

106. Band Heft 3, September 2004

D 20577



Jahresbericht

der Deutschen Mathematiker-Vereinigung

3 – 2004

Herausgegeben von A. Krieg
unter Mitwirkung von
U. Gather, E. Heintze, B. Kawohl,
H. Lange, H. Triebel



Jahresbericht

der Deutschen Mathematiker-Vereinigung

Der „Jahresbericht“ ist das offizielle Veröffentlichungsorgan der Deutschen Mathematiker-Vereinigung, für dessen inhaltliche Gestaltung im Auftrag des Präsidiums der jeweilige Herausgeber zuständig ist. Im „Jahresbericht“ sollen vornehmlich Überblicksartikel über Teilgebiete der reinen und angewandten Mathematik, Nachrufe sowie historische Artikel und Buchbesprechungen veröffentlicht werden.

Manuskripte:

Alle für die Schriftleitung des Jahresberichts bestimmten Briefe und Manuskripte sind an Prof. Dr. A. Krieg zu richten. Für Buchbesprechungen ist Prof. Dr. H. Lange zuständig. Bücher, von denen eine Besprechung erfolgen soll, werden bei den Verlagen angefordert. Autoren von Buchbesprechungen und Artikeln werden gebeten, die vorhandenen LATEX-style-files für den Jahresbericht zu verwenden. Somit kann der Aufwand für die Satzarbeiten erheblich reduziert werden. Sollten Illustrationen in die Arbeiten integriert werden, können diese auch in das Satzsystem übernommen werden. Dazu ist es erforderlich, dass die Daten der Abbildungen nochmals in separaten Dateien einzeln abgespeichert werden. Die LATEX-style-files sind neben weiteren Informationen im Internet verfügbar unter

<http://www.mathA.rwth-aachen.de/dmv/index.html>

Auf Anfrage können die style-files auch auf Diskette zugeschickt werden.

Grundsätzlich sollen nur solche Manuskripte eingereicht werden, die nicht gleichzeitig an anderer Stelle zur Veröffentlichung eingereicht oder bereits veröffentlicht worden sind. Mit der Annahme zur Veröffentlichung erwirbt der Verlag das Recht zur Vervielfältigung und Verbreitung sowie das Recht der Übersetzung in andere Sprachen.

Verlag:

B. G. Teubner Verlag/GWV Fachverlage GmbH
Abraham-Lincoln-Straße 46
65189 Wiesbaden
<http://www.teubner.de>
<http://www.gvv-fachverlage.de>

Geschäftsführer: Dr. Hans-Dieter Haenel
Verlagsleitung: Dr. Heinz Weinheimer
Gesamtleitung Anzeigen: Thomas Werner
Gesamtleitung Produktion: Reinhard van den Hövel
Gesamtleitung Vertrieb: Gabriel Göttlinger

Marketing/Sonderdrucke:

Eva Brechtel-Wahl
Telefon: (06 11) 78 78-3 79
Fax: (06 11) 78 78-4 39
E-Mail: eva.brechtel-wahl@gvv-fachverlage.de

Abonnentenverwaltung:

(Änderungen von Adressen und Bankverbindung, Rückfragen zu Rechnungen oder Mahnung)
VVA-Zeitschriftenservice, Abt. D6F6 / Jahresbericht der Deutschen Mathematiker-Vereinigung,
Postfach 7777, 33310 Gütersloh
Ursula Müller
Telefon: (0 52 41) 80-19 65
Fax: (0 52 41) 80-96 20
E-Mail: ursula.mueller@bertelsmann.de

Bezugsbedingungen:

Die Zeitschrift erscheint 4mal jährlich zum Jahresabonnementpreis von € 102,- (176,- sFr) inkl. Versandkosten. Der Bezug von Einzelheften ist nicht möglich. Schriftliche Kündigung des Abonnements spätestens sechs Wochen vor Ablauf des Bezugsjahres.
Für persönliche Mitglieder der DMV, die den Jahresbericht zu beziehen wünschen, ist der zwischen DMV und Verlag vereinbarte Bezugspreis maßgebend, der im Rahmen des Mitgliedsbeitrags erhoben wird.

Copyright ©

B. G. Teubner Verlag/GWV Fachverlage GmbH, Wiesbaden 2004. Printed in Germany. Der Verlag B. G. Teubner ist ein Unternehmen von Springer Science+Business Media. Alle Rechte vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages vervielfältigt oder verbreitet werden. Unter dieses Verbot fällt insbesondere die gewerbliche Vervielfältigung per Kopie, die Aufnahme in elektronischen Datenbanken und die Vervielfältigung auf CD-ROM und allen anderen elektronischen Datenträgern.

Satz: Fotosatz Behrens, D-68723 Oftersheim
Druck: Thomas Müntzer, Bad Langensalza

ISSN 0012-0456

Vorwort	91
--------------------------	----

Übersichtsartikel	Historischer Artikel	Buchbesprechungen
--------------------------	-----------------------------	--------------------------

Quantum Mechanics, The Stability of Matter and Quantum Electrodynamics

E. H. Lieb.	93
---------------------	----

Holomorphe Kurven, Hamiltonsche Dynamik und Symplektische Topologie

K. Mohnke.	111
--------------------	-----

Einige topologische Methoden in der nichtlinearen Analysis

M. Väth.	129
------------------	-----

Übersichtsartikel	Historischer Artikel	Buchbesprechungen
--------------------------	-----------------------------	--------------------------

M. Zeman: Inner Models and Large Cardinals

H.-D. Donder.	39
-----------------------	----

A. Holme: Geometry, Our Cultural Heritage

Ch. Birkenhake.	40
-------------------------	----

A. Barvinok: A Course in Convexity

J. M. Wills	41
-----------------------	----

E. Hairer, C. Lubich, G. Wanner: Geometric Numerical Integration

P. Deuflhard.	43
-----------------------	----

In den nächsten Heften erscheinende Arbeiten

J. H. Bruinier: Infinite Products in Number Theory and Geometry
V. Strassen: Komplexität und Geometrie bilinearer Abbildungen

Anschriften der Herausgeber

Prof. Dr. Aloys Krieg, Lehrstuhl A für Mathematik, RWTH Aachen, 52056 Aachen
E-Mail: krieg@mathA.rwth-aachen.de

Prof. Dr. Ursula Gather, Lehrstuhl für Mathematische Statistik und industrielle Anwendungen, Universität Dortmund, 44221 Dortmund
E-Mail: gather@statistik.uni-dortmund.de

Prof. Dr. Ernst Heintze, Institut für Mathematik, Universität Augsburg, 86135 Augsburg
E-Mail: heintze@math.uni-augsburg.de

Prof. Dr. Bernhard Kawohl, Mathematisches Institut, Universität zu Köln, 50923 Köln
E-Mail: kawohl@mi.uni-koeln.de

Prof. Dr. Herbert Lange, Mathematisches Institut, Friedrich-Alexander-Universität Erlangen-Nürnberg, Bismarckstraße 1 $\frac{1}{2}$, 91054 Erlangen
E-Mail: lange@mi.uni-erlangen.de

Prof. Dr. Hans Triebel, Mathematisches Institut, Friedrich-Schiller-Universität, Ernst-Abbe-Platz 1–4, 07740 Jena
E-Mail: triebel@minet.uni-jena.de

Bezugshinweis

Früher erschienene Bände (ab Band 68) des „Jahresberichts der Deutschen Mathematiker-Vereinigung“ können durch den Buchhandel oder den Verlag bezogen werden.

Nachdruck der Bände 41 bis 67 liefert: Swets & Zeitlinger, Heereweg 347 b, POB 810, NL-2160 SZ Lisse/Holland

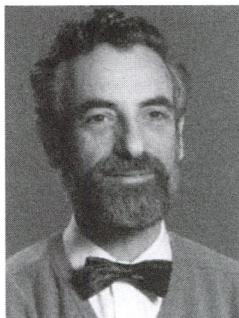
Vorwort

In dem vorliegenden Heft 3 des Jahresberichts können wir Ihnen drei Übersichtsartikel anbieten. Bei der Arbeit von Elliott Lieb handelt es sich um eine Ausarbeitung seines letztjährigen Hauptvortrags auf der DMV-Jahrestagung in Rostock, der der mathematischen Physik zuzuordnen ist.

Bereits auf der Jahrestagung 2002 in Halle hat Klaus Mohnke einen Plenarvortrag aus dem Bereich der Differentialgeometrie gehalten. Eine Ausarbeitung davon finden Sie als zweiten Beitrag.

Der Jahresbericht gibt auch in regelmäßigen Abständen Nachwuchswissenschaftlern die Chance, ihr Arbeitsgebiet zu präsentieren. Martin Väth ist Heisenberg-Stipendiat und hat einen Übersichtsartikel über topologische Methoden in der nichtlinearen Analysis verfasst.

A. Krieg



Quantum Mechanics, The Stability of Matter and Quantum Electrody- namics

Elliott H. Lieb

Abstract

- Mathematics Subject Classification: 81V10, 81V70, 81V45, 81T16
- Keywords and Phrases: Quantum-mechanics, Quantum-electrodynamics, Stability of Matter, Atoms, Radiation-field

Much progress has been made in the last few decades in developing the necessary mathematics for understanding the full implications of the quantum-mechanical many-body problem and why the material world appears to be as stable as it is despite the serious $-1/|x|$ singularity of the Coulomb potential that attracts negative electrons to positive atomic nuclei. Many problems remain, however, especially the understanding of the interaction of matter and the quantized radiation field discovered by Planck in 1900. A short review of some of the main topics is given.

This paper is an extended version of a talk at the DMV Jahrestagung in Rostock, 17 September, 2003 [1], © by Elliott H. Lieb. This article may be reproduced, in its entirety, for noncommercial purposes. Work partially supported by U.S. National Science Foundation grant PHY 0139984-A01.

Eingegangen: 01.01.2004

Departments of Mathematics and Physics, Princeton University,
P.O. Box 708, Princeton, NJ 08544-0708, USA, lieb@princeton.edu

DMV
JAHRESBERICHT
DER DMV
B. G. Teubner 2004

1 Introduction

This paper is a brief survey of the quantum-mechanical many-body problem, especially the question of the interaction of matter with radiation. The quantum-mechanical revolution of the 1920's brought with it many successes, but also a few problems that have yet to be resolved. The realization that there were a few problems with the simple textbook theory surfaced three or four decades ago. Since then some of the mathematical questions have been answered, but some big ones remain. This brief overview might, it is hoped, encourage some mathematicians to look into this fascinating topic.

We begin with a sketch of the topics that will concern us here.

1.1 Triumph of Quantum Mechanics

One of the basic problems of classical physics (after the discovery of the point electron by Thomson and of the (essentially) point nucleus by Rutherford) was the stability of atoms. Why do the electrons in an atom not fall into the nucleus? Quantum mechanics explained this fact. It starts with the classical Hamiltonian of the system (nonrelativistic kinetic energy for the electrons plus Coulomb's law of electrostatic energy among the charged particles). By virtue of the non-commutativity of the kinetic and potential energies in quantum mechanics the stability of an atom – in the sense of a finite lower bound to the energy – was a consequence of the fact that any attempt to make the electrostatic energy very negative would require the localization of an electron close to the nucleus and this, in turn, would result in an even greater, positive, kinetic energy.

Thus, the basic stability problem for an atom was solved by an inequality that says that $\langle 1/|x| \rangle$, the expected value of $1/|x|$, can be made large only at the expense of making the kinetic energy, which is proportional to $\langle p^2 \rangle$, even larger. A fundamental hypothesis of quantum mechanics is that p is represented by the differential operator $-i\hbar\nabla$ with $\hbar = h/2\pi$ and h = Planck's constant. In elementary presentations of the subject it is often said that the mathematical inequality that ensures this fact is the famous uncertainty principle of Heisenberg (proved by Weyl), which states that $\langle p^2 \rangle \langle x^2 \rangle \geq (9/8)\hbar^2$.

While this principle is mathematically rigorous it is actually insufficient for the purpose, as explained, e.g., in [19, 21], and thus gives only a heuristic explanation of the power of quantum mechanics to prevent collapse. A more powerful inequality, such as Sobolev's inequality (9), is needed (see, e.g., [23]). The utility of the latter is made possible by Schrödinger's representation of quantum mechanics (which earlier was a somewhat abstract theory of operators on a Hilbert space) as a theory of differential operators on the space of square integrable functions on \mathbb{R}^3 . The importance of Schrödinger's representation is sometimes underestimated by formalists, but it is of crucial importance because it permits the use of functional analytic methods, especially inequalities such as Sobolev's, which are not easily visible on the Hilbert space level. These methods are essential for the developments reported here.

To summarize, the understanding of the stability of atoms and ordinary matter requires a formulation of quantum mechanics with two ingredients:

- A Hamiltonian formulation in order to have a clear notion of a lowest possible (ground state) energy. Lagrangian formulations, while popular, do not always lend themselves to the identification of that quintessential quantum mechanical notion of a ground state energy. In quantum mechanics a Hamiltonian is not a function on phase space but rather a (pseudo-) differential operator.
- A formulation in terms of concrete function spaces instead of abstract Hilbert spaces so that the power of mathematical analysis can be fully exploited.

1.2 Some Basic Definitions

As usual, we shall denote the lowest energy (eigenvalue) of a quantum mechanical system by E_0 . (More generally, E_0 denotes the infimum of the spectrum of the Hamiltonian H in case this infimum is not an eigenvalue of H or is $-\infty$.) Our intention is to investigate arbitrarily large systems, not just atoms. In general we suppose that the system is composed of N electrons and K nuclei of various kinds. Of course we could include other kinds of particles but N and K will suffice here. $N = 1$ for a hydrogen atom and $N = 10^{23}$ for a mole of hydrogen. We shall use the following terminology for two notions of stability:

$$E_0 > -\infty \quad \text{Stability of the first kind,} \quad (1)$$

$$E_0 > C(N + K) \quad \text{Stability of the second kind} \quad (2)$$

for some constant $C \leq 0$ that is independent of N and K , but which may depend on the physical parameters of the system (such as the electron charge and mass). Usually, $C < 0$, which means that there is a positive binding energy per particle.

Stability of the second kind is absolutely essential if quantum mechanics is going to reproduce some of the basic features of the ordinary material world: The energy of ordinary matter is extensive (i.e., it is proportional to the number of particles), the thermodynamic limit exists (i.e., the $N \rightarrow \infty$ limit exists) and the laws of thermodynamics hold. Bringing two stones together might produce a spark, but not an explosion with a release of energy comparable to the energy in each stone. Stability of the second kind does not guarantee the existence of the thermodynamic limit for the free energy, but it is an essential ingredient [22] [19, Sect. V].

It turns out that stability of the second kind cannot be taken for granted, as Dyson discovered [9]. If Coulomb forces are involved, then *the Pauli exclusion principle is essential*. (This means that the L^2 functions of N variables, $\Psi(x_1, x_2, \dots, x_N)$, $x_i \in \mathbb{R}^3$, is antisymmetric under all transpositions $x_i \leftrightarrow x_j$. Particles, like electrons, whose functions Ψ obey this principle are called *fermions*. Particles whose Ψ functions are symmetric under permutations are called *bosons*.)

Charged bosons are *not stable* because for them $E_0 \sim -N^{7/5}$ (nonrelativistically) and $E_0 = -\infty$ for large, but finite N (relativistically, see Sect. 3.2). While positively charged bosons exist in the form of atomic nuclei, negatively charged, long-lived bosons do not exist in nature. This is a good thing in view of the instability just mentioned.

1.3 The Electromagnetic Field

A second big problem handed down from classical physics was the ‘electromagnetic mass’ of the electron. This poor creature has to drag around an infinite amount of electromagnetic energy that Maxwell burdened it with. Moreover, the electromagnetic field itself is quantized – indeed, that fact alone started the whole revolution [34].

While quantum mechanics accounted for stability with Coulomb forces and Schrödinger led us to think seriously about the ‘wave function of the universe’, physicists shied away from talking about the wave function of the particles in the universe *and* the electromagnetic field in the universe. It is noteworthy that physicists are happy to discuss the quantum mechanical many-body problem with external electromagnetic fields non-perturbatively, but this is rarely done with the quantized field. The quantized field cannot be avoided because it is needed for a correct description of atomic radiation, the laser, etc. However, the interaction of matter with the quantized field is almost always treated perturbatively or else in the context of highly simplified models (e.g., with two-level atoms for lasers).

The quantized electromagnetic field greatly complicates the stability of matter question. It requires, ultimately, mass and charge renormalizations. At present such a complete theory does not exist, but a theory *must* exist because matter exists and because we have strong experimental evidence about the manner in which the electromagnetic field interacts with matter, i.e., we know the essential features of a Hamiltonian that adequately accounts for the low energy processes that exist in every day life. In short, nature tells us that it must be possible to formulate a self-consistent quantum electrodynamics (QED) *non-perturbatively*, (perhaps with an ultraviolet, or high frequency, cutoff off the field at a few MeV). It should not be necessary to have recourse to quantum chromodynamics (QCD) or some other high energy theory to explain ordinary matter.

Physics and other natural sciences are successful because physical phenomena associated with each range of energy and other parameters are explainable to a good, if not perfect, accuracy by an appropriate self-consistent theory. This is true whether it be hydrodynamics, celestial dynamics, statistical mechanics, etc. If low energy physics (atomic and condensed matter physics) is not explainable by a self-consistent, non-perturbative theory on its own level one can speak of an epistemological crisis.

Some readers might say that QED is in good shape. After all, it accurately predicts the outcome of some very high precision experiments (Lamb shift, *g*-factor of the electron). But the theory does not really work well when faced with the problem, which is explored here, of understanding the many-body ($N \approx 10^{23}$) problem and the stable low energy world in which we spend our everyday lives.

1.4 Relativistic Mechanics

When the classical kinetic energy of a particle, $p^2/2m$, is replaced by its relativistic version $\sqrt{p^2c^2 + m^2c^4}$ the stability question becomes much more complicated, as will be seen later. It turns out that even stability of the first kind is not easy to obtain and it depends on the values of the physical constants, notably the fine structure constant

$$(3) \quad \alpha = e^2/\hbar c = 1/137.04 ,$$

where $-e$ is the electric charge of the electron.

For ordinary matter relativistic effects are not dominant but they are noticeable. In large atoms these effects severely change the innermost electrons and this has a noticeable effect on the overall electron density profile. Therefore, some version of relativistic mechanics is needed, which means, presumably, that we must know how to replace $p^2/2m$ by the Dirac operator (see (18)).

The combination of relativistic mechanics plus the electromagnetic field (in addition to the Coulomb interaction) makes the stability problem difficult and uncertain. Major aspects of this problem have been worked out in the last few years (about 35) and that is the subject of this paper.

2 Nonrelativistic Matter without the Magnetic Field

Maxwell's equations define the electric and magnetic fields in terms of potentials. While the equations determine the fields, the potentials are not determined uniquely; the choice of potentials is called the choice of gauge. We work in the 'Coulomb' gauge for the electromagnetic field. Despite the assertion that quantum mechanics and quantum field theory are gauge invariant, it seems to be essential to use this gauge, even though its relativistic covariance is not as transparent as that of the Lorentz gauge. The reason is the following.

The Coulomb gauge is the gauge in which electrostatic part of the interaction of matter with the electromagnetic field is just the conventional Coulomb "action at a distance" potential V_c given by (4) below (in energy units mc^2 and length units the Compton wavelength \hbar/mc). This part of the interaction depends only on the coordinates of the particles and not on their velocities. The dependence of the interaction on velocities, or currents, comes about through the magnetic part of the interaction. Despite appearances, this picture is fully Lorentz invariant (even if it is not gauge invariant).

$$(4) \quad V_c = - \sum_{i=1}^N \sum_{k=1}^K \frac{Z_k}{|\mathbf{x}_i - \mathbf{R}_k|} + \sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} + \sum_{1 \leq k < l \leq K} \frac{Z_k Z_l}{|\mathbf{R}_k - \mathbf{R}_l|} .$$

The first sum is the interaction of the electrons (with dynamical coordinates \mathbf{x}_i) and fixed nuclei located at \mathbf{R}_k of positive charge Z_k times the (negative) electron charge e . The second is the electron-electron repulsion and the third is the nucleus-nucleus repulsion. The nuclei are fixed because they are so massive relative to the electron that their motion is irrelevant. It could be included, however, but it would change nothing essential. Likewise, there is no nuclear structure factor because if it were essential for stability then the size of atoms would be the size of nuclei, about 10^{-13} cm, instead of about 10^{-8} cm, contrary to what is observed.

Although the nuclei are fixed points the constant C in the stability of matter (2) is required to be independent of the \mathbf{R}_k 's. Likewise (1) requires that E_0 have a finite lower bound that is independent of the \mathbf{R}_k 's.

For simplicity of exposition we shall assume here that all the Z_k are identical, i.e., $Z_k = Z$.

The magnetic field, which will be introduced later, is described by a vector potential $\mathbf{A}(x)$ which is a dynamical variable in the Coulomb gauge. The magnetic field is $\mathbf{B} = \text{curl } \mathbf{A}$.

There is a basic physical distinction between electric and magnetic forces which does not seem to be well known, but which motivates this choice of gauge. In electrostatics “like charges repel” while in magnetostatics “like currents attract”. A consequence of these facts is that the correct magnetostatic interaction energy can be obtained by minimizing the energy functional $\frac{1}{2} \int B^2 - \int \mathbf{j} \cdot \mathbf{A}$ with respect to the vector field \mathbf{A} , where \mathbf{j} is the electric current density. The electrostatic energy, on the other hand, *cannot* be obtained by a minimization principle with respect to the field (e.g., minimizing $\frac{1}{2} \int |\nabla \phi|^2 - \int \phi \varrho$ with respect to ϕ).

The Coulomb gauge, which puts in the electrostatics correctly, by hand, so to speak, and allows us to minimize the total energy with respect to the \mathbf{A} field, is the gauge that gives us the correct physics and is consistent with the “quintessential quantum mechanical notion of a ground state energy” mentioned in Sect. 1.1. In any other gauge one would have to look for a critical point of a Hamiltonian rather than a true global minimum.

The type of Hamiltonian that we wish to consider in this section is

$$(5) \quad H_N = T_N + \alpha V_c .$$

Here, T_N is the kinetic energy of the N electrons and has the form

$$(6) \quad T_N = \sum_{i=1}^N T_i ,$$

where T_i acts on the coordinate of the i^{th} electron. The nonrelativistic choice is $T = p^2$ with $\mathbf{p} = -i\nabla$ and $p^2 = -\Delta$ in appropriate units.

2.1 Nonrelativistic Stability for Fermions

The problem of stability of the second kind for nonrelativistic quantum mechanics was recognized in the early days by a few physicists, e.g., Onsager, but not by many. It was not solved until 1967 in one of the most beautiful papers in mathematical physics by Dyson and Lenard [10].

They found that the Pauli principle, i.e., Fermi-Dirac statistics, is essential. Mathematically, this means that the Hilbert space is the subspace of antisymmetric functions, i.e., $\mathcal{H}^{\text{phys}} = \wedge^N L^2(\mathbb{R}^3; \mathbb{C}^2)$. This is how the Pauli principle is interpreted post-Schrödinger; Pauli invented his principle a year earlier, however!

Their value for C in (2) was rather high, about -10^{15} eV (electron volts) for $Z = 1$. (The ground state energy of a hydrogen atom is -13 eV.) The situation was improved later by Thirring and myself [31] to about -20 eV for $Z = 1$ by introducing an inequality that holds only for the kinetic energy of fermions (not bosons) in an arbitrary state Ψ .

$$(7) \quad \langle \Psi, T_N \Psi \rangle \geq (\text{const.}) \int_{\mathbb{R}^3} \varrho_\Psi(\mathbf{x})^{5/3} d^3 \mathbf{x},$$

where ϱ_Ψ is the one-body density in the (normalized) fermionic wave function Ψ (of space and spin) given by an integration over $(N - 1)$ coordinates and N spins as follows.

$$(8) \quad \varrho_\Psi(\mathbf{x}) = N \sum_{\sigma_1, \dots, \sigma_N} \int_{\mathbb{R}^{3(N-1)}} |\Psi(\mathbf{x}, \mathbf{x}_2, \dots, \mathbf{x}_N; \sigma_1, \dots, \sigma_N)|^2 d^3 \mathbf{x}_2 \cdots d^3 \mathbf{x}_N.$$

Inequality (7) allows one simply to reduce the quantum mechanical stability problem to the stability of Thomas-Fermi theory, which was worked out earlier by Simon and myself [30].

The older inequality of Sobolev, mentioned in Sect. 1.1,

$$(9) \quad \langle \Psi, T_N \Psi \rangle \geq (\text{const.}) \left(\int_{\mathbb{R}^3} \varrho_\Psi(\mathbf{x})^3 d^3 \mathbf{x} \right)^{1/3},$$

is not as useful as (7) for the many-body problem because its right side is proportional to N instead of $N^{5/3}$. It is, however, strong enough to yield the stability of a system, like an atom, that has only a few electrons.

It is amazing that from the birth of quantum mechanics until 1967 none of the luminaries of physics had quantified the fact that electrostatics plus the uncertainty principle *do not suffice* for stability of the second kind, and thereby make thermodynamics possible (although they do suffice for the first kind). See Sect. 2.2. It was noted, however, that the Pauli principle was responsible for the large sizes of atoms and bulk matter (see, e.g., [9,10]).

2.2 Nonrelativistic Instability for Bosons

What goes wrong if we have charged bosons instead of fermions? Stability of the first kind (1) holds in the nonrelativistic case, but (2) fails. If we assume the nuclei are infinitely massive, as before, and $N = KZ$ then $E_0 \sim -N^{5/3}$ [10, 20]. To remedy the situation we can let the nuclei have finite mass (e.g., the same mass as the negative particles). Then, as Dyson showed [9], $E_0 \leq -(\text{const.})N^{7/5}$. This calculation was highly non-trivial! Dyson had to construct a variational function with pairing of the Bogolubov type in a rigorous fashion and this took several pages.

Thus, finite nuclear mass improves the situation, but not enough. The question whether $N^{7/5}$ is the correct power law remained open for many years. A lower bound of this type was needed and that was finally obtained in [6].

The results of this Section 2 can be summarized by saying that stability of the hydrogen atom is one thing but stability of many-body physics is something else !

3 Relativistic Kinematics (no magnetic field)

The next step is to try to get some idea of the effects of relativistic kinematics, which means replacing p^2 by $\sqrt{p^2 + 1}$ in non-quantum physics. (Recall that $mc^2 = 1$ in our units.) The simplest way to do this is to substitute $\sqrt{p^2 + 1}$ for T in (6). The Dirac operator will be discussed later on, but for now this choice of T will suffice. Actually, it was Dirac's choice before he discovered his operator and it works well in some cases. For example, Chandrasekhar used it successfully, and accurately, to calculate the collapse of white dwarfs (and later, neutron stars).

Since we are interested only in stability, we may, and shall, substitute $|\mathbf{p}| = \sqrt{-\Delta}$ for T . The error thus introduced is bounded by a constant times N since $|\mathbf{p}| < \sqrt{p^2 + 1} < |\mathbf{p}| + 1$ (as an operator inequality). Our Hamiltonian is now $H_N = \sum_{i=1}^N |\mathbf{p}_i| + \alpha V_c$.

3.1 One-Electron Atom

The touchstone of quantum mechanics is the Hamiltonian for 'hydrogen' which is, in our case,

$$(10) \quad H = |\mathbf{p}| - Z\alpha/|\mathbf{x}| = \sqrt{-\Delta} - Z\alpha/|\mathbf{x}| .$$

It is well known (also to Dirac) that the analogous operator with $|\mathbf{p}|$ replaced by the Dirac operator (18) ceases to make sense when $Z\alpha > 1$. Something similar happens for (10).

$$(11) \quad E_0 = \begin{cases} 0 & \text{if } Z\alpha \leq 2/\pi; \\ -\infty & \text{if } Z\alpha > 2/\pi . \end{cases}$$

The reason for this behavior is that both $|\mathbf{p}|$ and $|\mathbf{x}|^{-1}$ scale in the same way. Either the first term in (10) wins or the second does.

A result similar to (11) was obtained in [11] for the free Dirac operator $D(0)$ in place of $|\mathbf{p}|$, but with the wave function Ψ restricted to lie in the positive spectral subspace of $D(0)$. Here, the critical value is $\alpha Z \leq (4\pi)/(4 + \pi^2) > 2/\pi$.

The moral to be drawn from this is that relativistic kinematics plus quantum mechanics is a 'critical' theory (in the mathematical sense). This fact will plague any relativistic theory of electrons and the electromagnetic field – primitive or sophisticated.

3.2 Many Electrons and Nuclei

When there are many electrons is it true that the condition $Z\alpha \leq const.$ is the only one that has to be considered? The answer is no! One also needs the condition that α itself must be small, regardless of how small Z might be. This fact can be called a 'discovery' but actually it is an overdue realization of some basic physical ideas. It should have been realized shortly after Dirac's theory in 1927, but it does not seem to have been noted until 1983 [8].

The underlying physical heuristics is the following. With α fixed, suppose $Z\alpha = 10^{-6} \ll 1$, so that an atom is stable, but suppose that we have 2×10^6 such nuclei. By bringing them together at a common point we will have a nucleus with $Z\alpha = 2$ and one electron suffices to cause collapse into it. Then (1) fails. What prevents this from happening, presumably, is the nucleus-nucleus repulsion energy which goes to $+\infty$ as the nuclei come together. But this repulsion energy is proportional to $(Z\alpha)^2/\alpha$ and, therefore, if we regard $Z\alpha$ as fixed we see that $1/\alpha$ must be large enough in order to prevent collapse.

Whether or not the reader believes this argument, the mathematical fact is that there is a fixed, finite number $\alpha_c \leq 2.72$ ([32]) so that when $\alpha > \alpha_c$ (1) fails for *every* positive Z and for every $N \geq 1$ (with or without the Pauli principle).

The open question was whether (2) holds for *all* N and K if $Z\alpha$ and α are both small enough. The breakthrough was due to Conlon [5] who proved (2), for fermions, if $Z = 1$ and $\alpha < 10^{-200}$. The situation was improved by Fefferman and de la Llave [13] to $Z = 1$ and $\alpha < 0.16$. Finally, the expected correct condition $Z\alpha \leq 2/\pi$ and $\alpha < 1/94$ was obtained in [32]. (This paper contains a detailed history up to 1988.) The situation was further improved in [27]. The multi-particle version of the use of the free Dirac operator, as in Sect. 3.1, was treated in [18].

Finally, it has to be noted that charged bosons are *always* unstable of the first kind (not merely the second kind, as in the nonrelativistic case) for *every* choice of $Z > 0, \alpha > 0$. E.g., there is instability if $Z^{2/3}\alpha N^{1/3} > 36$ ([32]).

We are indeed fortunate that there are no stable, negatively charged bosons.

4 Interaction of Matter with Classical Magnetic Fields

The magnetic field \mathbf{B} is defined by a vector potential $\mathbf{A}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x}) = \text{curl } \mathbf{A}(\mathbf{x})$. In this section we take a first step (warmup exercise) by regarding \mathbf{A} as classical, but indeterminate, and we introduce the classical field energy

$$(12) \quad H_f = \frac{1}{8\pi} \int_{\mathbb{R}^3} B(\mathbf{x})^2 d^3\mathbf{x}.$$

The Hamiltonian is now

$$(13) \quad H_N(\mathbf{A}) = T_N(\mathbf{A}) + \alpha V_c + H_f,$$

in which the kinetic energy operator has the form (6) but depends on \mathbf{A} . We now define E_0 to be the infimum of $\langle \Psi, H_N(\mathbf{A}) \Psi \rangle$ both with respect to Ψ and with respect to \mathbf{A} .

4.1 Nonrelativistic Matter with Magnetic Field

The simplest situation is merely ‘minimal coupling’ without spin, namely,

$$(14) \quad T(\mathbf{A}) = |\mathbf{p} + \sqrt{\alpha} \mathbf{A}(\mathbf{x})|^2$$

This choice does not change any of our previous results qualitatively. The field energy is

not needed for stability. On the one-particle level, we have the ‘diamagnetic inequality’ $\langle \phi, |\mathbf{p} + \mathbf{A}(\mathbf{x})|^2 \phi \rangle \geq \langle |\phi|, p^2 |\phi| \rangle$. The same holds for $|\mathbf{p} + \mathbf{A}(\mathbf{x})|$ and $|\mathbf{p}|$. More importantly, inequality (7) for fermions continues to hold (with the same constant) with $T(\mathbf{A})$ in place of p^2 . (There is an inequality similar to (7) for $|\mathbf{p}|$, with $5/3$ replaced by $4/3$, which also continues to hold with minimal substitution [7].)

The situation gets much more interesting if spin is included. This takes us a bit closer to the relativistic case. The kinetic energy operator is the Pauli operator

$$(15) \quad T^P(\mathbf{A}) = |\mathbf{p} + \sqrt{\alpha} \mathbf{A}(\mathbf{x})|^2 + \sqrt{\alpha} \mathbf{B}(\mathbf{x}) \cdot \boldsymbol{\sigma},$$

where $\boldsymbol{\sigma}$ is the vector of 2×2 Pauli spin matrices and $L^2(\mathbb{R}^3)$ is replaced by $L^2(\mathbb{R}^3; \mathbb{C}^3)$.

4.1.1 One-Electron Atom

The stability problem with $T^P(\mathbf{A})$ is complicated, even for a one-electron atom. Without the field energy H_f the Hamiltonian is unbounded below. (For fixed \mathbf{A} it is bounded but the energy tends to $-\infty$ like $-(\log B)^2$ for a homogeneous field [2].) The field energy saves the day, but the result is surprising [14] (recall that we must minimize the energy with respect to Ψ and \mathbf{A}):

$$(16) \quad |\mathbf{p} + \sqrt{\alpha} \mathbf{A}(\mathbf{x})|^2 + \sqrt{\alpha} \mathbf{B}(\mathbf{x}) \cdot \boldsymbol{\sigma} - Z\alpha/|\mathbf{x}| + H_f$$

is bounded below if and only if $Z\alpha^2 \leq C$, where C is some constant that can be bounded as $1 < C < 9\pi^2/8$.

The proof of instability [33] is difficult and requires the construction of a zero mode (soliton) for the Pauli operator, i.e., a finite energy magnetic field and a *square integrable* ψ such that

$$(17) \quad T^P(\mathbf{A})\psi = 0.$$

The usual kinetic energy $|\mathbf{p} + \mathbf{A}(\mathbf{x})|^2$ has no such zero mode for any \mathbf{A} , even when 0 is the bottom of its spectrum.

The original magnetic field [33] that did the job in (17) is independently interesting, geometrically (many others have been found since then).

$$\mathbf{B}(\mathbf{x}) = \frac{12}{(1 + |\mathbf{x}|^2)^3} [(1 - x^2)\mathbf{w} + 2(\mathbf{w} \cdot \mathbf{x})\mathbf{x} + 2\mathbf{w} \wedge \mathbf{x}]$$

with $|\mathbf{w}| = 1$. The field lines of this magnetic field form a family of curves, which, when stereographically projected onto the 3-dimensional unit sphere, become the great circles in what is known as the Hopf fibration.

Thus, we begin to see that nonrelativistic matter with magnetic fields behaves like relativistic matter without fields – to some extent.

The moral of this story is that a magnetic field, which we might think of as possibly self-generated, can cause an electron to fall into the nucleus. The uncertainty principle cannot prevent this, not even for an atom!

4.1.2 Many Electrons and Many Nuclei

In analogy with the relativistic (no magnetic field) case, we can see that stability of the first kind fails if $Z\alpha^2$ or α is too large. The heuristic reasoning is the same and the proof is similar.

We can also hope that stability of the second kind holds if both $Z\alpha^2$ and α are small enough. The problem is complicated by the fact that it is the field energy H_f that will prevent collapse, but there there is only one field energy while there are $N \gg 1$ electrons.

The hope was finally realized, however. Fefferman [12] proved stability of the second kind for $H_N(\mathbf{A})$ with the Pauli $T^P(\mathbf{A})$ for $Z = 1$ and “ α sufficiently small”. A few months later it was proved [28] for $Z\alpha^2 \leq 0.04$ and $\alpha \leq 0.06$. With $\alpha = 1/137$ this amounts to $Z \leq 1050$. This very large Z region of stability is comforting because it means that perturbation theory (in \mathbf{A}) can be reliably used for this particular problem.

Using the results in [28], Bugliaro, Fröhlich and Graf [3] proved stability of the same nonrelativistic Hamiltonian – but with an ultraviolet cutoff, quantized magnetic field whose field energy is described below. (Note: No cutoffs are needed for classical fields.)

There is also the very important work of Bach, Fröhlich, and Sigal [4] who showed that this nonrelativistic Hamiltonian with ultraviolet cutoff, quantized field and with sufficiently small values of the parameters has other properties that one expects. E.g., the excited states of atoms dissolve into resonances and only the ground state is stable. The infrared singularity notwithstanding, the ground state actually exists (the bottom of the spectrum is an eigenvalue); this was shown in [4] for small parameters and in [15], [26] for all values of the parameters. (See Sect. 7.)

5 Relativity Plus Magnetic Fields

As a next step in our efforts to understand QED and the many-body problem we introduce relativity theory along with the classical magnetic field.

5.1 Relativity Plus Classical Magnetic Fields

Originally, Dirac and others thought of replacing $T^P(\mathbf{A})$ by $\sqrt{T^P(\mathbf{A}) + 1}$ but this was not successful mathematically and does not seem to conform to experiment. Consequently, we introduce the Dirac operator for T in (6), (13)

$$(18) \quad D(\mathbf{A}) = \alpha \cdot \mathbf{p} + \sqrt{\alpha} \alpha \cdot \mathbf{A}(\mathbf{x}) + \beta m ,$$

where α and β denote the 4×4 Dirac matrices and $\sqrt{\alpha}$ is the electron charge as before. (This notation of α and α is historical and is not mine.) The Hilbert space for N electrons is now changed to

$$(19) \quad \mathcal{H} = \wedge^N L^2(\mathbb{R}^3; \mathbb{C}^4) .$$

The well known problem with $D(\mathbf{A})$ is that it is unbounded below, and so we cannot hope to have stability of the first kind, even with $Z = 0$. Let us imitate QED (but with-

out pair production or renormalization) by restricting the electron wave function to lie in the positive spectral subspace of a Dirac operator.

Which Dirac operator?

There are two natural operators in the problem. One is $D(0)$, the free Dirac operator. The other is $D(\mathbf{A})$ that is used in the Hamiltonian. In almost all formulations of QED the electron is defined by the positive spectral subspace of $D(0)$. Thus, we can define

$$(20) \quad \mathcal{H}^{\text{phys}} = P^+ \mathcal{H} = \prod_{i=1}^N \pi_i \mathcal{H},$$

where $P^+ = \prod_{i=1}^N \pi_i$, and π_i is the projector of onto the positive spectral subspace of $D_i(0) = \boldsymbol{\alpha} \cdot \mathbf{p}_i + \beta m$, the free Dirac operator for the i^{th} electron. We then restrict the allowed wave functions in the variational principle to those Ψ satisfying

$$(21) \quad \Psi = P^+ \Psi \quad \text{i.e., } \Psi \in \mathcal{H}^{\text{phys}}.$$

Another way to say this is that we replace the Hamiltonian (13) by $P^+ H_N P^+$ on \mathcal{H} and look for the bottom of its spectrum.

It turns out that this prescription leads to disaster! While the use of $D(0)$ makes sense for an atom, it fails miserably for the many-fermion problem, as discovered in [29] and refined in [16]. The result is:

For all $\alpha > 0$ in (18) (with or without the Coulomb term αV_c) one can find N large enough so that $E_0 = -\infty$.

In other words, the term $\sqrt{\alpha} \boldsymbol{\alpha} \cdot \mathbf{A}$ in the Dirac operator can cause an instability that the field energy cannot prevent.

It turns out, however, that the situation is saved if one uses the positive spectral subspace of the Dirac operator $D(\mathbf{A})$ to define an electron. (This makes the concept of an electron \mathbf{A} dependent, but when we make the vector potential into a dynamical quantity in the next section, this will be less peculiar since there will be no definite vector potential but only a fluctuating quantity.) The definition of the physical Hilbert space is as in (20) but with π_i being the projector onto the positive subspace of the full Dirac operator $D_i(\mathbf{A}) = \boldsymbol{\alpha} \cdot \mathbf{p}_i + \sqrt{\alpha} \boldsymbol{\alpha} \cdot \mathbf{A}(\mathbf{x}_i) + \beta m$. Note that these π_i projectors commute with each other and hence their product P^+ is a projector.

The result [29] for this model ((13) with the Dirac operator and the restriction to the positive spectral subspace of $D(\mathbf{A})$) is reminiscent of the situations we have encountered before:

If α and Z are small enough stability of the second kind holds for this model.

Typical stability values that are rigorously established [29] are $Z \leq 56$ with $\alpha = 1/137$ or $\alpha \leq 1/8.2$ with $Z = 1$.

6 Quantized Electromagnetic Fields

Let us now try to analyze some of the problems connected with the quantization of the electromagnetic field. The great discovery of Max Planck [34], which was the first step in the new quantum theory, was that the energy of the electromagnetic field came in

quantized units. The energy unit of electromagnetic waves of frequency ν is $\hbar\nu$, and in terms of wave number k (i.e., the wave is proportional to $\exp(ik \cdot x)$) it is $\hbar c |k|$ since $2\pi\nu/|k| = c$ = speed of light.

We begin with the problem of generalizing the results in the previous subsection to the quantized field.

6.1 Relativity Plus Quantized Magnetic Field

The obvious next step is to try to imitate the strategy of Sect. 5.1 but with the quantized A field. This was done in [24]. The quantized A field is described by an operator-valued Fourier transform as

$$(22) \quad \mathbf{A}(\mathbf{x}) = \frac{1}{2\pi} \sum_{\lambda=1}^2 \int_{|\mathbf{k}| \leq \Lambda} \frac{\boldsymbol{\epsilon}_\lambda(\mathbf{k})}{\sqrt{|\mathbf{k}|}} \left[a_\lambda(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} + a_\lambda^*(\mathbf{k}) e^{-i\mathbf{k} \cdot \mathbf{x}} \right] d^3\mathbf{k},$$

where Λ is the ultraviolet cutoff on the wave-numbers $|\mathbf{k}|$. The operators a_λ, a_λ^* satisfy the usual canonical commutation relations

$$(23) \quad [a_\lambda(\mathbf{k}), a_\nu^*(\mathbf{q})] = \delta(\mathbf{k} - \mathbf{q}) \delta_{\lambda,\nu}, \quad [a_\lambda(\mathbf{k}), a_\nu(\mathbf{q})] = 0, \quad \text{etc}$$

and the vectors $\boldsymbol{\epsilon}_\lambda(\mathbf{k})$ are two orthonormal polarization vectors perpendicular to \mathbf{k} and to each other.

The field energy H_f is now given by a normal-ordered version of (12)

$$(24) \quad H_f = \sum_{\lambda=1,2} \int_{\mathbb{R}^3} |\mathbf{k}| a_\lambda^*(\mathbf{k}) a_\lambda(\mathbf{k}) d^3\mathbf{k}$$

The Dirac operator is the same as before, (18). Note that $D_i(\mathbf{A})$ and $D_j(\mathbf{A})$ still commute with each other (since $\mathbf{A}(\mathbf{x})$ commutes with $\mathbf{A}(\mathbf{y})$). This is important because it allows us to imitate Sect. 5.1.

In analogy with (19) we define

$$(25) \quad \mathcal{H} = \wedge^N L^2(\mathbb{R}^3; \mathbb{C}^4) \otimes \mathcal{F},$$

where \mathcal{F} is the Fock space for the photon field. We can then define the *physical* Hilbert space as before

$$(26) \quad \mathcal{H}^{\text{phys}} = \Pi \mathcal{H} = \prod_{i=1}^N \pi_i \mathcal{H},$$

where the projectors π_i project onto the positive spectral subspace of either $D_i(0)$ or $D_i(\mathbf{A})$.

Perhaps not surprisingly, the former case leads to catastrophe, as before. This is so, even with the ultraviolet cutoff, which we did not have in Sect. 5.1. Because of the cutoff the catastrophe is milder and involves instability of the second kind instead of the first kind. This result relies on a coherent state construction in [16].

The latter case (use of $D(\mathbf{A})$ to define an electron) leads to stability of the second kind if Z and α are not too large. Otherwise, there is instability of the first kind. The rigorous estimates are comparable to the ones in Sect. 5.1.

Clearly, many things have yet to be done to understand the stability of matter in the context of QED. Renormalization and pair production have to be included, for example.

The results of this section suggest, however, that a significant change in the Hilbert space structure of QED might be necessary. We see that it does not seem possible to keep to the current view that the Hilbert space is a simple tensor product of a space for the electrons and a Fock space for the photons. That leads to instability for many particles (or large charge, if the idea of 'particle' is unacceptable). The 'bare' electron is not really a good physical concept and one must think of the electron as always accompanied by its electromagnetic field. Matter and the photon field are inextricably linked in the Hilbert space $\mathcal{H}^{\text{phys}}$.

The following tables [24] summarize some of the results of this and the previous sections

Electrons defined by projection onto the positive subspace of $D(0)$, the free Dirac operator

	Classical or quantized field without cutoff Λ $\alpha > 0$ but arbitrarily small.	Classical or quantized field with cutoff Λ $\alpha > 0$ but arbitrarily small.
Without Coulomb potential αV_c	Instability of the first kind	Instability of the second kind
With Coulomb potential αV_c	Instability of the first kind	Instability of the second kind

Electrons defined by projection onto the positive subspace of $D(\mathbf{A})$, the Dirac operator with field

	Classical field with or without cutoff Λ or quantized field with cutoff Λ
Without Coulomb potential αV_c	The Hamiltonian is positive
With Coulomb potential αV_c	Instability of the first kind when either α or $Z\alpha$ is too large
	Stability of the second kind when both α and $Z\alpha$ are small enough

6.2 Mass Renormalization

In both classical and quantum electrodynamics there is a problem of mass renormalization. This means that when a charge is accelerated its accompanying electromagnetic field is also accelerated and acts like an additional mass. The ‘bare mass’ of the particle (which is the mass that appears in the Hamiltonian) must be chosen so that the final, physical mass (as measured in experiments) agrees with the physically measured value.

For a point particle, the additional mass is infinity, classically. For QED it is also infinite, but the divergence is less rapid as the radius of the charge goes to zero. In any case, with a finite ultraviolet cutoff Λ the additional mass is finite, but it is far from clear that, for each $\Lambda > 0$ one can adjust the bare mass (while keeping it positive) to give the correct physical mass. Opinions differ on this point and very little is known rigorously about the problem outside of perturbation theory. See [17].

There are two ways to define mass renormalization. Take one particle ($N = 1$) and then either

1. Find the bottom of the spectrum of $T + H_f$ under the condition that the total momentum of particle plus field is \mathbf{p} . Call it $E(\mathbf{p})$ and write, for small \mathbf{p} ,

$$E(\mathbf{p}) = E(\mathbf{p} = 0) + p^2/2m_{\text{physical}}$$

or else

2. Compute the binding energy of hydrogen ($N = 1, K = 1, Z = 1$). Call it E_0 and set

$$E_0 = m_{\text{physical}}c^2\alpha^2/2\hbar^2$$

The first way is the usual one; the second is motivated by the earliest experiment in quantum mechanics. These two definitions are not the same. In any case, we [25] can now obtain non-trivial bounds on the binding energy (in the context of the Schrödinger Hamiltonian or the Pauli Hamiltonian interacting with the quantized field) and thereby get some bounds on the renormalized mass using definition 2. For large cutoff Λ , these bounds differ in their Λ dependence from what might be expected from perturbation theory.

7 Existence of Atoms in Non-relativistic QED

One of the most recent topics concerns the seemingly trivial question of the existence of atoms. In some sense this question is the opposite of the stability of matter question.

The Hamiltonian we shall use to describe an atom or molecule with N electrons is

$$(27) \quad H_N = \sum_{i=1}^N T_i^P(\mathbf{A}) + \alpha V_c + H_f$$

where $T_i^P(\mathbf{A})$ is the Pauli kinetic energy operator (15), but \mathbf{A} is the quantized magnetic field given by (22), and H_f is the energy of the quantized field given by (24). As before, V_c is the Coulomb potential (4) of some fixed nuclei whose total nuclear charge is denoted by $Z = \sum Z_j$.

To show the existence of stable atoms we need to establish two things about H_N .

1. The ground state energy (bottom of the spectrum) of H_N is lower than that of $H_{N'}$ i.e., of a system with $N' < N$ electrons (with the remaining $N - N'$ electrons being allowed to escape to infinity). This is called *the binding condition*.
2. The bottom of the spectrum of H_N is actually an eigenvalue, i.e., Schrödinger's equation has a square integrable solution with $E =$ the bottom of the spectrum.

In the case of the Schrödinger equation without the field, problem 1. was solved by Zhislin in 1960 for the case $N < Z + 1$, which includes the neutral molecule. He did this by using a localization technique, whose positive localization energy (r^{-2}) is more than offset by the Coulomb attraction ($-r^{-1}$) of a positively charged system ($Z - N'$) to a negatively charged electron. The existence of the ground state (problem 2.) follows from standard arguments because in this case the bottom of the spectrum is negative while the bottom of the essential spectrum (which, in this case, is the bottom of the continuum) starts at zero. Thus, there is a gap in the spectrum and the technique of taking weak limits easily yields a non-zero eigenfunction [23].

When we turn on the interaction with the quantized magnetic field the situation changes significantly. One major difference is that the bottom of the essential spectrum is now the bottom of the spectrum because we can always create photons with arbitrarily small energy (recall that the energy of a photon with momentum k is $|k|$). Therefore, if a ground state exists it necessarily lies at the bottom of the essential spectrum and is not isolated. Eigenvalues in the continuum are notoriously difficult to handle, even for the simple Schrödinger operator.

A second major difference is that it is necessary to localize the A field as well as the electrons. This localization costs an energy r^{-1} , not r^{-2} as before, essentially because the field energy is proportional to $|k|$ instead of k^2 . Thus, the field localization competes with the Coulomb attraction.

Problems 1. and 2. were solved in [4] under the condition that α and Λ are small enough.

The first general result, valid for all values of the various constants, was in [15], where it was shown that 2. holds whenever 1. holds.

Finally, 1. was shown to hold for all values of the constants [26] under the same natural condition as Zhislin's, i.e., $N < Z + 1$.

References

- [1] This paper is an updated version of E.H. Lieb: *The Stability of Matter and Quantum Electrodynamics*, Milan Jour. Math. 71, 199–217 (2003). An earlier version appears in the proceedings of the Werner Heisenberg Centennial, December, 2001 in Munich. Springer Verlag (in press).
- [2] J. Avron, I. Herbst, B. Simon: *Schrödinger operators with magnetic fields III*, Commun. Math. Phys. 79, 529–572 (1981).
- [3] L. Bugliaro, J. Fröhlich, G. M. Graf: *Stability of quantum electrodynamics with nonrelativistic matter*, Phys. Rev. Lett. 77, 3494–3497 (1996).
- [4] V. Bach, J. Fröhlich, I. M. Sigal: *Spectral analysis for systems of atoms and molecules coupled to the quantized radiation field*. Comm. Math. Phys. 207, 249–290 (1999).

- [5] J. G. Conlon: *The ground state of a classical gas*, Commun. Math. Phys. **94**, 439–458 (1984).
- [6] J. G. Conlon, E. H. Lieb, H.-T. Yau: *The $N^{7/5}$ law for charged bosons*, Commun. Math. Phys. **116**, 417–448 (1988).
- [7] I. Daubechies: *An uncertainty principle for fermions with generalized kinetic energy*, Commun. Math. Phys. **90**, 511–520 (1983).
- [8] I. Daubechies, E. H. Lieb: *One electron relativistic molecules with Coulomb interaction*, Commun. Math. Phys. **90**, 497–510 (1983).
- [9] F. J. Dyson: *Ground state energy of a finite system of charged particles*, J. Math. Phys. **8**, 1538–1545 (1967).
- [10] F. J. Dyson, A. Lenard: *Stability of matter I and II*, J. Math. Phys. **8**, 423–434 (1967), **9**, 1538–1545 (1968).
- [11] W. D. Evans, P. P. Perry, H. Siedentop: *The spectrum of relativistic one-electron atoms according to Bethe and Salpeter*, Commun. Math. Phys. **178**, 733–746 (1996).
- [12] C. Fefferman: *Stability of Coulomb systems in a magnetic field*, Proc. Nat. Acad. Sci. USA, **92**, 5006–5007 (1995).
- [13] C. Fefferman, R. de la Llave: *Relativistic stability of matter. I*. Rev. Mat. Iberoamericana **2**, 119–213 (1986).
- [14] J. Fröhlich, E. H. Lieb, M. Loss: *Stability of Coulomb systems with magnetic fields I. The one-electron atom*, Commun. Math. Phys. **104**, 251–270 (1986). See also E. H. Lieb, M. Loss: *Stability of Coulomb systems with magnetic Fields II. The many-electron atom and the one-electron molecule*, Commun. Math. Phys. **104**, 271–282 (1986).
- [15] M. Griesemer, E. H. Lieb, M. Loss: *Ground states in non-relativistic quantum electrodynamics*, Invent. Math. **145**, 557–595 (2001).
- [16] M. Griesemer, C. Tix: *Instability of pseudo-relativistic model of matter with self-generated magnetic field*, J. Math. Phys. **40**, 1780–1791 (1999).
- [17] C. Hainzl and R. Seiringer: *Mass renormalization and energy level shifts in non-relativistic QED*, Adv. Theor. Math. Phys. **6**, 847–871 (2002).
- [18] G. Hoever, H. Siedentop: *The Brown-Ravenhall operator*, Math. Phys. Electronic Jour. **5**, no. 6 (1999).
- [19] E. H. Lieb: *The stability of matter*, Rev. Mod. Phys. **48**, 553–569 (1976).
- [20] E. H. Lieb: *The $N^{5/3}$ law for bosons*, Phys. Lett. **70A**, 71–73 (1979).
- [21] E. H. Lieb: *The stability of matter: From atoms to stars*, Bull. Amer. Math. Soc. **22**, 1–49 (1990).
- [22] E. H. Lieb, J.L. Lebowitz: *The existence of thermodynamics for real matter with Coulomb forces*, Phys. Rev. Lett. **22**, 631–634 (1969).
- [23] E. H. Lieb, M. Loss: *Analysis*, American Mathematical Society (1997).
- [24] E. H. Lieb, M. Loss: *Stability of a Model of Relativistic Quantum Electrodynamics*, Commun. Math. Phys. **228**, 561–588 (2002). arXiv math-ph/0109002, mp_arc 01-315.
- [25] E. H. Lieb, M. Loss: *A Bound on Binding Energies and Mass Renormalization in Models of Quantum Electrodynamics*, J. Stat. Phys. **108**, 1057–1069 (2002). arXiv math-ph/0110027.
- [26] E. H. Lieb, M. Loss: *Existence of Atoms and Molecules in Non-relativistic Quantum Electrodynamics*, Adv. Theor. Math. Phys. **7**, 667–710 (2003). arXiv math-ph/0307046.
- [27] E. H. Lieb, M. Loss, H. Siedentop: *Stability of relativistic matter via Thomas-Fermi theory*, Helv. Phys. Acta **69**, 974–984 (1996).
- [28] E. H. Lieb, M. Loss, J. P. Solovej: *Stability of matter in magnetic fields*, Phys. Rev. Lett. **75**, 985–989 (1995).
- [29] E. H. Lieb, H. Siedentop, J. P. Solovej: *Stability and instability of relativistic electrons in magnetic fields*, J. Stat. Phys. **89**, 37–59 (1997). See also *Stability of relativistic matter with magnetic fields*, Phys. Rev. Lett. **79**, 1785–1788 (1997).
- [30] E. H. Lieb, B. Simon: *Thomas-Fermi theory revisited*, Phys. Rev. Lett. **31**, 681–683 (1973).
- [31] E. H. Lieb, W. Thirring: *Bound for the kinetic energy of fermions which proves the stability of matter*, Phys. Rev. Lett. **35**, 687–689 (1975). Errata **35**, 1116 (1975).

- [32] E. H. Lieb, H.-T. Yau: *The stability and instability of relativistic matter*, Commun. Math. Phys. **118**, 177–213 (1988). See also *Many-body stability implies a bound on the fine structure constant*, Phys. Rev. Lett. **61**, 1695–1697 (1988).
- [33] M. Loss, H.-T. Yau: *Stability of Coulomb systems with magnetic fields III. Zero energy bound states of the Pauli operator*, Commun. Math. Phys. **104**, 283–290 (1986).
- [34] M. Planck: *Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum*, Verhandlung der Deutschen Physikalischen Gesellschaft **2**, 237–245 (1900).
- [35] G. Zhislin: *A study of the spectrum of the Schrödinger operator for a system of several particles*, Trudy Moscov Mat. Obsc. **9**, 81–120 (1960).



Holomorphe Kurven, Hamiltonsche Dynamik und symplektische Topologie

Klaus Mohnke

Abstract

- Keywords and Phrases: symplectic manifolds, Lagrangian embeddings, Hamiltonian dynamics, pseudoholomorphic curves
- Mathematics Subject Classification: 57R17; 53D12,40,45

Ziel dieses Artikels ist es zu zeigen, wie man ein altes Problem der symplektischen Topologie mit Hilfe neuer Techniken, die auf Eigenschaften gewisser nicht kompakter pseudoholomorpher Kurven beruhen, effektiv beweisen kann: Es gibt keine Lagrange Einbettung der Kleinschen Flasche in den \mathbb{C}^2 . Die allgemeine Fragestellung wird erläutert und eine kurze Einführung in die entsprechenden Invarianten gegeben.

We demonstrate how to prove an old problem of symplectic topology very effectively by means of new techniques which rely on properties of certain non-compact holomorphic curves: there exists no Lagrangian embedding of the Klein bottle into \mathbb{C}^2 . The general problem is outlined. We give a short introduction into the corresponding invariants.

Eingegangen: 07.04.2003, in revidierter Form am 22.12.2003

Klaus Mohnke, Humboldt-Universität zu Berlin, Institut für Mathematik,
10099 Berlin, mohnke@mathematik.hu-berlin.de

DMV
JAHRESBERICHT
DER DMV
© B. G. Teubner 2004

Symplektische Formen sind die grundlegenden Strukturinvarianten der klassischen Mechanik. Der Phasenraum, d. h. der Raum der Orts- und Impulskoordinaten (p_i, q_i) , ist mit einer kanonischen symplektischen Form versehen (siehe 2.1., Beispiel 5). Durch sie wird jeder differenzierbaren Funktion H das sogenannte Hamiltonsche Vektorfeld X_H zugeordnet. Insbesondere beschreibt das Hamiltonsche Feld der Totalenergie die infinitesimale Veränderung des mechanischen Systems, gegeben durch die bekannten Hamiltonschen Gleichungen:

$$\dot{p}_i = -H_{q_i}$$

$$\dot{q}_i = H_{p_i}$$

(siehe Arnolds „Klassiker“ [1] – auch für die folgende Diskussion). Der Fluss lässt die Energie H invariant. Diese Eigenschaft gilt nicht mehr für nicht-autonome Systeme, dass heißt solchen, bei denen die Hamiltonfunktion zeitabhängig ist. Die Diffeomorphismen der Familie, die man als Lösung zeitabhängiger Hamiltonscher Gleichungen erhält, nennt man Hamiltonsche Diffeomorphismen. Sie lassen immer noch die symplektische Struktur invariant. Allgemein nennt man Diffeomorphismen mit dieser Eigenschaft Symplektomorphismen. Systeme mit periodischer Hamiltonfunktion erhält man zum Beispiel bei der Untersuchung von Modellen der Himmelsmechanik, an denen mehr als zwei Himmelskörper beteiligt sind. Das autonome System mit zwei Himmelskörpern (Sonne–Erde oder Erde–Mond) und deren gegenseitige Anziehung ist integrierbar, und seine Lösungen werden durch Kepplers Gesetze beschrieben. Ein ganz brauchbarer Ansatz, um die Auswirkung zu studieren, die andere Himmelskörper auf dieses System haben, ist diese als periodische Störung der autonomen Hamiltonfunktion zu beschreiben, wobei die Rückwirkung des Systems auf die anderen Himmelskörper vernachlässigt wird. Poincaré stieß beim Studium solcher Systeme auf den Fakt, dass flächentreue, orientationserhaltende Diffeomorphismen des Annulus, die den äußeren Rand im positiven Sinn und den inneren im negativen Sinn drehen, wenigstens zwei Fixpunkte haben müssen. Birkhoff bewies dies mit strikt zwei-dimensionalen Methoden. Arnold entdeckte schließlich, dass das Poincaré–Birkhoff–Theorem ein Phänomen symplektischer Diffeomorphismen ist. Er machte darum seine berühmt gewordene Vermutung über untere Schranken für die Zahl von Fixpunkten Hamiltonscher Diffeomorphismen. In Hinblick auf periodische Hamiltonsche Systeme sind Fixpunkte aber gerade deren periodische Lösungen. Arnolds Vermutungen waren der Anlass für die Entwicklung vieler neuer Methoden aus der Variationsrechnung (z. B. die Entdeckung der Kapazitäten), aus der komplexen Analysis und der Theorie partieller Differentialgleichungen (z. B. pseudoholomorphe Kurven und Floer-Theorie) für das Studium der Hamiltonschen Dynamik.

Symplektische Strukturen findet man auch auf vielen Modulräumen der Lösungen von Feldgleichungen. Oft begegnet man dabei auf natürliche Weise Lagrangen Untermannigfaltigkeiten. Dies sind Untermannigfaltigkeiten halber Dimension, auf denen die symplektische Form verschwindet. Viele Eigenschaften von Objekten der symplektischen Geometrie können direkt durch Eigenschaften Lagranger Untermannigfaltigkeiten ausgedrückt werden. Beispielsweise ist ein Symplektomorphismus einer symplektischen Mannigfaltigkeit (M, ω) dadurch charakterisiert, dass sein Graph Lagrange bzgl. der symplektischen Struktur $\omega \oplus (-\omega)$ auf dem Produkt $M \times M$ ist. Der Null-

schnitt, jede Faser im Kotangentialbündel einer Mannigfaltigkeit sowie der Graph einer geschlossenen 1-Form darauf sind weitere wichtige Beispiele Lagranger Untermannigfaltigkeiten. Es scheint zudem so, dass viele Phänomene der symplektischen Geometrie auf oft nicht so offensichtliche Weise auf Eigenschaften Lagranger Untermannigfaltigkeiten beruhen; kurz gesagt: „Alles ist Lagrange“ (Weinstein, siehe [5] für einige Beispiele).

Das durch eine symplektische Form definierte Maß ist ebenfalls invariant unter Symplektomorphismen. Für klassische mechanische Systeme ist dies genau der Inhalt des Satzes von Liouville über die Invarianz des Phasenvolumens. Diese Invariante ist offenbar *robust*: Konvergiert eine Folge von Symplektomorphismen gleichmäßig, so lässt ihr Grenzwert, obwohl er nur noch stetig ist, das Maß ebenfalls invariant. Lange Zeit war unklar, ob es neben dem durch die symplektische Form definierten Maß und einigen offensichtlichen topologischen Invarianten überhaupt andere robuste Invarianten einer symplektischen Form gibt. Anders formuliert lautete die Frage: Was ist der Abschluss der Menge der Symplektomorphismen im Raum der stetigen Abbildungen? Besteht er aus allen das Maß erhaltenden Abbildungen oder einer echten Teilmenge? Schliesslich bewies Gromov in [10], dass sich das Volumen einer symplektischen Mannigfaltigkeit im Allgemeinen nicht mit einer symplektisch eingebetteten Kugel ausschöpfen lässt. Das Supremum der Radien von Kugeln in \mathbb{R}^{2n} , die sich symplektisch in eine $2n$ -dimensionale Mannigfaltigkeit einbetten lassen, ist somit eine solche Invariante, die sich *nicht* durch das Volumen ausdrücken lässt. Konkret besagt Gromov's Non-Squeeze-Theorem, dass für einen symplektisch eingebetteten Ball vom Radius r in einen symplektischen Zylinder vom Radius R , $B^{2n}(r) \hookrightarrow B^2(R) \times \mathbb{C}^{n-1}$, immer $r \leq R$ gilt. Das Volumen des Bildraumes ist jedoch unendlich. Daraus folgt, dass der gleichmäßige Grenzwert einer Folge von Symplektomorphismen selbst ein Symplektomorphismus ist, falls er differenzierbar ist. Das war die Geburtsstunde der symplektischen Topologie (siehe z. B. [16]). Sie ist heute eine beliebte Spielwiese vieler Mathematiker. Das hat sicher viel mit der eingangs erwähnten engen Verknüpfung mit physikalischen Fragestellungen zu tun. Andererseits liegt ihr Reiz darin begründet, dass sich viele der interessanten Probleme relativ elementar formulieren lassen. Für ihre Lösungen müssen hingegen oft alle Register der Geometrie und Analysis gezogen werden.

Eine Fülle von Beispielen hierfür findet sich in Gromovs bahnbrechender Arbeit [10]. Er studiert darin Abbildungen von kompakten Riemannschen Flächen in eine symplektische Mannigfaltigkeit, die holomorphe Funktionen einer komplexen Veränderlichen verallgemeinern (siehe 1.2.(2)). Diese *pseudoholomorphen Kurven* stellen sich als mysteriöse Dolmetscher zwischen symplektischer Topologie und Hamiltonscher Dynamik heraus. Beispielsweise übersetzen sie Gromovs oben beschriebene Non-Squeeze-Eigenschaft in die Existenz geschlossener Bahnen des Hamiltonfeldes auf fast allen (kompakten) Niveaumäppen einer eigentlichen Hamiltonfunktion (siehe [12]). Ein anderes prominentes Beispiel ist die Existenz einer holomorphen Scheibe mit Rand auf einer gegebenen geschlossenen Lagrangen Untermannigfaltigkeit $L \subset \mathbb{C}^n$ (siehe Theorem 1). Daraus folgt, dass die erste Homologie von L nicht verschwindet und man erhält eine neue Obstruktion gegen die Existenz von Lagrangen Einbettungen. In [18] wird mit Hilfe dieser holomorphen Scheibe die Existenz von Hamilton-Sehnen für sogenannte Legendre-Untermannigfaltigkeiten der Niveaumengen gezeigt. Die Anzahl solcher kompakter holomorpher Kurven führt schließlich zu den Gromov-Witten-Inva-

rianten und der Quantenkohomologie, die das Zählen auf effektive Weise organisieren (siehe [15]).

Für den Beweis von Arnolds Vermutung konstruierte Floer eine Homologie, die einer Art Morse-Theorie auf dem unendlich-dimensionalen Schleifenraum entspricht [8, 9]. Dafür wird die Klasse der betrachteten holomorphen Kurven auf Abbildungen vom (nicht kompakten) Zylinder in die symplektische Mannigfaltigkeit erweitert. Diese haben viele Eigenschaften der *punktierten holomorphen Kurven*, die hier besprochen werden sollen (siehe Abschnitt 2). Letztere sind pseudoholomorphe Abbildungen einer punktierten Riemannschen Fläche in eine symplektische Mannigfaltigkeit mit zylindrischen Enden und translationsinvarianten Strukturen darauf, den sogenannten *symplektischen Kobordismen*. Nahe der Punktierungen sind sie asymptotisch zu Zylindern über geschlossenen Bahnen des Hamiltonflusses der Höhenfunktion auf den Enden. Die Anzahl von solchen punktierten holomorphen Kurven definiert wieder Invarianten für diese symplektischen Kobordismen. Die von Givental, Hofer und Eliashberg in [7] entwickelte *Symplektische Feldtheorie (SFT)* beschreibt eine Möglichkeit zur kombinatorischen Berechnung der Gromov-Witten-Invarianten aus diesen neuen Invarianten. Geschlossene symplektische Mannigfaltigkeiten werden in (möglichst einfache) symplektische Kobordismen zerlegt. Den zylindrischen Enden ist ein linear-algebraisches Objekt zugeordnet, die *Kontakthomologie*. Dies ist eine graduierte superkommutative Algebra, die von allen periodischen Hamiltonbahnen auf der Basis des Zylinders frei erzeugt wird. Die Gromov-Witten-Invarianten der symplektischen Kobordismen sind dann algebraische Morphismen in der entsprechenden Kategorie. Die Philosophie ist also ganz analog zur Idee einer Topologischen Feldtheorie (siehe [2]), aber mit einem symplektischen Twist. Es gibt eine Dychotomie für die Ränder der Teile. Sie sind entweder (symplektisch) konvex oder konkav.

Wie in der Floer-Homologie ist die Hamiltonsche Dynamik bereits in die Definition der punktierten holomorphen Kurven eingebaut. Anders als bei kompakten holomorphen Kurven stellen punktierte holomorphe Kurven den Zusammenhang zwischen Hamiltonscher Dynamik und symplektischer Topologie also direkt her. Zum Beispiel folgt aus der Existenz nichtkompakter punktierter holomorpher Kurven *direkt* die Existenz von geschlossenen Hamilton-Bahnen. Dieser Umstand wurde erstmals von Hofer in [13] benutzt. Die Existenz der punktierten holomorphen Kurven wird z. B. aus dem Nichtverschwinden von Gromov-Witten-Invarianten geschlossen, d. h. der symplektischen Topologie der Systems (siehe auch 2.5., Theorem 15). Natürlich gibt es auch den umgekehrten Effekt. Kennt man beispielsweise die Gromov-Witten Invarianten der symplektischen Mannigfaltigkeit und die Dynamik des Geodätenflusses einer Riemannschen Mannigfaltigkeit halber Dimension, erhält man Aussagen über die Lagrangen Einbettungen letzterer in erstere. Dies soll hier am Beispiel der Kleinschen Flasche mit der flachen Metrik und der komplex projektiven Ebene mit der Fubini-Study-Form studiert werden (Abschnitt 3).

Danksagung. Ich möchte mich bei der Deutschen Mathematikervereinigung für die Einladung zum Vortrag auf ihrer Jahrestagung 2002 sowie beim Gutachter dieser Arbeit und Herrn Alois Krieg für die Geduld und die vielen konstruktiven Hinweise bedanken. Für ihre Hilfe möchte ich auch Janko Latschev, Janett Mohnke und Thomas Neukirchner herzlich danken.

1 Das Problem der Lagrangen Einbettungen

Wir beginnen die Exposition mit einer (stark verkürzten) Darstellung der Geschichte des Studiums Lagranger Untermannigfaltigkeiten, speziell im \mathbb{C}^n . Dabei soll ein wesentliches Merkmal der symplektischen Topologie herausgearbeitet werden. Es stellt sich nämlich heraus, dass die Eigenschaft, Lagrange zu sein, auf der Grenze zwischen differential-topologischer Flexibilität und geometrischer Starrheit liegt. Ersteres wird durch Gromovs Arbeiten zum *h-Prinzip* untermauert (siehe [11]), letzteres durch seine Konstruktion holomorpher Kurven in [10].

1.1 Lagrange Immersionen und h-Prinzip

Für ein detailliertes Verständnis der Diskussionen in diesem und im nächsten Abschnitt sei der interessierte Leser auf [4] verwiesen. Eine *symplektische Struktur* auf einer Mannigfaltigkeit W ist eine geschlossene 2-Form $\omega \in \Omega^2(W)$ ($d\omega = 0$), die auf jedem Tangentialraum eine nicht ausgeartete (antisymmetrische) Bilinearform definiert. Das bedeutet, dass durch

$$(1) \quad X \in T_x M \mapsto \omega(X, .) \in T^* M$$

ein Isomorphismus zwischen Tangential- und Kotangentialräumen erklärt wird, oder äquivalent dazu, dass die $2n$ -Form $\omega \wedge \dots \wedge \omega =: \omega^n$ eine Volumenform auf W ist. In jedem Fall folgt daraus, dass W notwendigerweise gerad-dimensional und orientierbar sein muss. Sei $H : W \rightarrow \mathbb{R}$ eine glatte Funktion. Wegen der Nichtdegeneriertheit, (1), von ω ist ihr sogenanntes Hamiltonsches Vektorfeld, X_H , durch $\omega(X_H, .) = -dH$ eindeutig bestimmt.

Wir nennen die Immersion $\iota : L^n \hookrightarrow (M^{2n}, \omega)$ einer Mannigfaltigkeit in eine symplektische Mannigfaltigkeit doppelter Dimension *Lagrange*, falls das pull-back der symplektischen Form darauf verschwindet: $\iota^*\omega = 0$. Für \mathbb{C}^n lässt sich die Lagrange-Bedingung alternativ auch einfach geometrisch charakterisieren: Die Rotation des Bildes des Tangentialraumes in $p \in L$, $\iota_*(T_p L)$, mittels Multiplikation mit $\sqrt{-1}$ steht senkrecht auf diesem. Es gibt einen guten Grund, dies nicht zur Definition zu machen. Die Bedingung suggeriert fälschlicherweise, dass Lagrange zu sein, eine Eigenschaft der Hermiteschen Geometrie des \mathbb{C}^n ist. Die ursprüngliche Definition ist aber allgemeiner: Jede komplexe Isometrie des \mathbb{C}^n ist natürlich ein Symplektomorphismus. Der Raum dieser Isometrien ist die endlich-dimensionale Liesche Gruppe $U(n)$. Andererseits ist der Fluss einer beliebigen glatten Funktion mit kompaktem Träger auf \mathbb{C}^n eine Familie von Symplektomorphismen. Demnach ist der Raum der Symplektomorphismen unendlich-dimensional und daher viel größer als der der Isometrien.

Die Lagrange-Bedingung im \mathbb{C}^n kann man nun abschwächen, indem man nur noch fordert, dass sich $\iota_*(T_p L)$ und $\sqrt{-1}\iota_*(T_p L)$ transversal schneiden. Solche Immersionen wollen wir *total reell* nennen. Es stellt sich heraus, dass sich Lagrange Immersionen rein differential-topologisch mittels des sogenannten h-Prinzips studieren lassen. Gromov und Lees bewiesen so, dass die folgenden Räume von Immersionen für eine gegebene geschlossene Mannigfaltigkeit L^n schwach homotopieäquivalent zu

einander sind:

$$\{\iota : L \hookrightarrow \mathbb{C}^n \mid L \text{ Lagrange}\} \cong \{\iota : L \hookrightarrow \mathbb{C}^n \mid \iota \text{ total reell}\} \cong \{\Phi : TL \otimes_{\mathbb{R}} \mathbb{C} \tilde{\rightarrow} L \times \mathbb{C}^n\}.$$

Φ sind komplex-lineare Trivialisierungen des komplexifizierten Tangentialbündels von L . Demnach ist dies die einzige Obstruktion gegen die Existenz Lagranger Immersionen von L in \mathbb{C}^n , und sie ist klassisch topologisch. Beispielsweise lassen alle Sphären S^n in den jeweiligen \mathbb{C}^n Lagranger Immersionen zu. Solche Probleme nennt Gromov auch „weich“. Man verwechsle dabei aber „weich“ nicht mit leicht. Die Lösungen weicher Probleme können trotzdem sehr hart oder besser gesagt schwer sein.

Diese Einheitlichkeit zwischen total reeller und Lagranger Bedingung geht verloren, sobald man Einbettungen betrachtet. Total reelle Einbettungen sind immer noch differential-topologisch vollständig beschreibbar, also immer noch weich. Audin beschrieb diese in vielen Fällen in [3]. Ganz anders verhält es sich mit Lagranger Einbettungen. Natürlich bilden diese eine Teilmenge total reeller Einbettungen. Aber es stellt sich heraus, dass die Existenz total reeller Einbettungen einer gegebenen geschlossenen Mannigfaltigkeit L in den \mathbb{C}^n nicht einmal notwendig die Existenz von Lagranger-Einbettungen impliziert, wie wir im nächsten Abschnitt sehen werden. Gromov bezeichnet solche Probleme als „hart“. Mittels h-Prinzip fand er eine Reihe von Fragestellungen, die durch eine kleine Abschwächung der Voraussetzungen weich werden. Seine Arbeit über pseudoholomorphe Kurven [10] lieferte schließlich den Beweis, dass die ursprünglichen Probleme jedoch tatsächlich hart waren.

1.2 Holomorphe Scheiben

Folgendes Resultat aus Gromovs Arbeit liefert neue, überraschende Obstruktionen gegen die Existenz Lagranger Einbettungen in den \mathbb{C}^n .

Theorem 1. Für jede geschlossene Lagranger Untermannigfaltigkeit $L \subset \mathbb{C}^n$ gibt es eine nicht-konstante holomorphe Abbildung auf der Einheitskreisscheibe $u = (u_1, \dots, u_n) : \Delta \rightarrow \mathbb{C}^n$ mit der Randbedingung $u(\partial\Delta) \subset L$. \square

Für die Existenz der holomorphenen Abbildung benutzte Gromov Invarianz-Eigenschaften von Räumen pseudoholomorpher Kurven in $\mathbb{C}^n \times \mathbb{C}$ zusammen mit seinem Kompaktheitsresultat für Folgen solcher Abbildungen mit „Bubbling“.

Dieses Theorem wird nun in der folgenden Weise ausgenutzt. Die symplektische Form $\omega_0 = \sum_{j=1}^n dx_j \wedge dy_j$ ist exakt, denn $\omega_0 = d\theta_0$ für die 1-Form $\theta_0 = \sum_{j=1}^n x_j dy_j$. Also folgt aus der Lagranger-Bedingung, $\omega_0|_L = 0$, dass die Einschränkung $\theta_0|_L$ eine geschlossene Form ist:

$$d(\theta_0|_L) = (d\theta_0)|_L = \omega_0|_L = 0.$$

Andererseits ist für die holomorphe Scheibe $u : \Delta \rightarrow \mathbb{C}^n$

$$\int_{\Delta} u^* \omega_0 = \int_{\Delta} \omega_0 \left(\frac{\partial u}{\partial s}, \frac{\partial u}{\partial t} \right) ds dt$$

in Koordinaten $z = s + \sqrt{-1}t \in \Delta \subset \mathbb{C}$. Die Cauchy-Riemann-Gleichungen besagen

nun, dass

$$(2) \quad \frac{\partial u}{\partial s} + J_0 \left(\frac{\partial u}{\partial t} \right) = 0$$

wobei $J_0 \in \text{End}(\mathbb{R}^{2n})$ die übliche komplexe Struktur auf \mathbb{C}^n ist, die eindeutig durch $J \left(\frac{\partial}{\partial x_k} \right) = \frac{\partial}{\partial y_k}$ und $J_0^2 = -Id$ charakterisiert wird. $z_k := x_k + \sqrt{-1}y_k$ sind dabei die komplexen Koordinaten (z_1, \dots, z_n) auf \mathbb{C}^n . Es ist leicht zu sehen, dass die Bilinearform

$$\omega_0(v, J_0 w) = v \cdot w$$

gleich dem Skalarprodukt auf $\mathbb{C}^n \cong \mathbb{R}^{2n}$ ist. Allgemein nennt man eine Familie $J = J(x) \in \text{End}(T_x W)$ auf einer Mannigfaltigkeit W mit $J^2 = -Id$ eine *fast-komplexe Struktur*. Ist W mit einer symplektischen Form ω versehen und wird durch $\omega(v, Jw)$ eine Riemannsche Metrik auf W definiert, so nennt man J ω -kompatibel. Die Cauchy-Riemann-Gleichung (2) lässt sich für jedes J hinschreiben. Deren Lösungen nennt Gromov *pseudoholomorph*.

In unserem Fall folgt, dass

$$\int_{\Delta} u^* \omega_0 = \int_{\Delta} \left| \frac{\partial u}{\partial s} \right|^2 ds dt > 0,$$

da das Verschwinden implizieren würde, dass $du \equiv 0$ und somit u konstant ist. Mit dem Satz von Stokes erhalten wir

$$\int_{\partial \Delta} u^* \theta_0 = \int_{\Delta} u^* \omega_0 > 0.$$

Damit beantwortete Gromov die Frage von Arnold, ob es eine geschlossene Lagrange Untermannigfaltigkeit L gibt, für die $\theta_0|_L$ exakt ist, d. h. für die es eine glatte Funktion $f : L \rightarrow \mathbb{R}$ gibt mit $df = \theta_0|_L$:

Folgerung 2. *Die Einschränkung $\theta_0|_L$ auf eine geschlossene Lagrange Untermannigfaltigkeit $L \subset \mathbb{C}^n$ ist nicht exakt. Insbesondere verschwindet die erste Betti-Zahl von L nicht.* \square

Da L insbesondere nicht einfach zusammenhängend sein kann, folgt zum Beispiel, dass es für keine n -Sphäre S^n eine Lagrange Einbettung in den \mathbb{C}^n geben kann. Für $n = 3$ blieb dies mit den „weichen“ Methoden ein offenes Problem: S^3 lässt nämlich (als einzige der Sphären) sogar eine total reelle Einbettung in den \mathbb{C}^3 zu.

1.3 Lagrange Flächen in \mathbb{C}^2

Die Frage, für welche geschlossene Flächen es eine Lagrange Einbettung in den \mathbb{C}^2 gibt, ist fast vollständig äquivalent zur Frage, ob es eine total reelle Einbettung gibt. Jede Fläche, die sich total reell in den \mathbb{C}^2 einbetten lässt, besitzt nämlich auch eine Lagrange Einbettung – mit einer Ausnahme: der Kleinschen Flasche. Diese lässt zwar eine total reelle Einbettung zu. Eine Lagrange Einbettung ließ sich aber nicht konstruieren, so dass man seit langem annahm, dass es keine gibt. Ihre erste Betti-Zahl ist aber gleich 1. Also ergibt sich aus Gromovs Theorem noch kein Widerspruch zur Existenz einer La-

grangen Einbettung. Mit den (nicht-kompakten) punktierten holomorphen Kurven kann man schließlich zeigen

Theorem 3 ([19]). *Es gibt keine Lagrange Einbettung der Kleinschen Flasche in den \mathbb{CP}^2 .*

Die bewiesene Aussage ist stärker als die Frage nach der Existenz einer solchen Einbettung in den \mathbb{C}^2 . Man kann nämlich durch Reskalierung annehmen, dass eine Lagrange Einbettung in den \mathbb{C}^2 in einer beliebig kleinen Kugel liegt. Andererseits besagt Darboux's Theorem, dass jeder Punkt einer symplektischen Mannigfaltigkeit (W^{2n}, ω) eine Umgebung besitzt, die symplektomorph zu einer (kleinen) Kugel im (\mathbb{C}^n, ω_0) ist. Also bekommt man für jede Lagrange Einbettung einer geschlossenen Mannigfaltigkeit in den \mathbb{C}^n eine Lagrange Einbettung in jede beliebige symplektische Mannigfaltigkeit der gleichen Dimension. Somit folgt die Nichtexistenz einer Lagrange Einbettung in den \mathbb{C}^2 aus der Nichtexistenz einer solchen Einbettung in irgend eine symplektische Mannigfaltigkeit.

Zum Beweis des Theorems 3 konstruieren wir eine Reihe von punktierten holomorphen Kurven, die einen Teil des Grenzwertes einer Folge pseudoholomorpher Kurven in \mathbb{CP}^2 bilden. Die zur Fubini-Study-Form kompatiblen fast-komplexen Strukturen degenerieren dabei in den Punkten der Lagrange-Untermannigfaltigkeit.

Satz 4. *Sei $L \subset \mathbb{CP}^2$ eine Lagrange-Einbettung einer flachen geschlossenen Fläche (Torus oder Kleinsche Flasche) bezüglich der Fubini-Study-Form ω . Dann gibt es eine ω -kompatible fast-komplexe Struktur J und drei J -holomorphe Sphären F, G und $H : S^2 \rightarrow \mathbb{CP}^2 \setminus L$. F, G und H kann man somit als 2-Zyklen verstehen. Deren Homologieklassen sind gleich und jede erzeugt $H_2(\mathbb{CP}^2; \mathbb{Z}) \cong \mathbb{Z}$. Außerdem gibt es noch zwei glatte Scheiben D und $E : \Delta \rightarrow \mathbb{CP}^2$ mit Rand auf L , die den folgenden homologischen Schnittbedingungen genügen:*

$$\begin{aligned} F \cdot D &= 1, & G \cdot D &= 0, & H \cdot D &= 0, \\ F \cdot E &= 0, & G \cdot E &= 1, & H \cdot E &= 0. \end{aligned}$$

Zuerst überzeugen wir uns einmal davon, dass dieser Satz das Problem der Kleinschen Flasche auch löst.

Beweis von Theorem 3. Wir zeigen, daß die Ränder $\partial D, \partial E$ linear unabhängig in $H_1(L; \mathbb{Q})$ sind. Sei also

$$k[\partial D] + l[\partial E] = 0 \in H_1(L; \mathbb{Z})$$

für ein Paar ganzer Zahlen (k, l) . Dass heißt, es gibt eine 2-Kette $C \in C_2(L; \mathbb{Z})$ mit $\partial C = k\partial D + l\partial E$. Somit ist $kD + lE - C \in Z_2(\mathbb{CP}^2; \mathbb{Z})$ ein 2-Zyklus, den wir jetzt gegen die pseudoholomorphen Sphären testen:

$$\begin{aligned} (kD + lE - C) \cdot F &= k \\ (kD + lE - C) \cdot G &= l \\ (kD + lE - C) \cdot H &= 0. \end{aligned}$$

Da $[F] = [G] = [H] \in H_2(\mathbb{CP}^2; \mathbb{Z})$ muß $k = l = 0$ sein, und die Behauptung ist gezeigt. Also ist $b_1(L) \geq 2$ und L kann somit nicht die Kleinsche Flasche sein. \square

In [19] wird auch noch geschlossen, dass alle Lagrange Tori in \mathbb{CP}^2 glatt isotop zueinander sind, d. h. dass es keine verknüpften Lagrange Tori in \mathbb{CP}^2 gibt.

Der Rest der Abhandlung dient der Einführung der Elemente der symplektischen Feldtheorie und der Demonstration ihrer Arbeitsweise anhand des Beweises von Satz 4.

2 Elemente der symplektischen Feldtheorie

Wir beschreiben hier die Zerlegung und Verklebung symplektischer Mannigfaltigkeiten, kompatible fast-komplexe Strukturen auf den symplektischen Kobordismen sowie punktierte holomorphe Kurven, deren Anzahl die Invarianten der symplektischen Feldtheorie (SFT) definiert. Für die Konstruktion und Beschreibung der SFT sei der Leser auf [7] verwiesen.

2.1 Symplektische Kobordismen

Sei (W, ω) eine symplektische Mannigfaltigkeit mit einer Produktstruktur auf den Enden: Das Komplement einer genügend großen kompakten Teilmenge besteht aus positiven und negativen Halbzylindern über Hyperflächen. Die Translation ist kein Symplektomorphismus, aber lässt die symplektische Form bis auf Skalierung invariant. Wir können annehmen, dass die Skalierung infinitesimal konstant ist. Wir wollen die Translationsrichtung so orientieren, dass dieser Faktor positiv wird. Insbesondere ist das Hamiltonfeld der Funktion, die diese Translation mit der auf \mathbb{R} vertauscht und die auf der Hyperfläche konstant ist, translationsinvariant. Dann nennen wir (W, ω) einen *symplektischen Kobordismus*. Die positiven Halbzylinder nennen wir konvexe, die negativen konkav Enden.

Eine fast-komplexe Struktur auf einem symplektischen Kobordismus heißt *kompatibel*, falls sie kompatibel zu ω und auf den Enden translationsinvariant ist sowie das zur Translation gehörige Vektorfeld mit dem obigen Hamiltonfeld vertauscht. Zu jedem symplektischen Kobordismus ist der Raum aller kompatiblen fast-komplexen Strukturen nicht-leer und kontrahierbar.

Beispiel 5. (1) Sei L eine beliebige Mannigfaltigkeit. Ihr Kotangentialbündel T^*L besitzt eine natürliche symplektische Struktur, die auf die folgende Weise gegeben wird. Die kanonische 1-Form $\theta \in \Omega^1(T^*L)$ wird durch

$$\theta_l(X) = l(\pi_* X)$$

definiert, wobei $l \in T^*L$ ein Kotangentialvektor ist. $X \in T_l(T^*L)$ ist ein Tangentialvektor an T^*L in l , und π_* ist das Differential der Projektion $T^*L \rightarrow L$. Man prüft leicht nach, z. B. durch Wahl lokaler Koordinaten auf L , dass das äußere Differential $d\theta \in \Omega^2(T^*L)$ symplektisch ist. Jetzt wählen wir noch eine Riemannsche Metrik g auf L . Bezeichne

$S^*L \subset T^*L$ das zugehörige Einheitssphärenbündel. Die Identifikation des offenen Endes von T^*L mit dem Halbzylinder $\mathbb{R}_+ \times S^*L$ sei so gewählt, dass die Multiplikation mit $\exp(r)$ in T^*L übergeht in die Translation um $r \in \mathbb{R}$ in der ersten Komponente im Zylinder. Das Kotangentialbündel T^*L ist dann ein symplektischer Kobordismus mit einem konvexen Ende der Form $\mathbb{R}_+ \times S^*L$. Die Flusslinien des zugehörigen translationsinvarianten Hamiltonfeldes (in T^*L) projizieren (natürlich) auf Geodäten in (L, g) ! Bahnen des Vektorfeldes entsprechen somit genau den orientierten Geodäten mit Längenparametrisierung.

(2) Sei $L \subset W$ eine geschlossene Lagrange Untermannigfaltigkeit in einer geschlossenen symplektischen Mannigfaltigkeit (W, ω) . Nach Weinstein lässt sich Darboux's Theorem wie folgt verallgemeinern: Eine hinreichend kleine Umgebung von L ist symplektomorph zu einer Umgebung des Nullschnittes in $(T^*L, d\theta)$. Nach Reskalierung einer gegebenen Metrik kann man annehmen, dass dies das Einheitskugelbündel bezüglich dieser ist. Demnach ist $W \setminus L$ ein symplektischer Kobordismus mit konkavem Ende $\mathbb{R}_- \times S^*L$.

(3) Über die Translationsinvarianz der symplektischen Form bis auf (translationsinvariante) Skalierungen bekommt man eine symplektische Struktur auf dem Zylinder $\mathbb{R} \times S^*L$, gegeben durch $d(e^\alpha)$, der sogenannten Symplektisierung. Dabei ist $\alpha = \theta|_{S^*L}$. Das ist (formal) ein trivialer symplektischer Kobordismus mit Halbzylinder-Enden über S^*L . Ist die fast-komplexe Struktur auf dem ganzen Zylinder translationsinvariant, sprechen wir von einer α -kompatiblen Struktur. Die Hamilton-Bahnen der entsprechenden Funktion $H(r, x) = r$ nennt man auch die (α)-Reeb-Bahnen.

2.2 Aufspaltung und Verklebung von symplektischen Kobordismen

Man kann eine symplektische Mannigfaltigkeit (W, ω) unter folgender Voraussetzung in symplektische Kobordismen „zerlegen“. Sei $M \subset W$ eine geschlossene Hyperfläche deren Kragenumgebung eine Produktstruktur ähnlich der Enden eines symplektischen Kobordismus besitzt. Wir setzen voraus, dass W durch M in $W \setminus M = W_0^+ \sqcup W_0^-$ zerlegt wird. Die Signatur der Teile W_0^\pm wird durch die Forderung bestimmt, dass das Translationsvektorfeld in W_0^+ hineinzeigt.

Beispiel 6. Die Menge $S^*L \subset T^*L$ erfüllt diese Bedingung. Wir haben bereits gesehen, dass man sie auch als Hyperfläche in W auffassen kann, falls $L \subset W$ Lagrange Untermannigfaltigkeit ist. Sie erbt dann die Produktstruktur auf einer Umgebung.

Man kann W_0^\pm durch Verkleben der Kragen der Ränder mit den Kragen der Ränder der Halbzylinder über M zu symplektischen Mannigfaltigkeiten (W^\pm, ω^\pm) „vervollständigen“. Sind wir mit einem symplektischen Kobordismus (W, ω) gestartet, so sind beide so erhaltenen (W^\pm, ω) wieder symplektische Kobordismen.

Beispiel 7. In unserem Beispiel ist also (W^+, ω^+) symplektomorph zu $(W \setminus L, \omega|_{W \setminus L})$, und (W^-, ω^-) zu $(T^*L, d\theta)$.

Eine Umkehrung dieser *Aufspaltung* einer symplektischen Mannigfaltigkeit kann wie folgt durch eine *Verklebung* von zwei symplektischen Kobordismen (W^\pm, ω^\pm) beschrieben werden. Nehmen wir an, dass ein konkaves Ende von (W^+, ω^+) und ein kon-

vexes Ende von (W^-, ω^-) Halbzyylinder über derselben Mannigfaltigkeit sind. Dann kann man nach Reskalierung von ω^- und Stutzen der beiden Enden beide Teile zu einer symplektischen Mannigfaltigkeit verkleben. Die beiden Enden sind durch einen symplektischen Zylinder endlicher Länge, sagen wir τ , ersetzt worden. Die symplektische Mannigfaltigkeit bezeichnen wir mit (W_τ, ω_τ) . Sind J^\pm kompatible fast-komplexe Strukturen auf W^\pm , so erhalten wir durch diese Verklebung eine Familie $\{J_\tau\}$ von ω_τ -kompatiblen fast komplexen Strukturen auf W_τ .

2.3 Punktisierte holomorphe Kurven

Ein Zylinder über einer geschlossenen α -Reeb-Bahn γ in der Symplektisierung einer Kontaktmannigfaltigkeit mit α -kompatibler fast komplexer Struktur ist eine komplexe Kurve. Dies ist der Prototyp einer punktierten holomorphen Sphäre mit einer positiven und einer negativen Punktierung, die wir einfach *Zylinder über der Reeb-Bahn* γ nennen werden.

Definition 8. Eine punktierte J -holomorphe Sphäre ist eine J -holomorphe Abbildung von einer punktierten Riemannschen Fläche nach W , die nahe der Punktierungen entweder in ein konkaves Ende (wir sprechen dann von einer positiven Punktierung) oder in ein konkavtes Ende (entsprechend negative Punktierung genannt) abbildet und dort asymptotisch (bezüglich r) zu dem entsprechenden Halbzyylinder über einer Reeb-Bahn ist.

Bemerkung 9. Punktisierte holomorphe Kurven (in Symplektisierungen) wurden erstmals von Hofer in [13] durch Analyse von nicht-konvergenten Folgen (kompakter) pseudoholomorpher Kurven konstruiert. Die wesentliche Eigenschaft, die für eine beliebige J -holomorphe Abbildung $f : \dot{\Sigma} \rightarrow W$ entweder die Hebbareit in den Punktierungen oder die Asymptotik garantiert, ist die Endlichkeit einer „Energie“.

Beispiel 10. Eine Riemannsche Struktur auf einer Mannigfaltigkeit L induziert eine fast-komplexe Struktur J_0 auf T^*L , die zwar kompatibel zu $d\theta$, nicht aber zur Kontaktform α ist. Sie ist insbesondere nicht translations-invariant unter der symplektischen Identifizierung $T^*L \setminus L \cong \mathbb{R} \times S^*L$. In ihr ist der Annullator der Normalenrichtungen einer Geodätischen γ ein komplexer Zylinder. Durch Streckung und Stauchung sind alle ρ -Sphärenbündel

$$\{\nu \in T^*L \mid \|\nu\| = \rho\}$$

mit dem Einheitssphärenbündel identifiziert. Somit kann man die dort definierte komplexe Struktur auf der Kontaktdistribution translationsinvariant auf alle ρ -Sphärenbündel transportieren. Mittels Abschneidefunktion in ρ auf T^*L definieren wir eine fast-komplexe Struktur J , die in der Nähe des Nullschnittes mit J_0 übereinstimmt. Mit Parameter $r := \ln \rho$ ist sie translationsinvariant außerhalb einer kompakten Umgebung des Nullschnittes 0_L . $J(\frac{\partial}{\partial r})$ ist das Reebvektorfeld (Geodätenfeld) auf dem Einheitskotangentialbündel für $\rho >> 0$. Dies ist überall parallel zu $J_0(\frac{\partial}{\partial r})$. Dann ist die vorher beschriebene J_0 -komplexe Kurve ebenfalls J -komplex und kann als 2-fach punktierte J -holomorphe Sphäre parametrisiert werden. Diese Kurve bezeichnen wir mit f_γ .

Lemma 11 ([19]). *Sei (L, g) der flache 2-Torus oder die flache Kleinsche Flasche. Dann sind alle J -holomorphen zweifach punktierten Sphären, die den Nullschnitt 0_L schneiden, von der in Beispiel 10 beschriebenen Form.* \square

2.4 Fortsetzungen punktierter holomorpher Kurven

Wir benutzen die Asymptotik punktierter holomorpher Kurven, um diese zu glatten Abbildungen auf kompakten Flächen mit Rand fortzusetzen.

Für eine punktierte Riemannsche Fläche definieren wir die glatte, kompakte Fläche mit Rand dadurch, dass wir jede Punktierung mit einer Randkomponente abschließen (aus einem Kreis S^1 bestehend). Wegen der Asymptotik bekommen wir eine Fortsetzung der holomorphen Abbildung auf diese Fläche zu einer differenzierbaren Abbildung. Wir kompaktifizieren dabei den Bildraum unter Benutzung der Produktstruktur auf den Enden.

Beispiel 12. *Sei $W^+ = W \setminus L$ das Komplement einer Lagrangen Untermannigfaltigkeit aus unserem Beispiel 7. Mit Hilfe der Projektion $\{-\infty\} \times S^*L \rightarrow L$ bekommen wir eine glatte Abbildung $\pi : \overline{W \setminus L} \rightarrow W$. Somit definiert $\pi \circ \bar{f} : \overline{\Sigma} \rightarrow W$ eine glatte Abbildung, die wir ebenfalls einfach mit \bar{f} bezeichnen werden. Es gilt $\bar{f}(\partial \overline{\Sigma}) \subset L$.*

Bemerkung 13. *Die soeben eingeführte Abbildung $\bar{f} : \overline{\Sigma} \rightarrow W$ ist keine pseudoholomorphe Kurve in W bezüglich einer kompatiblen fast-komplexen Struktur. Um für den mit dem Riemann-Hilbert-Problem vertrauten Leser Konfusion zu ersparen, werden im Folgenden einige wesentliche Unterschiede zu diesen aufgelistet.*

- (1) *Die konforme Struktur bezüglich derer f holomorph ist, degeneriert auf dem Rand von $\overline{\Sigma}$.*
- (2) *Ähnliches gilt für die kompatible fast-komplexe Struktur auf $W \setminus L$. Sie ist in den Punkten von L entartet.*
- (3) *\bar{f} erfüllt sehr starke „Randbedingungen“: Jede Komponente überlagert eine Geodäte auf L . Solche Bedingungen sind für pseudoholomorphe Kurven (mit Rand) nicht wohlgestellt, d. h. dass es sie generisch gar nicht gibt!*
- (4) *\bar{f} bildet die inneren Punkte $\dot{\Sigma} \subset \overline{\Sigma}$ in das Komplement von L ab. Es gibt keine einfache Möglichkeit, dies a priori für eine pseudoholomorphe Kurve mit Rand auf L zu kontrollieren.*

2.5 Grenzwerte holomorpher Kurven

Punktierte holomorphe Kurven erhält man beispielsweise als Grenzwerte von Folgen J_τ -holomorpher Kurven, wenn der Parameter τ dabei beliebig groß wird.

Satz 14 ([6]). *Sei die Fubini-Study-Form ω auf \mathbb{CP}^n so normiert, dass ihr Integral über die Erzeugende c von $H_2(\mathbb{CP}^n; \mathbb{Z}) \cong \mathbb{Z}$ gleich π ist. Sei $M \subset \mathbb{CP}^n$ eine geschlossene Hyperfläche, die \mathbb{CP}^n in zwei symplektische Kobordismen W^\pm wie oben beschrieben zerlegt, J^\pm, J_0 seien kompatible fast-komplexe Strukturen auf W^\pm bzw. $\mathbb{R} \times M$, J_τ sei*

die entsprechende Familie kompatibler fast-komplexer Strukturen auf \mathbb{CP}^n . Sei $\{f_n\}$ eine Folge J_{τ_n} -holomorpher Sphären

$$f_n : \mathbb{CP}^1 \longrightarrow \mathbb{CP}^n,$$

mit konstanter primitiver Fundamentalklasse, $[f_n] = c$ und $\tau_n \rightarrow \infty$. Dann bekommt man für eine Teilfolge als wohldefiniertes Grenzobjekt eine Aufspaltung in eine endliche Familie $F = (F^{(1)}, \dots, F^{(N)})$ punktierter holomorpher Kurven mit den folgenden Eigenschaften:

- (1) Jedes der $F^{(i)}$ ist die Vereinigung punktierter holomorpher Sphären in W^+ (für $i = 1$), in $\mathbb{R} \times M$ (für $i = 2, \dots, N-1$) sowie in W^- (für $i = N$)
- (2) Die negativen Asymptotiken von $F^{(i)}$ stimmen mit den positiven Asymptotiken von $F^{(i+1)}$ überein.
- (3) Die in diesem Sinne entlang den entsprechenden Randkomponenten der kompaktifizierten Flächen formal verklebten Definitionsbereiche ergeben eine Sphäre.
- (4) Die Homologieklassen der formalen Verklebung der Abbildungen stimmt mit c überein. Dabei identifizieren wir die Verklebung der (kompaktifizierten) Kobordismen entlang der entsprechenden Randkomponenten mit der ursprünglichen Mannigfaltigkeit, \mathbb{CP}^n .
- (5)

$$\int (F^{(1)})^* \omega^+ = \pi.$$

Wir erhalten sofort einen Beweis für folgendes

Theorem 15 ([14]). Jede reguläre Niveaumenge vom Kontakttyp einer Hamiltonfunktion auf \mathbb{CP}^n trägt wenigstens eine periodische Bahn des Hamiltonflusses.

Beweis: Jede solche Hyperfläche zerteilt \mathbb{CP}^n in zwei Teile. Wegen des Nichtverschwindens der entsprechenden Gromov-Witten-Invariante gibt es für zwei beliebige Punkte und jede kompatible fast komplexe Struktur wenigstens eine pseudoholomorphe Kurve, deren Fundamentalklasse primitiv ist und deren Bild jeden der beiden Punkte enthält. Wir wählen nun je einen Punkt in jeder der Komponenten des Komplements der Hyperfläche und wenden Satz 14 auf die entsprechende Familie J_{τ} -holomorpher Kurven an. Dann bekommen wir im Grenzwert punktierte holomorphe Kurven in beiden (mit Zylindern vervollständigten) Teilen. Diese müssen wenigstens eine Punktierung enthalten, da sie sich entlang dieser wieder zu einer Sphäre verkleben müssen. Eine Punktierung entspricht aber einer geschlossenen Hamilton-Bahn. \square

2.6 Schnittverhalten holomorpher Kurven in Dimension 4

Pseudoholomorphe Kurven haben viele Eigenschaften mit holomorphen Kurven gemeinsam. Eine sehr wichtige Eigenschaft ist das Schnittverhalten in Dimension 4:

Satz 16. [Theorems 7.1. aus [17]] Auf \mathbb{C}^2 sei eine fast-komplexe Struktur J gegeben mit $J(0) = \sqrt{-1}$ auf $T_0 \mathbb{C}^2 \cong \mathbb{C}^2$. Seien $f_i : \Delta \rightarrow \mathbb{C}^2$ zwei J -holomorphe Abbildungen mit $f_1(0) = f_2(0)$. Dann gibt es Umgebungen U_1 und U_2 der 0, so dass entweder

- $f_1(U_1) = f_2(U_2)$, oder
- $f_1(z_1) \neq f_2(z_2)$ für alle $(0, 0) \neq (z_1, z_2) \in U_1 \times U_2$.

Im zweiten Fall können wir 0 einen Schnittindex zuordnen, der genau der (wohl-definierten) transversalen Schnittzahl von Störungen der f_i in kleinen Umgebungen $V_i \subset \overline{V}_i \subset U_i$ entspricht. Dieser Index ist immer positiv. Er ist gleich 1 genau dann, wenn sich f_1 und f_2 in 0 transversal schneiden.

3 Holomorphe Sphären im Komplement von L

Wir kommen nun zum Beweis von Satz 4. Wir reskalieren die flache Metrik so, dass, wie in unseren Beispielen, das Einheitskugelbündel in T^*L symplektomorph zu einer Umgebung von $L \subset \mathbb{CP}^2$ wird. Wir fixieren eine fast-komplexe Struktur auf \mathbb{CP}^2 , die kompatibel zur Fubini-Study-Form sowie zur Kontaktform $\theta|_{S^*L}$ auf der Umgebung $S^*L \subset V \cong (-\epsilon, \epsilon) \times S^*L$ ist. Somit bekommen wir kompatible Strukturen J^\pm auf den symplektischen Bordismen $\mathbb{CP}^2 \setminus L, T^*L$.

3.1 Konstruktion der Scheiben

Wir wählen einen Punkt $x \in L$ (auf einer der beiden isolierten Geodäten, falls L die Kleinsche Flasche ist) und fixieren eine J -komplexe Gerade ξ im Tangentialraum, die keinen Tangentialvektor an L enthält. Die Wahl dieses konkreten Punktes wird erst im nächsten Schritt (Abschnitt 3.2) wichtig. Aus den Gromov-Witten-Invarianten für holomorphe Sphären (siehe [15]) folgt, dass es für jedes τ genau eine J_τ -holomorphe Sphäre $f_\tau : \mathbb{CP}^1 \rightarrow \mathbb{CP}^2$ gibt, wobei $[f] \in H_2(\mathbb{CP}^2; \mathbb{Z}) \cong \mathbb{Z}$ erzeugt, $f_\tau(0) = x$ sowie $(f_\tau)_*(T_0 \mathbb{CP}^1) = \xi$ ist. Eine Folge f_{τ_n} mit $\tau_n \rightarrow \infty$ wird nun im Sinne von Satz 14 gegen einen Grenzwert $F = (F^{(1)}, \dots, F^{(N)})$ konvergieren. $F^{(N)}$ enthält eine Komponente, deren Bild durch $x \in 0_L$ verläuft. Dieses hat aufgrund der Wahl mindestens drei Punktierungen: Die Wahl der tangentialen Richtung verhindert, dass diese vollständig in einer der Kurven f_γ einer Geodäten γ liegt. Da $\overline{\Sigma} \cong \mathbb{CP}^1$, gibt es zu jeder dieser Punktierungen eine (andere) Komponente in $F^{(1)}, f^{(i)} : \dot{\Sigma}^{(i)} \rightarrow \mathbb{CP}^2 \setminus L$, $i = 1, 2, 3, \dots$, die eine punktierte holomorphe Sphäre mit genau einer negativen Punktierung ist. Dann setzen wir

$$\begin{aligned} D &:= \bar{f}^{(1)}(\Delta) \\ E &:= \bar{f}^{(2)}(\Delta). \end{aligned}$$

Dieser erste Schritt ist die zentrale Idee bei der Konstruktion solcher Obstruktionen. In jeder der Situationen muss man dann Eigenschaften von F zeigen. Das soll durch die folgenden beiden Resultate illustriert werden.

Theorem 17 (Viterbo). *Sei $L \subset \mathbb{CP}^n$ ($n \geq 3$) eine geschlossene Lagrange Untermannigfaltigkeit. Dann gibt es auf L keine Riemannsche Metrik mit negativer Krümmung.*

Beweis: Durch jeden Punkt in \mathbb{CP}^n und jedes J_τ gibt es eine J_τ -holomorphe Sphäre mit primitiver Fundamentalklasse. Demnach gibt es durch jeden Punkt in T^*L eine punktierte J^- -holomorphe Kurve. Die Geodäten einer negativ gekrümmten Riemannschen Mannigfaltigkeit sind aber alle minimal und isoliert. Daraus kann man schließen, dass keine der punktierten holomorphen Kurven in T^*L deformiert werden kann. Darum kann man einen Punkt finden, der in keiner solchen enthalten ist und erhält einen Widerspruch. \square

Theorem 18 (Cieliebak/Mohnke). *Sei $L \subset \mathbb{CP}^n$ ein Lagranger n -Torus. Dann gibt es eine Scheibe $u : \Delta \rightarrow \mathbb{CP}^n$ mit Rand auf L , $u(\partial\Delta) \subset L$, so dass*

$$0 < \int_{\Delta} u^* \omega \leq \frac{\pi}{n+1}.$$

Beweis. Die beschriebene Konstruktion für beliebige Dimension liefert einen Grenzwert, dessen Komponente $F^{(1)}$ wenigstens $(n+1)$ einfach punktierte Sphären enthält. Nun ist nach Satz 14

$$\int (F^{(1)})^* \omega^+ = \pi.$$

Da die 2-Form $(F^{(1)})^* \omega^+$ auf allen Komponenten von $F^{(1)}$ nicht-negativ ist und nicht verschwindet, folgt die Ungleichung für eine dieser einfach punktierten Sphären $f : \mathbb{C} \rightarrow W^+$. Die Behauptung folgt dann mit

$$\int_{\mathbb{C}} f^* \omega^+ = \int_{\Delta} \bar{f}^* \omega.$$

 \square

3.2 Konstruktion der holomorphen Sphären

Mit jeder der punktierten holomorphen Kurven $f^{(i)}$ verfahren wir nun wie folgt: Wir wählen auf ihr einen glatten Punkt y und eine J -komplexe Richtung η derart, dass eine einfache punktierte holomorphe Sphäre, die durch y verläuft, und dort tangential an η ist, mindestens 4 reelle (effektive) Deformationsparameter hat. Wie schon bei der vorgenommenen Konstruktion benutzen wir, dass es für jedes τ genau eine J_τ -holomorphe Sphäre $g_\tau : \mathbb{CP}^1 \rightarrow \mathbb{CP}^2$ gibt, $[g_\tau] \in H_2(\mathbb{CP}^2; \mathbb{Z}) \cong \mathbb{Z}$ ist ein Erzeuger, $g_\tau(0) = y$ sowie $(g_\tau)_*(T_0 \mathbb{CP}^1) = \eta$. Eine Teilfolge g_{τ_n} konvergiert dann im Sinne von Satz 14 gegen einen Grenzwert $\mathbf{G} = (G^{(1)}, \dots, G^{(N)})$. Der zentrale Punkt des Arguments ist nun die folgende Beobachtung:

Lemma 19 ([19]). *Für den oben erhaltenen Grenzwert \mathbf{G} gilt $N = 1$, d. h. wir erhalten eine glatte J^+ -holomorphe Sphäre in $W^+ = \mathbb{CP}^2 \setminus L$.*

Zum Beweis: $G^{(1)}$ enthält eine Komponente, deren Bild durch y verläuft. Durch die generische Wahl von y und τ ist garantiert, dass diese einen Deformationsindex von wenigstens 4 hat. Allgemein gilt, dass der Index von \mathbf{G} (bzw. die Summe der Indizes

der Komponenten) nicht größer als 4 sein kann, wobei Gleichheit eintritt, genau dann wenn die Summe der Eulercharakteristiken der Komponenten $G^{(2)}, \dots, G^{(N)}$ verschwindet. Hier benutzt man, dass es keine einfach punktierten Sphären unter diesen Komponenten gibt, also jede einen nicht-positiven Beitrag liefert. Letzteres wiederum folgt aus der Nichtexistenz einer kontrahierbaren Geodäte für Torus bzw. Kleinsche Flasche. Da die formale Verklebung der Komponenten von \mathbf{G} eine Sphäre ergibt, gibt es wenigstens eine einfach punktierte Sphäre in $G^{(1)}$. Im Falle, dass T^2 der Torus ist, hat diese einen ungeraden Index. Dann wäre folglich für eine gebrochene holomorphe Kurve \mathbf{G} der Index aller Komponenten von $G^{(1)}$ nicht größer als 3, und die Aussage folgt in diesem Fall.

Für den Fall der Kleinschen Flasche schließt man hingegen zuerst, dass alle anderen Komponenten von $G^{(1)}$ vom Index 0 sind. Daraus folgert man, dass alle Punktierungen asymptotisch zu (den Reeb-Bahnen) von isolierten Geodäten sind. Somit gibt es in $G^{(N)}$ eine zwei-punktierte Sphäre f_γ für eine isolierte Geodäte γ . Mit Argumenten, die wieder Satz 16 benutzen, zeigt man, dass diese Geodäte einfach sein muß, und dass f_γ höchstens einmal in $G^{(N)}$ auftauchen darf. Es gibt jedoch nur zwei einfache isolierte geschlossene Geodäten. Deren Schnittindex mit dem Nullschnitt ist $f_\gamma \cdot 0_L = 1$ in \mathbb{Z}_2 . Andererseits ist $[L] = 0 \in H_2(\mathbb{CP}_2; \mathbb{Z}_2)$. Somit muss die Anzahl solcher f_γ mit isolierter Geodäte γ gerade sein. Ist $N > 1$, so müssen also beide möglichen f_γ auftauchen. Dann enthält aber $G^{(N)}$ auch f_γ für die Geodäte γ , auf der wir in 3.1 den Punkt x gewählt hatten. Da die Richtung ξ insbesondere transversal zu diesem f_γ in x ist, erhalten wir einen zweiten transversalen Schnittpunkt zwischen \mathbf{G} und \mathbf{F} . Daraus würde $f_n \cdot g_n \geq 2$ im Widerspruch zur Primitivität folgen. Somit muß auch für diesen Fall $N = 1$ sein und das Lemma folgt. \square

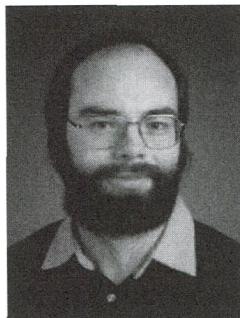
Nehmen wir nun an, dass diese Sphäre \mathbf{G} noch eine der anderen einfach punktierten Sphären irgendwo schneidet. Wegen der Schnitteigenschaft, Satz 16, geschieht dies mit einem endlichen, positiven algebraischen Index. Dann müssten sich aber in kleinen Umgebungen von y und diesem Punkt (für hinreichend große n) f_n und g_n mit einem Index größer als 1 schneiden. Wegen der Positivität aller eventuellen weiteren Beiträge zum algebraischen Schnittindex der beiden Kurven, wäre dieser somit größer als 1 im Widerspruch zur Primitivität der beiden Kurven.

Wir bezeichnen die Sphären, die ich auf diese Weise erhalte, nacheinander mit F , G und H . Sie genügen somit den Forderungen des Satzes 4.

Literatur

- [1] V.I. Arnold. *Mathematische Methoden der klassischen Mechanik*. Deutscher Verlag der Wissenschaften, 1988.
- [2] Michael F. Atiyah. *The geometry and physics of knots*. Cambridge Univ. Press, 1991.
- [3] Michèle Audin. Fibrés normaux d'immersions en dimension moitié, points doubles d'immersions lagrangiennes et plongements totalment réels. *Comment. Math. Helv.*, 63:593–623, 1988.
- [4] Michèle Audin, François Lalonde, and Leonid Polterovich. Symplectic rigidity: Lagrangian submanifolds. In *Holomorphic curves in symplectic geometry*, volume 117 of *Progress in Math.*, pages 271–321. Birkhäuser Verlag, 1994.

- [5] Paul Biran. Geometry of symplectic intersections. *ICM*, II:241–255, 2002.
- [6] F. Bourgeois, Y. Eliashberg, H. Hofer, K. Wysocki, and E. Zehnder. Compactness results in symplectic field theory. *preprint*, 2003.
- [7] Yakov Eliashberg, Alexander Givental, and Helmut Hofer. Introduction to symplectic field theory. In *Visions in mathematics-Towards 2000. Proceedings of a meeting, Tel Aviv, Israel, August 25-September 3, 1999. Part II*, pages 560–673. Birkhäuser, 1999.
- [8] Andreas Floer. Morse theory for Lagrangian intersections. *J. Diff. geom.*, 18:513–517, 1988.
- [9] Andreas Floer. The unregularised flow of the symplectic action. *Comm. Pure Appl. Math.*, 41:775–813, 1988.
- [10] Michail Gromov. Pseudo-holomorphic curves in symplectic manifolds. *Invent. math.*, 82:307–347, 1985.
- [11] Michail Gromov. *Partial differential relations*, volume 3 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1986.
- [12] H. Hofer and E Zehnder. *Symplectic Invariants and Hamiltonian Dynamics*. Birkhäuser Advanced Texts, Baseler Lehrbücher. Birkhäuser, 1994.
- [13] Helmut Hofer. Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three. *Invent. Math.*, 114:515–563, 1993.
- [14] Gang Liu and Gang Tian. Weinstein conjecture and GW-invariants, *Commun. Contemp. Math.*, 2:405–459, 2000.
- [15] Dusa McDuff and Dietmar Salamon. *J-holomorphic curves and Quantum Cohomology*, volume 6 of *University Lecture Series*. American Math. Soc., 1994.
- [16] Dusa McDuff and Dietmar Salamon. *Introduction to symplectic topology*. Oxford Mathematical Monographs. Clarendon Press, Oxford, 1995.
- [17] Mario J. Micallef and Brian White. The structure of branch points in minimal surfaces and in pseudoholomorphic curves. *Annals of Math.*, 139:35–85, 1994.
- [18] Klaus Mohnke. Holomorphic disks and the chord conjecture. *Annals of Math.*, 154:219–222, 2001.
- [19] Klaus Mohnke. Symplectic threadings through Lagrangian eyes. *preprint*, <http://xxx.lanl.gov/abs/math.SG/0106139>, 2001.



Einige topologische Methoden in der nichtlinearen Analysis

Martin Väth*

Abstract

- Keywords and Phrases: Abbildungsgrad, Fixpunktindex, Nichtkompaktheitsmaße, kondensierende Operatoren, ANR
- Mathematics Subject Classification: 47H11, 47H09, 47H10, 47H04, 47J05, 54H25

Anhand des Fixpunktsatzes von Schauder werden einige Ideen der nichtlinearen Funktionalanalyse aufgezeigt, die in der aktuellen Forschung eine Rolle spielen.

* Der Autor ist derzeit Heisenberg-Stipendiat (Az. VA 206/1–1) und möchte der DFG für die finanzielle Unterstützung danken.

Eingegangen: 10.07.2003

Martin Väth, Universität Würzburg, Math. Institut, Am Hubland,
97074 Würzburg; *derzeitige Adresse*: Freie Universität Berlin, Fachbereich
Mathematik und Informatik (WE1), Sekretariat Prof. B. Fiedler,
Arnimallee 2–6, 14195 Berlin. E-Mail: vaeth@mathematik.uni-wuerzburg.de

DMV
JAHRESBERICHT
DER DMV
© B. G. Teubner 2004

1 Einleitung

Die nichtlineare Analysis ist das Gebiet der Mathematik, das sich mit dem Lösen nichtlinearer Gleichungen beschäftigt, insbesondere von Gleichungen in unendlichdimensionalen Räumen, also etwa von (nichtlinearen) Differential- oder Integralgleichungen. Hierzu wird natürlich eine Vielzahl mathematischer Disziplinen herangezogen. Als besonders fruchtreich hat sich hier insbesondere die Topologie erwiesen. E. Zeidler hat im Übersichtswerk [62]–[66] die nichtlineare Analysis grob in drei Gebiete eingeteilt: Monotoniemethoden, Variationsmethoden und topologische Methoden. Obwohl zwar insbesondere die modernen Variationsmethoden auch viel Topologie benutzen, beschränken wir uns in diesem Artikel auf topologische Methoden im obigen Sinne. Der wohl bekannteste Satz dieses Gebiets dürfte der Fixpunktsatz von Schauder sein:

Satz 1 (Schauder). *Sei X normierter Raum und $K \subseteq X$ nichtleer und konvex. Sei $F: K \rightarrow K$ stetig. Weiterhin sei K kompakt, oder zumindest sei $F(K)$ relativkompakt in K , d. h. in einer kompakten Teilmenge von K enthalten¹. Dann hat F einen Fixpunkt, d. h. die Gleichung $x = F(x)$ hat eine Lösung.*

Für den Fall, dass K eine abgeschlossene Kugel in $X = \mathbb{R}^n$ ist, ist dies der Brouwersche Fixpunktsatz.

Als bekannteste und typische Anwendung des Satzes von Schauder sei der Satz von Peano genannt:

Satz 2 (Peano). *Falls $f: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig ist, hat das Anfangswertproblem $x'(t) = f(t, x(t))$, $x(0) = 0$, eine lokale Lösung.*

Beweis. Wende den Satz von Schauder an für $X := C([-T, T], \mathbb{R}^n)$, $K := \{x \in X : \|x\| \leq 1\}$, und

$$F(x)(t) := \int_0^t f(s, x(s)) ds.$$

Für genügend kleine T rechnet man leicht nach, dass $F: K \rightarrow K$, und nach Arzelá-Ascoli ist $\overline{F(K)} \subseteq K$ kompakt. Die Lösung der Gleichung $x = F(x)$ löst nach dem HDI das Anfangswertproblem auf $[-T, T]$. \square

Der Satz von Schauder stellt in gewissem Sinne das Herzstück der gesamten Theorie dar. Er ist eine gewisse Idealsituation: Wie einfache Beispiele zeigen, darf in diesem Satz keine einzige der Voraussetzungen ersatzlos gestrichen werden; aber andererseits stellt sich heraus, dass man jede einzelne der Voraussetzungen entscheidend abschwächen kann. Jede dieser Abschwächungen führt in ein anderes Gebiet der Theorie. Daher werden nun die verschiedenen Abschwächungen der Voraussetzungen in verschiedenen Abschnitten beschrieben.

Nur eine Voraussetzung scheint wesentlich zu sein, nämlich die Stetigkeit von F – schließlich basiert ja alles auf topologischen Methoden. Aber bei genauem Hinsehen lässt sich auch diese Voraussetzung abschwächen, wenn man sich damit zufrieden gibt, anstelle einer Lösung der Gleichung $x = F(x)$ nur eine Lösung der Ungleichung $\|x - F(x)\| < \varepsilon$ zu erhalten: In der Tat, der Beweis des Satzes von Schauder fußt auf der Beobachtung, dass man aufgrund der Kompaktheit die Abbildung F durch eine end-

lichdimensionale (nichtlineare) Abbildung approximieren kann; für diese Approximation wendet man dann den Fixpunktsatz von Brouwer an. Man benötigt also tatsächlich nur eine Stetigkeit von F auf gewissen endlichdimensionalen Teilräumen von X , was eine wesentliche Erleichterung sein kann.

2 Retrakte und absolute Retrakte

Zunächst soll diskutiert werden, inwiefern man die Voraussetzung abschwächen kann, dass die Menge K in Satz 1 eine konvexe abgeschlossene Teilmenge eines normierten Raums sein muss. Die erste Beobachtung, die man dabei wohl macht, ist, dass es vollkommen genügt, dass K homöomorph zu einer solchen Menge K_0 ist. In der Tat, wenn $h: K \rightarrow K_0$ ein entsprechender Homöomorphismus ist, so hat die Komposition $h \circ F \circ h^{-1}$ nach dem Fixpunktsatz von Schauder einen Fixpunkt $x_0 \in K_0$, und $h^{-1}(x_0)$ ist dann ein Fixpunkt von F .

Ebenso bleibt der Fixpunktsatz von Schauder aber auch gültig, wenn K nur ein Retrakt einer Menge M wie eben ist. Es sei kurz an die Definition erinnert:

Definition 1. Eine stetige Abbildung $\rho: M \rightarrow M$ eines topologischen Raums M heißt Retraktion, wenn $\rho \circ \rho = \rho$. Das Bild $\rho(M)$ ist ein Retrakt von M .

Anders formuliert: Eine Menge $K \subseteq M$ ist ein Retrakt von M , wenn man die Identitätsabbildung $id: K \rightarrow K$ zu einer stetigen Abbildung $\rho: M \rightarrow K$ fortsetzen kann.

Falls der Fixpunktsatz von Schauder für M gilt, so gilt er auch für jeden Retrakt K von M , denn für die zugehörige Retraktion ρ und stetiges $f: K \rightarrow K$ (mit relativkompaktem $f(K)$) hat die Abbildung $f \circ \rho: M \rightarrow M$ dann einen Fixpunkt $x_0 \in M$, also $x_0 = f(\rho(x_0)) \in K$, was $\rho(x_0) = x_0$ und damit $x_0 = f(x_0)$ impliziert. Diese Überlegungen zeigen, dass der Satz von Schauder sogar für jeden sog. AR-Raum K gilt:

Definition 2. Ein metrischer Raum K heißt *absoluter Retrakt (AR)*, wenn er homöomorph ist zu einem Retrakt einer konvexen Teilmenge M eines normierten Raumes. Ein metrischer Raum K heißt *absoluter Umgebungsretrakt (ANR)*, wenn er homöomorph ist zu einem Retrakt einer (in M) offenen Teilmenge einer konvexen Teilmenge M eines normierten Raumes.

Diese Begriffe wurden von Borsuk (s. z. B. [14]) eingeführt. Die Klasse der ANR-Räume ist immens groß: Jeder „lokal gutartige“ Raum ist ein ANR. Insbesondere ist jede Mannigfaltigkeit ein ANR. Außerdem ist jede endliche Vereinigung konvexer Teilmengen eines normierten Raumes (und jeder dazu homöomorphe Raum) ein ANR. Grob gesprochen kann man einen ANR als die topologische Version einer Mannigfaltigkeit interpretieren. Im Gegensatz zu einer Mannigfaltigkeit sind allerdings „Ecken und Kanten“ erlaubt, und die Dimension muss nicht konstant (oder auch nur definiert oder endlich) sein. Im wesentlichen schließt die Klasse der ANR-Räume nur einige pathologische Räume aus, die häufig zur Konstruktion von Gegenbeispielen herangezogen werden (wie etwa den Graphen der Funktion $x \mapsto \sin \frac{1}{x}$ und ähnliche Räume, bei denen vergleichbare „Häufungseffekte“ auftreten). In der Literatur findet man meist ei-

ne andere Definition von AR- und ANR-Räumen, die hier nur als Satz wiedergegeben wird:

Satz 3. Ein metrischer Raum K ist genau dann ein ANR, wenn es für jeden metrischen Raum X , jede abgeschlossene Teilmenge $A \subseteq X$ und jede stetige Abbildung $f: A \rightarrow K$ eine Umgebung U von A und eine stetige Fortsetzung $F: U \rightarrow K$ von f gibt. K ist genau dann ein AR, wenn man sogar stets eine stetige Fortsetzung $F: X \rightarrow K$ findet.

Beweis. Der Fortsetzungssatz von Dugundji [18] besagt, dass jede konvexe Teilmenge M eines normierten Raums die genannte Fortsetzungseigenschaft besitzt, also dass jede Abbildung $f: A \rightarrow M$ wie oben eine stetige Fortsetzung $F: X \rightarrow M$ besitzt. Ist nun K ein ANR, etwa $V \subseteq M$ offen, $\rho: V \rightarrow N$ Retraktion auf N und $h: N \rightarrow K$ ein Homöomorphismus, so kann man die Abbildung $g := h^{-1} \circ f: A \rightarrow N$ zu einer stetigen Abbildung $G: X \rightarrow M$ fortsetzen; für die Einschränkung G_0 von G auf $U := G^{-1}(V)$ gilt dann: $h \circ \rho \circ G_0: U \rightarrow K$ ist die gesuchte Fortsetzung. Falls K ein AR ist, kann man natürlich $V := M$ wählen und hat folglich $U = X$.

Umgekehrt ist nach dem Einbettungssatz von Arens-Eells [9] jeder metrische Raum K isometrisch zu einer abgeschlossenen Teilmenge N eines normierten Raums M , es gibt also insbesondere einen Homöomorphismus $h: N \rightarrow K$. Hat nun K die genannte Fortsetzungseigenschaft, so lässt sich h zu einer stetigen Abbildung $H: U \rightarrow K$ mit offenem $U \supseteq N$ (bzw. $U = M$) fortsetzen. Damit ist $\rho := h^{-1} \circ H: U \rightarrow N$ die gesuchte Retraktion auf N . \square

Es sei bemerkt, dass der Beweis des Fortsetzungssatzes von Dugundji (und damit des obigen Satzes) massiv vom Auswahlaxiom Gebrauch macht – nur für separables A und vollständiges metrisches M ist ein alternativer Beweis bekannt, für den das abzählbare Auswahlaxiom genügt. Da der Fortsetzungssatz von Dugundji nicht nur für normierte, sondern sogar für lokalkonvexe Räume richtig ist, folgt aus obigem Beweis sogar, dass wir in Definition 2 äquivalent auch lokalkonvexe Räume hätten zulassen dürfen.

Unsere Überlegungen vor Definition 2 haben gezeigt:

Satz 4 (Granas). Sei K ein AR, und $F: K \rightarrow K$ stetig mit relativkompaktem Bild. Dann hat F einen Fixpunkt.

Für Satz 4 hätten wir den Begriff von ANR-Räumen gar nicht gebraucht, sondern die Definition von AR-Räumen hätte ausgereicht. Die Tragweite dieses Satzes wird allerdings besonders deutlich, wenn man folgendes Ergebnis kennt:

Satz 5. Die AR-Räume sind genau die kontrahierbaren ANR-Räume.

Wir erinnern daran, dass ein Raum *kontrahierbar* heißt, wenn die Identität homotop zu einer konstanten Abbildung ist.

Beweis. Sei K ein AR, ohne Einschränkung $K = \rho(M)$ mit einer konvexen Menge M eines normierten Raums und einer Retraktion ρ . Für $x_0 \in M$ ist die Abbildung $H: [0, 1] \times K \rightarrow K$, $H(t, x) := \rho(tx_0 + (1-t)x)$ eine Kontraktion von K .

Sei umgekehrt K ein ANR mit einer Kontraktion $H: [0, 1] \times K \rightarrow K$ auf einen Punkt $x_0 \in K$. Sei X metrischer Raum, $A \subseteq X$ abgeschlossen, und $f: A \rightarrow K$ stetig. Da X nor-

mal und K ein ANR ist, gibt es ein offenes $U \supseteq A$ und eine stetige Fortsetzung $F_0: \overline{U} \rightarrow K$. Nach dem Lemma von Urysohn gibt es eine stetige Funktion $\lambda: X \rightarrow [0, 1]$ mit $\lambda|_A = 0$ und $\lambda|_{\partial U} = 1$. Definiere $F: X \rightarrow K$ durch

$$F(x) := \begin{cases} H(\lambda(x), F_0(x)) & \text{falls } x \in \overline{U}, \\ x_0 & \text{falls } x \notin U. \end{cases}$$

Da F eine stetige Fortsetzung von f definiert, ist K ein AR (Satz 3). \square

Grob gesprochen ist ein AR also ein ANR „ohne endlich-kodimensionale Löcher“. Dies ist besonders überraschend, wenn man bedenkt, dass unendlichdimensionale Sphären hingegen stets kontrahierbar sind [18]. Man erhält so sofort als frappierende Folgerung des Granasschen Fixpunktsatzes:

Korollar 1. *Sei S eine Sphäre in einem unendlichdimensionalen normierten Raum X , und $F: S \rightarrow S$ stetig mit relativkompaktem Bild. Dann hat F einen Fixpunkt.*

Beweis. S ist als Retrakt von $U := X \setminus \{0\}$ ein ANR-Raum, wegen der Kontrahierbarkeit also ein AR-Raum. \square

Korollar 1 gilt auch für die Sphäre S in einem endlichdimensionalen Raum, wenn man zusätzlich voraussetzt, dass $F(S) \neq S$ (was in unendlichdimensionalen Räumen bei kompaktem $\overline{F(S)}$ automatisch erfüllt ist). In der Tat, sei etwa $x_0 \in S \setminus F(S)$. Dann bildet F den Raum $S_0 := S \setminus \{x_0\}$ in sich ab (mit $\overline{F(S_0)} \subseteq F(S) \subseteq S_0$) und S_0 ist ein AR-Raum, denn offensichtlich ist S_0 kontrahierbar und als Retrakt von $U := X \setminus \{\lambda x_0 : \lambda \geq 0\}$ ein ANR-Raum.

Tatsächlich hat A. Granas den Fixpunktsatz sogar für ANR-Räume formuliert. In diesem Fall gilt der Satz aber natürlich nicht für beliebige kompakte Abbildungen (z. B. sind ja Drehungen des Kreises S^1 fixpunktfrei), sondern nur für solche, deren sog. *Lefschetz-Zahl* nicht verschwindet [31]. Auf die (mit Methoden der Homologietheorie definierte) Lefschetz-Zahl soll hier aber nicht weiter eingegangen werden. Es sei allerdings erwähnt, dass im Falle eines azyklischen Raumes (also eines Raumes mit der selben Homologie wie ein einpunktiger Raum – insbesondere ist jeder kontrahierbare Raum azyklisch), die Lefschetz-Zahl stets 1 ist, so dass unser obiger Satz 4 also tatsächlich ein Spezialfall des Satzes von Granas ist.

3 Abschwächung der Kompaktheitsvoraussetzung

Der mathematischen Folklore zufolge soll Schauder (sinngemäß) gesagt haben, dass die Kompaktheitsvoraussetzung zwar nicht schön, aber unverzichtbar sei. Tatsächlich ist bis heute noch kein nützlicher Satz der Theorie bekannt, in dem eine Kompaktheitsvoraussetzung keine wesentliche Rolle spielt: Letztlich basieren alle Beweise der Theorie auf einer Reduktion auf einen endlichdimensionalen Fall, der dann z. B. kombinatorischen Methoden zugänglich ist. Aber schon die obige Formulierung von Satz 1 zeigt, dass man die Kompaktheitsvoraussetzung „verschieben“ kann: Anstelle von Kompaktheit für K zu fordern, genügt es zu fordern, dass die Abbildung F kompakt ist. Zumindest falls K abgeschlossene konvexe Teilmenge eines Banachraum ist (was wir im fol-

genden annehmen), ist der Grund offensichtlich: Da die abgeschlossene konvexe Hülle $K_0 := \overline{\text{conv}}(F(K))$ dann konvex und kompakt ist, kann man die Version des Satzes für kompakte Mengen einfach auf $F: K_0 \rightarrow K_0$ anwenden.

Man könnte jetzt die Frage stellen, ob es z. B. mit einem ähnlichen Argument möglich ist, nur die Kompaktheit einer Iterierten F^n zu fordern (in Anlehnung an den Banachschen Fixpunktsatz). Aber dies ist ein noch ungelöstes Problem, das wohl erstmals von R. D. Nussbaum gestellt wurde (siehe z. B. [50]).

Wie G. Darbo viel früher bemerkt hatte, kann man allerdings eine positive Antwort geben, wenn man anstatt der reinen Iterierten der Abbildung F die Iteration von Mengen unter der Abbildung F betrachtet, wenn man also etwa einen Punkt $x_0 \in K$ festhält und iterativ $K_0 := K$ und $K_n := \overline{\text{conv}}(F(K_{n-1}) \cup \{x_0\})$ für $n = 1, 2, \dots$ definiert. Die Menge $K_\infty := \bigcap_{n=0}^{\infty} K_n$ ist nicht leer (wegen $x_0 \in K_\infty$), konvex, abgeschlossen und es gilt $F: K_\infty \rightarrow K_\infty$. Ist also insbesondere K_∞ kompakt, so besitzt F einen Fixpunkt. Es genügt also insbesondere, die Kompaktheit von einer der Mengen $\overline{F(K_n)}$ zu überprüfen.

Für Anwendungen ist zur Überprüfung der Kompaktheit der Menge K_∞ der Begriff eines *Nichtkompaktheitsmaßes* praktisch. Es gibt hiervon viele, es seien nur zwei erwähnt: Das *Hausdorffsche Nichtkompaktheitsmaß* $\chi_X(A)$ einer Menge A in einem Banachraum X ist das Infimum aller $\varepsilon > 0$, so dass A ein endliches ε -Netz $N \subseteq X$ hat, d. h. dass für jedes $a \in A$ die Ungleichung $\text{dist}(a, N) < \varepsilon$ gilt. Das *Kuratowskische Nichtkompaktheitsmaß* $\alpha(A)$ ist das Infimum aller $\delta > 0$, so dass A in endlich viele Teilmengen A_i mit $\text{diam } A_i < \delta$ zerlegt werden kann. Die beiden Definitionen sind qualitativ äquivalent, aber nicht quantitativ:

$$\chi_X(A) \leq \alpha(A) \leq 2\chi_X(A) \quad (A \subseteq X).$$

Grob gesprochen gilt: Je größer das Nichtkompaktheitsmaß ist, desto „weniger kompakt“ ist die Menge. Genauer kann man für $\gamma \in \{\chi_X, \alpha\}$ nachrechnen:

1. $\gamma(A) < \infty \iff A$ ist beschränkt.
2. $\gamma(A) = 0 \iff \overline{A}$ ist kompakt.
3. $A \subseteq B \implies \gamma(A) \leq \gamma(B)$.
4. $\gamma(A) = \gamma(\overline{\text{conv}} A)$.
5. $\gamma(A + B) \leq \gamma(A) + \gamma(B)$.
6. $\gamma(A \cup B) = \max\{\gamma(A), \gamma(B)\}$.

Seien nun X und Y Banachräume, $M \subseteq X$, und $F: M \rightarrow Y$. Häufig gibt es eine Zahl $L \geq 0$ mit

$$\gamma(F(A)) \leq L\gamma(A) \quad (A \subseteq M \text{ beschränkt}).$$

Es gibt dann stets eine kleinste solche Zahl, und diese nennen wir $[F]^\gamma$. Sie kann als quantitatives Maß für die Kompaktheit von F interpretiert werden:

1. $[F]^\gamma = 0 \iff F$ bildet beschränkte Mengen in relativkompakte ab.
2. $[F + G]^\gamma \leq [F]^\gamma + [G]^\gamma$.
3. $[F]^\alpha \leq L$ falls F Lipschitzstetig ist mit Lipschitzkonstanter L . Falls F eine lipschitzstetige Fortsetzung auf X (mit Konstanter L) besitzt, so ist auch $[F]^\chi \leq L$.

Falls $[F]^\gamma < 1$ heißt F *kondensierend*. Wir können nun den Fixpunktsatz von G. Darbo [16] beweisen:

Satz 6 [Darbo]. *Sei X Banachraum, und $K \subseteq X$ nichtleer, konvex, abgeschlossen und beschränkt. Sei $F: K \rightarrow K$ stetig und kondensierend. Dann hat F einen Fixpunkt.*

Beweis. Für die oben definierten Mengen K_n gilt:

$$\gamma(K_n) = \gamma(\overline{\text{conv}}(F(K_{n-1}) \cup \{x_0\})) = \gamma(F(K_{n-1}) \cup \{x_0\}) = \gamma(F(K_{n-1})) \leq [F]^\gamma \gamma(K_{n-1}).$$

Durch Induktion folgt $\gamma(K_n) \leq ([F]^\gamma)^n \gamma(K_0) \rightarrow 0$, also $\gamma(K_\infty) = 0$, d. h. K_∞ ist kompakt. Wie oben erklärt, folgt die Behauptung also aus dem Fixpunktsatz von Schauder für $F: K_\infty \rightarrow K_\infty$. \square

Beachte, dass jede Kontraktion $[F]^\alpha < 1$ erfüllt. Man erhält also auch den Fixpunktsatz von Banach als Spezialfall. Benutzt man die Abschätzung $[f + g]^\gamma \leq [f]^\gamma + [g]^\gamma$, erhält man aber sogar den folgenden Fixpunktsatz von M. A. Krasnoselskii [36] als Spezialfall, der sowohl den Banachschen als auch Schauderschen Fixpunktsatz auf natürliche Art verallgemeinert:

Satz 7 (Krasnoselskii). *Sei X Banachraum, und $K \subseteq X$ nichtleer, konvex, abgeschlossen und beschränkt. Es sei $F = f + g: K \rightarrow K$ mit einer Kontraktion f und kompaktem $\overline{g(K)}$. Dann hat F einen Fixpunkt.*

Für den ursprünglichen (elementaren) Beweis von Krasnoselskii ist zusätzlich die Voraussetzung $f(K) + g(K) \subseteq K$ notwendig – es ist bislang wohl noch kein Weg bekannt, wie man ohne Benutzung von Nichtkompattheitsmaßen einsehen könnte, dass diese Voraussetzung in Wirklichkeit überflüssig ist.

Tatsächlich gilt Satz 6 – anders als der Fixpunktsatz von Banach – auch noch im Falle $[F]^\gamma = 1$, wenn man voraussetzt, dass $\gamma(F(A)) < \gamma(A)$ für alle nichtkompakten abgeschlossenen Mengen A gilt. Dies hatte B. N. Sadovskii bewiesen, indem er die Folge K_n mit transfiniter Induktion weiter fortsetzte: Dieser Zugang benötigt allerdings das Auswahlaxiom. Es gibt aber einen alternativen Zugang, der sich als viel praktischer herausstellt: Man muss ja nur eine nichtleere kompakte konvexe Menge $K_0 \subseteq K$ mit $F(K_0) \subseteq K_0$ finden. Sei hierzu $x_0 \in K$ fest, und \mathcal{U} das System aller konvexen abgeschlossenen Mengen $U \subseteq K$ mit $x_0 \in U$ und $F(U) \subseteq U$. Dann ist $K_0 := \bigcap \mathcal{U}$ konvex, abgeschlossen und nichtleer, und es gilt

$$(1) \quad \overline{\text{conv}}(F(K_0) \cup \{x_0\}) = K_0.$$

In der Tat, wenn wir die Menge links mit M bezeichnen, so gilt für jedes $U \in \mathcal{U}$, dass $M \subseteq \overline{\text{conv}}(F(U) \cup \{x_0\}) \subseteq U$, mithin $M \subseteq K_0$. Dies impliziert auch $F(M) \subseteq F(K_0) \subseteq M$, also folgt $M \in \mathcal{U}$ und somit $K_0 \subseteq M$; insgesamt also $M = K_0$.

Die Menge K_0 ist nach Konstruktion sogar die kleinste Menge, die (1) erfüllt. Wenn wir sichern können, dass K_0 kompakt ist, hat F einen Fixpunkt (nach Schauder wegen $F(K_0) \subseteq K_0$).

Satz 8 (Sadovskii). *Sei X Banachraum, und $K \subseteq X$ nichtleer, konvex und abgeschlossen. Sei $F: K \rightarrow K$ stetig. Falls für jede abgeschlossene konvexe nichtkompakte Menge $A \subseteq K$ die Beziehung $\gamma(F(A)) \neq \gamma(A)$ für ein $\gamma \in \{\alpha, \chi_X\}$ gilt, dann hat F einen Fixpunkt.*

Beweis. Für K_0 wie oben und $\gamma \in \{\alpha, \chi_X\}$ gilt

$$\gamma(F(K_0)) = \gamma(F(K_0) \cup \{x_0\}) = \gamma(\overline{\text{conv}}(F(K_0) \cup \{x_0\})) = \gamma(K_0).$$

Dies ist nach unserer Annahme nur möglich, wenn K_0 kompakt ist. \square

Tatsächlich ist unser obiger Satz in gewissem Sinne allgemeiner als der Fixpunktsatz von Sadovskiĭ [51]: Es ist auch $\gamma(F(A)) > \gamma(A)$ zulässig, und die Wahl von γ darf von der Menge A abhängen.

Als eine typische Anwendung des Fixpunktsatzes von Darbo sei hier eine Verallgemeinerung des Satzes von Peano für Differentialgleichungen von Funktionen mit Werten in einem Banachraum $(E, |\cdot|)$ skizziert: Es ist bekannt, dass selbst für stetiges und beschränktes $f: [-T, T] \times E \rightarrow E$ das Anfangswertproblem

$$x'(t) = f(t, x(t)), \quad x(0) = 0$$

keine lokale Lösung besitzen muss. Wir werden allerdings im folgenden skizzieren, weshalb eine lokale Lösung immer existieren muss, falls man zusätzlich voraussetzt, dass es ein $L < \infty$ gibt mit $[f(t, \cdot)]^\gamma \leq L$.

Wie im Beweis von Satz 2 wird man hierzu natürlich versuchen, den Satz von Darbo im Raum $X := C([-T, T], E)$ auf die Menge $K := \{x \in X : \|x\| \leq 1\}$ und den Picard-Lindelöf-Operator

$$F(x)(t) := \int_0^t f(s, x(s)) ds$$

anzuwenden. Das Problem hierbei ist, dass die Menge $F(K) \subseteq X$ zwar gleichgradig stetig und beschränkt, aber im allgemeinen dennoch nicht relativkompakt ist. Für gleichgradig stetige beschränkte Mengen $A \subseteq X$ kann man allerdings – ähnlich wie im Beweis des Satzes von Arzelá-Ascoli – nachrechnen, dass

$$\chi_X(A) = \sup_{t \in [-T, T]} \chi_E(\{x(t) : x \in A\})$$

gilt; ohne die Voraussetzung der gleichgradigen Stetigkeit gilt stets „ \geq “. Falls man nun wüsste, dass

$$(2) \quad \chi_E(\left\{ \int_0^t f(s, x(s)) ds : x \in A \right\}) \leq \left| \int_0^t \chi_E(\{f(s, x(s)) : x \in A\}) ds \right|$$

gilt (den Integranden rechts ersetze man ggf. durch eine messbare Majorante), so könnte man damit dann also folgern, dass

$$\begin{aligned} \chi_X(F(A)) &= \sup_{t \in [-T, T]} \chi_E(\left\{ \int_0^t f(s, x(s)) ds : x \in A \right\}) \\ &\leq TL \sup_{s \in [-T, T]} \chi_E(\{x(s) : x \in A\}) \leq TL \chi_X(A), \end{aligned}$$

mithin $[F]^\gamma \leq TL$ gilt (und damit besitzt das Anfangswertproblem dann für genügend kleine T eine Lösung). Leider ist (2) i. a. nicht richtig, falls f nicht gleichmäßig stetig ist, siehe etwa [34]. Man kann jedoch mit einem technischen Aufwand nachweisen, dass (2) für abzählbare Mengen A gültig ist, wenn man rechts noch den Faktor 2 hin-

zufügt (auf diesen Faktor darf man unter gewissen Zusatzvoraussetzungen auch verzichten), siehe etwa [7], [34], [42], [45]. Die Einschränkung auf abzählbare Mengen A kann man auf zweierlei Weise umgehen: Entweder, man beobachtet, dass (2) auch für überabzählbares A richtig bleibt, wenn man rechts ein weiteres mal den Faktor 2 einfügt [7], oder man benutzt eine Verallgemeinerung des obigen Satzes von Sadovskiĭ, bei der man die Kompaktheitsvoraussetzung nur für abzählbare Mengen verifizieren muss, siehe etwa [44] oder [54] (siehe auch Satz 12). Letzteres hat natürlich den Vorteil, dass man ein größeres Existenzintervall für die Lösung nachweisen kann. Für weitere Details und Verallgemeinerungen sei auf [7] oder die Monographien [35], [56] verwiesen. Weitere Nichtkompaktheitsmaße und Anwendungen sind etwa in [1] zu finden.

4 Verzicht auf Selbstabbildungen

Wir werden nun die Voraussetzung $F(K) \subseteq K$ abschwächen. Ein wichtiges Ergebnis in diese Richtung ist der Satz der besten Approximation von Ky Fan [23]:

Satz 9 (Ky Fan). *Sei K konvexe kompakte Menge eines Banachraums X , und $F: K \rightarrow X$ stetig. Dann gibt es ein $x_0 \in K$ mit $\|F(x_0) - x_0\| = \text{dist}(F(x_0), K)$.*

Der Spezialfall $F(K) \subseteq K$ ist gerade der Fixpunktsatz von Schauder.

Eine andere Art der Verallgemeinerung hat sich jedoch als wichtiger herausgestellt, nämlich, wenn man einen Fixpunkt von $F: K \rightarrow X$ sucht, und $K = \overline{\Omega}$ der Abschluss einer offenen Menge $\Omega \subseteq X$ ist: In diesem Fall kann man Abbildungsgradtheorie (oder den sog. Fixpunktindex) benutzen; hierauf wird später eingegangen. Die Konvexität von Ω ist in diesem Fall überflüssig, aber die Konvexität von X ist bedeutsam. Wir beschränken uns hier auf den Fall eines Banachraums X .

Tatsächlich geht diese Theorie aber in vielen Fällen nicht weit genug, denn oft hat man nicht $F(\overline{\Omega}) \subseteq X$, sondern $F: \overline{\Omega} \rightarrow Y$ mit einem ganz anderen Banachraum Y . Dies ist beispielsweise bei allen Differentialoperatoren der Fall: Zwar wird häufig „künstlicherweise“ ein Differentialoperator als ein Operator von einem Banachraum in sich selbst aufgefasst, aber viel natürlicher ist es, z. B. einen Differentialoperator zweiter Ordnung als stetigen(!) Operator von C^2 (oder $H^{2,2}$) nach C (bzw. L_2) aufzufassen.

In dieser Situation ist es natürlich nicht sinnvoll, von Fixpunkten von F zu reden, sondern man ersetzt in der Fixpunktgleichung $x = F(x)$ die Identitätsabbildung auf der linken Seite durch eine andere Abbildung $J: \overline{\Omega} \rightarrow Y$, die sich „ähnlich gutartig“ wie die Identitätsabbildung verhalten soll. Man sucht also allgemeiner *Koinzidenzpunkte*, also Lösungen der Gleichung $J(x) = F(x)$. Die Essenz dieser Idee ist in folgender Definition zusammengefasst, die unabhängig von mehreren Autoren eingeführt wurde, z. B. in [25], [33] (leider auch unter verschiedenen Bezeichnungen; bei Granas z. B. heißt die folgende Klasse „essentiell“).

Definition 3. Sei X ein normaler topologischer Raum, und Y ein topologischer (Hausdorff) Vektorraum. Sei $\Omega \subseteq X$ offen und $J: \overline{\Omega} \rightarrow Y$ stetig. Die Abbildung J heißt

1. 0-zulässig, wenn $0 \notin J(\partial\Omega)$,
2. 0-epi, wenn für jede stetige Abbildung $F: \bar{\Omega} \rightarrow Y$ mit relativkompaktem Bild und $F|_{\partial\Omega} = 0$ die Gleichung $J(x) = F(x)$ eine Lösung besitzt.

Natürlich hat jede 0-epi Abbildung eine Nullstelle (setze $F(x) \equiv 0$). Wir werden bald sehen, dass die 0-epi Abbildungen genau jene sind mit „homotop stabilen“ Nullstellen.

Der Fixpunktsatz von Schauder besagt gerade, dass die Identitätsabbildung 0-epi ist.

Proposition 1 (Normalisierung). *Für einen normierten Raum $X = Y$ und offenes $\Omega \subseteq Y$ ist $J = id$ genau dann 0-zulässig und 0-epi, wenn $0 \in \Omega$.*

Beweis. Wenn $0 \notin \Omega$, so hat $J(x) = 0$ keine Lösung. Wenn $0 \in \Omega$ und $F: \bar{\Omega} \rightarrow Y$ wie oben ist, so setze F trivial fort. Da F eine (genügend große) Kugel kompakt und stetig in sich abbildet, hat F einen Fixpunkt $x_0 = F(x_0) \in \Omega$ nach Satz 1. \square

Die wichtigste Eigenschaft von 0-epi Abbildungen ist die Homotopieinvarianz:

Proposition 2 (Homotopieinvarianz). *Sei $J_1: \bar{\Omega} \rightarrow Y$ eine 0-zulässige homotope Störung von $J_0: \Omega \rightarrow Y$, d. h. es gelte $J_1 = J_0 - H(1, \cdot)$, wobei $H: [0, 1] \times \bar{\Omega} \rightarrow Y$ stetig mit relativkompaktem Bild und $H(0, \cdot) = 0$ sei, und jede der Abbildungen $J_0 - H(t, \cdot)$ sei 0-zulässig. Genau dann ist J_1 0-epi, wenn J_0 0-epi ist.*

Beweis. Sei ohne Einschränkung J_0 0-epi, und $F: \bar{\Omega} \rightarrow Y$ stetig mit relativkompaktem Bild und $F|_{\partial\Omega} = 0$. Wegen der Kompaktheit von $[0, 1]$ ist die Projektion $\pi: [0, 1] \times \bar{\Omega} \rightarrow \bar{\Omega}$ abgeschlossen (siehe etwa [15, Prop. I.8.2]). Daher ist

$$M := \bigcup_{t \in [0, 1]} \{x \in \bar{\Omega} : J_0(x) - H(t, x) = F(x)\} = \pi(\{(t, x) : J_0(x) - H(t, x) = F(x)\})$$

abgeschlossen. Wähle eine stetige (Urysohn-)Funktion $\lambda: X \rightarrow [0, 1]$ mit $\lambda|_{\partial\Omega} = 0$ und $\lambda|_M = 1$ und setze $G(x) := F(x) + H(\lambda(x), x)$. Da J_0 0-epi und $G|_{\partial\Omega} = 0$ ist, hat $J_0(x) = G(x)$ eine Lösung $x_0 \in \Omega$. Es folgt $x_0 \in M$ und daher $\lambda(x_0) = 1$, also $J_1(x_0) = F(x_0)$. \square

Korollar 2. *Eine 0-zulässige Abbildung $J: \bar{\Omega} \rightarrow Y$ ist genau dann 0-epi, wenn jede 0-zulässige homotope Störung eine Nullstelle besitzt.*

Also grob gesprochen: Eine Abbildung ist 0-epi, wenn sie eine „stabile“ Nullstelle besitzt. Diese Eigenschaft wird zur Definition in allgemeineren Situationen (etwa auf ANR-Räumen) benutzt.

Korollar 3. *Falls $J: \bar{\Omega} \rightarrow Y$ 0-epi und 0-zulässig mit abgeschlossenem $J(\partial\Omega)$ ist, so enthält $J(\Omega)$ eine Nullumgebung, nämlich mindestens die Zusammenhangskomponente der 0 in $Y \setminus J(\partial\Omega)$.*

Beweis. Da jedes y aus dieser Komponente mit 0 durch einen Weg f in $Y \setminus J(\partial\Omega)$ verbunden werden kann, ist $J - y$ eine 0-zulässige homotope Störung von J (setze $H(t, x) := J(x) - f(1-t)$), besitzt also eine Nullstelle. \square

Als typische Anwendung der Homotopieinvarianz beweisen wir nun einen Fixpunktsatz für Nicht-Selbstabbildungen:

Korollar 4. („Continuation Principle“ von Leray-Schauder). Sei $X = Y$ normiert, $\Omega \subseteq X$ offen mit $0 \in \Omega$, und $F: \bar{\Omega} \rightarrow Y$ sei stetig mit relativkompaktem Bild. Dann hat entweder F einen Fixpunkt in $\bar{\Omega}$, oder es gibt ein $x_0 \in \partial\Omega$ und ein $\lambda > 1$ mit $F(x_0) = \lambda x_0$.

Beweis. Betrachte die Homotopie $H(t, x) := tF(x)$. Da $J_0 := id$ 0-epi ist, muss entweder $J_1 := J_0 - H(1, \cdot) = J_0 - F$ 0-epi sein (dann hat F einen Fixpunkt), oder J_1 ist keine 0-zulässige homotope Störung (dann hat $x - H(t, x) = 0$ eine Lösung auf $\partial\Omega$ für ein $t \in (0, 1]$). \square

Das folgende einfache Ergebnis macht besonders deutlich, dass es bei 0-epi Abbildungen tatsächlich um die Existenz von Nullstellen geht:

Proposition 3 (Einschränkung). Es sei $J: \bar{\Omega} \rightarrow Y$ 0-epi, und $\Omega_0 \subseteq \Omega$ sei offen und enthalte alle Nullstellen von J . Dann ist $J: \bar{\Omega}_0 \rightarrow Y$ 0-epi.

Beweis. Sei $F: \bar{\Omega}_0 \rightarrow Y$ stetig mit kompaktem Bild und $F|_{\partial\Omega_0} = 0$. Setze F trivial fort. Dann hat $J(x) = F(x)$ eine Lösung $x_0 \in \Omega$, und die Voraussetzung impliziert $x_0 \in \Omega_0$. \square

Korollar 2 und Proposition 3 machen plausibel: Eine Abbildung J hat genau dann eine „physikalisch relevante“ Nullstelle, wenn eine geeignete Einschränkung von J 0-epi ist. Man kommt allerdings nicht umhin, Einschränkungen zu betrachten:

Beispiel 1. Sei $X = Y = \mathbb{R}$. Die Abbildung $J_1(x) := |x|$ ist in keinem Gebiet $\Omega \subseteq \mathbb{R}$ 0-epi (und hat auch keine Nullstelle, die nicht unter „beliebig kleinen Störungen“ verschwinden kann); die Abbildung $J_2(x) := |x| - 1$ hingegen hat etwa auf $\Omega_0 := (-2, 0) \cup (0, 2)$ gleich zwei „stabile“ Nullstellen, eine auf $\Omega_1 := (-2, 0)$, eine andere auf $\Omega_2 := (0, 2)$. Auf dem größeren Gebiet $\Omega := (-2, 2)$ ist J_2 allerdings nicht 0-epi.

Der Grund, weshalb J_2 in obigem Beispiel auf Ω nicht 0-epi ist, ist, dass J_2 in Umgebungen der beiden Nullstellen „entgegengesetzt orientiert“ ist. Topologisch wird dies dadurch ausgedrückt, dass der Abbildungsgrad verschiedenes Vorzeichen hat:

Für den Fall, dass $X = Y$ ein Banachraum, $\Omega \subseteq X$ offen, und J eine 0-zulässige kompakte Störung der Identität ist (d. h. $J = id - F$ wobei $F: \bar{\Omega} \rightarrow Y$ stetig mit relativkompakten Bild ist), so kann man den Abbildungsgrad von Leray-Schauder $\deg(J, \Omega)$ definieren, der ganz analoge Eigenschaften hat wie 0-epi Abbildungen (siehe z. B. [17], [22], [37], [63]):

Satz 10. \deg ist die eindeutig bestimmte Abbildung, die jedem (J, Ω) wie oben eine ganze Zahl zuordnet mit folgenden Eigenschaften:

1. (**Normalisierung**). Für $J = id - c$ ist

$$\deg(J, \Omega) = \begin{cases} 1 & \text{falls } c \in \Omega, \\ 0 & \text{falls } c \notin \bar{\Omega}. \end{cases}$$

2. (**Homotopieinvarianz**). Falls J_1 eine 0-zulässige homotope Störung von J_0 ist, so ist

$$\deg(J_1, \Omega) = \deg(J_0, \Omega).$$

3. (**Additivität**). Falls $\Omega_1, \Omega_2 \subseteq \Omega$ offen und disjunkt sind und $\Omega_1 \cup \Omega_2$ alle Nullstellen von J enthält, so ist

$$\deg(J, \Omega) = \deg(J, \Omega_1) + \deg(J, \Omega_2).$$

Die Additivität impliziert im Falle $\deg(J, \Omega) \neq 0$ insbesondere, dass J eine Nullstelle hat. Wegen der Homotopieinvarianz ist J dann also sogar 0-epi. Da $J = id - F$ automatisch abgeschlossene Mengen auf abgeschlossene Mengen abbildet (insbesondere ist $J(\partial\Omega)$ abgeschlossen), ist 0 dann sogar stets ein innerer Punkt des Bildes (Korollar 3).

In gewissem Sinne könnte man Abbildungsgradtheorie als eine „quantitative“ Version der „qualitativen“ Theorie der 0-epi Abbildungen auffassen: $\deg(J, \Omega)$ „zählt“ die (stabilen) Nullstellen von J in Ω „gemäß ihrer Vielfachheit“ (die allerdings auch negativ sein kann). In obigem Beispiel 1 ist der Abbildungsgrad von $J = J_2$ auf $(0, 2)$ gleich 1 (Normalisierung), und da J auf $(-2, 2)$ nicht 0-epi ist, muss der Abbildungsgrad dort 0 sein; zusammen mit der Additivität folgt, dass J auf $(-2, 0)$ den Abbildungsgrad -1 haben muss.

In ihrer Natur sind die Definitionen grundverschieden: Während 0-epi Abbildungen auf rein homotope Art definiert sind (Korollar 2), basiert Abbildungsgradtheorie auf Homologietheorie (obwohl es für den obigen Spezialfall auch rein analytische Zugänge gibt – diese versagen aber häufig beim Nachweis „feinerer“ Eigenschaften). Die Topologen wissen, dass ein Hauptunterschied zwischen Homotopietheorie und Homologietheorie darin besteht, dass Homotopietheorie nicht das „Ausschneidungssatz“ von Eilenberg-Steenrod [21] erfüllt. Daher ist es nicht allzu überraschend, dass der Abbildungsgrad die obige mächtige Additivitätseigenschaft hat, während für 0-epi Abbildungen nur Proposition 3 gilt. Insbesondere gilt nicht die Umkehrung dieser Proposition, wie das obige Beispiel 1 zeigt – es wäre daher falsch, 0-epi Abbildungen nur als einen „Abbildungsgrad modulo 2“ zu interpretieren.

Da eine triviale Homotopie eine triviale Homologie impliziert, ist es auch nicht allzu überraschend, dass ein nichttrivialer Abbildungsgrad impliziert, dass die entsprechende Abbildung 0-epi ist. Da andererseits der Satz von Hopf aber eine Beziehung zwischen Homotopie- und Homologietheorie darstellt, könnte man auch eine gewisse Umkehrung erwarten. Dies ist in der Tat unter erstaunlich schwachen Zusatzvoraussetzungen richtig (siehe [27], [57]):

Satz 11. Falls $\deg(J, \Omega)$ definiert und Ω zusammenhängend ist, dann gilt:

$$J \text{ ist 0-epi} \iff \deg(J, \Omega) \neq 0.$$

Es ist wesentlich, dass Ω zusammenhängend ist, wie das obige Beispiel 1 zeigt. Von größerer Bedeutung ist aber, dass selbst im Falle $X = Y$ nur für eine kleine Klasse von Abbildungen (nämlich gewisse Störungen der Identität) ein Abbildungsgrad definiert ist, und im Falle $X \neq Y$ gibt es bislang nur in Ausnahmefällen eine Abbildungsgradtheorie, während der Begriff einer 0-epi Abbildung natürlich *immer* definiert ist.

Der Fixpunktsatz von Schauder, der Ausgangspunkt unserer Überlegungen war, betrifft nur den Spezialfall $\deg(J, \Omega) = 1$; und umgekehrt benötigt man in diesem Spezialfall die komplizierte Abbildungsgradtheorie gar nicht, da man sie mit Hilfe der Theorie der 0-epi Abbildungen auf den Satz von Schauder reduzieren kann, ähnlich wie wir es in

Korollar 4 getan haben. Es gibt aber mindestens zwei weitere wichtige Fälle mit $\deg(J, \Omega) \notin \{0, 1\}$:

1. Falls J eine ungerade Abbildung auf Ω mit $0 \in \Omega = -\Omega$ (und $\deg(J, \Omega)$ definiert) ist, so ist $\deg(J, \Omega)$ stets ungerade (insbesondere ungleich 0, aber i. a. auch nicht 1). Dies ist ein Satz von Borsuk. Insbesondere muss jede zu J 0-zulässige homotope Störung im Bild eine Nullumgebung enthalten. Allein mit diesem Sachverhalt kann man ohne großen Aufwand (nur durch Betrachten geeigneter Homotopien) sowohl den berühmten Antipodensatz von Borsuk-Ulam als auch den Satz von der offenen Abbildung beweisen. Letzterer besagt, dass jede Injektion (für die der Abbildungsgrad definiert werden kann) offene Mengen in offene Mengen abbildet. Ein Spezialfall ist der nichttriviale Teil der Fredholmschen Alternative: Jede lineare injektive kompakte Störung der Identität ist offen und damit surjektiv.
2. Im Raum $X = Y = \mathbb{R}^n$ hat die Abbildung $(-id): \Omega \rightarrow Y$ den Abbildungsgrad $(-1)^n$ (falls $0 \in \Omega \subseteq X$), woraus für ungerades n folgt, dass 0-zulässige homotope Störungen von $(-id)$ keine 0-zulässigen homotopen Störungen von id sein können, und umgekehrt: Hieraus kann man den berühmten Satz vom Igel folgern.

Es ist bemerkenswert, dass man für die in 1. skizzierten Beweise der berühmten Sätze nur wissen muss, dass ungerade kompakte Störungen der Identität 0-epi sind: Alles andere folgt dann unmittelbar aus einem Homotopieargument. Dies macht deutlich, weshalb es sinnvoll ist, 0-epi Abbildungen unabhängig von Abbildungsgradtheorie zu studieren. Andererseits benötigt man die Abbildungsgradtheorie, um *nachzuweisen*, dass ungerade Abbildungen 0-epi sind.

5 Mehrwertige Abbildungen

Der Satz von Schauder bleibt auch richtig, wenn F eine oberhalbstetige mehrwertige Abbildung $F: K \multimap K$ ist, deren Bilder $F(x)$ kompakt und

1. konvex oder zumindest
2. azyklisch bzgl. der Čech-Kohomologietheorie (mit rationalen oder zumindest ganzzahligen Koeffizienten) sind.

Selbstverständlich enthält hier der zweite Fall den ersten, aber da die zugehörigen Beweise grundverschieden sind und verschiedene Erweiterungen besitzen, sind beide interessant. Die Beweisidee im ersten Fall ist einfach zu beschreiben: Man zeigt, dass man den Graph von F (oder zumindest einen geeigneten „Teilgraph“) durch Graphen von einwertigen Abbildungen approximieren kann und gewinnt das Ergebnis durch Anwendung des (einwertigen) Fixpunktsatzes von Schauder. Diesen Zugang nennen wir „Approximationsmethode“.

Der Zugang für den zweiten Fall wurde erstmals in [20] beschrieben: Sei $\Gamma \subseteq K \times X$ der Graph von $F: K \multimap K \subseteq X$, und $p: \Gamma \rightarrow K$ und $q: \Gamma \rightarrow X$ seien die kanonischen Projektionen. Genau dann hat F einen Fixpunkt $x \in F(x)$, wenn es ein $y \in \Gamma$ gibt mit $p(y) = q(y)$: Der Fixpunkt ist dann $x = p(y)$. Statt mehrwertige Abbildungen zu betrachten, sucht man also Koinzidenzpunkte der stetigen Abbildungen p und q . Hierbei ist es hilfreich, dass p eine sog. *Vietoris-Abbildung* ist, d. h. p ist stetig, *eigentlich* (d. h.

Urbilder kompakter Mengen sind kompakt) und surjektiv mit azyklischen Fasern $p^{-1}(\{x\})$. Aufgrund eines Satzes von Vietoris induziert p in der Čech-Homologie eine Bijektion. Durch Betrachten der Inversen dieser Bijektion kann man nun (zumindest im Rahmen der Homologietheorie) ganz analog argumentieren, als wenn $F = q \circ p^{-1}$ eine einwertige Abbildung (mit stetigem p^{-1}) wäre.

Dieser Zugang hat den Vorteil, dass er prinzipiell auch Aussagen über Koinzidenzen $p(y) = q(y)$ erlaubt, wenn $p: \Gamma \rightarrow K$ eine Vietoris-Abbildung (mit einem beliebigen Raum Γ) und $q: \Gamma \rightarrow X \supseteq K$ stetig ist. Dies erlaubt es insbesondere auch, zu beweisen, dass der Satz von Schauder gültig bleibt, wenn $F = g \circ G: K \rightarrow K$ die Komposition einer oberhalbstetigen kompakten azyklischen Abbildung G und einer stetigen Abbildung g ist: Betrachte dazu die kanonischen Projektionen p, q_0 des Graphen von G auf die Koordinaten, und setze $q := g \circ q_0$.

Diese Klasse von Abbildungen spielt eine wichtige Rolle bei Differentialgleichungen und -inklusionen, bei denen man keine Eindeutigkeit der Lösung voraussetzt: Es ist bekannt, dass z. B. für oberhalbstetiges beschränktes $f: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit konvexen kompakten Werten die Menge der Lösungen des Anfangswertproblems

$$x'(t) \in f(t, x(t)), \quad x(0) = x_0$$

eine sog. R_δ -Menge im Raum $C([0, T], \mathbb{R}^n)$ ist, d. h. der Durchschnitt einer absteigenden Folge kompakter AR-Räume. Insbesondere ist diese Menge azyklisch. Darüberhinaus ist die Abbildung G , die jedem x_0 die entsprechende Lösungsmenge zuordnet, oberhalbstetig und kompakt. Definiert man $g: C([0, T], \mathbb{R}^n) \rightarrow \mathbb{R}^n$ durch $g(x) := x(T)$, so kann man also den (mehrwertigen) Translationsoperator

$$F: x_0 \mapsto \{x(T) : x \text{ löst das AWP}\}$$

als Komposition $F = g \circ G$ schreiben und den Fixpunktsatz anwenden, um z. B. T -periodische Lösungen zu finden. Beachte, dass $F(x_0)$ selbst i. a. nicht azyklisch ist (aber $G(x_0)$ ist azyklisch).

6 Kombination der Zugänge

Die neuere Forschung in dem beschriebenen Gebiet beschäftigt sich u. a. mit der Frage, wie man die oben skizzierten Ideen kombinieren kann, um beispielsweise einen „Abbildungsgrad“ für nichtkompakte mehrwertige Abbildungen auf ANR-Räumen zu gewinnen. Man spricht im Zusammenhang mit ANR-Räumen meist von einem *Fixpunkt*- bzw. *Koinzidenzindex* statt von einem Abbildungsgrad: Der entscheidende technische Unterschied bei ANR-Räumen ist, dass die beim Abbildungsgrad betrachtete Menge Ω nicht mehr offen in einem Banachraum sein muss, sondern z. B. nur offen bzgl. der Relativtopologie eines invarianten ANR-Teilraums zu sein braucht. Dies ist etwa wichtig, wenn man eine Abbildung des positiven Kegels von L_p hat, da Teilmengen dieses Kegels niemals offen in L_p sind.

Wenn wir uns statt allgemeiner ANR-Teilräume zunächst auf konvexe Teilmengen normierter Räume beschränken, ist die Existenz eines Index für einwertige (nichtkompakte) Abbildungen weitgehend geklärt, siehe etwa [48], [51]; das wohl allgemeinste Ergebnis mit Kompaktheitsvoraussetzungen auf nur abzählbaren Mengen folgt aus [55].

Die Grundidee dabei ist es jeweils, durch geeignete Voraussetzungen (etwa, dass der Operator kondensierend ist), die Existenz einer sog. *fundamentalen* konvexen kompakten Menge zu sichern, auf der sich alles topologisch Relevante abspielt. Für den entsprechenden Abbildungsgrad kondensierender Abbildungen gilt ein analoges Ergebnis zu Satz 11 [27].

Auch der mehrwertige Fall wird in [55] weitgehend abgedeckt, allerdings gibt es eine immer noch ungeklärte technische Schwierigkeit hierbei: Man weiß nicht, ob der Index eindeutig ist, und insbesondere ist unklar, ob der Index eine topologische Invariante ist. Man kann zwar einen topologisch invarianten Index mit anderen Methoden definieren [19], [52], aber es ist ungeklärt, ob dieser Index stets ganzzahlige Werte annimmt und mit der „natürlicheren“ Definition übereinstimmt. Falls die Abbildung allerdings so einfach ist, dass man Approximationsmethoden benutzen kann, kann man damit die Eindeutigkeit eines Index nachweisen [40], [59], und man kann dann auch Satz 11 in gewissem Sinne übertragen [59] (beachte aber die Bemerkungen weiter unten hierzu). Es sei betont, dass diese Approximationsmethoden nicht nur für (mehrwertige) Abbildungen mit konvexen Werten, sondern auch für Abbildungen mit R_δ -Werten (ggf. komponiert mit einwertigen Abbildungen) sowie – unter gewissen Bedingungen an endliche (Überdeckungs-)Dimension – auch für Kompositionen azyklischer Abbildungen anwendbar sind, siehe etwa [2], [3], [10], [30], [38], [39], [41]. Da praktisch jede azyklische Menge, die in Anwendungen auftaucht, auch eine R_δ -Menge ist, rechtfertigt dies im Nachhinein eine Bemerkung von K. Deimling in [17, Remark 24.8(3)], wo er meinte, dass in praktisch allen Fällen azyklischer mehrwertiger Abbildungen ein Approximationssargument benutzt werden kann.

Man ist aber auch an einem Index auf allgemeinen (nichtkonvexen) ANR-Räumen interessiert. Dies ist insbesondere im Zusammenhang mit der *Nielsen-Zahl* bedeutsam, die es erlaubt, Aussagen über Vielfachheiten von Lösungen zu machen, aber die in kontrahierbaren (insbesondere konvexen) Räumen nichts Neues liefert. Eine Indextheorie auf ANR-Räumen wurde für kompakte Abbildungen von Granas [32] entwickelt (für den mehrwertigen Fall siehe die oben zitierten Referenzen [19], [40], [52]; die entsprechende Lefschetz-Zahl – die „Normierung“ des Index – wurde von L. Górniewicz [28] eingeführt). Für nichtkompakte Abbildungen sind die Ideen aus Abschnitt 3 aber kaum direkt übertragbar, da diese auf der konvexen Hülle basieren; beim Index wird die Konvexität der fundamentalen Mengen zusätzlich benötigt, um gewisse Homotopien zu konstruieren, die für die Definition des Index wesentlich sind. Dennoch lässt sich in einfachen Situationen die „Iterationsidee“ aus dem Beweis des Darboschen Fixpunktsatzes verallgemeinern [49] (für mehrwertige Abbildungen siehe z. B. [24], [60]), aber für den allgemeinen Fall und die mächtigere Idee aus dem Beweis des Fixpunktsatzes von Sadowskii bedarf es eines größeren technischen Apparats [5].

Für weiterführende Literatur der Indextheorie ein- und mehrwertiger Abbildungen möge der Leser etwa die Übersichtsartikel [11]–[13] oder die Monographien [29] oder [4] (und deren umfangreiche Referenzen) konsultieren.

Die Theorie der 0-epi Abbildungen ist noch weitaus weniger entwickelt. Es ist allerdings bekannt, dass auch für solche Abbildungen „nichtkompakte“ Ergebnisse existieren. Insbesondere sei hier ein Spezialfall eines Satzes aus [58] wiedergegeben:

Sei K vollständiger metrischer Raum und $\Omega \subseteq K$ offen. Sei Y ein Banachraum, und $J: \overline{\Omega} \rightarrow Y$ sei stetig. Es gebe ein $\delta > 0$ mit

$$(3) \quad \gamma(J(A)) \geq \delta \gamma(A) \quad (A \subseteq \text{abzählbar}).$$

Die Beziehung (3) impliziert, dass J eigentlich ist, und sie bedeutet in gewissem Sinne, dass diese Eigentlichkeit „gleichmäßig“ ist. Es gibt dann ein maximales solches $\delta \in (0, \infty]$, und dieses bezeichnen wir mit $[J]_\gamma$. Man sieht schnell, dass

$$[F + G]_\gamma \geq [F]_\gamma - [G]^\gamma,$$

insbesondere ist im Falle $J = id - F$ mit kondensierendem F stets $[J]_\gamma \geq 1 - [F]^\gamma$.

Satz 12. *Es sei $J: \overline{\Omega} \rightarrow Y$ stetig, 0-zulässig und 0-epi. Sei $F: \overline{\Omega} \rightarrow Y$ stetig mit $F|_{\partial\Omega} = 0$ und*

$$\gamma(F(A)) < [J]_\gamma \gamma(A)$$

für alle abzählbaren Mengen $A \subseteq \Omega$ mit $\gamma(A) > 0$ (was im Falle $\gamma(\Omega) < \infty$ insbesondere gilt wenn $[F]^\gamma < [J]_\gamma$). Dann hat die Gleichung $J(x) = F(x)$ eine Lösung $x \in \Omega$.

Wenn man benutzt, dass $J = id$ 0-epi ist, enthält dieses Ergebnis die früher erwähnte Verallgemeinerung des Satzes von Schauder mit Kompaktheitsbedingungen auf nur abzählbaren Mengen als Spezialfall. Wichtiger ist aber, dass Satz 12 auch impliziert, dass 0-epi Abbildungen J mit $[J]_\gamma > 0$ auch unter nichtkompakten 0-zulässigen homotopen Störungen H mit $[H]^\gamma < [J]_\gamma$ 0-epi bleiben. Insbesondere kann man dies benutzen, um Klassen von 0-epi Abbildungen zu definieren, die unter „kleinen“ (nichtkompakten) Störungen stabil bleiben: Dies spielt eine wichtige Rolle in der Spektraltheorie nicht-linearer Operatoren [6].

Es stellt sich auch heraus, dass eigentliche 0-epi Abbildungen J aufgrund von Approximationsmethoden auch automatisch eine entsprechende „0-epi Eigenschaft“ für mehrwertige Abbildungen F besitzen [8].

Leider sind immer noch fast keine homologen (Abbildungsgrad-)Methoden bekannt, die im Falle verschiedener Banachräume $Y \neq X \supseteq \Omega$ ein Kriterium dafür liefern könnten, dass eine gegebene Abbildung $J: \overline{\Omega} \rightarrow Y$ 0-epi ist. Auf den ersten Blick könnte man meinen, dass man für den Fall einer Vietoris-Abbildung J die in Abschnitt 5 beschriebenen Zugänge für die Koinzidenzgleichung $J(x) = F(x)$ benutzen könnte, aber dies ist nur bis zu einem gewissen Grad richtig: Die Voraussetzungen und Aussagen der Ergebnisse aus diesem Zugang betreffen stets die mehrwertige Abbildung $F \circ J^{-1}$, während man bei der Theorie der 0-epi Abbildungen in gewissem Sinne an der Abbildung $J^{-1} \circ F$ interessiert ist: Beispielsweise gilt Satz 12 i. W. bereits unter schwachen Kompaktheitsvoraussetzungen nur an $J^{-1} \circ \text{conv } F$. Frappierender ist, dass Satz 11 in diesem Zusammenhang zunächst sinnlos ist, da man in einem Fall Mengen $\Omega \subseteq Y$ und im anderen Fall Mengen $\Omega \subseteq X$ betrachtet. Man kann jedoch die Definition der 0-epi Abbildungen geeignet modifizieren, so dass ein Vergleich dennoch möglich ist und Satz 11 richtig bleibt, auch im nichtkompakten Fall [59]. Diese Modifikation scheint mit Kohomologie zusammenzuhängen, ebenso wie auch der Index für $F \circ J^{-1}$ eigentlich auf Kohomologietheorie basiert. Für viele Anwendungen wäre es aber viel nützlicher, wenn

man nicht die Definition von 0-epi Abbildungen modifizieren müsste, sondern auch einen entsprechenden (homologen) Index oder Abbildungsgrad hätte.

Von Skrypnik wurde ein Abbildungsgrad für den Fall $Y = X^*$ entwickelt (der i. W. für monotone Abbildungen anwendbar ist) [53], und von Mawhin [43] gibt es einen Koinzidenzindex für den Fall, dass J eine (kompakte Störung einer) Fredholm-Abbildung mit Index 0 ist. Nirenberg hat als erster auch Fredholm-Abbildungen mit positivem Index (also insbesondere $X = \mathbb{R}^m$ und $Y = \mathbb{R}^n$ mit $n < m$) zugelassen [46], [47], aber letztlich basiert auch diese Definition auf Homotopietheorie. Man kann diesen Mawhin-Nirenberg-Index auch mit den Ideen aus Abschnitt 5 kombinieren, was aber äußerst diffizil ist [26], [41].

Damit sind die bisher bekannten Abbildungsgradtheorien für $X \neq Y$ schon im Wesentlichen erschöpft (einige weitere Zugänge sind im Übersichtsartikel [61] zu finden, der auch eine umfangreiche Bibliographie enthält). Dass das Problem extrem schwer ist, sieht man schon an den obigen Dimensionsvoraussetzungen, die kein Zufall sind: Im Falle $m < n$ kann es keine 0-epi Abbildung geben (und somit auch keinen Abbildungsgrad), da die entsprechenden Homotopiegruppen $\pi_m(S^n)$ trivial sind.

Anmerkung

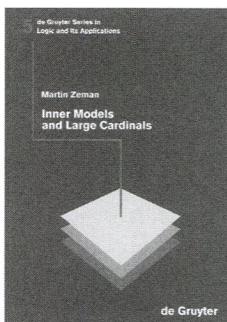
- 1 Anders ausgedrückt: Der Abschluss $\overline{F(K)}$ sei eine kompakte Teilmenge von K ; hierbei ist es egal, ob man den Abschluss in X oder in K nimmt.

References

- [1] Akhmerov, R. R., Kamenskiĭ, M. I., Potapov, A. S., Rodkina, A. E., and Sadovskiĭ, B. N., *Measures of noncompactness and condensing operators*, Birkhäuser, Basel, Boston, Berlin, 1992.
- [2] Andres, J. and Bader, R., *Asymptotic boundary value problems in Banach spaces*, J. Math. Anal. Appl. **274** (2002), 437–457.
- [3] Andres, J., Gabor, G., and Górniewicz, L., *Boundary value problems on infinite intervals*, Trans. Amer. Math. Soc. **351** (1999), no. 12, 4861–4903.
- [4] Andres, J. and Górniewicz, L., *Topological fixed point principles for boundary value problems*, Kluwer Academic Publishers, Dordrecht, 2003.
- [5] Andres, J. and Väth, M., *Coincidence index for noncompact mappings on nonconvex sets*, Nonlinear Funct. Anal. Appl. **7** (2002), no. 4, 619–658.
- [6] Appell, J., de Pascale, E., and Vignoli, A., *Nonlinear spectral theory*, DeGruyter, Berlin, 2003.
- [7] Appell, J., Väth, M., and Vignoli, A., *Compactness and existence results for ordinary differential equations in Banach spaces*, Z. Anal. Anwendungen **18** (1999), no. 3, 569–584.
- [8] —, *\mathcal{F} -epi maps*, Topol. Methods Nonlinear Anal. **18** (2001), 373–393.
- [9] Arens, R. F. and Eells, James, *On embedding uniform and topological spaces*, Pacific J. Math. **6** (1956), 397–403.
- [10] Bader, R. and Kryszewski, W., *Fixed-point index for compositions of set-valued maps with proximally ∞ -connected values on arbitrary ANR's*, Set-Valued Anal. **2** (1994), 459–480.
- [11] Borisovich, Yu. G., Gel'man, B. D., Myshkis, A. D., and Obukhovskii, V. V., *Topological methods in the fixed-point theory of multi-valued maps* (in Russian), Uspekhi Mat. Nauk **35** (1980), no. 1, 59–126, Engl. transl.: Russian Math. Surveys **35** (1980), no. 1, 65–143.
- [12] —, *Multivalued maps* (in Russian), Itogi Nauki-Seriya “Matematika” (Progress in Science-Mathematical Series) **19** (1982), 127–230, Engl. transl.: J. Soviet Math. **24** (1984), 719–791.

- [13] —, *Multivalued analysis and operator inclusions* (in Russian), Itogi Nauki-Seriya “Matematika” (Progress in Science-Mathematical Series) **29** (1986), 151–211, Engl. transl.: J. Soviet Math. **39** (1987), 2772–2811.
- [14] Borsuk, K., *Theory of retracts*, Polish Scientific Publ., Warszawa, 1967.
- [15] Bredon, G. E., *Topology and geometry*, Springer, New York, 1993.
- [16] Darbo, G., *Punti uniti in trasformazioni a codominio non compatto*, Rend. Sem. Mat. Univ. Padova **24** (1955), 84–92.
- [17] Deimling, K., *Nonlinear functional analysis*, Springer, Berlin, Heidelberg, 1985.
- [18] Dugundji, J., *An extension of Tietze’s theorem*, Pacific J. Math. **1** (1951), 353–367.
- [19] Dzedzej, Z., *Fixed point index theory for a class of nonacyclic multivalued mappings*, Dissertationes Math. (Rozprawy Mat.), no. 253, Polish Scientific Publ., Warszawa, 1985.
- [20] Eilenberg, S. and Montgomery, D., *Fixed point theorems for multi-valued transformations*, Amer. J. Math. **68** (1946), 214–222.
- [21] Eilenberg, S. and Steenrod, N., *Foundations of algebraic topology*, Princeton Univ. Press, Princeton, 1952.
- [22] Eisenack, G. and Fenske, C., *Fixpunkttheorie*, BI Wissenschaftsverlag, Mannheim, Wien, Zürich, 1978.
- [23] Fan, K., *Extensions of two fixed point theorems of F. E. Browder*, Math. Z. **112** (1969), 234–240.
- [24] Fournier, G. and Violette, D., *A fixed point index for compositions of acyclic multivalued maps in Banach spaces*, Operator Equations and Fixed Point Theorems (Singh, S. P., Sehgal, V. M., and Burry, J. H. W., eds.), vol. 1, The MSRI-Korea Publications, 1986, 203–224.
- [25] Furi, M., Martelli, M., and Vignoli, A., *On the solvability of nonlinear operator equations in normed spaces*, Ann. Mat. Pura Appl. **124** (1980), 321–343.
- [26] Gabor, D. and Kryszewski, W., *A coincidence theory involving Fredholm operators of nonnegative index*, Topol. Methods Nonlinear Anal. **15** (2000), no. 1, 43–59.
- [27] Giorgieri, E. and Väth, M., *A characterization of 0-epi maps with a degree*, J. Funct. Anal. **187** (2001), 183–199.
- [28] Górniewicz, L., *Homological methods in fixed-point theory of multi-valued maps*, Dissertationes Math. (Rozprawy Mat.), no. 129, Polish Scientific Publ., Warszawa, 1976.
- [29] —, *Topological fixed point theory of multivalued mappings*, Kluwer Academic Publishers, Dordrecht, 1999.
- [30] Górniewicz, L. and Kucharski, Z., *Coincidence of k-set contraction pairs*, J. Math. Anal. Appl. **107** (1985), 1–15.
- [31] Granas, A., *Generalizing the Hopf-Lefschetz fixed point theorem for non-compact ANRs*, Symposium on Infinite Dimension Topology, Bâton-Rouge, 1967, Ann. Math. Studies, vol. 27, 1972, 119–130.
- [32] —, *The Leray-Schauder index and the fixed point theory for arbitrary ANRs*, Bull. Soc. Math. France **100** (1972), 209–228.
- [33] —, *Continuation method for contractive maps*, Topol. Methods Nonlinear Anal. **3** (1994), 375–379.
- [34] Heinz, H.-P., *On the behaviour of measures of noncompactness with respect to differentiation and integration of vector-valued functions*, Nonlinear Anal. **7** (1983), no. 12, 1351–1371.
- [35] Kamenskii, M. I., Obukhovskii, V. V., and Zecca, P., *Condensing multivalued maps and semilinear differential inclusions in Banach spaces*, de Gruyter, Berlin, 2001.
- [36] Krasnoselskii, M. A., *Two remarks on the method of successive approximations* (in Russian), Uspekhi Mat. Nauk **10** (1955), 123–127.
- [37] Krasnoselskii, M. A. and Zabreiko, P. P., *Geometrical methods of nonlinear analysis* (in Russian), Nauka, Moscow, 1975, Engl. transl.: Springer, New York 1984.
- [38] Kryszewski, W., *Topological and approximation methods of degree theory of set-valued maps*, Dissertationes Math. (Rozprawy Mat.), no. 336, Polish Scientific Publ., Warszawa, 1994.
- [39] —, *Some homotopy classification and extension theorems for the class of compositions of acyclic set-valued maps*, Bull. Sci. Math. **119** (1995), 21–48.
- [40] —, *The fixed-point index for the class of compositions of acyclic set-valued maps on ANR’s*, Bull. Soc. Math. France **120** (1996), 129–151.
- [41] —, *Homotopy properties of set-valued mappings*, Univ. N. Copernicus Publishing, Toruń, 1997.

- [42] Kunze, M. and Schlüchtermann, G., *Strongly generated Banach spaces and measures of non-compactness*, Math. Nachr. **191** (1998), 197–214.
- [43] Mawhin, J. L. and Willem, M., *Critical point theory and Hamiltonian systems*, Springer, New York, Berlin, Heidelberg, 1989.
- [44] Mönch, H., *Boundary value problems for nonlinear ordinary differential equations in Banach spaces*, Nonlinear Anal. **4** (1980), no. 5, 985–999.
- [45] Mönch, H. and von Harten, G.-F., *On the Cauchy problem for ordinary differential equations in Banach spaces*, Arch. Math. (Basel) **39** (1982), 153–160.
- [46] Nirenberg, L., *An application of generalized degree to a class of nonlinear problems*, Contributions to Nonlinear Analysis. Proc. Symp. Univ. Wisconsin, Madison, 1971 (New York, London) (Zarantonello, E. H., ed.), Academic Press, 1971, 57–74.
- [47] –, *Generalized degree and nonlinear problems*, 3ieme Coll. sur l’Analyse fonction., Liege 1970 (Louvain, Belgique), Centre Belge de Recherches Matheématiques. Vander éditeur, 1971, 1–9.
- [48] Nussbaum, R. D., *The fixed point index for local condensing maps*, Ann. Mat. Pura Appl. **89** (1971), 217–258.
- [49] –, *Some fixed point theorems*, Bull. Amer. Math. Soc. **77** (1971), 360–365.
- [50] –, *The fixed point index and fixed point theorems*, Topological Methods in the Theory of Ordinary Differential Equations (Berlin, Heidelberg, New York) (Furi, M. and Zecca, P., eds.), Lect. Notes Math., Springer, 1993, 143–205.
- [51] Sadovskiĭ, B. N., *Limit-compact and condensing operators (in Russian)*, Uspekhi Mat. Nauk **27** (1972), no. 1, 81–146, Engl. transl.: Russian Math. Surveys **27** (1972), no. 1, 85–155.
- [52] Siegberg, H. W. and Skordev, G., *Fixed point index and chain approximations*, Pacific J. Math. **102** (1982), 455–486.
- [53] Skrypnik, I. V., *Nonlinear elliptic boundary value problems*, Teubner, Leipzig, 1986.
- [54] Väth, M., *Fixed point theorems and fixed point index for countably condensing maps*, Topol. Methods Nonlinear Anal. **13** (1999), no. 2, 341–363.
- [55] –, *An axiomatic approach to a coincidence index for noncompact function pairs*, Topol. Methods Nonlinear Anal. **16** (2000), no. 2, 307–338.
- [56] –, *Volterra and integral equations of vector functions*, Marcel Dekker, New York, Basel, 2000.
- [57] –, *On the connection of degree theory and 0-epi maps*, J. Math. Anal. Appl. **257** (2001), 223–237.
- [58] –, *Coincidence points of function pairs based on compactness properties*, Glasgow Math. J. **44** (2002), no. 2, 209–230.
- [59] –, *Coepi maps and generalizations of the Hopf extension theorem*, Topology Appl. **131** (2003), 79–99.
- [60] Violette, D. and Fournier, G., *Un indice de point fixe pour les composées de fonctions multivoques acycliques dans des espaces de Banach*, Ann. Sci. Math. Québec **22** (1998), no. 2, 225–244.
- [61] Zabreiko, P. P., *Rotation of vector fields: Definition, basic properties, and calculation*, Topological Nonlinear Analysis II. Degree, Singularity and Variations (Matzeu, M. and Vignoli, A., eds.), Birkhäuser, 1997, 445–601.
- [62] Zeidler, E., *Nonlinear functional analysis and its applications*, vol. III, Springer, New York, Berlin, Heidelberg, 1985.
- [63] –, *Nonlinear functional analysis and its applications*, vol. I, Springer, New York, Berlin, Heidelberg, 1986.
- [64] –, *Nonlinear functional analysis and its applications*, vol. IV, Springer, New York, Berlin, Heidelberg, 1988.
- [65] –, *Nonlinear functional analysis and its applications*, vol. II/A, Springer, New York, Berlin, Heidelberg, 1990.
- [66] –, *Nonlinear functional analysis and its applications*, vol. II/B, Springer, New York, Berlin, Heidelberg, 1990.



M. Zeman
**Inner Models and
Large Cardinals**

Berlin, New York, de Gruyter 2002, 369 S., € 138,-

Zeman gibt in seinem Buch eine Einführung in die Kernmodelltheorie, die ein wichtiger Zweig der modernen Mengenlehre ist. Die Darstellung beruht wesentlich auf handschriftlichen Manuskripten von Jensen, die seit längerer Zeit kursieren. Es ist erfreulich, dass diese Ergebnisse nun endlich in dieser ausgearbeiteten Form vorliegen.

In gewisser Weise ist die Kernmodelltheorie eine natürliche Fortsetzung klassischer Untersuchungen von Gödel. Ein inneres Modell ist eine transitive Klasse, die alle Ordinalzahlen enthält und in der die üblichen Axiome gelten. Gödel zeigte, dass ein kleinstes inneres Modell existiert, welches er das konstruktible Universum L nannte. Dieses ist kanonisch. Gödel konnte zeigen, dass in L die allgemeine Kontinuumshypothese gilt.

Später konnte Jensen den kanonischen Charakter von L wesentlich stärker bestimmen. Er entwickelte seine Feinstruktur von L. Mit ihrer Hilfe lässt sich fast jede kombinatorische Frage für das konstruktible Universum beantworten.

Nun hatte jedoch schon vorher Scott gezeigt, dass in L keine sehr großen Kardinalzahlen existieren können. Genauer gesagt gilt dies schon für messbare Kardinalzahlen. Ihre Existenz ist äquivalent dazu, dass auf einer Menge X ein nichttriviales zweiwertiges Maß existiert, welches auf allen Teilmengen von X definiert ist.

Es stellt sich daher die folgende natürliche Frage. Gibt es ein kanonisches inneres Modell, das eine Feinstruktur besitzt, und in welchem sehr große Kardinalzahlen existieren können? Solche Modelle werden Kernmodelle genannt.

Zeman beschreibt in seinem Buch drei Varianten solcher Modelle, die von wachsender Komplexität sind. Er möchte den Leser schrittweise in diesen komplizierten Gegenstand einführen. Am ausführlichsten behandelt er das Kernmodell K für Maße der Ordnung Null. In diesem kann die Klasse der messbaren Kardinalzahlen unbeschränkt sein, aber nicht viel mehr gelten. Allerdings tauchen in der Konstruktion von diesem K schon viele der Hauptideen auf. Sie wird in Kapitel 7 beschrieben.

Es beginnt mit einer natürlichen induktiven Definition von K. Diese ist analog zum Aufbau der konstruktiven Hierarchie. Es werden nur zusätzlich an geeigneten Stellen eindeutig bestimmte normale Maße auf dem bisher konstruierten Anfangsschritt hinzugenommen. Es ist aber nicht offensichtlich, dass diese Definition das Gewünschte liefert. Um dies zu erreichen, wird eine äquivalente Charakterisierung von K gegeben. Dann werden grundlegende Eigenschaften von K bewiesen, wie z. B. die generische Absolutheit und der schwache Überdeckungssatz.

In den ersten sechs Kapiteln wird der umfangreiche technische Apparat entwickelt, der notwendig ist, um die Ergebnisse von Kapitel 7 zu erlangen. Die Grundlagen in den ersten drei Kapiteln dienen allerdings in ihrer Allgemeinheit auch schon für die Untersuchung der noch höheren Kernmodelle.

Die Kapitel 4 bis 6 sind hingegen auf die Konstruktion des speziellen Kernmodells K aus Kapitel 7 zugeschnitten. Es werden die zugehörigen Mäuse untersucht, d. h. Strukturen, die wie Anfangsabschnitte von K aussehen und iterierbar sind. Iterierbarkeit bedeutet hierbei, dass man immer fundierte Strukturen erhält, wenn man sukzessive Ultrapotenzen und direkte Limites bildet. Diese Eigenschaft liefert die fundamentale Ver-

gleichbarkeit von Mäusen, die man durch Koiteration erhält. Hiermit erhält man das geeignete Kriterium für die eindeutige Wahl der Maße in der Konstruktion von K.

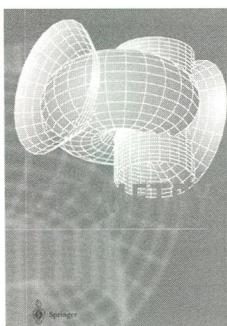
In Kapitel 8 wird ein Kernmodell konstruiert, welches eine starke Kardinalzahl besitzen kann. Hierzu müssen Maße durch Extender ersetzt werden. Da ansonsten viele Argumente analog verlaufen, werden einige Verallgemeinerungen dem Leser überlassen.

Schließlich erhält der Leser in Kapitel 9 einen Einblick in den gegenwärtigen Stand der Forschung. Hier werden Mäuse untersucht, die bei der Konstruktion sehr großer Kernmodelle benutzt werden. Dabei wird insbesondere die Iterierbarkeitsbedingung viel komplizierter. Die Feinstruktur dieser Mäuse wird vollständig behandelt. Die Konstruktion der zugehörigen Kernmodelle wird allerdings nicht mehr durchgeführt.

Insgesamt ist Zeman ein ausgezeichnetes Lehrbuch über Kernmodelltheorie gelungen. Es ist sehr gut zum Selbststudium geeignet. Zeman konzentriert sich ganz auf die Grundlagen der Theorie. Für Anwendungen verweist er aber auf die entsprechenden Originalarbeiten.

München

H.-D. Donder



A. Holme
Geometry, Our Cultural Heritage

Berlin u. a., Springer 2002, 378 S., € 34,95

Die vorliegende Monographie von A. Holme besteht aus zwei Teilen, überschrieben mit *A Cultural Heritage* und *Introduction to Geometry*.

In einem Bogen beginnend mit der Vor- und Steinzeit, über Ägypter, Babylonier und Griechen bis hin zur Neuzeit werden im ersten Teil die Ursprünge und Weiterentwicklungen der Geometrie aufgezeichnet. Dabei werden aber nicht einfach die zur jeweiligen Zeit neu erlangten Erkenntnisse aufgezählt, sondern es werden die mathematischen (damals zumeist geometrischen) Erkenntnisse in die politische bzw. gesellschaftliche Situation eingebettet und mit Anekdoten und Legenden verziert. In einem Abschnitt über den Ursprung der klassischen Probleme wird z. B. erzählt, wie Perikles' Lehrer Anaxagoras sich im Gefängnis sitzend mit der *Quadratur des Kreises* beschäftigt, oder, wie eine Delegation der von der Pest gebeilten und sich zugleich im Krieg mit Spartakus befindlichen Stadt Athen im Orakel von Delos, um Ausweg aus ihrer Miesere fragend den Rat bekommt, den Altar des Apollo zu verdopeln! Dieser ist würfelförmig und so ging die *Verdopplung des Würfels als Delisches Problem* in die Geschichte ein. Darüber hinaus werden die meisten der erwähnten mathematischen Resultate bewiesen, teils mit modernen Methoden. In den meisten Fällen jedoch gab sich der Autor viel Mühe die überlieferten Beweise bzw. Konstruktionen ausführlich in zeitgemäßer mathematischer Sprache vorzustellen.

Wie der Autor selber deutlich sagt, erhebt dieser erste Teil des Buches nicht den Anspruch, eine Abfassung über die Geschichte der Mathematik zu sein. Es soll vielmehr als Motivation und Quelle für Hintergrundwissen für den zweiten Teil des Buches dienen.

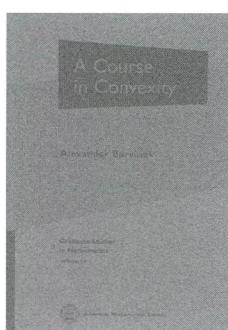
Inhalt des zweiten Teils, *Introduction to Geometry*, ist die mathematische Seite der Geometrie. Bekanntermaßen ist Geometrie ein weiter Begriff, das spiegelt sich auch in der Vielseitigkeit der im zweiten Teil abgehandelten Themen wieder. So finden sich dort Kapitel über Axiomatische Geometrie, Axiomatische Projektive Geometrie, Nicht-euklidische Geometrie, den Projektiven Raum, affine und projektive Geometrie in der Ebene, algebraische Kurven höheren Grades in \mathbb{R}^2 , höhere Geometrie in der pro-

je kiven Ebene, Fraktale Geometrie und Katastrophen Theorie. Das ist natürlich ein umfangreiches Programm für die ca 200 zur Verfügung stehenden Seiten. Dementsprechend werden die einzelnen Themen zügig bzw. nur punktuell abgehandelt. Erstaunlicherweise geht das aber nicht auf Kosten der Verständlichkeit.

Im Gegensatz zu vielen anderen Geometriebüchern wird in dem vorliegenden Band kein Aufgebot an speziellen Bezeichnungen verwandt. So kann man eigentlich in jedem Kapitel anfangen zu lesen, bzw. das Buch auch als Nachschlagewerk verwenden. Damit wird es auch dem vom Autor sich selbst gesetzten Ziel gerecht, die Leser nicht durch eine pedantische und formalistische Presentation von der Dynamik und Schönheit der Geometrie abzulenken. Wie Holme in seiner Einleitung schreibt, will er insbesondere unseren künfigen Lehrern und ebenso einer darüber hinaus gehenden an der Mathematik interessierten Gemeinschaft in seinem Buch ein umfangreiches und abgerundetes Bild der Geometrie präsentieren. Gleichzeitig soll die Monographie auch als Grundlage für Vorlesungen über Geometrie dienen. Meiner Ansicht nach ist der Autor seinem Ziel gerecht geworden.

Erlangen

Ch. Birkenhake



A. Barvinok
A Course in Convexity
Grad. Studies
in Math. 54

Providence, Am. Math. Soc., 2002, 366 S.,
\$ 59,-

Barvinoks Buch ist aus mehreren Graduate Courses an der University of Michigan in Ann Arbor entstanden. Auf seinen gut 360 Seiten enthält es Stoff für mehrere durchaus anspruchsvolle Vorlesungen über diverse Aspekte der Konvexität und Anwendungen dazu, vor allem in Optimierung, aber auch Zählentheorie.

Der übersichtliche und gut gegliederte Aufbau des Buches ermöglicht es auf vielfältige Weise, je nach Neigung und Zielrichtung, eine einsemestrig (4-stündige) Vorlesung daraus zu extrahieren, und Barvinok macht im Vorwort einige Vorschläge dazu.

Programm und Inhalt des Buches werden durch die Titel der acht Kapitel gut charakterisiert:

- 1) Konvexe Mengen allgemein, 2) Seiten und Extrempunkte, 3) Konvexe Mengen in topologischen Vektorräumen, 4) Polarität, Dualität und Linear Programming, 5) Konvexe Körper und Ellipsoide, 6) Seiten (-struktur) von Polytopen, 7) Gitter und konvexe Körper, 8) Gitterpunkte und Polyceder.

Naturgemäß haben die ersten beiden Kapitel einführenden Charakter: Das erste behandelt eher globale Aspekte (z. B. Helly-Typ-Sätze, Euler-Charakteristik), das zweite eher lokale Aspekte (Struktur des Randes) und Trennungssätze sowie Anwendungen auf einfache Optimierungsprobleme.

Bei den übrigen Kapiteln bestätigt sich, was der Autor schon im Vorwort erwähnt, dass es wegen der Stofffülle unmöglich ist, ein Textbook über Konvexität insgesamt zu schreiben. Also ist eine Auswahl notwendig, die naturgemäß subjektiv ist. In der Tat fällt sofort das Fehlen der Brunn-Minkowskischen Theorie auf; dem Herzstück klassischer analytischer Konvexität.

Eher rudimentär findet man dazu einiges in Kap. 1 und 5, z. B. etliche der klassischen Ungleichungen als Probleme und Aufgaben. Trotzdem ist dies kein wesentlicher Nachteil von Barvinoks Buch, denn einerseits gibt es gute Bücher zur Brunn-Minkowskischen Theorie (an der Spitze R. Schneiders gleichnamiges Werk, Cambridge 1993) und ande-

erseits sind diejenigen Bereiche, die Barvinok eingehender behandelt, zumindest teilweise bis an den neuen Stand der Forschung herangeführt.

Dies ist eine der Stärken von Barvinoks Buch: Obwohl es im Kern ein Lehrbuch ist, kommen immer wieder harte oder offene Probleme vor, mit Literaturhinweisen, Tips und gelegentlich auch Teilebeweisen; erkennbar als Apetizer für Studenten und junge (oder auch ältere) Wissenschaftler aufgemacht.

Ein weiteres wichtiges Plus des Buches ist die schon erwähnte enge Verknüpfung zwischen Theorie und Anwendung; der Studierende und Leser erlebt dies konkret an Beispielen und nicht nur als Hinweis.

Parade-Beispiele sind die Anwendung von Dualität auf Linear Programming in Kap. 4 sowie in Kap. 5, der Weg vom Löwner-John-Ellipsoid zur Ellipsoid-Methode von Shor und Khachian und in Kap. 3 der Weg von Konvexität in topologischen Vektorräumen mit dem Krein-Milman-Theorem zu Anwendungen in Kontrolltheorie und dem (Konvexitäts-)Satz von Lyapunov. Das Kap. 6 über die Seitenstruktur von Polytopen ist in sich abgeschlossen. Sein Kern sind die Euler-Poincarésche Identität, die Dehn-Sommerville-Gleichungen und McMullens Upper Bound Theorem.

Die auffälligsten Kapitel des Buches sind sicher die letzten beiden, da alle vorherigen zumindest teilweise zum Pflichtprogramm eines Buches oder einer Vorlesung über Konvexität gehören, während diese eher in Büchern über Diskrete Geometrie oder Geometrie der Zahlen zu finden sind. Diese beiden Kapitel, etwa ein Viertel des Buches, bilden quasi das Kürprogramm. Hier, insbesondere in Kapitel 8, hat der Autor im Laufe des letzten Jahrzehnts Bemerkenswertes beigetragen.

Das 7te Kapitel, anfangs eine Einführung in Geometrie der Zahlen, enthält neben den Standardbegriffen die Sätze von Blichfeldt, Minkowski und Minkowski-Hlawka (letzteren leider nicht mit dem eleganten Rogers-Davenport-Beweis), das „Flatness“ Theo-

rem und Reduktionstheorie, insbesondere die für diskretes Optimieren wichtige Lenstra-Lenstra-Lovasz Reduktion.

Während das 7te Kapitel eher metrische Aspekte behandelt, wird im letzten Kapitel die Gitterpunktanzahl konvexer Polyeder untersucht.

Hauptwerkzeug sind hier die erzeugenden Funktionen (Exponentialsummen), und der Autor entwickelt die in den letzten Jahren von Stanley, McMullen, Brion, Pommersheim und ihm selbst entwickelte Theorie der Algebra rationaler Polyeder, bei der Klassiker wie das Ehrhart-Polynom und das zugehörige Reziprozitätsgesetz beiläufig als Nebenergebnisse absfallen, und mit dieser kraft- und eindrucksvollen Theorie endet das Buch.

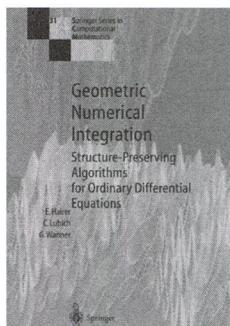
Zur Darstellung sei bemerkt, dass die einleitenden Überblicke zu Beginn eines jeden Kapitels oder Abschnitts sowie die historischen Bemerkungen mit Literaturhinweisen am Ende sehr hilfreich sind. Ebenso erleichtern die gut über das Buch verteilten Abbildungen das Verständnis und den Überblick.

Das Literaturverzeichnis ist angemessen, reichhaltig, doch nicht überbordend. Dagegen hätte man sich das Stichwortverzeichnis etwas detaillierter und ausführlicher gewünscht.

Insgesamt kann aber nach der Lektüre von Barvinoks Buch der Bemerkung des Autors im Vorwort nur zustimmen, dass das Schreiben eines Buches über Konvexität eine erfreuliche Erfahrung ist. Man spürt, dass er diese Freude am Thema an den Leser weitergeben möchte, was ihm auch weitgehend gelungen ist.

Siegen

J. M. Wills



E. Hairer, C. Lubich,
G. Wanner
**Geometric Numerical
Integration**
Comput. Math. 31

Berlin u. a., Springer, 2002, 515 S., \$ 84,95

Seit bahnbrechenden Arbeiten von Feng Kang (Beijing) und Sanz-Serna (Valladolid) haben sich symplektische Diskretisierungen Hamiltonscher Differentialgleichungen als äußerst aktives Forschungsfeld im Gebiet „Numerische Mathematik von Anfangswertproblemen bei gewöhnlichen Differentialgleichungen“ entfaltet. Aus Sicht der Anwendungen spielen Hamiltonsche Differentialgleichungen insbesondere in der Astronomie und der Moleküldynamik eine Rolle. Der vorgelegte Band erweitert das Spektrum auf symmetrische Diskretisierungen allgemeiner reversibler Differentialgleichungen. Diskretisierungen dieses Typs verdienen besonderes Augenmerk mit Blick auf die Erhaltung von Invarianten (wie etwa der Gesamtenergie eines mechanischen Systems), zumindest im asymptotischen Mittel über exponentiell lange Zeiten.

Das Buch gliedert sich in 14 Kapitel, ein umfangreiches Literaturverzeichnis und einen hilfreichen Index. Es enthält zahlreiche Übungsaufgaben.

Zunächst werden numerische Beispiele Hamiltonscher Differentialgleichungen vorgestellt und der Effekt symplektischer und nichtsymplektischer Diskretisierungen illustriert. Die Arnoldsche Katze veranschaulicht den diskreten Fluss im Vergleich mit dem kontinuierlichen. Sodann wird das Instrumentarium zur Konstruktion von Runge-Kutta-Integratoren höherer Ordnung bereitgelegt, wobei natürlicherweise partitionierte

Runge-Kutta-Methoden (PRK) besonders zum Zug kommen; erstaunlicherweise existiert hier eine Alternative zur Herleitung mittels PRK-Bäumen, nämlich die Erzeugung aus der Baker-Campbell-Hausdorff-Formel, wobei die erhaltenen Bedingungsgleichungen zwar unterschiedlich, aber natürlich äquivalent sind. Später werden die PRK auf symmetrische Verfahren für reversible Differentialgleichungen spezialisiert. Dieses Kapitel ist etwas technisch, jedoch für die Konstruktion von Verfahren höherer Ordnung unverzichtbar.

Ein zentraler Punkt ist naturgemäß die Theorie der Hamiltonschen Invarianten (nach Sophus Lie) und ihre Umsetzung ins Diskrete. Hier werden die schon erwähnten Arbeiten von Sanz-Serna und seiner Schule sowie der Autoren mit ihrem Anhang in einheitlichem Gewand dargestellt. Die ‚diskreten variationellen Integratoren‘ nach Marsden et al. werden kurz gestreift. Für den Geschmack des Rezensenten kommt dieses alternative Konzept etwas zu kurz, obwohl es in der Praxis extrem erfolgreich ist und gewisse Einschränkungen der konkurrierenden Verfahren zu umgehen erlaubt. Zum Thema Differentialgleichungen auf Mannigfaltigkeiten, insbesondere zu holonomen und nichtholonomen Zwangsbedingungen bei mechanischen Systemen, hätte sich ein Ausflug in die Theorie von Rabier und Rheinboldt sicherlich gelohnt (jüngst ebenfalls als Monographie erschienen, aber erstaunlicherweise nicht zitiert). Besonders verdienstvoll ist die Darstellung von Details der Implementierung symplektischer und symmetrischer Integratoren, die oft von Anwendern übersehen werden und dann die ganze schöne zugrundliegende Theorie zur Wirksamlosigkeit verdammten.

Aus Sicht der Numerischen Analysis wichtig ist die moderne Interpretation symplektischer Integratoren über eine Rückwärtsanalyse, bei der im Diskreten eine leicht gestörte schrittweitenabhängige Hamiltonfunktion generiert wird. (Allerdings sollte der Leser hierzu nicht die gleiche Erwartung hegen, die er vielleicht von Problemen der numerischen

linearen Algebra mitbringt: in der Moleküldynamik etwa können die in der Theorie implizit angegebenen Schrittweitenumgebungen verschwindend klein sein verglichen mit typischen Zeitkonstanten der Probleme; bei keplerartigen Problemen tritt dieses Phänomen allerdings in der Regel so nicht auf.) Mit dieser Interpretation können die Autoren sodann eine Störungstheorie Hamiltonscher Systeme auf die spezielle Störung durch symplektische Diskretisierungen anwenden. Dabei gehen sie von einer äußerst transparenten Darstellung der KAM-Theorie aus und modifizieren sie geeignet, um sie im Diskreten anwenden zu können. Anschließend spielen sie das gleiche Spiel nochmals im komplizierteren Fall symmetrischer Diskretisierungen reversibler Probleme. Dazu können die Autoren aus dem Vollen schöpfen, da diese Thematik durch jüngste eigene Arbeiten etabliert worden ist. Die Übertragung auf dissipative Störungen Hamiltonscher Systeme folgt schließlich demselben Muster – ganz offenbar ein äußerst fruchtbarer theoretischer Zugang!

Dem Kapitel über hochoszillatorische Differentialgleichungen kommt besondere Bedeutung mit Blick auf die diskretisierte Schrödinger-Gleichung zu. Hier öffnet sich ein interessantes neues Feld der Numerischen Mathematik, das in den nächsten Jahren gewiss noch mehr Aufmerksamkeit erlangen wird. Das letzte Kapitel über Mehrschrittverfahren zeigt deutlich, wie schwierig oft die Übertragung physikalischer oder analytischer Prinzipien bei diesem Typus von Diskretisierung sein kann: so bedarf schon die Frage, was symplektisch ist, in diesem Fall einer genaueren Überlegung. Dementsprechend ist dieses Kapitel etwas technisch und nur für Spezialisten geeignet.

Zusammenfassend ist zu sagen: Wie von diesen vielgelesenen und vielzitierten Autoren nicht anders erwartet, ist das Buch eine bibliophile Kostbarkeit (zahlreiche Illustrationen mit historischem Bezug), professionell geschrieben (äußerste Klarheit in Darstellung und Diction) und voller mathematisch interessanter Querbezüge. Wie schon in

ihren bisherigen Büchern, ist es den Autoren auch hier wiederum gelungen, ein international sichtbares Standardwerk zum Thema zu verfassen. Faszinierend ist die Breite des behandelten Stoffs, vom Kerngebiet Numerik weit hinein in die Analysis und in die klassische Mechanik. Der Band eignet sich als Grundlage von Spezialkursen der Numerischen Mathematik im Hauptstudium und als begleitendes Material für Standardkurse in Numerik und Analysis.

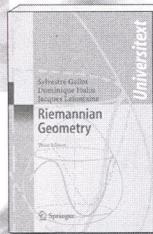
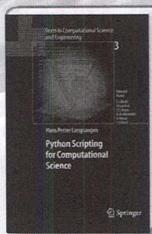
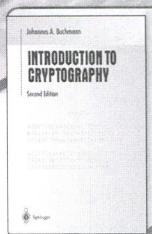
Hätte der Rezensent noch einen Wunsch frei gehabt, so hätte er sich gewünscht, dass sich die dargestellten theoretischen Konzepte gegen Ende des Buches weniger stark an integrable bzw. fast-integrable Systeme angelehnt hätten: solche Probleme treten zwar in der Astronomie, aber kaum in der Moleküldynamik auf; zum von den Autoren gewählten Blickwinkel passt denn auch, dass der größte Teil der zahlreichen illustrierten numerischen Beispiele vom Typ allgemeines Keplerproblem ist. Sowie man diesen engeren Problemkreis verlässt, ergeben sich neue zentrale Fragen, deren Diskussion in die Stochastik und die Ergodentheorie führt – vielleicht ohnehin ein zu weites Feld. Von diesem Blickwinkel aus hätte allerdings die Erhaltung von Invarianten im Diskreten eine ganz andere, zum Teil relativierte Bedeutung.

Dies trübt natürlich nicht den Gesamteinindruck: ein wunderschönes Buch zu einem wichtigen modernen Gebiet der Numerischen Analysis.

Berlin

P. Deuflhard

Textbooks from Springer



J. Buchmann

Introduction to Cryptography

From the reviews: "It gives a clear and systematic introduction into the subject whose popularity is ever increasing, and can be recommended to all who would like to learn about cryptography. The book contains many exercises and examples."

ZENTRALBLATT MATH

2nd ed. 2004. XVI, 335 p. (Undergraduate Texts in Mathematics) Hardcover **€ 79,95; sFr 135,50; £ 61,50**
ISBN 0-387-21156-X

Also available in softcover

€ 39,95; sFr 73,00; £ 30,50 ISBN 0-387-20756-2

H. P. Langtangen

Python Scripting for Computational Science

The author teaches you how to develop tailored, flexible, and efficient working environments built from small programs written in Python. The focus is on examples and applications of relevance to computational science.

2004. XX, 726 p. (Texts in Computational Science and Engineering. Vol. 3) Hardcover **€ 49,95; sFr 88,50; £ 38,50**
ISBN 3-540-43508-5

S. Gallot, D. Hulin, J. Lafontaine

Riemannian Geometry

This book, based on a graduate course on Riemannian geometry and analysis on manifolds, covers the topics of differential manifolds, Riemannian metrics, connections, geodesics and curvature, with special emphasis on the intrinsic features of the subject.

3rd ed. 2004. XVI, 322 p. (Universitext) Softcover
€ 34,95; sFr 64,00; £ 27,00
ISBN 3-540-20493-8

Please order from

Springer · Customer Service
Haberstr. 7 · 69126 Heidelberg, Germany
Tel.: +49 (0) 6221 - 345 - 0
Fax: +49 (0) 6221 - 345 - 4229
e-mail: SAG-bookorder@springer-sbm.com
or through your bookseller

All Euro and GBP prices are net-prices subject to local VAT, e.g. in Germany 7% VAT for books and 16% VAT for electronic products. Prices and other details are subject to change without notice. d&p - 011284x

D. Bump

Lie Groups

This book is intended for a one year graduate course on Lie groups. Rather than providing a comprehensive treatment, the author emphasizes the beautiful representation theory of compact groups.

2004. Approx. 410 p. 50 illus. (Graduate Texts in Mathematics. Vol. 225) Hardcover **€ 64,95; sFr 115,00; £ 50,00**
ISBN 0-387-21154-3

R. K. Guy

Unsolved Problems in Number Theory

About the first Edition: "...many talented young mathematicians will write their first papers starting out from problems found in this book."

András Sárközi, *MathSciNet*

3rd ed. 2004. XVIII, 437 p. 18 illus.

(Problem Books in Mathematics) Hardcover
€ 69,95; sFr 123,50; £ 54,00
ISBN 0-387-20860-7

J. D. Logan

Applied Partial Differential Equations

This text is written for the standard, one-semester, undergraduate course in elementary partial differential equations. The topics include derivations of some of the standard equations of mathematical physics and methods for solving those equations on bounded and unbounded domains.

2nd ed. 2004. XII, 209 p. 40 illus. (Undergraduate Texts in Mathematics) Hardcover **€ 79,95; sFr 135,50; £ 61,50**
ISBN 0-387-20935-2

Also available in softcover

€ 39,95; sFr 73,00; £ 30,50 ISBN 0-387-20953-0

springeronline.com

 Springer

Teubner Lehrbücher: einfach clever

Eberhard Zeidler (Hrsg.)

Teubner-Taschenbuch der Mathematik

2., durchges. Aufl. 2003. XXVI, 1298 S. Geb.

€ 34,90

ISBN 3-519-20012-0

Formeln und Tabellen - Elementarmathematik - Mathematik auf dem Computer - Differential- und Integralrechnung - Vektoranalysis - Gewöhnliche Differentialgleichungen - Partielle Differentialgleichungen - Integraltransformationen - Komplexe Funktionentheorie - Algebra und Zahlentheorie - Analytische und algebraische Geometrie - Differentialgeometrie - Mathematische Logik und Mengentheorie - Variationsrechnung und Optimierung - Wahrscheinlichkeitsrechnung und Statistik - Numerik und Wissenschaftliches Rechnen - Geschichte der Mathematik

Grosche/Ziegler/Zeidler/
Ziegler (Hrsg.)

Teubner-Taschenbuch der Mathematik. Teil II

8., durchges. Aufl. 2003. XVI, 830 S. Geb.

€ 44,90

ISBN 3-519-21008-8

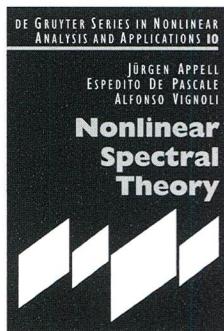
Mathematik und Informatik - Operations Research - Höhere Analysis - Lineare Funktionalanalysis und ihre Anwendungen - Nichtlineare Funktionalanalysis und ihre Anwendungen - Dynamische Systeme, Mathematik der Zeit - Nichtlineare partielle Differentialgleichungen in den Naturwissenschaften - Mannigfaltigkeiten - Riemannsche Geometrie und allgemeine Relativitätstheorie - Liegruppen, Liealgebren und Elementarteilchen, Mathematik der Symmetrie - Topologie - Krümmung, Topologie und Analysis

Stand Juli 2004.
Änderungen vorbehalten.
Erhältlich im Buchhandel
oder beim Verlag.



B. G. Teubner Verlag
Abraham-Lincoln-Straße 46
65189 Wiesbaden
Fax 0611.7878-400
www.teubner.de

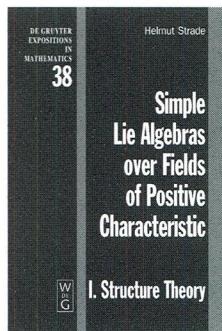
Just Released



Jürgen Appell, Espedito De Pascale, Alfonso Vignoli

■ Nonlinear Spectral Theory

2004. XI, 408 pages. 1 fig. Num. tab. Cloth.
€ [D] 148.00 / sFr 237.00 /
for USA, Canada, Mexico US\$ 148.95.
ISBN 3-11-018143-6
de Gruyter Series in Nonlinear Analysis and Applications 10

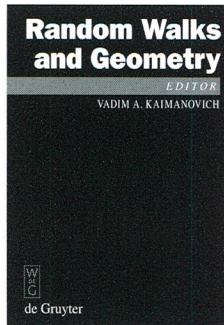


Helmut Strade

■ Simple Lie Algebras over Fields of Positive Characteristic

I. Structure Theory

2004. VIII, 540 pages. Cloth.
€ [D] 128.00 / sFr 205.00 /
for USA, Canada, Mexico US\$ 99.95
ISBN 3-11-014211-2
de Gruyter Expositions in Mathematics 38



■ Random Walks and Geometry Proceedings of a Workshop at the Erwin Schrödinger Institute, Vienna, June 18 - July 13, 2001

Ed. by Vadim A. Kaimanovich
In collab. with Klaus Schmidt,
Wolfgang Woess

2004. X, 532 pages. Cloth.
€ [D] 168.00 / sFr 269.00 /
for USA, Canada, Mexico US\$ 168.95.
ISBN 3-11-017237-2
de Gruyter Proceedings in Mathematics



de Gruyter
Berlin · New York

Prices are subject to change.

WWW.DEGRUYTER.DE

