantischem Gehalt.<sup>1</sup>

<sup>1</sup>Bei näherem Hinsehen erweisen sich die beiden formal reziproken Situationen jedoch als in der Substanz alles andere als symmetrisch: Im zweiten Fall ist die Zuordnung einer bestimmten Bewegung zu einer bestimmten Anweisung eine ausschließlich auf starre Regeln gegründete Operation. Ganz anders beim nicht-personifizierten Sender: Um die Streifen alternierender Magnetisierungsrichtung des Meeresbodenbasalts als ein Protokoll vergangener Umkehrungen des Erdmagnetfelds deuten zu können, muss der Beobachter über eine Anzahl einzelner Komponenten von Wissen verfügen und sie zueinander in Beziehung setzen können. Zu diesen Komponenten gehören - neben der Kenntnis des besagten Streifenmusters selbst:

- Kenntnis von Existenz und Stärke des Erdmagnetfelds.
- Kenntnis von Existenz und Mechanismus der Plattentektonik.

Die am Transfer einer Nachricht beteiligten Parteien brauchen also nicht notwendig Personen zu sein und man kann in Bezug auf die Sequenz der DNA den Sprachgebrauch von der „Genetischen Information“ mit zwei qualifizierenden Bedingungen als berechtigt verteidigen:

1. Man verzichtet auf die eigentlich gebotene Unterscheidung und benennt mit „Information“ etwas, was eigentlich allenfalls „Nachricht“ heißen sollte. Das kann man hinnehmen, denn zumindest umgangssprachlich wird die besagte Unterscheidung selten getroffen. Zudem ist man in guter Gesellschaft, wird doch selbst in Teilen der einschlägigen Fachwelt das berühmte, 1948 von Claude Elwood Shannon errichtete, mathematische Gebäude<sup>2</sup> hartnäckig weiterhin als „Informationstheorie“ bezeichnet und nicht, wie es korrekt heißen müsste (und bei Shannon auch hieß), als „Kommunikationstheorie“ – als Theorie der Nachrichtenübermittlung also. (Man beachte, dass diese Konvention die Frage nach dem semantischen Gehalt der genetischen Information bis auf weiteres vertagt.)
2. Man erkennt im Falle der genetischen „Information“ als einen Akt der Kommunikation auch einen solchen an, bei dem weder Sender noch Empfänger bewusst handelnde Personen sind,<sup>3</sup> sondern schlichte, einzelne Zellen oder Zellbestandteile, deren Dynamik irgendetwas anderes zu unterstellen als blinden Automatismus wir nicht den geringsten Grund haben.

- 
- Kenntnis von der chemischen Zusammensetzung des Meeresbodenbasalts.
  - Eine physikalische Theorie des Ferromagnetismus, einschließlich des Konzepts der Curie-Temperatur.

(Man beachte ferner, dass diese Einzelkomponenten ihrerseits zusammengesetzter Natur sind.) Wie all dies zu bündeln wäre in einer „Theorie höherer Ordnung“, die das Streifenmuster überzeugend deutet, war nicht vorgegeben: weder durch eine in den empirisch greifbaren Komponenten selbst angelegte Logik noch durch ein äußeres Regelwerk; ihre erstmalige Formulierung stellt somit einen im besten Sinne schöpferischen Akt dar.

<sup>2</sup>Shannon, C. E. 1948: A Mathematical Theory of Communication. Bell System Technical Journal 27, 379–423

<sup>3</sup>Allerdings ist bei der genetischen Information das Problem eines nicht personifizierten Senders wesentlich tiefer, als dies durch den Verweis auf Jahresringe und Streifen alternierender Richtung der Meeresbodenmagnetisierung erscheinen mag. Letztere nämlich sind bloße, in Materiemustern angelegte Protokolle geohistorischer Ereignisse mit höchst limitierter eigener Wirkmächtigkeit (außer der Tatsache, dass sie dem menschlichen Verstand als Ausgangspunkt scharfsinniger Interpretationen dienen). Demgegenüber ist die genetische Information eingebunden in vielfältige, weitreichende und subtil miteinander verwobene Kausalitätsketten – mit spektakulären Resultaten. Erst mit der Synthese von Evolutionstheorie und Molekularer Genetik wurde es möglich, einen Weg aufzuzeigen, wie ein so großer und differenzierter Fundus an Information, wie er in den Genomen der Organismen angelegt ist, in einem langen historischen Prozess sich durch ausschließlich natürliche Vorgänge bilden konnte und sich laufend weiterentwickelt.

## 2 Code

### Der Genetische Code

Ein Code ist umgangssprachlich zunächst ein schlichtes Regelwerk für die Transformation einer Sequenz, das heißt einer Abfolge von Zeichen, in eine andere – wobei sich die Elemente der zweiten Sequenz aus demselben Alphabet rekrutieren können oder aus einem anderen. Beispiele für den letzteren Fall sind zum einen der Morse-Code, der Buchstaben, Zahlen und Satzzeichen in Kombinationen von Punkten und Strichen umsetzt, zum anderen der „*American Standard Code for Information Interchange*“, kurz ASCII, der typographische und Steuerzeichen durch siebenstellige Binärzahlen repräsentiert. In beiden Fällen ist die Zuordnung in beide Richtungen eindeutig.

Transformationen unter *Beibehaltung* des Alphabets haben vor allem in der Kryptographie eine Rolle gespielt. Ein interessanter, in eine dritte Kategorie gehörender Fall ist die zu historischer Bedeutung gekommene Enigma-Maschine. Eingabe und Ausgabe rekurrieren hier zwar auf die Buchstaben des lateinischen Alphabets, jedoch bewirken spezielle Konstruktionsmerkmale der Maschine (drehbare Walzen), dass sich im laufenden Betrieb die Zuordnungen ständig ändern. Dies kommt – bei Verwendung des immer gleichen Zeichensatzes – der sequentiellen Benutzung vieler Alphabete gleich.

Wie beim Morse-Code oder ASCII geht es beim „Genetischen Code“ um die Transformation einer Zeichenkette in eine, die aus Elementen eines anderen Alphabets besteht – um die Translation,<sup>4</sup> wie man das in der Molekularen Genetik nennt, von der in Nukleotiden vorliegenden Sequenz der Nukleinsäuren in die Aminosäuresequenz der Proteine. Das Proteinalphabet verfügt über mehr Elemente als die vier des Alphabets der Nukleinsäuren; diese sind die zwanzig Standard-Aminosäuren. Bei der Translation (schematisch wiedergegeben in Abbildung 1) wird eine von der DNA abgeschriebene Teilsequenz (sogenannte Messenger-RNA oder mRNA)<sup>5</sup> in eine programmierbare, molekulare Synthesemaschine („Ribosom“) eingefädelt – ganz wie bei der heute obsoleten Fernschreibtechnik ein Lochstreifen in das Eingabe-/Ausgabe-Gerät. Das Ribosom interpretiert je drei Nukleotide der mRNA als Anweisung, an das wachsende Ende des entstehenden Proteinmoleküls eine ganz bestimmte der zwanzig verschiedenen Aminosäuren anzuheften. Zum Beispiel führt die Gegenwart des Monomerentriplets AAC im „Lesekopf“ (LK) der Maschine dazu, dass an ihrem „Schreibkopf“<sup>6</sup> (SK) die Aminosäure Asparagin (Asn) zu liegen kommt und an das Kettenende angefügt wird.

Die colineare Transformation vom Nukleinsäure- ins Proteinalphabet wird, Schritt für Schritt, von Adaptormolekülen (tRNAs) geleistet, die ihrerseits auch

<sup>4</sup>Keine ganz glückliche Metapher. Im Englischen, aus dem das Wort importiert wurde, bedeutet *translation* „Übersetzung“, und einen Text von einer Sprache in eine andere zu übersetzen ist mit seiner bloßen Transformation von einem Alphabet in ein anderes nicht zu vergleichen.

<sup>5</sup>RNA (Ribonukleinsäure) ist eine zweite, mit der DNA strukturell eng verwandte aber nicht identische Klasse von Nukleinsäuren. RNAs enthalten U (Uracil) anstelle von T.

<sup>6</sup>Die Wörter „Lesekopf“ und „Schreibkopf“ sind dem für die Beschreibung von Turing-Maschinen üblichen Sprachgebrauch entliehen. Dies erscheint hier durchaus angemessen.





Der Genetische Code ist somit nichts anderes als ein in einer einfachen Tabelle darstellbarer Satz von Regeln, der spezielle Abfolgen von je drei Nukleinsäurebausteinen mit je einer Aminosäure korreliert (siehe Abbildung 2) – ganz wie beim ASCII, nur dass dieser mit einer Wortbreite von sieben bit operiert, der Genetische Code mit sechs.<sup>7</sup> Das Wort „Genetischer Code“ bezogen auf die Translation – ist offenbar der Modellfall einer passgenauen Metapher.

Erste Position	Zweite Position				Dritte Position
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Abbildung 2. Zuordnungstabelle Nukleotidtripletts (mRNA) / Aminosäuren (Protein). Man beachte, dass Nukleinsäurefäden Polarität aufweisen; zum Beispiel sind CAG und GAC strukturell unterschiedlich und deshalb im Code verschieden zu belegen (mit Gln bzw. Asp). Die Abkürzungen sind wie folgt: A: Adenin, C: Cystein, G: Guanin, U: Uracil, Ala: Alanin, Asn: Asparagin, Arg: Arginin, Asp: Asparaginsäure, Cys: Cystein, Gly: Glycin, Gln: Glutamin, Glu: Glutaminsäure, His: Histidin, Ile: Isoleucin, Leu: Leucin, Lys: Lysin, Met: Methionin, Phe: Phenylalanin, Pro: Prolin, Ser: Serin, Thr: Threonin, Trp: Tryptophan, Tyr: Tyrosin, Val: Valin.

Unter den weiteren Unterschieden zwischen DNA und Proteinen sind zwei für die Funktion der letzteren von besonderer Bedeutung:

- Die Aminosäuren verfügen über ein weit reichhaltigeres Repertoire an molekularen Substrukturen („funktionellen Gruppen“), was sie wesentlich flexibler hinsichtlich Übernahme verschiedener chemischer Aufgaben macht. Aus der Substanzklasse der Proteine rekrutieren sich daher mehrere für die Aufrechterhaltung unterschiedlichster molekularer Lebensprozesse notwendige Stoffe – allen voran die Enzyme, heißt: Katalysatoren für je eine der vielen verschiedenen Stoffwechselreaktionen. In jeder Zelle gibt es von ungefähr eintausend an aufwärts verschiedene Enzyme.

<sup>7</sup> Anders als beim ASCII sind die Zuordnungen des Genetischen Codes nur in einer Richtung (Nukleinsäure → Protein) eindeutig und nur in dieser Richtung dient er der sequenzspezifischen biologischen Polymersynthese.

- Im Gegensatz zur DNA ist bei den Proteinen der Polymerfaden räumlich nicht monoton gefaltet (Doppelhelix). Jedes Protein wird zunächst zwar als eindimensionaler Faden vom Ribosom abgespult (siehe Abbildung 1), faltet sich anschließend aber zu einer dreidimensionalen Struktur, ohne die es seine biochemische Aufgabe nicht erfüllen könnte. Diese Faltung ist für alle Moleküle derselben Art gleich und der Sequenz eines bestimmten natürlichen Proteins ist (unter physiologischen Bedingungen) offenbar eine bestimmte, für seine Wirkung essentielle Raumstruktur eindeutig zugeordnet.<sup>8</sup>

An dieser Stelle hat die bequeme Übertragbarkeit menschlicher Vorstellungen von Text auf molekularbiologische Sachverhalte ein abruptes Ende: Kein Gedicht verdankt seine Wirkung einer bestimmten dreidimensionalen Faltung des Papiers, auf das es geschrieben ist. In einem per Analogie zum Genetischen Code versuchten, konzeptionellen Brückenschlag war dennoch in der Literatur lange von einem „Proteinfaltungscode“ die Rede, den es noch zu entdecken gäbe. Die Verhältnisse sind jedoch nicht zu vergleichen: Die Tabelle von Abbildung 2 könnte, anderen Ausgang einiger stochastischer Weichenstellungen in der ganz frühen Evolution vorausgesetzt, auch völlig anders aussehen. Ferner lassen sich unter Anwendung des Genetischen Codes über die entsprechenden Nukleinsäuren beliebige Proteinsequenzen „programmieren“ – auch experimentell per chemischer Nukleinsäuresynthese.<sup>9</sup> Für die Festlegung der Proteinfaltung durch die Sequenz ist jedoch ein ähnlicher Satz schlichter Korrelationsregeln auszuschließen, Was an seine Stelle zu treten hat, wird im übernächsten Abschnitt („Eine Alternative zum Bauplan“) näher ausgeführt.

### 3 Die DNA ein Bauplan?

So viel zu Codes als Zuordnungstabellen. Ein Softwareentwickler, der sagt: „Gestern habe ich 250 Zeilen Code geschrieben“, meint etwas ganz anderes. Mit dem sogenannten „Quellcode“, an dem er da gearbeitet hat, wird die Logik eines Algorithmus auf die Form eines geordneten Satzes konkreter Rechenanweisungen gebracht, die in ihrer Gesamtheit ein „Programm“ ausmachen. Hier geht es offenbar nicht um die Transformation von Zeichenketten und an „Code“ im ursprünglichen Sinn erinnert allenfalls die Tatsache, dass auch hier eine Sache für eine andere steht: Die Anforderungen, die der Prozessor eines Computers an die Struktur ihm zu vermittelnder Instruktionen stellt, weichen von menschlicher Sprache nämlich so weit ab, dass man – aus Gründen besserer Handhabbarkeit – die Aufgabe, ein Programm zu entwickeln, in zwei Stufen zerlegt: Die Anweisungen werden zunächst in einer intuitiv besser zugänglichen Programmiersprache wie zum Beispiel „C“ oder „Java“ geschrieben und das Ergebnis, eben der Quellcode, danach

<sup>8</sup>In diesem Sinne sind tRNAs – ihrer Adaptorfunktion angemessen – sowohl in der Welt der Nukleinsäuren zu Hause (Codon-Anticodon-Wechselwirkung mittels Basenpaarung) als auch in der der Proteine (spezifische Funktion begründet in 3D-Struktur – vergleiche oben).

<sup>9</sup>Heute eine Routineangelegenheit – zum Beispiel in der pharmazeutischen Industrie bei der Entwicklung proteinischer Wirkstoffe.

durch ein spezielles, für die jeweilige Programmiersprache allgemein einsetzbares Programm („Compiler“) für den Maschinengebrauch aufbereitet.<sup>10</sup> Dieser zweite Schritt verlangt keinen weiteren menschlichen Eingriff mehr.

Die Kompilation eines Quellcodes hat ihre Entsprechung unter den menschlichen Aktivitäten also nicht in der Chiffrierung/Dechiffrierung von Zeichenketten sondern in der Übersetzung eines Textes von einer Sprache in eine andere und da in Bezug auf letzteres niemand von „Code“ spricht, stellt der etablierte Sprachgebrauch geradezu eine Einladung zu Missverständnissen dar.<sup>11</sup>

Diese begriffliche Verwerfung ist bei der sich erst später durchsetzenden Zweitverwendung des Wortes „Code“ in der Molekularbiologie im Sinne eines Programms in diese mit importiert worden und hat zu dem weitverbreiteten, hartnäckigen und höchst schädlichen Missverständnis – auch *innerhalb* der Molekularbiologie! – geführt, die *gesamte* DNA, das „Genom“, eines Lebewesens, sei so etwas wie ein Satz von Anweisungen für dessen Bau.

Dieses Missverständnis hat die Molekularbiologie aber nicht gänzlich unvorbereitet getroffen, denn ironischerweise hat es seine Wurzeln sogar in einer Zeit noch vor der Aufklärung der DNA-Struktur: In einem Aufsatz unter dem Titel „*The information content and error rate of living things*“ haben Sidney M. Dancoff und Henry Quastler 1953 die folgende Analogie für die Reproduktion eines Organismus aufgestellt:<sup>12</sup>

*Given the following: a builder has an object standing in front of him; it is constructed out of building stones of which there is a large supply; the job is to build a second object identical with the first from the building stones. In order to accomplish this, the builder needs a certain large number of instructions about which type of building stone goes where. The total number of instructions needed will be defined as the information content of the object.*

Das ist eine gute Wiedergabe einer in algorithmischer Form angelegten technischen Bauanleitung.<sup>13</sup> und ähnlich ist es auch zugegangen, als beim Wiederaufbau der Dresdner Frauenkirche gut erhaltene, aus dem Trümmerhaufen geborgene Steinblöcke in den neu entstehenden Baukörper integriert wurden. Für eine ordentliche Anzahl dieser Blöcke ließ sich deren Position im ursprünglichen Bau rekonstruieren; entsprechend hat man sie markiert und in eine Liste eingetragen. Dann hat man die Blöcke in riesigen Regalen neben der Baustelle zwischengelagert und jeder einzelne wurde, wenn er an der Reihe war, an seine in der Liste verzeichnete Stelle gehievt und in das Mauerwerk eingefügt. Der fertigen Kirche sieht man diesen Mosaikcharakter noch heute an.

<sup>10</sup>Genau genommen spricht das kompilierte Programm nicht den Prozessor direkt an, sondern das Betriebssystem des Rechners, ein eigenes Element von Software, das in der heute üblichen Situation dazwischengeschaltet ist.

<sup>11</sup>Gewissermaßen das reziproke Problem zu dem in Fußnote 4 angesprochenen.

<sup>12</sup>Dancoff, S. M und Quastler, H. 1953: *The information content and error rate of living things*, in *Essays on the use of information theory in biology*, Henry Quastler ed., University of Illinois Press, Urbana, 1953, pp. 263–273

<sup>13</sup>R. Schaback: Definition von Algorithmischer Information, ab S. 11 dieses Bandes

Angewandt auf den Bau einer Zelle oder gar eines vielzelligen Organismus aus molekularen Modulen, wie der Stoffwechsel sie als chemische Zwischenprodukte liefert, könnte die Metapher von einer solchen Liste jedoch falscher gar nicht sein. Dennoch hat dieses Bild Jahrzehnte lang die Vorstellungen von der Rolle der genetischen Information und damit der DNA in der Entfaltung eines vielzelligen Organismus aus der befruchteten Eizelle, der Zygote, dominiert und den Durchbruch zu treffenderen Konzepten blockiert. Im Jahr 1985 immerhin hat Gunther Stent in einem kleinen Essay<sup>14</sup> folgendes geschrieben (durchaus auch mit einem selbstkritischen Unterton):

*... the tenet that the gene is a one-directional description of the primary structure [heißt: der Monomerenabfolge] of a particular protein molecule was turned, willy-nilly, into the doctrine that the genome is a one-dimensional description of the whole animal. In particular, it became to be believed that the genome embodies, not merely a protein catalog, but also a genetic program for development from zygote to adult.*

Demnach wäre ein Gen sehr wohl eine „*what goes where*“-Liste für den Bau eines Proteins (genauer: des linearen Fadens); das *Genom* aber mitnichten eine „*what goes where*“-Liste für den Bau eines ganzen Organismus und schon gar nicht ein Code (im Sinne eines Programms), welcher die hierfür erforderlichen, einzelnen Prozesse algorithmisch ausbuchstabiert. Mit einer solchen Aufgabe wäre das menschliche Genom in der Tat hoffnungslos überfordert: Schon seine Speicherkapazität mit – grob und großzügig geschätzt – gerade einmal einem Gigabyte wäre dafür viel zu gering.

## 4 Eine Alternative zum Bauplan

Der Versuch von Dancoff und Quastler, das Wesen menschlicher Ingenieurskunst auf die Gestaltbildung in der Biologie zu projizieren, ist offenbar gescheitert<sup>15</sup> und man sah sich gezwungen, für die Entstehung biologischer Muster auf organismischer Ebene nach Alternativen zu einem „festen Code“ Ausschau zu halten. Eine solche Alternative ist Gegenstand dieses Abschnitts.

<sup>14</sup>Stent, G. S. 1985: Thinking in one dimension: The impact of molecular biology on development. Cell 40, 1–2

<sup>15</sup>Man mag sich in der Rückschau fragen, wie der Erklärungsansatz von Dancoff und Quastler zu seiner großen ideengeschichtlichen Bedeutung gekommen sein mag, zumal beide Autoren ansonsten bei der Entwicklung der grundlegenden Konzepte der Molekularbiologie nicht in der ersten Reihe standen. Sicher hat bei der ursprünglichen Rezeption Originalität eine Rolle gespielt, und zwar insofern, als die Genetik erstmals aus dem Blickwinkel der damals brandneuen Kommunikationstheorie von Shannon betrachtet wurde. Die spätere zähe und letztlich kontraproduktive Behauptung dieses Bildes in den Köpfen der Leute hat wohl damit zu tun, dass sich der Ansatz in wohlvertrauten Bahnen menschlichen Denkens über technische Konstruktion bewegt, zu denen sich lange keine Alternative aufdrängte. Davon unabhängig jedoch ist es ein bleibendes Verdienst der Arbeit von Quastler und Dancoff, dass sie wohl als erste auf den Zusammenhang zwischen der Menge biologischer Information und der gerade noch tolerablen Fehlerrate der Reproduktion hingewiesen und für letztere eine quantitative Schätzung versucht haben.

Zunächst noch einmal zur Eingangsfeststellung des ursprünglichen Artikels von Dancoff und Quastler. Sie eröffnen ihr Argument folgendermaßen:

*A living being is a very complex thing. [...] To be complex, an object must necessarily contain many different parts. This implies that the specification of a very complex object will be of considerable length and contain a large amount of information; in other words, a very complex object must have a high information content.*

Das ganze Argument ist auf die Aussage des letzten Halbsatzes zugespitzt, und die trifft auf herkömmliche technische Objekte zu. Daraus – und dies ist der Stolperstein des Ansatzes – darf nicht geschlossen werden, die Aussage sei auf *beliebige* Situationen „eins zu eins“ übertragbar.<sup>16</sup> In den Jahrzehnten, die seit dem Erscheinen von „*Essays on the use of information theory in biology*“ vergangen sind, hat es auf mehreren Gebieten Entwicklungen gegeben, die illustrieren, dass die Verhältnisse auch ganz anders sein können. Dazu gehören vor allem jüngere Theorien der Dynamik komplexer Systeme und der Selbstorganisation von Materie.

Im folgenden werden nicht im *what-goes-where*-Sinn codierte Prozesse der Strukturbildung biologischer Objekte an vier Beispielen skizziert. Diese gehören zwei deutlich unterschiedlichen Ebenen der Komplexität an, auf denen fundamental unterschiedliche Prinzipien der Strukturbildung herrschen.

- Objekte der unteren Ebene, obwohl an Lebensprozessen beteiligt, verfügen nicht über die Fähigkeit zu eigenständigem Leben. Zu solchen Objekten gehören biologische Makromoleküle und nicht-kovalente makromolekulare Komplexe wie Ribosomen und Viren. Das auf dieser Ebene vorherrschende Prinzip der Strukturbildung ist das des Auffindens des energieärmsten Zustands des Systems – völlig entsprechend der Kristallisation anorganischer und organischer Substanzen wie Kochsalz oder Rohrzucker.
- Die höhere Komplexitätsebene wird von lebenden Systemen gebildet – angefangen bei einzelnen Zellen bis hin zu kompliziert gebauten, vielzelligen Organismen. Diese operieren fernab vom chemischen Gleichgewicht und ihrem Streben zum energieärmsten Systemzustand wird, solange ein Organismus lebt, durch ständige Zufuhr und Wandlung von Energie aktiv entgegengerichtet. Dies gilt auch für die Prozesse der Findung und des Erhalts der Struktur eines Lebewesens.

---

<sup>16</sup>Die Quastler'sche Vorstellung, eine Struktur hoher Komplexität verlange zu ihrem Aufbau viel Information (vgl. oben), hat noch am Beginn des 21. Jahrhunderts für großes Erstaunen (auch in der Fachwelt) darüber gesorgt, dass der Mensch kaum mehr Gene besitzt als das ungleich einfacher gebaute Fadenwürmchen *Caenorhabditis elegans*.

#### 4.1 Moleküle: 3D-Struktur von Proteinen

Die DNA-Sequenz bestimmt die Proteinsequenz in colinearer Codierung (vgl. oben). Um seine (zum Beispiel katalytische) Funktion im Stoffwechselfgeschehen auszuüben, muss sich das neu synthetisierte, fadenförmige Makromolekül aber erst in eine spezifische dreidimensionale Struktur falten.<sup>17</sup>

Diese Fähigkeit ist für synthetische Makromoleküle völlig unbekannt. Was also kann die Natur, was menschliche Chemiker nicht können? Man hat allen Grund zu der Annahme, dass nur ein sehr kleiner Teil aller denkbaren Proteinsequenzen die Fähigkeit eindeutiger Faltung vermittelt und die Natur auf das in der Evolution wirkende Wechselspiel von Mutation und Selektion angewiesen war, um diese kleinen, raren Inseln molekularer Ordnung in einem Meer von Chaos zu finden.

Damit drängt sich sofort eine zweite schwierige Frage in den Vordergrund: Was befähigt natürliche, frisch assemblierte Proteinketten bestimmter Sequenz, aus der ungeheuer großen Anzahl *möglicher* Faltungen („Konformationen“) diese eine ausgezeichnete innerhalb angemessener Zeit tatsächlich zu finden – und was bestimmt, *welche* Faltung dies ist? Die Triebfeder dafür sind abstandsabhängige anziehende und abstoßende Wechselwirkungen zwischen sämtlichen Teilen eines Proteinmoleküls. Zu diesen tragen lokalisierte elektrische Ladungen, Dipolmomente und anderes mehr bei. Im Verlauf der Faltung strebt die Summe aus allen sich paarweise ergebenden Energiebeiträgen einem Minimum zu, das nur mit *einer* einzigen Form der Faltung erreicht wird. Aus diesem Grund ist es nicht nötig, die Lage eines jeden Atoms in der Raumstruktur „per Code“ festzuhalten, sondern sie ergibt sich ohne weiteres Zutun aus der das ganze Molekül betreffenden Tendenz zur Energieminimierung. Das heißt, den Regeln, welche die Proteinfaltung bestimmen, fehlt jedes Element der dem Genetischen Code innewohnenden Kontingenz – sie sind in fundamentalen Eigenschaften der Materie naturgesetzlich festgelegt. Ihren bestimmenden Einfluss auf die dreidimensionale Struktur eines Proteins üben sie *via* den für jede Art von Protein unterschiedlichen Prozess der Faltung aus und sind somit nicht in einer universell und direkt anwendbaren, einfachen Tabelle einzufangen.

Auch ohne einen Faltungscodex zur Hand zu haben sind in jüngerer Zeit dennoch beachtliche Fortschritte gemacht worden, die Faltung eines Proteins ohne Vorwissen aus seiner Sequenz abzuleiten: Mit Hilfe leistungsfähiger Computer kann man die innermolekularen („konformationellen“) Bewegungen von Proteinmolekülen simulieren, dabei die intramolekulare Wechselwirkungsenergie aus all den vielen Einzelbeiträgen (vgl. o.) laufend bilanzieren und einen Weg („Trajektorie“) zum energieärmsten Zustand des Systems suchen.<sup>18</sup>

---

<sup>17</sup>Alle Proteinmoleküle, die unter Kontrolle desselben Gens synthetisiert wurden, bilden eine Population von Makromolekülen einheitlicher chemischer Konstitution – etwas, womit sich die von Menschen betriebene makromolekulare Chemie immer noch sehr schwer tut. Damit aber nicht genug: Alle Proteinmoleküle derselben Art nehmen zudem dieselbe, durch Faltung der Kette entstandene Raumstruktur (Konformation) ein – und dies schon kurz nach ihrer Synthese (bei einer unvorstellbar großen Anzahl alternativer Möglichkeiten).

<sup>18</sup>Einige für die Illustration der Pointe verzichtbare Nebenaspekte sind vernachlässigt.



Selbst wenn das skizzierte Sequenz/Strukturproblem einmal komplett gelöst ist (es ist immer noch weit entfernt davon), setzt darauf gleich ein weiteres auf: das vergleichbar schwierige Struktur/Funktionsproblem: Selbst mit dem besten biochemischen Sachverstand ist auch der präzisest beschriebenen 3D-Struktur eines Proteins unbekannter Funktion nur in Ausnahmefällen anzusehen, was genau diese natürliche Funktion ist – und das erst käme einem Verständnis des semantischen Gehalts der genetischen Nachricht gleich.

## 4.2 Makromolekulare Komplexe: Viren

Zur genetischen Festlegung der Struktur von Viren hat Sidney Brenner ungefähr folgenden, von Roger Lewin 1984 aufgegriffenen Gedanken entwickelt:<sup>19</sup> Einige besonders einfach gebaute Viren bestehen lediglich aus einem DNA-Genom, das in eine Hülle aus Proteinmolekülen eingebettet ist. Bei manchen Viren hat diese Hülle eine Gestalt mit den Symmetrieeigenschaften eines schlichten Ikosaeders.<sup>20</sup> Über beliebig viele Generationen hinweg haben alle Nachkommen eines solchen Virus dieselbe ikosaedrische Gestalt, also ist sie ein erbliches Merkmal und hängt folglich irgendwie an der Sequenz des viralen Genoms. Nirgends in dem Genom jedoch findet man Anweisungen für die Konstruktion eines Ikosaeders in einer Form, wie man sie für einen Montageroboter oder (heute) einen 3D-Drucker schreiben würde. Vielmehr entsteht diese Gestalt als eine Folge der oben bereits skizzierten Proteinfaltung. Für die Proteinmoleküle, aus denen die Virushülle besteht, hat die Faltung nämlich zwei Konsequenzen: Sie verleiht den gefalteten Molekülen nicht nur geometrische Passform im Sinne eines dreidimensionalen Puzzle, sondern statet sie außerdem an speziellen Stellen ihrer Oberfläche mit Kontaktpunkten aus, die gegenseitige Bindungsaffinität zwischen Nachbarn vermitteln. Beides zusammen lässt die Proteine sich auf räumlich definierte Art zusammenlagern. Wiederum also ist nur die Sequenz der Proteine kodiert, das dreidimensionale Objekt ergibt sich aus einem auf die Proteinfaltung aufgesetzten, spontanen Assemblierungsprozesses und das Merkmal „Ikosaeder“ ist auf distributive Weise im Genom insgesamt begründet.

## 4.3 Zwischenbemerkung

Will man daran festhalten, als den semantischen Gehalt der genetischen Nachricht die Funktion der von ihr codierten Makromoleküle zu betrachten (vergleiche oben), kann man ab hier leicht ins Straucheln geraten. Wohl ist die Monomerensequenz eines Proteins so etwas wie „Text“, aber erst dreidimensionale Faltung – und ggf. sich anschließende supramolekulare Assemblierung – schaffen die Grundlage für Funktion (enzymatische Katalyse, Infektion einer Wirtszelle oder anderes). In Kontrast dazu sind wir gewohnt, den semantischen Gehalt einer in menschlicher Sprache abgefassten Nachricht (verbal oder geschrieben) ungefähr colinear zur se-

<sup>19</sup>Lewin, L. 1984: Why is Development so Illogical?, Science 224, 1327–1329.

<sup>20</sup>Aus zwanzig gleichseitigen Dreiecken gebildeter Polyeder.



quentiellen Äußerung der einzelnen Wörter auszulesen:<sup>21</sup> „Abraham zeugte Isaak. Isaak zeugte Jakob. Jakob zeugte Juda und seine Brüder. Juda zeugte Perez . . . “ Auf einer höheren Ebene entstehender, zusätzlicher Gehalt ist literarischen Kunstwerken vorbehalten, zu denen die Verse Matthäus 1:2–16 nicht gehören. Treffende Vergleiche mit den Verhältnissen der Molekularbiologie bieten sich nicht an.

## 4.4 Zellen

Zellen sind dynamische Gebilde und permanente Ortszuweisungen für zelluläre Substrukturen sind eher die Ausnahme. Allerdings gibt es, zusätzlich zum bereits angesprochenen genetischen Material (in Chromosomen organisierte DNA), weitere distinkte Strukturen, die bei der Zellteilung reproduziert werden müssen, z. B. die Proteinfilamente des Zellskelettes und die Membranen, durch die verschiedene Zellkompartimente voneinander abgetrennt werden (und die Zelle insgesamt gegen die Außenwelt). Im typischen Fall werden solche Strukturen nicht durch kovalente chemische Bindungen zusammengehalten sondern sie bestehen, ähnlich wie eine Virushülle (vergleiche oben), aus molekularen Untereinheiten, die nur locker, wenn auch räumlich definiert, miteinander wechselwirken. Auch für den Bau solcher Strukturen gibt es keinen Code: Zum Beispiel wird im Laufe eines Zellzyklus durch Einfügen zusätzlicher Untereinheiten zweimal so viel Membran hergestellt als zu Beginn des Zyklus vorhanden war, dann das Ganze durch Teilung auf die beiden Tochterzellen verteilt.

## 4.5 Organismen

Vielzellige Organismen beziehen ihre äußere Gestalt und innere Organisation daraus, dass verschiedene Typen von Zellen in bestimmter Anzahl bestimmte Plätze einnehmen. Der Körper eines erwachsenen Menschen besteht aus insgesamt ungefähr  $10^{13}$  Zellen, von denen es 250 verschiedene Typen gibt. Bei der Ontogenese eines Menschen entstehen alle diese Zellen durch sukzessive Runden von Teilung aus einer einzigen „Gründerzelle“, der Zygote, die ihrerseits aus der Verschmelzung einer weiblichen und einer männlichen Keimzelle hervorgegangen ist. Neben der Zellteilung *per se* und deren Kontrolle besteht das Problem der Entwicklung also aus zwei Teilaspekten: Erstens der Auffächerung der Linien in die verschiedenen Zelltypen (Differenzierung) und zweitens der Zuweisung spezieller Plätze im Organismus für Zellen eines bestimmten Typs. Genau diese Prozesse glaubte man lange und fälschlicherweise von einem ins Genom geschriebenen, genetischen *Programm* kontrolliert (vergleiche oben: Fußnote 14). Erstaunlicherweise geht es ohne, wie im folgenden Abschnitt anhand eines stark vereinfachenden Modells beleuchtet wird.

---

<sup>21</sup>Eine Grundvoraussetzung jeglichen Simultandolmetschens.

## 5 Ein Gleichnis für die organismische Strukturbildung

Ein Embryo wächst anders als die wiedererstehende Frauenkirche gewachsen ist: Vom Legen des ersten Fundaments an spielt sich das Errichten eines Gebäudes auf dem endgültigen Größenmaßstab ab und ein Bauteil, das – gemäß Plan – an eine bestimmte Stelle gebracht wird, verbleibt dort unverändert und „*for good*“. Ein Embryo hingegen fängt als einzelne Zelle an, akquiriert Masse hinzu, wächst durch wiederholte Teilung zu einem zunächst scheinbar formlosen Zellklümpchen heran und durchläuft durch weiteren Zuwachs, Zellteilung und Zuordnung verschiedener Qualitäten zu distinkten Zellen einen Prozess von räumlicher und funktionaler Strukturierung zunehmender Verfeinerung. Einzelne Zellen können dabei mehrfach ihre Qualität und oft auch ihre Position im Gesamtverband wechseln.

Dieser mühelos festzustellende Unterschied in der Art des Wachstumsprozesses spiegelt die Art und Weise wider, wie die beiden Prozesse kontrolliert werden: Für die Realisierung eines Stücks Architektur oder eines anderen technischen Objekts ist Quastlers *what-goes-where*-Liste völlig adäquat – für das organische Wachstum unter Gestaltbildung, so wie es uns in der Embryogenese entgegentritt, braucht es nicht unbedingt mehr, aber etwas ganz anderes: Abhängig von Zeit und Raum sind in verschiedenen Zellen des reifenden Embryos unterschiedliche Gene aktiv. Die Produkte dieser differentiellen Genexpression (u. A. regulatorische Proteine und RNAs) modulieren ihrerseits die Aktivität anderer Gene und die wechselseitigen Einflüsse sind so miteinander verschränkt, dass einerseits lineare, hierarchisch sich verzweigende Kaskaden von Genexpression entstehen, andererseits aber auch Querverbindungen und Rückkopplungen. Hinzu kommen Signale von außerhalb der Zelle, das heißt solche, die von anderen Zellen des sich entfaltenden Embryos ausgehen oder von der „Welt da draußen“ (zum Beispiel solche über die gerade vorherrschende Umgebungstemperatur).

All dies ist mit beständigem Stoffwechsel und Dissipation<sup>22</sup> chemischer Energie verbunden und ergibt in der Summe einen dynamischen Prozess, der keine zentrale Steuerung kennt, bei dem vielmehr die Prozesskontrolle nicht zu trennen ist vom Prozess selbst. Diese letztere Aussage ist gleichbedeutend mit einer Umschreibung des Phänomens der „Selbstorganisation“. Wie auf der Hand liegt, kann man die Vorstellung der Selbstorganisation auf den Wiederaufbau der Frauenkirche ebenso wenig anwenden, wie umgekehrt die einer *what-goes-where*-Liste auf die Embryogenese.

Für eine Illustration dessen, wie komplexe Muster aus einem Minimum an Instruktion entstehen können, eignen sich sogenannte „Eindimensionale Zelluläre Automaten“ besonders gut.<sup>23</sup> Man stelle sich eine Reihe linear aufgereihter, mit den jeweiligen Nachbarn verbundener Zellen vor. Diese Zellen sind zunächst ganz abstrakter Natur, können aber unter Umständen Modelle für reale Zellen, zum Beispiel biologische, abgeben. Das Modell berücksichtigt ausschließlich den Aspekt der

<sup>22</sup>Umwandlung arbeitsfähiger Energie in (letztlich) Wärme.

<sup>23</sup>Wolfram, S. 2002: *A New Kind of Science*, Wolfram Media, Champaign, IL.

regelhaften Antwort auf Signale, die eine Zelle von benachbarten anderen Zellen empfängt und eben darin liegt die angekündigte, starke Vereinfachung.

Ein „Automat“ ist hier keine reale Maschine, sondern ein mathematisches Gedankenkonstrukt, das man folgendermaßen beschreiben kann: Die Zellen sollen in zwei alternativen Zuständen existieren können, die in Abbildung 3 mit schwarz und weiß symbolisiert werden.

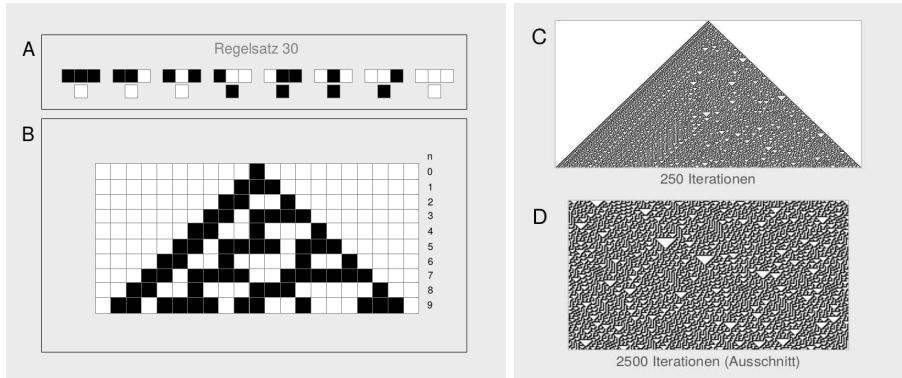


Abbildung 3. A: Ein Regelsatz (30) für die Entwicklung eines Eindimensionalen Zellulären Automaten. Obere Reihe: Die acht bei zwei alternativen Zellzuständen möglichen Zell-Triplets. Untere Reihe: Zustand der zum jeweiligen Triplet gehörenden zentralen Zelle nach Anwendung von Regelsatz 30. B: Schrittweise Entwicklung des Automaten nach  $n$  konsekutiven Anwendungen von Regelsatz 30. C und D: Ebenso nach 250 bzw. 2500 Iterationen.

Abhängig davon, in welchem der beiden Zustände sich eine jeweils betrachtete Zelle gerade befindet – und zusätzlich abhängig vom Zustand ihrer beiden Nachbarn links und rechts – wird der Zustand dieser Zelle in einem folgenden Schritt auf schwarz oder weiß gesetzt. Da es zwei verschiedene Zellzustände gibt und vier verschiedene Nachbarschaftssituationen (weiß/weiß, weiß/schwarz, schwarz/weiß und schwarz/schwarz) sind insgesamt acht verschiedene Konstellationen zu berücksichtigen, die man willkürlich und unabhängig voneinander mit den alternativen Regeln „setze auf schwarz“ oder „setze auf weiß“ belegen kann. Aus Gründen der Kombinatorik gibt es also  $2^8 = 256$  verschiedene Regelsätze. In Abbildung 3 ist einer davon (willkürlich „Regelsatz 30“ genannt) grafisch illustriert.

Durch gleichzeitige Anwendung des aus dem Angebot von 256 Alternativen ausgewählten Regelsatzes auf alle Zellen eines wie beschriebenen Schwarz-Weiß-Musters erzeugt der Zelluläre Automat in einem Schritt ein neues Muster – und daraus im übernächsten wieder ein neues. Reiht man diese Muster in chronologischer Folge aneinander, so ergibt sich ein zweidimensionales Protokoll der Entwicklung des Automaten in der Zeit, wie dies in Abbildung 3 B bis D für Regelsatz 30 und das minimalistische Ausgangsmuster einer einzigen schwarzen Zelle illustriert ist.

Es entsteht ein offenbar ziemlich komplexes Muster, das durch zahlreiche verschieden große, gleichschenklige Dreiecke dominiert wird, die sich teilweise überlappen und mit ihrer auf der Symmetrieachse liegenden Ecke alle in Richtung auf die Wachstumsfront des Musters weisen.

Dieses Muster vergleiche man mit der Pigmentierung des Gehäuses einer Meeresschnecke der Gattung *Conus* – ästhetisch reizvolle, aber auch höchst gefährliche Tiere, von denen einige Arten Jagd auf Fische machen und dabei ein sofortige Lähmung hervorrufendes Nervengift in ihre Beute injizieren. Selbst erwachsene Menschen können durch den Stich einer Conusschnecke getötet werden.

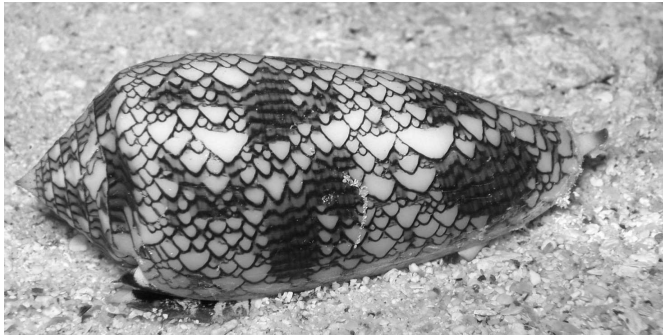


Abbildung 4. Ein Exemplar von *Conus textile*. Die bis zu ca. 10 cm langen Gehäuse von räuberischen Meeresschnecken der Gattung *Conus* zeichnen sich durch auffallende, Spezies-typische Pigmentierungsmuster aus. [https://en.wikipedia.org/wiki/Conus\\_textile#/media/File:Textile\\_cone.JPG](https://en.wikipedia.org/wiki/Conus_textile#/media/File:Textile_cone.JPG). Copyright (c) 2005 Richard Ling. Verwendung unter Creative Common Lizenz (Typ CC BY-SA 3.0).

Das Pigmentierungsmuster des Schneckenhauses ist offenbar nicht identisch mit dem Muster, das der als Beispiel gewählte Zelluläre Automat erzeugt hat, weist aber doch frappierende Ähnlichkeit auf – einschließlich der Orientierung der Dreiecke relativ zur Wachstumsfront. Das Schneckenhaus ist das auf Kalk geschriebene, wie auf einer Schriftrolle aufgewickelte Protokoll der das Gehäusewachstum begleitenden Pigmentproduktion.

In der Tat passen einige Aspekte der zum Bau des Schneckenhauses führenden Biomineralisation und Pigmentierung gut ins Bild eines Zellulären Automaten; dazu gehören:

1. Die Tatsache, dass diese Prozesse von wirklichen, lebenden Zellen betrieben werden.
2. Die lineare Aufreihung dieser Zellen entlang der Wachstumsfront (i. e. der Kante) des Schneckenhauses – wenn auch nicht zwingend in nur einer Reihe.
3. Die Plausibilität, mit der man sich die zuständigen Zellen in Bezug auf die Pigmentsynthese in zwei Zuständen („Schalterstellungen“), „AN“ und „AUS“, vorstellen kann.

4. Die Fähigkeit biologischer Zellen, miteinander zu kommunizieren, auch – und für das Modell des Zellulären Automaten besonders relevant – durch direkten Kontakt.
5. Die anhand vieler Beispiele dokumentierte Möglichkeit, genetische Schalter (Punkt 3) an Zelle-zu-Zelle-Kommunikation anzuhängen.

Suggestiv wie all dies klingen mag, ist dennoch Vorsicht geboten, die auffallenden Entsprechungen auf der phänomenologischen Ebene vorschnell für den Beleg eines vergleichbaren zugrunde liegenden Mechanismus zu halten, zumal anderes weit weniger gut passt – so zum Beispiel die in *diskreten* Zeitschritten synchronisiert sich vollziehende Entwicklung eines Zellulären Automaten.<sup>24</sup>

Es kommt hier aber gar nicht darauf an, ob hinter dem Pigmentierungsmuster eines *Conus*-Schneckenhauses tatsächlich ein in lebender Materie angelegter Zellulärer Automat am Werk ist; es reicht, den illustrierten Fall als ein Gleichnis zu begreifen, das in *allgemeiner* Form auf folgendes verweist: Komplexe aber geordnete Muster können ohne einen Bauplan („*what-goes-where*-Liste“) entstehen, dessen Ausführung durch eine zentrale Instanz koordiniert wird. Vielmehr reicht dazu ein distributiver Prozess aus, das heißt einer, der auf viele Entitäten verteilt ist, welche alle nur jeweils lokale Hinweise benutzen und diese nach einem kollektiv gültigen, potentiell sehr einfachen Regelwerk umsetzen.

Für die Embryogenese bedeutet dies: Veränderung von Genexpression und Musterbildung gehen weithin Hand in Hand, nur ist die DNA nicht mehr das *master molecule* sondern zunächst der *Empfänger* von Signalen, die vom sich entwickelnden Muster ausgehen. Dieses ist in eingewickelter Form bereits im Ausgangszustand und den Regeln angelegt und es entfaltet sich durch deren iterative Anwendung „von selbst“: Der Prozess benötigt keine Zufuhr zusätzlicher Information. Das Muster – die Architektur einer adulten Fliege zum Beispiel – hat, obwohl höchst komplex, einen geringen Gehalt an algorithmischer Information<sup>25</sup> – gering

<sup>24</sup>Die bei der ontogenetischen Entwicklung eines Lebewesens entstehenden (und zum Teil wieder vergehenden) Materiemuster lassen sich im Prinzip auch in *kontinuierlicher* Zeit simulieren – mit partiellen Differentialgleichungen, die Synthese, Ausbreitung und Abbau von Botenstoffen („Morphogene“) in lebendem Gewebe modellieren. Dieser mathematisch aufwendigere Ansatz geht auf Alan Turing und einen von ihm im Jahr 1952 publizierten Artikel zurück (Turing, A.M. 1952: The Chemical Basis of Morphogenesis. Phil. Trans. R. Soc. London B 237, 37–72). Später wurde er vor allem von Hans Meinhardt am MPI für Entwicklungsbiologie in Tübingen weiterentwickelt. Meinhardt hat speziell dem Problem der Pigmentierungsmuster bei Meeresschnecken und anderen Mollusken ein ganzes Buch gewidmet (Meinhardt, H. 1997: *Wie Schnecken sich in Schale werfen: Muster tropischer Meeresschnecken als dynamische Systeme*. Springer, Berlin.) Siehe auch Gunji, Y, 1990: Pigment color patterns of molluscs as an autonomous process generated by asynchronous automata. *BioSystems* 23, 317–334.

<sup>25</sup>Als Maß der algorithmischen Information eines Musters dient hier die Länge des kürzesten Algorithmus, der ausreicht, den Musterbildungsprozess festzulegen. Ein anderes, sehr einfaches Beispiel: Genome höherer Organismen besitzen in größerer Zahl sogenannte repetitive Sequenzen, z. B. solche der Form (TG)*n* – mit *n* in der Größenordnung von Dutzenden bis Hunderten. Häufigkeit und Ausdehnung solcher Muster stehen in flagrantem Gegensatz zu ihrem a priori sehr geringen statistischen Erwartungswert. Unter dem Aspekt algorithmischer Information schrumpfen sie zu einer Trivialität. Für den schrittweisen Aufbau einer solchen Sequenz (auf dem Papier!) gilt das folgende schlichte Regelpaar:

genug jedenfalls, um in das Fliegen genom zu passen. Weit und breit nichts zu sehen von einem „festen Code“.

## 6 Code – Zusammenfassung

Die Bezeichnung „Genetischer Code“ für die Gesamtheit der Korrelationsregeln zwischen Nukleinsäure- und Proteinsequenz ist gelungen und unkontrovers. Im Gegensatz dazu ist es höchst problematisch, den Code-Begriff, wie geschehen, über die Festlegung der Monomerabfolge in linearen Makromolekülen hinaus auf die genetische Bestimmung höherer Strukturen biologischer Objekte auszudehnen.

Für die Montage eines komplexen Objekts aus seinen Einzelkomponenten „nach Art von Quastler“ ist zunächst ein detailliertes Bild seiner Struktur notwendig, innerhalb der jeder Komponente individuell eine „Adresse“ zugewiesen wird (zum Beispiel als Triplet von Koordinaten im cartesischen Raum). Ferner muss für jede Komponente ein Prozess festgelegt werden, mittels dessen sie an den ihr zugewiesenen Ort gebracht wird. Die Länge der Baubeschreibung (als Code) ist somit eine Funktion der Komplexität des Objekts; sie kann als proportional zu  $N^x$  erwartet werden (mit  $N$  als Anzahl der Einzelkomponenten und  $x \geq 1$ ). In der Biologie verfehlt dieser Erklärungsansatz das Ziel – schon aus Gründen unzureichender Speicherkapazität des Genoms.

Für die Gestaltfindung biologischer Objekte ist vielmehr das Phänomen der *Selbstorganisation* bedeutend – und dies auf zwei fundamental verschiedene Weisen, die auf unterschiedlichen Ebenen der Komplexität operieren. Auf der unteren Ebene, der der räumlichen Struktur von Makromolekülen, herrscht als strukturbildendes Prinzip, ähnlich der Kristallisation, das Auffinden der niedrigsten Energie des Systems. Hierfür wird weder Zufuhr von Information noch Dissipation chemischer Energie benötigt.

Der Weg von der Nachricht (Sequenz eines Gens) zum semantischen Gehalt (biochemische Funktionen des dazugehörigen Proteins) ist auf der ersten Wegstrecke mit gängigen Metaphern der Textverarbeitung treffend zu beschreiben; den restlichen Weg finden die Moleküle alleine – geleitet lediglich durch im Prinzip einfache, in ihrem komplexen Zusammenspiel aber schwer zu durchschauende, chemische Prinzipien – und so zerrinnt dem menschlichen Betrachter der genetischen Information schon auf dieser einfachen Stufe der semantische Ansatz zwischen den Fingern.

- 
- Wenn der im vorausgegangenen Schritt angefügte Rest ein T ist, nimm als nächsten ein G.
  - Wenn der im vorausgegangenen Schritt angefügte Rest ein G ist, nimm als nächsten ein T.

Als dritte Regel könnte hinzukommen: Beende nach dem Hinzufügen eines Rests den Prozess mit einer Wahrscheinlichkeit von 0.x.

Soviel zu zwar biologischen, aber jeweils isoliert unbelebten Molekülen. Leben hingegen baut auf chemischer Dynamik in einem offenen, das heißt von Stoff- und Energieströmen durchflossenen System auf; gegen die Gleichgewichtseinstellung wird aktiv gearbeitet; die entstehenden Muster sind per se transienter Natur und ihr Erhalt von dauernder Energiedissipation abhängig. Die Struktur des Objekts wird durch Gesetzmäßigkeiten der Dynamik sowie deren Start- und Randbedingungen bestimmt. (Im engeren Sinn wird nur dieser zweite Weg der Gestaltbildung als Selbstorganisation bezeichnet.)

Die Struktur entwickelt sich dann – wie oben anhand eines Zellulären Automaten gleichnishaft beleuchtet – „von selbst“. Die Regeln der Dynamik können beliebig kompliziert, potentiell jedoch sehr einfacher Natur sein: Ein Zellulärer Automat ist mit wenigen Zeilen Computercode zu beschreiben und produziert – einen entsprechenden Regelsatz vorausgesetzt – dennoch ein komplexes Muster.<sup>26</sup>

Die Frage ist nicht: Was ist die kleinste Menge an Information, die die Struktur *beschreibt*, vielmehr ist die Frage: Was ist die kleinste Menge an Information, die den Prozess festlegt, der die Struktur (Vorliegen bestimmter Umstände vorausgesetzt) *erzeugt*?

Bezogen auf die ontogenetische Entwicklung eines vielzelligen Lebewesens sind selbst die Regeln ihrerseits nicht als algorithmischer Code angelegt; vielmehr haben sie ihre Wurzeln in subtilen und spezifischen Wechselwirkungen zwischen molekularen Komponenten der lebenden Materie. In dem dynamischen Geschehen spielt die DNA zwar eine bedeutende Rolle – in ihr sind die Sequenzen der Polymeren codiert, die Akteure im dynamischen Geschehen sind – sie hat aber nicht den Part des vielzitierten *master molecule* und angeblichen Ursprungs aller Kausalitätsketten im Zellgeschehen.<sup>27</sup>

---

<sup>26</sup>Stephen Wolfram hat es folgendermaßen ausgedrückt: „*Before the discoveries of this book, one might have thought that to create anything with a significant level of apparent complexity would necessarily require a procedure which itself had significant complexity. But what we have discovered in this book is that in fact there are remarkably simple programs that produce behaviour of great complexity.*“ Das liest sich wie eine direkte Entgegnung auf die von Dancoff und Quastler aufgestellte, apodiktische Behauptung (vergleiche oben) – mit dem zeitlichen Versatz eines halben Jahrhunderts. „*This book*“ ist das bereits zitierte Werk „*A New Kind of Science*“ (siehe Fußnote 23) und Wolfram ist dafür gescholten worden, mit Sätzen wie dem oben wörtlich wiedergegebenen zuviel an Entdeckung für sich persönlich zu reklamieren. Ersetzt man „*Before the discoveries of this book*“ durch „*Earlier*“, können wohl alle darin übereinstimmen, dass Wolfram eine wesentliche Einsicht der Komplexitätsforschung in angemessene Worte kleidet.

<sup>27</sup>Letzteres Missverständnis hat etliche Molekularbiologen sogar dazu gebracht, für Wesen und Wirken der DNA die altherwürdige Denkfigur des „unbewegten Bewegers“ zu bemühen: Siehe zum Beispiel Max Delbrücks Aufsatz „Aristotle – tole – tole“ in *Of Microbes and Life*, J. Monod und E. Borek eds. 1971, Columbia University Press, New York, pp. 50–55.



## 7 Gebrauch von Metaphern im Öffentlichen Diskurs

Zum Abschluss sollen anhand eines Beispiels mögliche praktische Konsequenzen des Gebrauchs schräger Metaphern für den *öffentlichen* Diskurs beleuchtet werden. Kaum ein Forschungsunternehmen hat in jüngerer Vergangenheit größere öffentliche Aufmerksamkeit auf sich gezogen als die Sequenzanalyse des Genoms des Menschen („*human genome project*“). Zur Zeit des lautesten Medienrummels startete einen aus fast jeder Tageszeitung das Wort von der „Entschlüsselung des menschlichen Genoms“ an. „Entschlüsselung“ ist eindeutig eine Metapher aus der Kryptographie und bezeichnet dort die Aufklärung des Regelwerks, mit der eine Botschaft verschlüsselt wurde oder die sich daran anschließende Anwendung dieses Regelwerks „rückwärts“ um eine verschlüsselte Botschaft in Klartext zu transformieren.

Als der Genetische Code (ursprüngliche Begriffsbelegung – vergleiche oben) entschlüsselt, die Enigma-Maschine enträtselt war, hat dies den jeweiligen *Schlusspunkt* eines umfangreichen und intellektuell höchst anspruchsvollen Forschungsprojekts markiert. Von Stund‘ an konnte man anhand einer auf jeden Bierdeckel passenden Zuordnungstabelle die Nukleotidsequenz eines beliebigen Strukturgens in Proteinsequenz transformieren bzw. – mit einigem mehr an Aufwand – den chiffrierten Funkverkehr der Deutschen Wehrmacht mitlesen.

Somit liegt im Wort von der Genomentschlüsselung die Verheißung, „das Genom“ habe nunmehr seine Geheimnisse preisgegeben und die Arbeit sei beendet. Da ferner das Missverständnis immer noch weit verbreitet ist, das Genom sei so etwas wie eine zentrale Leitstelle für *alle* biologischen Strukturen und Prozesse, schließt dies die noch weiterreichende Verheißung ein, man verstehe nunmehr die gesamte Biologie des untersuchten Organismus – und darin liegt ein weiteres Versprechen eingeschlossen: Wer die Physiologie durchdrungen hat, der beherrscht auch die Pathologie besser und somit werden wir alle – dank neuer Medikamente, die uns die „Genomentschlüsselung“ in den Schoß fallen lässt – ab morgen viel gesünder sein. Was da aber mit Genomentschlüsselung bezeichnet wird, ist in Wahrheit etwas ganz anderes: Es ist nicht mehr und nicht weniger als die experimentelle Feststellung welche Nukleotidsequenz ein ganzes Genom hat; es ist das Sichtbarmachen der bloßen Nachricht – vor jeder Interpretation. An diesem Punkt fängt die Forschung über biologische *Funktion* – die einen wirklich interessiert – erst an. Bei dieser Anstrengung ist die neu gewonnene Kenntnis der Sequenz sehr hilfreich, z. B. für die direkte Ableitung *in silico* von Proteinsequenzen in großer Zahl – keine kleine Errungenschaft der modernen Molekularen Genetik ganz ohne Zweifel, aber eben nur ein Startpunkt und auf keinen Fall mit der angestrebten Erkenntnis selbst gleichzusetzen. Wir haben es bei der „Genomentschlüsselung“ also mit einer Metapher zu tun, die durchaus in der Lage ist, eine klare Vorstellung in den Köpfen der Adressaten zu erzeugen<sup>28</sup> – diese Vorstellung aber ist *falsch*.

<sup>28</sup>Vergleiche einleitende Feststellungen.



Unbeschadet dessen feiert der besagte Sprachgebrauch innerhalb und außerhalb der Wissenschaft fröhlich Urständ – dies ist ein Fakt des Lebens und weitere Beispiele ähnlicher Art ließen sich anschließen. Wie solche Begriffsverwirrung in die Welt kommt, wurde hier zumindest angedeutet; wie sie sich dort jedoch hartnäckig perpetuiert und selbst verstärkt, ist eine interessante Frage eigener Berechtigung, die zu beleuchten den Rahmen dieses Aufsatzes sprengen würde.

Alles Haarspalterei? Wie das Beispiel des Worts „Code“ zeigt, bergen schlecht gewählte oder hinsichtlich ihres Geltungsbereichs unüberlegt erweiterte Metaphern die sehr reale Gefahr, das Denken *innerhalb* eines Fachs zu blockieren. In der Vermittlung von Wissenschaft nach *außen* verzerrt metaphorische Effekthascherei („Genomentschlüsselung“) leicht das Bild einer Disziplin. Häufig ist sie dazu ange-tan, der interessierten (und zahlenden!) Öffentlichkeit gegenüber nicht einlösbare Versprechungen zu implizieren – und die holen einen erfahrungsgemäß früher oder später wieder ein.



# Informationsgesteuerte Synthese – eine Blockbuster-, ‚Erfindung‘ der Evolution

*Hans-Joachim Fritz*

*Institut für Mikrobiologie und Genetik der Universität Göttingen  
Derzeitige Adresse: Institut für Organische Chemie und Biomolekulare Chemie  
der Universität Göttingen, Tammannstraße 2, D-37077 Göttingen*

*Leben, so wie wir es kennen, setzt zwingend informationsgesteuerte Synthese von Makromolekülen voraus; entsprechend markiert ihr erdgeschichtlich frühestes Auftreten einen Durchbruch der Evolution an der Schwelle zu voll ausgebildeten Organismen.*

Diese Behauptung wird mit dem vorliegenden Aufsatz untersucht - speziell durch Betrachtung der Bedürfnisse des in jeder lebenden Zelle ablaufenden, komplexen Stoffwechsels organischer Verbindungen.

Das Argument wird in 72 diskreten Schritten entwickelt, die einen Bogen von einer sehr einfachen Beobachtung zum Ziel der Erörterung schlagen. Dazu werden aus dem gegenwärtigen biologischen Wissen eine Reihe einzelner Komponenten herausgelöst und zu einer geradlinigen Argumentationslinie verknüpft.

So etwas ist nicht ganz unproblematisch: In der Biologie hängt alles an allem und manche Aussage kommt, ihrer Einbettung in den Gesamtzusammenhang beraubt, mit einem fälschlichen Anschein von Ad-hoc-Charakter daher. Um dies zumindest teilweise aufzufangen, wird mit Fußnoten eine unter der Ebene des Haupttextes liegende (ihrerseits auch lückenhafte) Sammlung von Erläuterungen, Einschränkungen, Präzisierungen etc. angeboten. Diese untermauern einzelne Punkte, sind aber für das Verfolgen des Gedankengangs entbehrlich. Der Aufsatz greift auf eine Reihe chemischer und biochemischer Grundtatsachen zurück, die mit Worten beschrieben werden. Graphische Illustrationen dazu sind bei Bedarf über *online* Nachschlagewerke sehr einfach zugänglich.

Aus direkter Anschauung glaubt sich jeder im Besitz einer soliden Vorstellung davon, was Leben ist. Geht man der Frage jedoch auf den Grund, lernt man bald, dass sich ‚Leben‘ beharrlich sträubt, definitorisch sauber von ‚Nicht-Leben‘ getrennt zu werden. Dies mahnt zu Vorsicht beim Versuch, ‚das Leben‘ mit einer anderen, begrifflich ebenfalls notorisch schwierigen Vorstellung wie ‚Information‘ in Beziehung zu setzen. Als Ausgangspunkt der zu entwickelnden Argumentationslinie wurde deshalb ein konkretes Objekt gesucht, das möglichst einfach, aber hinsichtlich seiner Qualität als etwas Lebendiges dennoch über alle Zweifel erhaben ist.

Einzellige Mikroorganismen, speziell Bakterien, erfüllen diese Voraussetzungen optimal und das den menschlichen Darm besiedelnde Bakterium *Escherichia coli* („*E. coli*“) ist von allen Organismen überhaupt im feinsten Detail verstanden.<sup>1</sup> Dadurch wird dieser bescheidene Organismus ein speziell geeignetes Objekt für die anstehende Betrachtung; die allgemeinen Schlüsse sind jedoch auf alle Lebewesen übertragbar.

## 1 Zweiundsiebzig Schritte

1. Eine einzelne *E. coli*-Zelle, in steriles Nährmedium eingebracht, erzeugt über Nacht eine makroskopisch sichtbare Trübung.
2. Die Trübung einer Flüssigkeit wie zum Beispiel Milch kommt durch Lichtstreuung an vielen, in Suspension gehaltenen, kleinen Partikeln (hier: Fetttropfchen) zustande. Unter dem Mikroskop erkennt man, dass im Falle einer Bakterienkultur die streuenden Partikel räumlich voneinander getrennte, zylindrische Zellen sind.<sup>2</sup>
3. Im Elektronenmikroskop wird deutlich: Die Zellen vermehren sich durch Längenwachstum, abgeschlossen durch eine Abschnürung in der Mitte des verlängerten Stäbchens – bis hin zur Trennung.
4. Aus Punkt (3) ergibt sich für die Population die Vorhersage eines exponentiellen Wachstumsgesetzes:  $N(n) = N_0 \cdot 2^n$ . ( $n$ : laufende Nummer der betrachteten Teilungsrunde;  $N(n)$ : Zellzahl nach jeweiliger Teilungsrunde;  $N_0$ : Zellzahl zu Beginn).
5. Eine diesem Gesetz folgende Vermehrungsrate wird in der Tat experimentell beobachtet.<sup>3</sup> Bei bestmöglicher Nährstoffversorgung<sup>4</sup> und Belüftung benötigt *E. coli* für die Vollendung eines kompletten Generationszyklus ca. 20 min (sog. ‚Verdopplungszeit‘).
6. In gut 12 h, also nach 36 Teilungsschritten, kommt man von einer zu  $2^{37} - 1 \approx 1.3 \cdot 10^{11}$  Zellen, wenn alle überleben.<sup>5</sup>

---

<sup>1</sup>Dies ist wissenschaftsgeschichtlicher Kontingenz ebenso sehr geschuldet wie einigen für molekularbiologische Untersuchungen besonders günstigen Eigenschaften von *E. coli*.

<sup>2</sup>Diese haben einen Durchmesser von ca.  $1 \mu\text{m}$  und eine Länge von ca.  $2 \mu\text{m}$  ( $1 \mu\text{m} = 10^{-3} \text{mm}$ ).

<sup>3</sup>Dies gilt solange Nährstoffkonzentration und Zelldichte in einem dies zulassenden Verhältnis stehen (exponentielle Phase der Vermehrung). Der Zuwachs an Zellzahl findet ein Ende (Beginn der ‚stationären‘ Phase), wenn die Nährstoffe zur Neige gehen oder die Zelldichte zu hoch wird.

<sup>4</sup>In der Frühzeit der Bakteriologie wurde als nahrhaftes ‚Vollmedium‘ oft Fleischbrühe zubereitet; heute nimmt man dafür eher Stoffmischungen wie eine Kombination aus Hefeextrakt und Milcheiweiß-Hydrolysat.

<sup>5</sup>Um auf die angegebene Zellzahl zu kommen, muss man die Kultur in mindestens 100 ml Nährlösung ansetzen.

7. Elektronenmikroskopische Inspektion der Zellstruktur lässt einen hohen (aber nicht kristallinen!) Ordnungsgrad erkennen. Jede Zelle ist in einen dreilagigen Sack eingehüllt: Eine robuste Zellwand von ca. 50 nm Dicke ist außen und innen von je einer dünnen Membran (jede 10 nm dick) umgeben. Im Innern der Zelle findet man als besonders auffällige Strukturen Fäden von 1 bis 2 nm Dicke und eine Vielzahl annähernd kugelförmiger Gebilde von ca. 25 nm Durchmesser.<sup>6</sup>
8. Im Kulturmedium enthaltene Substanzen sind offenbar in hoch organisierte Zellmasse überführt worden. Dieser Stoffumwandlung müssen chemische Prozesse zugrundeliegen.
9. Eine mit chemischen Stoffumwandlungen einhergehende Erhöhung der Ordnung eines Stoffgemenges („Systems“) ist ungewöhnlich.<sup>7</sup>
10. Alle aus der Vermehrung hervorgegangenen Zellen sehen gleich aus (untereinander und gleich wie die Ursprungszelle der Kultur). Dies gilt sowohl für ihre äußere Gestalt als auch für die nur elektronenmikroskopisch sichtbar zu machenden Teilstrukturen.
11. Die Uniformität aller Zellen legt nahe, dass die Ordnung schaffenden Mechanismen in allen Zellen identisch sind und von einer Zellgeneration zur nächsten stabil weitergegeben werden. Dies ist eine zusätzliche, eigene Manifestation von Ordnung.
12. Aus den Punkten (8), (9) und (11) ergibt sich Erklärungsbedarf für das Wesen der Stoffumwandlung an sich sowie für das mit ihr einhergehende Entstehen von Ordnung (einschließlich der genetischen Weitergabe der Ordnung schaffenden Mechanismen).
13. Um diesem Problem näherzukommen, betrachten wir zunächst Punkt (8) unter dem Aspekt des Gesetzes von der Erhaltung der Masse.<sup>8</sup>
14. Dazu vergleichen wir die Ausgangsstoffe der zellulären Chemie, d. h. die Substanzen, aus denen sich das Nährmedium zusammensetzt, mit den Produkten, d. h. den chemischen Verbindungen, aus denen die neu synthetisierte

---

<sup>6</sup>1 nm =  $10^{-6}$  mm.

<sup>7</sup>Von dieser Aussage nicht betroffen ist die manchmal zu beobachtende Kristallisation von Reaktionsprodukten. Die gehört aber in eine andere, nicht-dynamische Kategorie von Ordnung und ist für die gegenwärtige Betrachtung nicht relevant (siehe auch Punkt 30).

<sup>8</sup>Dieses Naturgesetz sagt aus, dass die Masse aller Stoffe, die eine chemische Reaktion eingehen, gleich ist der Masse aller Reaktionsprodukte. Es hat seine Wurzel darin, dass die Atomkerne chemisch reagierender Substanzen nicht angetastet, sondern lediglich die Bindungen zwischen Atomen umgruppiert werden. Dies spiegelt sich z. B. darin wider, dass in einer chemischen Gleichung links und rechts vom Reaktionspfeil immer gleich viele Atomsymbole eines jeden beteiligten Elements stehen.

Biomasse besteht.<sup>9</sup> Um mit der einen Seite des Vergleichs einen einfachen Referenzpunkt zu gewinnen, führen wir das Vermehrungsexperiment statt mit Vollmedium mit sogenanntem Minimalmedium durch.<sup>10</sup>

15. Die chemische Zusammensetzung eines Minimalmediums ist sehr einfach (siehe Tabelle 1).

NH <sub>4</sub> Cl	1.0 g
MgSO <sub>4</sub>	0.13 g
KH <sub>2</sub> PO <sub>4</sub>	3.0 g
Na <sub>2</sub> HPO <sub>4</sub>	6.0 g
Glucose*	4.0 g
Wasser	ad 1000 ml

Tabelle 1. Bakteriologisches Minimalmedium (für E. coli).

\*Glucose („Traubenzucker“) oder andere Zucker können bei einigen Mikroorganismen durch spezielle Kombinationen einfachster anorganischer Verbindungen ersetzt werden – z. B. durch eine Mischung von Kohlenstoffdioxid und molekularem Wasserstoff.

16. Für die chemische Charakterisierung der mit dem Minimalmedium zu vergleichenden Biomasse ernten wir die Zellen durch Zentrifugation, brechen sie anschließend auf<sup>11</sup> und trennen die Bestandteile der resultierenden Brühe zweckmäßigerweise zunächst durch eine zweite Runde von Zentrifugation in einen partikulären und einen löslichen Teil, anschließend den letzteren (das ‚geklärte Lysat‘) durch Dialyse<sup>12</sup> in eine nieder- und eine hochmolekulare Fraktion.
17. Eine Bestandsaufnahme (zunächst) der niedermolekularen Fraktion (sie ist technisch am einfachsten) zeigt, dass am Stoffwechselgeschehen in der Größenordnung von eintausend verschiedene niedermolekulare Verbindungen beteiligt sind. Diese bestehen aus organischen Verbindungen, deren Kohlenstoffgerüste zum Teil recht komplex gebaut sind<sup>13</sup> (Siehe auch Punkt 25).

<sup>9</sup>Für eine *vollständige* Stoffbilanz (auf die hier ohne Verlust für das Argument verzichtet werden kann), ist dabei auch Gasaustausch mit der Umgebung zu beachten, ebenso Stoffe im Medium, die bei Versuchsbeginn dort noch nicht anzutreffen waren (d. h. Ausscheidungsprodukte der Bakterien).

<sup>10</sup>Derartige ‚Diätahrung‘ verlängert die Verdopplungszeit um einiges – sonst aber bleibt alles wie gehabt.

<sup>11</sup>Zu diesem Zweck kann man zum Beispiel eine wässrige Aufschlämmung intakter Zellen unter hohem Druck durch eine sehr feine Düse pressen. Die bei der Passage durch die Düse eintretende, plötzliche Entspannung lässt die Zellen platzen.

<sup>12</sup>Dialyse ist die Entfernung niedermolekularer Stoffe aus einer wässrigen Lösung durch Passage durch eine Membran, die hochmolekulare Stoffe zurückhält. Die technisch einfachste Handhabung besteht im Einschluss der zu dialysierenden Probe in einem an beiden Enden zugeknoteten Stück Cellophan Schlauch, das man anschließend für mehrere Stunden in einem Becherglas mit gepuffertem Wasser belässt.

<sup>13</sup>Dazu gehören Endprodukte des niedermolekularen Stoffwechsels wie Aminosäuren, Nucleotide, Lipide, Coenzyme u. a. m. sowie deren (oft zahlreiche) synthetische Vorstufen.

18. Einige davon lassen sich bereits ohne spezielle Kenntnisse von Biochemie auf plausible Weise zu Reaktionsketten arrangieren; für die Aufklärung des größten Teils des metabolischen Reaktionsnetzwerks (das aus etwa gleich viel Reaktionen wie beteiligten Verbindungen besteht) waren und sind große Forschungsanstrengungen nötig.
19. Diese Verbindungen des Stoffwechsels sind nach Maßstäben Menschen-gemachter Laborchemie generell nicht sehr reaktionsfreudig,<sup>14</sup> zumal bei den in der Zelle vorherrschenden Bedingungen (wässriges Milieu, 37°C) und in ihren Reaktionsmöglichkeiten nicht besonders selektiv (heißt: Sie können alternativ mehrere verschiedene Reaktionen eingehen – zumal auch mit unterschiedlichen Reaktionspartnern).
20. Viele der Verbindungen besitzen zudem *mehrere* (schwach bis moderat) reaktive Atomgruppierungen, was zur Erhöhung der Anzahl alternativer Reaktionsmöglichkeiten auf unübersichtliche Weise beiträgt.
21. Sich selbst überlassen, würden die niedermolekularen chemischen Bestandteile der Zelle ein System träger, ungeordnet verlaufender Reaktionen bilden, noch wesentlich mehr als die tatsächlich beobachteten verschiedene Verbindungen bilden und dabei sehr langsam dem Gleichgewichtszustand entgegendümpeln („chemischer Morast“).
22. Dies steht in krassem Gegensatz zu den tatsächlich vorgefundenen Verhältnissen (nächste vier Punkte).
23. Aus energiearmen Ausgangsmaterialien wie anorganischem Phosphat werden energiereiche Verbindungen aufgebaut, wie – um im Beispiel zu bleiben – solche, die Phosphorsäureanhydridbindungen enthalten (Adenosin-triphosphat u. A. m.).
24. In derselben Zelle laufen reduktive und oxidative Prozesse nebeneinander ab: So wird zum Beispiel das im Medium vorgehaltene Sulfat (+VI) zum Teil zu Sulfid (-II) reduziert;<sup>15</sup> bei der gleich daneben stattfindenden Atmung wird der Kohlenstoff der Glukose zum Teil zu Kohlenstoffdioxid oxidiert.
25. Viele der vorgefundenen Verbindungen sind von sehr subtiler molekularer Architektur, z. B. besitzen sie häufig chirale Kohlenstoffzentren – und dies praktisch immer in isomerenreiner Form.<sup>16</sup>

---

<sup>14</sup>Viele der Verbindungen sind so reaktionsträge, dass man sie als praktisch inert gegenüber allen anderen Verbindungen der Zelle einordnen würde.

<sup>15</sup>Wie z. B. in der Mercaptogruppe der Aminosäure Cystein: HS-CH<sub>2</sub>-CH(NH<sub>2</sub>)-COOH.

<sup>16</sup>Mit ‚Chiralität‘ wird die Tatsache bezeichnet, dass es Verbindungen gibt, die in zwei subtil unterschiedlichen Formen auftreten. Deren Atome sind zwar auf dieselbe Weise verknüpft, verhalten sich zueinander aber wie Bild und Spiegelbild (linke und rechte Hand – daher die Bezeichnung ‚Chiralität‘). Wenn solche Verbindungen im Labor synthetisiert und dabei keine sehr speziellen Vorkehrungen getroffen werden, entstehen die beiden ‚isomeren‘ Formen gleichzeitig und im Verhältnis 1:1. ‚Isomerenrein‘ heißt eine Substanz, die Chiralität aufweist, aber nur aus der einen der beiden Formen besteht.

26. Eine Zelle schafft also, aus ein paar einfachen Chemikalien<sup>17</sup> eine breite Palette verschiedenster und nach Maßstäben menschlicher Chemie zum Teil schwer zugänglicher Verbindungen zu synthetisieren – und all dies mit einer Geschwindigkeit, die in weniger als einer Stunde ein komplettes Abbild ihrer selbst – „mit Haut und Haaren“ – entstehen lässt (vergleiche Punkte 5, 10 und 15).
27. Das Bild vom „chemischen Morast“ (vergleiche Punkt 21) könnte offenbar falscher nicht sein. Vielmehr ist die chemische Dynamik einer in Proliferation befindlichen Bakterienkultur nur so zu verstehen, dass der zelluläre Stoffwechsel ein Netzwerk schneller und sinnvoll ineinandergreifender Reaktionen darstellt, innerhalb dessen die Gleichgewichtseinstellung verhindert wird.
28. Dies verlangt die Erfüllung von zwei Voraussetzungen:
- A Beständige *Zufuhr von Energie*,<sup>18</sup> für die es folglich eine stetig sprudelnde Quelle geben muss.
  - B Drastische *Reaktionsbeschleunigung* – mithin die Existenz und Wirkung hochselektiver und äußerst effizienter Katalysatoren.<sup>19</sup>
29. Diese Katalysatoren sorgen für *Schnelligkeit und Ordnung*. Ordnung – der ohne Vorkenntnisse weniger offensichtliche Aspekt – kommt dadurch zustande, dass für die Katalyse jeweils eine einzige, ganz bestimmte Reaktion aus

<sup>17</sup>Siehe Punkt 15 und Tabelle 1: ‚Minimalmedium‘. Glucose ist ihrerseits eine organische Verbindung und Stoffwechselprodukt anderer Organismen (z. B. Pflanzen). Darauf (oder ähnliches) ist *Escherichia coli* zwingend angewiesen. Es gibt aber auch Mikroorganismen, die ihren kompletten Satz organischer Verbindungen ausgehend von Kohlenstoffdioxid oder einer anderen C-1 Verbindung aufbauen können.

<sup>18</sup>Genauer: Fluss von Energie durch das offene System Zelle, begleitet von Energiedissipation, d. h. Umwandlung arbeitsfähiger Energie in (letztlich) Wärme. In Gegenwart molekularer Sauerstoffs lebende Mikroorganismen nutzen als Energiequelle die Atmung, d. h. die Oxidation von Glucose (o.ä.) zu Kohlenstoffdioxid und Wasser. Die dabei gewonnene Energie wird zum Teil vorübergehend in chemische Energie von Stoffwechselprodukten, Aufbau von Konzentrationsgradienten und gezielte Bewegung (der ganzen Zelle oder von Zellbestandteilen relativ zueinander) „re-investiert“ zum anderen Teil direkt als Wärme in die Umgebung abgeführt. Dies ist einer Bakterienkultur auch anzumerken: Führt man das Vermehrungsexperiment in einem Kalorimeter durch, bemerkt man, dass Wärme freigesetzt wird. Auch wir halten unsere Körpertemperatur von 37°C mit Hilfe der metabolischen Abwärme aufrecht. „Zieht man den Stecker“, zerfällt die dynamische, energiegetriebene („dissipative“) Ordnung und das System strebt dem chemischen Gleichgewicht entgegen. (Der letzte Satz ist eine Beschreibung des Todes in chemischen Begriffen.)

<sup>19</sup>Ein Katalysator ist ein Stoff, der eine chemische Reaktion beschleunigt ohne selbst verändert zu werden (*bleibend* verändert zu werden, müsste man genauer sagen). Ein Katalysator tritt mit den Teilnehmern der Reaktion in eine transiente Wechselwirkung. Diese gibt den elektronischen Umgruppierungsprozessen, aus denen eine chemische Reaktion besteht, einen ohne Katalysator schwer zugänglichen Verlauf, der sich dadurch auszeichnet, dass Reaktionshindernisse umgangen oder verkleinert werden. (In Fachterminologie: Die Aktivierungsenergie der Reaktion wird abgesenkt.) Unter der Selektivität eines Katalysators versteht man die Trennschärfe, mit der er eine bestimmte Reaktion über andere bevorzugt und unter Effizienz den Faktor, um den er diese eine Reaktion beschleunigt.



einer größeren Anzahl möglicher Reaktionen ausgewählt wird. Die nicht beschleunigten verharren auf dem generell niedrigen Spontanniveau (vgl. Punkte 5, 19 und 26) und spielen für die chemische Dynamik des gesamten Stoffwechselgeschehens keine Rolle.<sup>20</sup>

30. Der Rest des Arguments beschränkt sich auf die Diskussion der Katalyse (Punkt 28B). Die für Entstehen und Aufrechterhalten dynamischer Ordnung ebenso wichtige Energiedissipation (Punkt 28A)<sup>21</sup> ist für das hier betrachtete Problem nicht direkt relevant und wird aus der weiteren Betrachtung ausgeklammert.
31. Einem biochemischen Katalysator muss man – vgl. Punkt 29 – also zwei Fähigkeiten zuweisen:
  - A) Er muss von mehreren alternativen Reaktionen, die ein Zwischenprodukt des Stoffwechsels eingehen kann (oder zwei solcher Zwischenprodukte miteinander) spezifisch eine einzige beschleunigen – und die um einen sehr großen Faktor.
  - B) Dazu muss er zunächst dieses Zwischenprodukt (sein ‚Substrat‘) mit hohem Unterscheidungsvermögen aus dem niedermolekularen Substanzenrepertoire der Zelle, auswählen (siehe auch Punkt 28B, Fußnote 19).
32. Beide Fähigkeiten müssen in der Struktur<sup>22</sup> des Katalysators verankert sein (passgenaue Anlagerung des Substrats, transiente Herstellung die Reaktion befördernder geometrischer und elektronischer Verhältnisse im Substrat).
33. Eine Verbindung, deren biochemische Reaktion zu einer anderen Verbindung beschleunigt werden soll, verhält sich offenbar zum Katalysator dieser Reaktion wie ein Werkstück zu einer dieses bearbeitenden Maschine.
34. Die in den Punkten 31 bis 33 zusammengefassten Überlegungen legen nahe, dass biochemische Katalysatoren wesentlich größer sind als ihre Substrate,

---

<sup>20</sup>In seltenen Ausnahmefällen spielen nicht-katalysierte Nebenreaktionen doch eine Rolle – entweder weil sie ungewöhnlich hohe Basisgeschwindigkeit haben oder weil ihre Produkte besonders wirkmächtig sind. Beispiele dafür sind die Entstehung von toxischem Methylglyoxal als Nebenprodukt des Zuckerstoffwechsels und die Methylierung von Guanosinresten in der DNA zu O<sup>6</sup>-Methylguanotin (Vorstufe einer Mutation). In beiden (und ähnlichen anderen) Fällen werden die unvermeidlichen schädlichen Produkte von speziellen Enzymen abgefangen und durch chemische Umsetzung entgiftet.

<sup>21</sup>Näheres in Fußnote 18. Vergleiche auch Punkt 9, Fußnote 7.

<sup>22</sup>Unter „Struktur“ ist hier nicht nur die dreidimensionale geometrische Gestalt des Moleküls zu verstehen; der chemische Begriff „Struktur“ umfasst auch die genaue Lokalisierung von Atomgruppen mit elektrischen Ladungen und anderen speziellen Eigenschaften auf dem durch die Molekülgeometrie vorgegebenen Gerüst.

also Makromoleküle:<sup>23</sup> jeder ein individuell aufgabenspezifisch geformtes Makromolekül – strukturstabil, wenn auch nicht völlig starr. Siehe auch Punkt 38.

35. Nach den biochemischen Katalysatoren („Enzymen“)<sup>24</sup> muss man also in der makromolekularen Zellfraktion (i. e. im Inhalt des Dialysebeutels – vgl. Punkt 16) suchen.<sup>25</sup> Dazu dienen zwei Werkzeuge:

(i) Ein Aktivitätstest („Assay“), der die Wirkung des Enzyms in ein makroskopisches Signal umsetzt.<sup>26</sup>

(ii) Fraktionierung der makromolekularen Zellbestandteile (hauptsächlich per Säulenchromatographie)<sup>27</sup> nach unterschiedlichen Kriterien. Nach jeder Runde von Fraktionierung wird mittels des Assays geprüft, in welcher Fraktion sich das Enzym befindet. Diese wird einer weiteren Fraktionierung (nach anderen Trennkriterien) unterworfen, die anderen verworfen.

36. Am Ende einer solchen mehrstufigen Prozedur steht ein Enzym als chemisch einheitlicher Stoff. Die Mengen, die man auf diese Weise von einem reinen Enzym gewinnen kann, sind generell sehr klein.<sup>28</sup>

37. Auf jedes molekular einheitliche Enzym kann man, wie zuvor bei den niedermolekularen Zellbestandteilen, die Methoden chemischer Analyse anwenden.

---

<sup>23</sup>Als günstiger Nebeneffekt kommt hinzu: Die nicht-kovalent verknüpften Aggregate aus Substrat und Katalysator sind einem Dauerbombardement durch die in thermischer Bewegung befindlichen Wassermoleküle der Umgebung ausgesetzt. Die so empfangenen Stöße sind einem geordneten Ablauf der Katalyse abträglich – durch die größere Masse des Katalysators werden sie gedämpft.

<sup>24</sup>Eine genaue Definition des Begriffs „Enzym“ wird erst mit Punkt 45 geliefert. Der Einfachheit halber wird er aber schon hier als vorläufiges Synonym für „biochemischer Katalysator“ eingeführt.

<sup>25</sup>Streng genommen gilt dies nur für homogene Katalyse, das heißt für solche, die durch molekular dispers in Lösung befindliche Katalysatoren bewirkt wird, wie dies für den weit überwiegenden Teil der biochemischen Katalyse der Fall ist. Träger heterogener Katalyse sind in der partikulären (durch Zentrifugation abgetrennten) Zellfraktion (siehe Punkt 16) zu suchen.

<sup>26</sup>Ein Enzym, zum Beispiel, das auf die Hydrolyse von Carbonsäureestern in schwach alkalischem Milieu spezialisiert ist (eine „Esterase“), lässt sich mit folgendem einfachen Assay nachweisen: Man stellt die Lösung, die man auf Anwesenheit des Enzyms untersuchen will, auf pH 8 ein und setzt eine kleine Menge der (farblosen) Lösung eines Carbonsäure-4-nitrophenylesters zu. Bei Anwesenheit der Esterase wird sehr schnell 4-Nitrophenol freigesetzt, das bei pH 8 zum überwiegenden Teil als intensiv gelbes 4-Nitrophenolat anion vorliegt.

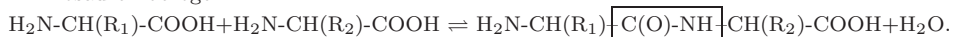
<sup>27</sup>Säulenchromatographie: In ein Rohr aus Glas, Kunststoff oder Metall wird ein jeweils spezieller, fein pulverisierter Feststoff („Matrix“) gefüllt. In dieses Rohr wird die in Wasser gelöste Substanzmischung gepumpt und danach mit einem stetigen Strom gepufferter Salzlösung wieder aus ihm ausgespült. Die Komponenten der Stoffmischung binden unterschiedlich locker (nicht-kovalent) an die Matrix und bewegen sich deshalb verschieden schnell durch die Feststoffpackung (die „Säule“) und kommen so zu verschiedenen Zeiten an deren Austrittsseite. Der austretende Puffer, samt darin gelösten Substanzen, wird portionsweise aufgefangen.

<sup>28</sup>Dies liegt an der hohen Effizienz von Enzymen (vgl. Punkt 28), von denen jedes individuelle Molekül pro Sekunde Zigtausende oder mehr Substratmoleküle umsetzen kann. Entsprechend braucht es in vielen Fällen nur einige Dutzend Enzymmoleküle einer bestimmten Sorte in jeder Zelle.

38. Dabei stellt sich (u. A.) heraus: Enzyme haben relative Molmassen von typischerweise 10 000 bis 100 000 (d. h. sie sind ca. 100 bis 1000-fach größer als typische kleine Moleküle der Zelle).
39. In Übereinstimmung mit den Forderungen von Punkt 31 ist jedes Enzym für die Beschleunigung je einer der besagten rund eintausend Reaktionen der Chemie der Zelle zuständig. Die Beschleunigungsfaktoren sind enorm: in Extremfällen bis  $10^{17}$ !
40. Mit dieser Feststellung sind – im Sinne von Punkt 29 – zwei Probleme (im Prinzip) gelöst:
- Man kann einsehen, dass eine Zelle sich innerhalb 20 min verdoppeln kann.
  - Die Chemie der Zelle ist – trotz hoher Komplexität – als ein sinnvolles Gefüge vorstellbar.
41. Im Gegenzug hat sich ein neues Problem aufgetan: Jetzt hat man es damit zu tun, die Existenz von ca. eintausend verschiedenen *Makromolekülen* zu erklären, die alle auch synthetisiert sein wollen.
42. Dieses neue Problem sieht auf den ersten Blick noch schwieriger aus als das alte. Auf der Suche nach einer Lösung lohnt sich ein Blick auf die chemische Natur der Enzyme.
43. Jedes Enzym liefert bei seiner Totalhydrolyse (Molekülspaltung durch Wasser) 20 verschiedene Aminosäuren vom selben Strukturtyp (siehe Abbildung 1) – immer dieselben 20, aber in individuell verschiedenen Mengenverhältnissen.
44. Die Hydrolyse ist die Umkehrung einer Kondensationsreaktion (Verschmelzung von zwei Molekülen zu einem unter Wasseraustritt.<sup>29</sup> Also kann man sich das Enzym-Polymer vorstellen als aus Aminosäuren durch wiederholte Kondensationsreaktion (‘Polykondensation’) hervorgegangen. Daraus und aus der allgemeinen Struktur der Aminosäuren folgt die Annahme, die Enzyme wären lineare Fadenmoleküle, innerhalb welcher einzelne Aminosäurereste durch Säureamidbindungen zusammengehalten werden.

---

<sup>29</sup>Kondensation und Hydrolyse in chemischer Notation:  $X-H + HO-Y \rightleftharpoons X-Y + H_2O$ , bzw. auf Aminosäuren bezogen:



Die charakteristische Säureamidgruppierung ist eingerahmt.

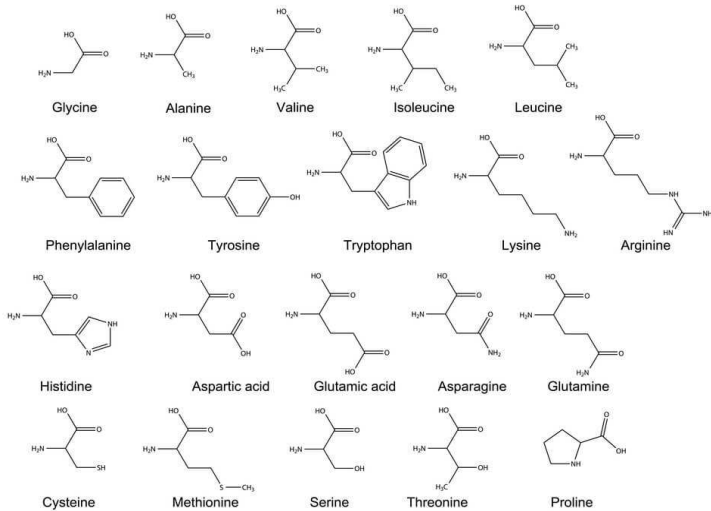


Abbildung 1. Die zwanzig Aminosäuren, aus denen Enzyme aufgebaut sind.

45. So wird das auch gefunden und die im vorigen Punkt skizzierte Strukturbeschreibung trifft allgemein auf die Substanzklasse der Proteine zu, die auch Vertreter mit anderen Funktionen als katalytische Reaktionsbeschleunigung umfasst. Proteinische Biokatalysatoren werden als Enzyme<sup>30</sup> bezeichnet.
46. Weiterer Befund der Analyse: Jedes Enzym ist einheitlich in Bezug auf seine Aminosäuresequenz;<sup>31</sup> untereinander unterscheiden sie sich in genau diesem Punkt.
47. Von einem Molekülfaden, der ausschließlich durch Einfachbindungen zusammengehalten wird, erwartet man – wegen der weitgehend ungehinderten Drehung der beiden durch die jeweilige Einfachbindung verknüpften Kettenteile relativ zueinander – hohe (,konformationelle‘) Kettenflexibilität.<sup>32</sup> Mit dieser Vorstellung steht die Forderung einer präzisen, stabil eingenommenen Raumstruktur (Punkt 34) in (scheinbarem) Widerspruch.

<sup>30</sup>In modernen Organismen sind die biochemischen Katalysatoren fast durchweg Enzyme (das heißt Proteine). Ein unbewiesenes aber plausibles Szenario (RNA-Welt) schlägt vor, dass in einer frühen Phase der Evolution Ribonukleinsäuren (RNA) statt Proteine Träger der biochemischen Katalyse waren – wie heute (noch?) in einzelnen Ausnahmefällen.

<sup>31</sup>In diese Aussage eingeschlossen ist die, dass alle Ketten eines bestimmten Enzyms gleiche Länge (das heißt gleiche Anzahl von miteinander verknüpften Aminosäureresten) haben.

<sup>32</sup>Das Ergebnis wäre eine zwar nach Kettenlänge und Sequenz (,konstitutionell‘) einheitliche Molekülpopulation, in der jedoch zu jedem Zeitpunkt jedes einzelne Molekül einen von den anderen Molekülen abweichenden dreidimensionalen Kettenverlauf (,Konformation‘) aufweist und diesen zudem mit der Zeit dauernd verändert. Den so beschriebenen, dynamischen Strukturzustand nennt man ‚Zufallsknäuel‘. Enzyme sind keine Zufallsknäuel.

48. Die postulierte, spezifische Raumstruktur eines Enzyms lässt sich jedoch – „trotz allem“ – experimentell nachweisen (am überzeugendsten durch Kristallisation und Röntgenkristallographie).
49. Zusätzlich zu seiner Sequenz (Punkt 46) ist ein Enzym also auch in Bezug auf seine Konformation einheitlich – in anderen Worten: Alle Ketten sind in derselben Weise dreidimensional gefaltet.<sup>33</sup>
50. Weiterer experimenteller Befund: Der spezifische Faltungszustand lässt sich durch ‚denaturierende‘ Agenzien (Harnstoff, Guanidiniumchlorid etc.) aufheben; die Kette geht dann in den Zustand eines Zufallsknäuels (siehe Fußnote 32) über, der keine katalytische Aktivität besitzt.
51. Durch Entzug des Denaturierungsmittels lässt sich der native (d. h. spezifisch gefaltete, katalytisch aktive) Zustand wiederherstellen.
52. Daraus muss man schließen, dass die Raumstruktur eines Proteins durch seine Sequenz eindeutig festgelegt ist<sup>34</sup> und ihm nicht etwa von irgendwelchen äußeren formgebenden Vorrichtungen aufgezwungen wird – während oder nach seiner Synthese.
53. Damit erweist sich die Konformationshomogenität als eine direkte Folge der Sequenzhomogenität. Letztere ist aber nur eine notwendige, keine hinreichende Voraussetzung für erstere. Wir haben allen Grund zu der Annahme, dass nur ein sehr kleiner Teil aller denkbaren Ketten über die Fähigkeit verfügt, einheitliche Konformation einzunehmen. Solche Ausnahmen zu finden und biochemisch nutzbar zu machen, ist eine Leistung der (molekularen) Evolution.
54. Die Konformationshomogenität ist die Wurzel der spezifischen katalytischen Aktivität: Sie verleiht Enzymen zum Beispiel die Fähigkeit, zwei miteinander zur Reaktion zu bewegende Moleküle in immer derselben Weise nebeneinander spezifisch und in einer für die Reaktion günstigen geometrischen Anordnung zu binden, den Reaktionsort (z. B.) von störendem Wasser abzusichern, und hilfreiche, auf dem Proteinfaden liegende Atomgruppierungen am Ort des Reaktionsgeschehens zur Verfügung zu stellen, so z. B. eine basische Gruppe oder ein komplexiertes Metallion.<sup>35</sup>

---

<sup>33</sup>Die Röntgenkristallographie kann von einer Substanz nur solche Strukturdetails abbilden, die an jedem seiner (sehr vielen) Gitterplätze im streuenden Kristall gleich ausgeprägt sind. Damit ist das Sichtbarmachen des Kettenverlaufs mit dieser Methode ein Beweis für die Konformationshomogenität des jeweiligen Proteins (Enzyms).

<sup>34</sup>Kein kleines Wunder für jeden Chemiker und als „Proteinfaltungsproblem“ eine derjenigen Fragen der Molekularbiologie, die sich – trotz unbestreitbarer, in jüngerer Zeit erzielter Fortschritte – besonders hartnäckig einer Lösung widersetzen.

<sup>35</sup>Das ist der Hintergrund für die Erfüllung der Postulate von Punkt 31.

55. *Zwischenfazit:* Für die Funktion eines Enzyms braucht dieses eine bestimmte Struktur, heißt: eine bestimmte, für alle individuellen Moleküle dieses Enzyms gleiche Kettenkonformation. Diese ist alleine durch die Aminosäuresequenz eindeutig festgelegt – es ist keine aktive Formgebung ‚von außen‘ nötig. Damit reduziert sich das Problem, einen bestimmten Katalysator herzustellen (Punkt 41) darauf, ein Protein der jeweiligen Sequenz zu synthetisieren.
56. Das aber ist schwer genug und dazu braucht es wiederum Katalyse – aus genau den gleichen Gründen, die für die Synthesen der niedermolekularen Zellinhaltsstoffe angeführt wurden (Punkt 29) – jetzt angewandt auf die 20 Aminosäuren und ihre Fähigkeit zur Polykondensation.
57. Man könnte zunächst versucht sein zu denken, die Katalyse der Synthese der Enzyme könne nach dem gleichen Prinzip gelöst sein, wie die Katalyse im niedermolekularen Stoffwechsel: Für jede einzelne Reaktion gäbe wiederum einen – und nur speziell für diese – zuständigen Katalysator.
58. Wie viele Katalysatoren würde man für die Synthese eines Enzyms brauchen? Es sind zwei unterschiedliche Arten des Aufbaus aus den Aminosäuren zu unterscheiden: sequentieller und hierarchischer Prozess.
59. Bei der sequentiellen Synthese wird an die wachsende Kette pro Schritt je ein Rest angefügt. Die Reaktionspartner in jedem einzelnen Verlängerungsschritt sind jeweils die teilfertige Kette auf der einen Seite und eine der zwanzig Aminosäuren auf der anderen: Für die Synthese einer Kette der Länge  $n$  braucht man also  $n - 1$  verschiedene Katalysatoren.<sup>36</sup>
60. Für die Synthese von eintausend verschiedenen Enzymen einer durchschnittlichen Kettenlänge von 500 Aminosäuren würden demnach 499 000 verschiedene Katalysatoren benötigt.<sup>37</sup>
61. An dieser Zahl ändert sich auch bei *hierarchischer* Synthese nichts wesentliches, bei der kleine Kettenabschnitte unabhängig vorgefertigt und danach sequenzspezifisch miteinander verknüpft werden.

---

<sup>36</sup>Die müssten deshalb alle verschieden sein, da mindestens *ein* Reaktionspartner – die wachsende Polypeptidkette – immer verschieden wäre.

<sup>37</sup>Die notwendige Gesamtzahl dieser (hypothetischen) Enzyme würde sich reduzieren, wenn bestimmte Kombinationen von wachsender Kette als erstes Substrat und neu anzuknüpfender Aminosäure als zweites in den Synthesen verschiedener Enzyme in gleicher Weise vorkämen. Die Einsparung wäre jedoch marginal, weil bei 20 verschiedenen Aminosäuren die Anzahl verschiedener Sequenzen sehr rasch mit der Kettenlänge ansteigt ( $20^n$ ,  $n$  = Anzahl Kettenglieder) und die Wahrscheinlichkeit der ‚Mehrfachverwendung‘ teilfertiger Ketten entsprechend gering ist und für längere Ketten rasch gegen Null geht. Für die Synthese aller aus drei Aminosäuren bestehenden Module bräuchte man bereits 8.000 verschiedene Enzyme: Es müssen dazu nämlich 400 verschieden Zweiermodule mit 20 verschiedenen Aminosäuren verknüpft werden. Für alle verschieden Vierermodule ist man bereits bei 160.000 Enzymen etc. Dies macht den Ansatz komplett unrealistisch.

62. Unausweichlich benötigt man für den Aufbau der Enzyme, die den niedermolekularen Stoffwechsel kontrollieren, wesentlich mehr Katalysatoren, als die herzustellenden Enzyme selbst – und an alle sind die gleichen Kriterien anzulegen wie an die Enzyme ihrerseits (Spezifität hinsichtlich Substraten und Reaktionstyp, Größe, Aufbau durch sequenzspezifische Polykondensation etc.).
63. *Schlussfolgerung*: Dieser Erklärungsversuch stürzt offenbar in einen unendlichen Regress ab – unter Explosion der Anzahl erforderlicher Katalysatoren – und führt sich somit selbst *ad absurdum*.
64. Somit gibt also die von den Enzymen bewirkte Katalyse mit ihrer unter Punkt 31 skizzierten, doppelten Spezifität hinsichtlich Substrat und Reaktion kein Modell ab für die Katalyse der chemischen Schritte, die zu ihrer eigenen Synthese führen.
65. Der Ausweg aus dem Dilemma, den gefunden zu haben, eben die im Titel genannte ‚Blockbuster-Erfindung der Evolution‘ darstellt, ist ein *allgemeiner, programmierbarer* Syntheseapparat mit der im folgenden beschriebenen Zerteilung der Aufgabe.
66. Mit einem solchen Syntheseapparat spielt die Natur der bereits synthetisierten Teilkette und die der gerade anzufügenden Aminosäure keine Rolle für die Auslösung der Kettenverlängerungsreaktion; das heißt, der dafür zuständige Katalysator braucht keine zwischen verschiedenen Teilketten und verschiedenen Aminosäuren differenzierende Substratspezifität – *und darf keine haben*.
67. Das Problem der notwendigen „Individualisierbarkeit“ des Gesamtprozesses, das heißt des Aufbaus von Ketten spezifischer Länge und spezifischer Sequenz, wird durch Zufütterung von Information gelöst – wie bei einer numerisch gesteuerten Werkzeugmaschine.
68. Die Natur hat dies – nach dem Schema der sequentiellen Synthese (Punkt 59) – eingerichtet: Der Katalysator hält die wachsende Kette durch den gesamten Kettenaufbauprozess hindurch dauerhaft fest und fügt ein Kettenglied nach dem anderen an: Siehe Abbildung 2.
69. Der chemischen Umsetzung jeweils vorgelagert, wird eine Instruktion folgender Natur in den Prozess eingespeist: Wähle für den jetzt gerade anstehenden Verlängerungsschritt aus den 20 Aminosäuren diejenige mit der Identität  $x$  aus und führe sie dem katalysierten Bindungsschluss mit dem wachsenden Kettenende zu – siehe Abbildung 2.





70. Diese Instruktionen müssen auf einem Datenträger abgelegt und von dort in geordneter Weise abrufbar sein. Der als Ergebnis des vorgetragenen Arguments zu postulierende biologische Datenträger ist die DNA und der Prozess des Abrufens der gespeicherten Information das Umschreiben von DNA-Teilsequenzen in RNA. Für weitere Details wird auf den Aufsatz ‚Informationstheoretische Metaphern im Vokabular der Molekularen Genetik‘ in diesem Band verwiesen.
71. Informationsgesteuerte Proteinsynthese ist ein komplexer Vorgang und ohne Zweifel das Ergebnis eines längeren Evolutionsprozesses. Zur Zeit des Übergangs eines Teils der Geochemie in Biochemie (heißt: vor ca. vier Milliarden Jahren) stand informationsgesteuerte Synthese nicht zur Verfügung und damit keine Vielzahl effizienter Katalysatoren, wie sie notwendig ist, um ein ausgedehntes Netzwerk chemischer Reaktionen in aufeinander abgestimmter Form und mit hohen Umsatzraten zu betreiben.
72. Damit war die „Prä-Biochemie“ auf einem vergleichsweise primitiven Niveau blockiert und die „Erfindung“ der informationsgesteuerten Synthese kann man daher mit einiger Berechtigung als die Öffnung des Tors zur Entfaltung des Lebens betrachten.

## 2 Kurzfassung

- Leben, so wie wir es kennen, benötigt geordneten Stoffwechsel.
- Dies benötigt effiziente Katalyse.
- Diese benötigt Makromoleküle als Träger der Aktivität.
- Die Makromoleküle müssen ihrerseits synthetisiert werden.
- Diese Synthese benötigt ihrerseits Katalyse (und mithin Katalysatoren).
- Es droht Absturz in unendlichen Regress.
- Dies ist nur zu vermeiden durch Trennung der Aufgabe in zwei separate Aspekte:<sup>38</sup>
  - (i) Immer gleiche Chemie beim schrittweisen Aufbau der makromolekularen Katalysatoren aus den Monomeren.

---

<sup>38</sup>Mit den entsprechenden präzisierenden Bedingungen für die Natur der Polymeren: Zum Beispiel muss die Synthese auf einen Satz von Monomeren zurückgreifen, die sich chemisch voneinander unterscheiden (nur so sind Moleküle mit individuellen Eigenschaften aufzubauen), die sich hinsichtlich ihrer Verknüpfungsschemie jedoch identisch verhalten. Folgendes zur Vermeidung eines Missverständnisses: Oben mag es so geklungen haben, als habe sich die informationsgesteuerte Synthese um die chemische Natur der Proteine herum evolviert. In Wirklichkeit ist es umgekehrt: Die Natur der Proteine, bzw. der Aminosäuren, ergab sich – nicht im chemischen Detail, aber in grundlegenden Aspekten – aus den Ansprüchen einer notwendig informationsgesteuerten Synthese.

- (ii) Festlegung der individuellen Reihenfolge, in der die verschiedenen Monomere zum Faden verknüpft werden.
- Die eigentliche Katalyse kann sich dann auf die sich zyklisch wiederholende, immer gleiche Teilaufgabe (i. e. die der Verknüpfung – ohne Rücksicht auf die Natur der Partner) konzentrieren und es reicht ein Standard-Katalysator für sämtliche anfallenden Schritte.
  - Als Folge kann die Identität der jeweils hinzuzufügenden Aminosäure nicht in der Struktur des Katalysators vorgegeben werden, sondern muss für jeden Schritt „per Anweisung“ zugefüttert werden.
  - Dementsprechend wird eine bestimmte Sequenz durch einen zu dieser Sequenz colinear geordneten Satz von Instruktionen<sup>39</sup> bestimmt.
  - Alle derartigen Instruktionssätze müssen als Daten auf einem materiellen Träger abgelegt und von dort für die praktische Umsetzung (Interpretation) abrufbar sein.
  - Die als Daten auf Träger (DNA) abgelegten Instruktionssätze sind identisch mit den Strukturgenen für die Proteine.
  - Da die Gene zweier aus einer Zellteilung hervorgegangenen Zellen offenbar gleich sind (untereinander und zu denen der Mutterzelle) müssen sie als materielle Substanz mit hoher Genauigkeit kopierbar sein.

### 3 Schlussbemerkung

Aus dem hier skizzierten evolutionären Durchbruch („Blockbuster“) ging die in allen heutigen Organismen etablierte Trias der Makromoleküle DNA/RNA/Proteine hervor (über eine mögliche Vorstufe mit ausschließlich RNA). Dies hat nicht nur die informationsgesteuerte Synthese einer Vielzahl von Biokatalysatoren ermöglicht, sondern auch eine besonders effiziente Form molekularer Evolution: Die DNA ist nicht nur einerseits chemisch recht stabil (Voraussetzung für die unverfälschte Weitergabe der in ihr gespeicherten Information von einer Generation an die nächste – vgl. Punkt 11) sondern auch instabil genug, um gelegentliche Strukturveränderungen zu erleiden. Diese (seltenen!) „Spontanmutationen“ werden ihrerseits wieder stabil weitergereicht und sorgen für entsprechend modifizierte Proteine (plus deren Funktionen) und somit für veränderte selektive Fitness des betroffenen Organismus. Bei näherem Hinsehen haben wir es also mit zwei Blockbuster-Erfindungen, angelegt in nur einem funktionalen System zu tun.

---

<sup>39</sup>Dies ist die konzeptionell einfachste Programmierung der Synthese eines aperiodischen („atak-tischen“), aber mit definierter Sequenz ausgestatteten, Co-Polymers. So hat es die Natur in der belebten Welt eingerichtet; es ist allerdings nicht notwendigerweise die einzig mögliche Art, dies zu erreichen.

# Daten, Prozesse und Information

*R. Schaback*

*Institut für Numerische und Angewandte Mathematik, Lotzestraße 16-18, 37083  
Göttingen*

**Zusammenfassung.** Im Anschluss an den Artikel *Definitionen von Information* [17] in diesem Band soll hier ein Weg skizziert werden, zu einem möglichst breiten Informationsbegriff zu kommen, obwohl, wie schon a. a. O. zitiert, C. E. Shannon 1953 [21] feststellte:

*It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field.*

In diesem Beitrag wird *Information* als Zustandsänderung eines *Prozesses* definiert, die durch *Daten* bewirkt ist. Dazu sind im wesentlichen nur die Begriffe *Daten* und *Prozess* zu definieren. Von Ereignissen, Entropie, Algorithmen, Komplexität, *beliefs*, Interpretation und Semantik braucht man nicht zu reden, diese Begriffe sind nachgeordnet.

Eine Reihe von Beispielen zeigt, wie sich dieser Informationsbegriff in verschiedenen Disziplinen anwenden lässt und wie er die anderen Begriffe von Information subsumiert. Das ist eine Möglichkeit, die Arbeit der Akademiekommission *Die Natur der Information* zusammenzufassen.

## 1 Definitionen

### Definition 1.

*Daten* sind *Strukturen*, entweder auf einem *physikalischen Medium* oder *abstrakt*, z. B. als Werte mathematischer Variablen.

Sie können zeitlich und räumlich variabel sein. Um nicht in eine infinite Rekursion von Definitionen zu kommen, bleiben die Begriffe *Struktur*, *abstrakt* und *physikalisches Medium* undefiniert, ebenso wie Raum und Zeit.

### Definition 2.

1. Ein *Prozess* ist ein Objekt, das *Zustände* hat und in zeitkontinuierlicher oder diskret sequentieller Abfolge *Änderungen* dieser Zustände im Rahmen eines Repertoires an möglichen Zuständen zulässt.
2. Ferner hat ein Prozess eine *Umgebung* als Kollektion von Daten.

3. Der Prozess hat eine *Eingabe* und eine *Ausgabe*. Er kann Daten aus der Umgebung als Eingabe verwenden, und er kann Daten aus der Umgebung als Ausgabe verändern, ebenso sich selbst.
4. Ist ein Prozess  $P$  in einem Zustand  $Z$ , so wählt er mit gewissen Wahrscheinlichkeiten Daten aus der Umgebung aus, deren Eingabe dann mit einer gewissen Wahrscheinlichkeit zu einem *Zustandsübergang*  $Z \rightarrow Z'$  führt.
5. Jeder Zustandsübergang  $Z \rightarrow Z'$  führt mit einer gewissen Wahrscheinlichkeit Änderungen an Daten der Umgebung durch.
6. Die oben genannten Wahrscheinlichkeiten hängen von  $Z$  und von früheren Zuständen des Prozesses und von den eingegebenen Daten ab.
7. Jeder Zustandsübergang  $Z \rightarrow Z'$  führt mit einer gewissen Wahrscheinlichkeit Änderungen dieser Wahrscheinlichkeiten und an dem Repertoire der Zustände des Prozesses durch.

Diese Definition lehnt sich eng an die der *Automaten* der Informatik an [9]. Sie erhalten eine Folge von Zeichen als Eingabe, und dadurch werden Zustandsübergänge ausgelöst. Automaten wie die *Turingmaschine* erlauben es, Begriffe wie *Algorithmus* oder *Komplexität* mathematisch sauber zu fassen, und definieren eine sehr wichtige Klasse von Prozessen. Ihre Prozessumgebung besteht nur aus je einem Schreib- und Leseband. Genaueres über Turingmaschinen und ihren Einfluss auf die Modellierung von Kognitionsprozessen findet man im Beitrag [12] von Gerd Lüer und Uta Lass in diesem Band ab Seite 89.

Weil Prozesse im Sinne der Disziplin *Betriebssysteme* der Informatik deutlich komplexere Umgebungen haben, wurde dieser nützliche Begriff auch hier verwendet. Die Umgebung eines Prozesses kann man als den Teil einer allgemeinen 'Umwelt' verstehen, zu dem der Prozess überhaupt in Kontakt treten kann. Prozesse können sich selbst und die Umwelt verändern, letztere aber nur durch ihre Ausgabe. Die Daten der Umwelt können laufend durch parallel ablaufende Prozesse verändert werden. Konkurrenz und Adaptation sind möglich. Eine direkte Kommunikation zwischen Prozessen wird nicht angenommen. Stattdessen tauschen Prozesse Daten über die Umwelt aus, wie bei Schallwellen, Post und Telefon.

Determinismus liegt vor, wenn die genannten Wahrscheinlichkeiten immer Null oder Eins sind. Indem man gewisse Wahrscheinlichkeiten auf Null setzt, kann man z. B. Ausgabe oder Eingabe unterbinden.

### Definition 3.

Gegeben sei ein Prozess  $P$  im Zustand  $Z$  und ein Datum  $D$  aus der Prozessumgebung  $U(P)$ . Die *Information des Datums  $D$  in Bezug auf den Prozess  $P$  im Zustand  $Z$*  besteht aus den sämtlichen durch Eingabe von  $D$  bewirkten Änderungen am Prozess und seinem Zustand.

Letzteres betrifft

1. die möglichen Zustandsübergänge  $Z \rightarrow Z'$  von  $P$  bei Eingabe von  $D$  und die Eintretenswahrscheinlichkeiten dieser Zustandsübergänge,
2. alle Änderungen am Prozess selbst, d. h.
  - Veränderungen am Zustandsrepertoire und
  - Veränderungen an den Übergangswahrscheinlichkeiten.

Information ist ein Attribut eines Eingabedatums, aber relativ zum Prozess und seinem aktuellen Zustand. Es gibt keine prozess- und zeitabsolute Information.

Es kann sein, dass ein Prozess  $P$  in einem Zustand  $Z$  ist, in dem das Datum  $D$  nur mit Wahrscheinlichkeit Null zur Eingabe benutzt wird. Dann hat  $D$  keine Information in Bezug auf  $P$  und  $Z$ , und dies ist ein Beispiel für die Prozessrelativität von Information. Das Datum wird 'nicht zur Kenntnis genommen', der Prozess ist in diesem Zustand 'auf  $D$  unaufmerksam' oder 'nimmt  $D$  nicht wahr'. Ignorierte Daten haben keine Information.

Um nicht einen versteckten Interpretationsprozess anzunehmen, der aus einer Ursache eine Wirkung macht, wird Information nicht als etwas definiert, was eine Änderung 'bewirkt', sondern mit der Änderung selbst gleichgesetzt. Das *Eingabedatum* bewirkt etwas, nicht die Information, die diesem Datum relativ zum Prozess und seinem aktuellen Zustand zukommt. Die durch das Eingabedatum ausgelöste Veränderung von Zuständen *ist* die Information.

Information verändert deshalb auch nicht die Umgebung. Das tun nur Prozesse, auf Grund der Informationen in den Eingabedaten.

Wenn zwei Eingabedaten  $D_1$  und  $D_2$  in Bezug auf den Prozess  $P$  im Zustand  $Z$  dieselben Änderungen zur Folge haben, haben sie dieselbe Information, auch wenn sie verschieden sind.

#### Definition 4.

Das Verhalten eines Prozesses in jedem Zustand  $Z$  und bei jedem Eingabedatum  $D$  ist festgelegt durch eine prozessinterne abstrakte Datenstruktur  $R_P(Z, D)$ , die als *interne Repräsentation* des Prozesses bezeichnet wird. Bei nichtdeterministischem Verhalten des Prozesses sind die entsprechenden Wahrscheinlichkeiten Bestandteil der internen Repräsentation.

Die *Information eines Datums*  $D$  in Bezug auf den Prozess  $P$  im Zustand  $Z$  besteht dann aus der durch  $D$  bewirkten Änderung  $R_P(Z', D') \rightarrow R'_P(Z', D')$  der internen Repräsentation des Prozesses, für alle möglichen Folgezustände  $Z'$  und Eingabedaten  $D'$ .

Die *internen Repräsentationen* von Prozessen sind aus den Sprach- und Kognitionswissenschaften entlehnt. Sie dienen lediglich dazu, die 'Änderungen am Prozess' der Definition 3 konkreter zu verorten und der Simulation auf einem Computer Vorschub zu leisten. Die interne Repräsentation sagt dem Prozess, 'was zu tun ist'.

Es wird nicht behauptet, dass die hier gegebene Informationsdefinition neu sei. Zum Beispiel beginnt K. Bachmann sein Manuskript [1] mit

*In Ermangelung einer allgemein akzeptierten Definition benutze ich Information hier im Sinne von Struktur, die von einem Empfänger erkannt wird und im Empfänger programmierte informations-abhängige („bedingte“) Reaktionen steuert. Die steuernde Einheit ist die Nachricht. Information impliziert eine selektive und subjektive Wahrnehmung und Interpretation von Strukturen, abhängig vom Empfänger.*

Hier ist zwar die Information mit den Eingabedaten gleichgesetzt, aber im Wesentlichen geht es um die dadurch bewirkte Steuerung des empfangenden Prozesses. Wie K. Bachmann mit diesem Informationsbegriff an Evolution herangeht, wird im Rahmen anderer Beispiele unten dargestellt.

## 2 Andere Informationsbegriffe

Die mathematische Informationsdefinition wird hier ignoriert, weil sie mit dem mathematischen Wahrscheinlichkeitsbegriff bis auf eine eindeutige Umrechnung übereinstimmt und somit wissenschaftlich redundant ist. Ebenso wird der algorithmische Informationsbegriff ausgeklammert, denn er beschreibt Komplexität, nicht Information.

Wie bei der bayesianischen Informationsdefinition in [17] (vgl. auch den Beitrag [20] von Kurt Schönhammer in diesem Band ab Seite 27) ist Information genau dann vorhanden, wenn sich die interne Repräsentation des interpretierenden Prozesses ändert, d. h. ein *change of beliefs* eintritt, und sie besteht aus diesem. Also ist die bayesianische Informationsdefinition ein Spezialfall.

Die semantische Informationsdefinition aus [17] rekuriert auf *Interpretation* und *Semantik*, und die Information kommt *strukturierten Daten* zu. Auch hier ist die Information ein Attribut eines Datums, aber Daten sind hier mit Strukturen synonym. Deshalb sind Daten immer *well-formed* im Vergleich zur Informationsbegriff von *Luciano Floridi*. [4]. Eingabedaten und Information sind hier immer relativ zu einem Prozess und seinem momentanen Zustand. Wenn man den Prozess als Interpretationsprozess eines Eingabedatums sieht, kommen sich die Definitionen sehr nahe, aber hier ist *semantic content* nicht explizit definiert. Der Interpretationsprozess verändert die Prozessrepräsentation und löst gegebenenfalls eine Aktion des Prozesses durch Modifikation der Umgebung oder seiner selbst aus. Mehr ‘Semantik’ oder *meaning* ist nicht da. Im zweiten Fall könnte man im Sinne von [4] von *instructional information* reden, im ersten Fall von *factual information*, und die Modifikation betrifft entweder das ‘prozedurale Gedächtnis’ oder das ‘deklarative Gedächtnis’ im Sinne kognitiver Architekturen, vgl. [23]. Mischformen sind erlaubt, und die *Interpretation durch Aktion* aus [17] ist ein Spezialfall.

Diese Sichtweise lässt die ‘Semantik’ eines Eingabedatums mit der zugehörigen Information zusammenfallen, weil beide aus den Änderungen am Prozess bestehen. Dadurch wird Semantik prozessrelativ, aber das ist nicht verwunderlich. Es

gibt keine absolute Semantik und kein absolutes Wissen. Beide sind immer relativ zum verarbeitenden Prozess und sind durch die prozessinterne Repräsentation determiniert.

Die Information verändert nur den Empfängerprozess. Wenn man Kommunikation modelliert, ist das Versenden von Daten durch einen Senderprozess ein Bestandteil des Senderprozesses allein, und es ist zu untersuchen, ob dieser so geartet ist, dass man ihm unterstellen kann, er habe ‘Bedeutung’ in die Daten gelegt, was ‘Bedeutung’ auch immer sein soll. Werden diese Daten zu Eingabedaten eines anderen Prozesses, und wird dieser durch diese Daten verändert, so liegt Information dieser Daten in Bezug auf den Empfängerprozess vor, und diese Prozessänderung kann definitorisch mit einer ‘Bedeutung’ für den Empfängerprozess identifiziert werden. So lässt sich ‘Semantik’ modellieren, aber primär nur im Empfängerprozess.

Den hier vorgestellten Informationsbegriff kann man nicht transitiv verwenden wie in

*If A carries the information that B, and B carries the information that C, then A carries the information that C [2]*

weil er nur aus der Wirkung eines Eingabedatums auf einen Prozess besteht und sonst nichts. Was das Eingabedatum mit anderen Dingen, Objekten oder Sachverhalten zu tun hat, ist irrelevant. Der Versuch, über ‘Information’ Epistemologie zu betreiben, geht bei diesem Informationsbegriff ins Leere. Man hat den Erkenntnisprozess und seine Modifikation durch Eingabedaten zu untersuchen.

### 3 Beispiele

Es ist klar, dass alle Computerprogramme und alle Automaten der Informatik als Prozesse im Sinne dieses Artikels verstanden werden können. Die Semantik von Programmiersprachen wird durch die Änderung der Verarbeitungsprozesse definiert, die sie beschreiben.

Den *zellulären Automaten* liegt ein Prozess zugrunde, bei dem die *Zellen* eines *Zellularraums* ihre Zustände in Abhängigkeit von den Zuständen benachbarter Zellen ändern, und damit lassen sich sehr komplexe Vorgänge beschreiben [14]. Man kann sogar postulieren, dass sich das ganze Universum [27] oder alle Wissenschaften [26] damit modellieren lassen.

Bei *künstlichen neuronalen Netzen* findet ein ähnlicher Prozess statt, der aber nicht nur die Veränderung des Prozessverhaltens, sondern auch der den Prozess tragenden Strukturen erlaubt. Sie sind frühe Beispiele für lernfähige Systeme. Das Gebiet *Computational Intelligence* [11] ist noch sehr viel weiter gefasst und zielt auf das inzwischen sehr erfolgreiche *Maschinelle Lernen*. Weitere Details über neuronale und allgemeinere Netze finden sich im Artikel „*Informationsverarbeitung in der Kognitionspsychologie und in den kognitiven Neurowissenschaften*“ von Gerd Lüer und Uta Lass ab S. 89 in diesem Band, bezogen auf Kognitionspsychologie und kognitive Neurowissenschaften.

In allen genannten Fällen liegen Daten und Prozesse vor, und die obige Definition von Information ist anwendbar. Die Prozesse sind sämtlich auf Computern ausführbar.

Bei der Veränderung thermodynamischer Prozesse (vgl. den Artikel „*Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information*“ [18] von Kurt Schönhammer ab S. 45 in diesem Band) gibt es einen Entropieanstieg, den man als Änderung des Informationsgehalts sehen kann.

Durch das Hinzutreten neuer Ereignisse ändert sich ein bayesianistisches System (siehe den Artikel „*Wahrscheinlichkeit und Information – wie Henne und Ei?*“ von Kurt Schönhammer ab S. 27 in diesem Band), weil sich alle Einzelwahrscheinlichkeiten ändern, indem sie zu bedingten Wahrscheinlichkeiten werden.

Durch Messung an einem quantenmechanischen System ändert sich das komplette System, inklusive Verschränkungen. Das wird im Artikel „*Quanteninformation: Versuch einer Einführung für Nicht-Physiker*“ [19] von Kurt Schönhammer ab S. 73 in diesem Band genauer erläutert.

Die *informationsgesteuerte Synthese* [5] in Zellen zur Replikation von DNA oder Produktion von Proteinen ist ein weiteres Beispiel. Die Prozesse sind biochemische Synthesen, und *jeder biochemischen Reaktion ist ein Enzym und jedem Enzym ein seine Synthese steuerndes Gen zugeordnet* [7]. Deshalb liegt die Information in den Genen, wenn es um Enzymsynthese geht, und in den Enzymen bei allgemeineren biochemischen Reaktionen.

Mikroben zeigen eine Art Sozialverhalten, ausgelöst durch chemische Signale.

*Die Antwort der Zelle auf ein Eingangssignal besteht, neben einem speziellen Verhalten, aus dem Senden eines Ausgangssignals als Input für andere Zellen* [6]

Die Information der Signale als Eingabedaten besteht genau aus den genannten Prozessveränderungen.

Signalen a priori eine tiefergehende und prozessinvariante Semantik zuzuschreiben, ist auch in der Ethologie eher fragwürdig als hilfreich. In der klassischen Sichtweise würde man behaupten,

1. ein Warnsignal<sup>1</sup>, z. B. das Pfeifen eines Murmeltiers, trüge die ‘Information’, dass ein spezifischer Fressfeind anwesend sei, und zwar womöglich durch eine Intention des warnenden Tieres, und
2. ein das Signal wahrnehmendes Tier unternehme eine Fluchtreaktion, *weil* das Signal diese Information trüge.

---

<sup>1</sup>Hinweis von Bernhard Ronacher: Hier sind keine aposematischen Warnsignale gemeint, d. h. Zeichen potentieller Beutetiere im Sinne von „rühr’ mich nicht an“ oder „ich bin nicht essbar“.



Diese sehr anthropomorphe Schlussweise sollte nach Stegmann [22] ersetzt werden durch: Das Tier ergreift die Fluchtreaktion, weil es evolutionär oder durch Lernen dazu disponiert ist, eine Fluchtreaktion zu ergreifen, wenn es das Signal wahrnimmt. Das Signal als Eingabedatum bewirkt eine Prozessänderung, und *diese* ist die Information, und dem Signal sollte nicht die Semantik ‘Hier ist ein Fressfeind’ unterlegt werden.

*Information agents* im Sinne von Floridi [3] sind Prozesse der hier gemeinten Art, aber auch autopoietische Systeme im Sinne von *Luisi* [13].

Der Beitrag [12] von Gerd Lüer und Uta Lass beschreibt auf Seite 89 in diesem Band detailliert das *Informationsverarbeitungsparadigma* der Kognitionspsychologie:

*Danach sind alle höheren geistigen Leistungen, über die Menschen verfügen, wie z. B. das Entscheiden, Urteilen und Problemlösen, die Aufmerksamkeit und der Wissenserwerb sowie die Gedächtnisnutzung am besten durch Prozesse der Informationsverarbeitung zu beschreiben und zu erklären.*

Dazu sind cognitive maps und *kognitive Architekturen* geeignet, und letztere ... *verfügen über Wissensbasen vom eigenen System sowie über Repräsentationen von der Umwelt.* Das entspricht ziemlich genau den hier gemeinten internen Prozessrepräsentationen, die auch in der ‘Diskursrepräsentationstheorie’ [10, 25] der Sprachwissenschaften auftreten.

In den kognitiven Neurowissenschaften stellt sich die Frage *Welche Hirnstrukturen ermöglichen kognitive Leistungen wie z. B. das Denken oder das Lernen?* [12]. Es liegen *zwei* Prozesse vor, ein biologisch-neuronaler und ein mental-kognitiver, und das Verhältnis dieser Prozesse ist das Problem. Jeder der beiden hat seine typischen Eingabedaten, die den Prozess verändern, und die jeweiligen Veränderungen sind die jeweiligen Informationen, die den Eingabedaten relativ zu den Prozessen zukommen. Schon im Bereich der Wahrnehmung tritt dieses ‘Leib-Seele’-Problem in voller Schärfe auf, und es kann nur durch Erforschung der Parallelität der Prozesse sinnvoll angegangen werden. Das wird in [12] detailliert beschrieben.

Mentale und neuronale Prozesse sind Spezialfälle der hier vorgeschlagenen Begriffssystematik, aber es ist weder zwingend noch beabsichtigt, der Definition von Prozessen einen Positivismus oder Reduktionismus zu unterlegen, indem mentale Prozesse auf neuronale reduziert werden. Beide sind wissenschaftlich interessant, und insbesondere ihr Verhältnis zueinander. Dazu zwei konträre Beispiele.

Bereits das Studium der Wahrnehmung [15] zeigt, dass neuronale Prozesse, die eng mit den Sinnesorganen verbunden sind, auf komplexe Weise auf höhere mentale Prozesse wirken (*Sensory System* versus *Perceptual System*), die komplexe *conceptual forms* (such as ‘perceptual object’, ‘surface’, ‘food’, ‘enemy’, ‘tool’, and ‘causal event’) als Objekte verarbeiten. Ähnlich geht Treue [24] vor, indem er zunächst Wahrnehmung auf der Basis einer vom Nervensystem erzeugten Repräsentation der Umwelt entstehen lässt, und *zudem müssen dann noch Entscheidungssysteme existieren, die eine[r] zweckmäßige[n] Auswahl der Handlungen des Organismus zu Grunde liegen.*

Hier ‘dienen’ die neuronalen Systeme den mentalen. Man kann aber mit Metzinger [16] auch fragen, ob die neuronalen Prozesse eine so starke Wirkung auf die mentalen Prozesse haben, dass unser mental erlebtes ‘Selbst’ ein Produkt der neuronalen Ebene sind. Jetzt beherrschen die neuronalen Systeme die mentalen.

Die mehrfach aufgeworfene Frage der Kommission nach dem Spannungsfeld zwischen Materie und Geist transformiert sich im obigen Kontext in die Frage, ob menschliches Bewusstsein sich rein auf materielle Repräsentationen im Nervensystem stützt oder andere ‘mentale’ Repräsentationen vorliegen. Man hat nach der ‘Natur der kognitiven Repräsentation’ zu fragen, wenn man Kognition sehr allgemein definiert. Und das ist ein weites Feld. Diverse Beiträge in der Kommissionsarbeit verfolgten Nervenimpulse von den Sinnesorganen in den Cortex, und es gibt offenbar gewisse Hirnregionen, die für gewisse Repräsentationen erforderlich sind, aber das ist keine Antwort auf die oben sehr prinzipiell gestellte Frage. Man hat zu untersuchen, welche Repräsentationen kognitive und neuronale Prozesse haben, und wie diese Repräsentationen korreliert sind.

In enger Anlehnung an K. Bachmann [1] kann die Anwendung des hier beschriebenen Informationsbegriffs auf Evolution damit beginnen, dass die lebenden Systeme bei ihrer Fortpflanzung auch eine Kopie ihres Prozesses weitergeben, und diese Kopien evolvieren dann. *Darwinsche Evolution ist ein unvermeidlicher Nebeneffekt des Lebensvorgangs.* Die Notwendigkeit der Lebenserhaltung zwingt die Prozesse, geeignete Rezeptoren für Umwelt-Eingabedaten zu entwickeln, ... *neben den Rezeptoren, mit denen sie sich selbst interpretier[en]. Aus den Mechanismen zur Wahrnehmung und Verarbeitung von Umwelt-Information ergibt sich eine Beeinflussbarkeit von Organismen durch Nachrichten (Signale) aus der Umwelt, bei der die Kommunikation zwischen Organismen mit der Zeit in der Evolution eine immer größere Bedeutung relativ zum Einfluss der unbelebten Umwelt bekommt. Organismen werden von anderen Organismen als Ressource oder Konkurrent wahrgenommen, und das wirkt als starker Darwinscher Selektionsfaktor.*

## 4 Erweiterungen

Will man Leben modellieren, braucht man als mögliche Zustände eines Prozesses

1. einen ‘Gebärzustand’, der einen neuen Prozess mit einem Standard-Zustandsrepertoire an die Umgebung abliefern, und
2. einen ‘Todeszustand’, in dem alle Wahrscheinlichkeiten Null sind.

Man kann mit K. Bachmann [1] die Vererbung als Übergabe eines Standard-Zustandsrepertoires an den neu erzeugten Prozess sehen. Das hat Ähnlichkeit mit genetischen oder evolutionären Algorithmen [8].

Bei Prozessen als belebte Organismen ist ein Teil des Prozesses (der ‘Körper’) Bestandteil der Umwelt und deshalb für andere Prozesse als Eingabedatum möglich, d. h. ‘wahrnehmbar’. Im Todeszustand bleibt nur dieser Teil im System und ist innerhalb dessen veränderbar. Verwesung und Gefressenwerden sind modellierbar.

Prozesse, die in der Lage sind, sich selbst bzw. ihre internen Repräsentationen in der gleichen Weise wie ihre Umgebung wahrzunehmen, könnte man ‘reflektorische’ Prozesse nennen. Dann sollten kognitive Prozesse reflektorisch sein und zu einer Selbstwahrnehmung führen, die möglicherweise nicht von der neuronalen Körperwahrnehmung determiniert ist, im Gegensatz zu Metzinger [16].

## 5 Fazit

Dieser Ansatz vermeidet die Begriffe *Semantik* und *Interpretation*. Sie werden durch *Prozess* ersetzt. Das kann man als billigen Taschenspielertrick sehen, der nichts erklärt. Stimmt, bei vordergründiger Sichtweise. Aber es ist wie ein Wechsel des Koordinatensystems oder des Bezugssystems in Mathematik oder Physik. Inhaltlich hat sich nichts getan, aber man sieht klarer und kann sich besser zurechtfinden. Genau wie wenn man eine neue Landkarte verwendet, die ja auch nicht inhaltlich die Landschaft verändert.

Die inhaltliche Frage nach der ‘Natur der Information’ verwandelt sich in die Frage nach Struktur und Verhalten von Prozessen, die Daten ‘von außen’ erhalten und damit irgendwie umgehen. Die verwirrende Vielfalt der Erscheinungsformen von Information wird etwas klarer benannt als Vielfalt von Prozessen, die Nachrichten ‘verarbeiten’, kognitiv, mental oder mechanisch, durch Aktion oder Zustandsänderung.

Die Analyse der einzelnen Prozesse bleibt als zentrale Aufgabe bestehen, aber der Informationsbegriff wird dabei zum Randthema.

## Literatur

- [1] K. Bachmann. Information und Evolution: Überblick. Begleitmanuskript zum Vortrag *Information und Evolution: Das C-Wert-Paradox und das Genom der Eukaryonten* in der Akademiekommission *Die Natur der Information*, 18.04.2008.
- [2] F. I. Dretske. *Knowledge and the Flow of Information*. MIT Press, 1981.
- [3] L. Floridi. *The Ethics of Information*. Oxford University Press, 2013.
- [4] L. Floridi. Semantic Conceptions of Information. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy* 2014. <http://plato.stanford.edu/archives/spr2014/entries/information-semantic> <2019-03-15>
- [5] H.-J. Fritz. Informationsgesteuerte Synthese - eine Blockbuster-Erfindung der Evolution. In H.-J. Fritz and R. Schaback, Herausgeber, *Die Natur der Information*, 167–182. Universitätsverlag Göttingen, 2019.
- [6] H.-J. Fritz. Nachricht und Information in der belebten Welt. Plenarvortrag in der Akademie, 21.01.2011.

- 
- [7] H.-J. Fritz. Datenträger und Informationsfluß in der Molekularen Biologie. Vortrag in der Akademiekommission *Die Natur der Information*, 30.03.2006.
- [8] I. Gerdes, K. Klawonn, and R. Kruse. *Evolutionäre Algorithmen: genetische Algorithmen - Strategien und Optimierungsverfahren - Beispielanwendungen*. Vieweg, Wiesbaden, 2004.
- [9] J. E. Hopcroft, R. Motwani, and J. D. Ullman. *Einführung in Automaten-theorie, Formale Sprachen und Berechenbarkeit*. Pearson Studium IT, 2011. Dritte Auflage.
- [10] H. Kamp and U. Reyle. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht, 1993.
- [11] R. Kruse, C. Borgelt, F. Klawonn, C. Moewes, G. Ruß, and M. Steinbrecher. *Computational Intelligence: Eine methodische Einführung in Künstliche Neuronale Netze, Evolutionäre Algorithmen, Fuzzy-Systeme und Bayes-Netze*. Vieweg+Teubner Verlag, 2011.
- [12] G. Lüer and U. Lass. Informationsverarbeitung in der Kognitionspsychologie und in den kognitiven Neurowissenschaften. In H.-J. Fritz and R. Schaback, Herausgeber, *Die Natur der Information*, 89–117. Universitätsverlag Göttingen, 2019.
- [13] Pier Luigi Luisi. Autopoiesis: a review and a reappraisal. *Naturwissenschaften*, 90:49–59, 2003.
- [14] K. Mainzer and L. Chua. *The Universe as Automaton. From Simplicity and Symmetry to Complexity*. Springer, Heidelberg, 2011.
- [15] R. Mausfeld. Intrinsic Multiperspectivity: Conceptual Forms and the Functional Architecture of the Perceptual System. In W. et. al. Welsch, editor, *Interdisciplinary Anthropology*. Springer-Verlag Berlin Heidelberg, 2011. Begleitmanuskript zum Vortrag *Information und Bedeutung in der Wahrnehmung* in der Akademiekommission *Die Natur der Information* am 18.01.2013.
- [16] T. Metzinger. *Empirische Perspektiven aus Sicht der Selbstmodell-Theorie der Subjektivität: Eine Kurzarstellung mit Beispielen*. Selbstverlag, ePUB, EAN 9783845016177, 2015. Begleitmanuskript zum Vortrag „Körperrepräsentation, das minimale phänomenale Selbst und die Erste-Person-Perspektive: Was ist die einfachste Form von Ichgefühl?“ in der Akademiekommission ‚Die Natur der Information‘ am 07.06.2013.
- [17] R. Schaback. Definitionen von Information. In H.-J. Fritz and R. Schaback, Herausgeber, *Die Natur der Information*, 11–25. Universitätsverlag Göttingen, 2019.

- 
- [18] K. Schönhammer. Der Entropiebegriff in der Physik und seine Beziehung zum Konzept der Information. In H.-J. Fritz and R. Schaback, Herausgeber, *Die Natur der Information*, 45–72. Universitätsverlag Göttingen, 2019.
- [19] K. Schönhammer. Quanteninformation: Versuch einer Einführung für Nicht-Physiker. In H.-J. Fritz and R. Schaback, Herausgeber, *Die Natur der Information*, 73–87. Universitätsverlag Göttingen, 2019.
- [20] K. Schönhammer. Wahrscheinlichkeit und Information – wie Henne und Ei? In H.-J. Fritz and R. Schaback, Herausgeber, *Die Natur der Information*, 27–43. Universitätsverlag Göttingen, 2019.
- [21] C. E. Shannon. The lattice theory of information. *Transactions of the IRE Professional Group on Information Theory*, 1:105–107, 1953.
- [22] U. Stegmann. Information in Ethology. Vortrag in der Akademiekommission *Die Natur der Information*, 20.04.2012.
- [23] W. H. Tack. Kognitive Architekturen. Vortrag in der Akademiekommission *Die Natur der Information*, 20.04.2007.
- [24] S. Treue. Die Verarbeitung von visueller Bewegungsinformation in der Großhirnrinde. Öffentlicher Plenumsvortrag am 21. Januar 2011.
- [25] G. Webelhuth. Diskursrepräsentationstheorie. Vortrag in der Akademiekommission *Die Natur der Information*, 02.11.2007.
- [26] S. Wolfram. *A new kind of science*. Wolfram Media Inc., 2002.
- [27] K. Zuse. *Rechnender Raum*. Schriften zur Datenverarbeitung Band 1. Friedrich Vieweg & Sohn Braunschweig, 1969.



Die Natur der Information  
H.-J. Fritz, R. Schaback (Hrsg.)  
© Universitätsverlag Göttingen 2019

## Vortragsreihe der Akademiekommission

16.12.2005

Prof. Dr. Robert Schaback, Göttingen  
Zum Begriff der Information in Mathematik und Informatik

06.01.2006

Prof. Dr. Kurt Schönhammer, Göttingen  
Der Entropiebegriff in der Thermodynamik und der Statistischen Mechanik I

20.01.2006

Prof. Dr. Werner Lehfeldt, Göttingen  
Anmerkungen zum Informationsbegriff in der Linguistik

10.03.2006

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Informationsgesteuerte Synthese – eine Blockbuster-Erfindung der Evolution

30.03.2006

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Datenträger und Informationsfluß in der Molekularen Biologie

28.04.2006

Prof. Dr. Gerd Lüer, Göttingen  
Kognitionspsychologie – eine Psychologie der Informationsverarbeitung

02.05.2006

Prof. Dr. Bernd Ronacher, HU Berlin  
Klassifikation akustischer Kommunikationssignale durch Heuschrecken:  
Probleme der Informationsverarbeitung

23.06.2006

Prof. Dr. Kurt Schönhammer, Göttingen  
Der Entropiebegriff in der Thermodynamik und der Statistischen Mechanik II

08.12.2006

Prof. Dr. Robert Schaback, Göttingen  
Zum Begriff der Information in Mathematik und Informatik

08.12.2006

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Informationstheoretische Metaphern im Vokabular der Molekularen Genetik

08.12.2006

Prof. Dr. Werner Lehfeldt, Göttingen  
Sprachliche Information als soziales Phänomen

15.12.2006

Prof. Dr. Norbert Elsner, Göttingen  
Bau und Funktionsweise natürlicher Neuronen

16.02.2007

Prof. Dr. Annette Zippelius, Göttingen  
Was sind neuronale Netzwerke?

20.04.2007

Prof. Dr. Werner H. Tack, Saarbrücken  
Kognitive Architekturen

04.05.2007

Prof. Dr. Theo Herrmann, Mannheim  
Sprachpsychologie

28.06.2007

Prof. Dr. Manfred Eigen, Göttingen  
Was ist Information?  
Vortrag in der Akademie

13.07.2007

Prof. Dr. Gerd Webelhuth, Göttingen  
Zur Modellierung der Sprachform und der Sprachbedeutung  
in der Gegenwartslinguistik

19.10.2007

Prof. Dr. Florentin Wörgötter, Göttingen  
Autonome, kognitive(?) Roboter:  
Heutige technische Möglichkeiten und die Frage nach der Einbettung  
solcher Systeme in unsere Welt

02.11.2007

Prof. Dr. Gerd Webelhuth, Göttingen  
Diskursrepräsentationstheorie II

09.11.2007

Prof. Dr. Luciano Floridi, Oxford  
The future development of the information society  
Vortrag in der Akademie

09.11.2007

Prof. Dr. Luciano Floridi, Oxford  
What is Bio-Information?



30.11.2007

Prof. Dr. Christof Biebricher, Göttingen  
Fragen der Entstehung des Lebens

18.01.2008

Prof. Dr. Werner Lehfeldt, Göttingen  
Subtile Mechanismen der sprachlichen Nachrichtenkodierung

15.02.2008

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Schnelle genetische Adaptation

18.04.2008

Prof. Dr. Konrad Bachmann, Gatersleben  
Information und Evolution: Das *C*-Wert-Paradox und das Genom der Eukaryonten

16.05.2008

Prof. Dr. Theo Geisel, Göttingen  
Neuronale Dynamik, Synchronisation und Informationsverarbeitung

30.05.2008

Prof. Dr. Ulrich Reimers, TU Braunschweig  
Moderne Kommunikationstechniken

13.06.2008

Prof. Dr. Helmuth Feilke, Gießen  
Sprachliche Vorgeformtheit – Zur Rolle von Muster und Prägung  
für Sprachverständnis und Sprachproduktion

17.10.2008

Frau Prof. Dr. Regine Eckardt, Göttingen  
Evolution menschlicher Sprache

31.10.2008

Prof. Dr. Konrad Bachmann, Gatersleben  
Biologische und nicht-biologische Information

14.11.2008

Prof. Dr. Robert Schaback, Göttingen  
Komplexität von Sprachen und sprachverarbeitenden Systemen  
aus Sicht der Informatik

12.12.2008

Prof. Dr. Michael Richter, Kaiserslautern und Calgary  
Der Ähnlichkeitsbegriff und die Suche nach Information

30.01.2009

Prof. Dr. Kurt Schönhammer, Göttingen  
Wahrscheinlichkeit und Information – wie Henne und Ei?

24.04.2009

Prof. Dr. Gerd Gigerenzer, Berlin  
Transparenz in der Risikokommunikation

12.05.2009

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Kommunikation und soziales Verhalten bei Mikroben

10.07.2009

Prof. Dr. Luciano Floridi, Hertfordshire und Oxford  
Semantic Information and the Correctness Theory of Truth

17.07.2009

Prof. Dr. Gunter Dueck, IBM Heidelberg  
Gedanken zur Mathematik des Instinktes

17.07.2009

Prof. Dr. Gunter Dueck, IBM Heidelberg  
Gedanken zur Mathematik des Instinktes

29.01.2010

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Biochemie der Unterdrückung von Rauschen  
bei der Weitergabe genetischer Information

16.04.2010

Prof. Dr. Gerald Spindler, Göttingen,  
Die rechtliche Behandlung von Informationen:  
Schutz (Urheberrecht, Patentrecht) und Haftung

28.05.2010

Prof. Dr. Bert Hölldobler, Würzburg und Arizona State,  
Der Superorganismus der Ameisen: Zivilisation durch Instinkt

11.06.2010

Prof. Dr. Konrad Cramer, Göttingen  
Eine Kritik des Naturalismus

09.07.2010

Prof. Dr. Pier Luigi Luisi, Rom III  
From the origin of life to cognition

10.12.2010

Prof. Dr. Andreas Dress, Bielefeld  
Manfred Eigen's Informationsraum und die Theorie der phylogenetischen Netze

21.01.2011

Prof. Dr. Stefan Treue, Göttingen  
Die Verarbeitung von visueller Bewegungsinformation in der Großhirnrinde  
öffentliche Sitzung der Akademie, Paulinerkirche

21.01.2011

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Nachricht und Information in der belebten Welt  
öffentliche Sitzung der Akademie, Paulinerkirche

21.01.2011

Prof. Dr. Robert Schaback, Göttingen  
Nachricht und Information, Text und Textverstehen:  
Ein Bericht über die Arbeit der Kommission  
öffentliche Sitzung der Akademie, Paulinerkirche

28.01.2011

Prof. Dr. Stefan Treue, Göttingen  
Vom Signal zur Wahrnehmung:  
sensorische Informationsverarbeitung als Abbild oder Repräsentation

06.05.2011

Priv. Doz. Dr. Meinard Müller, Saarbrücken  
Musik trifft Informatik:  
Automatisierte Methoden zur Verarbeitung von Musikdaten

20.05.2011

Prof. Dr. Peter Janich, Marburg  
Eine konstruktive Lösung des Widerspruchs  
zwischen natur- und geisteswissenschaftlichen Informationsbegriffen

10.06.2011

Prof. Dr. Kurt Schönhammer, Göttingen  
Quanteninformatik: Versuch einer Einführung für Nichtphysiker

01.07.2011

Prof. Dr. Wolfgang Künne, Hamburg  
Stufen des (Miss-) Verstehens

09.12.2011

Prof. Dr. Hans-Joachim Fritz, Göttingen  
Die Übertragung der Informationsspeicherung von RNA auf DNA –  
ein Fall von ‚Genetic Takeover‘

20.01.2012

Prof. Dr. Werner Lehfeldt, Göttingen  
Über die Schwierigkeiten eines Sprachwissenschaftlers,  
sich seines Gegenstandsbereichs zu vergewissern

03.02.2012

Prof. Dr. Rainer Kuhlen, Konstanz und HU Berlin  
Information als immaterielles Gemeingut – eine informationsethische Begründung

20.04.2012

Prof. Dr. Ulrich Stegmann, Aberdeen  
Information in Ethology

08.06.2012

Prof. Dr. Regine Eckardt, Göttingen

Textstrukturen: Von der wörtlichen Bedeutung zur Information in Texten

22.06.2012

Prof. Dr. Julia Fischer, Göttingen

Information und Inferenz in der Kommunikation von Tieren

7.12.2012

Professor Dr. Klaus Mainzer, TU München

Informationsdynamik in komplexen Systemen –  
Interdisziplinäre Grundlagen und Perspektiven

18.01.2013

Prof. Dr. Rainer Mausfeld, Kiel

Information und Bedeutung in der Wahrnehmung

07.06.2013

Prof. Dr. Thomas Metzinger, Mainz

Körperrepräsentation, das minimale phänomenale Selbst  
und die Erste-Person-Perspektive: Was ist die einfachste Form von Ichgefühl?

08.11.2013

Prof. Dr. Adrian Paschke, Berlin und Posen

Ein universelles Pragmatic Web 4.0: Auf dem Weg zur Intelligenz im Internet