D. Papenfuß, D. Lüst, W. P. Schleich (eds.)

100 Years Werner Heisenberg

Works and Impact

WILEY-VCH

100 Years Werner Heisenberg

Works and Impact

WILEY-VCH

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

100 Years Werner Heisenberg

Works and Impact

Edited by
D. Papenfuß, D. Lüst, W. P. Schleich

WILEY-VCH

Editors:

Dr. Dietrich Papenfuß
Alexander von Humboldt-Stiftung, Bonn

Professor Dr. Dieter Lüst
Humboldt-Universität zu Berlin

Professor Dr. Wolfgang P. Schleich
Universität Ulm

The publication of this work has been supported in part by the
Alexander von Humboldt-Foundation, Bonn/Germany.

This book was carefully produced. Nevertheless, editors, authors and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Cover Picture: Werner Heisenberg
Source: Rita Strothjohann/SV-Bilderdienst

Library of Congress Card No. applied for

A catalogue record for this book is available from the British Library

Die Deutsche Bibliothek – CIP Cataloguing-in-Publication Data
A catalogue record for this publication is available from
Die Deutsche Bibliothek

ISBN 3-527-40392-2

© Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2002

Printed on acid-free paper.

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such are not to be considered unprotected by law.

Composition: Druckhaus "Thomas Müntzer" GmbH, Bad Langensalza
Printing: betz-druck GmbH, Darmstadt
Bookbinding: Großbuchbinderei J. Schäffer GmbH & Co. KG, Grünstadt
Printed in the Federal Republic of Germany

Editorial

Two theories have dominated the physics of the Twentieth Century: Relativity and Quantum Mechanics. Whereas relativity was the work of a single person, Albert Einstein, Quantum Mechanics has many fathers: Max Born, Paul Adrian Maurice Dirac, Werner Heisenberg, Pascal Jordan, and Erwin Schrödinger. However, Quantum Mechanics as we know it today was triggered by Heisenberg's lonely night on the island of Helgoland where he invented matrix mechanics without knowing the concept of a matrix. Heisenberg is the undisputed leader in the new field of Quantum Physics. He not only wore the hat of the scientist, but also played a major role as a science organizer, eventually becoming the president of the Alexander von Humboldt-Foundation.

The year 2001 marked the Centennial of Heisenberg's birth. Honoring the close ties between the Alexander von Humboldt-Foundation and Heisenberg, the Foundation decided to organize the conference *100 Years Werner Heisenberg – Works and Impact* in Bamberg, Germany on September 26−30, 2001 The meeting emphasized his many major contributions to physics and at the same time illuminated the latest developments in Quantum Physics. It consisted of a popular talk by Anton Zeilinger (Vienna) for the public, plenary talks putting Heisenberg in the proper historical perspective and three parallel sessions dedicated to Elementary particles, Quantum Physics, and Quantum Field Theory and Gravitation. One of the many highlights of the meeting was an impressive rehearsed reading by the original London cast of the play "Copenhagen" by Michael Frayn which prompted a lively discussion between the audience, the author, and the actors, which lasted deep into the night.

The present issue contains selected papers presented at this Heisenberg Centenial meeting and which represent the spirit of the meeting. We would like to take the opportunity to thank the Alexander von Humboldt-Foundation for its generous suppport. Likewise, we are very grateful to Frau Ulrika Holdefleiß-Walter for her great efficiency and patience in organizing this meeting.

Dietrich Papenfuß
Dieter Lüst
Wolfgang P. Schleich

Contents

Opening-speech 1
Wolfgang Frühwald

Plenary Talks

Heisenberg and the Framework of Science Policy 3
Cathryn Carson

Werner Heisenberg: An Overview of His Life and Work 8
David C. Cassidy

Observability, Anschaulichkeit and Abstraction:
A Journey into Werner Heisenberg's Science and Philosophy 11
Jan Lacki

100 Years Werner Heisenberg – Works and Impact 30
Kazuo Yamazaki

Heisenberg-Einstein Context Principle and the Dynamic Core-context
of Discovery in Physics. 32
Giridhari Lal Pandit

Elementary Particles Theory

On the Large Nc Expansion in Quantum Chromodynamics 54
William A. Bardeen

Baryon Asymmetry, Dark Matter and the Hidden Sector 60
Luís Bento and Zurab Berezhiani

Time Asymmetric Quantum Theory and the Z-boson Mass and Width 67
Arno R. Bohm and Piotr Kielanowski

The Heisenberg Matrix Formulation of Quantum Field Theory 74
Stanley J. Brodsky

Nonlocal Structure of Higher Spin Fields 83
Anton Z. Capri

Fundamental Constants at High Energy 89
Harald Fritzsch

Quark Number Susceptibilities from Lattice QCD 96
Rajiv V. Gavai

Elementary Particles on a Dedicated Parallel Computer 102
Kazuyuki Kanaya

Why Extra Gauge Bosons Should Exist and How to Hunt Them 109
Arnd Leike

States of Strongly Interacting Matter 112
Helmut Satz

From Heisenberg to Supersymmetry 123
Mikhail A. Shifman

The Baryon Density Through the (Cosmological) Ages 133
Gary Steigman

Noncommutativity of Lepton Mass Matrices: Flavor Mixing and CP Violation 140
Zhi-zhong Xing

Quantum Physics

Heisenberg's Uncertainty Relations and Quantum Optics 146
Girish Saran Agarwal

Rubidium Condensate for Quantum Optics Studies 154
Ennio Arimondo

Interference of Spontaneously Emitted Photons 165
Almut Beige, Christian Schön, and Jiannis Pachos

Quantum Fictitious Forces 170
I. Białynicki-Birula, M. A. Cirone, J. P. Dahl, T. H. Seligman,
F. Straub, and W. P. Schleich

Quantum Teleportation 179
Samuel L. Braunstein

Heisenberg's Introduction of the "Collapse of the Wavepacket"
into Quantum Mechanics. 185
Raymond Y. Chiao and Paul G. Kwiat

The Hamiltonian Mass and Asymptotically Anti-de Sitter Space-times 195
Piotr T. Chruściel

The Bohr-Heisenberg Correspondence Principle Viewed from Phase Space 201
Jens P. Dahl

The Classical Atom: Stabilization of Electronic Trojan Wavepackets 207
David Farrelly, Ernestine A. Lee, and T. Uzer

Heisenberg's Equations in Laser Theory. A Historical Overview 213
Hermann Haken

Quantum Mechanics from a Heisenberg-type Equality 217
Michael J. W. Hall and Marcel Reginatto

Entanglers: Beam Splitters and Thermal Fields 223
Myung Shik Kim

Enhancing Otto-mobile Efficiency via Addition of a Quantum Carnot Cycle 228
Tomáš Opatrný and Marlan O. Scully

Atom Optics – from de Broglie Waves to Heisenberg Ferromagnets 235
Han Pu, Chris Search, Weiping Zhang, and Pierre Meystre

Quantum Field Theory and Gravitation

On Tachyon Condensation in String Theory: Worldsheet Computations 241
Oleg Andreev

Duality in the Low Dimensional Field Theory 246
Ivan Andrić and Danijel Jurman

Monopoles in Space-time Noncommutative Born-Infeld Theory 251
Paolo Aschieri

Canonical Quantization, Twistors and Relativistic Wave Equations 257
Iwo Białynicki-Birula

Geometry of Non-supersymmetric Strings 258
Anamaría Font

Observing Quanta on a Cosmic Scale 265
Craig J. Hogan

Supersymmetric Configurations in Gauged Supergravity 270
Swapna Mahapatra and Harvendra Singh

Anomalies and Schwinger Terms in NCG Field Theory Models 276
Jouko A. Mickelsson

On Gravitational Interaction of Fermions 282
Yuri N. Obukhov

New Einstein-Hilbert Type Action for Unity of Nature 288
Kazunari Shima

The Heisenberg Algebra and Spin 295
Dmitri Sorokin

Fortschr. Phys. **50** (2002) 5–7, 430–431

Opening-speech

Dear Humboldtians, Ladies and Gentlemen!

Welcome to Bamberg and to the 18th Symposium of the Alexander von Humboldt-Foundation. As part of its follow-up programme the Humboldt Foundation repeatedly organized symposia for selected Humboldt guest researchers from specific areas of specialization on given topics together with German scientists.

The Foundation was re-established by the Federal Republic of Germany in 1953 to support international scientific cooperation between German and foreign researchers by granting Humboldt Research Fellowships to highly qualified postdoctoral scientists and Humboldt Research Awards to internationally recognized scientists for their academic achievements. Since then the Foundation has built up an international network of more than 22000 scientists in some 130 countries maintaining academic, cultural, political and personal contacts with Germany.

The 18th Symposium is dedicated to Werner Heisenberg to commemorate his 100th birthday on December 5th this year. Werner Heisenberg was president of the Humboldt Foundation from 1953 to 1975. From own experience in his early career – Heisenberg had been postdoctoral fellow in Denmark – he had learned and appreciated the advantage and merits of doing research abroad in young years and he highly valued this experience. It was partly for that reason that he cherished the presidency of the Humboldt Foundation so much and for such a long time that it was the last office from which he resigned – only a few months before his death on February 11th, 1976.

Werner Heisenberg deserves well from sciences in Germany. He contributed an enormous part to the re-establishment of the Alexander von Humboldt-Foundation and the Max Planck Society: It was his scientific reputation which helped open the doors for many German researchers to the international science community.

Physics ranks quite high in the statistics of the Humboldt Foundation. In the years from 1972 to 2000 the Award Programme counts 1082 nominations which resulted in 724 Award Winners. Thus Physics is first followed by Chemistry with 431 Award winners and Biosciences third with 316 winners.In the Fellowship Programme Chemistry is first but followed by Physics in second place.

17 out of the 32 Humboldtians who were honoured with the Nobel Prize are physicists. And a special welcome now goes to the three Nobel Prize Winners who are among us today. With these facts it is evident and makes sense that we decided to dedicate the 18th Symposium to Werner Heisenberg.

The love of his country was never at stake – even under the most extreme situation and strained circumstances. During the Third Reich he went on with his research, led his family life and participated in musical events with friends. He never was a member of the Nazi Party. He never openly broke with the regime either. He continued to hold high scientific positions throughout the Third Reich. He was attacked in SS journals for his contribution to "Jewish Science". Courage and suffering can never be fully understood and appreciated by those who never had to live through dictatorship.

Aspects of his life will be under dispute this afternoon. We are very happy to have David Cassidy among the experts to speak to you. David Cassidy has devoted much of his academic career to Heisenberg's work and life during a turbulent century. His full scale biography and the recently published expanded edition of the bibliography of Werner Heisenberg's writings are essential to fully appreciate the first and longtime president of our Foundation.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Though the atmosphere we live in today no longer is that of a totalitarian regime we nevertheless have to be and are concerned with ethical issues and moral dilemmas in modern research. The responsibility of researchers and moral barriers in research – in different academic field, in genetics – are under discussion and being debated on both side of the Atlantic ocean. Who draws the lines? Who defines ethics in research? Who bans research activities and the production of atomic, biological and chemical weapons? Who is or will be in control, national control and global control? The number of dilemmas has not decreased.

Werner Heisenberg once told Mario Bunge, a Humboldt research fellow, that he was not a positivist and that in setting or constructing a theory, he had never been stimulated by an experiment. "My motivation and heuristic clues have always been of a philosophical nature". On that occasion he also complained about the new generation of theoretical physicists who wanted to go back to Ptolemy: "They only wish to describe and predict facts. On the other hand I am of a Newtonian cast of mind. I want to understand facts. Therefore, I appreciate the theories that explain the working of things, more than any phenomenological theories.".

Heisenberg died in 1976. 25 years later new generations of theoretical physicists have progressed in the field the Nobel Laureates Bohr, Born, Dirac, Heisenberg, Pauli and Schrödinger once paved the way for. Their work and development in Quantum Physics, Elementary Particles and Nuclear Physics has proven fundamental for the technological progress in the second half of the 20th century.

Internationalisation – for a long time a characteristic principle among physicists and of the Humboldt Foundation – has rightly become a key-concept in responding to the globalisation process when dealing with problems, whether of scientific or humanitarian nature – and perhaps also when peace is at stake.

Some 100 scientists from 27 countries – members of the worldwide Humboldt Network – have come to Bamberg to find out more about and discuss different aspects of Werner Heisenberg's 75 years' life span in the 20th century and the impact of his works in the 21th century.

My thanks go to the Chair-persons in the Plenary Sessions and Workshops. And I wish to express my thanks and appreciation to all participants. I am particularly grateful to the following scientists who gave their advice and helped to organize the symposium:

Wilfried Buchmüller
Dieter Lüst
Wolfgang Schleich
and from the chancery of the Foundation:
Dietrich Papenfuss and Ulrika Holdefleiss-Walter

I wish you an enjoyable stay in Bamberg, fruitful discussions in and outside the Plenary Sessions and the Workshops.

Wolfgang Frühwald
President of the Alexander von Humboldt-Foundation, Bonn

Fortschr. Phys. **50** (2002) 5–7, 432–436

Plenary Talks

Heisenberg and the Framework of Science Policy

CATHRYN CARSON

Office for History of Science and Technology
543 Stephens Hall, University of California
Berkeley, CA 94270-2350, USA

Abstract

In the decades after 1945, new structures were created for science policy in the Federal Republic. To the establishment of the postwar framework Heisenberg contributed as much as any other figure. This was true even though, on the whole, he took no great pleasure in the venture, nor was he always particularly adept at it. His conceptions revolved around certain key notions: autonomy and centralization, elite advisory bodies and relationships of trust, modernization and international standards. These show up at many levels of his activity, from the Max Planck Society to national and international advisory committees to the Humboldt Foundation itself. His opinions were shaped by encounters in the Federal Republic, but they also grew out of his experience of the Third Reich. At a moment like the present, when the postwar settlement is under review, it is interesting to reflect on the inherited system: on the extent to which it reflects the situation of the postwar decades and the intuitions of those who, like Heisenberg, created it.

Speaking on the history of science policy, which is often thought boring, is a dangerous undertaking in the late afternoon. Dangerous, too, is speaking about it before so many listeners with personal experience. Yet for a symposium for Heisenberg's centennial, the topic is important. In the decades after 1945, Heisenberg and his colleagues faced a one-time opportunity to shape a new system in creation. The science policy of the Third Reich, if such had existed, had been a disaster. At a moment when a new political order was being formed, science was the rising star on the international horizon. The science policy system was not created from scratch. Nonetheless, it confronted a new state and new options. Even though Heisenberg often thought of the task as a burden, he contributed to defining its framework as much as any other single person.

Thus there is historical interest in examining the postwar settlement under construction. The topic is also significant for us in the present, because that postwar settlement has since come under review. The world has changed since Heisenberg's times. Principles that seemed persuasive in his day have seemed less so to his successors. In the years since those criticisms were first articulated, we have seen the framework gradually being rethought. Thinking over those more recent experiences, we can gain by examining the structures he helped put into place.

This contribution will proceed by laying out Heisenberg's contribution to the framework of science policy [1, 2]. Moving quickly and schematically, it will focus on institutional and structural presuppositions and legacies, giving less emphasis to moments of decision. (History need not proceed chronologically.) The account will deal with Heisenberg's two main concerns, advising and institution-building, and offer reflections on what he accomplished.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

1. Advising

Heisenberg's first concern was science advising. What made it necessary? That seemed obvious. A science policy was to his mind the hallmark of a modern state. Scientists had a responsibility to contribute, as they were often expert where politicians were not. The obligation fell particularly on the famous names in the scientific community, who would be more readily listened to beyond it. Important to Heisenberg was a form of advising that was not simply lobbying for science or securing funding for research. That meant scientific input on broader concerns of society and politics. Important, too, was that initiative remain on the side of the scientists, serving the government but remaining autonomous.

Heisenberg's advisory efforts: an overview

Postwar optimism (late 1940s)
German Research Council
Becomes Senate of DFG
High ambitions basically fail

The nuclear era (1950s)
Federal ministry for atomic affairs
German Atomic Commission
And its subcommissions and panels
High ambitions disappointed, but practical work succeeds

Restructuring (1960s on)
Federal ministry for research
Hopes rise and fall as chancellors come and go
Advisory committees formed and dissolved
Advising becomes a desideratum, but not on Heisenberg's terms

Heisenberg looked to models of advising like those put into place by World War II in countries like Great Britain and the United States. The question was how these would translate to Germany. The governmental structure was different, the scientists were different, and the prehistory was different, too. From the start Heisenberg took federal leadership as the order of the day. This entangled him, of course, in the ongoing battles of the Bund versus the Länder.

How was advising to be carried out in practice? He imagined it emerging from small circles of first-rate individuals. Heisenberg was indeed a man of small circles and face-to-face meetings (also, one may note, quite skilled at leading them). He wanted to work with individuals who spoke their minds, independent of disciplinary or institutional affiliation. This seemed to him the model for how science policy should be made. It depended upon the building up of relationships of trust with the government, often based upon personal contacts between high-level scientists and high-level officials.

The outcome did not match Heisenberg's expectations. Science advising was established on a small scale, an example being the agenda-setting power around 1960 of the nuclear physics panel of the German Atomic Commission. On the larger scale, however, despite multiple tries, Heisenberg had to admit he never fully succeeded. This assessment held from his initial proposal of a German Research Council of the late 1940s through his arguments for advisory bodies to the chancellors of the 60s. His ambitions failed, first of all, to gain the support of his colleagues, then ran into contrary demands within the government. So while funding was secured (particularly for nuclear and particle physics), advising was

less well institutionalized. Exemplarily, the government set up its own atomic commission, only to bypass it when considering nuclear weapons for the federal armed forces. Still, the idea of high-level advising by scientific leaders has continued to exercise appeal.

2. Institution-building

The other domain in which Heisenberg invested his time was the construction and strengthening of the institutions of science. Institutions have some permanence; once they are in place, they are not easy to move, change, or dissolve. With an institutional landscape in flux after the war, individuals had opportunities to make long-lasting contributions. Heisenberg took full advantage. A few examples, selected from a wider spectrum, give a sense of his concerns.

His efforts began with laboratories and research facilities, with CERN as a prime example. (CERN had the added advantage of international connections, crucial for West Germany after the war.) Despite being a theorist, Heisenberg was the first head of CERN's Scientific Policy Committee. He proved adept at working out compromises when his own interests were not at stake. He also was good at collaborating with politicians to get the laboratory funded, even as it repeatedly exceeded expectations for its financial commitments. (That Heisenberg came to take the financing seriously was part of the reason why, by the late 1960s, he preferred storage ring plans to the SPS. Part of the reason was also his unified field theory, but that was not all of the story.)

A further locus of his efforts was the Max Planck Society, as it reconstituted itself out of the Kaiser Wilhelm Society. Heisenberg used his position as institute director to launch new institutions for big science: parts of the Karlsruhe Research Center, the Institute for Plasma Physics, and the Max Planck Institute for Extraterrestrial Physics. More generally, he grew by the late 1950s into a major power within the MPG. Along with others, he was in discussion to succeed Otto Hahn as president – which would have been for him a mistake, as he surely realized. Instead, in line with his personal inclinations, he continued to work behind the scenes, helping push through Adolf Butenandt's reforms of the early 1960s. Heisenberg's interest was in modernizing the MPG and ensuring its international standards. If German science was to hold onto such a remarkable institution – remarkable in quality, of course, also remarkable in structure and autonomy vis-à-vis the state – it would have to be defended and kept up to date by deliberate processes of self-renewal.

A final example of Heisenberg's activity is of course the Alexander von Humboldt Foundation. The Foundation was resurrected after the war, and starting in 1953 Heisenberg served it for more than two decades as its first postwar president. The pair of Heisenberg and Heinrich Pfeiffer, his long-time General Secretary, decisively shaped the Foundation's agenda and structures. Little needs to be said about its program, except to point out that the atmosphere with which Heisenberg helped endow it, an environment of personal relationships and face-to-face encounters, was of a piece with his views on science policy more generally.

The principal interest here is in the Humboldt Foundation's structure, which can be seen in many ways as a model. The Foundation was to operate in a free space outside the federal bureaucracy, a free space that the government had carved out voluntarily. The state's interest was in creating international scholarly connections, for which a government office seemed ill-suited in a world suspicious of Germany. Most of the Humboldt Foundation's funding came from the Foreign Office, but it was given to the scholars to dispose of by their own criteria. Thus the body was set up as a foundation, despite its limited endowment. This secured it leeway in defining its own organs, with fewer bureaucratic constraints. It also allowed it to accept government funding and public responsibilities while explicitly maintaining formal autonomy. In this respect, along with the Max Planck Society,

the Humboldt Foundation was the venture in which Heisenberg was most successful in realizing his ideals. It is no surprise that he took so much pleasure in it.

3. Reflections

Much of Heisenberg's thinking came from international models, so these developments belong to a global story. But at least as much of the impulse came from experiences of the Third Reich, from which Heisenberg learned in senses both positive and negative. Despite his troubles with the Nazis, Heisenberg also started on science policy in the Third Reich. This was when he began to learn how to work within a system, deal with politics, and make use of his prestige. Yet he distinctly reacted against the experience of subordination, which goes some way towards explaining his postwar stress on autonomy. Autonomy could be a direct response to the Third Reich, exhibited in a desire to maintain a balance vis–vis the state. It also resonated with liberal cultural values that were resurrected, too, after 1945. Those values stressed the autonomous personality given space to develop creativity and exercise talents. They were the same cultural values stressing personal contact that shaped the Humboldt Foundation, too.

In science policy, autonomy is a tricky notion. Scientific leaders and political theorists sometimes call it intrinsic to science. Sometimes, in a more concrete vein, they point to the clause anchoring freedom of research in the Federal Republic's Basic Law. Historians are skeptical about timeless arguments, and they want to understand how clauses like this one find their way into constitutions at all. In the fact of the matter, autonomy has been called into question for many observers since Heisenberg's day. With science dependent on public funding, government and society proved more inclined by the 1960s highlight accountability instead. That has meant that the choice of field and research method has been less and less left to the individual's wishes. One result has been a series of protests from the side of scientists objecting to interference with researchers' programs and the shrinking free space for research in general [3]. In these plaints, Heisenberg's era of visionary individuals appears as a golden age. But perhaps the principle of autonomy bears rethinking from within the community of researchers. Other notions from Heisenberg's day have already been reworked. An example is the idea that scientific advisory bodies are not also, simultaneously, the voice of an interest group; Heisenberg already took his leave of that notion. Similarly, the early insistence has faded that members of advisory boards are there purely as individuals, not as representatives of different fields or organizations. Institutions like the Humboldt Foundation have found ways to coordinate their independent initiative with the needs of their sponsors. Practices have adapted, even if basic notions have not.

The historian speaking here would not want to be misunderstood: I like autonomy; indeed, in my nostalgic moods, I wish I had more of it. But as a basis for argumentation it feels (to this historian) heavily imbued with the spirit of the late 1940s and 50s. In the late 1940s and 50s, in that historical setting, it may have been broadly compelling. It has not proved so compelling anymore for outsiders who do not get to share the individual's pleasure in independent scholarly work. The world has changed. This historian's experience, in academia and the world at large, has been that absolute autonomy is nowhere realized. When only the absolute notion is available, thoughtful reconsideration is blocked.

Positive suggestions are not easy to come by. Yet thinking about the postwar transformation of liberal individualism, the notion of autonomy could be refined, too. Very schematically, stressing a process instead of a state of affairs would allow a twofold rethinking. First, an emphasis on pluralism, a tolerance for idiosyncrasy, would still authorize individuals to follow their peculiar paths. This could be justified by a standard liberal presumption that a variety of starting points leads to a richer appreciation [4]. Second, the course and pace of research can be understood as developing by a logic set in part by the matter at

hand, constrained by the resistances of the domain under study. This shift offers alternatives to demands for short-term results or progress demonstrable by purely external criteria. So pluralism is allied to autonomy, but without its completely individualistic perspective. The logic of the matter at hand has similar allegiances, but locates the motor in the domain under study and outside the individual's wishes. These paired concepts give us room to rethink in a world that has changed. They do not offer absolute autonomy, but that strong notion may be a mirage.

All this is speculative, an attempt to put historical perspective to work. What it suggests, however, is that the historical constitution of the framework of science policy can open up questions for us today. Many of the structures of Heisenberg's era are with us still: some still honored, others now questioned. Trying to understand their origins, we gain some insight into their prospects.

References

[1] CATHRYN CARSON, *Heisenberg als Wissenschaftsorganisator*, forthcoming in Werner Heisenberg zum 100. Geburtstag, ed. Christian Kleint, Helmut Rechenberg, and Gerald Wiemers (Berlin: Akademie-Verlag, 2001).

[2] CATHRYN CARSON, *New models for science in politics: Heisenberg in West Germany*, Historical studies in the physical and biological sciences 30:1 (1999) 115–171.

[3] *Der schrumpfende Freiraum der Forschung:* Symposium der Max-Planck-Gesellschaft, Schloß Ringberg/Tegernsee, Mai 1994. Max-Planck-Gesellschaft, Berichte und Mitteilungen, H. 1/95.

[4] JOHN STUART MILL, *On liberty* (London: Penguin, 1974).

Fortschr. Phys. **50** (2002) 5–7, 437–439

Werner Heisenberg: An Overview of His Life and Work

DAVID C. CASSIDY

Hempstead, New York, USA

Professor Dr. Werner Karl Heisenberg was born on 5 December 1901 in Würzburg, Germany. He died at his home in Munich on 1 February 1976. Heisenberg lived during some of the most difficult and controversial years of modern Germany history. During his lifetime he experienced two lost world wars, three political upheavals, and four different political regimes – the Wilhelmian empire, the Weimar Republic, the Third Reich, and the postwar West German Federal Republic.

As a physicist, Heisenberg ranks with Niels Bohr, Paul Dirac, and Richard Feynman in his contributions to and impact upon contemporary physics. He was a key player in the development of quantum mechanics, the new physics of the atom in the 1920s and its presently accepted interpretation. He went on to formulate a quantum theory of ferromagnetism, the neutron-proton model of the nucleus, the S-matrix theory in particle scattering, and many other important advances in quantum field theory and high-energy particle physics. During his lifetime Heisenberg produced nearly 600 original research papers, philosophical essays, and explanations for general audiences [1]. All of these are now conveniently available in the nine volumes of the Gesammelte Werke/Collected Works [2].

Among his many scientific contributions Heisenberg is perhaps best known for the so-called uncertainty, or indeterminacy, principle of 1927, for the first breakthrough to quantum mechanics in 1925, and for his proposal of a unified field theory, the so-called "world formula". During the late 1950s. He received the Nobel Prize for Physics in 1932 at the age of 31 "for the creation of quantum mechanics, the application of which has, inter alia, led to the discovery of the allotropic forms of hydrogen".

Heisenberg remained in Germany during the nightmare years of the Hitler regime. He headed Germany's research effort on the applications of nuclear fission during World War II. After the war he played an important role in the reconstruction of West German science, in the shaping of the public culture of science in Germany, and in establishing the success of West Germany's nuclear and high-energy physics research programs.

Heisenberg was born into an academic family that was moving up the social ladder through academic achievement. Heisenberg's father, Dr. August Heisenberg, professor at Munich University, was a leading authority on Byzantine philology. Heisenberg's maternal grandfather, Dr. Nikolaus Wecklein, become rector of the famed Maximilians-Gymnasium (middle school) in Munich and a member of the Bavarian School Board. He was a leading authority on ancient Greek tragedy. Both of these men were determined that the two Heisenberg boys–Werner and his elder brother Erwin–would reach the next social stratum through academic achievement. (Erwin became a chemist and worked in industry.) After completing his studies at his grandfather's Gymnasium, Heisenberg entered the University of Munich in 1920. He received the doctorate in physics with the theoretician Arnold Sommerfeld in the record time of three years. Heisenberg had the good fortune of studying in the leading centers and with the leading researchers in the field of quantum atomic physics: Sommerfeld in Munich, Max Born in Gttingen and Niels Bohr in Copenhagen. He was also fortunate to have such outstanding fellow students and colleagues as Paul Dirac, Pasual Jordan, Hendrik Kramers, Friedrich Hund, and above all Wolfgang Pauli.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Heisenberg's breakthrough to a new quantum mechanics occurred in June 1925 while he was on the barren island of Helgoland recovering from an attack of hay fever. Cured of the malady, he returned to Göttingen with a strange-looking quantum multiplication rule that Max Born soon recognized as the rule for the multiplication of matrices. Utilizing the then little-known matrix algebra, Born, Heisenberg and Jordan quickly developed the version of quantum mechanics known as "matrix mechanics". Their work proceeded independently of Erwin Schrödinger's alternative "wave mechanics". In 1926 Schrödinger and Pauli showed that the two versions of quantum mechanics were in fact mathematically equivalent. Dirac and Jordan soon developed a unified quantum formalism then known as transformation theory. Modern, non-relativistic Quantum Mechanics was born, but the precise meanings of the symbols in applications to actual situations still required elaboration. While teaching and working with Bohr in Copenhagen, Heisenberg derived his famous uncertainty, or indeterminacy, relations from the new quantum formalism of Dirac and Jordan. The uncertainty relations formed the basis of the Unbestimmtheitsprinzip or, as commonly known today in English, the uncertainty principle. This principle, together with Bohr's complementarity principle and Born's probability interpretation of Schrödinger's wave function, constituted the famous Copenhagen Interpretation of quantum mechanics. This interpretation is still the main interpretive foundation for physical applications of quantum mechanics and for many of the philosophical implications of the new physics [3].

In 1927 Heisenberg was appointed Professor of Theoretical Physics at the University of Leipzig. He was celebrated at the time as Germany's youngest full professor. With Peter Debye as head of the experimental physics institute and Friedrich Hund as the second professor of theoretical physics, Leipzig University became one of the leading centers of education and research in quantum mechanics. Many of the brightest young physicists of the next generation visited, studied, or assisted in Heisenberg's institute. Among them were such luminaries as Rudolf Peierls, George Placzek, Viktor Weisskopf, Felix Bloch, Edward Teller, I. I. Rabi, C. F. von Weizsäcker, Ugo Fano, Richard Iskraut, Edoardo Amaldi, Eugene Feenberg, Shin-ichiro Tomonoga, Lew Landau, Gian Carlo Wick, Wang Foh-san, Laszlo Tisza, and Guido Beck [4]. Together with Heisenberg, they helped to develop quantum theories of nuclear structure, electrical conductivity, the Hall effect in metals, and the behavior of interactions of matter and radiation at extremely high energies. Many of these young people went on to work at Pauli's institute in Zürich and Bohr's institute in Copenhagen, with whom Heisenberg maintained a rich and lively correspondence on the latest problems in quantum physics.

In 1928 Heisenberg and Pauli published the framework for the first relativistic quantum field theory a quantum theory that encompassed both matter and non-material fields. As Heisenberg and his co-workers struggled with field theories and the applications of quantum mechanics to new areas during the 1930s, he found himself in the midst of another type of struggle, a struggle for the very survival of German physics.

After coming to power in 1933, the Hitler regime unleashed an assault on the academic profession, including physics, and a direct assault on Jewish teachers and professors. Heisenberg – the famous physicist, energetic young Nobel Prize winner, and German patriot – found himself at the center of the German physics community's response to the regime. Much has been written about his actions, motives, and strategies during those nightmare years [5]. In hindsight we can see that it was a losing battle almost from the very start.

Soon after the annexation of Austria late in 1938, Otto Hahn and Fritz Strassmann in Berlin, together with Lise Meitner in Sweden, discovered nuclear fission, the splitting of uranium nuclei with the release of enormous amounts of energy. This discovery, coming on the eve of the outbreak of war in Europe, lent a new and potentially dangerous dimension to nuclear research on both sides of the coming world war.

Heisenberg accepted scientific leadership of Germany's main nuclear-fission research effort in Berlin. In 1942 he was also appointed Professor of Physics at Berlin University, and

he was named director at the Kaiser Wilhelm Institute for Physics, the precursor of the present-day Max Planck Institute for Physics. Heisenberg's decision to remain in Germany, his reasons for working on Germany's nuclear-fission project, the project's lack of progress compared with the Allied effort, and Heisenberg's travels to occupied countries–including his visit to Bohr in Copenhagen in 1941 have all been the subjects of intensive scrutiny and debate, and even dramatic reconstruction, ever since. The issues are complex and profound and the debates are often highly emotional, even today. The positions taken cover the entire spectrum of possibilities [6].

Following the end of the war, Heisenberg and his closest colleagues remained in West Germany where they urged nuclear-reactor research and pressed for a more direct role for leading scientists as personal advisors to the new Federal Chancellor. Their efforts contributed to the modern-day coordination between the federal government and the Länder in matters of science policy [7].

As the long-time President of the Alexander von Humboldt Foundation, member of the founding committee of CERN, the European accelerator facility near Geneva, and energetic participant in other international scientific organizations, Heisenberg helped to re-establish international relations among German and foreign scientists and scholars.

Yet, while pursuing these important goals, Heisenberg never wandered far from his first concern – fundamental quantum physics. He continued his long-time correspondence with other physicists and especially with Wolfgang Pauli until Pauli's untimely death in 1957. In collaboration with Pauli he developed a new solution to the problem of infinities in high-energy interactions in a non-linear unified theory, his world formula. As originally formulated, the theory did not succeed and Pauli abruptly withdrew from the endeavor. But the work progressed in the following years, and it proceeds today with continuing prospects of success. Heisenberg's life and work and his impact upon his times during the century since his birth will no doubt keep scientists, historians, and many other scholars occupied for centuries to come.

References

[1] WERNER HEISENBERG, *A Bibliography of His Writings*, second edition, compiled by D.C. Cassidy (New York, 2001) http://www.aip.org/history/heisenberg/bibliography/contents.htm.

[2] WERNER HEISENBERG, *Gesammelte Werke*/Collected Works, 9 vols., W. Blum, H. P. Dürr, and H. Rechenberg, eds. (Berlin and Munich, 1985–1989).

[3] JAGISH MEHRA and HELMUT RECHENBERG, *The Historical Development of Quantum Theory*, 6 vols. (New York, 1982–2001).

[4] Christian Kleint and Gerald Wiemers, eds, Werner Heisenberg in Leipzig 1927–1942 (Leipzig, 1993), pp. 136-172.

[5] For instance, ALAN D. BEYERCHEN, *Scientists under Hitler: Politics and the Physics Community in the Third Reich (New Haven, 1977);* MARK WALKER, *Nazi Science: Myth, Truth, and the German Atomic Bomb (New York, 1995);* KARL HEINZ LUDWIG, *Technik und Ingenieure im Dritten Reich (Knigstein/Ts., 1979); and* D. CASSIDY, *Uncertainty: The Life and Science of Werner Heisenberg (New York, 1992).*

[6] Recent literature includes, in alphabetical order: CASSIDY, *Uncertainty;* THOMAS POWERS, *Heisenberg's War: The Secret History of the German Bomb* (London, 1993); PAUL LAWRENCE ROSE, *Heisenberg and the Nazi Atomic Bomb Project*, 1939–1945: *A Study in German Culture* (Berkeley, 1998); MARK WALKER, *German National Socialism and the Quest for Nuclear Power* 1939–1949 (New York, 1989).

[7] CATHRYN CARSON, *New Models for Science in Politics: Heisenberg in West Germany*, Historical Studies in the Physical Sciences, 30 (1999), 115–171.

Fortschr. Phys. **50** (2002) 5–7, 440–458

Observability, Anschaulichkeit and Abstraction: A Journey into Werner Heisenberg's Science and Philosophy

JAN LACKI

History and Philosophy of Science, Geneva University
24, quai Ernest Ansermet
CH-1211 Geneve 4, Switzerland
jan.lacki@physics.unige.ch

1. Introduction

Werner Heisenberg was one of the greatest physicists of the 20th century. He participated as a front rank actor in the shaping of a good part of XXth century physics and directly witnessed most of the intellectual struggles which led to what he called "Wandlungen in den Grundlagen der exakten Naturwissenschaft". This expression is borrowed from one of the many talks and writings he devoted to the analysis of the scientific and philosophical implications of his, and his fellows physicists, findings. Indeed, Heisenberg's scientific activity increasingly reflected his more general intellectual views. This makes him another magnificent representative of a glorious linage going from the remote times of modern science to Einstein, Bohr and the like. This "philosophical" vein started early in his scientific life, and got stronger with time, prompted by the highly demanding scientific, but also social and political context of his mature years.

My contribution is an attempt to review some of the most significative themes underlying Heisenberg's creativity illustrating, with some excerpts from his papers, the fruitful interplay of his technical mastery and physical insight, and his later philosophy. I have chosen three words which will serve as milestones in this journey into Heisenberg's science[1].

The first one, *observability*, can probably be related the most easily to Heisenberg's feats, since the requirement to concentrate on observable quantities served as the main "epistemological" motivation behind Heisenberg's celebrated matrix mechanics breakthrough.

Anschaulichkeit, a German word endowed with so many meanings that is defies any proper translation, makes one think of the painful groping of quantum scientists searching for a new intuition and meaning of the quantum formalism. I shall in particular refer here to the context of Heisenberg's celebrated paper on the indeterminacy relations.

The last milestone, *abstraction*, appears as the most arbitrary one since Heisenberg does not seem here to stand out more than any other of the founding fathers of quantum theory. Still, not only because Heisenberg enjoyed the mathematical methods of quantum mechanics, and used them with great mastery, but also because he deeply thought about the relations between abstract and intuitive appraisals of reality, this theme also fits my purpose.

There is actually a deeper connection between the themes chosen above. It relates to the very progression of Heisenberg's own philosophy and *Weltanschauung*. From the observa-

[1] For detailed and exhaustive accounts, see Hund 1967, Jammer 1966, Rechenberg and Mehra 1982–2000, Cassidy 1992.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

bility requirement of his 1925 paper, to his most accomplished philosophical statement of his 1942 manuscript, Heisenberg, under the clear influence of his teacher and friend Bohr, underwent a maturation process which led him to present a genuine philosophical work of great scope. In particular, Heisenberg's only recently published 1942 manuscript *Ordnung der Wirklichkeit*[2)] exposes an original position which goes far beyond the mere epistemological interests of a "philosophically" aware quantum scientist. It is a work of a genuine humanist which dares putting his scientific activity into a broader perspective, among other, equally valid, apprehensions of reality. This is where the tension, always observable in Heisenberg's research, between intuition, and abstraction, finds its philosophical resolution. It emphasizes the relevance of Heisenberg not only as one of the most impressive scientists of his time, but also as one of its most acute philosophical commentators.

2. Observability

Most of us know Heisenberg's celebrated paper which triggered quantum mechanics (Heisenberg 1925). It is a masterpiece of physical insight, formalistic boldness, and, considering the issue at stake, of impressive (but actually fake) candor. Remembering Heisenberg's declared guiding principle, when formulating the rules of the new calculatory quantum scheme, to concentrate on the observable quantities only, one can only marvel at the efficiency of this deceptively innocent positivistic precept. It is as if quantum mechanics was laying there from the very beginning, ready to be harvested.

Of course this is just an illusion. It took the whole power of analytical mechanics imported from celestial mechanics, and, on the other hand, of Bohr's physical insight, to bridge the yawning gap between classical physics of the 19th century, and the new quantum phenomenology of the early 20th century. Thanks to the theory of multiply-periodic systems, and Bohr's correspondence principle, simple atomic spectra and related phenomena found an acceptable description. From the beginning, Heisenberg was at the right school. He mastered classical techniques with the best teachers one could dream of, first Sommerfeld, and then Born. Finally in Copenhagen, when visiting Bohr during his winter 1924–25 stay, Heisenberg learned "physics". The stage was set.

But first, the use of the classical theory of multiply periodic systems supplemented with *ad hoc* quantization rules had to be seriously questioned. Apart from its congenial incoherence, the "old quantum theory" lost momentum because of its many failures among others the Helium atom, the complex spectra, and its inability to account for the interaction of atoms with radiation. Born and the Göttingen school started then the program of discretization of classical formulas, bypassing the classical mechanical models, as a step towards genuine "quantum mechanics". This approach worked nicely when applied to the problem of dispersion in the pioneering work of Kramers in Copenhagen (Kramers . . .). This was just the time when Heisenberg visited Bohr to learn from him.

This is the right place to emphasize the importance of Bohr in Heisenberg's intellectual development, surely as a physicist but, as we will shortly see, also as a philosopher. Under Sommerfeld's and then Born's guidance, Heisenberg grew as a *Wunder Kind*, a mathematical opportunist, playing with the formalism of classical mechanics and quantization rules for the sake to always come closer to experimental data. The overall physical picture, its soundness and ultimate meaning remained in the background. Pauli, who preceded Heisenberg in Copenhagen, often complained about the mathematical excesses of the Göttingen school. He rightly hoped that Bohr would know how to set Heisenberg's exceptional gifts

[2)] This manuscript has been published for the first time in Heisenberg's *Collected Works* after his death, see the explanatory note in Blum, Dürr and Rechenberg 1984, p. 217. It still deserves to be better known, and in what follows I shall quote extensive excerpts from it.

on a better track, the track of genuine physics. Least to say, Pauli's "plan" worked perfectly: Heisenberg, staying in Bohr's orbit, indeed became more "philosophical".

From his very first short stay in Copenhagen in spring 1924, Heisenberg got under Bohr's spell. His longer stay, later in the year, made him an ardent defensor of Bohr's physics and style. In Copenhagen, he also made some precious contacts, especially Kramers which made him get involved into dispersion theory.

It is interesting to compare the style of Heisenberg's papers before and after his Copenhagen "conversion". Prior to Bohr's influence, his papers are overly "technical", in the sense that the opening, but also closing statements appear as conventional, contrived pieces of scientific literature, where one mentions predecessors, pays dues, and finally states solved and still open problems. The overall impression is that of someone getting impatient to proceed with the real business: clever mathematical tricks, computation and check against experimental data. Starting with his paper on the correspondence principle and fluorescence, his first after meeting Bohr, his style seems to change. One finds there a long development on the possibility of sharpening the correspondence principle, and actually the whole paper hardly contains any formal part. In a word, it looks much "bohrian" in style.

The matrix mechanics 1925 paper could at first sight appear of the same trend. However, one should beware of hasty conclusions. Indeed, its proper assessment illustrates the necessity of an important distinction in the history of science: the distinction between the context of discovery and the context of justification. Heisenberg, from his first intent to compute the intensities of the hydrogen spectrum, till the famous Helgoland episode, followed a line of thought and was driven by motivations part of which only made it explicitly into the paper. On the other hand, some statements of the paper where most presumably not the real driving force behind his breakthrough. The observability principle might be just one of these. Be it as it may, Heisenberg chose to open his paper with the following words (Heisenberg 1925, p. 879):

Bekanntlich lässt sich gegen die formalen Regeln, die allgemein in der Quantentheorie zur Berechnung beobachtbarer Größen (z. B. der Energie im Wasserstoffatom) benutzt werden, der schwerwiegende Einwand erheben, dass jene Rechenregeln als wesentlichen Bestandteil Beziehungen enthalten zwischen Größen, die scheinbar prinzipiell nicht beobachtet werden können (wie z. B. Ort, Umlaufszeit des Elektrons), dass also jenen Regeln offenbar jedes anschauliche physikalische Fundament mangelt, wenn man nicht immer noch an der Hoffnung festhalten will, dass jene bis jetzt unbeobachtbaren Größen später vielleicht experimentell zugänglich gemacht werden könnten. Diese Hoffnung könnte als berechtigt angesehen werden, wenn die genannten Regeln in sich konsequent und auf einen bestimmt umgrenzten Bereich quantentheoretischer Probleme anwendbar wären. Die Erfahrung zeigt aber, dass sich nur das Wasserstoffatom und der Starkeffekt dieses Atoms jenen formalen Regeln der Quantentheorie fügen, dass aber schon beim Problem der „gekreuzten Felder" (Wasserstoffatom in elektrischem und magnetischem Feld verschiedener Richtung) fundamentale Schwierigkeiten auftreten, dass die Reaktion der Atome auf periodisch wechselnde Felder sicherlich nicht durch die genannten Regeln beschrieben werden kann, und dass schließlich eine Ausdehnung der Quantenregeln auf die Behandlung der Atome mit mehreren Elektronen sich als unmöglich erwiesen hat. Es ist üblich geworden, dieses Versagen der quantentheoretischen Regeln, die ja wesentlich durch die Anwendung der klassischen Mechanik charakterisiert waren, als Abweichung von der klassischen Mechanik zu bezeichnen. Diese Bezeichnung kann aber wohl kaum als sinngemäß angesehen werden, wenn man bedenkt, dass schon die (ja ganz allgemein gültige) Einstein-Bohrsche Frequenzbedingung eine so völlige Absage an die klassische Mechanik oder besser, vom Standpunkt der Wellentheorie aus, an die dieser Mechanik zugrunde liegende Kinematik darstellt, dass auch bei den einfachsten quantentheoretischen Problemen an eine Gültigkeit der klassischen Mechanik schlechterdings nicht gedacht werden kann. Bei dieser Sachlage scheint es geratener,

jene Hoffnung auf eine Beobachtung der bisher unbeobachtbaren Größen (wie Lage, Umlaufszeit des Elektrons) ganz aufzugeben, gleichzeitig also einzuräumen, dass die teilweise Übereinstimmung der genannten Quantenregeln mit der Erfahrung mehr oder weniger zufällig sei, und zu versuchen, eine der klassischen Mechanik analoge quantentheoretische Mechanik auszubilden, in welcher nur Beziehungen zwischen beobachtbaren Größen vorkommen.

We immediately see the effect of Bohr's influence. Since the very beginning, Heisenberg was impressed by the degree of Bohr's worried awareness of the breakdown of mechanics, and the latter's unwillingness to seek escape into fancy subterfuges. Bohr faced the problem in its full extension, and so did Heisenberg. However, the principle of observability was not Heisenberg's own discovery. We know that Pauli and some of the Göttingen people, especially Born played with similar ideas. Heisenberg appears to have introduced this requirement *a posteriori*, as a way of organizing and providing a methodological unity to his paper, only after having derived its crucial steps. Back in Göttingen, in contact with Born and Jordan, he might then have realized the whole relevance of this principle for his own endeavour. So, Born's role in Heisenberg's science actually goes beyond that of a mere mathematical pygmalion.

I shall present now Heisenberg's argumentation leading to his "matrix" multiplication rule. My reading is idiosyncratic and puts the whole problem into a perspective going possibly slightly beyond the strictly original motivations. However, I am convinced that this way of presentation displays best the originality of Heisenberg's approach. It will also provide a connection with our next milestone, *Anschaulichkeit*.

With historical distance, one can view Heisenberg's problem as basically one of providing a systematic way of associating (infinite) sequences of numbers (amplitudes) to a physical quantity in the quantum theory. Although Heisenberg himself did not recognize the proper mathematical meaning of his final achievement, namely matrix manipulation, he hit upon the right idea to try to organize the disparate amplitudes using as a guideline the consistency of the algebraic operations on the mathematical representatives of the associated quantities. To grasp this point, one has first to recall the working of the old quantum theory.

The paradigm of the old quantum theory was the case of a multiply periodic system which can be solved using uniformizing variables. Classically, such a system with s degrees of freedom, assumed non-degenerate, allows any physical quantity U to be expanded, using the uniformizing action-angle variables J_i, w_i, $i = 1, ..., s$, as a Fourier series:

$$U(t) = \sum_{\tau_1, \tau_2, ..., \tau_s} U_{\tau_1, \tau_2, ..., \tau_s}(J_1, J_2, \ldots, J_s)\, e^{2\pi i(\tau_1 \nu_{1+} \tau_2 \nu + ... + \tau_s \nu_s)\, t}\,, \tag{1}$$

where the τ' s run over all integers (both positive and negative) and the $U_{\tau_1, \tau_2, ..., \tau_s}$ are complex vectors depending only on the action variables J[3].

The prescription of the old quantum theory for the corresponding quantum problem was first to impose upon the action variables the following Bohr-Sommerfeld conditions[4]:

$$J_k = n_k h\,, \tag{2}$$

[3] Remember that in the uniformizing canonical action-angle variables, one has, for each action variable J_k, the equation for the corresponding angle $\frac{dw_k}{dt} = \frac{\partial H}{\partial J_k} = \nu_k$, where H is the Hamiltonian function of the system. This formula exhibits a generalized frequency ν_k corresponding to the uniform time evolution of the angle $w_k = \nu_k t$. See Jammer 1966, pp. 89–115.

[4] They have been proposed by Sommerfeld in the form

$$\oint p_k\, dq_k = n_k h$$

for each pair of coordinates, generalizing Bohr's condition of the quantization of the angular momentum of the orbit of the electron around the nucleus (Sommerfeld 1916).

where n_k are positive integers and h is the Planck's constant. Each choice of the n_k's defines a stationary state of the quantum system. The latter were understood as the only allowed classical motions. The other theoretical tool of the old theory was Bohr's correspondence principle. It enabled broadly speaking to "read off" the features of the eminently non-classical processes of quantum jumps between stationary states from the classical formulas (1) above. Indeed, let us remember that for a transition from a stationary state with energy $E = E(n_1, \ldots, n_s)$ to another one with (lower) energy $E' = E(n'_1, \ldots, n'_s)$, Bohr's frequency condition,

$$h\nu_q = \Delta E = E - E' \tag{3}$$

yields the frequency ν_q of the emitted radiation. Clearly, this frequency is not equal to any of the (mechanical) frequencies of the motion, which, classically, rule the frequency of the emitted radiation according to the laws of electrodynamics. However, if one observes that the energy difference ΔE can be expressed, in the limit of vanishing ΔJ_k, as

$$\Delta E = \sum_k \frac{\partial E}{\partial J_k}\, \Delta J_k = \sum \nu_k (n_k - n'_k)\, h\,,$$

then one recognizes that the quantum frequency corresponds in the expansion (1) above to the classical one:

$$\nu_q = \sum_k \nu_k (n_k - n'_k) \equiv \sum \nu_k \tau_k\,.$$

This correspondence, valid in the limit of high quantum numbers n_k, where the energy levels are assumed to coalesce (i.e. $(n_k - n'_k) \to 0$), was extended to hold approximately for all the other quantum numbers. Averaging over quantum numbers of the states involved in the transition was shown to provide some results also in those cases.

This extension of the correspondence principle suggested one to rewrite the Fourier series in the following form (for simplicity, I set here the number of degrees of freedom to one):

$$U(t) = \sum_\tau U_{n,n-\tau} e^{2\pi i(\tau\nu)\,t} \tag{4}$$

and even "replace", right form the start, the classical series by their "quantum" analogues, substituting for the classical frequency exponents those given by Bohr's frequency condition:

$$\sum_\tau U_{n,n-\tau} e^{2\pi i(\tau\nu)\,t} \rightsquigarrow \sum_\tau U_{n,n-\tau} e^{2\pi i\nu(n,n-\tau)\,t}\,. \tag{5}$$

The labelling emphasizes now the structure of the transitions from the stationary state defined by the n_k's (which fix the action variables $J_k = n_k h$, and the energy of the initial state), to the other ones, "τ"-distant from the latter.

The formulas above enabled thus one to read off the features of the quantum processes from the classical expression (1) or its quantum analogue (5). However, as far as a sound relation between physical concepts and their mathematical representatives is concerned, the formal expression (4), with its reinterpreted physical meaning, is quite incoherent. Indeed, let us remember that the series (1) has a physical meaning when applying to a classical multiply-periodic system. In this case the sum of the series *does indeed* correspond to the value of the physical quantity U. The reinterpreted series (4), not to say the quantum analo-

gue (5), although containing (harmonic) terms pertaining to the various transition processes making the quantum significance of U, *does not yield* a proper mathematical representation of U. What is actually essential is the set of amplitudes

$$U_{n,n-\tau} e^{2\pi i \nu(n,n-\tau)\, t} \tag{6}$$

as associated to the physical quantity U, and not the series itself; the latter becomes an obsolete reminiscence of the classical formal origin of the expressions. One can thus understand the challenge faced by Heisenberg in the following terms: to find a genuinely quantum equivalent which would gather the set of amplitudes (6) into a mathematically proper entity standing for a (quantum) mathematical representative of U (this representative would provide a way of handling the amplitudes "in one single piece" as was the case for the classical Fourier series (1)).

Heisenberg started noticing that in the problem of the emission of radiation by an electron submitted to an acceleration, the classical theory yields for the electric field E, emitted by the charge, the leading term of the form

$$E = \frac{e}{r^3 c^2}\; r \wedge r \wedge \dot{v}$$

(r is the position of the electron, v its velocity, e its charge, and $\wedge$ the wedge product of vectors). There are higher order contributions, all involving increasing powers of the velocity and its derivative, for instance $\frac{e\dot{v}v}{rc^3}$ and further $\frac{e\dot{v}v^2}{rc^4}$. Leaving aside the physical meaning of these expressions, one notices on the spot that they imply products of quantities, namely the velocity with its derivatives. This is why Heisenberg wrote (*ibid.* p. 880–881):

Man kann sich fragen, wie jene höheren Glieder in der Quantentheorie aussehen müßten. Da in der klassischen Theorie die höheren Näherungen einfach berechnet werden können, wenn die Bewegung des Elektrons bzw. ihre Fourierdarstellung gegeben ist, so wird man in der Quantentheorie Ähnliches erwarten. Diese Frage hat nichts mit Elektrodynamik zu tun, sondern sie ist, dies scheint uns besonders wichtig, rein *kinematischer* Natur; wir können sie in einfachster Form folgendermaßen stellen: Gegeben sei eine an Stelle der klassischen Größe $x(t)$ tretende quantentheoretische Größe; welche quantentheoretische Größe tritt dann an Stelle von $x(t)^2$?

In modern terms using somewhat loosely the language of commutative diagrams, one would phrase Heisenberg's question as finding the right rule or operation, say R such that the following diagram commutes (i.e. there are two paths from x to $\mathbf{x}^2$ yielding the same result, where x is any physical quantity):

$$\begin{array}{ccc} x & \longrightarrow & x^2 \\ \Downarrow & & \Downarrow \\ \mathbf{x} & \overset{R}{\Longrightarrow} & \mathbf{x}^2 \end{array} .$$

Here the double arrows denote the correspondence which associates the quantum mathematical expression to the classical one. The horizontal double arrow labelled R is the key issue. Its mathematical rule has to be coherent with the "commutativity" of the diagram.

On the road towards a solution, Heisenberg's first observation concerned the peculiar form of the additivity rule for quantum frequencies, which is a direct consequence of Bohr's condition (3) (*ibid.* p. 881):

Bevor wir diese Frage beantworten können, müssen wir uns daran erinnern, dass es in der Quantentheorie nicht möglich war, dem Elektron einen Punkt im Raum als Funktion

der Zeit mittels beobachtbarer Größen zuzuordnen. Wohl aber kann dem Elektron auch in der Quantentheorie eine Ausstrahlung zugeordnet werden; diese Strahlung wird beschrieben erstens durch die Frequenzen, die als Funktionen zweier Variablen auftreten, quantentheoretisch in der Gestalt:

$$\nu(n,\, n-\alpha) = h^{-1}\{E(n) - E(n-\alpha)\}\,,$$

in der klassischen Theorie in der Form:

$$\nu(n,\, \alpha) = \alpha\nu(n) = \left[\alpha\, \frac{dE}{dJ}\right] = \alpha\, \frac{1}{h}\, \frac{dE}{dn}\,.$$

(Hierin ist $nh = J$, einer der kanonischen Konstaten, gesetzt).

Als charakteristisch für den Vergleich der klassischen mit der Quantentheorie hinsichtlich der Frequenzen kann man die Kombinationsrelationen anschreiben:

Klassisch

$$\nu(n,\, \alpha) + \nu(n,\, \beta) = \nu(n,\, \alpha+\beta)$$

Quantentheoretisch

$$\nu(n,\, n-\alpha) + \nu(n-\alpha,\, n-\alpha-\beta) = \nu(n,\, n-\alpha-\beta)$$

$$\text{bzw. } \nu(n-\beta,\, n-\alpha-\beta) + \nu(n,\, n-\beta) = \nu(n,\, n-\alpha-\beta)\,.$$

Now (*ibid.* p. 881-883)[5)]:

Neben den Frequenzen sind zweitens zur Beschreibung der Strahlung notwendig die Amplituden; die Amplituden können als komplexe Vektoren (mit je sechs unabhängigen Bestimmungsstücken) aufgefaßt werden und bestimmen Polarisation und Phase. Auch sie sind Funktionen der zwei Variablen n und α, sodass der betreffende Teil der Strahlung durch den folgenden Ausdruck dargestellt wird:

Quantentheoretisch (7)

$$\text{Re}\,\{U(n,\, n-\alpha)\, e^{i\omega(n,\, n-\alpha)\, t}\} \qquad \text{(I)}$$

Klassisch (8)

$$\text{Re}\,\{U_\alpha\, e^{i\omega(n)\, \alpha t}\}\,. \qquad \text{(II)}$$

Der (in U enthaltenen) Phase scheint zunächst eine physikalische Bedeutung in der Quantentheorie nicht zuzukommen, da die Frequenzen der Quantentheorie mit ihren Oberschwingungen im allgemeinen nicht kommensurabel sind. Wir werden aber sofort sehen, dass die Phase auch in der Quantentheorie eine bestimmte, der in der klassischen Theorie analoge Bedeutung hat. Betrachten wir jetzt eine bestimmte Größe $x(t)$ in der klassischen Theorie, so kann man sie repräsentiert denken durch eine Gesamtheit von Größen der Form

$$U_\alpha\, e^{i\omega(n)\, \alpha t}\,,$$

[5)] The notations used by Heisenberg differ from mine: his $U(n,\, n-\alpha)$ corresponds to $U_{n,\, n-\alpha}$ and $\omega(n,\, n-\alpha)$ is $2\pi(\alpha_1\nu_1 + \alpha_2\nu_2 + \ldots + \alpha_s\nu_s)$.

die, je nachdem die Bewegung periodisch ist oder nicht, zu einer Summe oder zu einem Integral vereinigt $x(t)$ darstellen:

$$x(n,\, t) = \sum_{\alpha=-\infty}^{\infty} U_\alpha(n)\, e^{i\omega(n)\,\alpha t}$$

bzw.

$$x(n,\, t) = \int_{-\infty}^{+\infty} U_\alpha(n)\, e^{i\omega(n)\,\alpha t}\, d\alpha\,. \tag{IIa}$$

Eine solche Vereinigung der entsprechenden quantentheoretischen Größen scheint wegen der: Gleichberechtigung der Größen n, $n-\alpha$ nicht ohne Willkür möglich und deshalb nicht sinnvoll; wohl aber kann man die Gesamtheit der Größen

$$U(n,\, n-\alpha)\, e^{i\omega(n,\, n-\alpha)\, t}$$

als Repräsentant der Größe $x(t)$ auffassen und dann die oben gestellte Frage zu beantworten suchen: Wodurch wird die Größe $x(t)^2$ repräsentiert?

Die Antwort lautet klassisch offenbar so:

The answer in classical theory is obviously:

$$B_\beta(n)\, e^{i\omega(n)\,\beta t} = \sum_{\alpha=-\infty}^{\infty} U_\alpha U_{\beta-\alpha} e^{i\omega(n)\,(\alpha+\beta-\alpha)\, t} \tag{III}$$

bzw.

$$= \int_{-\infty}^{+\infty} U_\alpha U_{\beta-\alpha}\, e^{i\omega(n)\,(\alpha+\beta-\alpha)\, t}\, d\alpha \tag{IV}$$

wobei dann

$$x^2(t) = \sum_{\beta=-\infty}^{\infty} B_\beta(n)\, e^{i\omega(n)\,\beta t} \tag{V}$$

bzw.

$$= \int_{-\infty}^{+\infty} B_\beta(n)\, e^{i\omega(n)\,\beta t}\, d\beta\,. \tag{VI}$$

Quantentheoretisch scheint es die einfachste und natürlichste Annahme, die Beziehung (III, IV) durch die folgenden zu ersetzen:

$$B_\beta(n,\, n-\beta)\, e^{i\omega(n,\, n-\beta)\, t} = \sum_{\alpha=-\infty}^{\infty} U(n,\, n-\alpha)\, U(n-\alpha,\, n-\beta)\, e^{i\omega(n,\, n-\beta)\, t} \tag{VII}$$

bzw.

$$\int_{-\infty}^{\infty} U(n,\, n-.\alpha)\, U(n-\alpha,\, n-\beta)\, e^{i\omega(n,\, n-\beta)\, t}\, d\alpha \tag{VIII}$$

und zwar ergibt sich diese Art der Zusammensetzung nahezu zwangläufig aus der Kombinationsrelation der Frequenzen. Macht man diese Annahme (VII) und (VIII), so erkennt man auch, dass die Phasen der quantentheoretischen U eine ebenso große physikalische Bedeu-

tung haben wie die in der klassischen Theorie: nur der Anfangspunkt der Zeit und daher eine allen U gemeinsame Phasenkonstante ist willkürlich und ohne physikalische Bedeutung; doch die Phase der *einzelnen* U geht wesentlich in die Größe B ein. Eine geometrische Interpretation solcher quantentheoretischer Phasenbeziehungen in Analogie zur klassischen Theorie scheint zunächst kaum möglich.

We see that Heisenberg laid the core of his argumentation on the coherence of the mathematical representation of the physical quantities, a coherence that he naturally postulated to be valid in the unknown quantum case as well. The success of his reasoning was spectacular. As any student of quantum mechanics knows today, the combination rule (VII) corresponds to a matrix multiplication. It was Born's achievement to recognize it for the first time a couple of months later (Born and Jordan 1925). Once this becomes clear, the declared strategy followed by Heisenberg yields the key to a proper understanding of the nature of the double arrow correspondence in our diagram above. Physical quantities are mathematically associated in quantum theory to sets of coefficients (amplitudes), thus matrices, and as became clear later, to operators (von Neumann 1927, 1932). The writing of genuine quantum equations, which implied in general the ability to effect algebraic operations on the representatives of physical quantities, became possible. (Matrix) quantum mechanics was discovered founded upon the requirement of the coherence of the rule of correspondence between physical quantities and their mathematical representatives. It appears to me that one touches here the very core of the intimate link between physics and mathematics where the creativity of Heisenberg could function full steam.

3. Anschaulichkeit

The multiplication rule above which founded quantum mechanics was indeed a genuine *breakthrough*. The familiar "natural" link between the mathematical expressions of mechanics and what they stood for in the realm of physical processes was suddenly sharply cut. Thinking about it, Heisenberg's frontal clash with classical *Anschaulichkeit* took place just there. During the crucial phase of his work, Heisenberg's declared purpose was to "kill totally the concept of orbit". And indeed, formal classical *Anschaulichkeit* was given up as soon as the Fourier series lost its status of a (convergent) series expansion and became an kind of "generating" series. Somehow paradoxically, the requirement of concentrating on observable quantities eventually brought, with respect to the *Anschaulichkeit* of the overall formalism, quite an opposite effect. The formal carrier of the so-declared only observable quantities (amplitudes), namely the series, was itself classically related to an *anschaulich* entity, being for instance the expression of a dipole in space. Stripped off of his coefficients, it became irrelevant, and so its (spatio-temporal) reference; the amplitudes were left floating, so to speak, in the air, without an formal organizational support. This was indeed a radical step out of an *anschaulich* formalism. It took the whole development of quantum mechanics, especially transformation theory, to attach those amplitudes to a new formal support, this time an operator. But the *Anschaulichkeit* of the new formalism was gone forever. Or was it?

It is important to emphasize that the claim of *Anschaulichkeit* went in these times far beyond the requirement of the use of pictorial mechanical models in physics. It rather had generally to do with the representability of physical processes in space and time. As such, the requirement of *Anschaulichkeit* brought about the issue of kantian intuition and of its *a priori* forms. It is well known that the Copenhagen philosophy of quantum mechanics considered itself as a radical denial of Kantism, a second blow after the non-euclidean geometries of Einstein's relativity. Indeed, the Copenhagen philosophy even denied the often granted separation of reality into two realms, the objective and the subjective, according to the judgment of an impersonal, ahistorical subject.

This philosophical debate was for a good part triggered by Heisenberg's 1927 paper on the indeterminacy relations (Heisenberg 1927). The necessity to discuss the *Anschaulichkeit* of the new theory was a special personal matter to Heisenberg because of the challenge constituted by the recently proposed wave mechanics of Schrödinger (Schrödinger 1926a, b, c, d). The latter's scheme was quickly gaining popularity, and Heisenberg, together with the Göttingen group feared that the wave mechanics indisputable mathematical convenience would convert many to the underlying truth of Schrödinger's physical views. The episode of the indeterminacy relations set Heisenberg on a philosophical quest which found its most profound accomplishment in his 1942 *Ordnung der Wirklichkeit*. I shall come back to this work to bring together all the threads of my talk. For the moment, let me comment on some of the points raised by the 1927 paper.

As has been emphasized by many commentators, Heisenberg's original understanding of his relations was way remote from what became soon later Bohr's complementarity philosophy. Heisenberg, heated by the fierce fight against Schrödinger's all-wave continuous quantum mechanics, tended to overstress the particle, discontinuous aspect of quantum phenomena. His analysis of the celebrated γ-ray microscope rested happily with emphasis put on the recoil of the observed particle because of the collision with the agent of observation, the light quantum. Bohr found Heisenberg's analysis quite insufficient, almost missing the main point. It was Bohr who pointed out to Heisenberg the necessity to take into account the optical nature of the way of detecting the radiation scattered off the particle. More generally, Bohr insisted against Heisenberg on the capital importance of taking into account *both* aspects, particle and wave, for tracing back the physical reasons of the uncertainty relations. It was the necessary renunciation of having both under controlled experimental scrutiny which was the reason of the relations. Soon, Bohr was to erect this observation into a full-fledged philosophy of quantum physics which he would popularize as his dialectic of complementarity. For reasons not related solely to the physical issue, the situation between both men deteriorated, which did not speed up the process of reaching agreement. Still, Heisenberg eventually ended up converted to Bohr's broader views and became a believer a second time. Not only did he spread Bohr's views, he sharpened them on occasion, and even made them more radical for the sake of his own philosophy.

But what about *Anschaulichkeit*? In his indeterminacy relations paper Heisenberg proposed to modify the meaning of the *Anschaulichkeit* requirement put upon a theory, breaking resolutely up with the classical tradition of space-time representability. He wrote, in an apparent rebuttal of the claim of the exclusive *Anschaulichkeit* of Schrödinger's theory (Heisenberg 1927, p. 172):

> Eine physikalische Theorie glauben wir dann anschaulich zu verstehen, wenn wir uns in allen einfachen Fällen die experimentellen Konsequenzen dieser Theorie qualitativ denken können, und wenn wir gleichzeitig erkannt haben, dass die Anwendung der Theorie niemals innere Widersprüche enthält.

Needless to say, this was quite a departure from the implicit requirement of representability in space and time. A theory, claimed Heisenberg, need not be space-time representable in order to offer an *anschaulich* content. Thus, quantum mechanics, *his* matrix mechanics, could be *anschaulich* after all! Many years later though, the controversy about the exact meaning of this new *Anschaulichkeit* still continues. However, in 1927, transformation theory and its mathematical elaboration by John von Neuman on one side, and Bohr's complementarity philosophy on the other, achieved what Heisenberg was since long after: a universal calculatory scheme endowed with an exhaustive meaning. One finally could harvest without second thoughts the ripe fruits of the new understanding.

This brings me to one more aspect of this story. As observed above, it was to a certain extent the multiplicity of quantum formalisms, especially the competition between the wave and matrix mechanics, which prompted the interpretative quest. But it is transformation

theory, a unified formalism born out of this clash, which provided Heisenberg with just the possibility of making a statement on the joint indeterminacies of conjugate observables. Somehow, the claim of a better *Anschaulichkeit* of Schrödinger's theory triggered the confrontation, while, simultaneously, this theory prompted a formal unification which gave Heisenberg just the arms to defeat (at least in the majority's opinion) Schrödinger's views.

In the many talks and writings that Heisenberg started, as a young growing celebrity, to steadily dispense from 1927 on[6)], the problem of the apparent loss of *Anschaulichkeit* surfaces again and again. Heisenberg presents the issue in terms increasingly more related to general problems of knowledge, revisiting and criticizing the basic assumptions made at the beginning of modern science and philosophy. Thus, the whole tradition set upon Cartesian dualism and the ensuing object-subject cut had to be revised. In the *Ordnung der Wirklichkeit*, Heisenberg analyzed quite extensively the impossibility of the classical mode of objectivation in quantum theory. His analysis insists of course on the impossibility of classical spatio-temporal representability of quantum phenomena but, more interestingly, discusses as well the loss of *Anschaulichkeit* of the notion of a mechanical state. We will find here the connection with *abstraction*, our last keyword. Thus (p. 250–251):

Die Gesetze der Quantentheorie haben etwa die folgende Form: Der „Zustand" eines atomaren Systems kann durch bestimmte „Zustandsgrößen" oder „Zustandsfunktionen" beschrieben werden; diese Zustandsgrößen stellen aber nicht unmittelbar einen Vorgang oder eine Situation in Raum und Zeit dar, wie die der klassischen Mechanik; es sind nicht etwa einfach die Orte und die Geschwindigkeiten der Teilchen, die einen Zustand charakterisieren. [...] diese Zustandsgrößen [sind] vielseitiger als die der klassischen Theorie. Die Atome haben ja noch andere Eigenschaften als die von mechanischen Systemen der klassischen Physik, insbesondere Eigenschaften, die mit den „sinnlichen Qualitäten" zu tun haben. Die Zustandsgrößen enthalten auch noch Angaben über die Wahrscheinlichkeit bestimmter Werte dieser anderen Eigenschaften; z. B. die Wahrscheinlichkeit für eine bestimmte Farbe oder für eine chemische Affinität des Atoms.

Die Zustandsgröße kann nicht selbst mit einem anschaulichen Begriff etwa von der Art: Ort, Geschwindigkeit, Farbe, Temperatur verknüpft werden. Sondern sie kann nur sozusagen im Hinblick auf solche anschaulichen Eigenschaften analysiert werden und bezeichnet dann die Wahrscheinlichkeit dafür, dass jene Eigenschaft, die beobachtet werden soll, ganz bestimmte Werte annimmt. Dabei kann die Wahrscheinlichkeit in besonderen Fällen der Sicherheit beliebig nahe kommen und in diesen Fällen kann daher gesagt werden, dass die Zustandsgröße eine bestimmte objektive Eigenschaft des Systems bezeichne; aber auch dann enthält die Zustandsgröße *mehr* als die Bezeichnung der betreffenden Eigenschaft, und dieses „mehr" ist nicht ein „objektiver" Tatbestand.

An diesem Verhältnis sind zwei Züge besonders wichtig: Der Umstand, dass die Zustandsgröße mit der Gesamtheit ihrer Aussagen nicht selbst einen objektiven Sachverhalt in Raum und Zeit bezeichnet, und die Notwendigkeit, durch die Beobachtung jene Analyse der Zustandsgröße vorzunehmen, die ihren Zusammenhang mit der Wirklichkeit vermittelt.

Was zunächst den ersten Punkt betrifft, die Unmöglichkeit, die Zustandsgrößen im gewöhnlichen Sinne zu objektivieren, so folgt diese Unmöglichkeit am deutlichsten aus dem abstrakten mathematischen Charakter dieser Größen. Sie sind häufig formal dargestellt durch Funktionen in mehrdimensionalen abstrakten Räumen, die also sicher nicht unmittelbar einen Vorgang in Raum und Zeit unserer Anschauung bedeuten können. Der Zustandsbegriff der Quantentheorie ist wesentlich abstrakter, als etwa der Temperaturbegriff der Wärmelehre, der sich doch an eine sinnliche Vorstellung anlehnt. Aber erst durch diese

[6)] In his *Collected Works* (Heisenberg 1984) one counts, from 1927 until 1942, year of the *Ordnung der Wirklichkeit*, some twenty of them, most with a substantial epistemological and historical content.

Abstraktheit entsteht der Reichtum, der die Zustandsgröße befähig, zu sinnlichen Qualitäten ganz verschiedener Art in Beziehung zu treten, über sie etwas auszusagen. Die Zustandsgröße bezeichnet also im allgemeinen nicht eine bestimmte den Sinnen zugängliche Eigenschaft des Systems, sondern eine besonders geartete Fülle von Möglichkeiten für alle sinnlich wahrnehmbaren Eigenschaften.

I have quoted Heisenberg at some length, since one rarely finds in the writings of the tenants of the Copenhagen interpretation such explicit statement about the meaning of the wave-function. Bohr, whose philosophy underlies Heisenberg's, did not care, or, presumably, rather judged useless to offer an epistemological critique of the symbols used in the quantum formalism. Heisenberg just does it, and it seems to me that his position deserves to be better known when discussing the meaning of the wave function within the Copenhagen tradition. It appears that the issue is far more subtle than the usual dichotomy between its individual versus collective-ensemble meaning.

4. Abstraction

Given its radical departure from the *Anschaulichkeit* of classical physics, and the sophistication of its mathematical formalism, quantum theory is often considered as the most prominent example of the increasingly abstract turn of fundamental research. We just saw how Heisenberg himself acknowledged the highly abstract status of the symbols used in quantum mechanics. However, he emphasized as well the fundamental need for a (scientific) thinking which keeps close to intuition and analogy. How could he reconcile those apparently contradictory trends? The key is to be found in his overall philosophical stance presented in his 1942 manuscript.

But first, what is according to Heisenberg to be gained with abstraction? A specific answer can already be found in an excerpt from the quote above: „erst durch diese Abstraktheit entsteht der Reichtum, der die Zustandsgröße befähigt, zu sinnlichen Qualitäten ganz verschiedener Art in Beziehung zu treten, über sie etwas auszusagen". However, the dialectic of abstraction versus intuition took in Heisenberg's mind a much broader scope.

In a conference given in Budapest in 1941, Heisenberg introduced in full extent the important and highly epistemological issue of the theories of light of Goethe and Newton[7]. It served his purpose to illustrate and further discuss the often stigmatized opposition between the dry, artificial questioning of Nature through the method of exact sciences, and the intuitive inquiry of Goethe, closer to a global appraisal and respectful of inner experience. Throughout the text of his conference, Heisenberg definitely shows an understanding, even sympathy towards Goethe's fight against the "imperialism" of modern science represented by Newton. There is however no question of a fundamental contradiction created by the coexistence of both approaches, since they actually belong to different appraisals of reality. Modern science requires a renunciation with respect to just the kind of deeper knowing advocated by Goethe. This is the price to pay in order to gain mathematical clarity and a preliminary condition to access the realm of higher domains. In Heisenberg's words (p. 158–160):

Wenn Goethe sagt, dass das, was der Physiker mit seinen Apparaten beobachtet, nicht mehr die Natur sei, so meint er ja wohl auch, dass es weitere und lebendigere Bereiche der Natur gebe, die eben dieser Methode der Naturwissenschaft nicht zugänglich seien. In der Tat werden wir gerne glauben, dass die Naturwissenschaft dort, wo sie sich nicht mehr der leblosen, sondern der belebten Materie zuwendet, immer vorsichtiger werden muß mit den

[7] Die Goethe'sche und die Newton'sche Farbenlehre im Lichte der moderne Physik, in Heisenberg 1984, p. 146.

Eingriffen, die sie zum Zwecke der Erkenntnis an der Natur vornimmt. Je weiter wir unseren Wunsch nach Erkenntnis auf die höheren, auch die geistigen Bereiche des Lebens richten, desto mehr werden wir uns mit einer nur aufnehmenden, betrachtenden Untersuchung begnügen müssen. Von diesem Standpunkt aus erschiene die Einteilung der Welt in einen subjektiven und einen objektiven Bereich als eine all zu große Vereinfachung der Wirklichkeit. Vielmehr könnten wir an eine Einteilung in viele ineinandergreifende Bereiche denken, die sich durch die Fragen, die wir an die Natur richten und durch die Eingriffe, die wir bei ihrer Beobachtung zulassen, voneinander abschließen. Bei dem Versuch, eine solche Einteilung in einfachen Begriffen festzulegen, werden wir erinnert an eine verwandte Ordnung der Bereiche, die wir in Goethes Nachträgen zur Farbenlehre lesen. Goethe betont dabei, dass alle Wirkungen, die wir in der Erfahrung bemerken, auf die stetigste Weise zusammenhängen; dass es gleichwohl unvermeidlich sei, sie voneinander zu trennen.

[. . .]

Trennt man in dieser Weise die Wirklichkeit in verschiedene Gebiete, so löst sich der Widerspruch zwischen der Goetheschen und der Newtonschen Farbenlehre von selbst. Die beiden Theorien stehen an verschiedenen Stellen in jenem großen Gebäude der Wissenschaft. Sieher kann die Anerkennung der modernen Physik den Naturforscher nicht hindern, auch den Goetheschen Weg der Naturbetrachtung zu gehen und weiter zu verfolgen. Freilich wäre die Hoffnung, dass wir von dieser Erkenntnis aus schon bald zu einer lebendigeren und einheitlicheren Stellung zur Natur zurückkehren könnten, noch verfrüht; denn unserer Zeit scheint es aufgegeben, die niederen Bereiche der Natur durch die Experimente zu erkennen und durch die Technik uns anzueignen. Wir müssen bei diesem Vordringen auf den Gebieten der exakten Naturwissenschaft also einstweilen an vielen Stellen auf die lebendige Berührung mit der Natur verzichten, die Goethe als Vorbedingung für die tiefere Naturerkenntnis erschien. Wir nehmen diesen Verzieht auf uns, weil wir dafür ganz weite Zusammenhänge erkennen und in mathematischer Klarheit durchschauen können – Zusammenhänge, die doch wohl die Vorbedingung auch für völlig klares Verständnis der höheren Bereiche darstellen. Wem dieses Verlassen der unmittelbar lebendigen Region ein zu großes Opfer erscheint, der wird einstweilen den Weg der exakten Naturwissenschaft nicht verfolgen können. Nur dort, wo diese Wissenschaft an den äußersten Grenzen ihrer bisherigen Forschungsweise Beziehungen zum Leben selbst entdeckt, wird ihm ihr Sinn verständlich.

Reaching the end of his talk, Heisenberg offered this poetic account of abstraction (*ibid.* p. 160):

Aber vielleicht dürfen wir den Naturforscher, der das Gebiet der lebendigen Anschauung verläßt, um die großen Zusammenhänge zu erkennen, vergleichen mit einem Bergsteiger, der den höchsten Gipfel eines gewaltigen Gebirges bezwingen will, um von dort das Land unter ihm in seinen Zusammenhängen zu überschauen. Auch der Bergsteiger muß die von den Mensellen bewohnten fruchtbaren Täler verlassen. Je höher er kommt, desto weiter öffnet sich das Land seinem Blick, desto spärlicher wird aber auch das Leben, das ihn umgibt. Schließlich gelangt er in eine blendend klare Region von Eis und Schnee, in der alles Leben erstorben ist, in der auch er selbst nur noch unter großen Schwierigkeiten atmen kann. Erst durch diese Region hindurch führt der Weg zum Gipfel. Aber dort oben steht er in den Momenten, in deuen in vollster Klarheit das ganze Land unter ihm ausgebreitet liegt, doch vielleicht dem lebendigen Bereich nicht all zu fern. Wir verstehen, wenn frühere Zeiten jene leblosen Regionen nur als grauenvolle Öde empfanden, wenn ihr Betreten als eine Verletzung der höheren Gewalten erschien, die sich wahrscheinlich bitter an dem rächen werden, der sich ihnen zu nahen wagt. Auch Goethe hat das Verletzende in dem Vorgehen der Naturwissenschaft empfunden. Aber wir dürfen sicher sein, dass auch dem Dichter Goethe jene letzte und reinste Klarheit, nach der diese Wissenschaft strebt, völlig vertraut gewesen ist.

Given the argumentation above, one might then think that, within the domain (Bereich) of reality where scientific methods are operative, and accepting the ensuing collateral impoverishment, there is no limit to the power of abstract scientific methods. This conclusion is in fact incorrect. It takes more grounding into the philosophical conceptions of Heisenberg to learn his views on the question. It turns out that Heisenberg actually judged formal deductive thinking unable to bring by itself true novelty, for instance in launching bridges from know territories of knowledge towards new, yet unknown, but adumbrated ones. This can be properly understood only in the broader context of Heisenberg's philosophical statement on the nature and possibility of knowledge of reality that he developed in the months following his Budapest conference. The conception of the different domains of reality appears there as the key element of Heisenberg's philosophical position. Before going into more detail, it is worth recalling the historical context of this work.

The progressive shaping of Heisenberg's understanding of his science, his *Weltanschauung* and the tragic events of Nazi times of which he was more than a passive spectator triggered in him some sort of strive for a global appraisal of reality. It happened during the most gloomy period of Heisenberg's life, the middle of the war. Many scholars have commented Heisenberg's action at that time and I will not go any further into it. It must however be pointed out that the work which I want to discuss now, the 1942 manuscript, is a product of many necessities in Heisenberg's life, purely disinterested philosophical grounding being only one of them.

Prior to quantum mechanics, *Anschaulichkeit* was the norm of physical understanding and was often used as such in scientific controversies. One of the last episodes, surely the most tragic one, took place with the assault of the *Deutsche Physik* against theoretical physics, as a way to harm its prominent Jewish representatives. Under the banner of *Anschaulichkeit*, Lenard, Stark and the like tried to eradicate the very jewels of German 20th century theoretical physics, claiming the sole legitimacy of experimental physics, supposed to humbly, respectfully and faithfully collect the bare facts of genuine Nature. We know how Heisenberg fought to defend his discipline and how finally he won a bitter victory. However, Heisenberg's fight lead him on the path of a conception with a much broader scope than the quarrel of theoretical (mathematical) versus experimental physics.

Apart the desire to fight the *Deutsche Physik* from a philosophically consolidated point of view, the need of an assessment of the social and political turmoil of his time, as well as an endeavor to clarify his own stand were presumably Heisenberg's other important motivations[8)]. All this culminated in a new conception of reality which makes retrospectively Heisenberg on par with other contemporary thinkers. Heisenberg erected his vision on the basis of Bohr's philosophy, sharpened some of the latter's positions, and borrowed from other important sources of his time. The result is a surprising and appealing vision, a proposal for a genuinely new understanding of reality and of the manifold ways of the latter's appraisal by men. It is also a pathetic call for a new humanism, where all the accounts of reality coexist, mix, and mutually enrich each other without hierarchical nor reductionist trends. Science is just only one of many others, infinite in number.

It would be difficult to convey all the wealth of Heisenberg's manuscript in such a short space. Let me then just put forward the key elements which will enlighten and bring unity to my previous statements.

The 1942 manuscript builds around the theme of domains of reality already present in Heisenberg's Budapest conference. Thus, reality is conceived as made up of many intertwined domains, each defined by a different layer of connections put forth by a specific type of activity/inquiry (p. 233–234):

[8)] See Catherine Chevalley's long and insightful introduction to the french translation of the 1942 manuscript (Heisenberg 1998). In the sequel, I closely follow her account.

[...] wenn erkannt ist, dass es keinen Stoff gibt, der bestimmten Gesetzen folgt, sondern dass es eben nur gesetzmäßige Zusammenhänge gibt, die wir erfahren können und zu deren Beschreibung wir auch gelegentlich Wörter wie Stoff oder Materie verwenden – erst dann kann der Satz richtig verstanden werden, dass es sich bei der gesuchten Einteilung um eine Ordnung der Wirklichkeit nach gesetzmäßigen Zusammenhängen handeln soll. Unter einem Bereich der Wirklichkeit – soweit das Wort Bereich in diesem speziellen Sinne der Einteilung gebraucht wird – verstehen wir also eine Gesamtheit von gesetzmäßigen Zusammenhängen. Eine solche Gesamtheit muß einerseits eine feste Einheit bilden, sonst könnte man nicht mit Recht von einem sprechen, andererseits muß sie scharf gegen andere Gesamtheiten abgegrenzt werden können, damit eben eine Einteilung der Wirklichkeit möglich wird.

The new conception of reality entails then a layering of networks of connections – the domains – rather than a global unitary appraisal. Also, human experience suggest to arrange the domains following a progression which does not justify the traditional opposition between objective facts and subjective appearances (p. 235–236):

Die Ordnung der Wirklichkeit, die wir suchen, soll vom Objektiven zum Subjektiven aufsteigen. Sie soll also anfangen mit einem Teil der Wirklichkeit, den wir ganz außerhalb von uns stellen können, in dem wir ganz absehen können von den Methoden, mit Hilfe derer wir von seinem Inhalt Kenntnis erlangen. An der Spitze der Ordnung aber stehen, wie in dem Goethe'schen Entwurf, die schöpferischen Kräfte, mit deren Hilfe wir selbst die Welt erwandeln und gestalten. Dabei darf das Wort subjektiv nicht in dem Sinne mißverstanden werden, als handele es sich bei den höheren Bereichen etwa zum Teil um Sachverhalte, die nur für unser Gefühl oder nur für bestimmte Menschen existierten, oder um irgendeine Art von eingebildeter Wirklichkeit. Mit dem Wort subjektiv ist keineswegs der bloße Schein gemeint. Im Gegenteil enthält jeder dieser Bereiche zum größten Teil Zusammenhänge, die sich vollständig objektivieren lassen. Mit dem Wort subjektiv soll nur angedeutet werden, dass es bei einer vollständigen Beschreibung der Zusammenhänge eines Bereiches vielleicht nicht möglich ist, davon abzusehen, dass wir selbst in die Zusammenhänge verwoben sind.

The reevaluation of the "subjective", understood as referring to the degree of participation of the subject and its cognitive activity in the very realm of what is under scrutiny, is exemplified by the epistemological issues of atomic physics but has further implications (*ibid*):

So kann zum Beispiel in der Atomphysik, wenn ihre Gesetze vollständig formuliert werden sollen, nicht mehr davon abgesehen werden, dass unser Körper und die Apparate, mit denen wir beobachten, selbst den Gesetzen der Atomphysik unterworfen sind; ferner tritt in ihr unsere *Kenntnis* eines Sachverhaltes an Stelle eines physikalischen Faktums. In noch viel stärkerem Maße wird das Entsprechende etwa in der Psychologie gelten. Die Existenz der menschlichen Seele ist zwar ein objektiver Tatbestand, den wir ganz außerhalb von uns stellen können. In die Formulierung psychologischer Gesetze wird aber der Umstand wesentlich eingehen, dass wir selbst beseelte Wesen sind. Auch die Kräfte, von denen das Gleichnis der Religion spricht, sind objektiv.

Wenn also gesagt wird, dass es sich um eine Ordnung handeln soll, die vom Objektiven zum Subjektiven aufsteigt, so ist damit gemeint, dass in immer steigendem Maße das Erkenntnisverfahren, das uns von der Wirklichkeit Kunde gibt, selbst einen Bestandteil der Zusammenhänge bildet, die den betreffenden Bereich ausmachen.

[...]

Gegen den Plan der Einteilung, der in diesen Vorbemerkungen zum Vorschein kommt, wird man einwenden können, dass hier eigentlich nicht von Bereichen, sondern von bestimmten Idealisierungen der Wirklichkeit die Rede sei. Man kann z. B. sagen, dass die

klassische Physik diejenige Idealisierung darstellt, bei der wir von dem Erkenntnisverfahren, das uns über die Wirklichkeit unterrichtet, ganz absehen können. Es sieht also so aus, als ob der eingeschlagene Weg nicht zu einer Ordnung der Wirklichkeit, sondern zu einer Ordnung unseres Verständnisses oder unserer Kenntnis der Wirklichkeit führen würde. Nun hat allerdings schon der Begriff Ordnung nicht nur die zu ordnende Sache sondern auch uns selbst zur Voraussetzung, und insofern erscheint es nicht verwunderlich, wenn bei einer Ordnung nicht entschieden werden kann, ob sie uns als eine Ordnung der Wirklichkeit oder unseres Verständnisses der Wirklichkeit entgegentritt. Ferner versteht es sich von selbst, dass jede ordnende Einteilung der Wirklichkeit eine Idealisierung erfordert, denn die Wirklichkeit umfängt, uns zunächst als ein stetig in sich fließender Zusammenhang, aus dem wir erst durch den Eingriff unseres Denkens – das eben insofern idealisiert – bestimmte Vorgänge, Erscheinungen, Gesetze herauslösen. Schließlich aber muß man sich immer wieder klar machen, dass die Wirklichkeit, von der wir sprechen können, nie die Wirklichkeit an sich ist, sondern eine gewußte Wirklichkeit oder sogar in vielen Fällen eine von uns gestaltete Wirklichkeit.

How should one understand now the mechanism underlying the emergence of a domain of reality? Pattern of connections are distinguished within the framework of a practice of a specific language. We bring forth various patterns of connections using language as an exploratory tool, not only as a means of mere description, but as one of a genuine elaboration and structuring within the continuous manifold of experience. Language, instead of merely describing states of affairs exterior to our cognition, actually participates in their emergence. We ask questions about what we experience, we put forth tentative language acts, the latter's effect affecting back our usage and understanding of them. This continuous back and forth process, indeed the very process of knowledge, erects its object of inquiry while at the same time defining its own workings. In a sense, language is the world, and the world is just language. Obviously, the concept of language Heisenberg thinks of goes largely beyond the structures making speech possible, and relates to any conceptual, semantic framework humans are capable of. Science is just one of them.

Within the (infinite) multitude of various languages and of their working, corresponding to different types of behavior, and defining the different domains of reality, one can, according to Heisenberg, distinguish two modalities, the so-called static and dynamic ones. Again, a full characterization of the two types is beyond my present scope. Let us then stop at the following superficial one. The static type of language corresponds roughly to an attitude where emphasis is put on a systematic grounding of the connections, until such connections are reached which offer a character of clarity and constancy. This clearly characterizes science, in particular physics and other mathematical sciences (p. 223–225):

Dabei kann die Abbildung von Sachverhalten in der Sprache in zwei verschiedenen Weisen erfolgen, die man etwa als „statisch“ und „dynamisch“ unterscheiden, wenn auch nicht scharf trennen kann. Die Sprache kann einerseits versuchen, durch eine immer weitergehende Verschärfung der Begriffe zu einer immer genaueren Abbildung des gleichen gemeinten Sachverhalts zu kommen. Diese Verschärfung erfolgt durch eine ins Einzelne gehende Festlegung von Beziehungen zwischen den Begriffen – etwa Zurückführung spezieller Begriffe auf allgemeinere – oder durch die ad hoc vorgenommene Zuordnung der Begriffe zu ganz speziellen Erfahrungsinhalten. Die wissenschaftlichen Sprachen – etwa die der Rechtslehre oder die mathematisch formulierte Naturbeschreibung – geben Beispiel für solche Verschärfungen. Dabei kann schließlich ein völlig starres Schema von Verknüpfungsregeln zwischen den Begriffen, und von Begriffen zu Erfahrungsinhalten gebildet werden, so dass von jedem Satz, der dieses Begriffssystem benützt, eindeutig entschieden werden kann, ob er „richtig“ oder „falsch“ ist. Dabei wird freilich die Frage, wie genau dieses Begriffssystem den gemeinten Teil der Wirklichkeit abbildet, ausschließlich durch den Erfolg ent-

schieden. Ein vollständiges und exaktes Abbild der Wirklichkeit kann nie erreicht werden. Aber es wird erlaubt sein – wenn das betreffende Begriffssystem sich bewährt – von einem exakten Abbild des „wesentlichen Teiles“ des betreffenden Sachverhaltes zu sprechen; denn dadurch wird ja nur festgesetzt, auf welche Teile wir unser Augenmerk richten wollen.

The static language, because of the very rigidity of its logical conceptualizing, entails a necessary narrowing of its subject matter and an impoverishment of its concepts (p. 224):

Diese Verschärfung der Sprache, auf Grund deren dann von jedem Satz entschieden werden kann, ob er „richtig“ oder „falsch“ ist, geht freilich in vielen Fällen Hand in Hand mit einer Verarmung der in ihr vorkommenden Begriffe. Die Wörter einer solchen Kunstsprache beziehen sich – im Gegensatz zu den Wörtern der gewöhnlichen Sprache – nur noch auf ganz bestimmte Zusammenhangsbereiche. Dabei kann der Teil der Wirklichkeit, der durch die Kunstsprache abgebildet wird und der für den eingenommenen wissenschaftlichen Standpunkt „wesentlich“ ist, von anderen Gesichtspunkten aus als unwichtig erscheinen. Diese vorhin als „statisch“ bezeichnete Darstellung eines Teiles der Wirklichkeit ist also unvermeidbar mit einem schwerwiegenden Verzicht verknüpft: dem Verzicht auf jenes unendlich vielfache Bezogensein der Worte und Begriffe, das in uns erst das Gefühl erweckt, etwas von der unendlichen Fülle der Wirklichkeit verstanden zu haben.

The dynamic type of language corresponds on the other hand to an experience of reality where intuition, analogy, and, more generally, all that Goethe was missing in the Newtonian science, is effective (p. 224–225):

Der „statischen“ kann nun eine andere Art der Darstellung der Wirklichkeit gegenübergestellt werden, die eben durch das unendlich vielfache Bezogensein der Worte erst ermöglicht wird und die man als „dynamisch“ bezeichnen kann. In ihr soll der ausgesprochene Gedanke nicht ein möglichst getreues Abbild der Wirklichkeit sein, sondern er soll den Keim zu weiteren Gedankenreihen bilden; nicht auf die Genauigkeit, sondern auf die Fruchtbarkeit der Begriffe kommt es an. An einem Gedanken gliedern sich durch die vielfachen Bezüge neue Gedanken an, aus diesen entstehen wieder neue, bis schließlich durch die inhaltliche Fülle des von den Gedanken durchmessenen Raumes nachträglich ein getreues Abbild des gemeinten Wirklichkeitsbereichs entsteht. Diese Art der Darstellung beruht auf der Lebendigkeit des Wortes. Hier kann ein Satz im allgemeinen nicht „richtig“ oder „falsch“ sein. Aber man kann einen Satz, der fruchtbar zu einer Fülle weiterer Gedanken Anlaß gibt, als „wahr“ bezeichnen. Das Gegenteil eines „richtigen“ Satzes ist ein „falscher“ Satz. Das Gegenteil eines „wahren“ Satzes wird aber häufig wieder ein „wahrer“ Satz sein.

[...]

Im Bereich des „statischen“ Denkens wird erklärt – wie überhaupt die Klarheit das eigentliche Ziel dieser Denkform ist –, im Bereich der „dynamischen“ wird gedeutet. Denn hier werden unendlich vielfältige Beziehungen zu anderen Bereichen der Wirklichkeit gesucht, auf die wir deuten können.

[...]

Im Allgemeinen wird jeder Versuch, über die Wirklichkeit zu sprechen, gleichzeitig „statische“ und „dynamische“ Züge tragen. Dem klaren, rein statischen Denken droht die Gefahr, zur inhaltlosen Form zu entarten. Das dynamische Denken kann vage und unverständlich werden.

We are now in a better position to understand the respective parts of abstraction and intuition in scientific research, according to Heisenberg's views. The static modality accounts for the steady grounding into physical knowledge, the continuous sharpening of concepts, the strive for strict connections, laws, and other statements characteristic of a

scientific account of reality. Quantitative, precise accounts and subsequent formalizations are obtained at the price of a progressive draining of facts. But, none of the appraisals obtained that way can claim a universal range and validity. In the extension of (scientific) knowledge, intuition and other "non-logical" thought processes are crucial (p. 225):

Das Ziel der exakten Naturwissenschaft bilden zwar stets in sich geschlossene Systeme von Begriffen und Axiomen, die den gemeinten Teil der Wirklichkeit streng abbilden. Der Gang der Forschung aber, die von bekannten Begriffssystemen ausgehend die Ordnung eines neuen Erfahrungsbereichs anstrebt, kann nicht auf den durch logische Schlussketten vorgezeichneten Pfaden erfolgen. Der Abgrund zwischen den schon bekannten und den neuen Begriffssystemen kann durch intuitives Denken übersprungen, nicht durch formales Schließen überbrückt werden.

Because of the irreductible indeterminacy inherent to tany language, there can be no language with a universal scope[9]. Hence there is in particular no universal (scientific) theory. Genuine extension of knowledge cannot be obtained otherwise than by a switch to a different language, not by an extension of an existing one claiming universal validity. This is what happened in the quantum crisis and it is bound to happen again. But we should be confident in the everlasting human capacity to build up new languages fit to account for new patterns of connections, new domains (p. 226):

Obwohl hieraus die engen Grenzen deutlich werden, die jeder speziellen wissenschaftlichen Beschreibung der Wirklichkeit gesteckt sind, so besteht doch andererseits kein Grund dafür, prinzipielle Grenzen anzunehmen für die Fähigkeit des Menschen, irgendwelche Bereiche der Wirklichkeit schließlich zu verstehen. Im Gegenteil erscheint diese Fähigkeit der Menschen: zu verstehen, sich in der Wirklichkeit zurechtzufinden, durchaus unbegrenzt.

[. . .]

Obwohl also unser Denken stets gewissermaßen über einer grundlosen Tiefe schwebt — da wir nie von dem festen Grund klarer Begriffe aus Schritt für Schritt in das unbekannte Neuland vordringen können —, so wird dieses Denken doch schließlich jeder neuen Erfahrung, jedem zugänglichen Bereich der Welt gerecht werden können. Es wird sich immer wieder eine Sprache entwickeln, die eben zu dem ins Auge gefaßten Bereich der Wirklichkeit paßt und die Sachverhalte in diesem Gebiet genau abbildet.

Freilich wird, wie weit das Denken auch dringen mag, stets das Gefühl übrigbleiben, dass es jenseits des Erforschten noch andere Zusammenhänge gebe, die sich der sprachlichen Formulierung entziehen und deren Geltungsbereich jeweils mit dem Verständnis eines neuen Bereichs der Wirklichkeit noch einen Schritt weiter hinausgeschoben wird in das undurchdringliche Dunkel, das hinter den durch die Sprache formulierbaren Gedanken liegt. Dieses Gefühl bestimmt die Richtung, des Denkens, aber es gehört zu seinem Wesen, dass die Zusammenhänge, auf die es gerichtet ist, nicht in Worte gefaßt werden können.

Vielleicht kann man den Inhalt dessen, was in den letzten Abschnitten gesagt werden sollte, so zusammenfassen:

Jeder Bereich der Wirklichkeit kann schließlich in der Sprache abgebildet werden. Der Abgrund, der verschiedene Bereiche trennt, kann nicht durch logisches Schließen oder folgerichtiges Weiterentwickeln der Sprache überbrückt werden.

Die Fähigkeit des Menschen, zu verstehen, ist unbegrenzt. Über die letzten Dinge kann man nicht sprechen.

[9] The key issue of the indeterminacy of language is difficult to report here in sufficient detail. I prefer instead to refer the reader to Heisenberg's manuscript. Let it be sufficient to suggest here that the indeterminacy of language has at its roots a state of affairs similar to the one distinguished by Bohr in his complementarity. The very conditions which make the use of language effective have an essential impact on the content of what is expressed.

Clearly, the conception of human understanding presented by Heisenberg triggers many far reaching issues, to witness in particular the last sentences. Some of them may have a definite impact on the scientist's methodology and beliefs. As a way of conclusion, I shall merely point out just one, which might have played a role in Heisenberg's own research. He, as many other contributors and witnesses of the quantum revolution, did not later hesitate, when meeting deep difficulties, to invoke new breakdowns of established laws and theories. Bohr, confronted with the puzzle of β-decay, would hint at a violation of energy conservation and Heisenberg himself believed in a new fundamental length beyond which standard quantum mechanics would cease to be effective. This motivated him to start in the forties his far reaching S-matrix program, and later, in the fifties, to propose the non-linear field theories for a unified theory with an underlying scale. The philosophical conceptions of Heisenberg reported above might provide a clue as to why he was so prompt to expect, and even to personally trigger, new departures from established truths and ways of investigation. There simply was for him no end to human exploration of the "continuous flow of the content of human experience". Still other ways to carve out new patterns of connections, still new theories and new levels of phenomena were to be expected. Still new (scientific) revolutions were to come.

5. References

Born and Jordan 1925: M. Born and P. Jordan, *Zur Quantenmechanik*, Zeit. f. Phys. vol. **34** (1925), p. 858.

Born, Jordan and Heisenberg 1926: M. Born, P. Jordan and W. Heisenberg, *Zur Quantenmechanik*. II., Zeit. f. Phys. vol. **35** (1926), p. 557.

Cassidy 1992: D. Cassidy, *Uncertainty: the life and science of Werner Heisenberg*, New York, W. H. Freeman, 1992.

Heisenberg 1925: W. Heisenberg, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, Zeit. f. Phys. vol. **33** (1925), p. 879.

Heisenberg 1927: W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematic und Mechanik*, Zeit. f. Phys. vol. **43** (1927), p. 172.

Heisenberg 1984: W. Heisenberg, *Gesammelte Werke Collected Works*, W. Blum, H.-P. Dürr and H. Rechenberg (eds), Abt. C, Bd 1, Piper, 1984.

Heisenberg 1985: W. Heisenberg, *Gesammelte Werke Collected Works*, W. Blum, H.-P. Dürr and H. Rechenberg (eds), Series A, Part 1, Springer, 1985.

Heisenberg 1998: W. Heisenberg, *Philosophie, le manuscrit de 1942*, Catherine Chevalley (ed.), Paris, Seuil, 1998.

Hund 1967: F. Hund, *Geschichte der Quantentheorie*, Bibliographisches Institut, Hochschultaschenbücher Verlag, 1967.

von Meyenn and Weisskopf 1979: K. von Meyenn and V. F. Weisskopf (eds): *W. Pauli Wissenschaftlicher Briefwechsel 1919–1929*, New York, Springer, 1979.

Jammer 1966: Jammer, Max, *The Conceptual Development of Quantum Mechanics*, New York, McGraw-Hill Book Company; 1966.

Jammer 1974: Jammer, Max, *The Philosophy of Quantum Mechanics*, New York, 1974, John Wiley and Sons.

Mehra and Rechenberg 1982–2000: J. Mehra and H. Rechenberg, *The Historical Development of Quantum Theory*, New York, Berlin, Heidelberg, Springer, volumes 1–6.

von Neumann 1927: J. von Neumann, *Mathematische Begründung der Quantenmechanik*, Gött. Nach., (1927), pp. 1–57.

von Neumann 1932: J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Berlin, Springer 1932; *Mathematical Foundation of Quantum Mechanics*; english translation, *Mathematical Foundation of Quantum Mechanics*, Princeton, Princeton Univ. Press, 1955.

Schrödinger 1926a, b, c, d: E. Schrödinger, Quantisierung als Eigenwertproblem (Erste Mitteilung), Annalen der Physik, vol. 79 (1926), pp. 361–376; b: Zweite Mitteilung, vol. **79**, pp. 489–527; c: Dritte Mitteilung, vol. 80 (1926), pp. 437–490; d: Vierte Mitteilung, vol. **81** (1926), pp. 109–139.

Sommerfeld 1916: A. Sommerfeld, *Zur Quantentheorie der Spektrallinien*, Annalen der Physik, vol. **51** (1916), pp. 125–167.

Fortschr. Phys. **50** (2002) 5–7, 459–460

100 Years Werner Heisenberg – Works and Impact

KAZUO YAMAZAKI

Sakyoku, Kyoto 606-8287, Japan

So much we had heard about Heisenberg in this symposium, and it seems almost nothing left for me to talk about my distinguished teacher Werner Heisenberg, so I will restrict my talk about my personal memory, or episode or something like that. I was a fellow of Humboldt Foundation from 1957 to 1959 and also 1971 to 1972, and my Humboldt Betreuer was Werner Heisenberg himself. All together I collaborated with Heisenberg over ten years long at the Max Planck Institut for Physik in Goettingen and Munich, that director was also Heisenberg. Just at the time when I arrived at Goettingen in the beginning of October 1957, Heisenberg with Wolfgang Pauli engaged enthusiastically in his famous but now a notorious unified theory of elementary particles. So I was immediately absorbed in this very ambitious attempt. I liked and respected him for his talent of physics and also for his warm hearted, and so I collaborated with Heisenberg over 10 years as already spoke. One day he told me that after the war he once wanted and decided to become a Verwaltungsphysiker (administrative physicist), but unsuccessful, so after 1953 he returned to researcher. He liked and loved physics even in his late years, so after 1953 he once more persisted in to keep his time for physics. Thus he refused almost every offer of the important position or committee member in the governmental or academical or financial societies, which had asked him so much and even though he knew the importance of those. Almost the only exception was the presidency of the Humboldt Foundation from 1953 to 1975. His young years experiences in Copenhagen under Niels Bohr as a Rockefeller fellow was so important for him, and he believed the dream of solidarity of the international family of physicist built under Niels Bohr would be kept in any situation. His dream was mercilessly destroyed in World WarII, which shocked him seriously, still he believed in that the growing up of our international Humboldt Family would be decisive in order to keep the international understandings and the peace of world. This was the reason, I believe, why he kept the presidency of Humboldt Foundation so long.

For Heisenberg the most important point was the talent and nature of the person, irrespective of nationality, sex, race, religion, credo etc. So, for example, in the Nazi era in Leipzig, he adopted Hans Euler as his assistant, a firm and known communist. It was of course dangerous for Heisenberg himself, but Heisenberg wanted to save this gifted young and financially poor Euler. In the beginning of World War II the attack of Soviet against Finland and Poland acted on Euler grave influence, who lost his confidence and became a volunteer and a pilot of a reconnaissance plane and died in Soviet front. It was almost a suicide. Heisenberg was very sorry that he could not save Euler through the war. Probably the most promissing pupil of Heisenberg was this Hans Euler. The above Heisenbergs point of view still alive in the guide line for the election of fellowship in Humboldt Foundation, I believe.

Now let me tell something about the Heisenberg's unified field theory of elementary particles. In the beginning of it, from September 1957 to Feburuary 1958 about a half year long Pauli worked with Heisenberg still positively. Really was Pauli on the years change 1957–1958 enthusiastic and optimistic about the possibility to reach their common goal. This cooperation was essential for Heisenberg. For example, Pauli wrote in a letter of reply to the invitation for the summer school in Varena 1958 as follows "In the moment I am in close collaboration with Heisenberg on a very new form of the quantum theory of fields, including a new theoretical interpretation of the isospin space. The prospect looks bright at

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

present and until the summer we hope to know much more" and further in a letter of 16. Feb. 1958. to V. F. Weisskopf "In general I feel that my critical attitude and Heisenberg's optimism represent a good combination. Also I believe that Heisenberg has an excellent adviser for mathematical questions in Symanzik. We will soon see what comes out to all this. The main thing remains, that I still find the situation extremely interesting". In such a way Pauli behaved until Feb. 1958 to Heisenberg's Nonlinear Spinortheory.

In the theoretical physics group of Max Planck Institut for Physics in those days, our Heisenberg group were called sometimes as Prophets party while the others were called as Puritans party. We were also sometimes called as the "Rechensoldaten", means Heisenberg gave the command and we calculated it simply like an obediently soldier. But this was not the case, I believe, because Heisenbergs way of doing physics was based on the very severe discussion or debate between Heisenberg and one of us, and there was really no distinction under the participant of this battle. As it is well known, he was a great optimist like his father August Heisenberg and his teacher Arnold Sommerfeld. His considerations always point to positive sides of every new attempt or investigation or postulate and so on. So, in his recommendation, for example, he always emphasized positive or strong points for the candidates and kept away from weak points of those. He always emphasized the importance of the way of set up the problem (in German Fragestellung), and said like this "If we succeed to set up the correct way of approaching the problem, then we have already solved the problem half."

And when he failed to solve his problem by using the normal standard methods, then he always said that "Now we found the real essence of the problem". Once he settled the problem, he never gave up. And then he concentrated his force to this point, while the other person lost the courage and interest to attack the problem. His Fragestellung were always very intuitive and concrete and formulated mostly with mathematical expressions which contained no vague conjecture. As for example, Heisenbergs uncertainty principle is expressed in a form of mathematical inequality, while Bohrs complementarity principle is difficult to express in mathematical form. I have also learned very much through my Japanese translation of his famous book „Der Teil und das Ganze". This book is even now beloved and highly appreciated in Japan and published already over 50000. He emphasized the importance of the courage to bet at the decisive point by citing Columbus as follows." If I were asked what was Christopher Columbus' greatest achievement in discovering America, my answer would be . . . . His most remarkable feat was the decision to leave the known regions of the world and to sail westward, far beyond the point from which his provisions could have got him back home again. In science,too,it is impossible to open up new territory unless one is prepared to leave the safe anchorage of established doctrine and run the risk of a hazardous leap forward . . . ."

My son grew up in Munich and I remember in his Kindergarden days, he learned and said often like "Mein Herz ist klein, kann niemand hinein (My heart is small, can nobody break into) I think this was really Heisenberg. As Carl Friedrich von Weizsaecker aptly said "He was first of all a spontaneous man, then a genius scientist, ..." He was only responsible for his conscience. This was in every decision of Heisenberg accomplished.

Let me finish my talk with a small episode relating Heisenberg. After some years, about 1960, I finally begun to speak German. Once in a discussion with Heisenberg I said "Ich glaube, . . ." whereby I just meant like the English sentence "I think, . . ." Then excited Heisenberg said "We are here doing no Belief-science (Glaubenwissenschaft). Have you any reason or proof for your argument?" On the other hand he was so tolerant, that he had understood well my poor German and simply let me to speak so, but on the other hand he was sometimes so strict in the language. After we solved our problem and got some mathematical answer to the problem, then he always required us to express the results in the normal language. His advice in writing a scientific paper was to formulate first the Fragestellung and then our postulate clearly in the introduction, and to summarize our obtained concrete results, whereby we should write the text in detail, but in the main sections we should omit the technical details.

Fortschr. Phys. **50** (2002) 5–7, 461–482

Heisenberg-Einstein Context Principle and the Dynamic Core-context of Discovery in Physics

GIRIDHARI LAL PANDIT

Delhi University, Delhi, India

"Every discoverer of a new physical principle makes an important contribution to philosophy, even though he may not discuss it in philosophic terms"
Henry Margenau (1949)

Abstract

My main aim here is to present a general methodological principle, which I propose to designate as Heisenberg-Einstein context principle, as it emerges from their dialogue on physical theory, becoming their most important meeting point in the interpretation of quantum mechanics. I regard as context-principles those which can serve as guiding principles in theory-finding and theory-testing within physics. I think that they must be judged by their impact in making physics open-ended. As regards Heisenberg-Einstein dialogue, three important questions can be raised. The first question is which particular aspect of physical theory were they both concerned with. The second question is whether there emerged from their dialogue some general principle, as a guiding principle of sufficient methodological interest, on which they both showed significant agreement, resolving some of their methodological differences by removing some of their shared philosophical misunderstandings concerning physical theory. This gives rise to the third question whether their dialogue resulted in a significant impact on their respective positions on the interpretation, or foundations, of quantum mechanics. While answering the last two questions in the affirmative, I shall argue for the thesis that in their dialogue they shared a rather deep concern with the fundamental aspects of physical theory – with its mathematical formalism. Taken in this aspect, a physical theory, as and when it becomes mature after it has passed tests, contributes to context-building in *theory-finding* and *theory-testing*, generating new frontiers of discovery in physics and astronomy. In a nutshell, I am here concerned with the problem of reconnecting observation with theory, and discovery, or rather frontier of discovery, in physics with the core-context of discovery, trying to bring out the significance of the Heisenberg-Einstein context principle.

Heisenberg-Einstein Context Principle and the Dynamic Core-Context of Discovery in Physics[*)]
Giridhari Lal Pandit
Delhi University

Keyconcepts/abbreviations: context principles in physics; frontier of discovery (die vorderste Front der Forschung); core-context of discovery (**CCD**); (the methodology of) dynamic core-contexts of discovery (**DCCsD**); Heisenberg-Einstein context principle; asymmetry of theories and problems; search-and-discovery procedures (**SD-Procedures**); resolving power of a physical theory T – T**rp**(T); explanatory power of a physical theory T – T**ep**(T); the special theory of relativity (**STR**); the general theory of relativity (**GTR**); Einstein, Podolsky, Rosen, Gedankenexperiment (**EPRG**); Duhem-Heisenberg-Einstein methodology.

[*)] The invitation of the Alexander von Humboldt-Stiftung enabling the author to participate in its 18th Fachsymposium „100 Jahre Werner Heisenberg – Werke und Wirkung" (Bamberg: 26.–30. September, 2001) is gratefully acknowledged. It is a pleasure to acknowledge many fruitful discussions with Hans Günter Dosch, Institut für Theoretische Physik, Universität Heidelberg. For questions, critical comments and clarifications on the earlier versions of the text, from other colleagues at the Institut, notably Berthold Stech, Otto Nachtmann, Hans Jürgen Pirner, Dieter Gromes, Eduard Thommes and Klaus Hefft, among others, he expresses his sincere thanks and admiration to all of them. He is equally

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Contents

Abstract

1. The problem: reconnecting observation with theory and the frontier of discovery with the core-context of discovery 462

2. The quantum-context-building 467

3. The methodology of unification in physics 471

4. Theory and observation 472

5. The Heisenberg-Einstein context principle: the theory dominates 473

6. Turning around of the questions in understanding quantum mechanics 475

7. If it is always the theory which decides what can be observed, then what kind of fundamental theory can physics aim at? 477

8. The Duhem-Heisenberg-Einstein methodology 479

9. Conclusion 481

1. The Problem: Reconnecting Observation with Theory and the Frontier of Discovery with the Core-Context of Discovery

In early 1926 when Werner Heisenberg and Albert Einstein were engaged in a dialogue concerning the problem of reconnecting observation with theory, the leading philosophers of physics, notably Hans Reichenbach and Rudolf Carnap, among others from the Vienna Circle, and later on Karl R. Popper, were pre-occupied with the task of imparting a new direction to research in philosophy, which was defined by them as the logic of science. To

indebted to Erhard Scheibe and Helmut Rechenberg. They took the trouble of going through an earlier draft of the text.

The limitations of space here do not permit the author to discuss Heisenberg's (1989: 494) important philosophical reflections on environmental thought for improvement of environmental action. But what were his ethical concerns like where science policy and human and environmental interests are interconnected? Did he indicate a way of improvement, given that human activity has local and global impacts on environment? Briefly stated, he was one of the very few leading scientists of the world who were ahead of their times in conceptualizing this problem with remarkable insight and clarity. The problem, as he perceived it, has many levels of complexity, as it were, enfolded into it, *from* what he aptly described as the ambivalence of science *to* the paradox of all paradoxes that it is scientific and technological progress which controls moral progress and not the other way round. Approaching the *new* situation optimistically, as he did, he believed that it "reveals new and unexpected dangers, and these dangers should be considered as a challenge requiring our response." Heisenberg proposed then, as a guiding principle, "to consider any special progress in science or technology as a part of the whole, as something that cannot be separated from the general problems of our way of life, our environment, our political behaviour. This obligation – to keep in mind the unavoidable connection or interplay between *all* actions – will set and should set limits to our blind confidence in science and technology, but not necessarily to science and technology itself." For details see Werner Heisenberg (1989), *Gesammelte Werke* Band **V**, Piper: München/Zürich, p. 494. Herausgegeben von Walter Blum, Hans-Peter Dürr und H. Rechenberg. For detailed discussions see G. L. Pandit (1995), *Von der Ökologie des Bewußtseins zum Umweltrealismus: Wiederentdeckung menschlicher und nicht-menschlicher Interessenssphären.* Picus Verlag: Wien; G. L. Pandit, "Environmental Actions: Can a Part Control the Whole? in *Problems of Action and Observation.* Special Vol. **12** of **Systemica (2000: 351–368)**, Journal of the Dutch Systems Group edited by Ranulph Glanville and Gerard de Zeeuw; G. L. Pandit, "Emergence of Human Interest Studies – The Environmental Dimension" forthcoming in **Systemica.** Also turn to reference 13, S. 446, below.

carry out the programme of logical analysis of science, they thought it necessary to distinguish the context of justification from the context of discovery, declaring that philosophy was concerned with the former and that it had nothing to do with the latter (turn to reference 4, ch. 1 below). By the context of justification was meant the context of scientific rationality where theories as finished products of the process of discovery could be tested and then evaluated for rational choice, taking into account elaborate bodies of empirical evidence. Taken in their proper context of justification or confirmation, only theories themselves could be taken as the legitimate subjects of philosophical analysis by rational reconstruction. On the other hand, all that which fell within the context of discovery was to be excluded from philosophy. By context of discovery they meant *science in context.* Think, for example, of a context in which a scientist is able to conceive of new ideas which we never heard of before. The creative act of conceiving novel ideas belonged, according to them, to the context of discovery. Being external to science as a system of knowledge, and there being no logic of discovery, the context of discovery could not, they argued, form a legitimate subject of philosophical analysis by rational reconstruction. This particular concern with *exclusion* of the context of discovery from philosophy of science, in particular from the philosphy of physics of the 20th century, shaped quite decisively their entire programme of philosophical analysis of scientific knowledge as well as scientific changes taking place in physics and other natural sciences.

Neither Heisenberg nor Einstein, among other great physicists of the 20th century, probably thought of what Reichenbach and Carnap, or the logical positivists generally, and later on Popper, were upto when they ruled out any concern with the context of discovery. After a century of opposition to the context of discovery, the question today arises whether in doing so they did not throw away the baby along with the bath water. It is noteworthy that the context of discovery in the sense of the most dominant traditions of the 20th century philosophy of science is something which lies quite outside science. While ruling it out as a possible subject of philosophical rational reconstruction, the philosophers did not even consider the possibility of the context of discovery building itself up and playing an important role *within* science itself. It is against this background that I would like to pose the following question: Whether this is not the right occasion to take another look at the Heisenberg-Einstein context principle as it emerges from their dialogue on physical theory? Whether we should not explore the broader issues concerning physical theory which are still waiting for serious philosophical attention, the context of discovery *in* physics being one of them? With their context principle "that it is the theory which decides what we can observe", Heisenberg and Einstein *reconnected*, even re-integrated, observation with theory, and the frontier of discovery in physics with the core-context of discovery, thereby restoring to physics and astronomy what they had been robbed of under their earlier doctrine. Their earlier doctrine, a kind of Machian foundationalism, required that a theory must be founded on observable magnitudes alone, that in its entire structure it must include nothing but observable quantities. To pursue our discussion further, it is very important to ask the following questions: (1) Is not the concept of *context of discovery* rather systematically ambiguous, depending upon whether we are interested in looking at science and scientific change from outside or interested in looking at them from inside? And (2) what implications follow from the Heisenberg-Einstein context principle, reconnecting observation with theory rather in a manner which is quite at variance not just with their earlier convictions or beliefs but with those of the philosophers including the logical positivists? In answering the question (1) rather in the affirmative, I want strongly to argue that the dominant idea of *science in context* cannot be allowed to blind us to the important role which the *context principles* play inside science. In particular, it cannot be allowed to blind us to context-building inside physics, constraining development of physical theory as physics makes progress over historical time. But that is exactly what has happened during the past century, thanks to logical positivists, and to Reichenbach, Carnap and Popper particularly. I think

that it is most important today to distinguish between *science in context* - what the philosophers called the context of discovery – from *the context in science*. Philosophers may choose to deal, or not to deal, with the former. But they cannot ignore the latter, namely the context of discovery *in* physics and in other natural sciences. In order to differentiate it rather sharply from the philosophical conception of the context of discovery in the 20th century tradition of Carnap, Reichenbach and Popper, I propose to designate it as the core-context of discovery (**CCD**), entailing recognition of context-building within physics as the current theories mature sufficiently enough, not only generating new frontiers of research but demanding unification of the fundamental laws of nature. Quantum mechanics provides a good example of that which I am here hinting at.

It was Pierre Duhem (1906, 1914), the French physicist-philosopher, who first pointed out that there exist contexts of theory-testing within physics, thereby drawing our attention to a *context principle* at work in physics. He was referring to those contexts where clusters of theories, or hypotheses, participate while a newly proposed theory undergoes test. Given the theory under test, the test-results are to be jointly attributed to the theories participating in the test. If the test-results are negative because the predictions of the theory in question fail, no single theory can be held responsible. Like Duhem, I want to ask the question whether there do not exist **CCsD** within physics. Whether there do not build up such contexts where theories, which are already successful and confirmed, participate in the process of discovering the kind of new theory which physics can aim at? I think that there are at least three kinds of possibilities which we must consider. *First,* in physics a **CCD** emerges when the existing theories are not only successful but acquire sufficient maturity to become frontier-generating theories. Take for example Einstein's general theory of relativity which received confirmations already in 1919. It came back on the science in 1960s, generating a whole frontier of discovery known as big-bang cosmology. Of course, the participating theories or theory must also be successful in other important respects. The general theory of relativity amply fulfils this requirement. *Secondly*, a **CCD** may emerge when the participating theories are those which provide crucial inputs to the methodology of unification in physics. They must all participate in the search for a new theory in which the earlier theories would be allowed to live contradiction free. For example, this is true of quantum mechanics and the general theory of relativity. At present they cannot live together contradiction free. The search for a quantum theory of gravity is on in fundamental physics. If a theory is found which unifies all the four forces of nature, it may be possible for the existing theories to live in it contradition free. *Thirdly,* if we reconnect observation with theory in accordance with the Heisenberg-Einstein context principle, then physics can aim at those theories which fulfil this condition rather than those which do not. Thus, to think of the **CCD** in physics is to think of those theories about which we can say at least three things: (1) that they not only explain all that which they are expected to explain but they decide what can be observed or measured; (2) that, being highly successful, they participate in the methodology of unification; and (3) that, being sufficiently rich, mature and powerful, they generate new frontiers of discovery, making physics open-ended. Taken in these three dimensions, the **CCD** in physics deserves serious philosophical attention. I think that the future of the philosophy of science will depend upon how we go about the task of understanding the **CCD** in physics and, possibly, in other sciences.

While trying to place Heisenberg-Einstein dialogue on physical theory within a broader framework, it is not at all my intention here to go into the whole context which builds itself up over historical time in the debate about quantum mechanics that was initiated by the **EPRG** as early as 1935, or into that which builds up in the Bohr-Einstein-debate and in the more recent developments of the last half of the 20th century. I think that it is quite appropriate on this unique occasion, commemorating Werner Heisenberg on his 100th birthday, if I focus on a rather neglected aspect of his philosophy of physics. Very modestly speaking, my aim here is to take another look at some of its *salient features*, particularly those which

show themselves rather clearly and distinctly in the Heisenberg-Einstein dialogue on physical theory. Their dialogue had a beginning in early 1926 in Berlin, when Einstein himself, after listening to his lecture on quantum mechanics, took the initiative for opening a discussion with young Heisenberg.[1] If we want to know who it was exactly who took the *first look* at their dialogue, we will find that it was Heisenberg himself, who has left us an account while doing so, probably the only account which deserves serious study. There arise several important questions concerning their dialogue. The first question is which particular aspect of physical theory – or rather which aspect of theory-finding, theory-testing and theory-choice – were Heisenberg and Einstein in their dialogue concerned with. The second question is whether there emerged from their dialogue some context principle, as a guiding principle of sufficient methodological interest, on which they both showed significant agreement, resolving some of their methodological differences by removing some of their shared philosophical misunderstandings concerning physical theory. And the third question is whether their dialogue helped them in arriving at a better understanding of quantum mechanics, resulting in a significant impact on their respective positions on its foundations. Whether the principle which emerged from their dialogue could be placed in a wider methodological framework within physics itself? All of these questions are philosophical questions. As regards the second and third questions, which are rather closely linked with each other, I shall argue that they can be answered in the affirmative. In particular, as regards the third question, the impact in the case of Heisenberg appears, *at first sight*, to have been far-reaching, or rather far more dramatic than in the case of Einstein. But, in the final analysis, in his case it must be seen in the light of the mathematical formalism of quantum mechanics including the uncertainty relations, while in the case of Einstein one must look at it in the light of the developments that had already accompanied the origin, the formulation and the experimental confirmation of his GTR. In particular, we must look at it in the light of Einstein's own context principle (turn to reference 7 below) which says: "Es ist das schönste Los einer physikalischen Theorie, wenn sie selbst zur Aufstellung einer umfassenden Theorie den Weg weist, in welcher sie als Grenzfall weiterlebt." As a result of their dialogue, their earlier methodological positions underwent a significant change, with both of them embracing the same guiding principle: „Erst die Theorie entscheidet darüber, was man beobachten kann."[2] Rendered into English, it says that it is always the theory which decides what can be observed.[3] On the other hand, in my response to the first question, I shall frequently refer to this very principle as the Heisenberg-Einstein context principle, their most important methodological meeting point, which must have even shaped their later maturer understanding of the foundations of quantum mechanics. Thus, while focusing on Heisenberg-Einstein context principle, I shall argue that both of them were deeply concerned with those methodological aspects of physical theory which dominate the context-building both in *theory-finding* and *theory-testing*, generating new frontiers of discovery in physics and astronomy. Any change of theory, or its further development, can take place within relevant contexts that build themselves up within physics. But this has nothing to do with that which the leading philosophers of physics of the 20th century called the context of discovery, by which they meant *science in context,* distinguishing it from the *context of justification* of complete, possibly rival, theories. In a nutshell, my present project is to understand how the Heisenberg-Einstein context principle, and the entire dialogue on which it is based, contribute to our understanding of the context-building in physics, reconnecting observation and measurement with theory.

[1] Werner Heisenberg (1984), *Gesammelte Werke*. Piper: München/Zürich. Band **II**, 1956–1968, S. 428. Herausgegeben von Walter Blum, Hans-Peter Dürr und Helmut Rechenberg.

[2] Werner Heisenberg (1979), *Quantentheorie und Philosophie: Vorlesungen und Aufsätze*. Herausgegeben von Jürgen Busche. Philipp Reclam jun. Stuttgart. S. 13.

[3] Reference 2, S. 39.

Does not a science such as physics have some kind of structural identity, besides a strategy and methodology of research and a frontier of discovery? For example, how can we identify physics or astronomy, or any other natural science, structurally or methodologically, taking into account its successes or achievements as well as its open frontiers? While alternative approaches to answering this question may not be ruled out, I propose to offer a tentative answer along the following lines. At any moment of its developmental history, it must have enough structure – enough theories and their mathematical formalisms – with its past successes built into it to be able to claim that the directions of research which its practitioners follow are not without the guiding principles which may themselves be based on or rooted in that structure. The maturity and success of its best theories will be finally judged by their power to generate new frontiers of discovery. Whenever physics is able to identify a new frontier of discovery, which is generated or activitated by its sufficiently rich fundamental theories, we can refer to the participating theories as its **CCD.** What qualifies them for this role is their ability to build up inter-theoretical relations within physics thereby setting up the **CCD**. In other words, I want to put forward the following thesis: If physics develops its own frontiers of discovery which make it susceptible to scientific change, then closely correlated with them there must exist **CCsD** in physics. For the very existence of a frontier of discovery depends upon those successes and achievements which must be held constant in the search for new knowledge. The **CCD** in physics holds constant precisely those achievements which generate the new frontiers of discovery. Whenever there is a breakthrough at the frontier, say by inventing an altogether new theory, the fundamental past achievements should find a place in the new theory.

In order to put forward here my main claims, I shall proceed on the following assumption while relating it to the Heisenberg-Einstein context principle. From time to time, there develop from the very core of a physical theory – from its mathematical formalism – the contexts which can function as the *dynamic core-contexts of discovery* (**DCCsD**), taking *it* to the very frontier of discovery – die vorderste Front der Forschung. As we have seen, my choice of this designation for such specific contexts derives from several considerations, the most important among them being the methodology of *working forwards from physical theories to problems* whose solutions are not yet in sight.[4] In a nutshell, this is the methodology which requires that a good physical theory T must possess resolving power – **Trp**(T) – besides its explanatory power – **Tep**(T).[5] The *world* which the **Tep**(T) is by its very design intended to explain can be, for good reasons, approached as an object of scientific interrogation. Nature explained can at the same time continue to be nature interrogated as to its unknown and undiscovered aspects. The specific *contexts* which develop *within* physics for this purpose are determined by the **Trp**(T) and not by any of those factors that might be *external* to the scientific activity of world-making by explanatory theory-forming, theory-testing and theory-choice. As is quite possible, my enterprise here may sound not only ambitious but rather unconventional – one which is far removed from the mainstream philosophy of science. Should this be the case, I will be more than contented by staying away from all that which can be described as the scandal of the 20th century philosophy of science. I am here referring to the approach to science chosen from outside science, which led the eminent physicist Steven Weinberg to devote a whole chapter in his book *The Dreams of a Final Theory*[6] to the theme "Against Philosophy", contrasting the "unreasonable effectiveness of mathematics" with the "unreasonable ineffectiveness of philosophy" in the development of physics. The practice of the 20th century philosopher of science can be

[4] G. L. Pandit (1983), *The Structure and Growth of Scientific Knowledge: A Study in the Methodology of Epistemic Appraisal*. D. Reidl: Dordrecht. (**Boston Studies in the Philosophy of Science** 73).

[5] G. L. Pandit (1991), *Methodological Variance: Essays in Epistemological Ontology and the Methodology of Science*. Kluwer Academic Publishers: Dordrecht. (**Boston Studies in the Philosophy of Science** 131).

[6] Steven Weinberg (1992), *Dreams of a Final Theory.* Pantheon Books: New York.

compared with the practice of someone whose beliefs do not fit into the standard practice of the physicists and astrophysicists, as if one could do astrophysics by looking at stars as finished products, and not by looking at star-formation, star-birth, star-explosion, supernova-explosion and so on. In any case, I intend here to restrict myself to an enterprise which was already being hinted at as early as 1916 by Albert Einstein when he said: "Es ist das schönste Los einer physikalischen Theorie, wenn sie selbst zur Aufstellung einer umfassenden Theorie den Weg weist, in welcher sie als Grenzfall weiterlebt."[7] If a physical theory is at all to play this role in physics and astronomy, then it is most appropriate to study the *dynamic core-context building* that generates new frontiers of physical discovery.

2. The Quantum-Context-Building

The problem of understanding the role which inter-theoretical relations play in context-building in physics and in generating new frontiers of discovery presents itself as a challenging task for the philosophy of physics. In order to have a closer look at the core-context building in current fundamental physical theory, let us ask, with Eugene Wigner, 'How close are our present theories to perfection? ... are they, in particular our physics, even self-consistent? The answer to this question is rarely publicized, but it surely is: no. The general theory of relativity is based on the assumption of the meaningfulness of the space-time point concept defined by the crossing of two-world-lines, yet it is easy to show that quantum mechanics does not permit the definition of such points. The basic idea of the fundamental interpretation of quantum mechanics postulates the process of "measurement". Yet it is easy to show that the existence of this process is not consistent with the principles of quantum mechanics.[8] Wigner argues further: "Quantum mechanics is not in complete harmony with the theories of relativity, particularly not the general one. And even quantum mechanics alone still lacks the complete simplicity which we are striving for – in spite of the accomplishments of the past, particularly those of Salam and Weinberg, we still have several types of interactions, not united into a single equation, and there are other grave problems."[9] His following observations on the developments within physics are rather directly concerned with the subject of inter-theoretic relations: "The development of science, in particular physics, is miraculous for another reason: every step in its development shows that the preceding theory was valid only approximately, and valid approximately only under certain conditions. Newton's theory is valid with a high accuracy if only gravitational forces play a role, all pre-quantum theories are valid only for macroscopic bodies and for these only under certain conditions. It is lucky that such special conditions exist under which simplified approximate theories present a wonderfully good approximation. Surely, if macroscopic theories had not been developed, it would have been even more difficult, perhaps impossible, to develop quantum mechanics. General relativity would not have been invented, not even by Einstein, had the original theory of gravitation not existed."[10]

What are then the most important, or at least distinguishing, features of contemporary physics at the frontier of discovery? Does it in any way indicate that *revolutions* in physics are driven by observations and experiments alone? Can we not situate *them* in the context of a set of *core-theories* and *core-problems* which, together with the experimental situation,

[7] Albert Einstein (1919), *Über die spezielle und die allgemeine Relativitätstheorie.* Vierte Auflage: Fried. Vieweg & Sohn, Braunschweig, Germany, S. 52. (First German ed. 1916). See also Albert Einstein, 1960: *Theory of Relativity, Special and General*, p. 77.

[8] Eugene Wigner, "The Limitations of the Validity of Present-Day Physics" (pp. 118–133), in Richard Q. Elvee (ed.) *Mind in Nature: Nobel Conference* **XVII**. Harper & Row Publishers, San Francisco, 1982, p. 124.

[9] Reference 8, p. 125.

[10] Reference 8, 124.

could define physics at the frontier of discovery? What is it which decides what kind of fundamental theory can physics at the frontier of discovery aim at? I think that at any particular stage of scie-ntific progress the context in which revolutions in physics take place can be identified with the set of *core-theories* and *core-problems*, making physics open-ended at the frontier of discovery. Taking these as constituting the *theory-problem interactive systems*,[11] we may designate them as *dynamic core-contexts of discovery* in physics. At the end of the 19th century, physicists generally looked at (Newtonian) physics as complete, thinking that the task that still remained was a matter of filling in detail only. But at the beginning of the 20th century, it became clear that the mature theories of Maxwell and Newton played rather the role of the context of discovery for many new developments in physics, notably Einstein's special and general theory of relativity and the quantum theory. Again, think of Maxwell's unification of electric and magnetic forces in 1860's, predicting that light is an electromagnetic phenomenon. In the 1920s, Maxwell's theory and Einstein's general theory of relativity, with all its novel predictions and their confirmations, formed the **CCD** for attempts at further unifications, including Einstein's own unsuccessful search for a unified field theory. On the other hand, it is a well-known fact that the core-problem of black-body radiation shaped physics at the frontier of discovery from 1859 until 1926, not only in thermodynamics, but in electro-magnetism, in the old quantum theory, and in quantum statistics, in that order.[12] Thus, in the kind of development of fundamental physical theory which physics at the frontier of discovery can aim at, what is most important is how the **CCD** builds itsef up in *theory-testing* and *theory-finding*. Schematically, the following interplaying factors and levels of interaction are dinstinguishable here (block diagram 1):

Block diagram 1. A schematic representation of the core-context building and the frontier of discovery in fundamental physical theory

consistent background theories (e.g., successful unification of disparate phenomena within one theory) →
theory-testing, in which clusters of theories participate,
for example, the ongoing tests for changes in the values of the fundamental constants with time,
or for the detection of newly predicted fundamental particles,
or for establishing the mechanisms for generation of the masses of such particles, using the latest technology,
(e.g., the costliest high energy particle collider experiments that are planned at the Large Hadron Collider under construction at CERN) with lot of successful theory built into it →
sets of frontier-generating core-theories and core-problems →
core-context of discovery, as constituted by the core-theories and core-problems, in which the unification of successful theories, within a yet *new* theory, can be aimed at →
search-and-discovery procedures for finding the *path* to the *new* fundamental theory which physics aims at.

How strongly do theory-testing and context-building interplay with each other in shaping the frontier of discovery in physics? While *theory-testing* by checking a candidate theory

[11] Reference 4 and 5.

[12] See A. Pais (1982), *Subtle is the Lord: The Science and Life of Albert Einstein.* Oxford University Press: New York. p. 28.

by novel experiments can result either in its confirmation or in its falsification, *context-building* can develop the **CCD** from the very core of the existing theory, or theories, indicating the path to formulating the new theory which physics aims at. The stronger and the richer the confirmations resulting from *theory-testing*, the greater will be the maturity of the theory under test. The more mature a theory becomes by testing its consequences, the greater will be its ability to develop into a frontier generating **CCD.** Today, at the end of the 20th century, quantum mechanics itself serves as the most mature physical theory in this sense, where experimental confirmations of the theory are concerned. With Heisenberg, we can say that "The history of physics in our century was characterized by the interplay of two extra-ordinary developments those in experimental and in theoretical physics, independent with regard to the technique and the activity of the participating scientists, dependent of each other by a continuous dialogue. The experimental development has penetrated into quite new fields and has produced a great number of new and unexpected phenomena and last but not least most important practical applications, the theoretical development has given a new meaning to such old and fundamental concepts as space, time, state, smallest unit of matter. I am afraid that it might take another century, before one has become really well acquainted with all this new scientific material and its practical, political, ethical-philosophical consequences. But this will be the task for the younger generation."[13] The *quantum context-building* has already taken place. Now the most important task is to ask whether we can develop it further into a *dynamic core-context of discovery.*

Heisenberg's uncertainty principle states that the values of both the members of certain pairs of canonically conjugate variables, such as position and momentum, cannot be determined simultaneously to arbitrary precision. Heisenberg's uncertainty relations

$$\Delta q \cdot \Delta p \geq \hbar/2$$

give a mathematical expression to this irreducible level of uncertainty of the relevant pairs of dynamical variables when they are measured together. Thus, there is a reciprocal relationship between them, such that the product of their uncertainties must be greater than Planck's constant *h.* A universal statement of prohibition at the quantum level, this has to be understood as an important consequence of the formalism of quantum mechanics itself. As an intermediate step in the development of fundamental physical theory, consider Dirac's unification of quantum mechanics with the special theory of relativity and electrodynamics in the 1920s, resulting in the relativistic quantum mechanics of the electron. This belongs rather to the period of development of physical theory when most physicists were preoccupied with the fundamental difficulties of understanding the stability of atoms, the way electrons interact within an atom and the non-classical properties like the spin of a sub-atomic particle. Among other physicists, including Schrödinger, who clearly recognized the problem-situation developing within physics, it was Dirac who succeeded in finding a solution to the *core-contextual problem* of the existence of the negative energy states for the electron. It is necessary to ask what kind of *context-building* was at work here? For a physicist, like Dirac, wanting to write down a complete equation for the electron, there was a context-building from the discovery of the quantized angular momentum of the electron by O. Stern and W. Gerlach in 1922 and from the very core(s) of the special theory of relativity on the one hand and quantum mechanics on the other. Dirac knew that anyone grappling with the problem of the existence of negative energy states would have to work within this very context. The kind of theory-testing and context-building, which I am here referring to, got accelerated with the development of quantum mechanics, once it was already known since 1905 that "Einstein's formula for the energy of a system with a given momentum involves

[13] Werner Heisenberg (1985), *Gesammelte Werke* **III, 1969–1976,** Herausgegeben von Walter Blum, Hans-Peter Dürr und Helmut Rechenberg. Piper: München/Zürich 1985, S. 459–460.

a square root, and the result is that the value for the energy, mathematically, can be either positive or negative."[14] With such core-context-building, it was no longer possible to disregard the problem of the negative energy states. Quantum mechanics allows jumps discontinuously from one energy level to another. In the words of Dirac, "If we start off a particle in a positive energy state, it may jump into a negative energy state."[15] If we cannot exclude the negative energy states from our theory, Dirac thought, we must find a method of physical interpretation for them.[16] It was in May 1931 that Dirac proposed that there exists a new kind of particle "unknown to experimental physics, having the same mass and opposite charge to an electron. We may call such a particle an anti-electron."[17] Published in the *Proceedings of the Royal Society* in September of that year, it announced the birth of the modern idea of antimatter.[18] The far-reaching consequences of his theory's novel prediction (or requirement) that anti-electrons exist was confirmed in 1932 by the experimental discovery of the positron by C. D. Anderson.[19] Today we can look back to Dirac's relativistic quantum theory of the electron as a theory inseparable from *context-building* in physics at the frontier of discovery where the high energy collision experiments with particle creation (transmutation of energy into matter in accordance with the special theory of relativity) play quite a dominant role. Here it would be quite wrong to think of the concept of a fundamental field – or rather of a fundamental elementary particle – as if it was no longer well-defined simply because particles created by particle collisions, e.g., in electron-positron collusion experiments, have only a virtual existence. On the contrary, in fundamental physical theory, symmetries and fundamental fields are equally important.

I consider it an important task of the physicist and the philosopher of physics to find out how, from time to time, there develop **CCsD** within physics in which the physicist is constrained to make his/her choice between the possible paths for finding a new theory. Thus, here I am not at all concerned with *physics in context.* My sole concern is rather *the context within physics*. Normally, a physical theory itself matures under experimental testing until it itself becomes a **DCCD**, raising new problems and generating new frontiers. Dynamic core-contexts of discovery can build themselves up within physics even from the core of those fundamental theories which have, after acquiring maturity, become controversial either because of their mutual consistency being in question or the very path to their unification being a highly problematic and thorny one. For example, consider the fact that the GTR has still not been successfully incorporated into a consistent quantum mechanics. The physicists agree that quantum gravity is one of the great frontiers of discovery at the end of the 20th century. There may be alternative possibilities of approaching the current problem of formulating a quantum theory of gravity. But it is quite important to ask in this kind of situation what is the dynamic core-context of discovery for finding the path that leads to such a theory. What are the core-problems which contribute to the *context-building* for the physicist's search for such a theory? If we only think of the problem of understanding how the fundamental particles acquire their masses (the problem of the origin of mass)[20], or of the problem of the cosmological constant as the most serious of all pro-

[14] Paul A. M. Dirac (1983), "The origin of quantum field theory" (pp. 39–55), in Laurie. M. Brown and Lillian Hoddeson (eds.) *The Birth of Particle Physics*. Cambridge: Cambridge University Press, 1983, p. 51.

[15] Reference 14, p. 50.

[16] Paul A. M. Dirac (1978), "The Development of Quantum mechanics", in H. Hora and J. R. Shepanski (eds.) *Directions in Physics* (lectures delivered by P. A. M. Dirac August/September 1975). New York: John Wiley & Sons, 1978, p. 16.

[17] Reference 16.

[18] See Gordon Fraser (2000), *Antimatter: The Ultimate Mirror.* Cambridge: Cambridge University Press, p. 62.

[19] The similar anti-particle on the nuclear level, the antiproton, was discovered by Chamberlain, Segre, Wiegand, and Ypsilantis in 1955.

blems in particle physics and cosmology, or of the problem of under-standing the beginning of the universe and its present rate of expansion (including the dark matter problem — why does most of the matter that gravitates in the universe seem to be invisible?), a quantum theory of gravity seems to be the kind of theory which physics at the frontier of discovery can aim at.

3. The Methodology of Unification in Physics

The current physical theories always build up an implicate order of inter-theoretic clustering, partly harmonious, partly full of tension. The inter-theoretical relations which they build up activate the search for a new theory **T** not yet in sight, while restricting, at the same time, the choices available to physics at the frontier of discovery. Over historical time, they develop within physics what I have chosen to designate as **DCCsD**, which enable it then to find its path to **T**. I think that the best examples of **DCCsD** come from the history and the methodology of unification in physics, with a long record of success, from Maxwell's equations for electromagnetism to the electroweak theory, unifying the electromagnetic force with the weak force, and from the electroweak theory and Quantum Chromodynamics (QCD) to the grand unified theories (GUT). The current standard model of the elementary particle physics, a gauge theory of the strong and electroweak interactions, provides a fundamental theory of quarks and leptons, which has been tested up to energies approaching 1000 GeV. The important question about particle physics at very short distances, which pre-occupies the physicists at the frontier of discovery, is how one might go further from the standard model to its *possible* extension to supersymmetry, or superstring theory, or grand unification, making a more comprehensive understanding of the laws of particle physics and, therefore, of the *earliest* history of the universe possible. The methodology works by building up enough symmetry into the fundamental laws of nature as and when these are formulated, so that the underlying fundamental interactions and their force laws are unified, resulting in their description by a single unified theory. Interestingly enough, this methodology has itself evolved over historical time and acquired considerable complexity, possibly with a feedback from fundamental physical theory itself. For example, the 1960s saw a kind of shifting away from the earlier insistance on a reductionist research programme of unification by an ultimate explanation of nature in terms of a final theory of the simplest elementary particles, rather in the old tradition of Democritus. But the programme staged a comeback in the 1970s mainly through the development of the electroweak theory and the QCD, with the emphasis shifting to the local quantum field theories in which both the fundamental fields — or their particle representations — and fundamental symmetries interplay rather intimately. And this is the present scenario. As physics has moved from one frontier to another in search of the unifying theory, **Tu**, its **DCCsD** have exhibited a fundamental continuity in terms of the mathematical formalisms, from one successful unification to the other. On the other hand, they also have exhibited a strong undercurrent of continuity in the following equally non-trivial sense: As they have generated new frontiers of discovery, the older problems have been seen in a new light, as also the older theories which had either themselves generated those problems or provided their solutions. This type of continuity can be best described in terms of the earlier stable forms of knowledge being *reproblematized* by the **DCCsD**.

[20] The **SD**-procedures in this context have as their target the predicted Higgs boson, the electrically neutral spinless particle which is the most elusive object in the Standard Model. It is currently believed that the high energy and precision of an electron-positron linear collider holds an important key to the next step in a comprehensive understanding of the laws of particle physics.

4. Theory and Observation

What is it which makes physics an empirical science? Those who look at it from outside generally believe that in physics, "concepts and mathematical con-structs can simply be taken from experience" (i.e., from empirically established data).[21] Did either Einstein or Heisenberg hold this kind of view at any time during their life-time? Heisenberg had himself once argued as follows: "If this was the whole truth, when entering into a new field, we should introduce only such quantities that can directly be observed and formulate natural laws only by means of these quantities."[22] As a young man, recalls Heisenberg, he also believed "that this was just the philosophy which Einstein followed in his theory of relativity."[23] He is here referring to the special theory of relativity. What is still more interesting is that Heisenberg first thought, quite consistent with his belief, as if he too was following the same kind of philosophy – the philosophy of introducing only observable quantities in one's theory. Sooner or later one must ask how far is such a requirement appropriate for a physical theory. How far is it possible for a theory to be constructed solely in terms of quantities which can directly be observed, or which are closely related to such observable quantities? If you put the theory you construct in chains and expect it to save the phenomena, it reminds us more of Ernst Mach's overly restrictive phenomenalistic empiricism, rejecting atomism in physics. It even reminds us of logical positivism of the Wiener Kreis, demanding, in the tradition of the classical British empiricism and Ernst Mach's phenomenalistic empiricism, that a physical theory be verifiable in principle, as if the methodological rules developing from within the twentieth century physics were of no interest to us. It is remarkable that today we can say with certainty how ineffective the postulate of observability has been as it did not play any significant role in context-building in the major developments in physics during the past century. Had it done so, the development of physical theory would have been crippled rather from the very early stage. But the important question still remains: Did Einstein and Heisenberg have to change their views on the aim and structure of physical theory and on its connection to observation, or measurement? I consider it as the biggest drawback of the strategy of raising the observability of all quantities, e.g., energy, frequency, and so on, in one's equations to a postulate, as the young Heisenberg and, before him, Einstein and others did under the influence of the Mach-type phenomenalistic empiricism, that it encourages one to do physics in a style as if saving the phenomena were the sole aim of a physical theory. Since both of them rejected later on the view that the postulate had played a role in the construction of physical theory, what were the reasons possibly developing from within physics itself which led to such a development? To find an answer, we have to go back to early 1926, when under the auspices of the Physics Kolloquium of the University of Berlin, and in the tradition of Hermann von Helmholtz, Heisenberg was invited to report on the subject of the newly developed quantum mechanics. After the Kolloquium, which was attended by Einstein, among others, he and Heisenberg met at former's Berlin flat where an interesting discussion followed. Quite naturally, the discussion was started by Einstein with a question as follows: "What was the philosophy underlying your kind of very strange theory? The theory looks quite nice, but what did you mean by only observable quantities?"[24] Somewhat more specifically, he asked him „Was Sie uns da erzählt haben, klingt ja sehr ungewöhnlich. Sie nehmen an, dass es Elektronen im Atom gibt, und darin werden Sie sicher recht haben. Aber die Bahnen der Elektronen im Atom, die wollen Sie ganz abschaffen, obwohl man doch die Bahnen der

[21] Reference 13, S. 444.
[22] Reference 2, S. 13.
[23] Reference 2.
[24] Reference 1, S. 428. Here Heisenberg (S. 428) remarks: "But when I had to give a talk about quantum mechanics in Berlin in 1926, Einstein listened to the talk and corrected this view."

Elektronen in einer Nebelkammer unmittelbar sehen kann. Können Sie mir die Gründe für diese merkwürdigen Annahmen etwas genauer erklären?"[25] Recalling his own reply to this question, Heisenberg admitts that, "I did not believe any more in electronic orbits, in spite of the tracks in a cloud chamber. I felt that one should go back to those quantities which really can be observed and I also felt that this was just the kind of philosophy which he had used in relativity; because he also had abandoned absolute time and introduced only the time of the special coordiinate system and so on."[26] To this Einstein had responded by saying "That even if this may have been his philosophy and even if he may have used it, it was nonsense all the same, which meant that it was completely wrong."[27] The most significant remark which Einstein then made was that "It is never possible to introduce only observable quantities in a theory. It is the theory which decides what can be observed."[28]

5. The Heisenberg-Einstein Context Principle: the Theory Dominates

In the epistemology and methodology of the physicist's search for physical theory, if we want to single out at least one methodological principle of foundational importance to physics which has dominated the discussions between Heisenberg and Einstein, it is the one which Heisenberg has himself frequently focused on in his essays and lectures, almost invariably attributing it to Einstein during their intimate discussions on quantum mechanics, particularly on Heisenberg's uncertainty relations. The main aim of their discussions was to resolve their differences in order to improve their respective positions concerning the interpretation of quantum mechanics. The methodological principle (or the Heisenberg-Einstein context principle) as quoted in Reference 2, which they both came to agree upon, says: "Erst die Theorie entscheidet darüber, was man beobachten kann . . ." In the case of Heisenberg, one notices in the light of this principle that there is an interesting, a non-trivially significant, turning around of the fun-damental questions in quantum mechanics with far-reaching consequences for our understand-ing of its interpretational problems. What makes it all the more significant is how Heisenberg has himself acknowledged this. In his many

[25] Reference 2, S. 30. Here Heisenberg (S. 30) recalls Einstein's deeper concern when the latter asked him to clarify: „Aber Sie glauben doch nicht im Ernst, daß man in eine physikalische Theorie nur beobachtbare Größen aufnehmen kann." Also turn to reference 2, S. 91.

[26] Reference 1.

[27] Reference 13.

[28] Reference 13. For further details of their dialogue, under the same reference, see Werner Heisenberg, "Die Quantenmechanik und ein Gespräch mit Einstein", S. 85-100. Also turn to reference 2, S. 22–41. See also Werner Heisenberg (1969), *Der Teil und das Ganze: Gespräche im Umkreis der Atomphysik*. R. Piper & Co. Verlag: München. PP. 85–101. Evidently, as the Heisenberg-Einstein dialogue shows, both of them were far ahead of the philosophers of science lead by Rudolf Carnap and Moritz Schlick of the Wiener Kreis, during the 1920s and 1930s, as regards their abandoning of the philosophy that one should introduce only observable quantities in a theory. The philosophy of science communities world-wide have rather taken relatively much longer time to realise the futility of this philosophy and to turn around the questions concerning the relationship between theory and observation. One of the reasons why generally no attention has been paid by them to the frontier of discovery in physics is their rigid belief in analytic philosophy as a second-order discipline which can legitimately analyse only the logical structure of a theory once it is there as a finished product of scientific creativity. The security and stability of this belief is guaranteed by the distinction between the external context of discovery (= science in context) and the context of justification of theories which the logical positivists and other schools of philosophy drew at the very dawn of analytic philosophy in the 20th century. The strategic advantage of this distinction is believed to lie in the legitimacy it bestows upon analytic philosophy of science conceived as an enterprise which analyses a scientific theory in its logical relationship to the evidences in favour or against it. The context of scientific discovery conceived as science in context is none of the philosopher's concerns in the analytic tradition.

lectures and essays, he returns to this theme while clarifying the nature of the development of concepts in the physics of the 20th century, particul-arly in the fields of quantum mechanics and elementary particle theory.[29] In my opinion, Heisenberg's account of his dialogue with Einstein deserves a serious attention, at least for the following reasons. First, I think that it carries within itself a great methodological insight, provided we are interested in understanding what exactly the Heisenberg-Einstein context principle, as a guiding principle, says or implies in the context of the physicist's search for fundamental physical theory. If it is the theory which decides what can be observed, or measured, then it is methodologically imperative for the physicist working at the frontier of discovery to ask (i) what are the theories in the background which serve as, or provide, the core-context of discovery and (ii) what kind of fundamental theory can physics aim at? Thus, with the help of this principle, one can explore important correlations that may obtain between the *aim* and *structure* of a physical theory, or between the frontier of discovery and the **CCD**. Secondly, Heisenberg's (attitude to his) own discovery of the uncertainty relations, no less than the discoveries made earlier by Einstein, demonstrates the methodological importance of this principle as a guiding principle in physics. Thirdly, as regards Einstein, given this principle, the deeper epistemological and methodological reasons for his view that quantum mechanics is incomplete become rather clearly visible and accessible to rational reconstruction along the following lines of argument: Even before the physicists decide to legitimately subject the *extant* theories to severe criticism by experimental test, and by serious epistemic appraisal, they had better decide *first* what kind of fundamental theory should physics aim at? Given the kind of theory which physics should aim at, they can then decide what kind of methodology, not only of appraising but of finding a theory, should it develop and adopt. Both these steps seem essential to the building up of the *core-context of discovery* within physics for a better understanding of the *extant* theories and the role they themselves can play for seeking better alternatives to them, where possible or necessary. Thus, the very rationality of Einstein's radical opposition to (Copenhagen interpretation of) quantum mechanics, and to the general acceptance of the theory as if it were a final theory, could be explained with the help of this principle, whether or not it is taken together with the EPRG/ Bohr-Einstein debate. Fourthly, in my opinion, no rational reconstruction of Einstein's own research programme would be complete, if it did not find a fundamental role for this principle. Einstein's belief in this principle seems to have been so firm, as if it was that part of the Galilean methodology which a physicist could not easily give up. Fifthly, the real but largely unacknowledged impact of the Heisenberg-Einstein context principle on the developments in the 20th century physics and philosophy have yet to be properly assessed. Sixthly, and lastly, one might ask what was the exact nature of the impact of Heisenberg's Berlin lecture on Einstein? How far-reaching were the consequences of their subsequent discussion for Einstein himself and for Heisenberg, or for Bohr? Did their views stand corrected in some sense by this methodological principle which originally emerged from Einstein's remark? If we follow the details of their dialogue as reconstructed by Heisenberg, the answer is clearly that it made a most significant contribution by making an improved understanding of quantum mechanics possible, where Einstein, Heisenberg and Bohr were concerned. In Heisenberg's own words: "This remark of Einstein was very important for me later on when Bohr and I tried to discuss the interpretation of quantum theory . . ."[30] The task of clarifying its nature and implications assumed a great urgency for Heisenberg, since "The most conspicuous demonstration of this thesis by Einstein was the relations of uncertainty."[31] The question of the nature of the methodological principle expressed by Einstein's remark assumes importance when considered in the context of Heisen-

[29] Reference 2, S. 13.
[30] Reference 1, S. 429.
[31] Reference 13, S. 446.

berg's acceptance of its role in the interpretation of quantum mechanics. This question will be taken up for discussion in the next 6.

6. Turning Around of the Questions in Understanding Quantum Mechanics

Before considering how the Heisenberg-Einstein context principle, as it emerged from their dialogue, led to a change of view and to a change in attitude – to an attitude of turning those questions around which one usually asked those days in physics, especially in the context of quantum mechanics – let us pay some attention to Einstein's argument for his view that "whether you can observe a thing or not depends on the theory which you use. It is the theory which decides what can be observed."[32] As reconstructed by Heisenberg, the argument runs as follows: "Observation means that we construct some connection between a phenomenon and our realization of the phenomenon. There is something happening in the atom, the light is emitted, the light hits the photographic plate, we see the photographic plate and so on and so on. In this whole course of events between the atom and your eye and your consciousness you must assume that everything works as in the old physics. If you would change the theory concerning this sequence of events then of course the observation would be altered."[33]

Did Einstein's view that it is really dangerous in physics to say that one should only speak about observable quantities, which Heisenberg also now shared, have damaging conse-quences for those philosophies – recall Ernst Mach's attitude to the concept of atom – which are notorious not just for their instrumentalistic character but for their ineffectiveness in the development of science? The answer is clearly in the affirmative. Given the postulate of observability concerning how to structure a physical theory, what one could expect a theory to do was just to save the phenomena. Moreover, if Bohr and Heisenberg had held the view that the equations of quantum mechanics are a mathematical tool of calculating the probabilities for the various outcomes of measurements,[34] this did not necessarily imply that they were favouring instrumentalism in a rather philosophical sense, as some philosophers have suggested by way of criticism. I think that their view is not incompatible with Heisenberg-Einstein context principle which asserts that it is always the theory which decides what can be observed. On the contrary, it can very well serve as a demonstration of the principle when extended to the context of the interpretation of quantum mechanics. According to Heisenberg, "In quantum theory it meant, for instance, that when you have quantum mechanics then you cannot only observe frequencies and amplitudes, but for instance, also probability amplitudes, probability waves and so on, and these, of course, are quite different objects."[35] But when a change of theory results in change in observations, what happens to the pre-existing concepts of a pre-existing theory? This is precisely

[32] Reference 1, S. 429.
[33] Reference 1, S. 429.
[34] Reference 1, S. 429.
[35] Reference 1, S. 429. In an interesting statement regarding the role of the state vector, Eugene Paul Wigner (1971: 5–6) says: "There are two epistemological attitudes toward this. The first attitude considers the state vector to represent reality, the second attitude regards it to be a mathematical tool to be used to calculate the probabilities for various possible outcomes of observations. It is not easy to give an operational meaning to the difference of opinion which is involved because, fundamentally, the realities of objects and concepts are ill defined. One can adopt the compromise attitude according to which there is a reality to objects but quantum mechanics is not concerned therewith. It only furnishes the probabilities for the various possible outcomes of observations or measurements – in quantum mechanics these two words are used synonymously." See Eugene Paul Wigner (1971) "The Subject of Our Discussions", B. d'Espagnat (ed.) *Foundations of Quantum Mechanics*. International School of Physics "Enrico Fermi" 1970. Academic Press, New York, 1971, pp. 1–9.

the kind of question which Heisenberg himself asks as follows: "... when one has invented a new scheme which concerns certain observable quantities, then of course, the decisive question is: which of the old concepts can you really abandon? In the case of quantum theory it was more or less clear that you could abandon the idea of an electronic orbit."[36)]

According to Heisenberg, what Einstein must have meant by his remark is "that when we go from the immediate observation – a black line on a photographic plate or a discharge in a counter – to the phenomena we are interested in, we must make use of theory and of theoretical concepts. We cannot separate the empirical process of observation from the mathematical construct and its concepts."[37)] Interestingly enough, Heisenberg's own reformulation of this idea runs as follows: "In order to understand Nature we have to approach it by some concepts; and we try to establish an immediate connection between the observed phenomena and the concepts. If we are successful, we have both defined what we have observed and confirmed the validity of the concepts; if not, we may be forced to change the conceptual frame."[38)]

The question which one might ask here is whether an alternative way of looking at the Heisenberg-Einstein context principle is possible, which might bring us closer to understanding quantum mechanics, the uncertainty relations in particular, Einstein's attitude to it and the attitude of Heisenberg himself. In order to answer this question, we must go back to the year 1926 when the interpretational problems of quantum mechanics were still tormenting Heisenberg in Copenhagen. "... we felt", to quote Heisenberg, "that in the atom it seemed all right to abandon the concept of an electronic orbit. But what in a cloud chamber? In a cloud chamber you see the electron moving along the track; is this an electronic orbit or not?"[39)] How many nights must have he and Bohr spent discussing these problems, as Heisenberg recalls, with Bohr emphasizing the dominant role of the wave-particle dualism while he himself thought of the mathematical formalism as his starting point in search of a consistent interpretation.[40)] It was around this time that he remembered Einstein's remark that "It is the theory which decides what can be observed."[41)] There was, as a result, such a change in the whole approach, such a change in the very understanding of the problems, that everything fell in place. Let me quote Heisenberg as he describes this change in the following words: "From there it was easy to turn around our question and not to ask: 'How can I represent in quantum mechanics this orbit of an electron in a cloud chamber?' but rather to ask 'Is it not true that always only such situations occur in nature, even in a cloud chamber, which can be described by the mathematical formalism of quantum mechanics?' By turning around I had to investigate what can be described in this formalism; and then it was very easily seen, especially when one used the new mathematical discoveries of Dirac and Jordan about transformation theory, that one could not describe at the same time the exact position and the exact velocity of an electron; one had these uncertainty relations. In this way things became clear. When Bohr returned to Copenhagen, he had found an equivalent interpretation with his concept of complementarity, so finally we all agreed that now we had understood quantum theory."[42)] Turning around the questions which one usually asked, Heisenberg posed the question: "Well, if we want to know of a wave packet both its velocity and its position what is the best accuracy we can obtain, starting from the principle that only such situations are found in nature which can be represented in the mathematical scheme of quantum mechanics"[43)] Posing such type of questions

36) Reference 1, S. 429.
37) Reference 1, S. 429.
38) Reference 13.
39) Reference 13, S. 453.
40) Reference 1, S. 433.
41) Reference 1, S. 433.
42) Reference 1, S. 433.
43) Reference 1, S. 433.

resulted in the principle of uncertainty, which seemed to be compatible with the experimental situation — e.g., the fact that "the path of an electron in a cloud chamber was not an infinitely thin line with well-defined positions and velocities."[44] According to this principle,"the wave packet representing the electron is changed at every point of observation, that is at every water droplet in the cloud chamber. At every point we get new information about the state of the electron; therefore we have to replace the original wave packet by a new one, representing this new information."[45]

Let us now briefly consider where this kind of theory and this kind of questioning can lead us. Some of the consequences are described by Heisenberg as follows: "The state of the electron thus represented does not allow us to ascribe to the electron in its orbit definite properties like coordinates, momentum and so on. What we can do is only to speak about the probability to find, under suitable experimental conditions, the electron at a certain point or to find a certain value for its velocity. So finally we have come to a definition of state which is much more abstract than the original electronic orbit. Mathematically we describe it by a vector in Hilbert space."[46] The Hilbert space being a space of infinitely many dimensions, the concept of state here diverges from that in classical physics. The divergence does not yet mean any major departure from the method in physics as formulated by Galileo. We can still look at nature as the *object* of physical knowledge. Nature can still be described, or known, objectively by means of the mathematically formulated theories.

7. If it is Always the Theory which Decides what can be Observed, then what Kind of Fundamental Theory can Physics Aim at?

The most important question which has remained unasked in the context of the interpretational problems of quantum mechanics is how far did the Heisenberg-Einstein context principle finally shape Einstein's critical and radical attitude in that very context. During the period 1925—1931, his attitude was dominated by the following question: Is quantum mechanics a consistent theory?[47] The first time he is reported to have written approvingly about quantum mechanics was when he wrote in May 1926 to Schrödinger about those advances the latter had made in this field.[48] For many years, Einstein believed that the theory contained logical contradictions. And he must have done so mainly in the context of Heisenberg's uncertainty relations. By 1933 he had already given up this position, openly recognizing that the theory was a logically consistent theory.[49] By 1935, there was a turning around of the question concerning it resulting in the **EPRG** of 1935.[50] The question which has ever since dominated the foundational debate about the theory is the question which the authors of the **EPRG** raised in the very title of their paper: "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?"[51]

It can be very well argued that only someone who was already guided by his general conception of what makes a physical theory a really good theory — as being consistent, complete, and as possessing other virtues — could have chosen to make a new theory such as quantum mechanics the subject of a relentless critical debate. The chief advantages of the strategy

[44] Reference 13, S. 410-411.
[45] Reference 13, S. 410-411.
[46] Reference 13, S. 411.
[47] Reference 13, S. 411.
[48] See A. Pais (1982), *Subtle is the Lord: The Science and Life of Albert Einstein*. Oxford University Press: New York. P. 440.
[49] Reference 48, p. 442.
[50] Reference 48, p. 449.
[51] A. Einstein, B. Podolsky, and N. Rosen (1935), *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, Physical Review **47** (777—780).

followed by Einstein in his debates with Bohr, and others, seem then precisely to lie in those areas where he could simultaneously test his own general conception of a physical theory and his understanding of quantum mechanics, one against the other. In this sense, then, one might very well conclude, the whole debate on quantum mechanics as initiated by him, right from 1925 through 1935 to the end of his life, is a great dialogue on physical theory. In the context of a rather revolutionary scientific change, which quantum mechanics brought about, it raises the same question again and again: what kind of fundamental theory can physics aim at? And what kind of critical appraisal is appropriate in such a context?

It is here quite interesting to note how Einstein stated his points of agreement and disagreement on the interpretation of quantum mechanics in one of the last discussions Heisenberg had with him. In Heisenberg's own words: "I had a discussion with Einstein about this problem in 1954, a few months before his death. It was a very nice afternoon that I spent with Einstein but still when it came to the interpretation of quantum mechanics I could not convince him and he could not convince me. He always said 'well, I agree that any experiment the results of which can be calculated by means of quantum mechanics will come out as you say, but still such a scheme cannot be a final description of Nature'."[52] Thus, Einstein reiterated his belief in his general conception of a physical theory, including the Heisenberg-Einstein context principle, by accepting quanum mechanics as far as it seemed to him to go while insisting on its incompleteness. The latter theme has been at the core of the **EPRG** as well as the Bohr-Einstein debate.

If we seriously believe in the Heisenberg-Einstein context principle but are at the same time critical of a particular physical theory T, such as quantum mechanics, we could very well reject T's claim to being a final theory by arguing as follows: since it is always the theory which decides what can be observed, we had better always decide *first* what kind of fundamental theory T'should physics aim at. We had better do so before we decide which of the old concepts – wave, particle, position, velocity etc. – have a limited range of applicability in the domain of the extant theory, which limitations in the case of quantum mechanics are given by Heisenberg's relations of uncertainty. And we had better ask such an important question when it is openly a matter of finding a theory which may be taken to serve as a fundamental, even final, theory for the whole of physics.

To formulate the same argument differently, if it is always the theory which decides what can be observed, then this should have important consequences, *first* for our understanding of the extant physical theories – the STR, the GTR, and quantum field theories such as quantum electrodynamics, quantum chromodynamics, the standard model, supersymmetry as a possible extension of the standard model, and so on – and for our epistemic appraisal of those theories, and, *secondly*, for the methodology which the physicists must follow in their search for the kind of fundamental physical theory which physics can aim at. Thus, it should have general, or rather universal, methodological consequences in the following sense. Before the physicist subjects the extant theories to severe criticism, or to epistemic appraisal, he had better decide *first* what kind of fundamental theory should physics aim at, this being the question to which every generation of physicists, and methodologists, must return, thereby making physics essentially open-ended. And given the specifications for such a theory, he had better *decide* what kind of methodology of theory-finding should it adopt or develop. But the methodology of theory-finding, like the methodology of theory-testing, can work best when it is guided by the *context-building* from within physics in which all the extant successful theories participate.

Many specialists in the foundations of classical physics recognize rather only indirectly the enormous power of the optical metaphor which seems to be at work when we say that it is always the theory in physics which decides what can and what cannot be observed. But I think that there is more than an optical metaphor at work here. The methodological impor-

[52] Reference 13, S. 411.

tance of Heisenberg-Einstein context principle has rather to do with that fundamental aspect of physical theory, or its mathematical formalism, which performs best by making novel predictions and which decides whether all that which is being predicted is observable. One can here think of the quark theory which forbids observation of quarks and gluons in isolation. To quote Steven Weinberg in this context, "The idea that quarks and gluons can in principle never be observed in isolation has become part of the accepted wisdom of modern elementary particle physics, but it does not stop us from describing neutrons and protons and mesons as composed of quarks. I cannot imagine anything that Ernst Mach would like less.

The theory was only one step in a continuing process of reformulation of physical theory in terms that are more and more fundamental and at the same time farther and farther from everyday experience. How can we hope to make a theory based on observables when no aspect of our experience – perhaps not even space and time – appears at the most fundamental level of our theories? It seems to me unlikely that the positivist attitude will be of much help in the future."[53] While there need not be always a strictly positive correlation between that which a theory predicts for the first time and that whose observability it prohibits, the more admirable will be the theory if it itself suggests experiments to test the novel predictions which it makes. In this regard, the role of its mathematical formalism will be crucial. Moreover, if the observation of certain kinds of phenomena is forbidden by a particular theory T, the same kind of phenomena may become observable on the basis of another theory T′, where T′ is better than T in many other respects, besides being a successor theory in the same field. The bending of the rays of star light by the gravitational force of the sun was a phenomenon predictable from Newton's universal law of gravitation but not observable (a case of negative correlation). It became observable on the basis of Einstein's GTR which predicted it far more precisely than did Newton's law, even suggesting experiments to test its own novel predictions. In any case, by their very nature as universal statements of prohibitions, physical theories rule out one kind of phenomena by allowing some other kind of phenomena. It is not surprising if it is always the theory which decides what can, or what cannot, be observed.

8. The Duhem-Heisenberg-Einstein Methodology

I believe that the examples of Pierre Duhem, Werner Heisenberg, Albert Einstein and Niels Bohr offer themselves as excellent candidates for case-studies in the methodology of dynamic core-contexts of discovery for finding the path to the kind of fundamental theory which physics at the frontier of discovery can aim at. Duhem's work on the aim and structure of physical theory – (T) – remains a great milestone in the philosophy of physics of the last century. While reflecting on this theme, he argues that there are *two* major aspects, or parts, of (T), which call for careful attention and analysis by the physicists and philosophers of physics.[54] On the one hand, (T) has an explanatory part with which it proposes to take hold of the reality underlying the phenomena. On the other hand, (T) has a representative part, with which it proposes to bring about a natural classification of laws.[55] If we were to rationally reconstruct this distinction, the former aspect of (T) can be called its explanatory power in the standard sense of this term. The other important aspect of (T), to which he draws our attention for the first time, remains still a subject of great neglect at the hands of the experts in the philosophy of physics. I believe that it deserves a serious scholarly attention. Elsewhere I have argued for the methodology of theory-problem interactive systems, distinguishing between the explanatory power of

[53] Steven Weinberg (1992), *Dreams of a Final Theory.* Pantheon Books: New York. P. 184.

[54] Pierre Duhem (1906), *La theorie physique son object et sa structure.* 2nd edition 1914. See English translation: *The Aim and Structure of Physical Theory.* P. P. Wiener (transl.), Princeton University Press, Princeton, 1954, p. 32.

[55] Reference 54, pp. 31–33.

(T) and the resolving power of (T).[56] In this proposal, it is the resolving power of (T) which emerges as the most important aspect of a physical theory. A close similarity between Duhem's distinction and my own distinction cannot be ruled out. I think that in his analysis of physical theory, its aim and structure, Duhem might have been far ahead of his times, particularly in the context of the state of the art called logic amd methodology of scientific discovery. In so far as his methodology clearly recognizes *two* parts of (T), with the representative part being accorded by him the most important role in the dynamics of the growth of scientific knowledge within physics, my present methodological proposals – particularly the methodology of the dynamic core-contexts of discovery – can be seen as a further development of Duhemian methodology. On the other hand, I think that both the Bohr-Einstein debate and Heisenberg-Einstein dialogue on physical theory should be studied as further milestones in methodology leading to *quantum-context-building*. Together with Einstein's theories of relativity, directly or indirectly they help in building up the core-contexts of discovery in contemporary physics, if only by promoting an improved understanding of quantum mechanics. Equally important and relevant in this context are Heisenberg's and Einstein's conceptions of physical theory, in which there is no role for *ad hoc* hypotheses as a possible strategy to adjust the parameters of a theory which is faced with unfavourable experimental evidences. Thus, Heisenberg's geschlossene Theorien[57] and Einstein's complete theories[58] may even be regarded as variations on the same theme.[59] But it must be noted that they share nothing in common with the Kuhnian paradigms. Heisenberg's intention in introducing the concept of geschlossene Theorien is echoed by Steven Weinberg's following statement about quantum mechanics: "I simply do not know how to change quantum mechanics by a small amount without

[56] References 4 & 5. In a discussion going back to May–June 1986 at the Frei Universität Berlin, where I had just begun my research stay as a Fellow of the Alexander von Humboldt-Stiftung, Lorenz Krüger found himself in agreement with me, citing Diracs prediction of the positron to show how by its resolving power a physical theory is able to generate new problems. Dirac's relativistic theory of the electron predicted the existence of the positron before its experimental discovery by Anderson. In this case, it is clearly a whole cluster of theories – quantum mechanics, the special theory of relativity and the hypothesis of the electron spin – which determined the problem which Dirac was so successful in solving.

[57] Werner Heisenberg: "Der Begriff 'abgeschlossene Theorie' in der modernen Naturwissenschaft" in *Dialectica* **2**, 1948 (331–336). See also Werner Heisenberg (1984), *Gesammelte Werke.* Piper: München/Zürich. Band **I,** 1929–1955, S. 335–340. Herausgegeben von Walter Blum, Hans-Peter Dürr und Helmut rechenberg. For a detailed discussion, see Erhard Scheibe (2001), *Between Rationalism and Empiricism: Selected Papers in the Philosophy of Physics*. Springer Verlag: New York. PP. 136-141.

[58] Reference 51.

[59] I reject as untenable the view which regards the Kuhnian notion, or rather notions, of a paradigm as peculiarly reminiscent of Heisenberg's notion of a "closed theory". The pragmatic complexities and the semantic ambiguities inherent in the former, which have led Kuhn himself later on to distinguish the different senses of 'paradigm' present in his own work, do not allow any comparison between them. A Kuhnian paradigm in crisis, for example, is not just a theory in crisis. It is rather the normal scientific community and its disciplinary matrix which are in crisis when there is such a crisis at all. In other words, when a science, or any part of it, is undergoing test, it is not the theory which bears the main ordeal of the test. It is rather the puzzle-solving normal scientific community – *science in context* – which bears the ordeal. Thus, in Kuhn's picture, the *theory* suffers a displacement from that center which it still occupies in actual scientific practice and in the methodology of science. When Heisenberg argued that a closed theory is not amenable to small improvements by small changes in its structure – in its formalism – he was only following the best of the traditions in scientific practice, which disallows any scientific change by resort to *ad hoc* hypotheses. Many physicists would follow Heisenberg's own idea of a "closed theory" as a methodological rule facilitating rational scientific change within physics by replacing a theory as a whole whenever it is possible to do so. My criticism here applies particularly to the kind of account to be found in Mara Beller (1999), Quantum Dialogue. University of Chicago Press: Chicago & London. P. 288. See Thomas Kuhn (1962), *The Structure of Scientific Revolutions.* University of Chicago Press: Chicago. (2nd enlarged ed. 1970).

wrecking it altogether."[60] Heisenberg's also held the view that new problems in physics are inherited by the physicists from its historical development. They cannot be invented as theories are. I think that this asymmetry between problems and theories is of great methodological significance in so far as this too hints at the role of context-building in physics. The same is, I think, true of what Einstein said as early as 1916: "Es ist das schönste Los einer physikalischen Theorie, wenn sie selbst zur Aufstellung einer umfassenden Theorie den Weg weist, in welcher sie als Grenzfall weiterlebt."[61] To turn to quantum mechanics as a **CCD** in current physics, and not just as a tool of description or calculation, is a most challenging as well interesting task for physics at present. If there has been any turning around of fundamental physical theory to allow it to play such an important role, it is to be found in quantum mechanics, besides other theories. Did not Duhem[62] already hint at that very aspect of physical theory which can suggest discovery? When I proposed it myself (Pandit, 1983, 1991)[63] that serious attention be paid to the resolving power of a physical theory, I did not realise it then that I was only adding a variation on the Duhemian methodology, or rather on the Duhem-Heisenberg-Einstein methodology. To our question which was posed in § 7. above, there emerges an answer from the fore-going discussion as follows: The kind of fundamental theory which physics can aim at, as it moves from one frontier of discovery to the other, is that which not only explains what there is to explain in physics but also generates new problems on its way to interrogating nature as far as its own framework allows it to do.

9. Conclusion

According to an aphorism of Eddington "You cannot believe in astronomical observat-ions before they are confirmed by theory."[64] I think that what is true of astronomy is equally true of cosmology and the elementary particle physics. For in these fields the two enjoy a kind of strong reciprocal relationship. A sound cosmological model of the very early universe will receive important inputs from what happens at the frontier of discovery in high energy physics. Think of a fundamental theory in physics unifying all the four fundamental forces. Such a theory should find its testing ground in the very early universe as described by a plausible cosmological model of the various epochs of the universe.[65]From this, taken together with the Heisenberg-Einstein context principle, we can make a strong case for a rather strongly reciprocal relationship between theory and experiment in the disciplinary contexts of physics, astronomy and cosmology as follows: No only is it always the case that it is the theory which decides what can be observed, or measured, but also that it is the theory which is invited to interpret observations before those observations can be confirmed and allowed to play a rather crucial role in the frontier of scientific discovery. *On the other hand*, every *new* theory must be confirmed by novel experiments, such that the larger the number and variety of such experiments, the more mature will the theory be as core-context of discovery.[66]

[60] Reference 53, p. 88.
[61] Reference 7.
[62] Reference 54.
[63] References 4 & 5.
[64] As quoted in S. Chandrasekhar, "Development of General Relativity", *Nature* **252** (Nov. 1, 1974, 15–17), p. 17.
[65] The kind of reciprocity I am here hinting at is expressed in the following statement by Edward W. Kolb and Michael S. Turner (1990, p. 494), *The Early Universe,* Addison-Wesley Publishing Company: "Even a less than optimist person would have to conclude that the answers to many of the pressing questions must lie in understanding the earliest history of the universe, which in turn necessarily involves the application of physical theory at the most fundamental level to the cosmological setting." Turn to reference 5, pp. 265–326.
[66] The celebrated astrophysicist S. Chandrasekhar argued for a similar view in the context of astronomical observations. Turn to reference 64.

Bibliography

[1] M. Beller (1999), *Quantum Dialogue*. University of Chicago Press: Chicago & London.

[2] S. Chandrasekhar (1974), *Development of General Relativity* in Nature **252** (Nov. 1, 1974, pp. 15–17).

[3] P. A. M. Dirac (1978), *The Development of Quantum Mechanics* in H. Hora and J. R. Shepanski (eds.) *Directions in Physics* (Lectures delivered by P. A. M. Dirac August/September 1975). John Wiley & Sons: New York, 1978.

[4] P. A. M. Dirac (1983), *The origin of quantum field theory* (pp. 39–55) in Laurie M. Brown and Lillian Hoddeson (eds.) *The Birth of Particle Physics*. Cambridge University Press: Cambridge, 1983.

[5] P. Duhem (1906), *La theorie physique son object et sa structure.* 2nd French ed. 1914. English translation by P. P. Wiener: *The Aim and Structure of Physical theory.* Princeton University Press: Princeton, 1954.

[6] A. Einstein (1919), *Über die spezielle und die allgemeine Relativitätstheorie*. Vierte Auflage: Fried. Vieweg & Sohn, Braunschweig, Germany (first German ed. 1916).

[7] A. Einstein, B. Podolsky, and N. Rosen (1935), *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?* in Physical Review **47** (777–780).

[8] G. Fraser (2000), *Antimatter: The Ultimate Mirror.* Cambridge University Press: Cambridge.

[9] Heisenberg, W. (1948), *Der Begriff 'abgeschlossene Theorie' in der modernen Naturwissenschaft* in Dialectica **2**, 1948 (331–336).

[10] W. Heisenberg (1969), *Der Teil und das Ganze: Gespräche im Umkreis der Atomphysik.* R. Piper & Co. Verlag: München.

[11] W. Heisenberg (1971), *Physics and Beyond*. Harper & Row: London/New York. Translated from German by Arnold J. Pomerans.

[12] W. Heisenberg (1979), *Quantentheorie und Philosophie: Vorlesungen und Aufsätze.* Herausgegeben von Jürgen Busche. Philipp Reclam jun. Stuttgart.

[13] W. Heisenberg (1984, 1985, 1989), *Gesammelte Werke*. Band **I,** Band **II**, Band **III**, Band **V**. Herausgegeben von Walter Blum, Hans-Peter Dürr und Helmut Rechenberg. Piper: München/Zürich.

[14] E. W. Kolb and Michael S. Turner (1990), *The Early Universe.* Addison-Wesley Publishing Company.

[15] T. Kuhn (1962), *The Structure of Scientific Revolutions.* University of Chicago Press: Chicago. 2nd enlarged ed. 1070.

[16] A. Pais (1982), *Subtle is the Lord: The Science and Life of Albert Einstein.* Oxford University Press: New York.

[17] G. L. Pandit (1983), *The Structure and Growth of Scientific Knowledge: A Study in the Methodology of Epistemic Appraisal*. Boston Studies in the Philosophy of Science **73.** D. Reidel: Dordrecht.

[18] G. L. Pandit (1991), *Methodological Variance: Essays in Epistemological Ontology and the Methodology of Science.* Boston Studies in the Philosophy of Science **131**. Kluwer Akademic Publishers: Dordrecht.

[19] G. L. Pandit (1995), *Von der Ökologie des Bewußtseins zum Umweltrealismus: Wiederentdeckung menschlicher und nicht-menschlicher Interessanssphären*. Picus Verlag: Wien.

[20] G. L. Pandit (2000), *Environmental Actions: Can a Part Control the Whole?* in Systemica **12** (351–368).

[21] G. L. Pandit (2001), *Emergence of Human Interest Studies – The Environmental Dimension* in Systemica (forthcoming).

[22] E. Scheibe (2001), *Between Rationalism and Empiricism: Selected Papers in the Philosophy of Physics.* Springer Verlag: New York.

[23] S. Weinberg (1992), *Dreams of a Final Theory.* Pantheon Books: New York.

[24] E. P. Wigner (1971), */he Subject of Our Discussions*, in B. d'Espagnat (ed.) *Foundations of Quantum Mechanics*. International School of Physics *Enrico Fermi* 1970. Academic Press: New York.

[25] E. P. Wigner (1982), *The Limitations of the Validity of Present-Day Physics* (pp. 118–133) in Richard Q. Elvee (ed.) *Mind in Nature: Nobel Conference* **XVII**. Harper & Row Publishers: San Francisco.

Fortschr. Phys. **50** (2002) 5–7, 483–488

Elementary Particles Theory

On the Large Nc Expansion in Quantum Chromodynamics

WILLIAM A. BARDEEN

Theoretical Physics Department
Fermilab, MS 106, P.O. Box 500
Batavia, Illinois 60510, USA
bardeen@fnal.gov

Abstract

I discuss methods based on the large Nc expansion to study nonperturbative aspects of quantum chromodynamics, the theory of the strong force. I apply these methods to the analysis of weak decay processes and the nonperturbative computation of the weak matrix elements needed for a complete evaluation of these decays in the Standard Model of elementary particle physics.

1. Introduction

Field theories are frequently studied through a perturbative expansion in the interaction strength or coupling constant. In many cases, a nonperturbative analysis is required to apply these theories in physical situations. The large N expansion is a nonperturbative method of analysis that makes use of particular limits for parameters unrelated to the coupling constant. For example, O(N) spin systems where N is the number of spin components can be studied by mean field methods which become exact in the large N limit. Applications of perturbative QCD, such as the cross-sections for high p_t jets, are greatly simplified using color-ordered amplitudes and a large Nc expansion where Nc is the number of colors, the quantum number associated with the QCD gauge dynamics. The large Nc expansion can also be used to study nonperturbative aspects of quantum chromodynamics with applications to the structure of chiral symmetry breaking and hadronic string theory. More recently, it has been established that there exists a duality between the large N limit of SU(N) supersymmetric Yang-Mills theory and classical supergravity in a higher dimensional spacetime. In this talk, I will discuss the nature of the large Nc expansion in QCD and its application to the computation of weak decay amplitudes.

2. Large Nc Expansion in QCD

The large Nc expansion defines a nonperturbative reordering of the QCD coupling constant expansion. Each Feynman diagram can be classified by its dependence on the strong coupling constant, α_{strong}, and the number of colors Nc. In 1974, 't Hooft [1] argued that the structure of the theory simplified in the limit, $\text{Nc} \to \infty$, $\alpha_{\text{strong}} \to 0$, $\alpha_{\text{strong}} \cdot \text{Nc}$ fixed. The theory is summed to all orders in the rescaled coupling constant, $\alpha_{strong} \cdot \text{Nc}$, with corrections being formally suppressed by powers of 1/Nc. At leading order in this expansion, the

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

gluons are constrained to be in planar Feynman diagrams. Quarks lines form boundaries of the planar surfaces formed by the gluons. In this sense, the structure of large Nc QCD is similar to an open string theory with quarks attached to the ends of the string.

If we assume that large Nc QCD is confining, then the physical states are expected to consist of stringlike glueballs and mesons formed as quark–antiquark boundstates. The physical states are towers of meson resonances with increasing mass and spin. The lowest order meson scattering amplitude has the structure of a gluonic disk with meson boundstates attached to the edge of the disk. Since the color wavefunction of a meson boundstate is $O(1/\sqrt{Nc})$, the leading order meson amplitudes with more than two mesons boundstates attached to the edge of the disk are suppressed by a factor of (1/Nc) for each additional meson. This simple power counting implies that all mesons are stable at leading order and that the meson S-matrix is that of a trivial free meson theory. We may still consider the leading contribution to processes at any order in the large Nc expansion. A meson scattering amplitude at leading nontrivial order involves only meson tree amplitudes with simple poles at positions of the meson bound-states. Higher order diagrams involve either additional quark loops which are suppressed by a power of (1/Nc) or nonplanar diagrams which are suppressed by powers of $(1/Nc^2)$. The 1/Nc expansion is a topological diagram expansion with the complexity increasing at higher order. In two space-time dimensions, large Nc QCD has an exact solution while only its structure can be inferred in higher dimensions.

3. Weak Decay Amplitudes

Nonleptonic weak decays are mediated by the virtual exchange of W and Z bosons. These processes occur at a short distance scale, and their effects can be expressed in terms of effective local interactions between quark currents and densities at low energy [2]. The weak decay amplitudes are written as an expansion in terms of short distance Wilson coefficients, $C_i(\mu)$, and sets of local quark operators, $Q_i(\mu)$,

$$A(K \to \pi\pi) = \frac{G_F}{\sqrt{2}}\, V_{\mathrm{CKM}} \sum_i C_i(\mu)\, \langle K|\, Q_i(\mu)\, |\pi\pi\rangle\,. \tag{1}$$

where a normalization scale, μ, is introduced to separate the short and long distance physics contributions. Details of the short distance processes and perturbative QCD dynamics are used to compute the Wilson coefficients. Operator matrix elements are sensitive to long distance physics including the mechanisms of quark confinement and the formation of hadronic bound-states that can not be computed using perturbative QCD. The systematic computation of relevant Wilson coefficient functions for weak decays at two loops, (NLO), have been made by several groups [3]. The coefficient function may be expressed as

$$C_i(\mu) = Z_i(\mu) + i^* Y_i(\mu)\,, \tag{2}$$

where the coefficient, $Y_i(\mu)$, incorporates the short distance effects of CP violation.

Operator matrix elements are more difficult to analyze as perturbative methods can not be used. Several strategies have been employed to compute the matrix elements of the quark operators appearing in the expansion of the weak Hamiltonian. All of the operators being studied have the form of products of quark currents and densities which themselves are color-singlet quark bilinears. The simplest approximation invokes factorization where it is assumed the quark currents and densities independently couple to the hadrons in the initial and final states. The factorization approximation has inherent limitations as it fails to reproduce the scale dependence of the Wilson coefficents. A second method employs nu-

merical computations in the lattice formulation of QCD. The lattice method is constrained by the size of the numerical effort required and by problems associated with the chiral structure of lattice QCD. There has be recent progress in the direct computation of weak matrix elements using the domain wall formulation of chiral fermions on the lattice [4]. Another nonperturbative method for calculating the weak operator matrix elements invokes the large Nc expansion of QCD as discussed in the following section.

4. Large Nc Computation of Weak Matrix Elements

The large Nc analysis focuses on the computation of weak matrix elements of the four-quark operators, Q_i, generated by the renormalization group expansion.

$$\langle Q_i \rangle = \langle (\bar{\Psi}\Gamma_i\Psi)\,(\bar{\Psi}\Gamma_i\Psi) \rangle\,. \tag{3}$$

The analysis of the large Nc expansion of QCD relies on the toplogical structure of the quark diagrams and large Nc power counting to determine the various contributions to the weak matrix elements. At the leading order of the large Nc expansion, the matrix elements are factorized

$$\langle Q_i \rangle_F = \langle (\bar{\Psi}\Gamma_i\Psi) \rangle\, \langle (\bar{\Psi}\Gamma_i\Psi) \rangle\,. \tag{4}$$

We recall that the leading order quark amplitudes have the structure of a disk with the quark line forming the boundary of the disk with planar gluons filling in the surface of the disk. Since the quarks have Nc colors, the factorized quark densities must each be associated with separate disks and, therefore, two powers of Nc are generated, one for each disk. This factorized contribution is the leading order contribution (LO). Gluonic interactions between the two disks are nonplanar and therefore suppressed by at least two powers of Nc.

At next-leading-order (NLO) in the large Nc expansion, O(1/Nc), there are two possible contributions. An internal quark loop can be added to either disk of the leading order calculation giving a suppression of at least 1/Nc. Since this insertion does not involve the second disk it contributes only to the factorized matrix element. This correction represents a meson loop correction to the matrix element of the quark current or density. It will be assumed that we are able to measure these current matrix elements which include all order corrections phenomenologically.

The second contribution at NLO is nonfactorized. It arises when both quark bilinear operators are associated with a single disk. Because only one quark loop is involved instead of two, the amplitude is suppressed by a power of 1/Nc relative to the LO factorized contribution. At this order in the large Nc expansion, the two quark currents and the various meson bound-states are associated with the single quark line forming the boundary of the single disk. Hence, this amplitude is represented by a tree-level meson amplitude having only poles at the positions of the infinite tower of meson bound-states.

We can write this amplitude as the momentum integral of a meson amplitude with two external bilinear currents,

$$\langle Q_i \rangle_{\rm NF} = \int dk\, A_{\Gamma_i\Gamma_i}(k,\, -k,\, p_1 \ldots p_N)\,. \tag{5}$$

Knowledge of meson tree amplitudes may be used to compute integrand. At low momentum, chiral Lagrangian description is an exact representation of QCD dynamics for matrix elements involving low energy meson states. At high momentum, we can use large Nc version of the QCD renormalization group equation to compute integrand in terms of per-

turbative coefficient functions and factorized matrix elements. The entire integral is obtained by interpolating between the nonperturbative low energy approximation and the perturbative high energy amplitude. Calculations based on this duality have achieved some success in explaining the structure of weak decay amplitudes including the octet enhancement observed in $K \to 2\pi$ decays and the CP violation seen the ε'/ε measurement [5]. The precision of these methods depend on three ingredients:

- the phenomenological determination of the long distance meson amplitude,
- the order of the perturbative short distance calculation, the α_{strong} expansion,
- the accuracy of the interpolation between the long and short distance approximations.

The amplitude calculation requires the evaluation of the momentum integral in Eq. (5). At high momentum (short distance), $k \to \infty$, $p_1 \ldots p_N$ – small, the operator product expansion (PQCD) takes the form,

$$A_{\Gamma\Gamma}(k, -k;\, p_1, \ldots, p_N) \to C_{\Gamma\Gamma;\,Q}(k, \mu)\, \langle Q(\mu)\rangle_{\text{LO}}\,, \tag{6}$$

where the coefficient function, C, is O(1/Nc) and the operator matrix element is factorized. Note that all weak mixing processes are O(1/Nc). Using the renormalization group and the anomalous dimension matrices of perturbative QCD, we compute evolution of the NLO weak matrix elements,

$$\mu\, \partial\mu \langle Q_i(\mu)\rangle_{\text{NLO}} = -\gamma_{ij}(\alpha(\mu))\, \langle Q_j(\mu)\rangle_{\text{LO}}\,, \tag{7}$$

where the LO matrix element is factorized. The integrated form at NLO is

$$\langle Q_i\rangle_{\text{NLO}} = -\int d\mu^2/2\mu^2 \gamma_{ij}(\alpha(\mu))\, \langle Q_j(\mu)\rangle_{\text{LO}}\,. \tag{8}$$

The long distance contributions are described by tree-level meson amplitudes which involve nonperturbative parameters which are determined phenomenologically A number of approximations to the low energy meson amplitudes has been used in the study of weak decay amplitudes. They include chiral Lagrangians, chiral Lagrangians + vector mesons, chiral quark models, and extended Nambu–Jona-Lasinio models. The chiral Lagrangian is an exact description of QCD in the light quark limit and for low momenta. The various approximations given above are different attempts to parameterize the physical low energy dynamics. The models require the input of masses and coupling constants from data other than the weak decay processes. Further efforts are needed to improve extrapolation to higher momentum scales to more precisely match the perturbative QCD calculations of the high momentum part of the two-current correlation function. These efforts may involve adding additional mesons (vector, axial-vector, scalar, tensor, ...) or higher dimension operators in the effective field theory ($O(p^4)$, $O(p^6)$, ...).

In addition to the precise computation of the above tree-level meson amplitudes, matching conditions are also required to connect the momentum integral of the two-current correlation function with the standard perturbative analysis of the weak matrix elements. At NLO in the large Nc expansion, the weak decay amplitudes are divergent at high energy. The standard short distance analysis uses dimensional regularization, (NDR, HV), to regularize the quark amplitudes and define normalization scales for the coefficient functions and operator matrix elements. Both the chiral Lagrangian calculation and the short distance quark calculation may be regularized by a cutoff on momentum flowing through color singlet currents.

A consistent treatment of the short distance contributions matches the different regularization schemes by computing quark matrix elements of weak operators using both dimen-

sional regularization and momentum cutoff methods. To isolate the purely short distance effects, an infrared regularization is introduced,

$$\int dk \to \int dk\, k^2/(k^2 - M^2)\,, \qquad M = IR\,\text{cutoff}\,. \tag{9}$$

We may now compute the perturbative matrix elements of the various quark operators using the different UV cutoff schemes. The operator matrix elements will have a finite correction factor at one loop,

$$\langle Q_i^{\mathrm{NDR,HV}} \rangle = \langle Q_i^{\mathrm{mom}} \rangle - w_{ij}^{\mathrm{NDR,HV}} \left(\frac{\alpha}{4\pi}\right) \langle Q_j^{\mathrm{mom}} \rangle\,, \tag{10}$$

where w_{ij} is rotation matrix between dimensional regularization basis and momentum cutoff basis.

This rotation matrix has been computed for all ten four-fermion quark operators used to expand the weak Hamiltonian for both NDR and HV requlrization schemes [6]. The effective coefficient functions to use with momentum subtracted matrix elements can now constructed,

$$C_i^{\mathrm{mom}} = C_i^{\mathrm{NDR,HV}} - C_k^{\mathrm{NDR,HV}} w_{ki}^{\mathrm{NDR,HV}} \left(\frac{\alpha}{4\pi}\right)\,. \tag{11}$$

An example of effect of this shift on coefficient functions can be evaluated. Our numbers are based on the perturbative QCD analysis of Bosch et al. (1999) [7]. The normalization scale is $\mu = 1.3$ Gev and $\Lambda_{\mathrm{QCD}} = 0.340$. The result for certain coefficient functions are given below

CF	NDR	HV	$\mathrm{mom}_{\mathrm{NDR}}$	$\mathrm{mom}_{\mathrm{HV}}$
Z1	−0.425	−0.521	−0.669	−0.687
Z2	1.244	1.320	1.371	1.394
Y3	0.030	0.034	0.041	0.041
Y4	−0.059	−0.061	−0.063	−0.064
Y5	0.005	0.016	0.013	0.011
Y6	−0.092	−0.083	−0.091	−0.090

There is an enhancement of the Z1 and Z2 coefficients in the momentum basis over the values in either the NDR or HV schemes. The larger values of Z1 and Z2 will tend imply a larger octet enhancement and 27 supression factors than the conventional treatment. Of course, the important point of this aspect of the analysis is that we now have all the elements to make a consistent use of the 1/Nc expansion for computing the weak matrix elements:

- the standard renormalization group analysis of the weak coefficient functions is used to compute the short distance contributions and evolve the operators to low energy scales,
- a consistent matching between the dimensional regularization schemes and the momentum cutoff is used in the NLO analysis of the nonfactorized weak matrix elements,
- chiral Lagrangians or other effective field theories are used to describe the long distance contributions to the weak matrix elements and consistently matched to the short distance contributions.

The complete calculation has yet to be fully integrated. There are still some issues regarding matrix elements involving scalar and pseudoscalar densities that can affect the evaluation of the Q_6 and Q_8 operator matrix elements needed for ε'/ε analysis.

5. Conclusions

Precision tests of the Standard Model require knowledge of nonperturbative aspects of quantum chromodynamics, the strong dynamics. The Large Nc expansion combined with phenomenological knowledge of certain meson amplitudes provides one avenue for systematic estimates. The analysis outlined in this talk provides the elements for this systematic analysis of the weak decay matrix elements through next-leading-order However it is fundamentally limited by the ability to compute yet higher order terms in the large Nc expansion.

Direct numerical computation of weak matrix elements using lattice formulations of QCD provides another avenue. Recent developments related to the chiral symmetry structure of lattice QCD may lead to more realistic calculations of the necessary matrix elements.

The final picture remains unclear. As yet there is no clear violation of the Standard Model to be inferred by weak decay processes but the large value of ε'/ε and the large value of the $\Delta I = 1/2$ amplitude for $K \to 2\pi$ may yet challenge its validity.

Acknowledgements

I thank the Alexander von Humboldt Foundation for its hospitality and support. This research is supported by the Department of Energy under contract DE-AC02-76CHO3000.

References

[1] G. 't Hooft, Nucl. Phys. **B72** (1974) 461.
[2] M. Gaillard and B. Lee, Phys. Rev. Lett. **33** (1974) 108;
G. Altarelli and L. Maiani, Phys. Lett. **B52** (1974) 351;
F. Gilman and M. Wise, Phys. Rev. **D20** (1979) 2392.
[3] A. Buras, M. Jamin, and M. Lautenbacher, Nucl. Phys. **B400** (1993) 75;
Nucl. Phys. **B408** (1993) 209;
M. Ciuchini, E. Franco, G. Martinelli, and L. Reina, Nucl. Phys. **B415** (1994) 403;
A. Buras, P. Gambino, and U. Haisch, Nucl. Phys. **B570** (2000) 117.
[4] SPQcdR Collaboration, hep-lat/0110169 (2001); CP-PACS Collaboration hep-lat/0108013 (2001); RBC Collaboration hep-lat/0110075 (2001)
[5] W. Bardeen, A. Buras, and J.-M. Gérard, Nucl. Phys. **B293** (1987) 787;
W. Bardeen, Nucl. Phys. Proc. Suppl. **7A** (1989) 149;
A. Pich and E. de Rafael, Nucl. Phys. **B358** (1991) 311;
J. Bijnens, C. Bruno, and E. de Rafael, Nucl. Phys. **B390** (1993) 501;
T. Hambye, G. Köhler, E. Paschos, P. Soldan, and W. Bardeen, Phys. Rev. **D58** (1998) 014017;
Hambye, G. Köhler, and P. Soldan, Eur. Phys. J. **C10** (1999) 271.
[6] W. Bardeen, unpublished.
[7] S. Bosch, A. Buras, M. Gorbahn, S. Jäger, M. Jamin, M. Lautenbacher, and L. Silvestrini, Nucl. Phys. **B565** (2000) 3.

Fortschr. Phys. **50** (2002) 5–7, 489–495

Baryon Asymmetry, Dark Matter and the Hidden Sector

Luís Bento[1)] and Zurab Berezhiani[2,3)]

[1)] Centro de Física Nuclear da Universidade de Lisboa
Avenida Prof. Gama Pinto 2, Lisboa, 1649-003, Portugal
lbento@fc.ul.pt
[2)] Dipartamento di Fisica, Universitá di L'Aquila, I-67010 Coppito, AQ, Italy
INFN, Laboratori Nazionali del Gran Sasso, I-67010 Assergi, AQ, Italy
[3)] Andronikashvili Institute of Physics, GE-380077 Tbilisi, Georgia
berezhiani@fe.infn.it

Abstract

We propose a mechanism in which the baryon asymmetry of the Universe is produced in out-of-equilibrium, $B-L$ and CP violating scattering processes that convert ordinary particles into particles of some hidden sector. The same mechanism should produce a non-zero asymmetry also in the hidden sector particles and thus the latter could provide a natural candidate for the dark matter. We discuss the example of the hidden mirror sector, which also provides a natural framework for the sterile neutrinos, and consider the processes $l\phi \to l'\phi'$, $\bar{l}'\bar{\phi}'$ mediated by the heavy Majorana neutrinos N of the seesaw mechanism, where l and ϕ are ordinary lepton and Higgs doublets and l', ϕ' their hidden counterparts.

Long time ago Sakharov [1] has suggested that a non-zero baryon asymmetry (BA) can be produced in the initially baryon symmetric universe if three conditions are fulfilled: B-violation, C- and CP-violation and departure from thermal equilibrium. These conditions can be satisfied in the decays of heavy particles of grand unified theories. On the other hand, the nonperturbative sphaleron processes, which violate $B+L$ but conserve $B-L$, are effective at temperatures from about 10^{12} GeV down to 100 GeV [2]. Thus, one actually needs to produce a non-zero $B-L$ rather than just B, a fact that disfavors the simplest baryogenesis picture based on grand unification models like $SU(5)$. When sphalerons are in equilibrium, the baryon number and $B-L$ are related as $B = a(B-L)$, where a is a model dependent order one coefficient [3]. Hence, the observed baryon to entropy density ratio $B = n_B/s = (0.6-1)\times 10^{-10}$, needs to produce $B-L \sim 10^{-10}$.

The seesaw mechanism for neutrino masses offers an elegant possibility of generating non-zero $B-L$ in CP-violating decays of heavy Majorana neutrinos N into leptons and Higgses, the so called leptogenesis scenario [4]. Namely, due to complex Yukawa constants, the decay rates $\Gamma(N \to l\phi)$ and $\Gamma(N \to \bar{l}\bar{\phi})$ can be different from each other, so that leptons l and anti-leptons $\bar{l}$ are produced in different amounts.

Here we discuss an alternative mechanism that is based on scattering processes [5]. The main idea consists in the following. There may exist some hidden (shadow or mirror) sector of new particles which are not in thermal equilibrium with the ordinary particle world as far as the two systems interact very weakly e.g., if they only communicate via gravity. However, other messengers may well exist, namely, superheavy gauge singlets like right-handed neutrinos which can mediate very weak effective interactions between the ordinary and shadow leptons. Then, a net $B-L$ could emerge in the Universe as a result of CP-violating effects in the unbalanced production of shadow particles from ordinary particle collisions.

The simplest model of this type can be described as follows. Consider the standard $SU(3)\times SU(2)\times U(1)$ model, containing among other particles species, the leptons $l_i = (\nu, e)_i$ ($i = 1, 2, 3$ is the family index) and the Higgs doublet ϕ. Imagine that there is

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

also a hidden sector with gauge symmetry G', containing fermion and scalar fields that are singlets under the standard model gauge group, while the ordinary particles are instead singlets under G'. In general G' could be any gauge symmetry group containing, among other possible particles, some multiplets of scalar ϕ' and fermions l'_k possessing opposite gauge charges so that the products $l'_k\phi'$ are gauge invariant.

In the spirit of the seesaw mechanism, one can introduce some number heavy fermions N_a $(a = 1, 2, \ldots)$, the heavy neutrinos N, which are gauge singlets and thus can couple to l, ϕ as well as to l', ϕ' In this way, play the role of messengers between ordinary and shadow particles. The relevant Yukawa couplings have the form:

$$h_{ia} l_i N_a \phi + h'_{ka} l'_k N_a \phi' + \frac{1}{2}\, M_{ab} N_a N_b + \text{h.c.} \tag{1}$$

(charge-conjugation matrix C is omitted); all fermion states l, N, l' are taken to be left-handed while their C-conjugate, right-handed anti-particles are denoted as $\bar{l}$, $\bar{N}$, $\bar{l}'$. It is convenient to present the heavy neutrino mass matrix as $M_{ab} = g_{ab}M$, M being the overall mass scale and g_{ab} some typical Yukawa constants. (Without loss of generality, g_{ab} can be taken diagonal and real.)

The particular interesting candidate is the mirror world, an exact duplicate of the observable sector with the same gauge symmetry, which concept has attracted a significant interest over the last years, being motivated by various problems in particle physics and cosmology [6−10]. In this case we have a theory given by the product $G \times G'$ of two identical gauge factors with identical particle contents, which could naturally emerge e.g. in the context of $E_8 \times E'_8$ superstring theories. In particular, the G sector contains the standard gauge group $SU(3) \times SU(2) \times U(1)$ with the ordinary particles ϕ, l, etc., whereas G' contains their mirror partners ϕ', l', etc., in equivalent representations with respect to mirror group $G' = SU(3)' \times SU(2)' \times U(1)'$. It is natural to assume that both particle sectors are described by identical Lagrangians, that is, all coupling constants (gauge, Yukawa, Higgs) have the same pattern in both sectors and thus their microphysics is the same. In particular, one can impose a discrete symmetry under the exchange $\phi \to \phi'^{\dagger}$, $l \to \bar{l}'$, etc., the so-called mirror M-parity, which implies $h'_{ia} = h^*_{ia}$.

Integrating out the heavy states N in the couplings (1), we get the effective operators

$$\frac{A_{ij}}{2M}\, l_i l_j \phi\phi + \frac{D_{ik}}{M}\, l_i l'_k \phi\phi' + \frac{A'_{kn}}{2M}\, l'_k l'_n \phi'\phi' + \text{h.c.}\,, \tag{2}$$

with coupling constant matrices of the form $A = hg^{-1}h^T$, $A' = h'g^{-1}h'^T$ and $D = hg^{-1}h'^T$. Thus, the first operator in Eq. (2), due to the ordinary Higgs vacuum expectation value (VEV) $\langle\phi\rangle = v \sim 100\,\text{GeV}$, induces the small Majorana masses of the ordinary (active) neutrinos. In addition, if the shadow Higgs ϕ' also has a non-zero VEV $\langle\phi'\rangle = v' \ll M$, then the third operator provides the masses of the shadow neutrinos contained in l' (which in fact are sterile for the ordinary observer), while the second operator induces the mixing mass terms between the active and sterile neutrinos. The total mass matrix of neutrinos $\nu \subset l$ and $\nu' \subset l'$ reads as [8]:

$$M_\nu = \begin{pmatrix} m_\nu & m_{\nu\nu'} \\ m^T_{\nu\nu'} & m_{\nu'} \end{pmatrix} = \frac{1}{M} \begin{pmatrix} Av^2 & Dvv' \\ D^T vv' & A'v'^2 \end{pmatrix}. \tag{3}$$

Thus, this model provides a simple explanation of why sterile neutrinos could be light (on the same grounds as the active neutrinos) and could have significant mixing with the ordinary neutrinos. For example, if $v' \sim 10^2 v$ then, the shadow neutrinos ν' with masses of keV order could provide the warm dark matter component in the Universe [9]. Instead, if $\langle\phi'\rangle = 0$, the ν' are massless and unmixed with the ordinary neutrinos.

Let us discuss now the baryogenesis mechanism in our scenario. A crucial role in our considerations is played by the reheat temperature T_R, at which the inflaton decay and entropy production of the Universe is over and under which the Universe is dominated by a relativistic plasma of ordinary particle species. We assume that after the postinflationary reheating, different temperatures are established in the two sectors: $T'_R < T_R$, i.e. the hidden sector is cooler than the visible one, or ultimately, even completely "empty". This situation is motivated by the Big-Bang nucleosynthesis (BBN) constraints, and it can be achieved in the context of certain inflationary models [9−11]. In addition, the two particle systems should interact very weakly so that they do not come in thermal equilibrium with each other after reheating. The heavy neutrino masses are much larger than the reheat temperature T_R and thus cannot be thermally produced. As a result, the usual leptogenesis mechanism via $N \to l\phi$ decays is ineffective.

Now, the important role is played by lepton number violating scatterings mediated by the heavy neutrinos N. The "cooler" hidden world starts to be "slowly" occupied due to the entropy transfer from the ordinary sector through the $\Delta L = 1$ reactions $l_i\phi \to \bar{l}'_k\bar{\phi}'$, $\bar{l}_i\bar{\phi} \to l'_k\phi'$. In general these processes violate CP due to complex Yukawa couplings in Eq. (1), and so the cross-sections with leptons and anti-leptons in the initial state are different from each other. As a result, leptons leak to the hidden sector more (or less) effectively than antileptons and a non-zero $B - L$ is produced in the Universe. It is essential that these processes stay out of equilibrium. In other words, their rate should be less than the Hubble parameter $H = 1.66 g_*^{1/2} T^2/M_{Pl}$ (g_* being the effective number of particle degrees of freedom) for temperatures $T \leq T_R$.

For the rate of $\Delta L = 1$ reactions we have $\Gamma_1 = \sigma_1 n_{\rm eq}$, where $n_{\rm eq} \simeq (1.2/\pi^2)\, T^3$ is an equilibrium density per degree of freedom and σ_1 is the total cross section of $l\phi \to \bar{l}'\bar{\phi}'$ scatterings,

$$\sigma_1 = \sum \sigma(l\phi \to \bar{l}'\bar{\phi}') = \frac{Q_1}{8\pi M^2}\,. \tag{4}$$

The sum is taken over all flavor and isospin indices of initial and final states, and $Q_1 = {\rm Tr}(D^\dagger D) = {\rm Tr}[(h'^\dagger h')\, g^{-1} (h^\dagger h)^*\, g^{-1}]$. Hence, the out-of-equilibrium condition for this process reads as

$$\left(\frac{\Gamma_1}{2H}\right)_R \simeq 1.5 \times 10^{-3}\, \frac{Q_1 T_R M_{Pl}}{g_*^{1/2} M^2} < 1\,. \tag{5}$$

However, there are also scattering processes with $\Delta L = 2$ like $l\phi \to \bar{l}\bar{\phi}$ etc., which can wash out the produced $B - L$ unless they are out of equilibrium [4]. Their total rate is given as $\Gamma_2 \simeq (3Q_2/4\pi M^2)\, n_{\rm eq}$ where $Q_2 = {\rm Tr}\,(A^\dagger A) = {\rm Tr}\,[(h^\dagger h)\, g^{-1} (h^\dagger h)^*\, g^{-1}]$. Therefore, for a given reheat temperature T_R, the Eq. (5) and the analogous condition $K_2 = (\Gamma_2/2H)_R < 1$ translate into the lower limit on the heavy neutrino mass scale M:

$$M_{12} > 4.2\, g_*^{-1/4} Q_1^{1/2} T_9^{1/2}\,, \qquad M_{12} > 10.4\, g_*^{-1/4} Q_2^{1/2} T_9^{1/2}\,. \tag{6}$$

where $M_{12} \equiv (M/10^{12}\,{\rm GeV})$, $T_9 \equiv (T_R/10^9\,{\rm GeV})$ and $g_* \approx 100$ in the standard model. Clearly, if the Yukawa constants h_{ia} and h'_{ka} are of the same order, the out-of-equilibrium conditions for $\Delta L = 1$ and $\Delta L = 2$ processes are nearly equivalent to each other.

Let us turn now to CP violation. In $\Delta L = 1$ processes the CP-odd lepton number asymmetry emerges from the interference between the tree-level and one-loop diagrams of Fig. 1. The tree-level amplitude for the dominant channel $l\phi \to \bar{l}'\bar{\phi}'$ goes as $1/M$ and the radiative corrections as $1/M^3$. For the channel $l\phi \to l'\phi'$ instead, both tree-level and one-loop amplitudes go as $1/M^2$. As a result, the cross section CP-asymmetries are the same

Fig. 1. Tree-level and one-loop diagrams contributing to the CP-asymmetries in $l\phi \to \bar{l}'\bar{\phi}'$ (left column) and $l\phi \to l'\phi'$ (right column)

for both $l\phi \to \bar{l}'\bar{\phi}'$ and $l\phi \to l'\phi'$ channels (on the contrary, the diagrams with $l'\phi'$ inside the loops, not shown in Fig. 1, yield asymmetries, $\pm\Delta\sigma'$, symmetric to each other). However, CP violation takes also place in $\Delta L = 2$ processes (see Fig. 2). This is a consequence of the very existence of the hidden sector namely, the contribution of the hidden particles to the one-loop diagrams of Fig. 2. The direct calculation gives:

$$\sigma(l\phi \to \bar{l}'\bar{\phi}') - \sigma(\bar{l}\bar{\phi} \to l'\phi') = (-\Delta\sigma - \Delta\sigma')/2\,, \tag{7a}$$

$$\sigma(l\phi \to l'\phi') - \sigma(\bar{l}\bar{\phi} \to \bar{l}'\bar{\phi}') = (-\Delta\sigma + \Delta\sigma')/2\,, \tag{7b}$$

$$\sigma(l\phi \to \bar{l}\bar{\phi}) - \sigma(\bar{l}\bar{\phi} \to l\phi) = \Delta\sigma\,; \tag{7c}$$

$$\Delta\sigma = \frac{3J\,S}{32\pi^2 M^4}\,, \tag{7d}$$

where $J = \mathrm{Im\,Tr}\,[(h'^\dagger h')\,g^{-2}(h^\dagger h)\,g^{-1}(h^\dagger h)^* g^{-1}]$ is the CP-violation parameter and S is the c.m. energy square ($\Delta\sigma'$ is obtained from $\Delta\sigma$ by exchanging h with h').

This is in perfect agreement with CPT invariance that requires that the total cross sections for particle and anti-particle scatterings are equal to each other: $\sigma(l\phi \to X) = \sigma(\bar{l}\bar{\phi} \to X)$. Indeed, taking into account that $\sigma(l\phi \to l\phi) = \sigma(\bar{l}\bar{\phi} \to \bar{l}\bar{\phi})$ by CPT, we see

Fig. 2. Tree-level and one-loop diagrams contributing to the CP-asymmetry of $l\phi \to \bar{l}\bar{\phi}$. The external leg labels identify the initial and final state particles.

that CP asymmetries in the $\Delta L = 1$ and $\Delta L = 2$ processes should be related as

$$\sigma(l\phi \to X') - \sigma(\bar{l}\bar{\phi} \to X') + \sigma(l\phi \to \bar{l}\bar{\phi}) - \sigma(\bar{l}\bar{\phi} \to l\phi) = 0 \ , \tag{8}$$

where X' are the hidden sector final states, $\bar{l}'\bar{\phi}'$ and $l'\phi'$. That is, the $\Delta L = 1$ and $\Delta L = 2$ reactions have CP asymmetries with equal intensities but opposite signs. But, as L varies in each case by a different amount, a net lepton number decrease is produced, or better, a net increase of $B - L \propto \Delta\sigma$ (recall that the lepton number L is violated by the sphaleron processes, while $B - L$ is changed solely by the above processes).

As far as we assume that the hidden sector is cooler and thus depleted of particles, the only relevant reactions are the ones with ordinary particles in the initial state. Hence, the evolution of the $B - L$ number density is determined by the CP asymmetries shown in Eq. (7) and obeys the equation

$$\frac{dn_{B-L}}{dt} + 3Hn_{B-L} = \frac{3}{4}\ \Delta\sigma\, n_{\rm eq}^2 \, . \tag{9}$$

Since the CP-asymmetric cross section $\Delta\sigma$ is proportional to the thermal average c.m. energy square $S \simeq 17\, T^2$ and $H = 1/2\, t \propto T^2$, one integrates the above equation from $T = T_R$ to the low temperature limit and obtains the final $B - L$ asymmetry of the Universe as

$$B - L = \frac{n_{B-L}}{s} = \left[\frac{\Delta\sigma\, n_{\rm eq}^2}{4Hs}\right]_R \ , \tag{10}$$

where $s = (2\pi^2/45) g_* T^3$ is the entropy density.

The following remark is in order. In fact, the lepton number production starts as soon as the inflaton starts to decay and the particle thermal bath is produced, before the reheat temperature is established. (Recall that the maximal temperature at the reheating period is usually larger than T_R.) In this epoch the Universe is still dominated by the inflaton oscillations and therefore it expands as $t^{2/3}$ while the entropy of the Universe grows as $t^{5/4}$. The integration of Eq. (9) from some higher temperatures down to T_R gives an asymmetry 1.5 times larger than the estimation (10). Taking all these into account, the final result can be recasted as follows:[1)]

$$B - L \approx 2 \times 10^{-3}\ \frac{JM_{Pl}T_R^3}{g_*^{3/2}M^4} \approx 2 \times 10^{-8}\ \frac{JT_9^3}{M_{12}^4} \ , \tag{11}$$

where we have taken again $g_* \approx 100$. Taking also into account the lower limits (6), we obtain the upper limit on the produced $B - L$:

$$B - L < 10^{-8}\ \frac{JT_9}{Q^2} \ ; \qquad Q = \max\{Q_1, 6Q_2\} \, . \tag{12}$$

This shows that for Yukawa constants spread e.g. in the range $0.1 - 1$ one can achieve $B - L = \mathcal{O}(10^{-10})$ for a reheat temperature as low as $T_R \sim 10^9$ GeV. Interestingly, this co-

[1)] Observe that the magnitude of the produced $B - L$ strongly depends on the temperature — namely, larger $B - L$ should be produced in the patches where the plasma is hotter. In the cosmological context, this would lead to a situation where apart from the adiabatic density/temperature perturbations, there also emerge correlated isocurvature fluctuations with variable B and L which could be tested with the future data on the CMB anisotropies and large scale structure.

incidence with the upper bound from the thermal gravitino production, $T_R < 4 \times 10^9$ GeV or so [13], indicates that our scenario could also work in the context of supersymmetric theories.

The hidden sector may include coupling constants (e.g. gauge coupling constants of G') large enough to thermalize the hidden particles at a temperature T'. Once $K_1 < 1$, T' will be smaller than the parallel temperature of the ordinary system T. Obviously, the presence of the out-of-equilibrium hidden sector does not affect much the Big Bang Nucleosynthesis (BBN) epoch. Indeed, if the two sectors do not come into full thermal equilibrium at temperatures $T \sim T_R$ then, they evolve independently during the Universe expansion and approach the nucleosynthesis era with different temperatures. For $K_1 < 1$, the energy density transferred to the hidden sector will be crudely $\rho' \approx (8K_1/g_*)\,\rho$, where $g_*\,(\approx 100)$ is attained to the leptogenesis epoch. Thus, assuming that at the BBN epoch the shadow sector is dominated by relativistic degrees of freedom, we obtain an effective number of extra light neutrinos $\Delta N_\nu \approx K_1/2$.

It is worthwhile to stress that the leptogenesis mechanism we propose does not really rely on model dependent features of the hidden sector. They are however important and in principle testable to some extent if the hidden sector is to make up for the dark matter of the Universe. Namely, the same mechanism that produces the lepton number in the ordinary Universe, can also produce the lepton prime asymmetry in the hidden sector. The amount of this asymmetry will depend on the CP-violation parameter that replaces J in Eq. (7) and $\Delta\sigma'$ namely, $J' = \mathrm{Im\,Tr}\,[(h^\dagger h)\, g^{-2} (h'^\dagger h')\, g^{-1} (h'^\dagger h')^* g^{-1}]$. Then, if the shadow sector contains also some heavier particles of the lepton or baryon type, the shadow matter could provide a dark matter.

Depending on the gauge structure, field content and symmetry breaking scales in the hidden sector, one could have a shadow matter behaving as a cold, warm or self-interacting dark matter, or their combination. The possible marriage between dark matter and the lepto-baryogenesis mechanism is certainly an atractive feature of our scheme.

The interesting example can be provided by the mirror world, in which case the M-parity under the exchange $\phi \to \phi'^\dagger$, $l \to \bar{l}'$, etc., implies $h'_{ia} = h^*_{ia}$. In this case the CP-violation parameters are the same, $J' = J$. Then, one expects that $n_{B-L} = n'_{B-L}$ and the mirror baryon number density should be equal to the ordinary baryon density, $\Omega'_B = \Omega_B$.[2)] Generically, the mirror sector provides a sort of self-interacting dark matter, however, if it is significantly colder than the visible one, $T'/T < 0.3$ or so, the mirror photons decouple early and the mirror matter would behave as a cold dark matter as far as the large scale formation and the CMB anisotropies are concerned [10]. However, it would have interesting implications for the galaxy halo structures, microlensing experiments (Machos as mirror stars), the gravitational wave or the neutrino signals from the mirror supernovae explosions, etc.

References

[1] A. D. Sakharov, JETP Lett. **5** (1967) 24.
[2] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, Phys. Lett. **155B** (1985) 36.
[3] For a review, see A. D. Dolgov, Phys. Rep. **222** (1992) 309; A. Riotto and M. Trodden, Annu. Rev. Nucl. Part. Sci. **49** (1999) 35; W. Buchmüller and M. Plümacher, Int. J. Mod. Phys. A **15** (2000) 5047.
[4] M. Fukugita and T. Yanagida, Phys. Lett. B **174** (1986) 45; Phys. Rev. D **42** (1990) 1285.
[5] L. Bento and Z. Berezhiani, Phys. Rev. Lett. **87** (2001) 231304, hep-ph/0111116.

[2)] The mirror parity could be also spontaneously broken by the difference in weak scales $\langle\phi\rangle = v$ and $\langle\phi'\rangle = v'$, which would lead to somewhat different particle physics in the mrrror sector [8, 9], e.g. the mirror leptons and baryons could be heavier (or lighter) than the ordinary ones. But, as the mechanism only depends on the Yukawa constant pattern in (1), one still has $n_{B-L} = n'_{B-L}$, while $\Omega'_B \neq \Omega_B$.

[6] T. D. LI and C. N. YANG, Phys. Rev. **104**, 254 (1956); I. YU. KOBZAREV, L. B. OKUN, and I. YA. POMERANCHUK, Yad. Fiz. **3** (1966) 1154; S. BLINNIKOV and M. KHLOPOV, Sov. Astron. **27** (1983) 371; M. KHLOPOV et al., Sov. Astron. **35** (1991) 21; R. FOOT, H. LEW, and R. VOLKAS, Phys. Lett. B **272** (1991) 67; Mod. Phys. Lett. A **7** (1992) 2567; H. HODGES, Phys. Rev. D **47** (1993) 456; Z. SILAGADZE, Phys. At. Nucl. **60** (1997) 272; Z. BEREZHIANI, Phys. Lett. B **417** (1998) 287.
[7] R. FOOT and R. VOLKAS, Phys. Rev. D **52** (1995) 6595; For earlier work, R. FOOT, H. LEW, and R. VOLKAS, Mod. Phys. Lett. A **7** (1992) 2567.
[8] Z. BEREZHIANI and R. N. MOHAPATRA, Phys. Rev. D **52** (1995) 6607; For earlier work, E. AKHMEDOV, Z. BEREZHIANI, and G. SENJANOVIĆ, Phys. Rev. Lett. **69** (1992) 3013.
[9] Z. BEREZHIANI, A. DOLGOV, and R. N. MOHAPATRA, Phys. Lett. B **375** (1996) 26; Z. BEREZHIANI, Acta Phys. Pol. B **27** (1996) 1503; V. S. BEREZINSKY and A. VILENKIN, Phys. Rev. D **62** (2000) 083512.
[10] Z. BEREZHIANI, D. COMELLI, and F. L. VILLANTE, Phys. Lett. B **503** (2001) 362.
[11] E. KOLB, D. SECKEL, and M. TURNER, Nature **514** (1985) 415.
[12] M. A. LUTY, Phys. Rev. D **45** (1992) 455; M. PLÜMACHER, Z. Phys. C **74** (1997) 549; M. FLANZ and E. A. PASCHOS, Phys. Rev. D **58** (1998) 113009.
[13] J. ELLIS, A. LINDE, and D. V. NANOPOULOS, Phys. Lett. **118B** (1982) 59; D. V. NANOPOULOS, K. A. OLIVE, and M. SREDNICKI, ibid. **127B** (1983) 30.

Fortschr. Phys. **50** (2002) 5–7, 496–502

Time Asymmetric Quantum Theory and the Z-boson Mass and Width

ARNO R. BOHM[1)] and PIOTR KIELANOWSKI[2)]

[1)] Department of Physics, University of Texas at Austin, Austin, TX, USA
bohm@physics.utexas.edu2
[2)] Departamento de Física, CINVESTAV del IPN, Mexico
Piotr.Kielanowski@fis.cinvestav.mx

Abstract

A unified theory of resonance and decay phenomena is presented for the non-relativistic and the relativistic cases. It incorporates causality and time asymmetry. The relativistic quasistable particles are described by semigroup representations of the causal Poincaré transformations characterized by spin j and complex square mass $s_R = (M_R - i\Gamma_R/2)^2$, where Γ_R is the width of a relativistic Breit-Wigner scattering amplitude. The lifetime of the exponentially decaying state described by this semigroup representation $[j, s_R]$ is derived as $\tau = \hbar/\Gamma_R$. This provides a criterion for the appropriate definition of the mass and width of a relativistic resonance.

1. Introduction

Resonances and quasistable particles are characterized by some discrete quantum numbers such as (particle) species label and angular momentum j and additionally by two real numbers. These numbers are either resonance energy and width (E_R, Γ) or energy and (average) lifetime $(E_R, \tau = 1/R)$. In the relativistic case the elementary particles are labelled by either mass and width (M_R, Γ) or by mass and lifetime (M_R, τ) [1]. However the precise definition of these quantities when $\Gamma \neq 0$ has not been clear.

The width Γ is measured by the Breit-Wigner line-shape in the cross section

$$\sigma_j^{\mathrm{BW}}(E) \propto |a_j^{\mathrm{BW}}(E)|^2 = \left| \frac{r_\eta}{E - \left(E_R - i\,\dfrac{\Gamma}{2}\right)} \right|^2 \propto \frac{1}{(E - E_R)^2 + \left(\dfrac{\Gamma}{2}\right)^2}\,, \qquad \text{with} \quad 0 \le E < \infty\,. \tag{1}$$

The lifetime $\tau = \frac{1}{R}$ is measured by fitting the counting rate, $\frac{1}{N}\,\frac{\Delta N_\eta(t)}{\Delta t}$, for any decay product η to the exponential decay law for the partial decay rate $R_\eta(t)$

$$\frac{1}{N}\,\frac{\Delta N_\eta(t)}{\Delta t} \approx R_\eta(t) = R_\eta(0)\, e^{-t/\tau} = R_\eta(0)\, e^{-Rt}\,; \qquad R(t) = \sum_\eta R_\eta(t)\,. \tag{2}$$

The concept of width is thus based on the Breit-Wigner cross section (1) and the concept of lifetime is based on the exponential decay law (2). The decay rate $R \equiv \frac{1}{\tau}$ and the resonance width are experimentally as well as conceptually different quantities.

In non-relativistic quantum mechanics a common assumption is that

$$\frac{\hbar}{\Gamma} = \tau \equiv \frac{1}{R} \tag{3}$$

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

holds precisely, and many people think that resonances and decaying states are the same entities characterized alternatively by width Γ or by lifetime $\tau = \frac{1}{R}$. However this relation has been justified only in the Wigner-Weisskopf approximation [2], of which M. Levy wrote in 1959, "There does not exist ... a rigorous theory to which these various methods [Weisskopf-Wigner] can be considered as approximation" [3].

In relativistic physics a contrary opinion predominates. Based on the perturbation theoretical definition by the self energy of the propagators, resonances and decaying states are considered as complicated objects that cannot be described simply as an exponentially decaying state or as a state characterized by two numbers like (M_R, Γ).

The jth partial scattering amplitude in a relativistic resonance formation process, e.g. $e^+e^- \to Z \to \bar{f}f$, is a function of the invariant mass square $s = (p_1^\mu + p_2^\mu)^2 = (E_1^{\text{cm}} + E_2^{\text{cm}})^2$, where p_1^μ, p_2^μ are the momenta of two incoming (or outgoing) particles. One writes the amplitude of a resonance scattering process (with one resonance) as

$$a_j(s) = a_j^{\text{res}}(s) + B_j(s) \tag{4}$$

where $B_j(s)$ is the non-resonant background and $a_j^{\text{res}}(s)$ is to describe the contribution of the resonance per se. It is not possible to fix the functions $a_j^{\text{res}}(s)$ and $B_j(s)$ separately from the empirical data for $|a_j(s)|^2$, unless one has some theoretical arguments in favor of a particular functional form of $a_j^{\text{res}}(s)$ (or $B_j(s)$), e.g., the relativistic Breit-Wigner form of (18) below.

The questions then arise: 1.) Does it make sense to consider resonances as separate individual entities? 2.) Are resonances and quasistable particles different physical entities or are they only quantitatively different in the magnitude of Γ/E_R (for the non-relativistic case) or of Γ/M_R (for the relativistic case)?

If one wanted to differentiate between resonances and quasistable states one could try the following criterion: A resonance is that for which the width Γ can be measured, whereas a quasistable particle is that for which the lifetime τ can be measured. Another criterion one could use is: A resonance is what decays strongly, whereas a quasistable particle is what decays weakly. The second criterion would make the Z-boson a quasistable particle, the first criterion would make the Z-boson a resonance, and we want it to be an elementary particle.

Our experience suggests that stability or the value of lifetime is not a criterion of elementarity. A particle decays if it can and it remains stable if selection rules for exactly conserved quantum numbers (charge, baryon number) prevent it from decaying and a particle decays weakly, (with a long lifetime) if selection rules for approximately conserved quantities (hypercharge, charm) are involved.

Our phenomenological premise therefore is that resonances are not qualitatively different from quasistable states, and quasistable particles are not qualitatively different from stable particles, but only quantitatively by their values of widths Γ, and the lifetime-width relation (3) holds generally and exactly. Therefore, stable and quasistable states and resonances should be described in the same theoretical frame. This theory must relate an exact resonance amplitude characterized by (M_R, Γ) to an exponentially decaying state with the lifetime $\tau = \frac{1}{\Gamma}$. This theoretical frame needs also to include the (few) stable particles when $\Gamma = 0$.

2. From Resonance Phenomenology to Resonance Theory

In order to obtain a theory that unifies the resonances and decay phenomena we must introduce idealizations of the phenomenological concepts – as needs to be done for any

mathematical theory – and we must make modifications in the mathematics of the standard quantum theory.

The idealizations are that one extrapolates the Breit-Wigner in (1) to energy values on the negative real axis on the second Riemann sheet of the complex energy surface for the S-matrix, where the resonance poles are located at the complex energy value $z = z_R \equiv E_R - i\,\frac{\Gamma}{2}$. With this "ideal" Breit-Wigner scattering amplitude $a_j^{\mathrm{BW}}(E)$ one defines the "exact" Gamow vector ψ_j^{G} for the quasistable state

$$a_j^{\mathrm{BW}_i} = \frac{r_i}{E - z_{R_i}} \Longleftrightarrow \psi_j^G \equiv |z_{R_i}, j\cdots^-\rangle = \frac{i}{2\pi} \int_{-\infty_{II}}^{+\infty} \mathrm{d}E\, \frac{|E, j\cdots^-\rangle}{E - z_{R_i}} \qquad (5)$$
$$\text{for} \quad -\infty_{II} < E < \infty$$

where $-\infty_{II}$ indicates that $E < 0$ is on the second Riemann sheet.

There is of course always interference with non-resonant background and possibly with other resonances at z_{R_n}, but the i-th decaying state or resonance per se is defined by (5). The j-th partial wave amplitude $a_j(E)$ with a finite number of BW-resonances and non-resonant background $B(E)$ thus correspond to the state ϕ_j^+ which is a superposition of Gamow vectors and a vector ϕ^{bg} for background corresponding to $B(E)$:

$$a_j(E) = \sum_n \frac{r_n}{E - z_{R_n}} + B(E) \Longleftrightarrow \phi_j^+ = \sum_n |z_{R_n}, j\cdots^-\rangle\, c_n + \phi_j^{\mathrm{bg}}\,. \qquad (6)$$

The kets $|E^-\rangle$ under the integral of (5) (and similarly $|E^+\rangle$) are the out- (and in-) plane wave solutions of the Lippmann-Schwinger equation with $\mp i\epsilon$ in the denominator where $(-)$ refers to outgoing boundary conditions and $(+)$ to incoming boundary conditions. The $|E^-\rangle$ and the Gamow vectors (5), are not ordinary vectors in Hilbert space $\mathcal{H}$. They are, like the Dirac kets, basis vectors with respect to which the in-state vectors ϕ^- and the out-state vectors ϕ^+ can be expanded á la Dirac's basis vector expansion

$$\phi^\pm = \sum_j \int_0^\infty dE\, |E, j^\pm\rangle\, \langle {}^\pm E, j \mid \phi^\pm\rangle = \sum_j \phi_j^\pm\,. \qquad (7)$$

The kets suggest a mathematical modification of standard quantum mechanics, but there are additional reasons to revise the Hilbert space axiom: though one speaks of states ϕ^+ defined by a preparation apparatus (e.g., accelerator) and of observables $|\psi^-\rangle\,\langle\psi^-|$ defined by a registration apparatus (e.g., detector) as if they were different things, one mathematically identifies the set of the prepared in-states $\{\phi^+\}$ and the set of the registered observables $\{\psi^-\}$, and the Hilbert space axiom states that

$$\{\phi^+\} = \{\psi^-\} = \mathcal{H}\,. \qquad (8)$$

This is the cause of all kinds of trouble in the theory of scattering and decay (e.g., the impossibility of exponential decay). These troubles have not led to an impasse only because one ignored mathematical consequences of (8) and worked with undefined entities like Dirac and Lippmann-Schwinger kets.

We shall replace the Hilbert space axiom (8) with a new hypothesis which distinguishes between states and observables:

$$\begin{aligned} &\text{The prepared in-states are described by:} && \{\phi^+\} = \Phi_- \subset \mathcal{H} \subset \Phi_-^\times \\ &\text{and the registered observables (or out-states) by:} && \{\psi^-\} = \Phi_+ \subset \mathcal{H} \subset \Phi_+^\times \end{aligned} \qquad (9)$$

where $\Phi_\pm$ are Hardy spaces of the upper (+) and lower (−) complex energy half plane [4] and $\Phi_\pm^\times$ are their dual spaces of continuous antilinear functionals (kets).

The axiom (9) encapsulates the time asymmetry of causality: a state must be prepared by a time t_0 *before* the probability for an observable can be measured in it [5]. Further in addition to the apparatus-defined in-state vectors $\phi^+ \in \Phi_-$ and out-observable vectors $\psi^- \in \Phi_+$ the triplets (9) − called Gel'fand triplets [6] − also provide new generalized state vectors represented by the elements of the dual spaces $\Phi_\mp^\times$. These include familiar vectors like Dirac kets, but $\Phi_+^\times$ also provides us with the new states described by the Gamow vectors (5). These Gamow vectors are the central elements of the non-relativistic quantum theory of resonance scattering and decay. We review here only some results and refer for details to [7].

A vector $|F\rangle \in \Phi_+^\times$ is called a generalized eigenvector of the operator A if for some $\omega \in \mathbb{C}$

$$\langle A\psi \mid F\rangle = \langle \psi| A^\times |F\rangle = \omega\langle \psi \mid F\rangle \quad \text{for all} \quad \psi \in \Phi_+ \,. \tag{10}$$

This is usually written in Dirac's notation as $A^\times|F\rangle = \omega|F\rangle$ or $A|F\rangle = \omega|F\rangle$. Examples of generalized eigenvectors of $\Phi_+^\times$ are the Dirac-Lippmann-Schwinger kets:

$$H^\times|E^-\rangle = E\,|E^-\rangle\,, \qquad 0 \le E < \infty \tag{11}$$

fulfilling out-going boundary conditions (out-plane waves). Similarly the $|E^+\rangle$ (in-plane waves) are elements in $\Phi_-^\times$. (Due to the small imaginary part $\mp i\epsilon$ of the energy the $|E^\mp\rangle$ are not functionals on the Schwartz space). The space $\Phi_+^\times$ also contains, among others, the Gamow vectors defined in (5). Using the mathematical properties of the Hardy functions, one can prove [7] that $|E_R - i\Gamma/2^-\rangle$ of (5) is a generalized eigenvector of the self adjoint (semibounded) Hamiltonian H with a complex eigenvalue:

$$\langle \psi_\eta^-| H^\times |E_R - i\Gamma/2^-\rangle = (E_R - i\Gamma/2)\,\langle \psi_\eta^- \mid E_R - i\Gamma/2^-\rangle \quad \text{for all} \quad \psi_\eta^- \in \Phi_+^\times\,, \tag{12}$$

where $\psi_\eta^- \in \Phi_+$ represents the decay products η defined by the detector (out-states of a resonance scattering experiment or of a decay process). It is important to note that in order to prove (12) and (15) below, the integral in (5) has to extend to $-\infty_{II}$.

The expansion (6) (as an alternative to basis vector expansion (7)) of every state vector $\phi_j^+ \in \Phi_-$ with respect to the generalized eigenstates (12) plus a background integral is also a result of this theory (in the Weisskopf-Wigner approximation the $\phi^{\rm bg}$ is not present).

The probabilities to find in a state ϕ (6) the observable $|\psi\rangle\,\langle\psi|$ are in quantum theory given by the Born probabilities

$$\mathcal{P}_\psi(\phi(t)) = \mathrm{Tr}\,(|\phi(t)\rangle\,\langle\phi(t) \mid \psi\rangle\,\langle\psi|) = |\langle e^{iHt}\psi \mid \phi\rangle|^2 = |\langle \psi \mid e^{-iHt}\phi\rangle|^2\,. \tag{13}$$

We extend this interpretation to the elements of $\Phi_+^\times$ like the Gamow state $\psi^{\rm G} = |E_R - i\,\frac{\Gamma}{2}^-\rangle$. The probability per unit time to register as the detector counts ΔN_η in (2) the decay products described by the observable $|\psi_\eta^-\rangle\,\langle\psi_\eta^-|$ is − as a generalization of (13) − proportional to the absolute value square of the amplitude

$$\langle e^{iHt}\psi_\eta^- \mid E_R - i\Gamma/2^-\rangle = \langle \psi_\eta^-|\, e^{-iH^\times t}|E_R - i\Gamma/2^-\rangle \equiv \langle \psi_\eta^- \mid \psi_-^{\rm G}(t)\rangle\,. \tag{14}$$

This can be calculated for the Gamow vector (5) using the properties of the Hardy functions and one obtains [7]:

$$\langle \psi_\eta^- \mid \psi_-^{\rm G}(t)\rangle = e^{-iE_Rt}\,e^{-(\Gamma/2)t}\langle \psi_\eta^- \mid E_R - i\Gamma/2^-\rangle \quad \text{for} \quad t \ge \text{only}\,. \tag{15}$$

The conclusions from this mathematical consequence (15) are:

1. The Gamow state (5) with Breit-Wigner resonance width Γ decays exponentially in time since the counting rate of (2) $\Delta N_\eta(t)/\Delta t$ is proportional to:

$$|\langle \psi_\eta^- \mid \psi^{\mathrm{G}}(t)\rangle|^2 = e^{-\Gamma t}|\langle \psi_\eta^- \mid \psi^{\mathrm{G}}(0)\rangle|^2 . \tag{16}$$

   Therefore the lifetime τ of the Gamow state is given by $\tau = 1/\Gamma$ ($\hbar = 1$).
2. The time evolution of the Gamow vectors is asymmetric and it is given by the semigroup $U^\times(t) = e^{-iH^\times t}$, $t \geq 0$. Though this quantum mechanical irreversibility on the microphysical level appears at first shocking, it is consistent with the principle of causality and means that the Gamow state ψ^{G} must be prepared (at $t = t_0 = 0$) before an observable $|\psi_\eta^-\rangle\langle\psi_\eta^-|$ can be detected in it at $t > t_0$. Since one of the mathematical consequences of the Hilbert space axiom (8) is a unitary group evolution $U(t) = e^{-iHt}$, $-\infty < t < \infty$, the axiom (8) is in conflict with causality.

3. Relativistic Resonances and Causal Poincaré Semigroup Representations

For the relativistic resonance one starts with the resonance pole of the S-matrix at $s = s_R$ in the second sheet, but has no principle how to parameterize the complex pole position in terms of some real parameters that one could call resonance mass M and resonance width Γ. Some possible choices are [8]:

$$s_R = (M_Z^2 - iM_Z\Gamma_Z)\,(1 + (\Gamma_Z/M_Z)^2)^{-1} = \bar{M}_Z^2 - i\bar{M}_Z\bar{\Gamma}_Z = \left(M_R - i\,\frac{\Gamma_R}{2}\right)^2 . \tag{17}$$

Even after we fixed the ambiguity for $a_j^{\mathrm{res}}(s)$ in (4) by choosing the relativistic Breit-Wigner a_j^{BW} below, the ambiguity (17) in the choice of mass M and width Γ still remains. To select the physically motivated definition of M and Γ we associate to $a_j^{\mathrm{BW}}(s)$ a relativistic Gamow ket in analogy to (4)

$$a_j^{\mathrm{BW}} = \frac{r}{s - s_R} \Longleftrightarrow |[j,\, s_R]\, j_3,\, \hat{\boldsymbol{p}}\rangle = \frac{i}{2\pi} \int_{-\infty_{II}}^{\infty} ds\, |[j,\, s]\, j_3,\, \hat{\boldsymbol{p}}^-\rangle\, \frac{1}{s - s_R} . \tag{18}$$

$$-\infty < s < \infty \qquad -\infty < \hat{\boldsymbol{p}} < \infty , \qquad -j \leq j_3 \leq +j$$

Here the degeneracy label $\hat{\boldsymbol{p}}$ is the space component of the 4-velocity $\hat{\boldsymbol{p}} = \gamma \boldsymbol{v} = \boldsymbol{p}/\sqrt{s}$, $\gamma = \sqrt{1 - \boldsymbol{v}^2} = \hat{p}^0$.

Stable relativistic particles, the free asymptotic as well as the interacting [9], are according to Wigner described by representations of the Poincaré transformations; the same should hold for resonances. Indeed, one can show [10] that the relativistic Gamow kets $|[j,\, s_R]\, j_3,\, \hat{\boldsymbol{p}}^-\rangle$ span (i.e. are the basis vectors in the sense of (7)) an irreducible representation space $[j,\, s_R]$ of the Poincaré *semigroup* $\mathcal{P}_+$. This semigroup consists of all proper orthochronous Lorenz transformations and space-time translations into the forward cone (the causal Poincaré transformations)

$$\mathcal{P}_+ = \{(\Lambda,\, x) \mid \Lambda \in \overline{SO(3,\, 1)},\ \det \Lambda = +1,\ \Lambda_0^0 \geq 1,\ x^2 = t^2 - \boldsymbol{x}^2 \geq 0,\ t \geq 0\} . \tag{19}$$

The transformation of the Gamow kets $|[j,\, s_R]\, j_3, \hat{\boldsymbol{p}}^-\rangle$ under $(\Lambda,\, x) \in \mathcal{P}_+$ is very similar to the transformation of the Wigner basis vectors of the unitary representation $[j\,, M^2 > 0]$ and is given by [10]

$$U^\times(\Lambda,\, x)\, |[j,\, s_R]\, j_3,\, \hat{\mathbf{p}}^-\rangle = e^{-i\gamma\sqrt{s_R}(t-\mathbf{xv})} \sum_{j_3'} D^j_{j_3' j_3}(W(\Lambda^{-1},\, \hat{p}))\, |[j,\, s_R]\, j_3',\, \Lambda^{-1}\hat{\boldsymbol{p}}^-\rangle$$
$$\text{only for} \quad x^2 \geq 0 \quad \text{and} \quad t \geq 0 \tag{20}$$

where $e^{-i\gamma\sqrt{s_R}(t-\boldsymbol{xv})} = e^{-ipx}$ and $W(\Lambda,\, \hat{p})$ are the Wigner rotations.

The states at rest are eigenvectors of the Hamiltonian $H = P^0$ with complex eigenvalue $\sqrt{s_R}$ (the analogue of the non-relativistic case (12)):

$$H^\times |[j,\, s_R]\, j_3,\, \mathbf{0}^-\rangle = \sqrt{s_R}\, |[j,\, s_R]\, j_3,\, \mathbf{0}^-\rangle\,. \tag{21}$$

For the time evolution at rest $\boldsymbol{v} = 0$ one obtains from (20)

$$\begin{aligned}\psi^{\mathrm{G}}_{s_R}(t) &\equiv e^{-iH^\times t}|[j,\, s_R]\, j_3,\, \mathbf{0}^-\rangle = e^{-i\sqrt{s_R}t}|[j,\, s_R]\, j_3,\, \mathbf{0}^-\rangle \\ &= e^{-iM_R t} e^{-\Gamma_R t/2}|[j,\, s_R],\, \mathbf{0}^-\rangle \quad \text{for} \quad t \geq 0 \quad \text{only}\,.\end{aligned} \tag{22}$$

This is the analogue of (15) for the relativistic case and one draws from it the same conclusions 1 and 2 of Section 2. In particular the Born probability for measuring the out-observable $|\psi^-_\eta\rangle\, \langle\psi^-_\eta| = |f\bar{f}\rangle\, \langle f\bar{f}|$ in the quasistable state $|Z\rangle = \psi^{\mathrm{G}}_{[j,\, s_R]}(t)$ is proportional to

$$|\langle\psi^-_\eta \mid \psi^{\mathrm{G}}_{[j,\, s_R]}(t)\rangle|^2 = e^{-\Gamma_R t}|\langle\psi^-_\eta \mid \psi^{\mathrm{G}}_{[j,s_R]}(0)\rangle|^2 \quad \text{for} \quad t \geq 0 \quad \text{only}\,. \tag{23}$$

This means that the lifetime derived for the relativistic Gamow states (18) is $\tau = \hbar/\Gamma_R$, where Γ_R is the real parameter of the resonance pole position s_R in (17). The Poincaré semigroup transformation of the relativistic Gamow vectors (20), (22) therefore imply for (17) that $(M_R,\, \Gamma_R)$ is the right definition of mass and width of a relativistic resonance. For this mass the experimental line shape data for the Z-boson provide the value $M_R = (91.1626 \pm 0.0031)$ GeV which differs by 10 times the quoted experimental error from the standard value of M_Z listed in [1].

Wigner unitary group representation $[j,\, M^2]$ for a stable particle is somehow a special case $\Gamma_R \to 0$ of (20). However this is not the limit $i\Gamma \to i0$ of the semigroup representation $[j,\, (M - i\Gamma/2)^2]$ which leads to the relativistic Lippmann-Schwinger scattering states and these also have only semigroup transformations.

Acknowledgment

This paper is based on the talk at the Heisenberg Symposium, Sept 26–30, 2001; the kind hospitality of the Alexander von Humboldt Foundation is gratefully acknowledged. We are also grateful to the Welch Foundation for the additional support. We are much obliged to S. Wickramasekara for proofreading and correcting the manuscript.

References

[1] D. GROOM et al., *Review of Particle Properties*, Eur. Phys. J. **C15**, 1 (2000).
[2] V. WEISSKOPF and E. P. WIGNER, Z. f. Physik **63** (1930) 54; **65** (1930) 18. See also, for example, M. L. GOLDBERGER, K. M. WATSON, *Collision Theory* (Wiley, New York, 1964), chap 8.
[3] M. LEVY, Nuovo Cimento **13** (1959) 115.

[4] This means the energy wave functions $\langle {}^{-}E \mid \psi^{-} \rangle \equiv \psi^{-}(E) \in \mathcal{H}^2_{+} \cap S|_{\mathbb{R}^+}$ of an observable or out-state $\psi^{-} \in \Phi_{+}$ is a smooth rapidly decreasing function of E that can be analytically continued into the *upper* half of the complex energy plane, whereas the energy wave function $\langle {}^{+}E \mid \phi^{+} \rangle \equiv \phi^{+}(E) \in \mathcal{H}^2_{-} \cap S|_{\mathbb{R}^+}$ is a function that can be analytically continued into the lower half complex energy plane. In contrast, under the assumption (8), $\{\phi^{+}(E)\} = \{\psi^{-}(E)\} = L^2(\mathbb{R}^+) =$ Lebegues square integrable functions. The space S is the Schwartz space of smooth functions and hypothesis (9) says that the wave functions $\phi^{+}(E)$ and $\psi^{-}(E)$ are "slightly better" than Schwartz space functions but they are differently better for in-states $\phi^{+}(E)$ than for out-observables $\psi^{-}(E)$. The two Hardy spaces $\mathcal{H}^2_{+} \cap S|_{\mathbb{R}^+}$ and $\mathcal{H}^2_{-} \cap S|^{+}_{\mathbb{R}}$ as well as the Schwartz space S are dense in the Hilbert space L^2. As far as the energy distribution in the accelerator beam $|\phi^{+}(E)|^2$ and the energy resolution of the detector $|\psi^{-}(E)|^2$ are concerned one cannot distinguish between smooth Schwartz functions and smooth and analytic Hardy functions. But the consequences for these two choices differ significantly, for the Schwartz space one obtains time symmetric group evolution and for the Hardy spaces (9) one obtains causal time asymmetric semigroup evolution. Mathematical results see P. L. Duren, $\mathcal{H}^p$ *Spaces*, (Academic Press, New York, 1970) or Appendix A2 of [7] below. One can prove that the two triplets of spaces in (9) form two Rigged Hilbert spaces with the same Hilbert space $\mathcal{H}$, M. Gadella, J. Math. Phys. **24** (1983) 1462.

[5] A. Bohm, I. Antoniou, and P. Kielanowski, Phys. Lett. **A189** (1994) 552.

[6] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions*, Volume 4, (Academic Press, New York, 1964); A. Bohm, M. Gadella, *Dirac Kets, Gamow Vectors and Gel'fand Triplets*, (Springer Verlag, New York 1989).

[7] A. Bohm, S. Maxson, M. Loewe, and M. Gadella, Physica **A 236** (1997) 485; A. Bohm, J. Math. Phys. **22** (1981) 2813; Lett. Math. Phys. **3** (1978) 455.

[8] A. Sirlin, Phys. Lett., **B267** (1991) 240; R. G. Stuart, Phys. Rev. **D56** (1997) 1515; S. Willenbrock and G. Valencia, Phys. Lett. **B259** (1991) 373. The definition (M_Z, Γ_Z) gives the standard values quoted in [1].

[9] S. Weinberg, *The Quantum Theory of Fields*, vol. 1, chap. 3, (Cambridge University Press, 1995).

[10] A. Bohm, H. Kaldass, and S. Wickramasekara, *Relativistic Gamow Vectors*, Preprint, The University of Texas at Austin.

Fortschr. Phys. **50** (2002) 5–7, 503–511

The Heisenberg Matrix Formulation of Quantum Field Theory*

STANLEY J. BRODSKY

Stanford Linear Accelerator Center
Stanford University, Stanford, CA 94309, USA
sjbth@slac.stanford.edu

Abstract

Heisenberg's matrix formulation of quantum mechanics can be generalized to relativistic systems by evolving in light-front time $\tau = t + z/c$. The spectrum and wavefunctions of bound states, such as hadrons in quantum chromodynamics, can be obtained from matrix diagonalization of the light-front Hamiltonian on a finite dimensional light-front Fock basis defined using periodic boundary conditions in x^- and $x_\perp$. This method, discretized light-cone quantization (DLCQ), preserves the frame-independence of the front form even at finite resolution and particle number. Light-front quantization can also be used in the Hamiltonian form to construct an event generator for high energy physics reactions at the amplitude level. The light-front partition function, summed over exponentially-weighted light-front energies, has simple boost properties which may be useful for studies in heavy ion collisions. I also review recent work which shows that the structure functions measured in deep inelastic lepton scattering are affected by final-state rescattering, thus modifying their connection to light-front probability distributions. In particular, the shadowing of nuclear structure functions is due to destructive interference effects from leading-twist diffraction of the virtual photon, physics not included in the nuclear light-front wavefunctions.

1. Introduction

One of the challenges of relativistic quantum field theory is to compute the wavefunctions of bound states, such as the amplitudes which determine the quark and gluon substructure of hadrons in quantum chromodynamics. However, any extension of the Heisenberg-Schrödinger formulation of quantum mechanics $H|\psi\rangle = i\,\frac{\partial}{\partial t}\,|\psi\rangle = E|\psi\rangle$ to the relativistic domain has to confront seemingly intractable hurdles: (1) quantum fluctuations preclude finite particle-number wavefunction representations; (2) the charged particles arising from the quantum fluctuations of the vacuum contribute to the matrix element of currents – thus knowledge of the wavefunctions alone is insufficient to determine observables; and (3) the boost of an equal-time wavefunction from one Lorentz frame to another not only changes particle number, but is as complicated a dynamical problem as solving for the wavefunction itself.

In 1949, Dirac [1] made the remarkable observation that ordinary "instant" time t is not the only possible evolution parameter. In fact, evolution in "light-front" time $\tau = t + z/c = x^+$ has extraordinary advantages for relativistic systems, stemming from the fact that a subset of the Lorentz boost operators becomes purely kinematical. In fact, the Fock-state representation of bound states defined at equal light-front time, i.e., along the light-front, provides wavefunctions of fixed particle number which are independent of the eigenstate's four-momentum P^μ. Furthermore, quantum fluctuations of the vacuum are absent if one uses light-front time to quantize the system, so that matrix elements such as the

* Work supported by Department of Energy contract DE–AC03–76SF00515.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Fig. 1. Dirac's three forms of Hamiltonian dynamics. From Ref. [2]

electromagnetic form factors only depend on the currents of the constituents described by the light-front wavefunctions.

In Dirac's "Front Form", the generator of light-front time translations is $P^- = i\,\frac{\partial}{\partial\tau}$. Boundary conditions are set on the transverse plane labeled by $x_\perp$ and $x^- = z - ct$. See Fig. 1. Given the Lagrangian of a quantum field theory, P^- can be constructed as an operator on the Fock basis, the eigenstates of the free theory. (This method is also called "light-cone" quantization in the literature.) Since each particle in the Fock basis is on its mass shell, $k^- \equiv k^0 - k^3 = \frac{k_\perp^2 + m^2}{k^+}$, and its energy $k^0 = \frac{1}{2}(k^+ + k^-)$ is positive, only particles with positive momenta $k^+ \equiv k^0 + k^3 \geq 0$ can occur in the Fock basis. Since the total plus momentum $P^+ = \sum_n k_n^+$ is conserved, the light-front vacuum cannot have any particle content. The operator $H_{\rm LC} = P^+P^- - P_\perp^2$, the "light-front Hamiltonian", is frame-independent.

The Heisenberg equation on the light-front is

$$H_{\rm LC}|\Psi\rangle = M^2|\Psi\rangle\,. \tag{1}$$

This can in principle be solved by diagonalizing the matrix $\langle n|\, H_{\rm LC}\, |m\rangle$ on the free Fock basis: [2]

$$\sum_m \langle n|\, H_{\rm LC}\, |m\rangle\, \langle m \mid \psi\rangle = M^2 \langle n \mid \Psi\rangle\,. \tag{2}$$

The eigenvalues $\{M^2\}$ of $H_{\rm LC} = H_{\rm LC}^0 + V_{\rm LC}$ give the squared invariant masses of the bound and continuum spectrum of the theory. For example, the light-cone gauge interaction terms of QCD which are important for a meson are illustrated in Fig. 2. The projections $\{\langle n \mid \Psi\rangle\}$ of the eigensolution on the n-particle Fock states provide the light-front wavefunctions. Thus solving a quantum field theory is equivalent to solving a coupled many-body quantum mechanical problem:

$$\left[M^2 - \sum_{i=1}^{n} \frac{m_{\perp i}^2}{x_i}\right] \psi_n = \sum_{n'} \int \langle n|\, V_{\rm LC}\, |n'\rangle\, \psi_{n'} \tag{3}$$

Fig. 2. The front-form matrix of QCD interactions in light-cone gauge. A maximum of 8 constituents in a meson are shown. From Ref. [2]

where the convolution and sum is understood over the Fock number, transverse momenta, plus momenta, and helicity of the intermediate states. Here $m_\perp^2 = m^2 + k_\perp^2$.

In QCD, the wavefunction of a hadron describes its composition in terms of the momenta and spin projections of quark and gluon constituents. For example, the eigensolution of a negatively-charged meson QCD, projected on its color-singlet $B = 0$, $Q = -1$, $J_z = 0$ eigenstates $\{|n\rangle\}$ of the free Hamiltonian $H_{\rm LC}^{\rm QCD}(g = 0)$ at fixed $\tau = t - z/c$ has the expansion:

$$|\Psi_M;\, P^+,\, \vec{P}_\perp,\, \lambda\rangle = \sum_{n\geq 2,\lambda_i} \int \prod_{i=1}^{n} \frac{d^2k_{\perp i}\, dx_i}{\sqrt{x_i}\, 16\pi^3}\; 16\pi^3 \delta\left(1 - \sum_j^n x_j\right) \delta^{(2)}\left(\sum_\ell^n \vec{k}_{\perp\ell}\right)$$
$$|n;\, x_i P^+,\, x_i \vec{P}_\perp + \vec{k}_{\perp i},\, \lambda_i\rangle\; \psi_{n/M}(x_i,\, \vec{k}_{\perp i},\, \lambda_i)\,. \tag{4}$$

The set of light-front Fock state wavefunctions $\{\psi_{n/M}\}$ represent the ensemble of quark and gluon states possible when the meson is intercepted at the light-front. The light-front momentum fractions $x_i = k_i^+/P_\pi^+ = (k^0 + k_i^z)/(P^0 + P^z)$ with $\sum_{i=1}^n x_i = 1$ and $\vec{k}_{\perp i}$ with $\sum_{i=1}^n \vec{k}_{\perp i} = \vec{0}_\perp$ represent the relative momentum coordinates of the QCD constituents and are independent of the total momentum of the state. The actual physical transverse momenta are $\vec{p}_{\perp i} = x_i \vec{P}_\perp + \vec{k}_{\perp i}$. The λ_i label the light-front spin S^z projections of the quarks and gluons along the quantization z direction. The spinors of the light-front formalism automatically incorporate the Melosh-Wigner rotation. The physical gluon polarization vectors $\epsilon^\mu\,(k, \lambda = \pm 1)$ are specified in light-cone gauge by the conditions $k \cdot \epsilon = 0$, $\eta \cdot \epsilon = \epsilon^+ = 0$. The parton degrees of freedom are thus all physical; there are no ghost or negative metric states. A detailed derivation of light-front quantization of non-Abelian gauge theory in light-cone gauge is given in Ref. [3]. Explicit examples of light-front wavefunctions in QED are given in Ref. [4].

Angular momentum has simplifying features in the light-front formalism since the projection J_z is kinematical and conserved. Each light-front Fock wavefunction satisfies the angular momentum sum rule: $J^z = \sum_{i=1}^{n} S_i^z + \sum_{j=1}^{n-1} l_j^z$. The sum over S_i^z represents the contribution of the intrinsic spins of the n Fock state constituents. The sum over orbital angular momenta

$$l_j^z = -i\left(k_j^1 \frac{\partial}{\partial k_j^2} - k_j^2 \frac{\partial}{\partial k_j^1}\right) \tag{5}$$

derives from the $n-1$ relative momenta. This excludes the contribution to the orbital angular momentum due to the motion of the center of mass, which is not an intrinsic property of the hadron [4]. The numerator structure of the light-front wavefunctions is in large part determined by the angular momentum constraints.

The most important feature of light-front Fock wavefunctions $\psi_{n/p}(x_i, \vec{k}_{\perp i}, \lambda_i)$ is the fact they are Lorentz invariant functions of the relative coordinates, independent of the bound state's physical momentum $P^+ = P^0 + P^z$, and $P_\perp$ [5]. The light-front wavefunctions represent the ensembles of states possible when the hadron is intercepted by a light-front at fixed $\tau = t + z/c$. The light-front representation thus provide a frame-independent, quantum-mechanical representation of a hadron at the amplitude level, capable of encoding its multiquark, hidden-color and gluon momentum, helicity, and flavor correlations in the form of universal process-independent hadron wavefunctions.

If one imposes periodic boundary conditions in $x^- = t + z/c$, then the plus momenta become discrete: $k_i^+ = \frac{2\pi}{L} n_i$, $P^+ = \frac{2\pi}{L} K$, where $\sum_i n_i = K$ [6, 7]. For a given "harmonic resolution" K, there are only a finite number of ways positive integers n_i can sum to a positive integer K. Thus at a given K, the dimension of the resulting light-front Fock state representation of the bound state is rendered finite without violating Lorentz invariance. The eigensolutions of a quantum field theory, both the bound states and continuum solutions, can then be found by numerically diagonalizing a frame-independent light-front Hamiltonian $H_{\rm LC}$ on a finite and discrete momentum-space Fock basis. Solving a quantum field theory at fixed light-front time τ thus can be formulated as a relativistic extension of Heisenberg's matrix mechanics. The continuum limit is reached for $K \to \infty$. This formulation of the non-perturbative light-front quantization problem is called "discretized light-front quantization" (DLCQ) [7]. Lattice gauge theory has also been used to calculate the pion light-front wavefunction [8].

The DLCQ method has been used extensively for solving one-space and one-time theories [2], including applications to supersymmetric quantum field theories [9] and specific tests of the Maldacena conjecture [10]. There has been progress in systematically developing the computation and renormalization methods needed to make DLCQ viable for QCD in physical spacetime. For example, John Hiller, Gary McCartor, and I [11] have shown how DLCQ can be used to solve $3+1$ theories despite the large numbers of degrees of freedom needed to enumerate the Fock basis. A key feature of our work is the introduction of Pauli Villars fields to regulate the UV divergences and perform renormalization while preserving the frame-independence of the theory. A recent application of DLCQ to a $3+1$ quantum field theory with Yukawa interactions is given in Ref. [11]. Representative plots of the one-boson one-fermion light-front Fock wavefunction of the lowest mass fermion solution of the Yukawa $(3+1)$ theory showing spin correlations and the presence of non-zero orbital angular momentum are shown in Fig. 3.

There has also been important progress using the transverse lattice, essentially a combination of DLCQ in $1+1$ dimensions together with a lattice in the transverse dimensions [2−14]. One can also define a truncated theory by eliminating the higher Fock states in favor of an effective potential [15]. Spontaneous symmetry breaking and other nonperturba-

Fig. 3. DLCQ results for the one-boson one-fermion wavefunction in a fermion system with parallel and antiparallel fermion helicity, as a function of longitudinal momentum fraction y and one transverse momentum component q_x in the $q_y = 0$ plane. The parameter values for the DLCQ resolution are $K = 29$, $N_\perp = 7$. Further details are given in Ref. [11].

tive effects associated with the instant-time vacuum are hidden in dynamical or constrained zero modes on the light-front. An introduction is given in Refs. [16, 17].

2. General Features of Light-front Wavefunctions

The maximum of a light-front wavefunction occurs when the invariant mass of the partons is minimal; i.e., when all particles have equal rapidity and are all at rest in the rest frame. In fact, Dae Sung Hwang and I [18] have noted that one can rewrite the wavefunction in the form:

$$\psi_n = \frac{\Gamma_n}{M^2\left[\sum_{i=1}^{n} \frac{(x_i - \hat{x}_i)^2}{x_i} + \delta^2\right]} \tag{6}$$

where $x_i = \hat{x}_i \equiv m_{\perp i}/\sum_{i=1}^{n} m_{\perp i}$ is the condition for minimal rapidity differences of the constituents. The key parameter is $M^2 - \sum_{i=1}^{n} m_{\perp i}^2/\hat{x}_i \equiv -M^2\delta^2$. One can interpret $\delta^2 \simeq 2\epsilon/M$ where $\epsilon = \sum_{i=1}^{n} m_{\perp i} - M$ is the effective binding energy. This form shows that the wavefunction is a quadratic form around its maximum, and that the width of the distribution in $(x_i - \hat{x}_i)^2$ (where the wavefunction falls to half of its maximum) is controlled by $x_i\delta^2$ and the transverse momenta $k_{\perp_i}$. Note also that the heaviest particles tend to have the largest $\hat{x}_i$, and thus the largest momentum fraction of the particles in the Fock state, a feature familiar from the intrinsic charm model. For example, the b quark has the largest momentum fraction at small $k_\perp$ in the B meson's valence light-front wavefunction, but the distribution spreads out to an asymptotically symmetric distribution around $x_b \sim 1/2$ when $k_\perp \gg m_b^2$.

The fall-off the light-front wavefunctions at large $k_\perp$ and $x \to 1$ is dictated by QCD perturbation theory since the state is far-off the light-front energy shell. This leads to counting rule behavior for the quark and gluon distributions at $x \to 1$. Notice that $x \to 1$ corresponds to $k^z \to -\infty$ for any constituent with nonzero mass or transverse momentum.

The above discussion suggests that an approximate form for the hadron light-front wavefunctions could be constructed through variational principles and by minimizing the expectation value of $H_{\rm LC}^{\rm QCD}$.

3. A Light-front Event Amplitude Generator

The light-front formalism can be used as an "event amplitude generator" for high energy physics reactions where each particle's final state is completely labeled in momentum, helicity, and phase. The application of the light-front time evolution operator P^- to an initial state systematically generates the tree and virtual loop graphs of the T-matrix in light-front time-ordered perturbation theory in light-front gauge. The loop integrals only involve integrations over the momenta of physical quanta and physical phase space $\prod d^2k_{\perp i}\, dk_i^+$. Renormalized amplitudes can be explicitly constructed by subtracting from the divergent loops amplitudes with nearly identical integrands corresponding to the contribution of the relevant mass and coupling counter terms (the "alternating denominator method") [19]. The natural renormalization scheme to use for defining the coupling in the event amplitude generator is a physical effective charge such as the pinch scheme [20]. The argument of the coupling is then unambiguous [21]. The DLCQ boundary conditions can be used to discretize the phase space and limit the number of contributing intermediate states without violating Lorentz invariance. Since one avoids dimensional regularization and nonphysical ghost degrees of freedom, this method of generating events at the amplitude level could provide a simple but powerful tool for simulating events both in QCD and the Standard Model.

4. The Light-front Partition Function

In the usual treatment of classical thermodynamics, one considers an ensemble of particles $n = 1, 2, \ldots N$ which have energies $\{E_n\}$ at a given "instant" time t. The partition function is defined as $Z = \sum_n \exp -\frac{E_n}{kT}$. Similarly, in quantum mechanics, one defines a quantum-statistical partition function as $Z = tr \exp -\beta H$ which sums over the exponentiated-weighted energy eigenvalues of the system.

In the case of relativistic systems, it is natural to characterize the system at a given light-front time $\tau = t + z/c$; i.e., one determines the state of each particle in the ensemble as its encounters the light-front. Thus we can define a light-front partition function

$$Z_{\mathrm{LC}} = \sum_n \exp - \frac{p_n^-}{kT_{\mathrm{LC}}}$$

by summing over the particles' light-front energies $p^- = p^0 - p^z = \frac{p_\perp^2 + m^2}{p^+}$. The total momentum is $P^+ = \sum p_n^+$, $\vec{P}_\perp = \sum_n \vec{p}_{\perp n}$, and the total mass is defined from $P^+P^-P_\perp^2 = M^2$. The product $\frac{M}{P^-} T_{\mathrm{LC}}$ is boost invariant. In the center of mass frame where $\vec{P} = 0$ and thus $P^+ = P^- = M$. It is also possible to consistently impose boundary conditions at fixed $x^- = z - ct$ and $x_\perp$, as in DLCQ. The momenta p_n^+, $\vec{p}_{\perp n}$ then become discrete. The corresponding light-front quantum-statistical partition function is $Z = tr \exp -\beta_{\mathrm{LC}} H_{\mathrm{LC}}$ where $H_{\mathrm{LC}} = i \frac{\partial}{\partial \tau}$ is the light-front Hamiltonian.

For non-relativistic systems the light-front partition function reduces to the standard definition. However, the light-front partition function should be advantageous for analyzing relativistic systems such as heavy ion collisions, since, like true rapidity, $y = \ln \frac{p^+}{P^+}$, light-front variables have simple behavior under Lorentz boosts. The light-front formalism also takes into account the point that a phase transition does not occur simultaneously in t, but propagates through the system with a finite wave velocity.

5. Light-front Wavefunctions and QCD Phenomenology

There have been extensive applications of light-front wavefunctions to QCD phenomenology [22]; for example, form factors [23] and the handbag contribution to deeply virtual Compton scattering $\gamma^* p \to \gamma p$ can be expressed as overlaps of the light-front wavefunctions [24, 25]; quark and gluon distributions are light-front wavefunction probabilities. The distributions measured in the diffractive dissociation of hadrons are computed from transverse derivatives of the light-front wavefunctions. Progress in measuring the basic parameters of electroweak interactions and CP violation will require a quantitative understanding of the dynamics and phase structure of B decays at the amplitude level. The light-front Fock representation is specially advantageous in the study of exclusive B decays. For example, Dae Sung Hwang [26] and I have derived an exact frame-independent representation of decay matrix elements such as $B \to D\ell\bar{\nu}$ from the overlap of $n' = n$ parton-number conserving wavefunctions and the overlap of wavefunctions with $n' = n - 2$ from the annihilation of a quark-antiquark pair in the initial wavefunction.

One can also express the matrix elements of the energy momentum tensor as overlap integrals of the light-front wavefunctions [4]. An important consistency check of any relativistic formalism is to verify the vanishing of the anomalous gravito-magnetic moment $B(0)$, the spin-flip matrix element of the graviton coupling and analog of the anomalous magnetic moment $F_2(0)$. For example, at one-loop order in QED, $B_f(0) = \frac{\alpha}{3\pi}$ for the electron when the graviton interacts with the fermion line, and $B_\gamma(0) = -\frac{\alpha}{3\pi}$ when the graviton interacts with the exchanged photon. The vanishing of $B(0)$ can be shown to be exact for bound or elementary systems in the light-front formalism [4], in agreement with the equivalence principle [27–29].

6. Structure Functions are Not Parton Distributions

The quark and gluon distributions of hadrons can be defined from the probability measures of the light-front wavefunctions. For example, the quark distribution in a hadron H is

$$P_{q/H}(x_{Bj}, Q^2) = \sum_n \int^{k_{i\perp}^2 < Q^2} \left[\prod_i dx_i \, d^2k_{\perp i}\right] |\psi_{n/H}(x_i, k_{\perp i})|^2 \sum_{j=q} \delta(x_{Bj} - x_j) \,. \tag{7}$$

It has been conventional to identify the leading-twist structure functions $F_i(x, Q^2)$ measured in deep inelastic lepton scattering with the light-front probability distributions. For example, in the parton model, $F_2(x, Q^2) = \sum_q e_q^2 P_{q/H}(x, Q^2)$. However, Paul Hoyer, Nils Marchal, Stephane Peigne, Francesco Sannino, and I [30] have recently shown that the leading-twist contribution to deep inelastic scattering is affected by diffractive rescattering of a quark in the target, a coherent effect which is not included in the light-front wavefunctions, even in light-cone gauge. The gluon propagator in light-cone gauge $A^+ = 0$ is singular:

$$d_{\rm LC}^{\mu\nu}(k) = \frac{i}{k^2 + i\varepsilon}\left[-g^{\mu\nu} + \frac{n^\mu k^\nu + k^\mu n^\nu}{n \cdot k}\right] \tag{8}$$

has a pole at $k^+ \equiv n \cdot k = 0$, which has to be defined by an analytic prescription such as the Mandelstam-Liebbrandt prescription [31]. In final-state scattering involving on-shell intermediate states, the exchanged momentum k^+ is of $O(1/\nu)$ in the target rest frame, which enhances the second term of the light-cone gauge propagator. This enhancement allows rescattering to contribute at leading twist even in LC gauge.

Thus diffractive contributions to the deep inelastic scattering $\gamma^* p \to Xp'$ cross sections, which leave the target intact, contribute at leading twist to deep inelastic scattering. Diffractive events resolve the quark-gluon structure of the virtual photon, not the quark-gluon structure of the target, and thus they give contributions to structure functions which are not target parton probabilities. Our analysis of deep inelastic scattering $\gamma^*(q)\, p \to X$, when interpreted in frames with $q^+ > 0$, also supports the color dipole description of deep inelastic lepton scattering at small x_{bj}. For example, in the case of the aligned-jet configurations, one can understand $\sigma_T(\gamma^* p)$ at high energies as due to the coherent color gauge interactions of the incoming quark-pair state of the photon interacting, first coherently and finally incoherently, in the target.

The distinction between structure functions and target parton probabilities is also implied by the Glauber-Gribov picture of nuclear shadowing [32–35]. In this framework, shadowing arises from interference between complex rescattering amplitudes involving on-shell intermediate states. In contrast, the wave function of a stable target is strictly real since it does not have on energy-shell configurations. Thus nuclear shadowing is not a property of the light-front wavefunctions of a nuclear target; rather, it involves the total dynamics of the γ^*-nucleus collision. A strictly probabilistic interpretation of the deep inelastic cross section cross section is thus precluded.

Acknowledgment

I thank Professor Harald Fritzsch and the Von Humboldt Foundation for inviting me to the Heisenberg Symposium. Much of the work reported here was done in collaboration with others, especially, Markus Diehl, John Hiller, Paul Hoyer, Dae Sung Hwang, Nils Marchal, Bo-Qiang Ma, Gary McCartor, Hans Christian Pauli, Stephane Peigne, Francesco Sannino, Ivan Schmidt, and Prem Srivastava. This work was supported by the Department of Energy under contract number DE-AC03-76SF00515.

References

[1] P. A. Dirac, Rev. Mod. Phys. **21** (1949) 392.
[2] For a review and references, see S. J. Brodsky, H. C. Pauli, and S. S. Pinsky, Phys. Rept. **301** (1998) 299 [arXiv:hep-ph/9705477].
[3] P. P. Srivastava and S. J. Brodsky, Phys. Rev. D **64** (2001) 045006 [arXiv:hep-ph/0011372].
[4] S. J. Brodsky, D. S. Hwang, B. Ma, and I. Schmidt, Nucl. Phys. B **593** (2001) 311 [hep-th/0003082].
[5] G. P. Lepage and S. J. Brodsky, Phys. Rev. D **22** (1980) 2157.
[6] T. Maskawa and K. Yamawaki, *Method of Quantization*, Prog. Theor. Phys. **56** (1976) 270.
[7] H. C. Pauli and S. J. Brodsky, Phys. Rev. D **32** (1985) 1993.
[8] A. Abada, P. Boucaud, G. Herdoiza, J. P. Leroy, J. Micheli, O. Pene, and J. Rodriguez-Quintero, Phys. Rev. D **64** (2001) 074511 [arXiv:hep-ph/0105221].
[9] Y. Matsumura, N. Sakai, and T. Sakai, Phys. Rev. D **52** (1995) 2446 [arXiv:hep-th/9504150].
[10] J. R. Hiller, S. Pinsky, and U. Trittmann, *supersymmetric Yang-Mills theory*, Phys. Rev. D **64** (2001) 105027 [arXiv:hep-th/0106193].
[11] S. J. Brodsky, J. R. Hiller, and G. McCartor, hep-ph/0107038.
[12] W. A. Bardeen, R. B. Pearson, and E. Rabinovici, Phys. Rev. D **21** (1980) 1037.
[13] S. Dalley, Phys. Rev. D **64** (2001) 036006 [arXiv:hep-ph/0101318].
[14] M. Burkardt and S. K. Seal, arXiv:hep-ph/0105109.
[15] H. C. Pauli, arXiv:hep-ph/0111040.
[16] G. McCartor, in *Proc. of New Nonperturbative Methods and Quantization of the Light Cone,* Les Houches, France, 24 Feb–7 Mar 1997.
[17] K. Yamawaki, arXiv:hep-th/9802037.
[18] S. J. Brodsky and D.-S. Hwang, in preparation.
[19] S. J. Brodsky, R. Roskies, and R. Suaya, *Momentum Frame*, Phys. Rev. D **8** (1973) 4574.

[20] J. M. CORNWALL and J. PAPAVASSILIOU, Phys. Rev. D **40** (1989) 3474.
[21] S. J. BRODSKY and H. J. LU, Phys. Rev. D **51** (1995) 3652 [arXiv:hep-ph/9405218].
[22] S. J. BRODSKY, arXiv:hep-ph/0102051.
[23] S. J. BRODSKY and S. D. DRELL, Phys. Rev. D **22** (1980) 2236.
[24] S. J. BRODSKY, M. DIEHL, and D. S. HWANG, Nucl. Phys. B **596** (2001) 99 [arXiv:hep-ph/0009254].
[25] M. DIEHL, T. FELDMANN, R. JAKOB, and P. KROLL, Nucl. Phys. B **596** (2001) 33 [Erratum-ibid. B **605** (2001) 647] [arXiv:hep-ph/0009255].
[26] S. J. BRODSKY and D. S. HWANG, Nucl. Phys. B **543** (1999) 239 [hep-ph/9806358].
[27] L. OKUN and I. YU. KOBSAREV, ZhETF, 43 (1962) 1904 (English translation: JETP 16 (1963) 1343); L. OKUN, in proceedings of the 4th International Conference on Elementary Particles, Heidelberg, Germany (1967), edited by H. Filthuth, North-Holland (1968).
[28] I. YU. KOBSAREV and V. I. ZAKHAROV, Ann. Phys. **60** (1970) 448.
[29] O. V. TERYAEV, hep-ph/9904376 (1999).
[30] S. J. BRODSKY, P. HOYER, N. MARCHAL, S. PEIGNE, and F. SANNINO, hep-ph/0104291.
[31] G. LEIBBRANDT, Rev. Mod. Phys. **59** (1987) 1067.
[32] V. N. GRIBOV, Sov. Phys. JETP **29** (1969) 483 [Zh. Eksp. Teor. Fiz. **56** (1969) 892].
[33] S. J. BRODSKY and J. PUMPLIN, Phys. Rev. **182** (1969) 1794.
[34] S. J. BRODSKY and H. J. LU, Phys. Rev. Lett. **64** (1990) 1342.
[35] G. PILLER and W. WEISE, Phys. Rept. **330** (2000) 1 [hep-ph/9908230].

Fortschr. Phys. **50** (2002) 5–7, 512–517

Nonlocal Structure of Higher Spin Fields

ANTON Z. CAPRI

Department of Physics and Theoretical Physics Institute, University of Alberta,
Edmonton, AB T6G 2J1, Canada
capr@phys.ualberta.ca

Abstract

We first review the problem of superluminal propagation or even loss of propagation for higher spin fields minimally coupled to an external electromagnetic field, discovered by Velo and Zwanziger in the 1960's. Then, we show that these fields have a nonlocal structure. This indicates that these fields have internal structure and should not be considered as elementary. The main tool in this investigation is the use of a Peirce decomposition of the algebra of the matrices occuring in the first order wave equations for these fields.

1. Background

Higher spin fields, that is fields with spin $s \geq 1$, attacted a lot of interest in the 1960's due to the proliferation of new particles that were discovered. There were also several indications that there was something sick about these fields. One of the earliest such indications is in a paper by Johnson and Sudarshan [2] in which they showed that if a massive spin 3/2 field ψ is coupled to an electromagnetic field, the anticommutator of ψ and its adjoint $\psi^\dagger$ becomes indefinite for a sufficiently strong magnetic field. The actual result was that for

$$\left(\frac{2e}{3m^2}\right)^2 \vec{B}^2 > 1$$

the anticommutator is negative. This leads to the contradiction that the intrinsically non-negative opertor $\psi^\dagger\psi + \psi\psi^\dagger$ becomes negative for the condition quoted.

Eight years later Velo and Zwanziger [1] showed that the same condition causes a massive spin 3/2 field to cease propagating. However, even for smaller values of the external magnetic field $\vec{B}$ the spin 3/2 field ψ propages at superluminal velocities. Their analysis was based on finding the characteristic surfaces for the field equations.

The situation is not improved by introducing additional couplings. In fact, in general, the onset of problems may occur for even weaker magnetic fields as shown in a series of papers by Capri and Shamaly as well as Capri and Kobes. [3] In these papers we not only considered minimal coupling of spin 1 and spin 3/2 fields to an external electromagnetic field, but also looked for possible compensating effects due to local couplings involving all possible Pauli terms as well as self-couplings and couplings to spin 1/2 fields.

In all of these studies one simply writes out the field equations, eliminates the constraints and examines the characteristic equation.

A different line of attack was to use equations that do not describe fields of pure spin and so do not involve constraints. Examples are the Bhabha equations which in addtion to spin s contain also spin $s-1$, $s-2, \ldots$. In this case causal propagation is maintained but,

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

as shown by Pais and Uhlenbeck, [4] an indefinite metric is required. As far as I know, this line of investigation has not been carried much further.

This is a very brief and incomplete summary of the problem encountered with higher spin fields. It was realized that the fact that constraints were required to get pure spin equations was at the heart of the problem, but just what this implied was not clear.

2. The Peirce Decomposition and Nonlocal Structure

I shall now describe what is meant by a Peirce decomposition and how it is applied. To begin with, we always start with the free higher spin wave equations in the form

$$(\beta^{\mu} p_{\mu} - m)\, \psi = 0 \tag{1}$$

where the β-matrices satisfy an algebra of the form

$$(\beta \cdot p)^n = p^2 (\beta \cdot p)^{n-2} \tag{2}$$

for any four-vector p. The value of n depends on the spin. Thus, for spin 1/2 the value of n is 2, although so-called "barnacled" equations with higher values of n are also possible. For spin 1 we have $n = 3$ and for spin 3/2 we have $n = 4$. The smallest possible value of n for a field of spin s is $2s + 1$. Again higher values of n are possible in all these cases, but this does matter as we soon see. Clearly for $n > 2$ the matrix β^0 is singular, as can be seen by choosing

$$p = (1, 0, 0, 0)$$

to get

$$\beta_0^n = \beta_0^{n-2}\,. \tag{3}$$

This implies the existence of constraints. By performing a Peirce decomposition one is able to eliminate the constraints and extract the pure spin s part say ψ_s of the field ψ and get a propagation equation of the form

$$i\,\frac{\partial \psi_s}{\partial t} = H_s \psi_s\,. \tag{4}$$

The evolution operator H_s contains no time derivatives and evolves the pure spin s part of the field.

To see how this Peirce decomposition relates to constraints we begin with the Duffin-Kemmer algebra for the spin 0 and spin 1 fields and obtain an equation of the form 4. In this case we have $n = 3$. The field ψ satisfying 1 can now be split into

$$\psi = \psi_s + \psi_r \tag{5}$$

where ψ_s is the pure spin $s = (0, 1)$ field and ψ_r are the auxiliary or constrained fields. These auxiliary fields have to be eliminated to give a pure spin s field. The projection operator onto spin s is β_0^2 and onto the auxiliary fields is $1 - \beta_0^2$. That this last operator leads to constraints is evident from applying it to 1 to get

$$(1 - \beta_0^2)\, [\vec{\beta} \cdot \vec{p} + m]\, \psi = 0\,. \tag{6}$$

This is clearly an equation of constraint since it does not involve any time derivatives and must therefore hold for all times. So, writing

$$\psi_s = \beta_0^2 \psi\,, \qquad \psi_r = (1-\beta_0^2)\,\psi \tag{7}$$

we get after a little algebra and the use of the constraint equation 6 that

$$\psi_r = -\frac{1}{m}\,(\beta_0 \not{\vec{p}} + \not{\vec{p}}\beta_0)\,\psi_s\,. \tag{8}$$

Here we have written $\not{\vec{p}}$ for $\vec{\beta}\cdot\vec{p}$. Then, multiplying 1 from the left by β_0 we get

$$\begin{aligned}\beta_0^2 i\,\frac{\partial}{\partial t}(\psi_s+\psi_r) &= i\,\frac{\partial \psi_s}{\partial t}\\ &= \beta_0(\not{\vec{p}} + m)\,(\psi_s+\psi_r)\\ &= \beta_0(\not{\vec{p}} + m)\,\psi_s - \frac{\beta_0 \not{\vec{p}}}{m}\,(\beta_0\not{\vec{p}} + \not{\vec{p}}\beta_0)\,\psi_s\,.\end{aligned} \tag{9}$$

So, finally we have

$$i\,\frac{\partial \psi_s}{\partial t} = \beta_0\left[\not{\vec{p}} + m - \frac{\not{\vec{p}}}{m}\,(\beta_0\not{\vec{p}} + \not{\vec{p}}/\beta_0\right]\psi_s\,. \tag{10}$$

This is now the propagation equation for pure spin $s = (0,\,1)$ with the constraints eliminated and is of the form 4.

The general idea behind the specific Peirce decomposition that we just performed is to take any matrix A in the algebra of β-matrices and split it into submatrices as follows

$$A = PAP + PA(1-P) + (1-P)\,AP + (1-P)\,A(1-P) \tag{11}$$

where, P is the projection operator onto spin s. In the case above we had

$$P = \beta_0^2\,. \tag{12}$$

After such a decomposition the wave equation 1 may be rewritten in the form

$$i\,\frac{\partial \psi_s}{\partial t} = H_s \psi_s\,, \tag{13}$$

where H_s is a matrix of the form PAP. In a representation adapted to this decomposition H_s then assumes a matrix form

$$H_s = \begin{pmatrix} H' & 0 \\ 0 & 0 \end{pmatrix}. \tag{14}$$

3. Spin 1 with Quadrupole Coupling

To further illustrate the technique we now apply it to a spin 1 field with a Pauli term corresponding to an electromagnetic quadrupole coupling. In this case, Velo and Zwanziger

showed that acausal propagation occurs. We begin with the coupled Proca equations in the form

$$p^\mu U^\nu - p^\nu U^\mu = mG^{\mu\nu}\,, \qquad p_\nu G^{\nu\mu} - mU^\mu = \frac{q}{m}\,(Q^\mu_{\nu\lambda} - Q^\mu_{\lambda\nu})\,U^\lambda\,, \tag{15}$$

where

$$Q_{\lambda\nu\mu} = \partial_\lambda F_{\nu\mu} \tag{16}$$

and $F_{\nu\mu}$ is the electromagnetic field tensor. We are leaving out the minimal coupling terms since they neither produce difficulties nor help to eliminate them. If we write the equations in terms of the Duffin-Kemmer algebra they read

$$[\beta_\mu + (q/m)\,\sigma^{\nu\lambda}(Q_{\nu\mu\lambda} + Q_{\lambda\nu\mu})]\,p^\mu\psi = m\psi\,. \tag{17}$$

Here,

$$\sigma^{\nu\lambda} = \beta^\nu\beta^\lambda + ag^{\nu\lambda} \tag{18}$$

with a a constant.

For simplicity we choose a static external electromagnetic field with only magnetic field components. This means that

$$Q_{j0k} = Q_{jk0} = Q_{0\mu\nu} = 0 \tag{19}$$

and leads to great simplifications in the computations, yet allows us to illustrate the essential points. From the algebra of the β-matrices we find

$$\beta_0^2\sigma^{ij} = \sigma^{ij}\beta_0^2\,, \qquad \beta_0^2\sigma^{0j} = \sigma^{0j}\,, \qquad \beta_0^2\sigma^{j0} = 0\,, \qquad \sigma^{0j}\beta_0^2 = 0\,. \tag{20}$$

We now multiply 17 by $1 - \beta_0^2$ and rearrange terms using the identies just listed to express ψ_r in terms of ψ_s.

$$[m - (q/m)\,(Q_{ikj} + Q_{jik})\,\sigma^{ij}p^k)]\,\psi_r = [(\beta_0\vec{\boldsymbol{p}} + \vec{\boldsymbol{p}}\beta_0) - (q/m)\,Q_{ik0}\beta_i\beta_0 p^k]\,\psi_s\,. \tag{21}$$

To finish expressing ψ_r in terms of ψ_s requires the nonlocal operator

$$[1 - (q/m^2)\,(Q_{ikj} + Q_{jik})\,\sigma^{ij}p^k)]^{-1}\,.$$

This shows that the evolution operator H_s in the equation

$$i\,\frac{\partial}{\partial t}\,\psi_s = H_s\psi_s \tag{22}$$

is nonlocal and makes it clear that this field has a form factor for such a magnetic quadrupole interaction. Thus, this field has internal structure.

4. Peirce Decomposition of Spin 3/2 Field

As stated earlier, the minimal algebra for the β-matrices in this case is

$$(\beta_\mu p^\mu)^4 = p^2(\beta_\mu p^\mu)^2\,. \tag{23}$$

To proceed we write out a few special cases of this equation.

$$\beta_0^3\vec{p} + \beta_0^2\vec{p}\beta_0 + \beta_0\vec{p}\beta_0^2 + \vec{p}\beta_0^3 = \beta_0\vec{p} + \vec{p}\beta_0\,, \tag{24}$$

$$\beta_0^2\vec{p}^2 + \beta_0\vec{p}\beta_0\vec{p} + \vec{p}\beta_0^2\vec{p} + \beta_0\vec{p}^2\beta_0 + \vec{p}\beta_0\vec{p}\beta_0 + \vec{p}^2\beta_0^2 = \vec{p}^2 - \beta_0^2\vec{p}^2\,. \tag{25}$$

Next, we proceed as for the spin 1 equation and multiply the field equation 1 by

$$(1-\beta_0^2)\,\beta_0$$

to get the constraint equation.

$$(\beta_0 - \beta_0^3)\,(\vec{p} + m)\,\psi = 0\,. \tag{26}$$

We also define

$$\psi_{3/2} = \beta_0^2\psi\,, \qquad \psi_{1/2} = (1-\beta_0^2)\psi\,. \tag{27}$$

After some boring algebra we then find that we can eliminate $\psi_{1/2}$ and get the evolution equation for the pure spin 3/2 field.

$$p^0\psi_{3/2} = \beta_0^3(\vec{p} + m)\,\{1 + [1 - (\vec{\alpha}\cdot\vec{p}/m)^2]^{-1}\,Q\}\,\psi_{3/2}\,, \tag{28}$$

where

$$\vec{\alpha}\cdot\vec{p} = (\beta_0\vec{p} + \vec{p}\beta_0)\,\beta_0 \tag{29}$$

and

$$Q = m^{-2}(1-\beta_0^2)\,[\vec{p}^2 + (\beta_0\vec{p} + \vec{p}\beta_0)\,(\vec{p} - m)\,\beta_0]\,. \tag{30}$$

So, we have achieved our aim and written

$$p^0\psi_{3/2} = H_{3/2}\psi_{3/2}\,. \tag{31}$$

However, the evolution operator $H_{3/2}$ is clearly nonlocal. A somewhat tedious calculation now shows that the 16×16 matrix $H_{3/2}$ has 8 eigenvalues 0 and 4 eigenvalues $+(\vec{p}^2+m^2)^{1/2}$ and 4 eigenvalues $-(\vec{p}^2+m^2)^{1/2}$ as is to be expected for a spin 3/2 field. To show that the field really has internal structure one needs to do coupling to some other field. The result involves tedious computations that can not be reduced to an explicit form for the resultant evolution operator $H_{3/2}$. However, the results do show that $H_{3/2}$ is again nonlocal. It thus appears that the source of all the difficulties in the higher spin fields arises from the fact that they have internal structure and are not elementary. This gives a physical reason for all the difficulties that have plagued higher spin theories.

References

[1] K. JOHNSON and E. C. G. SUDARSHAN, Ann. Phys. (N.Y.) **13** (1961) 126.
[2] G. VELO and D. ZWANZIGER, Phys. Rev. **186** 1337 (1969); **188** (1969) 2218.
See also A. S. WIGHTMAN, *Troubles in the external-field problem for invariant wave equations, in Coral Gables Conference on Fundamental Interactions at High Energy,* Vol 4, (New York 1971).
[3] A. SHAMALY and A. Z. CAPRI, Ann. Phys. (N.Y.) **74** (1972) 503;
A. Z. CAPRI and A. SHAMALY, Can. J. Phys. **52** (1974) 919;
ibid **54**, 1089, (1976);
A. Z. CAPRI and R. L. KOBES, Phys. Rev. **D22** (1980) 1967.
[4] A. PAIS and G. E. UHLENBECK, Phys. Rev. **79** (1950) 145.

Fortschr. Phys. **50** (2002) 5–7, 518–524

Fundamental Constants at High Energy[1,2)]

HARALD FRITZSCH

Theory Division, CERN, CH-1211 Geneva 23, Switzerland
and Ludwig-Maximilians-Universität München, Sektion Physik
Theresienstr. 37, 80333 München, Germany
fritzsch@mppmn.mpg.de

Abstract

The progress of Particle Physics is closely linked to the progress in the understanding of the fundamental constants, like the finestructure constant, the mass of the electron or nucleon, or the electroweak mixing angle. The relation between the 18 fundamental constants of the Standard Model and the elementary units used in other fields like quantum optics or solid state physics is far from trivial and will be discussed. Relations between the various constants might exist, providing signals for the physics beyond the Standard Model. Recent observations in astrophysics indicate a slight time variation of the finestructure constant. If true, it has profound implications for many particle and nuclear physics phenomena. In particular the nuclear mass scale should change in time, a phenomenon which could be observed in the laboratory using advanced methods of quantum optics.

The Standard Model of particle physics [1] is a superposition of QCD for the strong interactions and the electroweak gauge theory for the electromagnetic and weak interactions. It gives a nearly complete description of all observed phenomena in atomic, nuclear and particle physics. It is only nearly complete, since certain phenomena, among them the increasing evidence for neutrino oscillations and the dominance of matter as compared to antimatter in the universe, cannot be described within the framework of the Standard Model. The major drawback exhibited by the Standard Model is the fact that a large number of constants, in particular many mass parameters, have to be adjusted according to the experimental measurements and cannot be predicted within the theory. Many theoreticians in particle physics believe for this reason that the Standard Model must be regarded only as a first step towards a more complete understanding, and that in the future it will be embedded in a larger and more complete theoretical framework.

In this talk I cannot offer a solution of the problem of the many fundamental constants, however I shall give a critical overview and describe possible directions one might go. Let me first point out that the Standard Model, which is based on a description of the fundamental forces by the theoretical framework of quantum theory, is a description of the local laws of nature. It does not say anything about the possible boundary conditions which are imposed on these local laws from the outside, in particular from cosmological boundary conditions. It could well be that at least certain fundamental constants are subject to such boundary conditions. If this is the case, it would not help to find a description of these constants by embedding the Standard Model in a more fundamental theory. Some of the constants appearing in the Standard Model could indeed be cosmic accidents, i.e. quantities which were fluctuating wildly at the time of the creation of the universe, but were frozen immediately afterwards. These constants could be called "frozen accidents" [2]. Other con-

1) Partially supported by VW-Stiftung Hannover (I-77495).
2) Invited talk given at the Heisenberg symposium of the Alexander von Humboldt-Foundation (Bamberg, September 2001).

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

stants could eventually be determined by dynamical laws, which go beyond the laws of the Standard Model. Such constants could indeed be calculated in a future theory. It might also be that some of the elementary constants are not constants at all, but are slowly changing in time. Such ideas were pioneered by Dirac [3], who once proposed that the gravitational constant is a function of the cosmic time. This idea, however, faded away during the course of the last century, since no time variation of the gravitational constant has been observed. Nevertheless one should keep in mind that not only the gravitational constant, but also other constants appearing in the Standard Model might turn out to be slowly varying functions of time.

The constant which plays the most significant role in atomic physics is the finestructure constant α introduced by Arnold Sommerfeld in 1917. Sommerfeld noted that the finestructure of the atomic levels was determined by a dimensionless number whose value today is given by:

$$\alpha^{-1} = 137.03599976(50)\,. \tag{1}$$

Numerically it turned out that the inverse of α is quite close to be an integer number. Sommerfeld himself did not indulge in philosophical speculations about the nature of α. Such speculations were started in the thirties by Arthur Eddington, who speculated about an intrinsic relation between the inverse of α and the total number of different charged objects [4]. He introduced a specific counting of these objects, including their spin and came up with the astonishing number 136, which in Eddington's view was sufficiently close to the observed number of 137.

Shortly after Eddington Werner Heisenberg came up with a proposal to describe α by an algebraic formula, which works up to an accuracy of 10^{-4}:

$$\alpha = 2^{-4}3^{-3}\pi\,, \tag{2}$$

Wyler found in 1971 an algebraic formula based on group-theory arguments which works up to the level of 10^{-6} [4]:

$$\alpha = \frac{9}{8\pi^4}\left(\frac{\pi^5}{2^4 5!}\right)^{\frac{1}{4}}. \tag{3}$$

Today we must interpret such attemps as useless. In particular in the underlying theory of Quantum Electrodynamics the actual value of the coupling constant changes if one changes the reference point, i.e. changing the energy scale.

The theory of Quantum Electrodynamics is still the most successful theory in science. It brings together Electrodynamics, Quantum Mechanics and Special Relativity. QED is a renormalizable theory and has been tested thus far up to a level of one in 10 Million. Physical quantities like the anomalous magnetic moment of the electron can be calculated in terms of powers of α, and α has been determined that way to a high degree of accuracy.

The value of α describes the coupling strength of electrodynamics at distances which are large compared to the Compton wavelength of the electron. At smaller distances α changes slowly. In pure QED, i.e. in the presence of only the photon field and the electron, the effective value of α is given in momentum space by:

$$\alpha_{\text{eff}}\left(q^2\right) = \frac{\alpha}{1 - \dfrac{\alpha}{3\pi}\ln\left(\dfrac{-q^2}{Am_e^2}\right)}\,, \tag{4}$$

$$A = \exp\left(5/3\right), \qquad -q^2 > 0\,. \tag{5}$$

The infinitesimal change of α is dictated by the renormalization group equation:

$$\frac{d}{d\ln(q/M)}\, e(q;\, e_r) = \beta(e)\,, \qquad e(M;\, e_r) = e_r \tag{6}$$

(M: renormalization point) .

At high energy not only virtual electron–positron pairs contribute, but also myon pairs, τ pairs, quark-antiquark-pairs etc. At smaller distances α is becoming larger. This effect can also be seen directly in the experiments. At LEP the effective value of α given at an energy scale of 91 GeV (mass of Z-boson) is:

$$\alpha(M_z) \cong (127.5)^{-1}\,. \tag{7}$$

The renormalization group requires that the strength of the electromagnetic coupling increases at increasing energy and eventually reaches a point, where perturbation theory breaks down. In pure QED, i.e. in the theory of a photon field, interacting only with one charged fermion, the electron, the critical energy (Landau singularity) is extremely high and far above the energy scale given by gravity. Of course, this theory is not realistic, since in the real world there are 3 charged leptons and 6 charged quarks, and therefore the increase of the coupling constant happens at a much higher rate. Already at energies, which were reached by the LEP-Accelerator, of the order of 200 GeV, the associated value of the finestructure constant is more than 10% higher than at low energy. In any case this signifies that one should not attach a specific fundamental meaning to the numerical value of the finestructure constant.

The fact that in particle physics all phenomena can be described in terms of a number of fundamental constants is, of course, directly related to the Standard Model. The Standard Model is a superposition of the quark–gluon gauge field theory (QCD) and of the electroweak gauge theory, based on the gauge group $SU(2) \times SU(1)$. The Standard Model is not merely a gauge field theory, like many others. It aims at a complete description of all particle physics phenomena, and it is extremely successful in doing so. Let me remind you that the experimental program, using the LEP Accelerator at CERN, came to an end in the year 2000. The outcome of the research done with the LEP Accelerator constitutes a triumph in particular for the standard electroweak gauge theory. The parameters of the theory, most notably the mass of the Z-boson and the coupling parameters have been determined with an impressive accuracy.

In the Standard Model the masses of the weak bosons are generated by the coupling of the boson fields to the thus far hypothetical Higgs-field. The model requires the existence of a Higgs particle, whose mass, if it exists, is one of the basic parameters of the Standard Model. The present limit of the mass of the Higgs particle, given by the CERN experiments, is about 110 GeV.

In the Standard Model the number of basic parameters is 18, including the three gauge coupling constants. Thirteen of these constants are directly related to the fermion masses.

The 18 basic constants of the Standard Model can be listed as follows:

$$\begin{array}{ccccccccc} m_e & m_u & m_d; & m_\mu & m_c & m_s; & m_\tau & m_t & m_b \\ & & \theta_u & \theta_d & \theta & \delta & & & \\ & & & M_w & M_h & & & & \\ & & & \alpha & \alpha_s & \alpha_w & & & \end{array}\,. \tag{8}$$

The nature of most of the fundamental constants seems to be intrinsically related to the generation of masses. One of the peculiar features of the Standard Model is the fact that

two different types of mass generation mechanisms seem to operate. On the one hand the masses of the weak bosons and of the fermions are given by the coupling of these fields to the scalar boson. On the other hand the masses of the nucleons and moreover the masses of all nuclei are predominantly due to a dynamical mass generation. The generation of mass in QCD could be described as "mass from no-mass". In lowest order the behavior of the QCD coupling constant α_s is given by:

$$\alpha_s(q^2) = \frac{2\pi}{b_0 \ln\left(\frac{q}{\Lambda}\right)}\,, \qquad b_0 = 11 - \frac{2}{3}\, n_f \tag{9}$$

(n_f: number of flavors, $q = \sqrt{q^2}$), Λ: scale parameter).

Formally the coupling constant becomes infinite, if the energy scale involved approaches the critical value Λ. Through "dimensional transmutation" the functional dependence of the coupling constant on the energy leads to the appearance of a mass scale. In the limit in which all the quark masses are set to zero, the masses of the bound states (nucleons etc.) are proportional to Λ. Using the experimental value $\alpha_s(M_z) = 0.1184 \pm 0.0031$, as given by the LEP-experiments, one obtains $\Lambda = 213 + 38/-35$ MeV [5].

In principle the nucleon mass, one of the fundamental parameters of atomic physics, can be calculated in terms of Λ, if the effects of the quark masses are neglected. Thus far an exact determination of the nucleon mass in terms of Λ has not been possible, due to the complexity of the calculations, e.g. within the approach of lattice QCD. However, simpler quantities, for example the pion decay constant, have been calculated with success. The pion decay constant is given by the matrix element of the axial vector current:

$$\langle 0|\, A_\mu\, |\pi\rangle = i p_\mu F_\pi\,. \tag{10}$$

It has the dimension of mass. The theoretical result is [6]:

$$F_\pi/\Lambda = 0.56 \pm 0.05 \tag{11}$$

while the experiments give:

$$F_\pi/\Lambda = 0.62 \pm 0.10\,. \tag{12}$$

The good agreement between experiment and theory indicates that QCD is able to describe not only perturbative features of the strong interaction physics, but also indicates that in the future one might be able to calculate more complicated quantities like the nucleon mass with a good precision.

One must keep in mind that the quark masses are non−zero and will influence the numerical value of the nucleon mass. Unfortunately the uncertainties imposed by our ignorance about the contribution of the quark mass terms to the nucleon mass is high. The matrix element of the non−strange quark mass term, the σ-term, is only poorly known:

$$\langle p|\, m_u \bar{u}u + m_d \bar{d}d\, |p\rangle \approx 45\ \text{MeV} \pm 25\%\,. \tag{13}$$

Also the mass term of the strange quarks plays an important role for the nucleon mass. Typical estimates give:

$$\langle p|\, m_s \bar{s}s\, |p\rangle \sim 40\ \text{MeV}\,, \tag{14}$$

with an error which is not less than about 50%. Note that the $u-d$ contribution and the s-contribution to the nucleon mass are of similar order.

In general we can say that the nucleon mass is a dual entity. The dominant part of it (about 90%) is due to the dynamical mechanism offered uniquely by QCD, i.e. due to the field energy of the confined quarks and gluons. About 10% of the nucleon mass arises due to the nonvanishing masses of the u, d and s-quarks. The strange part of this contribution is about as large as the non-strange part. Moreover there is a small electromagnetic term of about 2% (of order $\alpha \cdot \Lambda$).

In particle physics we are confronted at the beginning of the new millenium with the unsolved problem of the spectrum of the lepton and quark masses. The mass eigenvalues show a remarkable mass hierarchy. As an example I mention the masses of the charge 2/3 quarks: $u : c : t \approx 5 : 1150 : 174000$ (the masses are given in MeV).

Most of the quark masses and all of the lepton masses are much smaller than the mass scale of the weak interactions given in the Standard Model by the vacuum expectation value v of the scalar field $v \approx 246$ GeV. Only the mass of the t-quark is of the same order of magnitude as the weak interaction mass scale. It is remarkable that the mass of the t-quark is within the allowed errors equal to the vacuum expectation value divided by $\sqrt{2}$:

$$v/\sqrt{2} \approx 174\,\mathrm{GeV} = m_t\,. \tag{15}$$

Such a mass relation might be a hint towards an interpretation as a Clebsch-Gordan-relation, related to an internal symmetry. However, no such symmetry has been identified thus far, and the question remains whether the relation above is an accident or not. Another interesting feature of the quark mass spectrum is the fact that for each charge channel the mass ratios seem to be universal:

$$\begin{aligned} m_d : m_s &= m_s : m_b\,, \\ m_u : m_c &= m_c : m_t\,. \end{aligned} \tag{16}$$

Again a deeper understanding of this scaling feature is missing.

In the Standard Model the transitions between the various families of quarks (and possibly also of the leptons) arise because the states entering the weak interactions are not identical to the mass eigenstates. The transition strengths are in general given by complex amplitudes, however, it is well-known that the multitude of the flavor transitions is given by 3 mixing angles, which I like to denote by θ, θ_u and θ_d, and a complex phase parameter δ. All transition strengths can be expressed in terms of these four parameters. For example, the Cabbibo transition between the up quark and the strange quark, often denoted as V_{12}, is given in the complex plane by θ_u, θ_d and the phase δ, which is the relative phase between the two angles

$$V_{12} \cong \Theta_u - \Theta_d\, e^{-i\delta}\,. \tag{17}$$

In the complex plane V_{12}, θ_u and θ_d form a triangle, which is congruent to the so called "unitarity triangle". Since the absolute value of V_{12} is given with very high precision, a good determination of the angles θ_u and θ_d would allow us to determine the shape of the triangle. Thus far only one of the angles of the triangle, denoted usually by β, has been determined by the experiments, since it is related to the observed strength of the CP violation in the decay of B-mesons. However, the allowed ranges are still large: $\sin 2\beta \approx 0.45 \ldots 1$.

One can show that the angle Θ_u, which describes essentially the mixing between the u- and the c-quarks is essentially 0 in the limit $m_u \to 0$. Likewise Θ_d is essentially 0 for $m_d \to 0$. In simple models for the mass generation based on symmetries beyond the Standard Model one finds simple relations between the mass eigenvalues and the mixing angles [7].

$$\tan \Theta_u \approx \sqrt{\frac{m_u}{m_c}} \qquad \tan \Theta_d \approx \sqrt{\frac{m_d}{m_s}} \,. \tag{18}$$

If these relations hold the unitarity triangle is determined with rather high precision. In particular the angle α which is equal to the phase parameter δ is essentially $\frac{\pi}{2}$, which would imply the CP violation in nature to be maximal [7]. Relations between the mass eigenvalues and the mixing angles are of high interest since such relations would reduce the number of fundamental parameters of the Standard Model. It is conceivable that all the 3 mixing angles as well as the phase parameters are fixed by such relations, although the exact structure of the relations is still unclear. Further relations, in particular mass relations between the leptons and quarks and relations among the coupling constants can be obtained if the Standard Model is viewed as a low energy limit of a grand unified theory, based on large symmetry groups, e.g. $SO(10)$.

The coupling constants of the gauge groups $SU(3)$ and $SU(2)$ both decrease at high energies, while the coupling constant of the $U(1)$-sector decreases. At very high energies they become of comparable magnitude. If one uses the observed magnitudes of the coupling constants, one finds that they do converge at high energies, however do not meet exactly at one point. The energy scale were they approach each other is about 10^{15} GeV. If the gauge groups of the Standard Model are indeed subgroups of a bigger symmetry group and if the symmetry breaking of the grand unified theory happens at one specified energy, one would expect that the three coupling constants meet at one particular point on the energy scale. One but not the only possibility to reach a convergence of the coupling constants is to introduce supersymmetry. In supersymmetry for each fermion of the Standard Model a corresponding boson is introduced, and each boson of the Standard Model is accompanied by a corresponding fermion. Since the partners of the fermions and bosons have not been observed, their masses must be sufficiently high, typically above about 200 GeV.

The supersymmetric partners of the fermions and bosons do contribute to the renormalization of the coupling constants. If one choses a symmetry breaking for the supersymmetry at an energy scale of the order of about one TeV, one finds that the two coupling constants converge at an energy scale of $1.5 \cdot 10^{16}$ GeV. In such a theory the three different coupling constants for the strong, electromagnetic and weak interactions are fixed just by one coupling constant, the unified gauge coupling constant at high energies.

In grand unified theories one typically finds also a parellelism between the quarks of charge $-\frac{1}{3}$ and the charged leptons, implying that at the grand unified energy scale the mass of the charged lepton and of the corresponding quark should be equal, e.g.: $m_b = m_\tau$. Indeed, such a relation works quite well for the $\mathrm{b} - \tau$-system. The observed fact that the b-quark mass is about a factor of 3 larger than the τ-lepton mass comes from the renormalization effect, mostly due to the QCD interaction. Similar relations between the μ-mass and the s-quark mass or between the electron mass and the d-mass do not seem to hold. The relations between these masses must be more complicated than the one given above. Nevertheless it is conceivable that such mass relations exist.

Taking into account the relations between the fundamental parameters of the Standard Model discussed above, one may ask how many independent parameters might finally remain. The most optimistic answer is 7: one coupling constant for the unified interaction, the 3 masses of the charged leptons and the 3 masses of the charge $\frac{2}{3}$ quarks. Note that the t-mass is supposed to describe also the energy scale of the weak interaction, fixing at the same time the $W - Z$-masses and the mass of the scalar boson. The strength of the unified coupling constant can be related to the scale parameter Λ of QCD.

Gravity does not have a place in the Standard Model. The gravitational interaction is characterized by a critical energy scale, the Planck-mass: $\Lambda_p = 1.221047 \times 10^{19}$ GeV. The interplay between the gravitational interaction and the Standard Model gauge interactions

can be described by dimensionless ratios like:

$$\Lambda/\Lambda_p = 0.17 \cdot 10^{-19}\,. \tag{19}$$

Such ratios are not fixed by the considerations made above. They are candidates for a time variation on a cosmological time scale. Furthermore the unified coupling constant might also depend on the time, implying that the finestructure constant α becomes a function of time.

Recently one has found indications that the finestructure constant was perhaps smaller in the past. Studying the finestructure of various lines in distant gas clouds, one found [8]: $\delta\alpha/\alpha = (-0.72 \pm 0.18) \cdot 10^{-5}$. It remains to be seen whether this effect holds up in future observations. If the finestructure constant α undergoes a cosmological shift, one should expect similar shifts also to affect the strong interaction coupling constant, in other words the Λ-scale, which in turn would affect the magnitude of the proton mass. Furthermore the neutron-proton-mass difference, which has an electromagnetic contribution, would change. This would be important for the nucleosynthesis of the light elements. Thus far a systematic study of all effects of a time variation of α has not been carried out.

If one takes the idea of a Grand Unification of the gauge forces seriously, a time shift of α would make sense only if the unified coupling constant undergoes a time shift as well. But this would imply, as recently pointed out [9] that the QCD scale Λ would also change in time. As a result the nucleon mass and all nuclear mass scales would be time-dependent. Grand unification implies that the relative change of the nucleon mass is about two orders of magnitude larger than the relative change of α. Using advanced methods of quantum optics, a time variation of α and of the nuclear mass scale could be observed by monitoring the atomic finestructure and molecular rotational or vibrational frequencies [9].

In general we can expect that the problem of the fundamental constants of the Standard Model will remain in the focus of research in particle physics at least for the ten years. The success of any new direction in theoretical research should be measured in terms of its power to make predictions about the fundamental constants or about relations among them.

References

[1] See e.g.: C. Quigg, Gauge Theories of the Strong, Weak and Electromagnetic Interactions (Advanced Book Classics) (Reprint Series); M. Peskin and U. Schroeder, Elementary Particle Physics: Concepts and Phenomena (Texts and Monographs in Physics); O. Nachtmann, A. Lahee, and W. Wetzel, Elementary Particle Physics, Berlin Heidelberg (1990).
[2] M. Gell-Mann, private communication.
[3] P. Dirac, Nature **192** (1987) 235.
[4] For a discussion see: D. Gross, Phys. Today, **42** (1989) 12.
[5] S. Bethke, J. Phys. **G26** (2000) R27.
[6] M. Luescher, Phys. Bl. **56** (2000) 65.
[7] H. Fritzsch and Z. Xing, Prog. Part. and Nucl. Phys. **45** (2000) 1–81 [hep-ph/9912358].
[8] J. K. Webb et al., Phys. Rev. Lett. **87** (2000) 091301.
[9] X. Calmet und H. Fritzsch, CERN–TH/2001-33 [hep-ph 0112110].

Fortschr. Phys. **50** (2002) 5–7, 525–530

Quark Number Susceptibilities from Lattice QCD*

RAJIV V. GAVAI

Department of Theoretical Physics
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India
gavai@tifr.res.in

Abstract

Results from our recent investigations of quark number susceptibilities in both quenched and 2-flavour QCD are presented as a function of valence quark mass and temperature. A strong reduction ($\sim$40%) is seen in the strange quark susceptibility above T_c in both the cases. A comparison of our isospin susceptibility results with the corresponding weak coupling expansion reveals once again the non-perturbative nature of the plasma up to $3T_c$. Evidence relating the susceptibility to another non-perturbative phenomena, pionic screening lengths, is presented.

1. Introduction

Quantum chromodynamics (QCD) is the theory of interactions of quarks and gluons which are the basic constituents of strongly interacting particles such as protons, neutrons and pions. A complete lack of any experimental observation of a free quark or gluon led to the hypothesis of their permanent confinement in the observable particles. As pointed out by Satz [3] in his talk, relativistic heavy ion collision experiments at BNL, New York, and CERN, Geneva offer the possibility of counting the number of degrees of freedom of strongly interacting matter at high temperatures, thereby providing a strong argument for Heisenberg to accept the physical reality of quarks and gluons in spite of their confinement at lower temperatures. The basic idea here is to look for the production of quark-gluon plasma (QGP), predicted by QCD to exist beyond a transition temperature T_c, in the heavy ion collisions and when found, determine its thermodynamical properties, such as, e.g., its energy density. At very high temperatures, these properties can be computed theoretically and are seen to be directly proportional to the number of quarks and gluons. In order to test this idea, however, one has to face the fact that the temperatures reached in the current, and near future, experiments are likely to lie below about $5T_c \ldots 10T_c$. The current best theoretical estimates for thermodynamic observables in this temperature range are provided by numerical simulations of QCD defined on a discrete space-time lattice [4]. It seems therefore prudent to evaluate as many independent observables as possible and compare them with both experiments and approximate analytical methods. Quark number susceptibilities constitute a useful independent set of observables for testing this basic idea of counting the degrees of freedom of strongly interacting matter.

Investigations of quark number susceptibilities from first principles can have direct experimental consequences as well since quark flavours such as electric charge, strangeness or baryon number can provide diagnostic tools for the production of flavourless quark-gluon plasma in the central region of heavy ion collisions. It has been pointed out recently [5, 6] on the basis of simple models for the hadronic and QGP phases that the fluctuations of

* Based on work done [1, 2] with Sourendu Gupta and Pushan Majumdar.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

such conserved charges can be very different in these two phases and thus can act as probes of quark deconfinement. Indeed, excess strangeness production has been suggested as a signal of quark-gluon plasma almost two decades ago [7]. Lattice QCD can provide a very reliable and robust estimate for these quantities in *both* the phases since in thermal equilibrium they are related to corresponding susceptibilities by the fluctuation-dissipation theorem:

$$\langle \delta Q^2 \rangle \propto \frac{T}{V} \frac{\partial^2 \log Z}{\partial \mu_Q^2} = \chi_Q(T, \mu_Q = 0) \,. \tag{1}$$

Here μ_Q is the chemical potential for a conserved charge Q, and Z is the partition function of strongly interacting matter in volume V at temperature T. Unfortunately, the fermion determinant in QCD becomes complex for any nonzero chemical potential for most quantum numbers including those mentioned above. Consequently, Lattice QCD is unable to handle finite chemical potential satisfactorily at present, and cannot thus yield any reliable estimates of any number density. However, the susceptibility above, i.e., the first derivative of the number density at zero chemical potential, can be obtained reasonably well using conventional simulation techniques, facilitating thereby a nontrivial extension of our theoretical knowledge in the nonzero chemical potential direction.

Quark number susceptibilities also constitute an independent set of observables to probe whether quark-gluon plasma is weakly coupled in the temperature regime accessible to the current and future planned heavy ion experiments (say, $1 \leq T/T_c \leq 10$). A lot of the phenomenological analysis of the heavy ion collisions data is usually carried out assuming a weakly interacting plasma although many lattice QCD results suggest otherwise. It has been suggested [8] that resummations of the finite temperature perturbation theory may provide a bridge between phenomenology and the lattice QCD by explaining the lattice results starting from a few T_c. As we will see below, quark number susceptibilities can act as a cross-check of the various resummation schemes. Earlier work on susceptibilities [9] did not attempt to address this issue and were mostly restricted to temperatures very close to T_c. Furthermore, the quark mass was chosen there to vary with temperature linearly. We improve upon them by holding quark mass fixed in physical units ($m/T_c = \text{constant}$). We also cover a larger range of temperature up to $3T_c$ and the accepted range of strange quark mass in our simulations.

2. Formalism

After integrating the quarks out, the partition function Z for QCD at finite temperature and density is given by

$$Z = \int \mathcal{D}U \, e^{-S_g} \det M(m_u, \mu_u) \det M(m_d, \mu_d) \det M(m_s, \mu_s) \,. \tag{2}$$

Here $\{U_\nu(x)\}$, $\nu = 0 - 3$, denote the gauge variables and S_g is the gluon action, taken to be the standard [4] Wilson action in our simulations. Due to the well-known "fermion doubling problem" [4], one has to face the choice of fermion action with either exact chiral symmetry or violations of flavour symmetry. The Dirac matrices M depend on this choice. Since we employ staggered fermions, the matrices, M, are of dimensions $3N_s^3N_t$, with $N_s(N_t)$ denoting the number of lattice sites in spatial(temporal) direction. These fermions preserve some chiral symmetry at the expense of flavour violation. Although, they are strictly defined for four flavours, a prescription exists to employ them for arbitrary number of flavours which we shall use. m_f and μ_f are quark mass and chemical potential (both in lattice units) for flavour f, denoting up(u), down(d), and strange(s) in Eq. (2). The chemical

potential needs to be introduced on lattice as a function $g(\mu)$ and $g(-\mu)$ multiplying the gauge variables in the positive and negative time directions respectively, such that [10] i) $g(\mu) \cdot g(-\mu) = 1$ and ii) the correct continuum limit is ensured. While many such functions g can be constructed, $\exp(\mu)$ being a popular choice, the results for susceptibilities at $\mu = 0$ can easily be shown to be independent of the choice of g even for finite lattice spacing a. From the Z in Eq. (2), the quark number densities and the corresponding susceptibilities are defined as

$$n_f \equiv \frac{T}{V} \frac{\partial \ln Z}{\partial \mu_f}, \qquad \chi_{ff'} \equiv \frac{\partial n_f}{\partial \mu_{f'}} = \frac{T}{V} \left[\frac{1}{Z} \frac{\partial^2 Z}{\partial \mu_f \, \partial \mu_{f'}} - \frac{1}{Z} \frac{\partial Z}{\partial \mu_f} \frac{1}{Z} \frac{\partial Z}{\partial \mu_{f'}} \right], \tag{3}$$

where the volume $V = N_s^3 a^3$ and the temperature $T = (N_t a)^{-1}$. To lighten the notation, we shall put only one subscript on the diagonal parts of χ.

In order to obtain information for quark-gluon plasma in the central region, we evaluate the susceptibilities at the point $\mu_f = 0$ for all f. In this case, each n_f vanishes, a fact that we utilize as a check on our numerical evaluation. Moreover, the product of the single derivative terms in Eq. (3) vanishes, since each is proportional to a number density. We set $m_u = m_d < m_s$. Noting that staggered quarks have four flavours by default, $N_f = 4$, and defining $\mu_3 = \mu_u - \mu_d$, one finds from Eq. (3) that the isotriplet and strangeness susceptibilities are given by

$$\chi_3 = \frac{T}{2V} \, \mathcal{O}_1(m_u), \qquad \chi_s = \frac{T}{4V} \left[\mathcal{O}_1(m_s) + \frac{1}{4} \, \mathcal{O}_2(m_s) \right], \tag{4}$$

where $\mathcal{O}_1 = \langle \mathrm{Tr}\,(M'' M^{-1} - M' M^{-1} M' M^{-1}) \rangle$, $\mathcal{O}_2 = \langle (\mathrm{Tr}\, M' M^{-1})^2 \rangle$, $M' = \partial M / \partial \mu$ and $M'' = \partial^2 M / \partial \mu^2$. The angular brackets denote averaging with respect to the Z in Eq. (2). One can similarly define baryon number and charge susceptibilities. We refer the reader for more details on them to Ref. [2].

In the discussion above, quark mass appears as an argument of $\mathcal{O}_i$ and implicitly in the Boltzmann factor of Z. Let us denote it by m_{val} and m_{sea} respectively. While the two should ideally be equal, we evaluated the expressions above in steps of improving approximations (and increasing computer costs) by first setting $m_{\mathrm{sea}} = \infty$ for all flavours (quenched approximation [1]) and then simulating two light dynamical flavours, by setting $m_{\mathrm{sea}}/T_c = 0.1$ (2-flavour QCD [2]). In each case we varied m_{val} over a wide range to cover both light u, d quarks as well as the heavier strange quark. Details of our simulations as well as the technical information on how the thermal expectation values of $\mathcal{O}_i$ were evaluated are in Refs. [1, 2].

3. Results

Based on our tests [1] of volume dependence, made by varying N_s from 8 to 16, we chose $N_s = 12$, as the susceptibilities differed very little from those obtained on an $N_s = 16$ lattice. Choosing $N_t = 4$, the temperature is $T = (4a)^{-1}$, where the lattice spacing a depends on the gauge coupling β. Using the known $\beta_c(N_t')$ for $N_t' = 6$, 8, 12, where β_c is the gauge coupling at which chiral (deconfinement) transition/cross-over takes place for lattices with temporal extent N_t', we obtained results at $T/T_c = N_t'/N_t = 1.5$, 2 and 3 in both the quenched approximation and the 2-flavour QCD. From the existing estimates [11] of T_c for 2-flavour QCD, one finds that the sea quark mass in our dynamical simulations corresponds to 14−17 MeV. Figure 1 shows our results, normalized to the free field values on the same size lattice, χ_{FFT}. These can be computed by setting gauge fields on all links to unity:

Fig. 1. $\chi/\chi_{\rm FFT}$ as a function of temperature for various valence quark masses

$U_\nu(x) = 1$ for all ν and x. Note that due this choice of our normalization, the overall factor N_f for degenerate flavours cancels out, permitting us to exhibit both χ_3 and χ_s of Eq. (4) on the same scale in Fig. 1. The continuous lines in Fig. 1 were obtained from the interpolation of the data obtained in the quenched approximation (the data are not shown for the sake of clarity), while the results for two light dynamical flavours are shown by the data points. In each case the value of $m_{\rm val}/T_c$ is indicated on the left. Due to the fact [2] that the contribution of $\mathcal{O}_2$ to χ_s turns out to be negligibly small for $T > T_c$, χ_3 and χ_s appear coincident in Fig. 1. For the real world QCD, the low valence quark mass results are relevant for χ_3 and those for moderate valence quark masses are for χ_s.

Although T_c differs in the quenched and 2-flavour QCD by a factor of 1.6−1.7, the respective susceptibilities shown in Fig. 1 as a function of the dimensionless variable T/T_c change by at most 5−10% for any $m_{\rm val}/T$. Thus the effect of "unquenching", i.e., making 2 flavours of quarks (u and d) light enough to include the contribution of the corresponding quark loops, appears to be primarily a change of scale set by T_c. Since the strange quark is a lot heavier than the up and down quarks, this suggests further that including its loop contributions, i.e., including a dynamical but heavier strange quark, may not change the results in Fig. 1 significantly. For a wide range for strange quark mass of 75 to 170 MeV, the strangeness susceptibility can be read off from the shaded region. It is smaller by about 40% compared to its ideal gas value near T_c and the suppression shows a strong temperature dependence. This has implications for the phenomenology of particle abundances, where an ideal gas model without such a suppression of strangeness is employed and could therefore result in underestimates of the temperatures reached.

For the smallest $m_{\rm val}/T$ and highest temperature we studied, the ratio $\chi/\chi_{\rm FFT}$ in Fig. 1 is seen to be 0.88 (0.85) for 2-flavour (quenched) QCD, with a mild temperature variation in the large T-region. Since its variation with valence quark mass is negligibly small for small $m_{\rm val}/T$, one can assume the results for massless valence quarks to be essentially the same as those for $m_{\rm val}/T = 0.03$ in Fig. 1. In order to check whether the de-

Fig. 2. $\chi/\chi_{\rm FFT}$ as a function of α_s/π for N_f dynamical massless quarks

grees of freedom of QGP can really be counted using these susceptibilities, one needs to know whether the deviation from unity can be explained in ordinary perturbation theory or its improved/resummed versions. Usual weak coupling expansion [12] yields $\chi/\chi_{\rm FFT} = 1 - 2\,\frac{\alpha_s}{\pi}\,[1 - 4\sqrt{\frac{\alpha_s}{\pi}\,(1+\frac{N_f}{6})}]$. Figure 2 shows these predictions for various N_f along with the leading order N_f-independent prediction. Using a scale $2\pi T$ for the run-

Fig. 3. $4\chi_3/T^2$ (open symbols) and $\chi_\pi/10T^2$ (filled symbols) as a function of M_π/T at $2T_c$ (circles) and $3T_c$ (boxes)

ning coupling and $T_c/\Lambda_{\overline{MS}} = 0.49(1.15)$ for the $N_f = 2(0)$ theory [11], the values $T/T_c = 1.5$ and 3 are marked on the figure as the second (first) set. As one can read off from the Fig. 2, the ratio decreases with temperature in *both* the cases in the range up to $3T_c$ whereas our results in Fig. 1 display an increase. Furthermore, the perturbative results lie significantly above in each case, being in the range 1.027−1.08 for 2-flavour QCD and 0.986−0.994 for quenched QCD. Although the order of magnitude of the degrees of freedom can be gauged from these results and their eventual comparison with experiments, they do call for clever resummations of perturbation theory for a more convincing and precise count.

Alternatively, the deviations from free field theory could stem from non-perturbative physics. One known indicator of non-perturbative physics in the plasma phase is the screening length in the channel with quantum numbers of pion. While it exhibits chiral symmetry restoration above T_c by being degenerate with the corresponding scalar screening length, its value is known to be much smaller than the free field value unlike that for other screening lengths. Figure 3 shows our results for χ_3 and χ_π (defined as a sum of the pion correlator over the entire lattice) as a function of the inverse pionic screening length, M_π/T. It suggests the non-perturbative physics in the two cases to be closely related, if not identical.

Acknowledgments

It is a pleasure to thank my collaborators Sourendu Gupta and Pushan Majumdar. I am also thankful to the Alexander von Humboldt Foundation for organizing this wonderful symposium in its characteristic perfect style. It is a delight to acknowledge the warm hospitality of the Physics Department of the University of Bielefeld, especially that from Profs. Frithjof Karsch and Helmut Satz.

References

[1] R. V. Gavai and S. Gupta, Phys. Rev. **D64** (2001) 074506.
[2] R. V. Gavai, S. Gupta, and P. Majumdar, hep-lat/0110032, Phys. Rev. **D**, in press.
[3] H. Satz, in these proceedings.
[4] M. Creutz, *Quarks, Gluons and Lattices*, 1985, Cambridge University Press.
[5] M. Asakawa, U. W. Heinz, and B. Müller, Phys. Rev. Lett. **85** (2000) 2072.
[6] S. Jeon and V. Koch, Phys. Rev. Lett. **85** (2000) 2076.
[7] J. Rafelski and B. Müller, Phys. Rev. Lett. **46** (1982) 1066;
erratum-ibid **56** (1986) 2334.
[8] J. P. Blaizot et al., Phys. Rev. **D63** (2001) 065003;
J. O. Andersen et al., Phys. Rev. **D63** (2001) 105008;
K. Kajantie et al., Phys. Rev. Lett. **86** (2001) 10.
[9] S. Gottlieb et al., Phys. Rev. Lett. **59** (1987) 1513;
R. V. Gavai et al., Phys. Rev. **D40** (1989) 2743;
S. Gottlieb et al., Phys. Rev. **D55** (1997) 6852.
[10] R. V. Gavai, Phys. Rev. **D32** (1985) 519.
[11] A. Ali Khan et al., Phys. Rev. **D63** (2001) 034502;
F. Karsch et al., Nucl. Phys. **B605** (2001) 579;
S. Gupta, Phys. Rev. **D64** (2001) 034507.
[12] J. I. Kapusta, *Finite-temperature Field Theory*, 1989, Cambridge University Press, Cambridge, UK, pp. 132−133.

Fortschr. Phys. **50** (2002) 5–7, 531–537

Elementary Particles on a Dedicated Parallel Computer

KAZUYUKI KANAYA for the CP-PACS Collaboration[1)]

Institute of Physics, University of Tsukuba, Tsukuba, Ibaraki 305-8571, Japan
kanaya@rccp.tsukuba.ac.jp

Abstract

Numerical simulation of lattice QCD is now an indispensable tool for investigating non-perturbative properties of quarks at low energies. In this paper, I introduce a recent lattice QCD calculation of light hadron spectrum and quark masses performed on a dedicated parallel computer, CP-PACS. From a simulation of QCD with two flavors of dynamical quarks, we find $m_{ud}^{\overline{\rm MS}}(2\,{\rm GeV}) = 3.44^{+0.14}_{-0.22}$ MeV for the averaged mass of u and d quarks to reproduce the experimental masses of π and ρ mesons, and $m_s^{\overline{\rm MS}}(2{\rm GeV}) = 88^{+4}_{-6}$ MeV or 90^{+5}_{-11} MeV to reproduce the K or ϕ meson mass. These values are about 20–30% smaller than the previous estimates using the quenched approximation.

1. Introduction

Quarks are among the most fundamental building blocks of nature. Unlike the previous elementary constituents, quarks are found to possess a remarkable property, *confinement*, i.e., we cannot isolate quarks by dissociating hadrons. Historically, fundamental theories of nature have emerged through reactions of isolated elements to weak external perturbations. Quark confinement, however, has the consequence that we cannot directly measure the fundamental properties of quarks by experiment.

Fortunately, quarks are found to possess another remarkable property, *asymptotic freedom*, whereby interactions among quarks become weak at high energies. This enabled us to identify quantum chromodynamics (QCD) as the fundamental theory of quarks. QCD has been quite successful in explaining high energy processes of particles.

On the other hand, because the coupling parameter of QCD becomes large at low energies, basic properties of hadrons, such as mass spectrum, spatial sizes, decay constants, etc., are not calculable by conventional analytic methods. In other words, we cannot reconstruct hadrons from quarks based on the dynamics of QCD yet. In order to answer the fundamental question of whether QCD is correct also at low energies, we have to carry out non-perturbative calculations of hadrons directly from the first principles of QCD.

Calculation of low-energy hadron properties is also required for determination of fundamental parameters of nature. Due to quark confinement, quark masses and coupling parameters have to be inferred indirectly from a comparison of experimental results for hadron masses etc. with a theoretical calculation of them as functions of the fundamental parameters of QCD. Also for electroweak interactions, although the core part of the reaction can be reliably calculated in perturbation theory, uncertainties in QCD corrections to the quark currents propagate to the determination of the Cabibbo-Kobayashi-Maskawa parameters.

[1)] CP-PACS Collaboration: A. Ali Khan, S. Aoki, G. Boyd, R. Burkhalter, S. Ejiri, M. Fukugita, S. Hashimoto, K. Ide, K.-I. Ishikawa, N. Ishizuka, Y. Iwasaki, T. Izubuchi, K. Kanaya, T. Kaneko, Y. Kuramashi, V. Lesk, T. Manke, Y. Namekawa, K.-I. Nagai, J.-I. Noaki, M. Okawa, M. Okamoto, Y. Taniguchi, H. P. Shanahan, A. Ukawa, T. Umeda, T. Yamazaki, and T. Yoshié.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Numerical simulation of quarks based on the lattice formulation of QCD is currently the only method to calculate low-energy non-perturbative properties of hadrons. Recently, predictions with big impact are beginning to be produced through development of dedicated parallel computers [1]. In particular, a big breakthrough was achieved by the CP-PACS computer developed at the University of Tsukuba [2]. In this paper, I summarize major results from CP-PACS, focusing on the most fundamental topics of QCD, i.e., the light hadron spectrum and light quark masses.

In Section 2, I introduce the lattice formulation of QCD and explain why dedicated computers have been developed by physicists. Results for the light hadron spectrum and light quark mass are presented in Sections 3 and 4. Conclusions are given in Section 5.

2. Lattice QCD and Dedicated Parallel Computers

We formulate QCD on a four-dimensional hypercubic lattice with a finite lattice spacing a and finite lattice size Na. The real world is defined by the limit of vanishing lattice spacing $a \to 0$ keeping the lattice size Na finite or sufficiently large (the continuum limit). Before taking this limit, the theory is finite and mathematically well-defined. Therefore, we can apply various non-perturbative techniques. Here, we perform numerical simulations on finite lattices to calculate hadrons.

To accomplish this, the most time-consuming part of the calculation is the inversion of quark propagation kernels. Numerically, this is an inversion of large sparse complex matrices. For the case of Wilson-type lattice quark actions, the typical size of the matrix is $12V \times 12V$, where $12 = 3 \times 4$ is the freedom of color and spin, and $V = N^4$ is the lattice volume. Because the condition number is inversely proportional to the quark mass, the inversion is numerically more intensive when we decrease the quark mass. Even with the latest supercomputers, it is difficult to simulate the light u, d quarks directly. Therefore, in addition to the continuum extrapolation, we need to extrapolate the results at typically around the s quark mass to the physical u, d quark mass point. This procedure is called "the chiral extrapolation".

In summary, in order to extract a prediction for the real world, we have to perform continuum and chiral extrapolations. To get a precise and reliable result, it is essential to have good control of these extrapolations. This requires a large-scale systematic calculation,

Fig. 1. Development of the computer speed

and thus huge computer power. The power of vector supercomputers in 1980's was not sufficient, which motivated several groups of lattice physicists to construct parallel computers dedicated to lattice calculations [1].

The historical development of computer speed is shown in Fig. 1. The frontier of speed has been advanced by parallel computers since around 1990. Dedicated machines developed by physicists contributed much to this trend. At the University of Tsukuba, we have developed two machines, QCDPAX (1990) [3] and CP-PACS (1996) [2]. CP-PACS is a parallel computer achieving peak performance of 614.4 GFLOPS with 2048 single-processor nodes. Since the first power-on in 1996, intensive lattice QCD calculations have been made on CP-PACS. In the following, I focus on the studies of hadron spectrum and light quark masses.

3. Light Hadron Mass Spectrum

Precise calculation of the hadron mass spectrum directly from the first principles of QCD is one of the main goals of lattice QCD. Because the computer power required is enormous, we study the issue step by step as follows:

Step 1: Calculate in the quenched approximation, in which the effects of dynamical pair creation and annihilation of quarks are neglected.
Step 2: Include the dynamical u, d quarks in a degenerate approximation, while heavier quarks are treated in the quenched approximation (*two-flavor full QCD*).
Step 3: Include the dynamical s quark ($2 + 1$ *flavor full QCD*).
Steps 4, 5, ...: Introduce the u, d mass difference, dynamical c quark, etc.

3.1. Quenched studies

With the quenched approximation, we can reduce the computer time by a factor of several hundred preserving the basic properties of QCD (confinement, asymptotic freedom, and spontaneous breakdown of chiral symmetry). The effects of the approximation are expected to be about 10% in the spectrum.

The first studies in this approximation had already been made in early 80's. The issue turned out to be quite tough and computationally demanding. Actually, it took about ten years until the first systematic study, performing all the extrapolations, was attempted in 1993 [4]. From this study, the quenched light hadron spectrum was found to be consistent with experiment within the errors of about 10%. However, the quality of data was insufficient to test the accuracy of numerical extrapolations, and also the final errors were too large to resolve the quenching artifact.

The status was significantly improved by CP-PACS [5]. From an intensive computation involving about 100 times more floating-point calculations than the previous studies, it became possible to perform systematic tests on the quality of the extrapolations. Errors in the final results for the light hadron spectrum are now confidently estimated to be less than about 1% for mesons and about 3% for baryons. These errors include statistical and all systematic errors except for those from the quenched approximation itself.

The quenched light hadron spectrum is shown in Fig. 2. Experimental values of meson masses $-$ M_π, M_ρ, and either M_K or M_ϕ $-$ are used as inputs to fix the lattice spacing a (or the strong coupling constant), the average u, d quark mass m_{ud}, and the s quark mass m_s. Masses for other hadrons are among the predictions of QCD. From Fig. 2, we find that the global pattern of the spectrum is well reproduced. At the same time, we clearly see that the quenched spectrum deviates from experiment by about 10% (7 standard deviations) for

Fig. 2. Quenched light hadron spectrum [5]

the worst case: Results from the M_K-input and M_ϕ-input are discrepant by, and the hyperfine splitting between K and K^* mesons is smaller than experiment by about 10%. Decuplet baryon splittings are also small. Because all other systematic errors are well controlled, we identify the discrepancy as quenching artifact.

3.2. Two-flavor full QCD

From the quenched simulation we find that, to calculate hadronic quantities with precision better than 10%, we have to incorporate dynamical quarks. A naive extension of the quenched simulation to full QCD is difficult because several hundred times more computer time is required. A partial solution is given by improvement of the lattice theory, whereby continuum properties are realized on coarser lattices. From a preparatory study [6], we find that the combination of renormalization-group improved glue action and clover-improved

Fig. 3. Continuum extrapolation of vector meson masses M_ϕ and M_{K^*} in two-flavor full QCD ($N_f = 2$) and quenched QCD ($N_f = 0$), using M_K as input [7]. Open circles are recent results using a different improved lattice action [8].

Wilson quark action is effective in removing major lattice artifacts. This reduces the computer time by a factor of about ten.

Although more might still be hoped for, we can start the first systematic studies of two-flavor full QCD on the CP-PACS performing both chiral and continuum extrapolations [7]. To identify dynamical quark effects clearly, we carried out another quenched simulation using the same improved action.

Figure 3 shows the continuum extrapolation of vector meson masses. We find that the two quenched results (open and filled squares) lead to universal values in the continuum limit $a = 0$. They deviate from the experimental values (diamonds) as noted in the previous subsection. On the other hand, the full QCD results (circles) extrapolate to values much closer to experiment. Accordingly, we find no big discrepancies between the results from M_K and M_ϕ inputs. This means that the quenching artifacts in the spectrum are mostly removed by introducing dynamical u, d quarks.

The whole light hadron spectrum can now be approximately reproduced from QCD by adjusting just three parameters; the strong coupling constant, m_{ud} and m_s. This provides us with a strong confirmation that QCD is correct also at low energies. Remaining small deviations from experiment may be explained by the quenched approximation of the s quark. In order to confirm this, however, uncertainties from the chiral and continuum extrapolations should be reduced to the level of our quenched study. This goal is reserved for future work.

4. Light Quark Masses

Adjustment of m_{ud} and m_s in the calculation of the hadron spectrum provides us with the most direct determination of quark masses from QCD. Figure 4 shows m_s in the $\overline{\text{MS}}$ scheme at $\mu = 2$ GeV as functions of the lattice spacing a. Results from the quenched improved action, which are consistent with the quenched standard action in the continuum limit, are omitted for clarity.

On the lattice, there are several alternative definitions for the quark mass, with the difference being $O(a)$. In Fig. 4, they are denoted as VWI and AWI (vector and axial-vector

Fig. 4. Continuum extrapolation of the s quark mass m_s in the $\overline{\text{MS}}$ scheme at 2 GeV from two-flavor [7] and quenched QCD [5]. Quenched masses are shown by triangles. Filled and open symbols are for the M_K and M_ϕ inputs, respectively.

Table 1
Light quark masses in MeV in the $\overline{\text{MS}}$ scheme at 2 GeV [7]

	m_{ud}	m_s (K-input)	m_s (ϕ-input)	m_s/m_{ud}
$N_f = 0$ Stand.	4.57 ± 0.18	116 ± 3	144 ± 6	$\approx 25-31$
$N_f = 0$ Impr.	$4.36^{+0.14}_{-0.17}$	110^{+3}_{-4}	132^{+4}_{-6}	$\approx 25-30$
$N_f = 2$	$3.44^{+0.14}_{-0.22}$	88^{+4}_{-6}	90^{+5}_{-11}	26 ± 2

Ward-Takahashi identity). See [5, 7] for details. As shown in Fig. 4, different definitions lead to different values of m_q at finite lattice spacings. This has been a big source of uncertainty in the previous calculations. From Fig. 4, we see that they converge to a universal value in the continuum limit, as theoretically expected. This confirms the quality of our calculations.

The quenched value for m_s, however, differs by about 20% between M_K-input and M_ϕ-input. This is equivalent to the quenched artifact discussed in Section 3.1. We note that the quenched artifact is larger in the light quark masses than in the hadron spectrum. When we turn on the dynamical u, d quarks, the discrepancy between the inputs disappears within our errors, in accord with the fact that the full QCD reproduces the hadron spectrum better (Sec. 3.2).

The light quark masses in the continuum limit are summarized in Table 1. The errors include our estimates for systematic errors from chiral and continuum extrapolations and renormalization factors. We note that the masses from two-flavor QCD are 20−30% smaller than those from the quenched QCD. In particular, our s quark mass in $N_f = 2$ QCD is about 90 MeV, which is significantly smaller than the value ≈ 150 MeV often used in phenomenology. Our results are, however, consistent with recent estimates $m_s = 83-130$ MeV and $m_{ud} = 3.4-5.3$ MeV from QCD sum rules [9] and $m_s/m_{ud} = 24.4 \pm 1.5$ from one-loop chiral perturbation theory [10].

5. Conclusions

An intensive calculation of the light hadron spectrum in the quenched approximation has revealed discrepancies of about 10% from experiment. With two flavors of dynamical u, d quarks, the first systematic study performing both the continuum and chiral extrapolations has shown that the discrepancies are mostly removed, providing us with a strong confirmation that QCD is the correct theory of quarks at low energies as well.

Precision calculation of hadron masses enables us to determine fundamental parameters of quarks directly from the first principles of QCD. We find that the dynamical quark effect is as large as 20−30% in light quark masses. Noticeable and sizable dynamical quark effects are observed also in B meson decay constants [11], the equation of state at high temperatures [12], and the topological structure of the QCD vacuum [13]. At the same time, new types of lattice fermions have begun to be applied to study hadronic matrix elements relevant to the $\Delta I = 1/2$ rule and the CP violation parameters [14].

Because a shift in fundamental parameters can have significant implications to phenomenological studies of the standard model, it is urgent to evaluate dynamical quark effects in other hadronic quantities as well. The influence of dynamical s quark should also be studied. Studies in these directions are under way to clarify the precise structure of the standard model.

The studies presented in this report have been performed by the CP-PACS Collaboration. I thank the members of the Collaboration for discussions and comments. This work is in part supported by the Grants-in-Aid of Ministry of Education, Science and Culture (Nos. 12304011 and 13640260) and JSPS Research for Future Program.

References

[1] For recent reports, see Proc. Lattice 2000, Nucl. Phys. **B** (Proc. Suppl.) **94** (2001).
[2] Y. IWASAKI, Nucl. Phys. **B** (Proc. Suppl.) **60A** (1998) 246, For further details, see http://www.rccp.tsukuba.ac.jp.
[3] Y. IWASAKI, et al., Comp. Phys. Commun. **49** (1988) 363.
[4] F. BUTLER, et al., Nucl. Phys. **B430** (1994) 179.
[5] S. AOKI, et al. (CP-PACS Collab.), Phys. Rev. Lett. **84** (2000) 238.
[6] S. AOKI, et al. (CP-PACS Collab.), Phys. Rev. **D60** (1999) 114508.
[7] A. ALI KHAN, et al. (CP-PACS Collab.), Phys. Rev. Lett. **85** (2000) 4674; hep-lat/0105015.
[8] S. AOKI, et al. (JLQCD Collab.), hep-lat/0110179; T. KANEKO, hep-lat/0111005.
[9] R. GUPTA and K. MALTMAN, DPF2000 (2001) [hep-ph/0101132].
[10] H. LEUTWYLER, Phys. Lett. **B378** (1996) 313.
[11] A. ALI KHAN, et al. (CP-PACS Collab.), Phys. Rev. **D64** (2001) 034505; *ibid.* **D64** (2001) 054504.
[12] M. OKAMOTO, et al. (CP-PACS Collab.), Phys. Rev. **D60** (1999) 0094510;
A. ALI KHAN, et al. (CP-PACS Collab.), ibid. **D63** (2001) 034502;
ibid. **D64** (2001) 074510;
Y. NAMEKAWA, et al. (CP-PACS Collab.),
ibid. **D64** (2001) 074507.
[13] A. ALI KHAN, et al. (CP-PACS Collab.), Phys. Rev. **D64** (2001) 114501.
[14] A. ALI KHAN, et al. (CP-PACS Collab.), Phys. Rev. **D63** (2001) 114504;
ibid. **D64** (2001) 114506;
J-I. NOAKI, et al. (CP-PACS Collab.), hep-lat/0108013

Fortschr. Phys. **50** (2002) 5–7, 538–540

Why Extra Gauge Bosons Should Exist and How to Hunt Them

ARND LEIKE

Ludwig-Maximilians-Universität München, Sektion Physik
Theresienstr. 37, D-80333 München, Germany
leike@theorie.physik.uni-muenchen.de

Abstract

Werner Heisenberg's work is the foundation for many topics of present research. This is also true for the search for extra gauge bosons. The prospects of future colliders in this search are shortly mentioned.

During the last years of his life, Werner Heisenberg tried to found a unified theory of all interactions. Today, 25 years later, this unified theory is still lacking. The minimum unification group of strong and electroweak interactions is $SU(5)$ [1]. It is ruled out by several experiments. LEP has measured [2] $\sin^2 \hat{\theta}(M_Z)\,(\overline{MS}) = 0.23117(16)$, while $SU(5)$ predicts $\sin^2 \theta(M_Z) = 0.21$. Proton decay is not observed in experiment [2] leading to $\tau(p \to e^+\pi^0) > 1.6 \cdot 10^{33}$ years, while $SU(5)$ predicts $\tau(p \to e^+\pi^0) = 10^{28}$–$10^{30}$ years. The observed neutrino mixing in solar and atmospheric neutrino experiments indicates that this particle must have a non-zero mass, which is not the case in $SU(5)$.

The next extension of the unification group is $SO(10)$ [3]. It is consistent with experimental data. All unification groups larger than $SU(5)$ contain extra neutral and extra charged gauge bosons. Experimental signals of these particles would give interesting information about the underlying grand unified theory. The search for extra gauge bosons is therefore a part of the physical programme of every present and future collider.

In the vacuum of a quantum field theory, virtual particles are continuously emitted and absorbed. The energy needed to create these virtual particles can be "borrowed" from the vacuum due to the tunnel effect. However, these effects happen at length scales of the order of the de Broglie wavelengths of the involved particles. Therefore, one must resolve very short distances to be sensitive to heavy particles. This can be done by high energy experiments or by precision experiments.

The higher the energy of an particle the higher is its momentum. According to Heisenberg's uncertainty relation [4], test particles with higher momentum can resolve smaller scales. This is well known from electron microscopes.

High precision measurements can also give valuable information on heavy particles because they are sensitive to rare processes. Consider, for example, atomic parity violation. In the conservative atomic model, one assumes that only photons are exchanged between the atomic nucleus and the electrons. However, the interaction between an electron and a quark of the nucleus can also be mediated by the Z Boson. This is a rare process because the Z Boson is heavy. Photon exchange does not violate parity but Z Boson exchange does. Parity violating transitions are observed in experiments. They prove quantitatively that Z Bosons *are* exchanged in atoms. The atomic parity violation experiments are so precise that they also set limits on extra neutral gauge bosons, although these particles must be considerably heavier than the Z [2].

Extra gauge bosons are not observed. This means that they are either not existing or heavy. If extra gauge bosons are not too heavy they will show up in future experiments. I

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

will concentrate here on limits on extra gauge bosons from a future linear e^+e^- collider and compare them with possible limits from LHC. The expected experimental signal of these particles in future experiments is a small deviation of observables from the Standard Model prediction. Future experiments will either see deviations from the SM or not. In the first case it must be decided whether these deviations can be due to extra gauge bosons. If yes, the corresponding constraints on their parameters must be found. In the second case, the data can be interpreted as exclusion limits to extra gauge bosons. In both cases, it is necessary to calculate the measured observables in theories with extra gauge bosons. A recent review about these calculations for extra neutral gauge bosons can be found in [5].

Collision experiments are calculated with the help of the S-matrix. Heisenberg founded the S-matrix theory in 1943 [6]. The results of recent investigations on the search potential for extra gauge Bosons at a future linear e^+e^- collider can be found in the TESLA technical design report [7].

The discovery limits for extra gauge bosons are shown in Tables 1 and 2. The numbers in the tables also show the dependence of the limits on the expected systematic errors. Case A and B in Table 1 refer to the following assumptions on systematic errors [8]:

$$\begin{aligned}
&\text{case A: } \Delta P_{e^\pm} = 1.0\%\,, \qquad \Delta L = 0.5\%\,, \qquad \Delta^{\text{sys}}\epsilon_{\text{lepton}} = 0.5\%\,,\\
&\Delta^{\text{sys}}\epsilon_{\text{hadron}} = 0.5\%\,,\\
&\text{case B: } \Delta P_{e^\pm} = 0.5\%\,, \qquad \Delta L = 0.2\%\,, \qquad \Delta^{\text{sys}}\epsilon_{\text{lepton}} = 0.1\%\,,\\
&\Delta^{\text{sys}}\epsilon_{\text{hadron}} = 0.1\%\,.
\end{aligned} \tag{1}$$

The beam polarizations are $P_{e^+} = 0.6$ and $P_{e^-} = 0.8$ in both tables.

Table 1
Discovery limits on extra neutral gauge bosons in TeV for an e^+e^- collider. The numbers are taken from Fig. 5.2.1. of [7]. See the original reference [8] for more details.

	$\sqrt{s} = 0.5TeV$, $L_{\text{int}} = 1000fb^{-1}$ $e^+e^- \to f\bar{f}$		$\sqrt{s} = 1TeV$, $L_{\text{int}} = 1000fb^{-1}$ $e^+e^- \to f\bar{f}$		LHC $L_{\text{int}} = 100fb^{-1}$
case (syst.err.)	A	B	A	B	
Model					
χ	6.2	6.9	9.5	12.2	4.3
ψ	2.1	3.4	4.0	5.5	4.2
η	3.3	4.0	5.3	6.6	4.3
LR	7.2	9.3	13.7	15.6	4.6

Table 2
Discovery limits on extra charged gauge bosons in TeV for an e^+e^- collider. The numbers are selected from table 5.2.1. of [7]. See the original references [9] for more details. The discovery limits for LHC, $L_{\text{int}} = 100fb^{-1}$ are between 5.1 and $5.9TeV$ [10].

	$\sqrt{s} = 0.5TeV$, $e^+e^- \to \nu\bar{\nu}\gamma$		$L_{\text{int}} = 1000fb^{-1}$ $e\gamma \to \nu q + X$		$\sqrt{s} = 1TeV$, $e^+e^- \to \nu\bar{\nu}\gamma$		$L_{\text{int}} = 1000fb^{-1}$ $e\gamma \to \nu q + X$	
syst. err, %	0.1	2.0	0.5	2.0	0.1	2.0	0.5	2.0
Model								
SSM W′	4.8	1.7	4.0	2.7	5.9	2.2	5.8	4.6
LRM	1.3	0.9	0.7	0.6	1.7	1.2	1.2	1.1
KK	5.0	1.8	5.7	3.8	6.4	2.3	8.2	6.5

The numbers in Table 2 are sensitive not only to systematic errors but also to kinematical cuts, which take into account the detector acceptance. These cuts are also needed to suppress the SM background and to remove singularities arising due to soft or collinear photons. The limits from $e\gamma$ collisions assume backscattered laser photons. See references [9] for details.

As we see from the tables, hadron and lepton colliders are complementary in a search for extra gauge bosons. In the case of a positive signal, a lepton collider is especially strong in model discrimination. See references [8, 9] for a discussion of that point.

Acknowledgement

I would like to thank the Alexander von Humboldt Foundation for inviting me to this perfectly organized stimulating conference.

References

[1] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32** (1974) 438.
[2] *Review of Particle Physics*, Eur. Phys. J. **C15** (2000) 1.
[3] H. Fritzsch and P. Minkowski, Ann. Phys. (NY) **93** (1975) 193.
[4] W. Heisenberg, Z. Phys. **43** (1927) 172.
[5] A. Leike, Phys. Rep. **317** (1999) 143–250.
[6] W. Heisenberg, Z. Phys. **120** (1943) 513;
W. Heisenberg, Z. Phys. **120** (1943) 673.
[7] http://tesla.desy.de/.
[8] S. Riemann, LC-TH-2001-007.
[9] S. Godfrey, P. Kalyniak, B. Kamal, and A. Leike, Phys. Rev. **D61** (2000) 113009;
M. A. Doncheski, S. Godfrey, P. Kalyniak, B. Kamal, and A. Leike, Phys. Rev. **D63** (2001) 053005.
[10] T. Rizzo, in Proceedings of the 1996 DPF/DPB Summer Study on High-Energy Physics; hep-ph/9612440.

Fortschr. Phys. **50** (2002) 5–7, 541–551

States of Strongly Interacting Matter

HELMUT SATZ

Fakultät für Physik, Universität Bielefeld, D-33501 Bielefeld, Germany
satz@physik.uni-bielefeld.de

Abstract

I discuss the phase structure of strongly interacting matter at high temperatures and densities, as predicted by statistical QCD, and consider in particular the nature of the transition of hot hadronic matter to a plasma of deconfined quarks and gluons.

1. Prelude

To speak about quark matter in a meeting dedicated to Heisenberg is somewhat problematic. In one of his last talks, Heisenberg noted: "There exists the conjecture that the observable hadrons consist of non-observable quarks. But the word 'consist' makes sense only if it is possible to decompose a hadron into these quarks with an energy expenditure much less than the rest mass of a quark" [1]. Therefore I asked myself what arguments might have convinced Heisenberg to revise his opinion. On philosophical level, which after all played a significant role in Heisenberg's argumentation, one might remember what Lucretius pointed out more than two thousand years earlier: "So there must be an ultimate limit to bodies, beyond perception by our senses. This limit is without parts, is the smallest possible thing. It can never exist by itself, but only as primordial part of a larger body, from which no force can tear it loose" [2]. This must be one of the earliest formulations of confinement; curiously enough, it was generally ignored by all 'atomists' before the advent of QCD. Lucretius argues that the ultimate building blocks of matter cannot have an independent existence, since otherwise one could ask what they are made of.

On more physical grounds, we note that the energy density of an ideal electromagnetic plasma, consisting of electrons, positrons and photons, is given by the Stefan-Boltzmann law

$$\epsilon_{\rm QED} = \frac{\pi^2}{30}\left[2 + \frac{7}{8}\ 2 \times 2\right] T^4\,, \tag{1}$$

which counts the number of constituent species and their degrees of freedom (two spin orientations each for electrons, positrons and photons). For a hot and hence asymptotically free quark-gluon plasma, the corresponding form is

$$\epsilon_{\rm QCD} = \frac{\pi^2}{30}\left[2 \times 8 + \frac{7}{8}\ 2 \times 2 \times 3 \times 3\right] T^4\,, \tag{2}$$

which again counts the number of otherwise confined constituents and their degrees of freedom (eight colors of gluons, three colors and three flavors for quarks and antiquarks, and two spin orientations for quarks and gluons). The energy density of the hot QGP thus provides direct information on what it is made of.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

A similar argument was in fact used by Zel'dovich even before the advent of the quark model [3]. He notes that if dense stellar or pre-stellar media should not obey the equation of state of neutron matter, this might be an indication for other types of elementary particles: "It will be necessary to consider as many Fermi distributions as there are elementary particles. The problem of the number of elementary particles may be approached in this way, since if some particle is in reality not elementary, it would not give rise to a separate Fermi distribution". So the behavior of matter in the limit of high constituent density seems to be a good way to address the question of its ultimate building blocks.

2. Hadronic Matter and Beyond

Hadrons have an intrinsic size, with a radius of about 1 fm. Hence a hadron needs a volume $V_h = (4\pi/3)\, r_h^3 \simeq 4\ \mathrm{fm}^3$ to exist. This implies an upper limit n_c to the density of hadronic matter, $n_h < n_c$, with $n_c = V_h^{-1} \simeq 0.25\ \mathrm{fm}^{-3} \simeq 1.5 n_0$, where $n_0 \simeq 0.17\ \mathrm{fm}^{-3}$ denotes standard nuclear density. Fifty years ago, Pomeranchuk pointed out that this also leads to an upper limit for the temperature of hadronic matter [4]. An overall volume $V = NV_h$ causes the grand canonical partition function to diverge when $T \geq T_c \simeq 1/r_h \simeq 0.2$ GeV.

This conclusion was subsequently confirmed by more detailed dynamical accounts of hadron dynamics. Hagedorn proposed a self-similar composition pattern for hadronic resonances, the statistical bootstrap model, in which the degeneracy of a given resonant state is determined by the number of ways of partitioning it into more elementary constituents [5]. The solution of this classical partitioning problem [6] is a level density increasing exponentially with mass, $\rho(m) \sim \exp\{am\}$, which leads to a diverging partition function for an ideal resonance gas once its temperature exceeds the value $T_H = 1/a$, which turns out to be close to the pion mass. A yet more complete and detailed description of hadron dynamics, the dual resonance model, confirmed this exponential increase of the resonance level density [7, 8]. While Hagedorn had speculated that T_H might be an upper limit of the temperature of all matter, Cabbibo and Parisi pointed out that T_H could be a critical temperature signalling the onset of a new quark phase of strongly interacting matter [9]. In any case, it seems clear today that hadron thermodynamics, based on what we know about hadron dynamics, contains its own intrinsic limit [10].

On one hand, the quark infrastructure of hadrons provides a natural explanation of such a limit; on the other hand, it does so in a new way, different from all previous reductionist approaches: quarks do not have an independendent existence, and so reductionalism is at the end of the line, in just the way proposed by Lucretius.

The limit of hadron thermodynamics can be approached in two ways. One is by compressing cold nuclear matter, thus increasing the baryon density beyond values of one baryon per baryon volume. The other is by heating a meson gas to temperatures at which collisions produce further hadrons and thus increase the hadron density beyond values allowing each hadron its own volume. In either case, the medium will undergo a transition from a state in which its constituents were colorless, i.e., color-singlet bound states of colored quarks and gluons, to a state in which the constituents are colored. This end of hadronic matter is generally referred to as deconfinement.

The colored constituents of deconfined matter

- could be massive *constituent quarks*, obtained if the liberated quarks dress themselves with gluon clouds;
- or the liberated quarks could couple pairwise to form bosonic colored *diquarks*;
- or the system could consist of unbound quarks and gluons, the *quark-gluon plasma* (QGP).

One of the tasks of statistical QCD is to determine if and when these different possible states can exist.

In an idealized world, the potential binding a heavy quark-antiquark pair into a color-neutral hadron has the form of a string,

$$V(r) \sim \sigma r\,, \tag{3}$$

where σ specifies the string tension. For $r \to \infty$, $V(r)$ also diverges, indicating that a hadron cannot be dissociated into its quark constituents: quarks are confined. In a hot medium, however, thermal effects are expected to soften and eventually melt the string at some deconfinement temperature T_c. This would provide the string tension with the temperature behavior

$$\sigma(T) = \begin{cases} \sigma(0)\,[T_c - T]^a & T < T_c\,, \\ 0 & T > T_c\,, \end{cases} \tag{4}$$

with a as critical exponent for the order parameter $\sigma(T)$. For $T < T_c$, we then have a medium consisting of color-neutral hadrons, for $T > T_c$ a plasma of colored quarks and gluons. The confinement/deconfinement transition is thus the QCD version of the insulator/conductor transition in atomic matter.

In the real world, the string breaks when $V(r)$ becomes larger than the energy of two separate color singlet bound states, i.e., when the 'stretched' hadron becomes energetically more expensive than two hadrons of normal size. It is thus possible to study the behavior of Eq. (4) only in quenched QCD, without dynamical quarks and hence without the possibility of creating new $q\bar{q}$ pairs. We shall return to the case of full QCD and string breaking in Section 4.

The insulator-conductor transition in atomic matter is accompanied by a shift in the effective constituent mass: collective effects due to lattice oscillations, mean electron fields etc. give the conduction electron a mass different from the electron mass in vacuum. In QCD, a similar phenomenon is expected. At $T = 0$, the bare quarks which make up the hadrons 'dress' themselves with gluons to form constituent quarks of mass $M_q \simeq 300 - 350$ MeV. The mass of a nucleon then is basically that of three constituent quarks, that of the ρ meson twice M_q. With increasing temperature, as the medium gets hotter, the quarks tend to shed their dressing. In the idealized case of massless bare quarks, the QCD Lagrangian $\mathcal{L}_{\rm QCD}$ possesses chiral symmetry: four-spinors effectively reduce to two independent two-spinors. The dynamically created constituent quark mass at low T thus corresponds to a spontaneous breaking of this chiral symmetry, and if at some high $T = T_\chi$ the dressing and hence the constituent quark mass disappears, the chiral symmetry of $\mathcal{L}_{\rm QCD}$ is restored. Similar to the string tension behavior of Eq. (4) we thus expect

$$M_q(T) = \begin{cases} M_q(0)\,[T_\chi - T]^b & T < T_\chi\,, \\ 0 & T > T_\chi\,, \end{cases} \tag{5}$$

for the constituent quark mass: T_χ separates the low temperature phase of broken chiral symmetry and the high temperature phase in which this is restored, with b as the critical exponent for the chiral order parameter $M_q(T)$.

An obvious basic problem for statistical QCD is thus the clarification of the relation between T_c and T_χ. In atomic physics the electron mass shift occurs at the insulator-conductor transition; is that also the case in QCD?

The deconfined QGP is a color conductor; what about a color superconductor? In QED, collective effects of the medium bind electrons into Cooper pairs, overcoming the repulsive Coulomb force between like charges. These Cooper pairs, as bosons, condense at low tem-

Fig. 1. Four-phase and three-phase structure for strongly interacting matter

peratures and form a superconductor. In contrast to the collective binding effective in QED, in QCD there is already a microscopic qq-binding, coupling two color triplet quarks to an antitriplet diquark. A nucleon can thus be considered as a bound state of this diquark with the remaining third quark,

$$[3 \oplus 3 \oplus 3]_1 \sim [(3 \oplus 3 \to \overline{3}) \oplus 3]_1 \,, \tag{6}$$

leading to a color singlet state. Hence QCD provides a specific dynamical mechanism for the formation of colored diquark bosons and thus for color superconductivity. This possibility [11] has created much interest and activity over the past few years [12].

We thus have color deconfinement, chiral symmetry restoration and diquark condensation as possible transitions of strongly interacting matter for increasing temperature and/or density. This could suggest a phase diagram of the form shown on the left in Fig. 1, with four different phases. The results of finite temperature lattice QCD show that at least at vanishing baryochemical potential ($\mu_B = 0$) this is wrong, since there deconfinement and chiral symmetry restoration coincide, $T_c = T_\chi$, as the corresponding transitions in atomic physics do. In Section 4 we shall elucidate the underlying reason for this.

A second guess could thus be a three-phase diagram as shown on the right in Fig. 1, and this is in fact not in contradiction to anything so far. In passing, we should note, however, that what we have here called the diquark state is most likely more complex and may well consist of more than one phase [12].

After this conceptual introduction to the states of strongly interacting matter, we now turn to the quantitative study of QCD at finite temperature and vanishing baryochemical potential. In this case, along the $\mu_B = 0$ axis of the phase diagram 1, the computer simulation of lattice QCD has provided a solid quantitative basis.

3. Statistical QCD

The fundamental dynamics of strong interactions is defined by the QCD Lagrangian

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4}\left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^a_{bc} A^b_\mu A^c_\nu\right)^2 - \sum_f \bar{\psi}^f_\alpha (i\gamma_\mu \partial^\mu + m_f - g\gamma_\mu A^\mu)\, \psi^f_\beta \,, \tag{7}$$

in terms of the gluon vector fields A and the quark spinors ψ. The corresponding thermodynamics is obtained from the partition function

$$\mathcal{Z}(T, V) = \int \mathcal{D}A\, \mathcal{D}\psi\, \mathcal{D}\bar{\psi} \exp\left\{-S(A, \psi, \bar{\psi}; T, V)\right\} , \tag{8}$$

here defined as functional field integral, in which

$$S(A,\,\psi,\,\bar{\psi};\,T,\,V) = \int_0^{1/T} d\tau \int_V d^3x\,\mathcal{L}_{\rm QCD}(\tau = ix_0,\,\boldsymbol{x}) \tag{9}$$

specifies the QCD action. As usual, derivatives of log $\mathcal{Z}$ lead to thermodynamic observables; e.g., the temperature derivative provides the energy density, the volume derivative the pressure of the thermal system.

Since this system consists of interacting relativistic quantum fields, the evaluation of the resulting expressions is highly non-trivial. Strong interactions (no small coupling constant) and criticality (correlations of all length scales) rule out a perturbative treatment in the transition regions, which are of course of particular interest. So far, the only *ab initio* results are obtained through the lattice formulation of the theory, which leads to something like a generalized spin problem and hence can be evaluated by computer simulation. A discussion of this approach is beyond the scope of this survey; for an overview, see e.g. [13]. We shall here just summarize the main results; it is to be noted that for computational reasons, the lattice approach is so far viable only for vanishing baryochemical potential, so that all results given in this section are valid only for $\mu_B = 0$.

As reference, it is useful to recall the energy density of an ideal gas of massless pions of three charge states,

$$\epsilon_\pi(T) = \frac{\pi^2}{30}\,3T^4 \simeq T^4\,, \tag{10}$$

to be compared to that of an ideal QGP (see Eq. (2)), which for three massless quark flavors becomes

$$\epsilon_{\rm QCD}(T) \simeq 16T^4\,. \tag{11}$$

The corresponding pressures are obtained through the ideal gas form $3P(T) = \epsilon(T)$. The main point to note is that the much larger number of degrees of freedom of the QGP as compared to a pion gas leads at fixed temperatures to much higher energy densities and pressures.

The energy density and pressure have been studied in detail in finite temperature lattice QCD with two and three light dynamical quark species, as well as for the more realistic case of two light and one heavier species. The results are shown in Fig. 2, where it is seen that in all cases there is a sudden increase from a state of low to one of high values, as expected at the confinement-deconfinement transition. To confirm the connection between the transition and the increase of energy density or pressure, we make use of the order

Fig. 2. Energy density and pressure in full QCD with light dynamical quarks

parameters for deconfinement and chiral symmetry restoration; these first have to be specified somewhat more precisely than was done in the more conceptual discussion of Section 2.

In the absence of light dynamical quarks, for $m_q \to \infty$, QCD reduces to pure $SU(3)$ gauge theory; the potential between two static test quarks then has the form shown in Eq. (3) when $T < T_c$ and vanishes for $T \geq T_c$. The Polyakov loop expectation value defined by

$$\langle |L(T)| \rangle \equiv \lim_{r \to \infty} \exp\{-V(r, T)/T\} = \begin{cases} 0\,, & \text{confinement} \\ L(T) > 0\,, & \text{deconfinement} \end{cases} \tag{12}$$

thus also constitutes an order parameter for the confinement state of the medium, and it is easier to determine than the string tension $\sigma(T)$. In lattice QCD, $L(T)$ becomes very similar to the magnetization in spin systems; it essentially determines whether a global $Z_3 \in SU(3)$ symmetry of the Lagrangian is present or is spontaneously broken for a given state of the medium.

In the other extreme, for $m_q \to 0$, $\mathcal{L}_{\text{QCD}}$ has intrinsic chiral symmetry, and the chiral condensate $\langle \psi\bar{\psi} \rangle$ provides a measure of the effective mass term in $\mathcal{L}_{\text{QCD}}$. Through

$$\langle \psi\bar{\psi} \rangle = \begin{cases} K(T) > 0\,, & \text{broken chiral symmetry}\,, \\ 0\,, & \text{restored chiral symmetry}\,, \end{cases} \tag{13}$$

we can determine the temperature range in which the state of the medium shares and in which it spontaneously breaks the chiral symmetry of the Lagrangian with $m_q = 0$.

There are thus two *bona fide* phase transitions in finite temperature QCD at vanishing baryochemical potential.

For $m_q = \infty$, $L(T)$ provides a true order parameter which specifies the temperature range $0 \leq T \leq T_c$ in which the Z_3 symmetry of the Lagrangian is present, implying confinement, and the range $T > T_c$, with spontaneously broken Z_3 symmetry and hence deconfinement.

For $m_q = 0$, the chiral condensate defines a range $0 \leq T \leq T_\chi$ in which the chiral symmetry of the Lagrangian is spontaneously broken (quarks acquire an effective dynamical mass), and one for $T > T_\chi$ in which $\langle \psi\bar{\psi} \rangle\,(T) = 0$, so that the chiral symmetry is restored. Hence here $\langle \psi\bar{\psi} \rangle(T)$ is a true order parameter.

In the real world, the (light) quark mass is small but finite: $0 < m_q < \infty$. This means that the string breaks for all temperatures, even for $T = 0$, so that $L(T)$ never vanishes. On the other hand, with $m_q \neq 0$, the chiral symmetry of $\mathcal{L}_{\text{QCD}}$ is explicitly broken, so that $\langle \psi\bar{\psi} \rangle$ never vanishes. It is thus not clear if some form of critical behavior remains, and we are therefore confronted by two basic questions:

- how do $L(T)$ and $\langle \psi\bar{\psi} \rangle\,(T)$ behave for small but finite m_q? Is it still possible to identify transition points, and if so,
- what if any relation exists between T_c and T_χ?

In Fig. 3 we show the lattice results for two light quark species; it is seen that $L(T)$ as well as $\langle \psi\bar{\psi} \rangle\,(T)$ still experience very strong variations, so that clear transition temperatures can be identified through the peaks in the corresponding susceptibilities, also shown in the figure. Moreover, the two peaks occur at the same temperature; one thus finds here (and in fact for all small values of m_q) that $T_c = T_\chi$, so that the two 'quasi-critical' transitions of deconfinement and chiral symmetry restoration coincide.

Although all lattice calculations are performed for non-vanishing bare quark mass in the Lagrangian, results obtained with different m_q values can be extrapolated to the chiral limit

Fig. 3. The temperature dependence of the Polyakov loop L and the chiral condensate $\psi\bar{\psi}$, as well as of the corresponding susceptibilities

$m_q = 0$. The resulting transition temperatures are found to be $T_c(N_f = 2) \simeq 175$ MeV and $T_c(N_f = 3) \simeq 155$ MeV for two and three light quark flavors, respectively. The order of the transition is still not fully determined. For $N_f = 3$ light quark species, one obtains a first order transition. For two light flavors, a second order transition is predicted [14], but not yet unambiguously established.

4. The Nature of Deconfinement

In this last section I want to consider in some more detail two basic aspects which came up in the previous discussion of deconfinement:

- Why do deconfinement and chiral symmetry restoration coincide for all (small) values of the input quark mass?
- Is there still some form of critical behavior when $m_q \neq 0$?

Both features have recently been addressed, leading to some first and still somewhat speculative conclusions which could, however, be more firmly established by further lattice studies.

In the confined phase of pure gauge theory, we have $L(T) = 0$, the Polyakov loop as generalized spin is disordered, so that the state of the system shares the Z_3 symmetry of the Lagrangian. Deconfinement then corresponds to ordering through spontaneous breaking of this Z_3 symmetry, making $L \neq 0$. In going to full QCD, the introduction of dynamical quarks effectively brings in an external field $H(m_q)$, which in principle could order L in a temperature range where it was previously disordered.

Since $H \to 0$ for $m_q \to \infty$, H must for large quark masses be inversely proportional to m_q. On the other hand, since $L(T)$ shows a rapid variation signalling an onset of deconfinement even in the chiral limit, the relation between H and m_q must be different for $m_q \to 0$. We therefore conjecture [15, 16] that H is determined by the effective constituent quark mass M_q, setting

$$H \sim \frac{1}{m_q + c\langle\psi\bar{\psi}\rangle}\,, \tag{14}$$

since the value of M_q is determined by the amount of chiral symmetry breaking and hence by the chiral condensate. From Eq. (14) we obtain

- for $m_q \to \infty$, $H \to 0$, so that we recover the pure gauge theory limit;
- for $m_q \to 0$, we have

$$\langle \psi\bar{\psi} \rangle = \begin{cases} \text{large, } H \text{ small, } L \text{ disordered}, & \text{for} \quad T \leq T_\chi\,; \\ \text{small, } H \text{ large, } L \text{ ordered}, & \text{for} \quad T > T_\chi\,. \end{cases} \tag{15}$$

In full QCD, it is thus the onset of chiral symmetry restoration that drives the onset of deconfinement, by ordering the Polyakov loop at a temperature value below the point of spontaneous symmetry breaking [16]. Comparing the behavior of $L(T)$ in pure gauge theory to that in the chiral limit of QCD, we have in both cases a rapid variation at some temperature T_c. This variation is for $m_Q \to \infty$ due to the spontaneous breaking of the Z_3 symmetry of the Lagrangian at $T = T_c^\infty$; for $m_q \to 0$, the Lagrangian retains at low temperatures an approximate Z_3 symmetry which is explicitly broken at T_χ by an external field which becomes strong when the chiral condensate vanishes. For this reason, the peaks in the Polyakov loop and the chiral susceptibility coincide and we have $T_\chi = T_c < T_c^\infty$.

A quantitative test of this picture can be obtained from finite temperature lattice QCD. It is clear that in the chiral limit $m_q \to 0$, the chiral susceptibilities (derivatives of the chiral condensate $\langle \psi\bar{\psi} \rangle$) will diverge at $T = T_\chi$. If deconfinement is indeed driven by chiral symmetry restoration, i.e., if $L(T, m_q) = L(H(T), m_q)$ with $H(T) = H(\langle \psi\bar{\psi} \rangle\,(T))$ as given in Eq. (14), than also the Polyakov loop susceptibilities (derivatives of L) must diverge in the chiral limit. Moreover, these divergences must be governed by the critical exponents of the chiral transition.

Preliminary lattice studies support our picture [16]. In Fig. 4 we see that the peaks in the Polyakov loop susceptibilities as function of the effective temperature increases as m_q decreases, suggesting divergences in the chiral limit. Further lattice calculations for smaller m_q (which requires larger lattices) would certainly be helpful. The question of critical exponents remains so far completely open, even for the chiral condensate and its susceptibilities.

Next we want to consider the nature of the transition for $0 < m_q < \infty$. For finite quark mass neither the Polyakov loop nor the chiral condensate constitute genuine order parameters, since both are non-zero at all finite temperatures. Is there then any critical behavior? For pure $SU(3)$ gauge theory, the deconfinement transition is of first order, and the associated discontinuity in $L(T)$ at T_c cannot disappear immediately for $m_q < \infty$. Hence in a certain mass range $m_q^0 < m_q \leq \infty$, a discontinuity in $L(T)$ remains; it vanishes for m_q^0 at the

Fig. 4. The Polyakov loop susceptibilities with respect to temperature χ^L_κ (left) and to quark mass χ^L_m (right) as function of the temperature variable $\kappa = 6/g^2$ for different quark massses

Fig. 5. The phase structure of QCD (left) and of the 3-d 3-state Potts model (right)

endpoint $T_c(m_q^0)$ in the $T - m_q$ plane; see Fig. 5. For $m_q = 0$, we have the genuine chiral transition (perhaps of second order [14]) at T_χ, which, as we just saw, leads to critical behavior also for the Polyakov loop, so that here $T_c(m_q = 0) = T_\chi$ is a true critical temperature. What happens between $T_c(m_q^0)$ and $T_c(m_q = 0) = T_\chi$? The dashed line in Fig. 5 separating the hadronic phase from the quark-gluon plasma is not easy to define unambiguously: it could be obtained from the peak position of chiral and/or Polyakov loop susceptibilities [17], or from maximizing the correlation length in the medium [18]. In any case, it does not appear to be related to thermal critical behavior in a strict mathematical sense.

An interesting new approach to the behavior along this line could be provided by cluster percolation [19]. For spin systems without external field, the thermal magnetization transition can be equivalently described as a percolation transition of suitably defined clusters [20, 21]. We recall that a system is said to percolate once the size of clusters reach the size of the system (in the infinite volume limit). One can thus characterize the Curie point of a spin system either as the point where with decreasing temperature spontaneous symmetry breaking sets in, or as the point where the size of suitably bonded like-spin clusters diverges: the critical indices of the percolation transition are identical to those of the magnetization transition.

For non-vanishing external field H, there is no more thermal critical behavior; for the 2d Ising model, as illustration, the partition function now is analytic. In a purely geometric description, however, the percolation transition persists for all H, but the critical indices now are those of random percolation and hence differ from the thermal (magnetization) indices. For the 3d three state Potts' model (which also has a first order magnetization transition), the resulting phase diagram is shown on the right of Fig. 5; here the dashed line, the so-called Kertész line [22], is defined as the line of the geometric critical behavior obtained from cluster percolation. The phase on the low temperature side of the Kertész line contains percolating clusters, the high temperature phase does not [23]. Comparing this result to the $T - m_q$ diagram of QCD, one is tempted to speculate that deconfinement for $0 < m_q < \infty$ corresponds to the Kertész line of QCD [24]. First studies have shown that in pure gauge theory, one can in fact describe deconfinement through Polyakov loop percolation [25, 26]. It will indeed be interesting to see if this can be extended to full QCD.

5. Summary

We have seen that at high temperatures and vanishing baryon density, hadronic matter becomes a plasma of deconfined colored quarks and gluons. In contrast, at high baryon densities and low temperatures, one expects a condensate of colored diquarks. The quark-gluon plasma constitutes the conducting, the diquark condensate the superconducting phase of QCD.

For vanishing baryon density, the deconfinement transition has been studied extensively in finite temperature lattice QCD. In pure $SU(N)$ gauge theory (QCD for $m_q \to \infty$), deconfinement is due to the spontaneous breaking of a globel Z_N symmetry of the Lagrangian and structurally of the same nature as the magnetization transition in Z_N spin systems. In full QCD, deconfinement is triggered by a strong explicit breaking of the Z_N symmetry through an external field induced by the chiral condensate $\langle \psi \bar{\psi} \rangle$. Hence for $m_q = 0$ deconfinement coincides with chiral symmetry restoration.

For finite quark mass, $0, <, m_q, <, \infty$, it does not seem possible to define thermal critical behavior in QCD. On the other hand, spin systems under similar conditions retain geometric cluster percolation as a form of critical behavior even when there is no more thermal criticality. It is thus tempting to speculate that cluster percolation will allow a definition of color deconfinement in full QCD as genuine but geometric critical behavior.

Acknowledgements

It is a pleasure to thank S. Digal, J. Engels, S. Fortunato, F. Karsch and E. Laermann for many helpful discussions. The financial support of the German Ministry of Science (contract 06BI902) and of the GSI Darmstadt (contract BI-SAT) is gratefully acknowledged.

References

[1] W. Heisenberg, *Was ist ein Elementarteilchen*, Talk at the Meeting of the German Physical Society (DPG), March 1975.
[2] Titus Lucretius Carus, *De rerum natura*, Liber primus, Lines 599–608, ∼ 55 B. C.
[3] Ya. B. Zel'dovich, *The Number of Elementary Baryons and the Universal Baryon Repulsion Hypothesis*, JETP 37 (1959) 569.
[4] I. Ya. Pomeranchuk, Dokl. Akad. Nauk SSSR **78** (1951) 889.
[5] R. Hagedorn, Nuovo Cim. Suppl. **3** (1965) 147.
[6] L. Euler, Novi Commentarii Academiae Scientiarum Petropolitanae **3** (1753) 125;
E. Schröder, Z. Math. Phys. **15** (1870) 361;
G. H. Hardy and S. Ramanujan, Proc. London Math. Soc. **17** (1918) 75.
[7] S. Fubini and G. Veneziano, Nuovo Cim. **64** A (1969) 811.
[8] K. Bardakçi and S. Mandelstam, Phys. Rev. **184** (1969) 1640.
[9] N. Cabibbo and G. Parisi, Phys. Lett. **59** B (1975) 67.
[10] H. Satz, Fortschr. Phys. **33** (1985) 259.
[11] D. Bailin and A. Love, Phys. Rep. **107** (1984) 325.
[12] K. Rajagopal, Nucl. Phys. A **642** (1998) 26c and Nucl. Phys. A **661** (1999) 150c;
E. Shuryak, Nucl. Phys. A **642** (1998) 14c;
Th. Schäfer, Nucl. Phys. A **642** (1998) 45c.
[13] F. Karsch, *Lattice QCD at High Temperature and Density*, Schladming Lectures 2001, hep-lat/0106019.
[14] R. D. Pisarski and F. Wilczek, Phys. Rev. D **29** (1984) 338.
[15] R. V. Gavai, A. Gocksch, and M. Ogilvie, Phys. Rev. Lett. **56** (1986) 815.
[16] S. Digal, E. Laermann, and H. Satz, Eur. Phys. J. C **18** (2001) 583.
[17] S. Fortunato and H. Satz, Phys. Lett. B **509** (2001) 189.

[18] M. CASELLE et al., *Bound States in the 3d Ising Model and Implications for QCD at Finite Temperature and Density*, hep-lat/0110160, to appear in the Proceedings of *Lattice 2001*, Oct. 2001, Berlin, Germany.
[19] For a recent survey, see D. STAUFFER and A. AHARONY, *Introduction to Percolation Theory*, Taylor & Francis, London 1994.
[20] C. M. FORTUIN and P. W. KASTELEYN, J. Phys. Soc. Japan **26** (Suppl.), 11 (1969); Physica **57** (1972) 536.
[21] A. CONIGLIO and W. KLEIN, J. Phys. A **13** (1980) 2775.
[22] J. KERTÉSZ, Physica A **161**, 58 (1989).
[23] S. FORTUNATO and H. SATZ, *Cluster Percolation and First Order Phase Transitions in the Potts Model*, hep-ph/0108058; Nucl. Phys. B, in press.
[24] H. SATZ, Nucl. Phys. A **642** (1998) 130c.
[25] S. FORTUNATO and H. SATZ, Phys. Lett. B **475**, (2000) 311; Nucl. Phys. A **681**, (2001) 466c.
[26] H. SATZ, *Cluster Percolation and Thermal Critical Behavior*, hep-lat/0110013, to appear in the Proceedings of the Europhysics Conference on Computational Physics, Aachen, Germany, September 2001.

Fortschr. Phys. **50** (2002) 5–7, 552–561

From Heisenberg to Supersymmetry

MIKHAIL A. SHIFMAN

Theoretical Physics Institute, University of Minnesota
116 Church Str. S.E., Minneapolis, MN 55455, USA
shifman@physics.spa.umn.edu

Abstract

There is a little-known connection between Heisenberg's research in the 1960's and supersymmetry. First, I try to explain, in laymen terms, what supersymmetry is and why it gives rise to such high expectations in high-energy physics. After a brief excursion into origins of supersymmetry I present evidence that a section in an essentially forgotten Heisenberg's book, which contained unorthodox ideas, was inspirational for Dmitry Volkov, one of the co-discoverers of supersymmetry.

1. Introduction

I am honored to give this talk at the Heisenberg Centennial Symposium. The theme of this Symposium is the impact of Heisenberg's work on modern theoretical physics. In the previous talks we have heard how deeply various Heisenberg's ideas are intertwined in the fabric of today's theory. I will focus on supersymmetry. This might seem surprising. You might say that supersymmetry has nothing to do with Heisenberg's work, it does not belong here.

In fact, there is a little-known connection between one of Heisenberg's late physics projects and supersymmetry, a direction of research which gained dominant positions in theoretical high-energy physics in the 1990's, twenty years after Heisenberg's death. I will try to tell you this curious story.

The outline of my talk is as follows. First, I will try to explain, in laymen terms, what supersymmetry is. My explanations will be deliberately brief, primitive and, hopefully, graphic. I will keep in mind that non-experts comprise a significant part of this audience. Second, I will outline why supersymmetry is so cherished and why it gives rise to such high expectations in the community. Third, I will have to say a few words about the triple discovery of supersymmetry. This will bring me to my main story. In 1966 Heisenberg published a book "Introduction to the Unified Field Theory of Elementary Particles." A section in this book gave inspiration to Dmitry Volkov, one of the co-discoverers of supersymmetry. Heisenberg's remark served as an initial impetus for Volkov's work which eventually led to a breakthrough discovery. Thus, this remark turned out to be visionary.

I will conclude my talk by recommending a few publications which might be useful to those who are interested in the issue, both physicists and historians of science.

2. Supersymmetry in Field Theory

To visualize conventional (non-supersymmetric) field theory one usually thinks of a space filled with a large number of coupled anharmonic oscillators. For instance, in the case of $1+1$ dimensional field theory, with a single spatial dimension, one can imagine an infinite

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Fig. 1. A mechanical analogy for the scalar field theory

chain of penduli connected by springs (Fig. 1). Each pendulum represents an anharmonic oscillator. One can think of it as of a massive ball in a gravitational field. Each spring works in the harmonic regime, i.e. the corresponding force grows linearly with the displacement between the penduli. Letting the density of penduli per unit length tend to infinity, we return to field theory.

If a pendulum is pushed aside, it starts oscillating and initiates a wave which propagates along the chain. After quantization one interprets this wave as a scalar particle.

Can one present a fermion in this picture? The answer is yes. Imagine that each pendulum acquires a spin degree of freedom (i.e. each ball can rotate, see Fig. 2). Spins are coupled to their neighbors. Now, in addition to the wave that propagates in Fig. 2, one can imagine a spin wave propagating in Fig. 3. If one perturbs a single spin, this perturbation will propagate along the chain.

Our world is $1+3$ dimensional, one time and three space coordinates. In this world bosons manifest themselves as particles with integer spins. For instance, scalar (spin-0) particle from which we started is a boson. Photon (spin-1 particle) is a boson too. At the same time particles with semi-integer spins – electrons, protons, etc. – are fermions.

Conventional symmetries, such as, say, isotopic invariance, do not mix bosons with fermions. Isosymmetry tells us that the proton and neutron masses are the same. It also tells us that the masses of π^0 and π^+ are the same. However, no prediction for the ratio of the pion to proton masses emerges.

Supersymmetry is a very unusual symmetry. It connects masses and other properties of bosons, on the one hand, and fermions on the other. Thus, each known particle acquires a superpartner: photon's superpartner is *photino* (spin1/2), electron's superpartner is *selectron* (spin 0). Since spin is involved, and spin is related to geometry of space-time, it is clear that supersymmetry has a deep geometric nature. Unfortunately, I have no time to dwell on further explanations. Instead, I'd like to present here a quotation from Witten which nicely summarizes the importance of this concept for modern physics. Witten writes [1]:

> "... One of the biggest adventures of all is the search for supersymmetry. Supersymmetry is the framework in which theoretical physicists have sought to answer some of the questions left open by the Standard Model of particle physics.

Fig. 2. A mechanical analogy for the spinor field theory

Fig. 3. Superspace

> Supersymmetry, if it holds in nature, is part of the quantum structure of space and time. In everyday life, we measure space and time by numbers, "It is now three o'clock, the elevation is two hundred meters above see level," and so on. Numbers are classical concepts, known to humans since long before Quantum Mechanics was developed in the early twentieth century. The discovery of Quantum Mechanics changed our understanding of almost everything in physics, but our basic way of thinking about space and time has not yet been affected.
> Showing that nature is supersymmetric would change that, by revealing a quantum dimension of space and time, not measurable by ordinary numbers. ... Discovery of supersymmetry would be one of the real milestones in physics."

I have tried to depict "a quantum dimension of space and time" in Fig. 3. Two coordinates, x and y represent the conventional space-time. I should have drawn four coordinates, x, y, z and t, but this is impossible – we should try to imagine them.

The axis depicted by a dashed line (going in the perpendicular direction) is labeled by θ (again, one should try to imagine four distinct θ's rather than one). The dimensions along these directions cannot be measured in meters, the coordinates along these directions are very unusual, they anticommute,

$$\theta_1\theta_2 = -\theta_2\theta_1 \,, \tag{1}$$

and, as a result, $\theta^2 = 0$. This is in sharp contrast with ordinary coordinates for which 5 meters $\times$ 3 meters is, certainly, the same as 3 meters $\times$ 5 meters. In mathematics θ's are known as Grassmann numbers, the square of every given Grassmann number vanishes. These extra θ directions are purely quantum structures. In our world they manifest themselves through the fact that every integer spin particle has a half-integer spin superpartner.

A necessary condition for any theory to be supersymmetric is the balance between the number of the bosonic and fermionic degrees of freedom, having the same mass and the same "external" quantum numbers, e.g. electric charge. To give you an idea of supersymmetric field theories let us turn to the most familiar and simplest gauge theory, quantum electrodynamics (QED). This theory describes electrons and positrons (one Dirac spinor, four degrees of freedom) interacting with photons (an Abelian gauge field, with two physical degrees of freedom). Correspondingly, in its supersymmetric version, SQED, one has to add one massless Majorana spinor, photino (two degrees of freedom), and two complex scalar fields, selectrons (four degrees of freedom).

Balancing the number of degrees of freedom is the necessary but not sufficient condition for supersymmetry in dynamically nontrivial theories, of course. All interaction vertices must be supersymmetric too. This means that each line in every vertex can be replaced by that of a superpartner. Say, we start from the electron-electron-photon coupling (Fig. 4a).

Fig. 4. Interaction vertices in QED and its supergeneralization, SQED. (a) $\bar{e}e\gamma$ vertex; (b) selectron coupling to photon; (c) electron-selectron-photino vertex. All vertices have the same coupling constant. The quartic self-interaction of selectrons is also present, but not shown.

Now, as we already know, in SQED the electron is accompanied by two selectrons. Thus, supersymmetry requires the selectron-selectron-photon vertices, (Fig. 4*b*), with the same coupling constant. Moreover, the photon can be replaced by its superpartner, photino, which generates the electron-selectron-photino vertex (Fig. 4*c*), with the same coupling.

With the above set of vertices one can show that the theory is supersymmetric at the level of trilinear interactions, provided that the electrons and selectrons are degenerate in mass, while the photon and photino fields are both massless. To make it fully supersymmetric one should also add some quartic terms, describing self-interactions of the selectron fields. Historically, SQED was the first supersymmetric theory in four dimensions to be discovered [2].

3. Promises of Supersymmetry

Supersymmetry has yet to be discovered experimentally. In spite of the absence of the direct experimental evidence, immense theoretical effort was invested in this subject in the last thirty years, over 30,000 papers published. The so-called Minimal Supersymmetric Standard Model became a generally accepted paradigm in high-energy physics. In this respect the phenomenon is rather unprecedented in the history of physics. Einstein's general relativity, the closest possible analogy one can give, was experimentally confirmed within several years after its creation. Only in one or two occasions, have theoretical predictions of a comparable magnitude had to wait for experimental confirmation that long. For example, the neutrino had a time lag of 27 years. A natural question arises: why we believe this concept to be so fundamental?

Supersymmetry may help us solve the deepest mysteries of nature – the cosmological term problem and the hierarchy problem.

3.1. Cosmological term

An additional term in the Einstein action of the form

$$\Delta S = \int d^4x \sqrt{g}\, \Lambda \tag{2}$$

goes under the name of the cosmological term. It is compatible with the general covariance and, therefore, can be added freely; this fact was known to Einstein. Empirically Λ is very small, see below. In classical theory there is no problem with fine-tuning Λ to any value.

The problem arises at the quantum level. In conventional quantum field theory it is practically inevitable that

$$\Lambda \sim M_{\rm Pl}^4\,, \tag{3}$$

where M_{Pl} is the Planck scale, $M_{\mathrm{Pl}} \sim 10^{19}$ GeV. This is to be confronted with the experimental value of the cosmological term,

$$\Lambda_{\mathrm{exp}} \sim (10^{-12}\ \mathrm{GeV})^4 . \tag{4}$$

The divergence between theoretical expectations and experiment is 124 orders of magnitude! This is probably the largest discrepancy in the history of physics.

Why supersymmetry may help? In supersymmetric theories Λ is strictly forbidden by supersymmetry, $\Lambda \equiv 0$. Of course, supersymmetry, even if it is there, must be broken in

Fig. 5. SUSY time arrow

nature. People hope that the breaking might occur in a way ensuring splittings between the superpartners' masses in the ball-park of 100 GeV, with the cosmological term in the ball-park (4).

3.2. Hierarchy problem

Masses of spinor particles (electrons, quarks) are protected against large quantum corrections by chirality ("handedness"). For scalar particles the only natural mass scale is $M_{\rm Pl}$. Even if originally you choose this mass in the "human" range of, say, 100 GeV, quantum loops will inevitably drag it to $M_{\rm Pl}$. A crucial element of the Standard Model of electroweak interactions is the Higgs boson (it has not been yet discovered). Its mass has to be in the ball-park of 100 GeV. If you let its mass to be $\sim M_{\rm Pl}$, this will drag, in turn, the masses of the W bosons. Thus, you would expect $(M_W)_{\rm theor} \sim 10^{19}$ GeV while $(M_W)_{\rm exp} \sim 10^2$ GeV. The discrepancy is 17 orders of magnitude.

Again, supersymmetry comes to the rescue. In supersymmetry the notion of chirality extends to bosons, through their fermion superpartners. There are no quadratic divergences in the boson masses, at most they are logarithmic, just like in the fermion case. Thus, the Higgs boson mass gets protected against large quantum corrections.

Having explained that supersymmetry may help solve two most challenging problems in high-energy physics, I hasten to add that it does a lot of other good things, right now. It proved to be a remarkable tool in dealing with previously "uncrackable" issues in gauge theories at strong coupling. Let me give a brief list of achievements: (i) first finite four-dimensional field theories; (ii) first exact results in four-dimensional gauge theories [3]; (iii) first fully dynamical (toy) theory of confinement [4]; (iv) dualities in gauge theories [5]. The latter finding was almost immediately generalized to strings which gave rise to a breakthrough discovery of *string dualities*. This topic is beyond the scope of my talk, however.

To conclude this section, I present an *arrow of time in supersymmetry* (Fig. 5). It will also serve as a natural bridge to the main point of my today's talk – and may, I please, remind you that this main point is the impact of a virtually forgotten Heisenberg's book on the birth of supersymmetry.

4. Heisenberg and Supersymmetry

In the late 1950's Heisenberg found himself in an almost complete scientific isolation. Initially it was probably caused by alienation due to Heisenberg's role in the German atomic project during WWII. I think that later it was Heisenberg's deliberate choice to distance himself from "popular" topics on which the major efforts of the theoretical community were focussed. Naturally, such topics were also points of attraction for young and ambitious researches.

Instead, Heisenberg embarked on the investigation of a "unified field theory of elementary particles". In 1966 he published a book [6] where he summarized his work on this subject. I do not think that many people in the world studied this book carefully. In fact, it was to a large extent obsolete by the time of its publication.

Heisenberg's idea was to use a fundamental Dirac field, endowed with a four-fermion interaction of a special type, to write and nonperturbatively solve the emerging nonlinear field equations. He hoped to get in this way a complete set of "elementary" particles known at that time (both, hadrons and leptons) and dynamically describe their properties in terms of one or two input constants. From today's perspective it is absolutely obvious that this direction was a dead end. Not only it could not be made viable theoretically, it contradicted experimental data which started appearing in the 1960's. (For instance, Heisenberg did not

believe that Gell-Mann's $SU(3)$ could be a fundamental symmetry and insisted that $SU(3)$-based relations were a numerical coincidence.) Moreover, if you open the book, you will see that even in the framework of the postulates accepted by Heisenberg, his derivations were sometimes inconsistent, sometimes not compelling, with wide gaps and lots of handwaving. The man who wrote this book does not seem to be the Heisenberg of the pre-WWII time.

One might say that the whole program was a waisted effort. Let us not make hasty conclusions, however. As it happens with great minds, Heisenberg's intuition (or a direct line to God, if you wish) worked by itself, with no factual basis in the construction he was trying to develop. In this book, Heisenberg put forward a deep idea that got a life of its own, in spite of the fact that it was absolutely unsubstantiated in the framework of Heisenberg's own postulates. The idea was that massless particles must emerge from nonlinear field equations as a result of a spontaneous breaking of a continuous global symmetry. He conjectured, in particular, that *neutrinos* could be Goldstone particles. The continuous symmetry considered by Heisenberg was isosymmetry which he himself had introduced in the early 1930's. Although Heisenberg certainly new the Goldstone theorem and discussed it more than once in his book, he somehow got himself confused as to the quantum numbers of the corresponding Goldstone bosons. Sometimes he speaks of pions in this context, and in Chapter 8 he suggests that photons might be Goldstones emerging from the spontaneous breaking of isosymmetry. Of course, we now know that the isotopic invariance is a vectorial symmetry which cannot be spontaneously broken, and even if it were, this breaking could produce only a scalar Goldstone, definitely not pseudoscalar and not vector.

Heisenberg's approach was wrong – this is beyond any doubt. What is important for us now, is that on the other side of the Iron Curtain it ignited the imagination of Dmitry Volkov, a theoretical physicist from Kharkov (Ukraine) who was also very isolated, but for totally different reasons.

The Kharkov theoretical physics community was small. It was far from the main scientific centers, which automatically meant, under perverted Soviet conditions, that outside scientific contacts were cut to a minimum. Not only travel abroad was virtually beyond reach, even participation in conferences and seminars in Moscow was quite a rare occasion. Journals and preprints from the West arrived sporadically, from time to time, and the collection was far from being complete. Under these circumstances, physics books published in Russian became one of the most important sources of information.

Heisenberg's book [6] was translated in Russian and published by *Mir Publishers* in 1968. It was widely available. Its price was only 1 ruble and 1 kopek. I do not know the origin of this weird number, but I well remember that I bought this book at the time of publication in a kiosk (I was the third year student at that time), and then made several attempts to go through it with a pencil in my hand. After a few weeks I gave up, having decided that it was of no concern to me – it gave no appropriate food for my imagination. Apparently, this was a wrong conclusion (which is certainly explained by the lack of experience on my side).

Dmitry Volkov, who read this book carefully at the very same time, came to a different understanding. He posed himself a question: "OK, Heisenberg's idea of neutrino as a Goldstone particle does not follow from his construction. Is there a way to make it right?" In search of the answer, Volkov, together with his student Vladimir Akulov, invented a nonlinear realization of supersymmetry.

A few comments on the birth of supersymmetry are in order here, but before making them, I think it is time to listen to Volkov himself. This will give us a much better idea of his thinking process. In his last interview Prof. Volkov said [7]:

> "... a couple of words about how I came to the discovery of supersymmetry and supergravity. The starting point was the idea of W. Heisenberg. Let me tell you what this idea was. At the end of the 1960's, Heisenberg proposed the idea that all elementary particles

can be described by a uniform theory on the basis of the nonlinear equation formulated by him [6]. He started this work together with W. Pauli, but sometime later Pauli dissociated himself from this theory and even sharply criticized Heisenberg for his assumptions. But, nevertheless what did Heisenberg contribute to the discovery of supersymmetry?

Heisenberg tried to explain the spectrum of all elementary particles on the basis of his theory. And thanks to his great intuition, he guessed the places of the particles and predicted how the photon should appear in this theory. Moreover, he found a place for the neutrino within his theory and assumed that the neutrino emerges as a result of spontaneous symmetry breaking. This was a very unusual assumption, because all known Goldstone particles emerging as a result of spontaneous symmetry breaking had spin zero in all theories known at that time. This idea of Heisenberg was revolutionary, because he was the first to formulate the thought that there might exist Goldstone particles in nature with spin one-half. To tell the truth, he found this particle by an incorrect method, but nonetheless it was an idea that had a strong impact on me and from that time I often, although not constantly, would think about whether this idea could be realized.

And when I started to consider how such particle might appear in the theory, I understood that this requires an extension of the usual physical groups, which is the basis for all relativistic processes. That means an extension of the Lorenz group and an extension of the Poincaré group so that new operators would be present, which would correspond to a quantum number of the neutrino. Thus, the main result that we obtained was that we managed to create such an extension of the Poincaré group, which is now called the Poincaré supergroup.

Later on, we encountered certain difficulties and, if one speaks about the direct application to the neutrino, this idea nonetheless did not work. Why? Because the group that we were considering contained the Poincaré group and it is known that the Poincaré group leads eventually to the general theory of relativity, if one considers the local transformations of the group. That means that the same local transformations of the supergroup would lead to the emergence of certain superpartners of the ordinary gravity field, and these superpartners would totally absorb the Goldstone particles.

And thus, when I understood such ideas, I, together with my collaborators, proposed an extension even of Einstein's general theory of relativity, which would include just what we call supersymmetry. So, in fact, certain superpartners would emerge. Moreover, in the extension of this theory arose not only superpartners – particles with spin 3/2 – but also particles with spin 1, with spin 1/2 and with spin zero. That is, there arose the idea that all elementary particles might be included into the system of supergravity. [. . .]

Heisenberg's idea was, if one can say so, a physical idea, but in order to give this idea shape in a new precise mathematical theory, a certain mathematical formalism was required. And here I am very thankful to J. Schwinger, whom I personally knew, and who in a number of cases could have discovered supersymmetry. . . ."

This is a very instructive testimony which teaches us a few things: that the development of thought is very nonlinear; that conjectures of deep minds may have more truth in them than it might seem at first sight, that it is of paramount importance to maintain a versatile scientific community working on a variety of topics, not necessarily those constituting the fashion of the day; that personal and other scientific contacts are indispensable, and so on.

To put Volkov's work in a proper perspective, I have to say a few words about the earliest stages of supersymmetry. I will not dwell on details of this complex issue – when, where, how and why supersymmetry was discovered. An extensive account of this story can be found in the book [8] which has just been published. There you will find both, testimonies of the discoverers and pioneers, and historical essays.

The history of supersymmetry is unique because it was discovered practically *simultaneously and independently* on the both sides of the Iron Curtain *three times* in the early

1970's (see Fig. 5). As so often when exploring new ground, much of the early work on supersymmetry was hit and miss. Golfand and Likhtman of the Lebedev Institute in Moscow were the first – they reported a construction of the super-Poincaré algebra and a version of massive super-QED in 1971. The formalism contained a massive photon and photino, a charged Dirac spinor and two charged scalars (spin-0 particles). Likhtman found algebraic representations that could be viewed as chiral and vector supermultiplets, and he observed the vanishing of the vacuum energy in supersymmetric theories.

Subsequent to the work of Golfand and Likhtman, contributions from the East were made by Volkov and Akulov (Kharkov Institute for Physics and Technology) who in 1972 found a massless fermion, appearing due to the spontaneous supersymmetry breaking, and tried to associate it with the neutrino [9]. Within a year, Volkov and Soroka gauged the super-Poincar'e group which led to elements of supergravity [10]. They suggested that a spin 3/2 graviton's superpartner becomes massive upon "eating up" the Goldstino that Akulov and Volkov had discussed earlier: the existence of this "super-Higgs mechanism" in full-blown supergravity was later established in the West.

In the West, a completely different approach was taken. As is well-known, a breakthrough into the superworld was made by Wess and Zumino in 1973. This work was done independently because western researchers knew little if anything about the work done in the Soviet Union. The pre-history on which Wess and Zumino based their inspiration, has common roots with string theory – another pillar of modern theory – which in those days was referred to as the "dual model".

Subsequent progress followed rapidly. In a series of papers, Wess and Zumino, Iliopoulos and Zumino, and Ferrara, Iliopoulos, and Zumino described miraculous cancelations concerning the renormalizability of supersymmetric theories. The superspace/superfield formalism was worked out by Salam and Strathdee. Key non-renormalization theorems were proven by West and others. Non-Abelian gauge theories were supersymmetrized, simultaneously, by Ferrara and Zumino, and Salam and Strathdee in 1974. Mechanisms for the spontaneous breaking of supersymmetry through F-and D-terms were found by O'Raifeartaigh, and Fayet and Iliopoulos. Foundations of what is now known as the minimal supersymmetric standard model were laid by Fayet. In parallel, local supersymmetry was being developed too. This process culminated in 1976 with the publication of two papers (Ferrara, Freedman, and van Nieuwenhuizen, and Deser and Zumino). These authors assembled together various "super-elements" which were in circulation at that time, completing the elegant construction of modern supergravity.

5. What to Read?

I conclude my talk with a few recommendations on the literature. People often ask: is there anything to read about supersymmetry for the general audience? I should say that this theoretical development is rather fresh and the literature for non-experts is not abundant. I may refer you to Gordy Kane's book quoted in Ref. [1]. As for history of supersymmetry, several relevant publications appeared recently [8, 11, 12]. They present an extensive and, to my mind, very instructive source of information for curious physicists and historians of science. The articles collected there contain exhaustive references covering all stages of the development of supersymmetry.

Acknowledgments

I am deeply grateful to Alexander von Humboldt-Stiftung for organizing this marvelous and stimulating meeting and for inviting me to give this talk. Financial support from Alexander von Humboldt-Stiftung is gratefully acknowledged. This work was also supported in part by DOE grant DE-FG02-94ER408.

References

[1] E. WITTEN, in G. KANE, *Supersymmetry: Unveiling the Ultimate Laws of Nature* (Perseus Books, 2000).

[2] YU. A. GOLFAND and E. P. LIKHTMAN, JETP Lett. **13**, (1971) 323 [Reprinted in *Supersymmetry*, Ed. S. Ferrara, (North-Holland/World Scientific, Amsterdam – Singapore, 1987), Vol. 1, page 7].

[3] V. A. NOVIKOV, M. A. SHIFMAN, A. I. VAINSHTEIN, and V. I. ZAKHAROV, Nucl. Phys. B **229** (1983) 381; Phys. Lett. B **166** (1986) 329.

[4] N. SEIBERG and E. WITTEN, Nucl. Phys. B **426** (1994) 19, (E) B **430** (1994) 485; [hep-th/9407087]. Nucl. Phys. B **431** (1994) 484 [hep-th/9408099].

[5] N. SEIBERG, Proceedings 4th Int. Symposium on Particles, Strings, and Cosmology (PASCOS 94), Syracuse, New York, May 1994, Ed. K. C. WALI (World Scientific, Singapore, 1995), p. 183 [hep-th/9408013]; Int. J. Mod. Phys. A **12** (1997) 5171 [hep-th/9506077].

[6] W. HEISENBERG, *Introduction to the Unified Field Theory of Elementary Particles*, (Interscience Publishes, London, 1966).

[7] D. VOLKOV, in *Supersymmetry and Quantum Field Theory*, Eds. J. Wess and V. Akulov (Springer, Berlin, 1998), p. 1, see also Ref. [8].

[8] G. KANE and M. SHIFMAN (eds.), *The Supersymmetric World: The Beginnings of the Theory* (World Scientific, Singapore, 2000).

[9] D. V. VOLKOV and V. P. AKULOV, Phys. Lett. **B46** (1973) 109.

[10] D. V. VOLKOV and V. SOROKA, JETP Lett. **18** (1973) 312; Theor. Math. Phys. **20** (1974) 829.

[11] *The Many Faces of the Superworld*, Yuri Golfand Memorial Volume, Ed. M. Shifman (World Scientific, Singapore, 2000).

[12] *SUSY 30*, Proc. Int. Symp. Celebrating 30 years of Supersymmetry, Minneapolis, MN, USA, 13–17 October 2000, Eds. K. Olive, S. Rudaz and M. Shifman (North-Holland, Amsterdam, 2001).

Fortschr. Phys. **50** (2002) 5–7, 562–568

The Baryon Density Through the (Cosmological) Ages

GARY STEIGMAN

Departments of Physics and Astronomy, The Ohio State University
174 West 18th Avenue, Columbus, OH 43210, USA
steigman@mps.ohio-state.edu

Abstract

The light element abundances probe the baryon density of the universe within a few minutes of the Big Bang. Of these relics from the earliest universe, deuterium is the baryometer of choice. By comparing its primordial abundance inferred from observations with that predicted by Big Bang Nucleosynthesis (BBN), the early universe baryon density is derived. This is then compared to independent estimates of the baryon density several hundred thousand years after the Big Bang and at present, more than 10 billion years later. The excellent agreement among these values represents an impressive confirmation of the standard model of cosmology.

1. Introduction

In the new, precision era of cosmology, **redundancy** will play an increasingly important role, permitting multiple, independent, tests of and constraints on competing cosmological models, and providing a window on systematic errors which can impede progress or send us off in unprofitable directions. To illustrate the value of this approach, here the baryon density of the Universe is tracked from the first few minutes (as revealed by BBN), through the first few hundred thousand years later (as coded in the fluctuation spectrum of the Cosmic Microwave Background – CMB), and up to the present epoch, approximately 10 Gyr after the expansion began. Theory suggests and terrestrial experiments confirm that the baryon number is preserved throughout these epochs, so that the number of baryons ($\equiv$ nucleons) in a comoving volume *should* be unchanged from BBN to today. As a surrogate for tracking a comoving volume, the nucleon density may be compared to the density of CMB relic photons. Except for the additional photons produced when $e^{\pm}$ pairs annihilate, the number of photons in any comoving volume is also preserved. As a result, the baryon density may be traced throughout the evolution of the universe utilizing the nucleon-to-photon ratio $\eta \equiv n_{\rm N}/n_\gamma$. Since the temperature of the CMB fixes the present number density of relic photons, the fraction of the critical mass/energy density in baryons (nucleons) today ($\Omega_{\rm B} \equiv \rho_{\rm B}/\rho_{\rm crit}$) is directly related to η by, $\eta_{10} \equiv 10^{10}\eta = 274\Omega_{\rm B}h^2$, where the Hubble parameter is $H_0 \equiv 100h$ km s^{-1} Mpc^{-1}. According to the HST Key Project, $h = 0.72 \pm 0.08$ [1].

For several decades now the best constraints on η have come from the comparison of the predictions of BBN with the primordial abundances of the relic nuclides D, ^{3}He, ^{4}He, and ^{7}Li, as inferred from observational data. New data, of similar accuracy, which will soon be available, will enable quantitative probes of the baryon density at later epochs in the evolution of the universe. Indeed, recent CMB data has very nearly achieved this goal. Although a comparable level of precision is currently lacking for the present universe estimates, the data do permit comparisons of independent estimates of η (or $\Omega_{\rm B}h^2$) at three widely separated eras in the evolution of the universe.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

2. The Baryon Density during the First Few Minutes

During the first few minutes in the evolution of the universe the density and temperature were sufficiently high for nuclear reactions to occur. As the universe expanded, cooled, and became more dilute, the universal nuclear reactor ceased to create or destroy nuclei. The abundances of the light nuclei formed during this epoch are determined by the competition between the time available (the universal expansion rate) and the density of the reactants: neutrons and protons. The abundances of D, ^{3}He, and ^{7}Li are "rate limited", being determined by the competition between the nuclear production/destruction rates and the universal expansion rate. As such, they are sensitive to the nucleon density and have the potential to serve as "baryometers".

Of the light, relic nuclides whose primordial abundances may probe the baryon density, deuterium is the **baryometer of choice**. Its predicted primordial abundance varies sigificantly with the nucleon density (D/H $\propto \eta^{-1.6}$) (see Fig. 1). As a result, a primordial abundance known to, say, 10%, determines the baryon density (η) to $\sim 6\%$; truly precision cosmology! Equally important, BBN is the only astrophysical site where an "interesting" abundance of deuterium may be produced (D/H $\gtrsim 10^{-5}$) [2]; the relic abundance is not enhanced by post-BBN production. Furthermore, as primordial gas is cycled through stars, deuterium is completely destroyed. Because of the small binding energy of the deuteron, this destruction occurs during pre-main sequence evolution, when stars are fully mixed. As a result, the abundance of deuterium will only have decreased (or, remained close to its primordial value) since BBN.

After some false starts, as of early 2001 there were three high-redshift (z), low-metallicity (Z) QSO absorption line systems where deuterium had been reliably detected [3, 4, 5]. Earlier observations [6] of a system with very high D/H are widely agreed to have had insufficient velocity data to rule out contamination of the deuterium absorption by a hydrogen interloper [7]. More recently, detection of deuterium has been claimed for another QSO absorption line system [8, 9]. However, the large variations in the inferred D/H, due to the uncertain velocity structure, render this line of sight inappropriate for determining the primordial deuterium abundance. An apparently more reliable, recent determination along a different line of sight yields a surprisingly small D/H [10]. These current data, along with the D/H values for the local interstellar medium [11] and the pre-solar nebula [12] are shown in Fig. 2 as a function of metallicity. At low

Fig. 1. The band stretching from upper left to lower right is the BBN-predicted deuterium abundance. The horizontal band is the observational estimate of the primordial abundance (see the text). The vertical band provides an estimate of the BBN-derived baryon density.

Fig. 2. The deuterium abundance by number relative to hydrogen derived from observations of high-z, low-Z QSO absorption line systems as a function of metallicity (relative to solar). Also shown are the deuterium abundances inferred for the local interstellar medium [11] and for the presolar nebula [12].

metallicity there should be a deuterium "plateau", whose absence, so far, is notable – and very puzzling. Clearly more data is needed. For the purpose of deriving the early universe baryon density, I will rely on the deuterium abundance proposed by O'Meara et al. [5]: $(\mathrm{D/H})_{\mathrm{BBN}} \equiv 3.0 \pm 0.4 \times 10^{-5}$.

From a careful comparison between the BBN predicted abundance and the adopted primordial value, the baryon density when the universe is less than a half hour old is determined (see Fig. 1): $\eta_{10}(\mathrm{BBN}) = 5.6 \pm 0.5 (\Omega_{\mathrm{B}} h^2 = 0.020 \pm 0.002)$. This is truly a "precision" determination; only time will tell if it is accurate. The likelihood distribution of this BBN-derived baryon density is the curve labelled "BBN" in Figure 5.

3. The Baryon Density a Few Hundred Thousand Years Later

The early universe is radiation dominated and the role of "ordinary" matter (baryons) is subdominant. As the universe expands and cools, the density in non-relativistic matter becomes relatively more important, eventually dominating after a few hundred thousand years. At this stage perturbations begin growing under the influence of gravity and, on scales determined by the relative density of baryons, oscillations in the baryon-photon fluid develop. At redshift ~ 1100, when the electron-proton plasma combines to form neutral hydrogen, the CMB photons are freed to propagate thoughout the universe ("last scattering"). These CMB photons preserve the record of the baryon-photon oscillations through very small temperature fluctuations in their spectrum. These fluctuations, or anisotropies, have been detected by the newest generation of experiments, beginning with COBE [13] and continuing with the exciting early results from BOOMERANG [14, 15] and MAXIMA-1 [16], providing a tool for constraining the baryon density at last scattering. In Fig. 3 the status quo ante is shown. The relative heights of the odd and even peaks in the CMB angular fluctuation spectrum depend on the baryon density and these early BOOMERANG and MAXIMA-1 data favored a "high" baryon density (compare the "BBN case", $\eta_{10} = 5.6$, in the upper panel of Fig. 3 with that for a baryon density some 50% higher shown in the lower panel). These data posed a challenge to the consistency of the standard models of cosmology and particle physics, suggesting that the baryon number may have changed (increased) since BBN. Subsequently, new data from the DASI experiment [17],

Fig. 3. The CMB angular fluctuation spectra for two models which differ only in the adopted baryon density. The BBN inferred baryon density is shown in the upper panel and, for comparison, a higher baryon density model is shown in the lower panel. The data are from the "old" BOOMERANG and MAXIMA-1 observations; see the text for references.

along with the revised and expanded BOOMERANG [18] and MAXIMA [19] data appeared. The new and revised data eliminated the challenge posed by the older results.

Although the extraction of cosmological parameters from the CMB anisotropy spectra can be very dependent on the priors adopted in the analyses (see e.g. [20]), the baryon density inferred when the universe was a few hundred thousand years old is robust. For example, we (Kneller et al. [20]) find, $\eta_{10}(\mathrm{CMB}) = 6.0 \pm 0.6(\Omega_{\mathrm{B}}h^2 = 0.022 \pm 0.002)$. The likelihood distribution for this CMB-determined baryon abundance is shown in Fig. 5 by the curve labelled "CMB". The excellent agreement between the two independent estimates, BBN at a few minutes, and CMB at a few hundred thousand years, represents a spectacular success for the standard model of cosmology and illustrates the great potential for future precision tests of cosmology.

4. The Baryon Density at 10 Gyr

There are a variety of approaches to measuring the baryon density today, or during the very recent past. Many of these depend on assumptions concerning the relation between mass and light or, they require the adoption of a specific model for the growth of structure. The

Fig. 4. The 68% (solid) and 95% (dotted) contours in the $\Omega_\Lambda - \Omega_M$ plane allowed by the SNIa magnitude-redshift data (see the text for references). Geometrically flat models lie along the diagonal line labelled $k = 0$.

approach utilized here attempts to avoid such model-dependent assumptions. Instead, the data from the SNIa magnitude-redshift surveys [21−23], is used along with the *assumption* of a flat universe (this latter receives strong support from the newest CMB data [17−19]) to pin down the total matter density (Ω_M). This is then combined with an estimate of the universal baryon **fraction** ($f_B \equiv \Omega_B/\Omega_M$) derived from studies of the X-ray emission from clusters of galaxies. For more details on this approach, see [24] and [25].

In Fig. 4 are shown the SNIa-constrained 68% and 95% contours in the $\Omega_\Lambda - \Omega_M$ plane. The expansion of the universe is currently accelerating for those models which lie above the (dashed) $q_0 = 0$ line. The $k = 0$ line is for a "flat" (zero 3-space curvature) universe. As shown in [25], adopting the assumption of flatness and assuming the validity of the SNIa data, leads to a reasonably accurate ($\sim$25%) estimate of the present matter density: $\Omega_M(\text{SNIa; Flat}) = 0.28^{+0.08}_{-0.07}$.

As the largest collapsed objects, rich clusters of galaxies provide an ideal probe of the baryon **fraction** in the present/recent universe. Observations of the X-ray emission from clusters of galaxies permit constraints on the hot gas content of such clusters which, when corrected for the baryons in stars (but, unfortunately, not for any dark baryons!), may be used to estimate (or bound) f_B. From observations of the Sunyaev-Zeldovich effect in X-ray clusters, the hot gas fraction may be constrained [26] and used to estimate f_B [27]. The combination of f_B and Ω_M is then used to derive a present-universe ($t_0 \approx 10$ Gyr; $z \lesssim 1$) baryon density: $\eta_{10}(\text{SNIa; Flat}) = 5.1^{+1.8}_{-1.4}(\Omega_B h^2 = 0.019^{+0.007}_{-0.005})$. In Figure 5 the corresponding likelihood distribution for the present universe baryon density is shown labelled by "SNIa". Although the uncertainties are largest for this present-universe value, it is in excellent agreement with the other, independent estimates.

5. Summary – Concordance

The abundances of the relic nuclides produced during BBN reflect the baryon density present during the first few minutes in the evolution of the universe. Of these relics from the early universe, deuterium is the baryometer of choice. Although more data is to be desired,

Fig. 5. The likelihood distributions (normalized to equal areas under the curves) for the baryon-to-photon ratios derived from BBN, from the CMB, and for the present universe (SNIa)

the current data permit reasonably constrained estimates of the abundance of primordial deuterium, leading to a tight constraint on the early universe baryon-to-photon ratio $\eta_{10}(\mathrm{BBN}) = 5.6 \pm 0.5$ ($\Omega_{\mathrm{B}}h^2 = 0.020 \pm 0.002$).

Several hundred thousand years later, when the universe became transparent to the CMB radiation, the baryon density was imprinted on the temperature fluctuation spectrum. In determining the baryon density, the current CMB data have a precision approaching that of BBN: $\eta_{10}(\mathrm{CMB}) = 6.0 \pm 0.6$ ($\Omega_{\mathrm{B}}h^2 = 0.022 \pm 0.002$). The excellent agreement between the BBN and CMB values (see Fig. 5) provides strong support for the standard model of cosmology.

In the present universe most baryons are dark ($\eta(\mathrm{LUM}) \ll \eta(\mathrm{BBN}) \approx \eta(\mathrm{CMB})$), so that estimates of the baryon density some 10 billion years after the expansion began are more uncertain and, often model-dependent. In § 4 an estimate of the total matter density (Ω_{M}) derived from the SNIa magnitude-redshift data was combined with the assumption of a flat universe to derive $\eta_{10}(\mathrm{SNIa;Flat}) = 5.1^{+1.8}_{-1.4}$ ($\Omega_{\mathrm{B}}h^2 = 0.019^{+0.007}_{-0.005}$). Although this estimate is of lower statistical accuracy than those from BBN or the CMB, it is in agreement with them (see Fig. 5). Note that if the mass of dark baryons in clusters were similar to the stellar mass, the present-universe baryon density estimate would increase by $\sim 10\%$, bringing it into essentially perfect overlap with the BBN and CMB values.

Increasingly precise observational data have permitted us to track the baryon density from the big bang to the present. At widely separated epochs, from the first few minutes, through the first few hundred thousand years, to the present, $\sim$ ten billion year old universe, a consistent value for the baryon abundance is revealed. This remarkable concordance of the standard, hot, big bang cosmological model is strikingly revealed by the overlapping likelihood distributions for the universal baryon abundances shown in Fig. 5.

References

[1] W. L. FREEDMAN et al., ApJ **553** (2001) 47.
[2] R. EPSTEIN, J. LATTIMER, and D. N. SCHRAMM, Nature **263** (1976) 198.
[3] S. BURLES and D. TYTLER, ApJ **499** (1998a) 699.
[4] S. BURLES and D. TYTLER, ApJ **507** (1998b) 732.
[5] J. M. O'MEARA et al., ApJ **552** (2001) 718.
[6] J. K. WEBB, R. F. CARSWELL, K. M. LANZETTA, R. FERLET, M. LEMOINE, A. VIDAL-MADJAR, and D. V. BOWEN, Nature **388** (1997) 250.
[7] D. KIRKMAN et al., ApJ **559** (2001) 23.

[8] S. D'Odorico, M. Dessauges-Zavadsky, and P. Molaro, A & A **338** (2001) L1.
[9] S. A. Levshakov, M. Dessauges-Zavadsky, S. D'Odorico, and P. Molaro, ApJ **565** (2002) 696.
[10] M. Pettini and D. V. Bowen, ApJ **560** (2001) 41.
[11] J. L. Linsky and B. E. Wood, Proceedings of IAU Symposium 198, The Light Elements and Their Evolution (L. da Silva, M. Spite, and J. R. Medeiros eds.; ASP Conference Series), (2000) p. 141.
[12] G. Gloeckler and J. Geiss, Proceedings of IAU Symposium 198, The Light Elements and Their Evolution (L. da Silva, M. Spite, and J. R. Medeiros eds.; ASP Conference Series), (2000) p. 224.
[13] C. L. Bennett et al., ApJ **464** (1996) L1.
[14] P. de Bernardis et al., Nature **404** (2000) 955.
[15] A. E. Lange et al., Phys. Rev. D **63** (2001) 042001.
[16] S. Hanany et al., ApJ **545** (2000) L5.
[17] N. W. Halverson et al., preprint (2001) [astro-ph/0104489].
[18] C. B. Netterfield et al., preprint (2001) [astro-ph/0104460].
[19] A. T. Lee et al., ApJ **561** (2001) L1.
[20] J. P. Kneller, R. J. Scherrer, G. Steigman, and T. P. Walker, Phys. Rev. D **64** (2001) 123506.
[21] S. Perlmutter et al., ApJ **483** (1997) 565.
[22] B. P. Schmidt et al., ApJ **507** (1998) 46.
[23] S. Perlmutter et al., ApJ **517** (1999) 565.
[24] G. Steigman, N. Hata, and J. E. Felten, ApJ **510** (1999) 564.
[25] G. Steigman, T. P. Walker, and A. Zentner, preprint (2000) [astro-ph/0012149].
[26] L. Grego et al., ApJ **552** (2001) 2.
[27] G. Steigman, J. P. Kneller, and A. Zentner, Revista Mexicana de Astronomia y Astrofysica **12** (2002) 265.

Fortschr. Phys. **50** (2002) 5–7, 569–574

Noncommutativity of Lepton Mass Matrices: Flavor Mixing and CP Violation

ZHI-ZHONG XING

Institute of High Energy Physics, P.O. Box 918 (4), Beijing 100039, China
xingzz@mail.ihep.ac.cn

Abstract

I follow a lesson learnt from Heisenberg's *matrix* Quantum Mechanics to study the property of lepton flavor mixing and CP violation. I show that the commutator of lepton mass matrices defined in vacuum keeps invariant under terrestrial matter effects within the four-neutrino mixing scheme. A set of model-independent sum rules for neutrino masses are obtained, and they can be generalized to hold for an arbitrary number of neutrino families. Useful sum rules for the rephasing-invariant measures of leptonic CP violation have also been found. Finally I present a generic formula of T-violating asymmetries applicable to the future long-baseline neutrino oscillation experiments.

A lesson learnt from Heisenberg's *matrix* Quantum Mechanics is that the observables are represented by Hermitian operators and their commutator yields a description of their compatibility, i.e., whether the observables can be measured simultaneously or not. As already noticed by some authors [1], the commutator of up- and down-quark mass matrices describes whether they can be diagonalized simultaneously or not. The noncommutativity of fermion mass matrices in mathematics is indeed a consequence of flavor mixing and CP violation in physics. In this talk I am going to demonstrate that the commutator of lepton mass matrices is particularly useful to establish the realtions between the observables of neutrino mixing in vacuum and those in matter.

The motivation to study lepton flavor mixing comes, first of all, from the robust Super-Kamiokande evidence for atmospheric and solar neutrino oscillations [2]. In addition, the $\nu_\mu \to \nu_e$ oscillation has been observed by the LSND Collaboration [3]. A simultaneous interpretation of solar, atmospheric and LSND neutrino oscillation data has to invoke the existence of a light sterile neutrino [4], because they involve three distinct mass-squared differences ($\Delta m^2_{\rm sun} \ll \Delta m^2_{\rm atm} \ll \Delta m^2_{\rm LSND}$). In the four-neutrino mixing scheme, CP violation is generally expected to manifest itself. To measure leptonic CP- and T-violating effects needs a new generation of accelerator neutrino experiments with very long baselines [5]. In such long-baseline experiments the terrestrial matter effects, which are likely to deform the neutrino oscillation patterns in vacuum and to fake the genuine CP- and T-violating signals, must be taken into account. To pin down the underlying dynamics of lepton mass generation and CP violation relies crucially upon how accurately the fundamental parameters of lepton flavor mixing can be measured and disentangled from matter effects [6]. It is therefore desirable to explore possible model-independent relations between the effective neutrino masses in matter and the genuine neutrino masses in vacuum. It is also useful to establish some model-independent relations between the rephasing-invariant measures of CP violation in matter and those in vacuum [7].

This talk aims to show that the commutator of lepton mass matrices is invariant under terrestrial matter effects. As a consequence, some concise sum rules of neutrino masses can be obtained model-independently. Such sum rules can even be generalized to hold for an arbitrary number of neutrino families. Another set of sum rules are derived as well for the

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

rephasing-invariant measures of leptonic CP violation in matter. Finally I present a generic formula of T-violating asymmetries, which is applicable in particular to the future long-baseline neutrino oscillation experiments.

The phenomenon of lepton flavor mixing arises from the mismatch between the diagonalization of the charged lepton mass matrix M_l and that of the neutrino mass matrix M_ν in an arbitrary flavor basis. Without loss of generality, one may choose to identify the flavor eigenstates of charged leptons with their mass eigenstates. In this specific basis, where M_l is diagonal, the lepton flavor mixing matrix V links the neutrino flavor eigenstates directly to the neutrino mass eigenstates. For the admixture of one sterile (ν_s) and three active (ν_e, ν_μ, ν_τ) neutrinos, the explicit form of V reads

$$\begin{pmatrix} \nu_s \\ \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} V_{s0} & V_{s1} & V_{s2} & V_{s3} \\ V_{e0} & V_{e1} & V_{e2} & V_{e3} \\ V_{\mu 0} & V_{\mu 1} & V_{\mu 2} & V_{\mu 3} \\ V_{\tau 0} & V_{\tau 1} & V_{\tau 2} & V_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_0 \\ \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{1}$$

where ν_i (for $i = 0, 1, 2, 3$) denote the mass eigenstates of four neutrinos. The effective Hamiltonian responsible for the propagation of neutrinos in vacuum can be written as

$$\mathcal{H}_{\rm eff} = \frac{1}{2E} \left(M_\nu M_\nu^\dagger \right) = \frac{1}{2E} \left(V D_\nu^2 V^\dagger \right), \tag{2}$$

where $D_\nu \equiv {\rm Diag}\,\{m_0, m_1, m_2, m_3\}$, m_i are the neutrino mass eigenvalues, and $E \gg m_i$ denotes the neutrino beam energy. When active neutrinos travel through a normal material medium (e.g., the earth), which consists of electrons but of no muons or taus, they encounter both charged- and neutral-current interactions with electrons. The neutral-current interaction is universal for ν_e, ν_μ and ν_τ, while the charged-current interaction is associated only with ν_e. Their effects on the mixing and propagating features of neutrinos have to be taken into account in all long-baseline neutrino oscillation experiments. Let us use $\tilde{M}_\nu$ and $\tilde{V}$ to denote the effective neutrino mass matrix and the effective flavor mixing matrix in matter, respectively. Then the effective Hamiltonian responsible for the propagation of neutrinos in matter can be written as

$$\tilde{\mathcal{H}}_{\rm eff} = \frac{1}{2E} \left(\tilde{M}_\nu \tilde{M}_\nu^\dagger \right) = \frac{1}{2E} \left(\tilde{V} \tilde{D}_\nu^2 \tilde{V}^\dagger \right), \tag{3}$$

where $\tilde{D}_\nu \equiv {\rm Diag}\,\{\tilde{m}_0, \tilde{m}_1, \tilde{m}_2, \tilde{m}_3\}$, and $\tilde{m}_i$ are the effective neutrino mass eigenvalues in matter. The deviation of $\tilde{\mathcal{H}}_{\rm eff}$ from $\mathcal{H}_{\rm eff}$ is given by

$$\Delta\mathcal{H}_{\rm eff} \equiv \tilde{\mathcal{H}}_{\rm eff} - \mathcal{H}_{\rm eff} = \begin{pmatrix} a' & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{4}$$

where $a = \sqrt{2}\, G_{\rm F} N_e$ and $a' = \sqrt{2}\, G_{\rm F} N_n / 2$ with N_e and N_n being the background densities of electrons and neutrons [8], respectively.

Now let me introduce the commutators of 4×4 lepton mass matrices to describe the flavor mixing of one sterile and three active neutrinos. Without loss of any generality, I continue to work in the afore-chosen flavor basis, where M_l takes the diagonal form $D_l = {\rm Diag}\,\{m_s, m_e, m_\mu, m_\tau\}$ with $m_s = 0$. Note that I have assumed the $(1, 1)$ element of D_l to be zero, because there is no counterpart of the sterile neutrino ν_s in the charged

lepton sector. We shall see later on that our physical results are completely independent of m_s, no matter what value it may take. The commutator of lepton mass matrices in vacuum and that in matter can then be defined as

$$\begin{aligned} C &\equiv i[M_\nu M_\nu^\dagger,\, M_l M_l^\dagger] = i[V D_\nu^2 V^\dagger,\, D_l^2]\,, \\ \tilde{C} &\equiv i[\tilde{M}_\nu \tilde{M}_\nu^\dagger,\, M_l M_l^\dagger] = i[\tilde{V} \tilde{D}_\nu^2 \tilde{V}^\dagger,\, D_l^2]\,. \end{aligned} \tag{5}$$

Obviously C and $\tilde{C}$ are traceless Hermitian matrices. In terms of neutrino masses and flavor mixing matrix elements, I obtain the explicit expressions of C and $\tilde{C}$ as follows:

$$\begin{aligned} C &= i \begin{pmatrix} 0 & \Delta_{es} Z_{se} & \Delta_{\mu s} Z_{s\mu} & \Delta_{\tau s} Z_{s\tau} \\ \Delta_{se} Z_{es} & 0 & \Delta_{\mu e} Z_{e\mu} & \Delta_{\tau e} Z_{e\tau} \\ \Delta_{s\mu} Z_{\mu s} & \Delta_{e\mu} Z_{\mu e} & 0 & \Delta_{\tau\mu} Z_{\mu\tau} \\ \Delta_{s\tau} Z_{\tau s} & \Delta_{e\tau} Z_{\tau e} & \Delta_{\mu\tau} Z_{\tau\mu} & 0 \end{pmatrix}, \\ \tilde{C} &= i \begin{pmatrix} 0 & \Delta_{es} \tilde{Z}_{se} & \Delta_{\mu s} \tilde{Z}_{s\mu} & \Delta_{\tau s} \tilde{Z}_{s\tau} \\ \Delta_{se} \tilde{Z}_{es} & 0 & \Delta_{\mu e} \tilde{Z}_{e\mu} & \Delta_{\tau e} \tilde{Z}_{e\tau} \\ \Delta_{s\mu} \tilde{Z}_{\mu s} & \Delta_{e\mu} \tilde{Z}_{\mu e} & 0 & \Delta_{\tau\mu} \tilde{Z}_{\mu\tau} \\ \Delta_{s\tau} \tilde{Z}_{\tau s} & \Delta_{e\tau} \tilde{Z}_{\tau e} & \Delta_{\mu\tau} \tilde{Z}_{\tau\mu} & 0 \end{pmatrix}, \end{aligned} \tag{6}$$

where $\Delta_{\alpha\beta} \equiv m_\alpha^2 - m_\beta^2$ for $\alpha \neq \beta$ running over $(s,\, e,\, \mu,\, \tau)$, and

$$Z_{\alpha\beta} \equiv \sum_{i=0}^{3} (m_i^2 V_{\alpha i} V_{\beta i}^*)\,, \qquad \tilde{Z}_{\alpha\beta} \equiv \sum_{i=0}^{3} (\tilde{m}_i^2 \tilde{V}_{\alpha i} \tilde{V}_{\beta i}^*)\,. \tag{7}$$

One can see that $\Delta_{\beta\alpha} = -\Delta_{\alpha\beta}$, $Z_{\beta\alpha} = Z_{\alpha\beta}^*$ and $\tilde{Z}_{\beta\alpha} = \tilde{Z}_{\alpha\beta}^*$ hold.

To find out how $\tilde{Z}_{\alpha\beta}$ is connected with $Z_{\alpha\beta}$, I need to establish the relation between $\tilde{C}$ and C. Taking Eqs. (2−4) into account, I immediately obtain

$$\tilde{C} = 2iE[\tilde{\mathcal{H}}_{\rm eff},\, D_l^2] = C + 2iE[\Delta\mathcal{H}_{\rm eff},\, D_l^2] = C\,. \tag{8}$$

This interesting result indicates that the commutator of lepton mass matrices in vacuum is invariant under terrestrial matter effects. As a straightforward consequence of $\tilde{C} = C$, I arrive at $\tilde{Z}_{\alpha\beta} = Z_{\alpha\beta}$ from Eq. (6). Then a set of concise sum rules of neutrino masses emerge [7]:

$$\sum_{i=0}^{3} (\tilde{m}_i^2 \tilde{V}_{\alpha i} \tilde{V}_{\beta i}^*) = \sum_{i=0}^{3} (m_i^2 V_{\alpha i} V_{\beta i}^*)\,; \tag{9}$$

or equivalently

$$\sum_{i=1}^{3} (\tilde{\Delta}_{i0} \tilde{V}_{\alpha i} \tilde{V}_{\beta i}^*) = \sum_{i=1}^{3} (\Delta_{i0} V_{\alpha i}\, V_{\beta i}^*)\,, \tag{10}$$

where $\Delta_{i0} \equiv m_i^2 - m_0^2$ and $\tilde{\Delta}_{i0} \equiv \tilde{m}_i^2 - \tilde{m}_0^2$ for $i = 1,\, 2,\, 3$. It becomes obvious that the validity of Eq. (9) or Eq. (10) has nothing to do with the assumption of $m_s = 0$ in the charged lepton sector. Although I have derived these sum rules in the four-neutrino mixing scheme, they may simply be generalized to hold for an arbitrary number of neutrino families.

It should be noted that $Z_{\alpha\beta}$ and $\tilde{Z}_{\alpha\beta}$ are sensitive to a redefinition of the phases of charged lepton fields. The simplest rephasing-invariant equality is of course $|\tilde{Z}_{\alpha\beta}| = |Z_{\alpha\beta}|$. For the description of CP or T violation in neutrino oscillations, we are more interested in the following rephasing-invariant relationship:

$$\tilde{Z}_{\alpha\beta}\tilde{Z}_{\beta\gamma}\tilde{Z}_{\gamma\alpha} = Z_{\alpha\beta}Z_{\beta\gamma}Z_{\gamma\alpha}\,, \tag{11}$$

for $\alpha \neq \beta \neq \gamma$ running over (s, e, μ, τ). As one can see later on, the imaginary parts of $Z_{\alpha\beta}Z_{\beta\gamma}Z_{\gamma\alpha}$ and $\tilde{Z}_{\alpha\beta}\tilde{Z}_{\beta\gamma}\tilde{Z}_{\gamma\alpha}$ are related respectively to leptonic CP violation in vacuum and that in matter.

It should also be noted that the results obtained above are only valid for neutrinos propagating in vacuum and in matter. As for antineutrinos, the corresponding formulas can straightforwardly be written out from Eqs. (3)–(11) through the replacements $V \Rightarrow V^*$, $a \Rightarrow -a$ and $a' \Rightarrow -a'$.

In terms of the matrix elements of V or $\tilde{V}$, one may define the rephasing-invariant measures of CP violation as follows:

$$J^{ij}_{\alpha\beta} \equiv \mathrm{Im}\,(V_{\alpha i}V_{\beta j}V^*_{\alpha j}V^*_{\beta i})\,, \qquad \tilde{J}^{ij}_{\alpha\beta} \equiv \mathrm{Im}\,(\tilde{V}_{\alpha i}\tilde{V}_{\beta j}\tilde{V}^*_{\alpha j}\tilde{V}^*_{\beta i})\,, \tag{12}$$

where the Greek subscripts ($\alpha \neq \beta$) run over (s, e, μ, τ), and the Latin superscripts ($i \neq j$) run over (0, 1, 2, 3). Of course, $J^{ii}_{\alpha\beta} = J^{ij}_{\alpha\alpha} = 0$ and $\tilde{J}^{ii}_{\alpha\beta} = \tilde{J}^{ij}_{\alpha\alpha} = 0$ hold by definition. With the help of the unitarity of V or $\tilde{V}$, one may straightforwardly obtain some correlation equations of $J^{ij}_{\alpha\beta}$ and $\tilde{J}^{ij}_{\alpha\beta}$ [9]. Therefore there are only nine independent $J^{ij}_{\alpha\beta}$ (or $\tilde{J}^{ij}_{\alpha\beta}$) in the four-neutrino mixing scheme under discussion. If only the flavor mixing of three active neutrinos is taken into account, there will be a single independent $J^{ij}_{\alpha\beta}$ (or $\tilde{J}^{ij}_{\alpha\beta}$), redefined as J (or $\tilde{J}$). It is a unique feature of the three-family flavor mixing scenario, for either leptons or quarks [1], that there exists a universal CP-violating parameter.

To establish the relation between $\tilde{J}^{ij}_{\alpha\beta}$ and $J^{ij}_{\alpha\beta}$, I make use of the equality in Eq. (11). The key point is that the imaginary parts of the rephasing-invariant quantities $Z_{\alpha\beta}Z_{\beta\gamma}Z_{\gamma\alpha}$ and $\tilde{Z}_{\alpha\beta}\tilde{Z}_{\beta\gamma}\tilde{Z}_{\gamma\alpha}$,

$$\begin{aligned}
\mathrm{Im}\,(Z_{\alpha\beta}Z_{\beta\gamma}Z_{\gamma\alpha}) &= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\left[\Delta_{i0}\Delta_{j0}\Delta_{k0}\,\mathrm{Im}\,(V_{\alpha i}V_{\beta j}V_{\gamma k}V^*_{\alpha k}V^*_{\beta i}V^*_{\gamma j})\right]\,,\\
\mathrm{Im}\,(\tilde{Z}_{\alpha\beta}\tilde{Z}_{\beta\gamma}\tilde{Z}_{\gamma\alpha}) &= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\left[\tilde{\Delta}_{i0}\tilde{\Delta}_{j0}\tilde{\Delta}_{k0}\,\mathrm{Im}\,(\tilde{V}_{\alpha i}\tilde{V}_{\beta j}\tilde{V}_{\gamma k}\tilde{V}^*_{\alpha k}\tilde{V}^*_{\beta i}\tilde{V}^*_{\gamma j})\right]\,,
\end{aligned} \tag{13}$$

which do not vanish unless leptonic CP and T are good symmetries, amount to each other. The right-hand side of Eq. (13) can be expanded in terms of $J^{ij}_{\alpha\beta}$ and $\tilde{J}^{ij}_{\alpha\beta}$. In doing so, one needs to use Eq. (12) as well as the unitarity conditions of V and $\tilde{V}$ frequently. After some lengthy but straightforward algebraic calculations, I arrive at the following sum rules of CP- or T-violating parameters:

$$\begin{aligned}
&\tilde{\Delta}_{10}\tilde{\Delta}_{20}\tilde{\Delta}_{30}\sum_{i=1}^{3}\left(\tilde{J}^{0i}_{\alpha\beta}\,|\tilde{V}_{\gamma i}|^2 + \tilde{J}^{0i}_{\beta\gamma}\,|\tilde{V}_{\alpha i}|^2 + \tilde{J}^{0i}_{\gamma\alpha}\,|\tilde{V}_{\beta i}|^2\right)\\
&\quad + \sum_{i=1}^{3}\sum_{j=1}^{3}\left[\tilde{\Delta}_{i0}\tilde{\Delta}^2_{j0}\left(\tilde{J}^{ij}_{\alpha\beta}\,|\tilde{V}_{\gamma j}|^2 + \tilde{J}^{ij}_{\beta\gamma}\,|\tilde{V}_{\alpha j}|^2 + \tilde{J}^{ij}_{\gamma\alpha}\,|\tilde{V}_{\beta j}|^2\right)\right]\\
&= \Delta_{10}\Delta_{20}\Delta_{30}\sum_{i=1}^{3}\left(J^{0i}_{\alpha\beta}\,|V_{\gamma i}|^2 + J^{0i}_{\beta\gamma}\,|V_{\alpha i}|^2 + J^{0i}_{\gamma\alpha}\,|V_{\beta i}|^2\right)\\
&\quad + \sum_{i=1}^{3}\sum_{j=1}^{3}\left[\Delta_{i0}\Delta^2_{j0}\left(J^{ij}_{\alpha\beta}\,|V_{\gamma j}|^2 + J^{ij}_{\beta\gamma}\,|V_{\alpha j}|^2 + J^{ij}_{\gamma\alpha}\,|V_{\beta j}|^2\right)\right].
\end{aligned} \tag{14}$$

I would like to remark that this result is model-independent and rephasing-invariant. It may be considerably simplified, once the hierarchy of neutrino masses and that of flavor mixing angles are theoretically assumed or experimentally measured. If one "switches off" the mass of the sterile neutrino and its mixing with active neutrinos (i.e., $a' = 0$, $\Delta_{i0} = m_i^2$, $\tilde{\Delta}_{i0} = \tilde{m}_i^2$, $J_{\alpha\beta}^{0i} = 0$, and $\tilde{J}_{\alpha\beta}^{0i} = 0$), then Eq. (14) turns out to take the form $\tilde{J}\tilde{\Delta}_{21}\tilde{\Delta}_{31}\tilde{\Delta}_{32} = J\Delta_{21}\Delta_{31}\Delta_{32}$. This elegant relationship has been derived in Refs. [7, 10] with the help of the equality $\mathrm{Det}\,(\tilde{C}) = \mathrm{Det}\,(C)$, instead of Eq. (11), in the three-neutrino mixing scheme.

The matter-corrected CP-violating parameters $\tilde{J}_{\alpha\beta}^{ij}$ can, at least in principle, be determined from the measurement of CP- and T-violating effects in a variety of long-baseline neutrino oscillation experiments. The conversion probability of a neutrino ν_α to another neutrino ν_β is given in matter as

$$\tilde{P}(\nu_\alpha \to \nu_\beta) \;=\; \delta_{\alpha\beta} - 4 \sum_{i<j} \left[\mathrm{Re}\,(\tilde{V}_{\alpha i}\tilde{V}_{\beta j}\tilde{V}_{\alpha j}^{*}\tilde{V}_{\beta i}^{*}) \sin^2 \tilde{F}_{ji}\right] - 2 \sum_{i<j} \left[\tilde{J}_{\alpha\beta}^{ij} \sin\,(2\tilde{F}_{ji})\right], \tag{15}$$

where $\tilde{F}_{ji} \equiv 1.27\tilde{\Delta}_{ji}L/E$ with $\tilde{\Delta}_{ji} \equiv \tilde{m}_j^2 - \tilde{m}_i^2$, L stands for the baseline length (in unit of km), and E is the neutrino beam energy (in unit of GeV). The transition probability $\tilde{P}(\nu_\beta \to \nu_\alpha)$ can directly be read off from Eq. (15), if the replacements $\tilde{J}_{\alpha\beta}^{ij} \Rightarrow -\tilde{J}_{\alpha\beta}^{ij}$ are made. To obtain the probability $\tilde{P}(\bar{\nu}_\alpha \to \bar{\nu}_\beta)$, however, both the replacements $J_{\alpha\beta}^{ij} \Rightarrow -J_{\alpha\beta}^{ij}$ and $(a, a') \Rightarrow (-a, -a')$ need be made for Eq. (15). In this case, $\tilde{P}(\bar{\nu}_\alpha \to \bar{\nu}_\beta)$ is not equal to $\tilde{P}(\nu_\beta \to \nu_\alpha)$. The difference between $\tilde{P}(\bar{\nu}_\alpha \to \bar{\nu}_\beta)$ and $\tilde{P}(\nu_\beta \to \nu_\alpha)$ is a false signal of CPT violation, induced actually by the matter effect [6]. Thus the CP-violating asymmetries between $\tilde{P}(\nu_\alpha \to \nu_\beta)$ and $\tilde{P}(\bar{\nu}_\alpha \to \bar{\nu}_\beta)$ are in general different from the T-violating asymmetries between $\tilde{P}(\nu_\alpha \to \nu_\beta)$ and $\tilde{P}(\nu_\beta \to \nu_\alpha)$. The latter can be explicitly expressed as follows:

$$\begin{aligned}\Delta\tilde{P}_{\alpha\beta} &\equiv \tilde{P}(\nu_\beta \to \nu_\alpha) - \tilde{P}(\nu_\alpha \to \nu_\beta) \\ &= 4[\tilde{J}_{\alpha\beta}^{01} \sin\,(2\tilde{F}_{10}) + \tilde{J}_{\alpha\beta}^{02} \sin\,(2\tilde{F}_{20}) + \tilde{J}_{\alpha\beta}^{03} \sin\,(2\tilde{F}_{30}) \\ &\quad + \tilde{J}_{\alpha\beta}^{12} \sin\,(2\tilde{F}_{21}) + \tilde{J}_{\alpha\beta}^{13} \sin\,(2\tilde{F}_{31}) + \tilde{J}_{\alpha\beta}^{23} \sin\,(2\tilde{F}_{32})]\,.\end{aligned} \tag{16}$$

If the hierarchical patterns of neutrino masses and flavor mixing angles are assumed, the expression of $\Delta\tilde{P}_{\alpha\beta}$ may somehow be simplified [11]. Note that only three of the twelve nonvanishing asymmetries $\Delta\tilde{P}_{\alpha\beta}$ are independent, as a consequence of the unitarity of $\tilde{V}$ or the correlation of $\tilde{J}_{\alpha\beta}^{ij}$. Since only the transition probabilities of active neutrinos can be realistically measured, we are more interested in the T-violating asymmetries $\Delta\tilde{P}_{e\mu}$, $\Delta\tilde{P}_{\mu\tau}$ and $\Delta\tilde{P}_{\tau e}$. These three measurables, which are independent of one another in the four-neutrino mixing scheme under discussion, must be identical in the conventional three-neutrino mixing scheme. In the latter case, where $a' = 0$, $\tilde{J}_{\alpha\beta}^{01} = \tilde{J}_{\alpha\beta}^{02} = \tilde{J}_{\alpha\beta}^{03} = 0$, and $\tilde{J}_{\alpha\beta}^{12} = -\tilde{J}_{\alpha\beta}^{13} = \tilde{J}_{\alpha\beta}^{23} = \tilde{J}$ for (α, β) running over (e, μ), (μ, τ) and (τ, e), Eq. (16) can be simplified to $\Delta\tilde{P}_{\alpha\beta}^{(3)} = 16\tilde{J} \sin \tilde{F}_{21} \sin \tilde{F}_{31} \sin \tilde{F}_{32}$ [12]. The overall matter contamination residing in $\Delta\tilde{P}_{\alpha\beta}$ is usually expected to be insignificant. The reason is simply that the terrestrial matter effects in $\tilde{P}(\nu_\alpha \to \nu_\beta)$ and $\tilde{P}(\nu_\beta \to \nu_\alpha)$, which both depend on the parameters (a, a'), may partly (even essentially) cancel each other in the T-violating asymmetry $\Delta\tilde{P}_{\alpha\beta}$. In contrast, $\tilde{P}(\nu_\alpha \to \nu_\beta)$ and $\tilde{P}(\bar{\nu}_\alpha \to \bar{\nu}_\beta)$ are associated respectively with $(+a, +a')$ and $(-a, -a')$, thus there should not have large cancellation of matter effects in the corresponding CP-violating asymmetries.

To summarize, let me remark that the strategy of this talk is to formulate the sum rules of neutrino masses and CP violation in a model-independent way. An obvious sum rule of

neutrino masses is, of course,

$$\tilde{m}_0^2 + \tilde{m}_1^2 + \tilde{m}_2^2 + \tilde{m}_3^2 = m_0^2 + m_1^2 + m_2^2 + m_3^2 + a + a' \,, \tag{17}$$

which arises straightforwardly from Eqs. (2), (3) and (4). Following a lesson learnt from Heisenberg's *matrix* Quantum Mechanics, I have introduced the commutators of lepton mass matrices to describe the flavor mixing phenomenon of three active and one sterile neutrinos. It has been shown that the commutator defined in vacuum is invariant under terrestrial matter effects. An important consequence of this interesting result is the emergence of a set of model-independent sum rules for neutrino masses in two different media. I have also presented some useful sum rules for the rephasing-invariant measures of leptonic CP violation in the four-neutrino mixing scheme. A generic formula of T-violating asymmetries, which is applicable in particular to the future long-baseline neutrino oscillation experiments, has been derived and discussed.

References

[1] C. Jarlskog, Phys. Rev. Lett. **55** (1985) 1039;
H. Fritzsch and Z. Z. Xing, Phys. Lett. B **353** (1995) 114; Nucl. Phys. B **556** (1999) 49; and references therein.

[2] Super-Kamiokande Collaboration, Y. Fukuda et al., Phys. Rev. Lett. **81** (1998) 1562; **81** (1998) 4279; http://www-sk.icrr.u-tokyo.ac.jp/dpc/sk/;

[3] LSND Collaboration, C. Athanassopoulos et al., Phys. Rev. Lett. **81** (1998) 1774.

[4] For a recent review with extensive references, see: H. Fritzsch and Z. Z. Xing, Prog. Part. Nucl. Phys. **45** (2000) 1; hep-ph/9912358.

[5] See, e.g., B. Autin et al., Report No. CERN-SPSC-98-30 (1998); M. G. Catanesi et al., Report No. CERN-SPSC-99-35 (1999); D. Ayres et al., physics/9911009; C. Albright et al., hep-ex/0008064; and references therein.

[6] Z. Z. Xing, Phys. Lett. B **487** (2000) 327;
Phys. Rev. D **64** (2001) 073014.

[7] Z. Z. Xing, Phys. Rev. D **64** (2001) 033005.

[8] L. Wolfenstein, Phys. Rev. D **17** (1978) 2369;
S. P. Mikheyev and A. Yu. Smirnov, Yad. Fiz. (Sov. J. Nucl. Phys.) **42** (1985) 1441.

[9] V. Barger, Y. B. Dai, K. Whisnant, and B. L. Young, Phys. Rev. D **59** (1999) 113010.

[10] P. F. Harrison and W. G. Scott, Phys. Lett. B **476** (2000) 349.

[11] A. Kalliomäki, J. Maalampi, and M. Tanimoto, Phys. Lett. B **469** (1999) 179.

[12] H. Fritzsch and Z. Z. Xing, Phys. Lett. B **517** (2001) 363; Phys. Rev. D **61** (2000) 073016; Phys. Lett. B **440** (1998) 313; Phys. Lett. B **372** (1996) 265.

Fortschr. Phys. **50** (2002) 5–7, 575–582

Quantum Physics

Heisenberg's Uncertainty Relations and Quantum Optics

GIRISH SARAN AGARWAL

Physical Research Laboratory
Navrangpura, Ahmedabad 380 009, India

Abstract

We present a brief review of the impact of the Heisenberg uncertainty relations on quantum optics. In particular we demonstrate how almost all coherent and nonclassical states of quantum optics can be derived from uncertainty relations.

1. Introduction: Heisenberg Uncertainty Relations

Heisenberg's uncertainty relations had tremendous impact in the field of quantum optics particularly in the context of the construction of coherent and other classes of states for different physical systems and in the reconstruction of quantum states. We present very general arguments based on the equality sign in the Heisenberg uncertainty relations to demonstrate a very large class of coherent and nonclassical states for a wide variety of quantum systems such as single mode and two mode radiation fields, quantized motion of trapped ions, collection of spins, and two level atoms. The resulting states for spin systems are especially interesting as these have varied applications starting from the dynamics of a collection of two level atoms to two component Bose condensates. Additionally, spins in such states have strong entanglement.

Consider a quantum mechanical system with two physical observables represented by hermitain operators A and B satisfying the commutation relation

$$[A, B] = iC\,. \tag{1}$$

One then makes use of the Schwarz inequality

$$\langle(\Delta A)^2\rangle\,\langle(\Delta B)^2\rangle \geq |\langle\Delta A\,\Delta B\rangle|^2\,, \tag{2}$$

and Eq. (1) to obtain the uncertainty relation

$$\langle(\Delta A)^2\rangle\,\langle(\Delta B)^2\rangle = \left|\left\langle\frac{1}{2}\,[\Delta A, \Delta B]\right\rangle\right|^2 + \left|\left\langle\frac{1}{2}\,\{\Delta A, \Delta B\}\right\rangle\right|^2 \geq \frac{1}{4}\,|\langle C\rangle|^2\,. \tag{3}$$

One finds that the equality sign holds only if the state of the system satisfies the eigenvalue equation,

$$(\Delta A)|\psi\rangle = -i\lambda(\Delta B)|\psi\rangle, \tag{4}$$

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

where λ is a complex number. We will refer to such states as minimum uncertainty states. Using Eq. (4) we can derive some very general results

$$\langle(\Delta A)^2\rangle = |\lambda|^2 \langle(\Delta B)^2\rangle\,, \tag{5}$$

$$\langle\{\Delta A, \Delta B\}\rangle = \frac{\mathrm{Im}\,\lambda}{\mathrm{Re}\,\lambda}\,\langle C\rangle\,. \tag{6}$$

Thus for $|\lambda| = 1$, the observables have equal variance and there is no correlation between A and B if $\mathrm{Im}\,\lambda = 0$. In what follows we show how the different coherent and nonclassical states introduced [1−9] in quantum optics follow from the eigenvalue Eq. (4). The Eq. (4) also enables us to define in a natural way the squeezed states associated with bosons but also for spin systems and other systems such as those described by $SU(1,1)$ algebra.

System of Bosons: Let us first consider Heisenberg algebra for the position and momentum operators x and p. Then Eq. (4) leads to a wavefunction $\psi(x)$ which is Gaussian. On setting $\lambda = 1$, we recover coherent states as introduced by Glauber, [4] whereas for $\lambda \neq 1$, we recover the squeezed coherent states as introduced by Yuen and others [3, 5, 6]. Note further that the eigenvalue Eq. (4) in terms of the harmonic oscillator annihilation and creation operators a and $a^\dagger([a, a^\dagger] = 1)$ can be written as

$$\begin{aligned} &A = \sqrt{\frac{\hbar}{2}}\,(a + a^\dagger)\,, \qquad B = \sqrt{\frac{\hbar}{2}}\,(a - a^\dagger)/i\,, \\ &(\mu a + \nu a^\dagger)\,|\psi\rangle \equiv \alpha\,|\psi\rangle\,, \\ &\mu = \frac{1}{2}\left(\sqrt{\lambda} + \frac{1}{\sqrt{\lambda}}\right), \qquad \nu = \frac{1}{2}\left(\sqrt{\lambda} - \frac{1}{\sqrt{\lambda}}\right), \qquad \mu^2 - \nu^2 = 1\,. \end{aligned} \tag{7}$$

Thus squeezed coherent state is an eigenstate of the operator obtained by using Bogoliubov transformed annihilation operator. The states as defined by (7) for both $\nu = 0$ and $\nu \neq 0$ have been extensively used in quantum optics in the study of radiation fields and the quantized center of mass motion of trapped ions.

In many situations in quantum optics, for example, in dealing with the spontaneous noise in amplifiers, the state of the field is no longer pure and hence one has to introduce a mixed state. In such situations it turns out especially useful to work with Gaussian Wigner functions of the form [10]

$$\begin{aligned} &w(q, p) = \frac{1}{\sqrt{(2\pi)^2\,(\alpha\beta - \gamma^2)}}\,\exp\left[-\frac{\alpha q^2 + \beta p^2 - 2\gamma qp}{2(\alpha\beta - \gamma^2)}\right], \\ &\alpha = \langle p^2\rangle\,, \qquad \beta = \langle q^2\rangle\,, \qquad \gamma = \frac{1}{2}\,\langle qp + pq\rangle\,. \end{aligned} \tag{8}$$

Here we assume that $\langle p\rangle = \langle q\rangle = 0$. An important feature of (8) is its close connection to the uncertainty relations. This is because every well defined Gaussian is *not* a bonafide Wigner function. The parameters α, β and γ must satisfy the stronger form of the uncertainty inequality

$$\sigma = \left[(\alpha\beta - \gamma^2)^{\frac{1}{2}} - \frac{1}{2}\right] \geq 0\,; \tag{9}$$

Note that for a field in thermal equilibrium the correlation parameter $\gamma = 0$, and in addition the uncertainties in q and p are equal (setting $m = \omega = 1$). Further one finds a very interesting result for the entropy associated with the mixed state (8)

$$S = K_B[(\sigma + 1)\,\mathrm{In}\,(\sigma + 1) - \sigma \ln \sigma]\,. \tag{10}$$

The entropy thus depends in an important way on the correlation parameter γ. The state (8) in fact represents what is now called squeezed thermal state [11], the study of which became very popular in the mid eighties. Furthermore such a state represents squeezed vacuum for $\sigma = 0$. The photon number distributions corresponding to (8) exhibited a number of interesting features [11]. We further note that states like (8) are beginning to be used in quantum information theory, where results like (10) have been rediscovered for the purpose of calculation of the quantum capacity of a channel [12].

2. $SU(1,1)$ Algebra – Entangled States for Photons, Trapped Ions etc.

Consider a set of operators K_i satisfying the algebra

$$[K_+, K_-] \equiv -2K_0\,, \qquad [K_0, K_\pm] = \pm K_\pm\,. \tag{11}$$

This algebra in its various realizations is extremely useful in quantum optics. In its single mode realization $K_- = \frac{1}{2}a^2,\ K_0 = a^\dagger a + \frac{1}{2}$, the eigenvalue Eq. (4) for $A = K_x,\ B = K_y$ and $\lambda = -1$ leads to even and odd states, $|\alpha\rangle \pm |-\alpha\rangle$ with $|\alpha\rangle$ standing for a coherent state. Such states have been extensively studied for their non-classical properties. These have been experimentally realized. In fact the eigenvalue Eq. (4) for the algebra (11)

$$(\mu K_- + \nu K_+)\,|\psi\rangle = \lambda\,|\psi\rangle \tag{12}$$

leads to a very wide variety of nonclassical fields associated with the radiation fields as well as the quantized motion of the center of mass of trapped ions. For a two mode realization of the $SU(1,1)$ group $K_- = ab,\ K_0 \equiv \frac{1}{2}(a^\dagger a + bb^\dagger)$ we get pair coherent states [7],

$$\begin{aligned} &ab\,|\Phi\rangle = \xi\,|\Phi\rangle\,, \qquad (a^\dagger a - b^\dagger b)\,|\Phi\rangle = q\,|\Phi\rangle\,, \\ &|\Phi(\xi, q)\rangle = N(\xi, q)\sum_{n=0}^{\infty} \frac{\xi^n}{\sqrt{n!(n+q)!}}\,|n+q, n\rangle\,, \\ &N(\xi, q) = \left\{\sum_{n=o}^{\infty} \frac{|\xi|^{2n}}{n!(n+q)!}\right\}^{-1/2}. \end{aligned} \tag{13}$$

These pair coherent states have very remarkable nonclassical properties such as entanglement, sub-Poissonian statistics, quadrature squeezing etc. These states have also been considered [13] for the study of EPR-like correlations and violation of Bell inequalities. In Fig. 1 we show the contour plots for the quadrature distribution $\Phi(x, y) = \langle x, y\,|\,\Phi\rangle$ for $\xi = 3$. Note that the quadrature distribution is far from a Gaussian distribution [14]. Originally such states were shown to occur in the competition of two nonlinear optical processes, viz; four-wave mixing and two-photon absorption [15]. More recently a proposal for producing pair coherent states for ionic motion was outlined [16].

Fig. 1. The contour plots of the distribution $|\Phi(x, y)|^2$ for a pair coherent state for two different values of the parameter $q = 0$ (left); $q = 1$ (right)

3. $SU(2)$ Algebra – Coherent, Squeezed and Entangled States for a Collection of Spins

We devote rest of the paper to the problem of different types of states for a collection of spin systems [8, 9, 17, 18]. This would cover a very wide class of systems in physics. We give some typical examples:

(a) Quantum Properties of Partially Polarized Light
Consider a two mode radiation field with the two modes representing two orthogonal polarizations of light. The field is in general, partially polarized. We use Schwinger boson representation to express two modes in terms of spin operators $\vec{S} : S^\dagger = a^\dagger b, S^z = \frac{1}{2}(a^\dagger a - b^\dagger b)$. The usual optical elements like beam splitters, wave plates are described by unitary transformations of the form $\exp(i\theta\vec{S}.\vec{n})$, where $\vec{n}$ is a unit vector. Note further that the traditional Stokes parameters will be given by the expectation value of $\vec{S}$ in the state of the radiation field. Since the state of the field could be a nonclassical state, the Stokes parameters could exhibit very significant quantum fluctuations and these could be studied in terms of variances like $\langle S_\alpha S_\beta\rangle - \langle S_\alpha\rangle\langle S_\beta\rangle$. Clearly Heisenberg uncertainty relations will be especially relevant in the study of quantum fluctuations in Stokes parameters.

(b) Collection of Two Level Atoms
Such a system plays a very fundamental role in quantum optics. This system is equivalent to a collection of spins. It has been shown that squeezed states of a collection of spins can be produced by interaction with squeezed light i.e. we have a transfer of squeezing from field to atoms [9]. Furthermore the dispersive interaction of a collection of atoms in a cavity leads to an effective interaction $S^z S^z$ which can lead to the generation of cat like states for spin systems.

(c) Quantum Dots
Here electron-hole interactions are especially relevant and a study of collective effects can be done by introducing spin operators [19]

$$S^+ = \sum_{n=0}^{N} e_n^+ h_n^+ . \tag{14}$$

The inter dot coupling can be reduced approximately to an effective interaction of the form $S^z S^z$, which as we will see leads to the production of entangled states.

(d) Two component Bose Condensates

If we approximate each component of the condensate by a single mode, then the effective Hamiltonian reduces to [20]

$$H = E_a a^\dagger a + E_b b^\dagger b + \alpha(a^\dagger a)^2 + \beta(b^\dagger b)^2 + \gamma a^\dagger a b^\dagger b^+ (g a^\dagger b + g^\dagger a b)\,. \tag{15}$$

Using the conservation laws, this Hamiltonian can be written in terms of spin operators as the sum of a quadratic term $\eta S^z S^z$ and linear terms. The quadratic term gives us the possibility of the introduction of squeezed and other states for a two component Bose condensate.

Other examples where states of a collection of spin systems are especially useful include: (i) Heisenberg exchange interaction, (ii) a suitably oriented state of an atomic system corresponding to a given F value [21]. Such a oriented state can be produced by suitable optical pumping methods.

We start with the eigenvalue equation (4) for the spin problem written in the form

$$((\vec{a}.\hat{S}) + i\eta(\vec{b}.\hat{S}))\,|\psi\rangle = \lambda\,|\psi\rangle\,, \tag{16}$$

where we choose $\vec{a}$ and $\vec{b}$ to be perpendicular to each other. In particular we can choose $\vec{a} = (\hat{\theta}\cos\phi - \hat{\phi}\sin\phi)$, $\vec{b} = \hat{\theta}\sin\phi + \hat{\phi}\cos\phi$, where $\hat{\theta}$ and $\hat{\phi}$ are the unit vectors in three dimensional spherical polar coordinate system. Let $\cup(\theta,\,\phi)$ be the unitary operator

$$\cup(\theta,\,\varphi) = \exp\left(\xi S^\dagger - \xi^* S^-\right), \qquad \xi = \frac{\theta}{2}\,e^{-i\phi}\,. \tag{17}$$

The general solution of (16) for $\eta = \pm 1$ is given by

$$|\psi,\,\pm\rangle = \cup(\theta,\,\varphi)\,|S,\,\pm S\rangle\,, \tag{18}$$

and for $\eta \neq \pm 1$ by

$$|\Psi_m\rangle = \mathcal{N}_m \cup(\theta,\,\phi)\,e^{\mu S_z}\,e^{-i\frac{\pi}{2}S_y}\,|S,\,m\rangle\,, \qquad \eta = \tanh\mu\,, \tag{19}$$

where $\mathcal{N}_m$ is a normalization constant. The states defined by (18) are the atomic coherent states which were studied extensively by Arecchi et al. [8]. The states defined by (19) are the squeezed states for a system of spins. These states were studied extensively by Agarwal and Puri [9, 22]. It is important to note that both coherent states and the squeezed states for a system of spins follow from the Heisenberg uncertainty relations for an arbitrary combination of spin operators $\vec{a}.\vec{S}$ and $\vec{b}.\vec{S}$ with $\vec{a}.\vec{b} = 0$.

The squeezed states for the spin system as defined by (19) has some remarkable properties which are similar to the properties for the squeezed states for the radiation field. In particular the state $|\psi_0\rangle$ for $\theta = 0$ is the analog of the squeezed vacuum. We summarize some of the important properties of the state $|\psi_0\rangle$:

(i) Strong entanglement between different spins in contrast to the states (18) which show no spin-spin correlations. In particular the state $|\psi_0\rangle$ for $S = 1$ has the form $\alpha\,|\uparrow\uparrow\rangle + \beta\,|\downarrow\downarrow\rangle$.
(ii) For integral values of S, the probability $p(\ell)$, of finding the spins in the collective state $|S,\,l\rangle$ is an oscillatory function of l and is zero for odd values of l.
(iii) The phase distribution associated with the state $|\psi_0\rangle$ exhibits a kind of bifurcation [23] which is similar to that exhibited by squeezed vacuum associated with the radiation field.

(iv) The atomic squeezed states can be produced by transferring squeezing from the radiation field to the atomic system.
(v) An important parameter ξ, which is a measure of how much spectroscopic resolution can be achieved by using squeezed states, defined by [22]

$$\xi = \sqrt{2S}\,(\Delta S_x)/|\langle S_z\rangle| \tag{20}$$

can be made much smaller than 1.

The atomic squeezed states as indicated above exhibit oscillatory population distributions [9]. The origin of these oscillations can be traced back to different pathways that contribute to the amplitude of oscillations [24]. The zeros in the population distribution are due to complete destructive interference between two pathways. This is easily seen by using the relation between atomic system and an equivalent bosonic system. The atomic states can be written in terms of an equivalent two mode bosonic system as follows: $S^+ = a^\dagger b$, $S^- = ab^\dagger$, $S^z = \frac{a^\dagger a - b^\dagger b}{2}$, $a^\dagger a + b^\dagger b =$ conserved,

$$|S, m\rangle \leftrightarrow |S+m, S-m\rangle_p\,, \qquad a^\dagger a\,|S+m, S-m\rangle_p = (S+m)\,|S+m, S-m\rangle_p\,. \tag{21}$$

The amplitude for finding the atoms in the state $|S, m\rangle$ in an atomic squeezed state is proportional to

$$\langle S, m \mid \psi_0\rangle \propto_p \langle S, S|\exp\left\{\frac{\pi}{4}\,\left(a^\dagger b - ab^+\right)\right\}|S+m, S-m\rangle_p\,. \tag{22}$$

The right hand side of Eq. (22) is easily interpreted in terms of a beam splitter picture – involving incident photons or two ports of the beam splitter. It is the probability amplitude of finding the outgoing field in the state $|S, S\rangle_p$ given that the input field was in the state $|S+m, S-m\rangle_p$. The complete destructive interference can be understood by examining the various pathways leading to S photon on each of the two output ports.

4. Superposition of Atomic Coherent States: Cat like states

We next discuss properties and production of cat like states for a collection of spins in analogy to similar states for the radiation field. i.e. we consider a superposition of atomic coherent states of the form $|\theta_1, \phi_1\rangle + |\theta_2, \phi_2\rangle$. The cat like states for bosons have been produced in different ways, for example a coherent field passing through a Kerr medium leads to the production of cat like states for certain values of the propagation length (time).

We already know that atomic coherent states can be produced by applying a coherent drive on a collection of spins in ground state. For producing atomic cat like states, we therefore search for an analog of Kerr medium for the spin system [17]. Let us consider the dynamics of a collection of atoms in a dispersive cavity described by

$$H = \hbar\omega_0 S^z + \hbar\omega_c a^\dagger a + \hbar g(S^+ a + a^\dagger S^-)\,. \tag{23}$$

We assume the cavity to be of high quality and possibly a small number of thermal photons. We assume large detuning: $\omega_0 - \omega_c \equiv \delta_c \gg g$. Under these conditions the cavity modes can be adiabatically eliminated and one obtains an effective interaction h

$$h \equiv \hbar\eta\left[\frac{N}{2}\left(\frac{N}{2}+1\right) - S^{z^2} + (2\bar{n}+1)S^z\right], \qquad \eta = \frac{g^2\delta_c}{k^2+\delta_c^2}\,, \tag{24}$$

which is quadratic in S^z. Such a quadratic term can be interpreted as arising from the shift of energy levels by cavity vacuum (Lamb shift). It may be added that a quadratic interaction for an extended system can lead to atomic solitons.

We next demonstrate how (24) leads to the formation of atomic or cat states. Consider the time evolution for a system initially in a coherent state $(\bar{n} = 0)$

$$|\Psi(t)\rangle \equiv e^{-iht/\hbar}\,|\theta,\,\phi\rangle = \sum_{K=0}^{N} \binom{N}{K}^{1/2} \times e^{ik\phi}\,\sin^{N-K}\left(\frac{\theta}{2}\right)\cos^{K}\left(\frac{\theta}{2}\right) e^{-[i(N-K)\,(K+1)\,\eta t]}\left|\frac{N}{2},\,\frac{N}{2}-K\right\rangle. \tag{25}$$

In order to simplify (25) we note that the exponent which is quadratic in K has certain periodicity properties. We can write this exponent as a Fourier series and then each term in the Fourier series can be shown to result in a coherent state. Calculations show that

$$|\Psi\rangle = e^{\frac{-i\pi N}{m}} \sum_{q=0}^{m-1} f_q^{(0)} \left|\theta,\,\phi+\pi\,\frac{2q-N}{m}\right\rangle, \qquad t=\frac{\pi}{m\eta}\,, \qquad m=\text{odd}\,, \tag{26}$$

$$|\Psi\rangle = e^{\frac{-i\pi N}{m}} \sum_{q=0}^{m-1} f_q^{(0)} \left|\theta,\,\phi+\pi\,\frac{2q-N+1}{m}\right\rangle, \qquad t=\frac{\pi}{m\eta}\,, \qquad m=\text{even}\,, \tag{27}$$

where the $f's$ are the Fourier coefficients. We have thus shown how an effective hamiltonian which is quadratic in the collective operator S^z can produce superpositions of atomic or spin coherent states. Such states exhibit strong entanglement among individual spins. Note that the production of such entangled states is also very relevant to systems like two component Bose condensates and quantum dots, as such systems, under reasonable approximations, can be described by an effective Hamiltonian which is quadratic in S^z. In particular for $m = 2$, we generate a superposition of two coherent states with the same θ, but different ϕ,

$$|\Psi(t)\rangle = \frac{e^{-iN\pi/2}}{\sqrt{2}}\left[e^{i\pi/4}\left|\theta,\,\phi-\pi\,\frac{N-1}{2}\right\rangle + e^{-i\pi/4}\left|\theta,\,\phi-\pi\,\frac{N-3}{2}\right\rangle\right]. \tag{28}$$

Recently considerable success has been achieved with regard to spin squeezing [25] both in the context of Bose Condensates and atoms in cells. Finally, we mention that a general phase space theory for spin systems has been developed [26]. This theory takes into account the conservation of the angular momentum S^2 and its finite length. The applications of this theory to various problems like EPR correlations have been considered. Besides it is understood how to reconstruct [27] such phase space distributions and the density matrix from a variety of measurements.

Clearly Heisenberg's uncertainty relations had a deep impact on the field of quantum optics and many independent developments can be viewed from the point of a central theme viz. the uncertainty relations.

The author thanks some of his long time collaborators and in particular R. R. Puri for participating in many projects.

References

[1] R. R. PURI, *Mathematical Methods of Quantum Optics*, (Springer, Berlin 2001).
[2] JOHN R. KLAUDER and BO-STURE SKAGERSTAM, *Coherent States*, (World Scientific, Singapore, 1984).

[3] A. Perelomov, *Generalized Coherent States and their Applications*, (Springer-Verlag, London, 1985).
[4] R. J. Glauber, Phys. Rev. **131** (1963) 2766.
[5] H. P. Yuen, Phys. Rev. A **13** (1976) 2226.
[6] C. M. Caves and B. L. Schumaker, Phys. Rev. A **31** (1985) 3068, 3093.
[7] G. S. Agarwal, J. Opt. Soc. Am. B **5** (1988) 180.
[8] F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, Phys. Rev. A **6** (1972) 2211.
[9] G. S. Agarwal and R. R. Puri, Phys. Rev. A **41** (1990) 3782.
[10] G. S. Agarwal, Phys. Rev. A **3** (1971) 828.
[11] G. S. Agarwal and G. Adam, Phys. Rev. A **38** (1988) 750; ibid, **39** (1989) 6259.
[12] A. S. Holevo and R. F. Werner, Phys. Rev. A **63** (2001) 032312.
[13] A. Gilchrist, P. Deuar, and M. Reid, Phys. Rev. Lett. **80** (1998) 3169;
A. Gilchrist and W. I. Munro, J. Opt. B **2** (2000) 47.
[14] J. Banerji and G. S. Agarwal, Phys. Rev. A **59** (1999) 4777.
[15] G. S. Agarwal, Phys. Rev. Lett. **57** (1986) 827.
[16] S. C. Gou, J. Steintach, and P. L. Knight, Phys. Rev. A **54** (1996) R1014.
[17] G. S. Agarwal, R. R. Puri, and R. P. Singh, Phys. Rev. A **56** (1997) 2249.
[18] M. Kitagawa and M. Ueda, Phys. Rev. A **47** (1993) 5138.
[19] Luis Quiroga and Neil F. Johnson, Phys. Rev. Lett. **83** (1999) 2270.
[20] S. Raghavan, H. Pu, P. Meystre, and N. P. Bigelow, Opt. Commun. **188** (2001) 149.
[21] G. Klose, G. Smith, and P. S. Jessen, Phys. Rev. Lett. **86** (2001) 4721.
[22] G. S. Agarwal and R. R. Puri, Phys. Rev. A **49** (1994) 4968.
[23] G. S. Agarwal and R. P. Singh, Phys. Lett. **217A** (1996) 215.
[24] G. S. Agarwal, J. Mod. Opt. **43** (1996) 223.
[25] A. Sorensen, L-M. Duan, J. I. Cirac, and P. Zoller, Nature **409** (2001) 63;
A. Kuzmich, L. Mandel, and N. P. Bigelow, Phys. Rev. Lett. **85** (2000) 1594;
B. Julsgaard, A. Koghekin, and E. S. Polzik, Nature **413** (2001) 400;
A. Sorensen and K. Molmer, quant-ph/0011035.
[26] G. S. Agarwal, Phys. Rev. A **24** (1981) 2889;
J. Dowling, G. S. Agarwal, and W. P. Schleich, Phys. Rev. A **49** (1994) 4101.
[27] G. S. Agarwal, Phys. Rev. A **57** (1998) 671.

Fortschr. Phys. **50** (2002) 5–7, 583–593

Rubidium Condensate for Quantum Optics Studies

ENNIO ARIMONDO

INFM, Dipartimento di Fisica, Università di Pisa, Via Buonarroti 2, I-56127 Pisa, Italy
and Abteilung für Quantenphysik, Universität Ulm, D-89069, Ulm, Germany
arimondo@df.unipi.it

Abstract

Different experiments with Bose-Einstein rubidium condensates are presented in order to evidence the impact of condensate samples for the research in quantum optics and quantum physics. Measurements of the condensate free expansion, of the geometric forces acting on the condensate trapped into a time-orbiting-potential triaxial magnetic trap, of the acceleration of the condensate through a one-dimensional, off-resonant optical lattices and finally of the photoionization of the rubidium atoms composing the condensate are presented.

1. Introduction

Since the first experimental realizations, many aspects of Bose-Einstein condensed atomic clouds (BECs) have been studied [1], ranging from collective excitations to superfluid properties and quantized vortices. Bose-Einstein condensates constitute a special class of samples. Their preparation, based on an accumulated and precise knowledge of atomic physics, requires a large quantum optics effort for manipulating atoms by light. Also the handling and detection of the condensates is based on laser spectroscopy techniques. However the properties of the final product often resembles those of solid state physics so that different knowledge should be combined in their study. The different facets of this research area will be illustrated through experiments performed on a rubidium condensate.

BEC in alkali atoms within magnetic traps has been achieved in two different kinds of magnetic traps: static traps of the Ioffe-Pritchard type, and dynamic ones, also called time-orbiting-potential (TOP) traps. In the latter a quadrupole magnetic field is modified by a superimposed rotating bias field which eliminates the zero of the magnetic field at the origin and creates an effective harmonic potential when averaged over the fast rotation of the instantaneous field. This time-averaged approximation accounts for most of the observed features of these traps. There are, however, residual effects, leading to a micromotion of the atoms at the frequency of the rotating bias field and a dependence of their equilibrium position in the trap on the sense of rotation of the bias field. Describing the condensate confinement of TOP trap, the temporal evolution of fast coordinates (the atomic spin variables) can be separate from the slow coordinates (the atomic center of mass motion). As result of such a temporal separation, reaction, or geometric, forces acting on the slow dynamics influence the motion of a condensate within the trap. The evidence for geometric forces was made clear in precise measurements of the atomic equilibrium position in the trap when inverting the sense of rotation of the magnetic trap bias field, and of the dependence of the rotation-sensitive equilibrium position on the applied magnetic fields [2]. Those measurements will be here briefly reported.

The properties of ultra-cold atoms in periodic light-shift potentials in one, two and three dimensions have been studied extensively in the past ten years [3, 4]. Optical lattices represent a perfect periodic potential that confine the condensate and can be modified with large

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

degree of freedom. In our recent experimental investigation [5] a rubidium condensate was loaded into one-dimensional, off-resonant optical lattices and accelerated by chirping the frequency difference between the two laser beams forming the optical lattice. For small values of the condensate acceleration, Bloch oscillations were observed. When the acceleration was increased, Landau-Zener tunneling out of the lowest lattice Bloch energy band, leading to a breakdown of the oscillations, was also studied. The presence of mean-field interactions within the condensate can be described through an optical lattice effective potential [6]. From the Landau-Zener tunneling, the effective potential was measured for various condensate densities and trap geometries, yielding good agreement with theoretical calculations.

Photoionization of cold atomic sample is an interesting topic at the interface between quantum optics, atomic physics, plasma physics and quantum statistics. The photoionization process of a Bose-Einstein condensate near the atomic threshold would produce a sample with a bosonic character interacting with fermionic gases of electron and positive ions. The positive and negative charges moving through the condensate produce a local deformation that modifies their motion. Furthermore a theoretical analysis of photoionization of a BEC by monochromatic laser light [7] has predicted that because of the coherent nature of the initial atomic ensemble and the narrow spectral width of the laser ionization source, the occupation number of the electron/ion final states can become close to unity, especially for excitation close to threshold. Under those conditions the condensate ionization rate should be slowed down by a Pauli blockade process and finally determined by a balance between the laser light acting on the atoms and the rate of escape of the charged products from the system. Preliminary experiments of irradiation a BEC of cold rubidium atoms with laser pulses ionizing through one-photon and two-photon absorption have been performed, and their progress will be reviewed.

In order to make a close connection with the quantum optics, the expansion of a condensate following the switching off of confining the magnetic trap will be also discussed.

Sections II discusses the experimental apparatus developed for the Bose-Einstein condensation of rubidium and the observations of the condensate properties. The following four Sections report the results the experimental investigations on the expansion of the condensate, on the geometric forces, on the condensate properties within a one-dimensional (1-D) optical lattice, and finally on the laser photoionization of the condensate. A conclusion completes the present work.

2. Experimental Apparatus

Our experimental apparatus and procedures for creating BECs are described in detail elsewhere [8]. In our apparatus 5×10^7 atoms captured in a magneto-optical trap (MOT) were transferred into a triaxial time-orbiting potential trap (TOP) [9]. Subsequently, the atoms were evaporatively cooled down to the transition temperature for Bose-Einstein condensation, and after further cooling we obtained condensates of $\approx 2 \times 10^4$ atoms without a discernible thermal component in a magnetic trap with 15−30 Hz frequencies.

The confining TOP magnetic trap consists of a pair of quadrupole coils oriented along a horizontal axis and two pairs of bias-field coils, one incorporated into the quadrupole coils and the other along a horizontal axis perpendicular to that of the quadrupole coils. The instantaneous magnetic field seen by the atoms can be written as

$$\vec{B} = \hat{\boldsymbol{i}}[2b'x + B_0 \sin(\Omega_0 t)] + \hat{\boldsymbol{j}}[b'y + B_0 \cos(\Omega_0 t)] + \hat{\boldsymbol{k}}[b'z], \tag{1}$$

where b' is the quadrupole gradient along the z-axis, and B_0 and Ω_0 are the magnitude and the frequency of the bias field, respectively.

In the experiments on optical lattices, either the magnetic trap was switched off and a horizontal 1-D optical lattice was switched on, or the interaction between the condensate and the lattice took place inside the magnetic trap, which was subsequently switched off to allow time-of-flight imaging. The lattice beams were created by a 50 mW diode slave-laser injected by a grating-stabilized master-laser blue-detuned by $\Delta \approx 28-35$ GHz from the ^{87}Rb resonance line. We realized a counter-propagating lattice geometry with $\theta = 180$ deg and an angle-tuned geometry with $\theta = 29$ deg. The lattice constant is $d = \pi/\sin(\theta/2)\, k_l$, with k_l the laser wavenumber, and θ the angle between the two laser beams creating the 1-D optical lattice. For the counterpropagating and angle-tuned configurations the lattice constants d are respectively 0.39 and 1.56 μm, a large lattice constant also implying that only a few potential wells were occupied by the condensate. An acceleration of the lattice was effected by applying a linear ramp to the δ detuning between the two lattice beams.

3. Condensate Expansion

As characteristic test applied to probe the formation of a Bose-Einstein condensate, several groups examined the ballistic evolution of the condensate cloud after its release from the magnetic trap [1]. The presence of strong repulsive interactions between the condensate atoms leads to a deformation in the aspect ratio of the trap during the ballistic expansion. A similar test was initially applied also to our condensate [8, 10], with results different from those of other groups because our system operates in conditions close to the Heisenberg indetermination limit. Experimental results for the expansions of condensate are reported in Fig. 1. The dimensions Δz and Δx along the vertical axis and along the strong quadrupole axis have been measured for different ballistic expansion times t_e. Owing to the limiting resolution of our imaging system at the time of the test, the cloud size was limited by the pixel resolution on the CCD camera. Three

Fig. 1. In (a) and (b) time dependencies of the squares of the condensate dimensions Δz and Δx versus the ballistic expansion time t_e for $N = 1000$ atoms. The theoretical curves for the condensate expansion correspond to the Gross-Pitaevskii numerical integration (continuous line), the Thomas-Fermi limit (dashed line) and the harmonic oscillator limit (dashed line).

theoretical curves are reported in the figure. A numerical solution of the Gross-Pitaevskii equation for the expanding magnetic trap reproduces the experimental values. However a physical understanding of the phenomenon regulating the condensate expansion is obtained examining two different limiting cases. For the Thomas-Fermi limit, where the kinetic energy term is neglected in respect to the atomic mean field interaction, the expansion has been described through scaling laws in Ref. [11]. That limit predicts an expansion larger than the measured one. Instead for the expansion of a ground state harmonic oscillator, with the spatial distribution described by a Gaussian wavefunction corresponding to the Heisenberg uncertainty limit, the time evolution of the spatial distribution is given by

$$\zeta^2(t) = \zeta^2(0) + v_\zeta^2 t_e^2 \,, \tag{2}$$

where $\zeta(t) = (\Delta x, \Delta y, \Delta z)$ at the time t_e and v_ζ the initial velocity [12]. A better agreement with the experimental results is obtained using this formula for the time dependence.

4. Geometric Forces

In order to describe analytically the motion of atoms in a TOP-trap, two approximations are usually made. The first approximation, which we shall call the 'time-averaging' approximation, is based on the assumption that the frequency of oscillation of the atoms in the (to first order) harmonic potential created by the superposition of the quadrupole and bias fields is much smaller than the frequency Ω_0 of the bias field. The atoms will, therefore, experience many periods of the bias field before appreciably changing their position. This allows one to average over a whole period of the bias field at a fixed position of the atoms, leading to a time-averaged potential $\langle U \rangle$ in the adiabatic approximation [8]

$$\langle U \rangle = \frac{M_{\mathrm{Rb}}}{2} \left(\omega_{\mathrm{x}}^2 x^2 + \omega_{\mathrm{y}}^2 y^2 + \omega_{\mathrm{z}}^2 z^2\right) , \tag{3}$$

where M_{Rb} is the mass of the rubidium atom, and

$$\omega_{\mathrm{x}} = \sqrt{\frac{2\mu b'}{M_{\mathrm{Rb}} B_0}} \,, \tag{4}$$

$$\omega_{\mathrm{y}} = \sqrt{\frac{\mu b'}{2 M_{\mathrm{Rb}} B_0}} \,, \tag{5}$$

$$\omega_{\mathrm{z}} = \sqrt{\frac{\mu b'}{M_{\mathrm{Rb}} B_0}} \,, \tag{6}$$

with μ_{m} the atomic magnetic moment. This results in a triaxial time-orbiting-potential with a harmonic frequency ratio $\omega_{\mathrm{x}} : \omega_{\mathrm{z}} : \omega_{\mathrm{y}}$ of $2 : 1 : \sqrt{2}$.

When the time-averaging approximation is lifted, the first correction to the dynamics of the trapped atoms is the atomic micromotion. This motion originates because the atoms experience a time-dependent potential at frequency Ω_0 and its harmonics. An expansion of the potential energy for atoms close to the trap centre gives additional linear terms ΔU modulated at the frequency Ω_0 on top of the effective harmonic potential of Eq. (3)

$$\Delta U = \mu b' (2x \cos \Omega_0 t - y \sin \Omega_0 t) \,. \tag{7}$$

Balancing the restoring force from the ΔU potential with the centrifugal force of the resulting orbital motion yields a periodic micromotion with amplitudes

$$x(t) = -\frac{2\mu b'}{M_{\mathrm{Rb}}\Omega_0^2}\cos\Omega_0 t \quad \text{and} \quad y(t) = \frac{\mu b'}{m\Omega_0^2}\sin\Omega_0 t\,. \tag{8}$$

This solution corresponds to an elliptical motion of the condensate with a very small spatial amplitude, around 100 nm. However from the point of view of the geometric forces to be discussed in the following produces also an atomic motion within a spatially inhomogeneous magnetic field.

The second approximation, called the adiabatic approximation, concerns the orientation of the atomic magnetic dipole with respect to the local direction of the magnetic field. As the atom moves and the magnitude and direction of the local magnetic field changes, the magnetic moment of the atom will precess around the direction of the field. In the adiabatic approximation, it is assumed that the Larmor precession frequency of the magnetic dipole is much larger than the rotational frequency of the bias field, and hence the magnetic moment μ_{m} can be thought of as being aligned with the local field at all times. In this case, the interaction energy between the atomic magnetic moment and the total magnetic field (quadrupole plus bias field) is simply $\mu_{\mathrm{m}}B$, with $B(\vec{r})$ the amplitude of the total local magnetic field and its $\hat{\boldsymbol{b}}(\vec{r})$ ist versor. This approximation cannot describe the motion of a magnetic momentum within an inhomogeneous magnetic field. In fact because the atomic spin $\vec{S}$ obeys the following equation:

$$\frac{\delta\vec{S}}{\delta t} = \vec{\mu}_{\mathrm{m}} \times \vec{B}\,, \tag{9}$$

if the magnetic moment $\vec{\mu}_{\mathrm{m}}$ is parallel to the local magnetic field $\vec{B}(\vec{r})$, the atomic spin cannot follow adiabatically the external magnetic field. The non-adiabatic orientation of the magnetic moment in respect of the applied magnetic field was described in Ref. [13]. Applying that description to our magnetic field configuration, the i-th component of the force acting on the atoms contained in the magnetic trap becomes

$$F_{\mathrm{i}} = \mu_{\mathrm{m}}\left[\frac{\partial B(\vec{r})}{\partial x_{\mathrm{i}}} + \frac{1}{\gamma}\,\frac{\partial\hat{\boldsymbol{b}}(\vec{r})}{\partial x_{\mathrm{i}}}\cdot\hat{\dot{\boldsymbol{b}}}\times\hat{\boldsymbol{b}}\right], \tag{10}$$

where the gyromagnetic ratio γ is given by

$$\gamma = \frac{\mu}{S}\,. \tag{11}$$

The last term in Eq. (10), originating from the non-adiabatic evolution of the magnetic moment within the local magnetic field, is a geometric force, i.e. a force whose amplitude does not depend on the amplitude $B(\vec{r})$ of the magnetic field, but only on the local orientation of the field [14]. The geometric force produces an additional contribution to the action on the TOP trap on the condensate. Because of the $\hat{\boldsymbol{b}}$ dependence, the geometric force changes its sign with the sign of Ω_0, i.e. inverting the rotation sense of the rotating field in Eq. (1). By measuring accurately the vertical equilibrium position of the condensate under the action of the force of Eq. (10) and of the gravity for both rotation directions of the radiofrequency magnetic field, in Ref. [2] the geometric forces acting on the condensate were accurately determined.

5. BEC in 1D Optical Lattice

We have investigated the motion of a condensate within a periodic potential with spatial dependent depth $V_0 \sin^2 kx$ and wavenumber $k = \pi/d$, with d the optical lattice constant. In a preliminary optical lattice experiment aimed to determine the depth of the periodic optical potential, we flashed on the counterpropagating lattice with a detuning $\delta = 4E_{\mathrm{rec}}/h = 2\hbar k_1^2/M$ between the two lasers making up the optical lattice for a variable time. The chosen detuning corresponded to the first Bragg resonance, causing the condensate to undergo Rabi oscillations between the momentum states $|p = 0\rangle$ and $|p = 2\hbar k_1 *\rangle$. From the measured Rabi frequency we could determine the lattice depth V_0.

In order to accelerate the condensate, we adiabatically loaded it into the lattice by switching one of the lattice beams on suddenly and ramping the intensity of the other beam from 0 to its final value V_0 in a time between 200 µs and few ms. Thereafter, the linear increase of the detuning δ provided a constant acceleration $a = \frac{\lambda}{2}\frac{d\delta}{dt}$ of the optical lattice. After a few milliseconds of acceleration, the lattice beams were switched off and the condensate was imaged after another 10−15 ms of free fall flight. Since for our magnetic trap parameters the initial momentum spread of the condensate was much less than a recoil momentum of the optical lattice (the Bloch momentum $p_{\mathrm{B}} = h/d$) and since the adiabatic switching transfers the spread into lattice quasimomentum, the different momentum classes $|p = \pm n p_{\mathrm{B}} = \pm 2nh/d\rangle$ (where $n = 0, 1, 2, \ldots$) occupied by the condensate wavefunction could be resolved directly after the time-of-flight. As described in [15], the acceleration process within a periodic potential can also be viewed as a succession of adiabatic rapid passages between momentum states $|p = 2nh/d\rangle$ and $|p = 2(n+1)\, h/d\rangle$. Up to $6\,p_{\mathrm{B}}$ momentum could be transferred to the condensate preserving the phase-space density of the condensate. The average velocity of the condensate was derived from the occupations of the different momentum states. Figure 2 shows the results of the acceleration of a condensate. In the rest-frame of the lattice the evolution of the atomic velocity correspond to Bloch oscillations of the condensate velocity with a Bloch-period $\tau_{\mathrm{B}} = h(M_{\mathrm{Rb}} a d)^{-1}$.

When we increased the acceleration of the lattice, not all of the condensate was coherently accelerated up to the final velocity of the lattice. This condensate velocity loss can be interpreted in terms of Landau-Zener tunneling of the condensate out of the lowest band when the edge of the Brillouin zone is reached. The Landau-Zener probability for tunneling into the first excited band (and, therefore, effectively to the continuum, as the gaps between higher bands are negligible for the shallow potentials used here) is [15]

$$P_{\mathrm{LZ}} = e^{-\frac{a_c}{a}} \tag{12}$$

with the critical acceleration a_c given by

$$a_c = \frac{\pi V_0^2}{16\hbar^2 k}\,. \tag{13}$$

The average velocity of the condensate after acceleration to the Bloch velocity $v_{\mathrm{B}} = p_{\mathrm{B}}/M_{\mathrm{Rb}}$ for a final velocity v_{B} of the lattice depends on the Landau-Zener probability P_{LZ}

$$\frac{v_m}{v_{\mathrm{B}}} = 1 - P_{\mathrm{LZ}}\,. \tag{14}$$

We have verified that this equation describes properly the dependence of the measured v_m mean velocity on the a and V_0 parameters.

Fig. 2. Bloch oscillations of a Bose-Einstein condensate in an optical lattice. (a) Acceleration in the counter-propagating lattice with $d = 0.39\,\mu\text{m}$, $V_0 \approx 0.29 E_B$ and $a = 9.81\ \text{ms}^{-2}$. Dashed line: theory. (b) Bloch oscillations in the rest frame of the lattice, along with the theoretical prediction (solid line) derived from the shape of the lowest Bloch band. (c) Acceleration in a lattice with $d = 1.56\,\mu\text{m}$ and $V_0 \approx 1.38 E_B$. In this case, the Bloch oscillations are much less pronounced. Dashed and solid lines: theory for $V_0 = 1.38 E_B$ and $V_{\text{eff}} \approx 0.88 E_B$. Reproduced from Ref. [5] with the permission of the American Physical Society.

The properties of a Bose-Einstein condensate located in a periodic optical lattice with depth U_0 are described through the Gross-Pitaevskii equation valid for the single-particle wavefunction. The nonlinear interaction of the condensate may be described through a dimensionless parameter [6, 16, 17]

$$C = \frac{g}{E_B} = \frac{n_p a d^2}{\pi} \tag{15}$$

corresponding to the ratio of the nonlinear interaction term $g = 4\pi n_p \hbar^2 a/M$ and the lattice Bloch energy $E_B = \hbar^2 (2\pi)^2 / M d^2$. The parameter C contains the peak condensate density n_p in the Thomas-Fermi limit, the s-wave scattering length a, the atomic mass M, the lattice constant $d = \pi / \sin(\theta/2)\, k$. From this it follows that a small angle θ in the angle-tuned configuration of the optical lattice should result in a large interaction term C.

The role of the nonlinear interaction term of the Gross-Pitaevskii equation may be described through an effective potential in a non-interacting gas model [6, 16]. In the perturbative regime of Ref. [6] the effective potential is

$$V_{\text{eff}} = \frac{V_0}{1 + 4C}\,, \tag{16}$$

with the same spatial structure as the optical lattice. We therefore expect that for large values of C, i.e. large mean-field effects, the effective optical lattice potential acting on the condensate should be significantly reduced.

In order to demonstrate unequivocally the reduction of the effective lattice potential by the interaction term, however, it would be necessary to vary n_p appreciably holding all

other parameters constant. As the interaction term is expected to distort the band structure of the condensate in the lattice [18], it should affect all measurable quantities (Rabi frequency, amplitude of Bloch oscillations, and tunneling probability) in the same way. In order to measure more accurately the variation of the effective potential V_{eff} with the interaction parameter C, we studied Landau-Zener tunneling out of the lowest Bloch band for small lattice depths in both geometries. To this end, as discussed above, the acceleration of the lattice was increased in such a way that an appreciable fraction of the atoms tunneled across the band gap into the first excited band. We fitted the measured condensate velocity v_m through Eq. (14) imposing for the Landau-Zener probability the following dependence on the effective potential:

$$P_{\mathrm{LZ}} = e^{-\frac{\pi V_{\mathrm{eff}}^2}{8\hbar p_{\mathrm{B}} a}}. \tag{17}$$

Thereafter, we studied the variation of the final mean velocity v_m as a function of the condensate density for the two lattice geometries. The condensate density was varied by changing the mean frequency of the magnetic trap (from $\approx 25\,\mathrm{Hz}$ to $\approx 100\,\mathrm{Hz}$). Figure 3 shows the ratio V_{eff}/V_0 as a function of the C parateter for the counter-propagating geometry and the angle-geometry. As expected, the reduction of the effective potential is much larger in the angle geometry. Although we could realize densities up to $4 \times 10^{14}\,\mathrm{cm}^{-3}$ by using larger trap frequencies, data points for $n_p > 1 \times 5 \times 10^{13}\,\mathrm{cm}^{-3}$ were not included in out analysis as the resulting diffraction patterns were not easily interpretable within the simple model of Eq. (16), implying for the condensate to follow the spatial periodicity of the optical lattice.

The theoretical predictions of Eq. (16) are also shown in the figure, with the potential V_0 determined from the Rabi oscillations described previously. Within the experimental uncertainties, agreement with the simple theory of Eq. (16) was reasonable. An effective potential may be derived also within the frame of a tight-binding approximation, i.e. describing the condensate within each potential minimum potential through the solution of the Gross-Pitaevskii equation while neglecting the overlap with neighboring potential minima [19]. A weak tunneling of the condensate towards the neighboring wells should be supposed pre-

Fig. 3. Dependence of the effective potential V_{eff} on C parameter for the two lattice geometries. The experimental results for the counter-propagating (closed triangles) and angle geometries (open circles) are plotted together with the theoretical predictions, dashed and solid and lines, for Eq. (16) and the tight-binding approach respectively. Parameters in these experiments were $a = 23.4\,\mathrm{m\,s}^{-2}$ and $V_0 = 0.29E_{\mathrm{B}}$ for the counter-propagating lattice and $a = 3.23\,\mathrm{m\,s}^{-2}$ and $V_0 = 0.88E_{\mathrm{B}}$ for the angle geometry.

sent in order to preserve the phase relation of the condensate wavefunction between the neighboring optical lattice wells. The tight-binding predictions presented in Fig. 3 appear to provide a better agreement with experimental results than those of Eq. (16). However the validity of the tight-binding approach for the parameters of the experiment requires an additional investigation.

6. Condensate Ionization

Suppose that one introduces into the condensate a charged particle of atomic size, as an ion or an electron, that can readily be manipulated by externally applied forces. Such particle could then serve as a microscopic probe of the medium in which it is immersed, for instance a superfluid liquid or a condensate. In fact this approach was used in a large series of experimental investigations with ^{4}He superfluid liquid. As a manifestation of superfluidity, objects travelling below a critical velocity through the superfluid should propagate without dissipation, as tested in liquid ^{4}He studies dragging negative ions through an external electric field. Similarly, evidence for the superfluidity in gaseous Bose-Einstein condensates could be obtained by measuring the motion of the ions produced by laser ionization in presence of an external electric field. Charge hopping processes, with electrons jumping from a neighboring atom onto the positive ion, may also contribute to the conductivity in an ultracold gas. Due to the coherent nature of the initial atomic ensemble and narrow spectral width of the light source producing the photoionization process, the occupation numbers of the final states n_e for electrons and n_i for positive ions, with $n_e = n_i$, may be close to unity. Because the ionization products (electrons and ions) obey Fermi-Dirac statistics, the rate of the laser driven photoionization depends on the final state occupation through $1 - n_e$. Thus, in presence of n_e values close to unity, the ionization process is blocked by the previous electron production.

In order to produce rubidium ions, we irradiated the condensate atoms confined by the magnetic trap with pulsed 534 nm or 296 nm ionizing lasers, studying the induced losses from the magnetic trap after a single laser pulse. The 296 nm radiation produces a direct ionization process, while absorption of two 534 nm photons in a non resonance process is required for the rubidium photoionization In order to determine the photoionization rate as a loss process for the atoms confined within the magnetic trap, we monitored the atoms remaining in the magnetic trap.

The experiment aims to test the predictions on the Pauli blockade of Ref. [7]. However, the major result obtained so far [20] is the observation that the ionization process does not destroy the condensate itself. We observed that even after several laser ionization processes a condensate remains formed, of course with a smaller number of atoms, other properties being preserved. Furthermore we measured losses much larger than those predicted on the basis of the ground state cross-section theoretical value for the individual atoms. Moreover we observed no clear evidence of a photoionization threshold behavior when the pulsed laser wavelength was scanned around the photoionization threshold value. We measured also the density profile modifications of the condensed cloud produced by the photoionization laser. We noticed that the photoionization pulse ejected a secondary condensate from the original one, with the secondary condensate separating in space from the original one. By measuring the position of the secondary condensate as a function of the evolution time within the magnetic trap, we realized that the secondary cloud is constituted by atoms in the $F = 2$, $m = 1$ Zeeman state. We also discovered that a fit of the measured oscillating condensate position could be obtained only by supposing that the secondary condensate was created with an initial velocity pointing downwards the vertical axis. The process taking place within the condensate could be described as an $F = 2$, $m = 1$ atom laser emission from the original condensate following the excitation by the photoionizing laser. We don't

know which process has precisely produced the coherent transfer of condensate atoms from the initial Zeeman level to the final one. However it should noted that the losses from the magnetic trap are not necessarily associated to the ionization of the cold atoms. An escape from the magnetic trap is also produced by the processes where the atomic internal state is not modified, but the atom acquires a kinetic energy larger than the magnetic trap depth.

7. Conclusions

In summary, we have use a Rb condensate to investigate several properties of atoms interacting with electromagnetic fields in different configurations. Measurements on condensate samples have tested the expansion of a harmonic potential wavepacket in conditions close to the Heisenberg uncertainty limit. Moreover the high spatial resolution provided by the condensate has allowed us to perform a measurement of the geometric forces acting on a magnetic moment moving within a spatially inhomogeneous magnetic field. The coherent acceleration of Bose-Einstein condensates adiabatically loaded into a 1-D optical lattice as well as Bloch oscillations and Landau-Zener tunneling out of the lowest Bloch band. We have verified that Bloch oscillations and Landau-Zener tunneling are produced in a condensate without a noticeable modification of the phase-space and condensate depletion. Our investigation did not test the evolution of the condensate phase, but from the Bragg scattering experiments where a condensate interacted with an optical lattice for a 0.1 μs, we infer that the longer interaction times of our experiment should not destroy the condensate phase.

Acknowledgements

The contributions of M. Anderlini, D. Ciampini, M. Cristiani F. Fuso, J. H. Müller, O. Morsch, R. Mannella, G. Smirne and J. Thomsen to the experimental parts of the presetnt research is gratefully acknowledged. The author thanks M. Kalinski, G. Metikas and W. P. Schleich for the analysis of the condensate effective potential. This research was supported by the Sezione A of the INFM-Italy through a PAIS Project, by the MURST-Italy through a PRIN Progetto, and by the EU through the Cold Quantum-Gases Network, contract HPRN-CT-2000-00125. The visit to the University of Ulm was supported by the Alexander von Humboldt Foundation scientific program.

References

[1] See reviews by W. Ketterle, R. S. Durfee, D. M. Stamper-Kurn, and by E. Cornell et al. in *Bose-Einstein condensation in atomic gases*, edited by M. Inguscio, S. Stringari and C. Wieman (IOS Press, Amsterdam) (1999).
[2] J. H. Müller, O. Morsch, D. Ciampini, M. Anderlini, R. Mannella, and E. Arimondo, Phys. Rev. Lett. **85** (2000) 4454.
[3] For reviews see P. S. Jessen and I. H. Deutsch, Adv. At. Mol. Opt. Phys. **37** (1996) 95;
G. Grynberg and C. Triché, in *Coherent and Collective Interactions of Particles and Radiation Beams* edited by A. Aspect, W. Barletta, and R. Bonifacio (IOS Press, Amsterdam) (1996);
D. R. Meacher, Cont. Phys. **39** (1998) 329.
[4] M. Raizen, C. Salomon, and Q. Niu, Phys. Today **50** (7) (1997) 30.
[5] O. Morsch, J. H. Müller, M. Cristiani, D. Ciampini, and E. Arimondo, Phys. Rev. Lett. **87** (2001) 140402.
[6] Dae-Il Choi and Qian Niu, Phys. Rev. Lett. **82** (1999) 2022.
[7] I. E. Mazets, Quantum Semiclass. Opt. **10** (1998) 675.
[8] J. H. Müller, D. Ciampini, O. Morsch, G. Smirne, M. Fazzi, P. Verkerk, F. Fuso, and E. Arimondo, J. Phys. B: Atom. Mol. Opt. Phys. **33** (2000) 4095.
[9] E. W. Hagley, L. Deng, M. Kozuma, J. Wen, K. Helmerson, S. L. Rolston, and W. D. Phillips, Science **283** (1999) 1706.

[10] G. SMIRNE, Tesi di Laurea, Università di Pisa, March 2000, unpublished.
[11] Y. CASTIN and R. DUM, Phys. Rev. Lett. **77** (1996) 5315.
[12] The temporal evolution of the atomic wavefunction may be derived from the propagator of the free particle, as in W. P. SCHLEICH *Quantum Optics in Phase Space* (Wiley-VCH, Berlin 2001).
[13] Y. AHARONOV and A. STERN, Phys. Rev. Lett. **69** (1992) 3593.
[14] M. BERRY, Proc. Roy. Soc. London Ser. A **452** (1996) 1207.
[15] M. BEN-DAHAN, E. PEIK, J. REICHEL, Y. CASTIN, and C. SALOMON, Phys. Rev. Lett. **76** (1996) 4508;
E. PEIK, M. BEN DAHAN, I. BOUCHOULE, Y. CASTIN, and C. SALOMON, Phys. Rev. A **55** (1997) 2989.
[16] M. L. CHIOFALO, M. POLINI, and M. TOSI, Eur. Phys. J. D **11** (2000) 371.
[17] A. TROMBETTONI and A. SMERZI, Phys. Rev. Lett. **86** (2001) 2353.
[18] K. BERG-SØRENSEN and K. MØLMER, Phys. Rev. A **58** (1998) 1480.
[19] M. KALINSKI, G. METAKIS, and W. P. SCHLEICH, private communication.
[20] E. ARIMONDO, in Proc. Int. E. Fermi Summer School on *Quantum Information and Computation*, in press.

Fortschr. Phys. **50** (2002) 5–7, 594–598

Interference of Spontaneously Emitted Photons

ALMUT BEIGE, CHRISTIAN SCHÖN, and JIANNIS PACHOS

Max-Planck-Institut für Quantenoptik, D-85748 Garching, Germany
a.beige@mpq.mpg.de

Abstract

We discuss an experimental setup where two laser-driven atoms spontaneously emit photons and each and every photon causes a "click" at a point on a screen. By deriving the probability density for an emission into a certain direction from basic quantum mechanical principles we predict a spatial interference pattern. Similarities and differences with the classical double-slit experiment are discussed.

1. Introduction

In 1930 Werner Heisenberg wrote in his book *The physical principles of the quantum theory*, "It is very difficult for us to conceive the fact that the theory of photons does not conflict with the requirements of the Maxwell equations. There have been attempts to avoid the contradiction by finding solutions of the latter which represent 'needle' radiation (unidirectional beams), but the results could not be satisfactorily interpreted until the principles of the quantum theory had been elucidated. These show us that whenever an experiment is capable of furnishing information regarding the direction of emission of a photon, its results are precisely those which would be predicted from a solution of the Maxwell equations of the needle type (...)" [1].

A simple quantum mechanical experiment that combines wave and particle features of emitted light is sketched in Fig. 1. It shows two two-level atoms at a fixed distance r that are both coupled to the same free radiation field. In addition, the atoms are excited by a resonant laser field, so that they can continuously emit photons. Each emitted photon causes a "click" at a certain point on a screen far away from the atoms. These "clicks" add up to an interference pattern with a spatial intensity distribution very similar to a classical double-slit experiment.

This setup has been proposed by Scully and Drühl [2] in 1982 as a quantum eraser concerning observation and delayed choice phenomena in quantum mechanics. Their predictions were verified experimentally by Eichmann et al. in 1993 [3]. Since then the interpretation of this experiment has attracted continuous interest (see for instance Refs. [4–7]

Fig. 1. Experimental setup. Two two-level atoms are placed at a fixed distance r from each other. Both are coupled to the same free radiation field and are continuously driven by a resonant laser. This leads to spontaneous photon emissions. Each photon causes a "click" at a point on a screen.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

and references therein). This is mainly due to its simplicity together with the fundamental quantum mechanical issues it deals with.

Recently, Schön, and Beige [6] presented a quantum mechanical description of the experiment based on the assumption of environment-induced measurements. Here we first summarise the main ideas of Ref. [6] and then focus on the nature of the spatial interference of the spontaneously emitted photons. Main features of the interference pattern can as well be predicted by a corresponding classical model in which the atoms are replaced by dipole sources of coherent light. Finally, similarities and differences of the two descriptions are discussed.

2. The Spontaneous Emission of Photons

The setup shown in Fig. 1 consists of two components – the quantum mechanical system including the atoms together with the free radiation field and its environment represented by the screen. In the following we describe the time evolution of the quantum mechanical system by the Schrödinger equation. Let us denote the energy difference between the ground state $|1\rangle_i$ and the excited state $|2\rangle_i$ of atom i at the position $\boldsymbol{r}_i$ by $\hbar\omega_0$ whilst $a^\dagger_{\boldsymbol{k}\lambda}$ is the creation operator of a single photon with wave vector $\boldsymbol{k}$ and polarisation λ. The Hamiltonian that describes the interaction between the atoms and the quantised free radiation field within the rotating wave approximation equals

$$H_{\text{int}} = \hbar \sum_{i=1,2} \sum_{\boldsymbol{k}\lambda} e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i}\, g_{\boldsymbol{k}\lambda} a^\dagger_{\boldsymbol{k}\lambda} S_i^- + \text{h.c.}\,, \tag{1}$$

where $g_{\boldsymbol{k}\lambda}$ describes the coupling strength of the field mode $(\boldsymbol{k}, \lambda)$ and S_i^- is the atomic lowering operator $|1\rangle_i \langle 2|$.

The occurrence of spontaneous "clicks" at points on the screen suggest that the effect of the screen onto the quantum mechanical system can be described with the projection postulate for ideal measurements. The screen measures whether a photon has been created in the free radiation field or not [8]. If a photon is detected, then also the direction $\hat{\boldsymbol{k}}_0$ [9] of its wave vector is determined [6]. The projector that describes this measurement outcome is given by

$$\mathsf{P}_{\hat{\boldsymbol{k}}_0} = \sum_{k,\lambda} |1_{k\hat{\boldsymbol{k}}_0\lambda}\rangle \langle 1_{k\hat{\boldsymbol{k}}_0\lambda}|\,. \tag{2}$$

If no "click" is observed, then the system is projected onto a state with the free radiation field in the vacuum state $|0_{\text{ph}}\rangle$. Here the atoms are continuously driven by a laser field and we assume that the screen performs rapidly repeated ideal measurements. The time Δt between two successive measurements should not be too short, $\Delta t \gg 1/\omega_0$, to allow for a substantial time evolution, but also not too big so that the excitation of two-photon states in Δt is negligible [8]. As these measurements are caused by the presence of the screen we call them environment-induced measurements.

For simplicity, let us assume that the state of the atoms at a time t is known and equals $|\psi\rangle$ while the free radiation field is in the ground state $|0_{\text{ph}}\rangle$. Letting the system evolve for a short time Δt and applying the projector (2) we find

$$|\psi\rangle\,|0_{\text{ph}}\rangle \xrightarrow{\Delta t} \sum_{i=1}^{2} \left[\frac{3A}{8\pi}\,(1 - |\hat{\boldsymbol{D}}_{21}\cdot\hat{\boldsymbol{k}}_0|^2)\right]^{1/2} e^{-\mathrm{i}k_0\hat{\boldsymbol{k}}_0\cdot\boldsymbol{r}_i}\, S_i^-\,|\psi\rangle \left|\begin{array}{l}\text{normalised}\\ \text{field state } (\hat{\boldsymbol{k}}_0)\end{array}\right\rangle, \tag{3}$$

where the right hand side is the unnormalised state of the system in case of a "click" in the $\hat{\boldsymbol{k}}_0$ direction away from the atoms at time $t + \Delta t$. Here we assumed that the dipole moment ${}_i\langle 2|\,\boldsymbol{D}\,|1\rangle_i = \boldsymbol{D}_{21}$ is for both atoms the same and A denotes the spontaneous decay rate of a single atom in free space. Equation (3) has been derived with the help of first order perturbation theory. For details see Section II in Ref. [6].

The norm squared of the state (3) gives the probability density $I_{\hat{\boldsymbol{k}}_0}(\psi)$ for the corresponding measurement outcome,

$$I_{\hat{\boldsymbol{k}}_0}(\psi) = \frac{3A}{8\pi}\,\left(1 - |\hat{\boldsymbol{D}}_{21}\cdot\hat{\boldsymbol{k}}_0|^2\right)\,\left\|\sum_i e^{-\mathrm{i}k_0\hat{\boldsymbol{k}}_0\cdot\boldsymbol{r}_i}\,S_i^-\,|\psi\rangle\right\|^2 . \tag{4}$$

The probability density for any "click" to occur can be obtained by integrating over all orientations of $\hat{\boldsymbol{k}}_0$, leading to the product of the spontaneous decay rate A and the population of the excited atomic states, namely $I(\psi) = A\sum_i \|S_i^-\,|\psi\rangle\|^2$. The "clicks" on the screen are caused by the *spontaneously emitted* photons. Immediately after the measurement, the excitation in the free radiation field vanishes and its state becomes again $|0_{\mathrm{ph}}\rangle$.

3. The Interference Pattern on the Screen

Let us now specify the spatial probability density for a photon emission. If the atoms are continuously driven by a resonant laser field, their state $|\psi\rangle$ shortly before an emission is not known. To apply the result of Eq. (4) to this situation we have to describe the atoms by their steady state density matrix ρ_{ss}, so $I_{\hat{\boldsymbol{k}}_0}$ becomes

$$\begin{aligned} I_{\hat{\boldsymbol{k}}_0}(\rho_{\mathrm{ss}}) = \frac{3A}{8\pi}\,&\left(1 - |\hat{\boldsymbol{D}}_{21}\cdot\hat{\boldsymbol{k}}_0|^2\right)\,\left[2\,\langle 22|\,\rho_{\mathrm{ss}}\,|22\rangle + \langle 12|\,\rho_{\mathrm{ss}}\,|12\rangle + \langle 21|\,\rho_{\mathrm{ss}}\,|21\rangle\right. \\ &\left. + 2\,\mathrm{Re}\left(\langle 12|\,\rho_{\mathrm{ss}}\,|21\rangle\,e^{-\mathrm{i}k_0\hat{\boldsymbol{k}}_0\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)}\right)\right] . \end{aligned} \tag{5}$$

In the following we denote the Rabi frequency of the laser field with respect to atom i by Ω_i. Proceeding as in Ref. [6] to calculate ρ_{ss} we find that

$$\begin{aligned} I_{\hat{\boldsymbol{k}}_0}(\rho_{\mathrm{ss}}) = \frac{3A}{8\pi}\,&\frac{1 - |\hat{\boldsymbol{D}}_{21}\cdot\hat{\boldsymbol{k}}_0|^2}{\left(A^2 + 2\,|\Omega_1|^2\right)\left(A^2 + 2\,|\Omega_2|^2\right)}\,\left[4\,|\Omega_1|^2\,|\Omega_2|^2 + A^2\,|\Omega_1|^2 + A^2\,|\Omega_2|^2\right. \\ &\left. + 2A^2\,\mathrm{Re}\left(\Omega_1^*\Omega_2\,e^{-\mathrm{i}k_0\hat{\boldsymbol{k}}_0\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)}\right)\right] . \end{aligned} \tag{6}$$

Figure 2 shows the spatial intensity distribution for the case where both laser fields are in phase and have the same intensity, i.e. $\Omega_1 = \Omega_2$. One easily recognises an interference pattern which results from the last term in Equation (6).

To understand the origin of the interference let us assume again that the state of the atoms shortly before an emission is known and equals $|\psi\rangle$ while the free radiation field is in the vacuum state. During the time evolution Δt with respect to the Hamiltonian (1) each atom transfers excitation into all modes $(\boldsymbol{k}, \lambda)$ of the free radiation field. In case of a photon detection the state of the field is projected onto a photon state with the wave vector towards the direction $\hat{\boldsymbol{k}}_0$ of the "click". To calculate the probability density for this event, one had to determine the norm of the reset state given in Eq. (3). This state is the sum of the contributions from both atoms which differ only by the phase factor $e^{-\mathrm{i}k_0\hat{\boldsymbol{k}}_0\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)}$. In addition, each contribution contains a different atomic state, namely $S_1^-\,|\psi\rangle$ or $S_2^-\,|\psi\rangle$. Therefore the visibility of the interference pattern depends on the overlap of the atomic states $S_1^-\,|\psi\rangle$ and $S_2^-\,|\psi\rangle$. If these two states are orthogonal, i.e. when the atoms are prepared in $|22\rangle$, then the probability den-

Fig. 2. Density plot of the emission rate $I_{\hat{\boldsymbol{k}}_0}(\rho_{\rm ss})$ for two continuously driven two-level atoms with $r = 20\,\lambda_0$ and $\Omega_1 = \Omega_2 = 0.3\,A$. White areas correspond to spatial angles ϑ and φ with maximal intensity.

sity to find a "click" at a certain point on the screen is just the sum of the probabilities of two one-atom cases. Maximum interference takes place if $S_1^- \,|\psi\rangle = S_2^- \,|\psi\rangle$. This is the case when $|\psi\rangle$ is a non-trivial superposition of the states $|12\rangle$ and $|21\rangle$.

In general both atoms contribute to the spontaneous emission of a photon. Nevertheless, one could ignore this and ask from which atom the photon originated. To answer this one has to perform a measurement on the atomic state and determine whether the atoms are either in $S_1^- \,|\psi\rangle$ or $S_2^- \,|\psi\rangle$. These two states are the reset states of the two one-atom cases where one atom emits and the state of the other one remains unchanged. Only if the reset states are orthogonal, one can find out with certainty which atom emitted the photon and the *which way* information is available in the experiment. In this case the interference vanishes. The more overlap the states have, the stronger becomes the visibility of the interference pattern [2, 3, 5]. For a more detailed discussion of interference criteria see also Refs. [4, 6, 7] and references therein.

4. Comparison with the Classical Double-Slit Experiment

Let us now consider a classical double-slit experiment with two dipole sources at positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Both dipoles have the same direction $\hat{\boldsymbol{D}}$ of the dipole moment and simultaneously emit electromagnetic waves with frequency ω_0. The resulting electric field at a certain point $\boldsymbol{R}$ on a far away screen ($|\boldsymbol{R} - \boldsymbol{r}_i| \gg |\boldsymbol{r}_1 - \boldsymbol{r}_2|$) is then

$$\boldsymbol{E}(\boldsymbol{R}, t) = \sum_{i=1}^{2} \frac{E_0^{(i)}}{|\boldsymbol{R} - \boldsymbol{r}_i|} \, [\hat{\boldsymbol{D}} - (\hat{\boldsymbol{k}}_0 \cdot \hat{\boldsymbol{D}})\, \hat{\boldsymbol{k}}_0] \, e^{-\mathrm{i}\boldsymbol{k}_0 \cdot (\boldsymbol{R} - \boldsymbol{r}_i)} \, e^{-\mathrm{i}\omega_0 t} \,, \tag{7}$$

where $E_0^{(1)}$ and $E_0^{(2)}$ characterise the strength of the dipoles. The phase of the electric field produced by each source i is sensitive to its position $\boldsymbol{r}_i$ while the wave vector $\boldsymbol{k}_0 = k_0 \hat{\boldsymbol{k}}_0$ has in both cases, to a very good approximation, the same direction as $\boldsymbol{R} - \boldsymbol{r}_1 \approx \boldsymbol{R} - \boldsymbol{r}_2$ and the amplitude k_0. The intensity of the produced light is given by

$$\begin{aligned} I_{\hat{\boldsymbol{k}}_0}(E_0^{(1)}, E_0^{(2)}) = \frac{\epsilon_0 c}{2} \, (1 - |\hat{\boldsymbol{D}} \cdot \hat{\boldsymbol{k}}_0|^2) \, [|E_0^{(1)}|^2 + |E_0^{(2)}|^2 \\ + 2\,\mathrm{Re}\,(E_0^{(1)*} E_0^{(2)} \, e^{-\mathrm{i}\boldsymbol{k}_0 \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2)})] \,. \end{aligned} \tag{8}$$

The last term describes the interference pattern on the far away screen.

If we compare this spatial dependence of the interference with the spatial dependence in the quantum mechanical double-slit experiment, we see that they can in both cases be exactly the same. Nevertheless the visibility in the quantum model is in general lower than in the classical model. The reason is that population of the atomic state $|22\rangle$ does not contribute to the amplitude of the interference term. In the quantum case, the more *which way* information is available in the experiment, the less is the visibility of the interference pattern [4]. This reduction of the visibility is a purely quantum mechanical effect.

The analogies of the quantum system with an equivalent classical one is not a surprising feature [10]. Each atom populates the free radiation field with an excitation of a certain effective frequency, ω_0, as dictated from the energy difference between the atomic levels 1 and 2. In addition, when the measurement of the radiation field takes place and a photon is detected then the state of the field is projected towards a certain direction $\hat{\boldsymbol{k}}_0$, as can be seen from Equation (3). These two deterministic characteristics are not imposed in the initial theory as the atoms are allowed to emit with any frequency and the radiation field around them has no preference in the direction of propagation of its photon states. Nevertheless, the average along all possible photon frequencies as well as the act of detection, which projects the free radiation field onto the photon state with direction $\hat{\boldsymbol{k}}_0$, gives effectively the equivalence between the spontaneously emitting atoms and the classical sources.

5. Conclusions

Here we presented a quantum mechanical description of a two-atom double-slit experiment. We showed that spontaneous emission of photons can be derived from basic quantum mechanical principles by assuming environment-induced measurements on the free radiation field. By calculating the probability density for an emission into a certain direction $\hat{\boldsymbol{k}}_0$ we found spatial interference. The interference fringes appearing in the quantum case result from the spontaneously emitted photons and are not just produced by the interference of a scattered laser field. This is particularly evident in Eq. (4) where the atoms are initially prepared in an atomic state $|\psi\rangle$ without the need that this state is prepared by continuous laser radiation.

In general, the spatial dependence of the interference pattern and its visibility could be described also by an equivalent classical double-slit experiment. It is worth noticing that the superposition of the two amplitudes in the quantum and in the classical case produces standing waves throughout space. Their pattern is observed on the screen as seen in Fig. 2.

References

[1] W. Heisenberg, *The physical principles of the quantum theory* (The University of Chicago Press, Chicago 1930), Chapter V, §5.
[2] M. O. Scully and K. Drühl, Phys. Rev. A **25** (1982) 2208.
[3] U. Eichmann, J. C. Berquist, J. J. Bollinger, J. M. Gilligan, W. M. Itano, and D. J. Wineland, Phys. Rev. Lett. **70** (1993) 2359.
[4] B.-G. Englert, Phys. Rev. Lett. **77** (1996) 2154.
[5] W. M. Itano, J. C. Berquist, J. J. Bollinger, D. J. Wineland, U. Eichmann, and M. G. Raizen, Phys. Rev. A **57** (1998) 4176.
[6] C. Schön and A. Beige, Phys. Rev. A **64** (2001) 023806.
[7] C. Skornia, J. von Zanthier, G. S. Agarwal, E. Werner, and H. Walther, Phys. Rev. A **64** (2001) 063801.
[8] G. C. Hegerfeldt, Phys. Rev. A **47** (1993) 449;
G. C. Hegerfeldt and D. G. Sondermann, Quantum Semiclass. Opt. **8** (1996) 121.
[9] In the following we mark normalised vectors with a hat.
[10] For another example of equivalence between a quantum field theoretical model with a classical one see, e.g. J. Pachos, Phys. Rev. D **58** (1998) 125016.

Fortschr. Phys. **50** (2002) 5–7, 599–607

Quantum Fictitious Forces

I. Białynicki-Birula[1, 2], M. A. Cirone[2], J. P. Dahl[2, 3], T. H. Seligman[4], F. Straub[2], and W. P. Schleich[2]

[1] Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland
[2] Abteilung für Quantenphysik, Universität Ulm, D-89069 Ulm, Germany
[3] Chemical Physics, Department of Chemistry, Technical University of Denmark, DTU 207, DK-2800 Lyngby, Denmark
[4] Centro de Ciencias Físicas, University of Mexico (UNAM), Cuernavaca, Mexico

Abstract

We present Heisenberg's equation of motion for the radial variable of a free non-relativistic particle in D dimensions. The resulting radial force consists of three contributions: (i) the quantum fictitious force which is either attractive or repulsive depending on the number of dimensions, (ii) a singular quantum force located at the origin, and (iii) the centrifugal force associated with non-vanishing angular momentum. Moreover, we use Heisenberg's uncertainty relation to introduce a lower bound for the kinetic energy of an ensemble of neutral particles. This bound is quadratic in the number of atoms and can be traced back to the repulsive quantum fictitious potential. All three forces arise for a free particle: "Force without force".

1. Three Pillars of Quantum Mechanics

"*This paper tries to lay the foundations for a quantum theoretical mechanics, which is based solely on observable quantities.*"[1]

reads the abstract of Heisenberg's fundamental paper [1] entitled "On the quantum theoretical translation of kinematical and mechanical relations"[2]. What an ambitious plan! What a daring approach! What a dramatic departure from classical notions:

"*...we recall that in quantum theory it is not possible to associate with the electron a point in space as a function of time. Nevertheless, in quantum theory, we can also associate radiation with the electron: this radiation is first described by the frequencies which appear to be functions of two variables ...*"[3].

With these lines matrix mechanics enters the stage of physics. At this moment Heisenberg does not yet know that his quantities with two indices are matrices. However, he already

[1] Unless indicated differently, the English translation is always provided by the authors. The German original reads: „In der Arbeit soll versucht werden, Grundlagen zu gewinnen für eine quantentheoretische Mechanik, die ausschließlich auf Beziehungen zwischen prinzipiell beobachtbaren Größen basiert ist."

[2] The original title is: „Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen".

[3] The German text reads: „... müssen wir uns daran erinnern, daß es in der Quantentheorie nicht möglich war, dem Elektron einen Punkt im Raum als Funktion der Zeit mittels beobachtbarer Größen zuzuordnen. Wohl aber kann dem Elektron auch in der Quantentheorie eine Ausstrahlung zugeordnet werden; diese Strahlung wird beschrieben erstens durch die Frequenzen, die als Funktionen zweier Variablen auftreten ..."

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

recognizes that two such quantities $x(t)$ and $y(t)$ do not necessarily commute in quantum theory:

"*Whereas classically $x(t)\,y(t)$ is always equal to $y(t)\,x(t)$, this is in general not necessarily the case in quantum theory.*"[4)]

The importance of the commutativity of matrices stands out most clearly in Heisenberg's equation of motion

$$\frac{d\hat{A}}{dt} = \frac{i}{\hbar}\,[\hat{H}, \hat{A}] \tag{1}$$

for an arbitrary operator $\hat{A}$ and the Hamiltonian $\hat{H}$.

It is remarkable that this equation does not appear in Heisenberg's original paper [1] submitted on July 29, 1925. It emerges for the first time in the paper "On quantum mechanics"[5)] by M. Born and P. Jordan [2] submitted on September 27, 1925. The central role of (1) for quantum theory is emphasized in the "Dreimänner Arbeit" [3] and by Heisenberg in his Chicago lectures [4] in the spring of 1929. In the first textbook [5] on matrix mechanics by Born and Jordan it takes a slightly different form.

Another manifestation of the non-commutativity of matrices is the uncertainty relation proposed by Heisenberg in his paper [6] entitled: "The physical content of quantum kinematics and mechanics"[6)]. Here the uncertainty relation does not take the familiar form

$$\Delta q \cdot \Delta p \sim h \tag{2}$$

which now even appears on the stamp issued by the German postal service. Heisenberg states:

"*Let q_1 be the precision with which the value q is known ... Let p_1 be the precision with which the value p is determinable ... then, according to the elementary laws of the Compton effect p_1 and q_1 stand in the relation*

$$p_1 \cdot q_1 \sim h\,. \tag{3}$$

That this relation is a straightforward mathematical consequence of the rule $\boldsymbol{pq} - \boldsymbol{qp} = h/(2\pi i)$ will be shown below."[7)]

Two years later H. P. Robertson derived the generalized uncertainty relation [8]

$$\Delta\hat{A} \cdot \Delta\hat{B} \geq \frac{1}{2}\,|\langle[\hat{A}, \hat{B}]\rangle| \tag{4}$$

for the standard deviations

$$\Delta A \equiv \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2} \tag{5}$$

[4)] The corresponding German original is: „Während klassisch $x(t)\,y(t)$ stets gleich $y(t)\,x(t)$ wird, braucht dies in der Quantentheorie im allgemeinen nicht der Fall zu sein."

[5)] German original: „Zur Quantenmechanik".

[6)] German title: „Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". For the english translation we refer to Ref. [7].

[7)] The German text reads: „Sei q_1 die Genauigkeit, mit der der Wert q bekannt ist, ... p_1 die Genauigkeit, mit der der Wert p bestimmbar ist, ... so stehen nach elementaren Formeln des Comptoneffekts p_1 und q_1 in der Beziehung $p_1 \cdot q_1 \sim h$. Daß diese Beziehung in direkter mathematischer Verbindung mit der Vertauschungsrelation $\boldsymbol{pq} - \boldsymbol{qp} = h/(2\pi i)$ steht, wird später gezeigt werden."

and

$$\Delta B \equiv \sqrt{\langle \hat{B}^2 \rangle - \langle \hat{B} \rangle^2} \tag{6}$$

of two hermitian operators $\hat{A}$ and $\hat{B}$.

Observables as non-commuting operators, time evolution from the commutation relations with the Hamiltonian and the uncertainty relation summarize the essential ingredients of consequences of quantum mechanics. They form the three pillars of our understanding of the microscopic world form.

2. Outline of Paper

In a series of papers [9−13] we have shown that the quantum kinematics and dynamics of a free particle strongly depends on the number D of dimensions accessible to the particle. In particular, we have demonstrated [9] that a shell-like wave packet in two space dimensions first implodes and then explodes. The corresponding wave packet in three dimensions never implodes but only explodes. These phenomena are consequences of the quantum fictitious force [10−12] which in 2D is attractive − quantum anticentrifugal force [10, 11] − but which vanishes for 3D. Another consequence of this force is the dimensional enhancement [12] of the kinetic energy of an ensemble of particles. Indeed, for an appropriately prepared ensemble the kinetic energy of the total system is not proportional to the total number N of particles but to N^2. This effect which is reminiscent of superradiance [14] is due to the fact that interference depends strongly on the number of space dimensions [13]. This feature reflects itself in the quantum fictitious force which for dimensions larger than three is always repulsive. Moreover, its strength is quadratic in D.

So far our analysis has rested on the Schrödinger picture. In the present paper we use the Heisenberg picture to shine new light on these phenomena. In particular, we discuss Heisenberg's equation of motion for the radial coordinate. Moreover, we use the generalized uncertainty relation (4) to show that indeed the lower bound of the kinetic energy depends on the square of the number of dimensions. Throughout the paper we do not derive but rather motivate the results. The derivations are rather lengthy and will be published elsewhere.

Our paper is organized as follows. In Section 3 we formulate the problem and then turn in Section 4 to Heisenberg's equation of motion for the radial coordinate of a free particle in D dimensions. We show that a part of the radial force arises from the quantum fictitious potential which depends quadratically on D. In addition to this force there is one that is non−zero only at the origin. It results from the singular nature of the radial variable at the origin. Both forces are a consequence of the non-commutativity of position and momentum and exist even when the total angular momentum vanishes. The third contribution is the familiar quantum−modified centrifugal force due to a non-vanishing angular momentum. In Section 5 we turn to the problem of dimensional enhancement of kinetic energies. We start from the commutation relation between the operators of the momentum $\hat{\vec{p}}$ and the radial unit vector $\hat{\vec{e}}_r \equiv \hat{\vec{r}}/\hat{r}$ in D dimensions and derive the corresponding generalized uncertainty relation. This procedure allows us to put a lower bound on the total kinetic energy which depends quadratically on D. In Section 6 we briefly summarize our results.

3. Formulation of Problem

It is quite amusing to realize that Heisenberg's equation of motion and the uncertainty relation still hold surprises even for simple quantum systems. To bring this out most clearly

we focus on the motion of a free non-relativistic particle in D dimensions described by the Hamiltonian

$$\hat{H} \equiv \frac{\hat{\vec{p}}^2}{2M} \equiv -\frac{\hbar^2}{2M}\, \Delta^{(D)} \,. \tag{7}$$

Here M denotes the mass and the momentum

$$\hat{\vec{p}} \equiv \frac{\hbar}{i}\, \vec{\nabla}^{(D)} \tag{8}$$

is proportional to the D-dimensional gradient $\vec{\nabla}^{(D)}$ expressed in cartesian coordinates $x_1, \ldots, x_D$ giving rise to the D-dimensional Laplacian $\Delta^{(D)}$.

Two questions immediately come to mind: (i) What is the physical significance of $D > 3$ dimensions? (ii) How could there be anything interesting in the physics of a free particle?

The answer to the first question lies in the fact that we interprete the D dimensional hyper space of the quantum state of a single hypothetical particle as the configuration space of two or more non-interacting real particles in one two or three dimensions. In this way we arrive at $D = d \times N$ where d denotes the dimensions of the space in which N particles are allowed to move. Even when the particles do not interact with each other they can be entangled, that is the total wave function does not separate into the individual one-particle wave functions.

The answer to the second question rests on the special form of the state of the particle. It is an s-wave, which depends only on the hyper radius

$$r \equiv \sqrt{x_1^2 + x_2^2 + \ldots + x_D^2} \,. \tag{9}$$

Indeed, for such a choice of the wave function interesting effects occur as we show now.

4. Heisenberg's Equations of Motion

In this section we discuss the time evolution of the radial operator. In particular, we present the operators $M\dot{\hat{r}}$ and $M\ddot{\hat{r}}$ of the radial momentum and the force.

4.1. Radial Momentum Operator

We start our discussion by considering first the time evolution of the radial operator

$$\hat{r} \equiv \sqrt{\hat{x}_1^2 + \hat{x}_2^2 + \ldots + \hat{x}_D^2} \,. \tag{10}$$

For this purpose we substitute the Hamiltonian (7) and the operator $\hat{r}$ into Heisenberg's equation of motion (1) and arrive at

$$M\, \frac{d\hat{r}}{dt} = \frac{\hbar}{i}\, \frac{1}{2}\, [\Delta^{(D)}, \hat{r}] \,. \tag{11}$$

The action of the Laplacian on $\hat{r}$ creates a singularity at the origin. In order to keep track of this singularity it is advantageous to introduce the quantity

$$r_\epsilon \equiv \sqrt{\epsilon^2 + r^2} \tag{12}$$

perform all calculations for non-zero values of ϵ and then let ϵ tend to zero.

This approach leads to the formula

$$M\,\frac{d\hat{\boldsymbol{r}}}{dt}=\frac{\hbar}{i}\,\left(\frac{D-1}{2r}+\frac{\partial}{\partial r}\right)+\frac{\hbar}{i}\,\mathcal{D}(\hat{\boldsymbol{r}})=\frac{\hbar}{i}\,\frac{1}{r^{(D-1)/2}}\,\frac{\partial}{\partial r}\,r^{(D-1)/2}+\frac{\hbar}{i}\,\mathcal{D}(\hat{\boldsymbol{r}})\,, \quad (13)$$

where

$$\mathcal{D}(r)\equiv\lim_{\epsilon\to 0}\,\frac{1}{2}\,\frac{\epsilon^2}{\left(\epsilon^2+r^2\right)^{3/2}} \quad (14)$$

is a distribution [15].

Hence, the time rate of change of the radial operator $\hat{\boldsymbol{r}}$ consists of two parts: (*i*) The first term is a differentiation with respect to the radial variable, (*ii*) the second contribution is nonvanishing only at the origin and depends crucially on the number D of dimensions.

4.2. Radial Force Operator

We now turn to the second derivative of the radial variable, that is the radial acceleration. When we apply again Heisenberg's equation of motion, that is

$$\frac{d^2\hat{\boldsymbol{r}}}{dt^2}=\left(\frac{i}{\hbar}\right)^2\,[\hat{\boldsymbol{H}},\,[\hat{\boldsymbol{H}},\,\hat{\boldsymbol{r}}]]=-\frac{\hbar^2}{4M^2}\,[\Delta^{(D)},\,[\Delta^{(D)},\,\hat{\boldsymbol{r}}]] \quad (15)$$

we arrive after a lengthy algebra at the quantum force

$$M\,\frac{d^2\hat{\boldsymbol{r}}}{dt^2}\equiv\hat{\boldsymbol{F}}=-\frac{\partial V(\hat{\boldsymbol{r}})}{\partial\hat{\boldsymbol{r}}}+F_Q(\hat{\boldsymbol{r}})+F_l(\hat{\boldsymbol{r}}) \quad (16)$$

in radial direction on a free particle. Here we have introduced the abbreviation

$$V(r)\equiv\frac{\hbar^2}{2M}\,\frac{(D-1)\,(D-3)}{4r^2} \quad (17)$$

for the quantum fictitious potential and the contribution

$$F_Q(r)\equiv-\frac{\hbar^2}{2M}\,\left\{2\,\frac{D-1}{r}\,\frac{\partial\mathcal{D}}{\partial r}+\frac{\partial^2\mathcal{D}}{\partial r^2}+4\,\frac{1}{r^{D-1}}\,\frac{\partial}{\partial r}\,\left(r^{D-1}\mathcal{D}\,\frac{\partial}{\partial r}\right)\right\} \quad (18)$$

is a combination of distributions [15].

The quantum modified centrifugal force

$$F_l(r)\equiv\lim_{\epsilon\to 0}\,\left\{-\frac{\hbar^2}{M}\,\frac{1}{\left(\epsilon^2+r^2\right)^{1/2}}\,\frac{\hat{\boldsymbol{\Lambda}}^2}{r^2}\right\} \quad (19)$$

results from the square $-\hbar^2\hat{\boldsymbol{\Lambda}}^2$ of the angular momentum operator.

We emphasize that the first two terms are always present irrespective of the angular momentum. Even for *s*-waves, that is for the case of a vanishing angular momentum, do we find these two forces in the radial direction.

The potential V defined in (17) vanishes for $D = 1$ and $D = 3$ and is negative for $D = 2$. Hence, for two dimensions we have an attractive potential, that is an attractive force.

It is due to this attractive force that a shell-like wave function in $D = 2 \times 1 = 2$ space dimensions first implodes and then explodes. Since in $D = 3 \times 1 = 3$ dimensions the potential and the quantum fictitious force vanishes no such effect exists for this choice of space.

For $D \geq 4$ the potential is positive and the corresponding force is repulsive. This feature would be easy to understand if the angular momentum were non-zero. Indeed, in this case there would be a centrifugal force pushing the particles apart. However, we emphasize that this force is present irrespective of angular momentum. Moreover, it depends quadratically on the number D of dimensions. It becomes important for a large ensemble of particles where $D = d \times N$ with $1 \ll N$ gives rise to the dimensional enhancement [12] of kinetic energy discussed in the next section.

Moreover, we recognize that due to the term F_Q the radial force depends on the behaviour of the wave function at the origin. Indeed, the distributions in F_Q focus on the properties of the wave function, such as the values of the function and its derivatives at the origin. This term is the reason why a Gaussian initial wave function does not display the implosion effect but a wave function with a hole at the origin, such as a Gaussian multiplied by r^2 does [9].

5. Dimensional Enhancement of Kinetic Energy

We now show that under appropriate conditions the average kinetic energy of N non-interacting free particles increases quadratically with N. Here we again assume that the particles move in a d-dimensional space. Hence, the wave function lives in a $D = d \times N$-dimensional hyper space.

5.1. Lower Bound for Kinetic Energy

We first recall the commutation relation

$$[\hat{\vec{p}}, \hat{\vec{e}}_r] = \hat{\vec{p}} \cdot \hat{\vec{e}}_r - \hat{\vec{e}}_r \cdot \hat{\vec{p}} = \frac{\hbar}{i} \frac{D-1}{r} + 2 \frac{\hbar}{i} \mathcal{D}(\hat{r}) \tag{20}$$

between the momentum operator $\hat{\vec{p}}$ and the unit vector $\hat{\vec{e}}_r \equiv \hat{\vec{r}}/\hat{r}$ in radial direction.

When we substitute this expression into the right-hand side of the generalized uncertainty relation (4) and focus on wave functions that behave at the origin such that the distribution $\mathcal{D}$ cannot make a contribution we arrive at

$$(\Delta \vec{p})^2 \cdot (\Delta \hat{\vec{e}}_r)^2 \geq \frac{\hbar^2}{4} (D-1)^2 \langle r^{-1} \rangle^2 . \tag{21}$$

We emphasize that in this derivation we have used a slight generalization of the generalized uncertainty relation (4). In Robertson's formulation the operators $\hat{A}$ and $\hat{B}$ are scalars, whereas in our case they are vectors. Nevertheless, Robertson's derivation holds true.

Due to the isotropy of the s-wave we find for the variances

$$(\Delta \vec{p})^2 = \langle \vec{p}^{\,2} \rangle - \langle \vec{p} \rangle^2 = \langle \vec{p}^{\,2} \rangle \tag{22}$$

and

$$(\Delta \hat{\vec{e}}_r)^2 = \langle \hat{\vec{e}}_r^{\,2} \rangle - \langle \hat{\vec{e}}_r \rangle^2 = \langle \hat{\vec{e}}_r^{\,2} \rangle = \langle 1 \rangle = 1 . \tag{23}$$

In the last step we have assumed that the state is normalized.

Hence, we arrive at the inequality

$$\langle \hat{\vec{p}}^{\,2} \rangle \geq \frac{\hbar^2}{4}\,(D-1)^2\,\langle r^{-1}\rangle^2 \tag{24}$$

which implies for the average kinetic energy

$$E_{\rm kin} \equiv \left\langle \frac{\hat{\vec{p}}^{\,2}}{2M} \right\rangle \equiv \sum_{j=1}^{N} \left\langle \frac{\hat{\vec{p}}_j^{\,2}}{2M} \right\rangle \geq \frac{\hbar^2}{2M}\,\frac{(D-1)^2}{4}\,\langle r^{-1}\rangle^2 \equiv \frac{1}{4}\,(D-1)^2\,E_0\,, \tag{25}$$

where we have introduced the energy

$$E_0 \equiv \frac{\hbar^2}{2M}\,\langle r^{-1}\rangle^2\,. \tag{26}$$

When we recall that $D = d \cdot N$ and take the limit of large number of particles, that is, $N \gg 1$, this inequality takes the form

$$E_{\rm kin} \geq \frac{1}{4}\,(dN-1)^2\,E_0 \sim \frac{1}{4}\,d^2N^2\,E_0\,. \tag{27}$$

Hence, the lower bound of the average kinetic energy is determined by the square of the number d of space dimensions and the square of the number N of particles.

5.2. Examples of Wave Functions

Unfortunately the inequality (27) is slightly misleading since it gives the impression that the quadratic dependence is independent of the form of the wave function. Indeed, the energy E_0 depends on the dimensions of hyper space through the expectation value of $1/r$.

5.2.1. Thermodynamic case

We illustrate this statement for a Gaussian wave function

$$\psi(r) = \mathcal{N}\,\frac{1}{\sqrt{S^{(D)}}}\,\exp\left[-\frac{1}{2}\left(\frac{r}{\delta r}\right)^2\right]. \tag{28}$$

Here $\mathcal{N} \equiv \sqrt{2/\Gamma(D/2)}\,\delta r^{-D/2}$ is the normalization constant arising from the condition

$$1 = \int_0^\infty dr\,r^{D-1}\int d\Omega^{(D)}\,|\psi(r)|^2 \tag{29}$$

and $S^{(D)} \equiv \int d\Omega^{(D)}$ denotes the surface of the D-dimensional unit sphere. The quantity δr is a measure of the width of the Gaussian. For the special choice of the Gaussian wave-packet (28) the expectation value $\langle r^{-1}\rangle$ takes the form

$$\left\langle \frac{1}{r} \right\rangle = \mathcal{N}^2 \int_0^\infty dr\,r^{(D-2)}\,\exp\left[-\left(\frac{r}{\delta r}\right)^2\right] = \frac{1}{\delta r}\,\frac{\Gamma\left(\dfrac{D-1}{2}\right)}{\Gamma\left(\dfrac{D}{2}\right)}\,. \tag{30}$$

where Γ denotes the Gamma function.

With the help of the Stirling formula [16]

$$\Gamma(z) \cong e^{-z}\, z^{z-1/2}\, \sqrt{2\pi} \tag{31}$$

we find for $D = dN \gg 1$ the asymptotic formula

$$\frac{\Gamma\left(\dfrac{D-1}{2}\right)}{\Gamma\left(\dfrac{D}{2}\right)} \cong \sqrt{\frac{2}{D}} \tag{32}$$

which yields the expression

$$\left\langle \frac{1}{r} \right\rangle \simeq \frac{1}{\delta r}\, \sqrt{\frac{2}{D}}\,. \tag{33}$$

for the expectation value.

When we substitute this result into the expression (26) for the minimal energy we find from the inequality (27)

$$\frac{\hbar^2}{2M}\, \frac{1}{\delta r^2}\, \frac{D}{2} \leq E_{\text{kin}}\,. \tag{34}$$

From this analysis we recognize that the square of the expectation value cancels a power D in the lower bound creating only a linear dependence in D.

It is interesting to note that for the Gaussian wave function (28) the average kinetic energy satisfies the equal sign as shown in [12]. In this case the kinetic energy is proportional to the number of particles in complete agreement with thermodynamics.

5.2.2. Enhancement case

In contrast, the wave function

$$\psi(r) = \mathcal{N}\, \frac{1}{\sqrt{S^{(D)}}}\, \frac{\exp\left[-\dfrac{1}{2}\left(\dfrac{\beta}{r} + \kappa r\right)\right]}{r^{(D-1)/2}} \tag{35}$$

with the two parameters β and κ has a normalization constant $\mathcal{N}$ which according to (29) is determined by the condition

$$1 = \mathcal{N}^2 \int_0^\infty dr\, \exp\left[-\left(\frac{\beta}{r} + \kappa r\right)\right] \tag{36}$$

and is therefore independent of D. Hence, the expectation value

$$\left\langle \frac{1}{r} \right\rangle = \mathcal{N}^2 \int_0^\infty dr\, \frac{1}{r}\, \exp\left[-\left(\frac{\beta}{r} + \kappa r\right)\right] \tag{37}$$

is independent of D. The lower bound for the kinetic energy given by (25) and (27) depends quadratically on the number of particles.

6. Conclusions

Quantum mechanics has come a long way. From Heisenberg's deep insight that it is not the individual Bohr orbits in the atom that matter but their transition frequencies giving rise to matrices and a non-commutative algebra, via the uncertainty relation to the formalism we apply today. More than 75 years have passed since Heisenberg's lonely night at Helgoland giving birth to his fundamental paper. Despite its age and the multitude of examples quantum mechanics is still full of surprises and open questions. Here we do not allude to the question of the measurement process or the recent most sophisticated developments of quantum communication, quantum computing or quantum cryptography. Even the most elementary system of a free particle in D dimensions can display surprising features as discussed in this paper and best summarized in the spirit of J. A. Wheeler by the phrase: "Force without force". We are confident that the wonders of quantum mechanics will never cease to exist.

Acknowledgements

We thank G. Alber, M. V. Berry, J. Botero, A. Delgado, M. Fedorov, R. Glauber, D. Greenberger, M. J. W. Hall, R. F. O'Connell, G. Metikas, W. C. Schieve, G. Süßmann, E. C. G. Sudarshan, S. Varro, K. Vogel and J. A. Wheeler for many fruitful discussions. Moreover, we are grateful to the Humboldt-Stiftung for organizing this splendid and stimulating conference in honor of W. Heisenberg. I. B.-B., J. P. D. and T. S. acknowledge the most generous support of the Humboldt-Stiftung which made this work possible. M. A. C. and W. P. S. acknowledge financial support from the network 'QUEST', HPRN-CT-2000-00121, of the IHP program of the European Union.

References

[1] W. HEISENBERG, Z. Phys. **33** (1925) 879.
[2] M. BORN and P. JORDAN, Z. Phys. **34** (1925) 858.
[3] M. BORN, W. HEISENBERG, and P. JORDAN, Z. Phys. **35** (1926) 557.
[4] W. HEISENBERG, *The Physical Principles of the Quantum Theory* (University of Chicago Press, Chicago, 1930) (reprinted in an unabridged and unaltered version by Dover, New York, 1949).
[5] M. BORN and P. JORDAN, *Elementare Quantenmechanik* (Zweiter Band der Vorlesungen über Atommechanik) (Springer, Berlin, 1930).
[6] W. HEISENBERG, Z. Phys. **43** (1927) 172.
[7] J. A. WHEELER and W. H. ŻUREK, *Quantum Theory and Measurement* (Princeton University Press, Princeton, 1983).
[8] H. P. ROBERTSON, Phys. Rev. **34** (1929) 163.
[9] I. BIAŁYNICKI-BIRULA, M. A. CIRONE, J. P. DAHL, M. FEDOROV, and W. P. SCHLEICH, Phys. Rev. Lett. (submitted).
[10] M. A. CIRONE, G. METIKAS, and W. P. SCHLEICH, Z. Naturforsch. A **56** (2001) 48.
[11] M. A. CIRONE, K. RZĄŻEWSKI, W. P. SCHLEICH, F. STRAUB, and J. A. WHEELER, Phys. Rev. A **65** (2002) 022101.
[12] W. P. SCHLEICH and J. P. DAHL, Phys. Rev. A (submitted).
[13] M. A. CIRONE, J. P. DAHL, M. FEDOROV, D. GREENBERGER, and W. P. SCHLEICH, J. Phys. B **35** (2002) 191.
[14] R. H. DICKE, Phys. Rev. **93** (1954) 99.
[15] I. M. GEL'FAND and G. E. SHILOV, *Generalized Functions* (Academic Press, New York, 1964).
[16] M. ABRAMOWITZ and I. A. STEGUN, *Handbook of Mathematical Functions* (Dover Publications, New York, 1965).

Fortschr. Phys. **50** (2002) 5–7, 608–613

Quantum Teleportation

SAMUEL L. BRAUNSTEIN

Informatics, Bangor University, Bangor LL57 1UT, United Kingdom
schmuel@sees.bangor.ac.uk

Abstract

Given a single copy of an unknown quantum state, it is impossible in principle to identify it. The mostly inaccessible information carried by the state is termed quantum information. In contrast to classical information, it cannot be copied or cloned. This concept provides a theoretical underpinning for all aspects of quantum communication and quantum computation. Here we use it to consider quantum teleportation.

1. Introduction

During this symposium we have heard many things about Heisenberg's life as a brilliant physicist, including his struggles with both scientific and moral questions. While preparing to give this talk I looked at information about his life and found a very surprising fact: he nearly failed his final Ph.D. orals. The difficulty arose when Wien, who was on his examining committee, moved from mathematical questions to those on experimental physics. Heisenberg was unable to derive the resolving power of such devices as microscopes and telescopes. In the end, Heisenberg received a III, equivalent perhaps to a C, as the overall grade for his doctorate. This was in 1923.

It is perhaps ironic then, that this young man had within 4 years understood the limits to resolution better than anyone else of his time. The Heisenberg microscope gedanken experiment helped lay the foundations of quantum mechanics and the uncertainty principle. This principle has re-emerged as a corner stone of our understanding of the limitations of quantum information processing. Here we will consider its role in our understanding of quantum teleportation.

2. What's in a Name

Before we begin we shall discuss the term 'teleportation' which has only recently appeared within the scientific literature [1].

Having seen sufficient science-fiction movies in my time, I can make an initial stab at defining teleportation as:

`Some kind of instantaneous 'disembodied' transport.`

Now, this is inconsistent with relativity, so let's immediately change our definition to

`Some kind of 'disembodied' transport.`

The term disembodied requires some explanation, however, we shall prefer to be vague about it: The object moves from one place to another without manifestly appearing within the intervening space.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

In the classical day-to-day world we already have many examples which could fit this definition: a fax machine transports an image as electricity; a telephone transporting sound waves; overhead projectors, etc. Should these really count as teleportation? After all they are all really copying processes. They leave the image, the sound, what-have-you behind and send the copy across space in some disembodied way. An object is scanned, the information is transmitted in a different form and a copy is reconstructed.

Let us keep away from philosophical questions at the moment. However, we note that the above procedure has intrinsic limitations. It will not work, in general, for quantum systems. Consider for example a single-photon wavepacket prepared in a state of linear polarization down to a 1° accuracy. This means that the preparation procedure requires almost 8 bits worth of polarization orientation information.

Any attempt at measuring this orientation, will yield at most one bit's worth of information. For example, simply running this photon through a polarization-dependent beamsplitter will cause the photon to be transmitted or reflected. Yielding at most one bit. This demonstrates the principle, that in general, it is impossible to accurately determine an *unknown* quantum state. This is essentially a general statement of the Heisenberg uncertainty principle for quantum systems. A consequence of this is that it is likewise impossible to perfectly copy an unknown quantum state. For if a device were capable of perfect copying, we could apply it many times to our original system, yielding a vast ensemble. Then with this ensemble we could then measure the system's state to any desired accuracy.

This latter form of the quantum limitation is today called the no-cloning theorem [2] and embodies many of the key features of Heisenberg's original principle.

3. Teleportation Protocol

If no-cloning appears to forbid a copying mechanism for performing disembodied transport, what can teleportation of an unknown quantum state be? Clearly, it cannot be a copying procedure. Nonetheless, the aim of quantum teleportation is to:

`move an unknown quantum state across a classical communication channel.`

Let us call the sender Alice and the receiver Bob. There is no in-principle restriction to the eavesdropping on a classical communication channel (classical information may be freely copied). So it would seem that another agent Bob′ could receive a copy of Alice's message and follow the same protocol as the real Bob. His actions would produce a copy of Alice's state! However, this cannot be, as it would violate the no-cloning theorem.

To get around this problem quantum teleportation, distinguishes the 'real Bob' as special, since only he shares one-half of a maximally entangled state with Alice, whose half is somehow used in her part of the protocol. Even if another agent were to share such a state with her, it would not participate in her actions. This 'quantum link' between Alice and Bob is sufficient to get around the limitations placed on transport through a classical communication channel.

To see how this works, we shall give the mathematics here only for the teleportation of a state from a two-dimensional Hilbert space. A key feature which allows the protocol to work is that there is, in general, a basis of maximally entangled states

$$\begin{aligned}
|\mathrm{Ent}_1\rangle &\propto |\updownarrow\leftrightarrow\rangle - |\leftrightarrow\updownarrow\rangle \\
|\mathrm{Ent}_2\rangle &\propto |\updownarrow\leftrightarrow\rangle + |\leftrightarrow\updownarrow\rangle \\
|\mathrm{Ent}_3\rangle &\propto |\updownarrow\updownarrow\rangle - |\leftrightarrow\leftrightarrow\rangle \\
|\mathrm{Ent}_4\rangle &\propto |\updownarrow\updownarrow\rangle + |\leftrightarrow\leftrightarrow\rangle \ .
\end{aligned} \tag{1}$$

Let us label the incoming state given to Alice as $|\psi\rangle_{\text{in}}$ and suppose that Alice and Bob share between themselves the first of these entangled states. The total state would then be described as

$$|\Phi\rangle_{\text{in},A,B} = |\psi\rangle_{\text{in}} \otimes |\text{Ent}_1\rangle_{A,B} \,. \tag{2}$$

Since Eqs. (1) describe a complete set for a four-dimensional Hilbert space, we may decompose the states in Alice's hand. Formally we have

$$|\Phi\rangle_{\text{in},A,B} = \frac{1}{2} \sum_{j=1}^{4} |\text{Ent}_j\rangle_{\text{in},A} \otimes U_j \, |\psi\rangle_B \,, \tag{3}$$

where U_j are *fixed* unitary operators on a two-dimensional Hilbert space suitably chosen for this decomposition to hold. It is clear from the linearity of this equation that the U_j cannot depend on the state $|\psi\rangle$ itself.

Let us now suppose that Alice makes a measurement in this basis identifying her result by the associated subscripts $j_0 = 1, 2, 3, 4$. Conditioned on knowing this value the best description of the state remaining in Bob's hand is then given by

$$\rightarrow U_{j_0} \, |\psi\rangle_B \,. \tag{4}$$

If Alice now communicates this identification j_0 to Bob (requiring two bits of classical communication) then Bob knows which basis state Alice found. Based on this knowledge, he then operates $U_{j_0}^{\dagger}$ on his state, yielding

$$\rightarrow U_{j_0}^{\dagger} U_{j_0} \, |\psi\rangle_B = |\psi\rangle_B \,. \tag{5}$$

At this stage the protocol is complete and the state has been transferred between Alice and Bob. This procedure utilizes two channels: one potentially pre-existing channel of shared quantum entanglement; and one of classical communication used for Alice to transmit her measurement result to Bob. Neither by themselves contains any information about the original state. This form of transport would seem qualify as disembodied. Further, at no stage did we produce a copy of the original, nor has either Alice or Bob discovered any details about the state they participated in transporting.

4. Continuous Quantum Variables

Here we shall discuss one particular implementation of quantum teleportation based, not on states from a finite-dimensional Hilbert space, but on coherent and squeezed states of the electromagnetic field. An idealized version was first discussed by Vaidman [3]. This assumed maximally entangled states which in infinite Hilbert space dimensions corresponds to infinite energy states. Here we discuss a realizable scheme as originated by Braunstein and Kimble [4].

Optical entanglement is easily created. It turns out that almost any pure states (excluding coherent states) combined at a beamsplitter will yield entanglement. The simplest choice is of a pair of squeezed states with opposite squeezing, as shown in Fig. 1.

In this case, a pair of so-called twin beams are generated. To detect this kind of entanglement, one simply reverses the procedure: twin beams of this sort recombined at a beamsplitter resolve into unentangled squeezed states. The identity of the specific twin-beam state may be determined by measuring the 'displacement' of the position-squeezed beam and the

Fig. 1. Twin-beam entanglement created by combining squeezed states on a beamsplitter. Entanglement is 'resolved' into unentangled states by the time-reversed procedure.

'kick' of the momentum squeezed one. As we saw in the previous section the ability to create and detect entanglement are the key requirements in implementing teleportation. This also holds for teleportation in infinite dimensional Hilbert spaces. More details about the scheme for continuous quantum variables may be found in Ref. [4].

5. Criteria and Evidence for Teleporting Unknown Coherent States

In any realistic experiment we will not be able to consider teleporting *completely* unknown states. We will have some partial information about them. For example, we may wish to consider the teleportation of a limited class of easily constructed states, such as coherent states with a limited peak amplitude [5].

What criteria shall we test in such a circumstance? The limited set of states may be described as an alphabet of states $\{|\psi_\alpha\rangle\}$ with some associated probabilities P_α for their selection in any given run. If Alice and Bob follow a teleportation protocol, Bob hopes to have recreated the original state $|\psi_\alpha\rangle$ during each run. However, in a real experiment he will only achieve an approximate (probably mixed) state ρ_α compared to the original.

The average overlap probability

$$F = \int d\alpha\, P_\alpha \langle\psi_\alpha|\rho_\alpha|\psi_\alpha\rangle\,, \tag{6}$$

will be a measure of the performance of the teleportation. It is called the average fidelity of the teleportation.

The criteria we consider requires that this mean fidelity be better than Alice and Bob could achieve using a classical communication channel alone. Such a performance cannot be reproduced by a conventional copy-and-send strategy. Demonstrating a quantum state transfer protocol with mean fidelity better than could be achieved without entanglement guarantees that a quantum channel has been utilized. In teleportation protocols the only quantum channel is due to the entanglement.

For an alphabet of coherent states it can be shown that no more than a 50% mean fidelity is possible in the absence of shared entanglement [5]. This criteria assumes that Alice and Bob know only the *form* of the alphabet, but not the actual state being transmitted during each run. It makes no other assumption about their actions provided only that they communicate via a classical communication channel.

Fig. 2. Data from Ref. [6]. The lower plot shows the mean fidelity achievable in the absence of shared entanglement. Even at optimal feed-forward gain it never exceeds 50%. The upper plot shows the mean fidelity with twin-beam entanglement shared between sending and receiving stations.

Furusawa et al. implemented just such a protocol and criteria in their experimental demonstration of teleportation [6]. Data from that experiment is shown in Fig. 2. The bound of 50% mean fidelity was beaten only when sender and receiver had access to shared entanglement.

6. Conclusion

Quantum teleportation is a wholly new protocol for the transmission of quantum states. No experiments so far have demonstrated all of the features which teleportation has to offer. A preliminary list of these features reads:

- Quantum transport through a classical channel,
- Entanglement assisted transport of a quantum state,
- Many bits can be 'sent' via few bits,
- Can teleport to an *unknown* location,
- Can teleport half of an entangled state,
- Can teleport an entire space or filter a subspace.

Ultimately, which features are most important to achieve will depend on the applications to which teleportation is put. After all these years of our studying quantum mechanics, it still has beautiful surprises for us. There is no uncertainty about that.

References

[1] C. H. BENNETT, G. BRASSARD, C. CREPEAU, R. JOZSA A. PERES, and W. K. WOOTTERS, Phys. Rev. Lett. **70** (1993) 1895.
[2] W. K. WOOTTERS and W. H. ZUREK, Nature **299** (1982) 802.
[3] L. VAIDMAN, Phys. Rev. A **49** (1994) 1473.
[4] S. L. BRAUNSTEIN and H. J. KIMBLE, Phys. Rev. Lett. **80** (1998) 869.
[5] S. L. BRAUNSTEIN, C. A. FUCHS, and H. J. KIMBLE, J. Mod. Opt. **47** (2000) 267.
[6] A. FURUSAWA, J. L. SØRENSEN, S. L. BRAUNSTEIN, C. A. FUCHS, H. J. KIMBLE, and E. S. POLZIK, Science **282** (1998) 706.

Fortschr. Phys. **50** (2002) 5–7, 614–623

Heisenberg's Introduction of the "Collapse of the Wavepacket" into Quantum Mechanics

RAYMOND Y. CHIAO[1)] and PAUL G. KWIAT[2)]

[1)] Department of Physics, University of California, Berkeley, CA 94720-7300, USA
chiao@physics.berkeley.edu
[2)] Department of Physics, University of Illinois, Urbana, IL 61801, USA
kwiat@uiuc.edu

Abstract

Heisenberg in 1929 introduced the "collapse of the wavepacket" into quantum theory. We review here an experiment at Berkeley which demonstrated several aspects of this idea. In this experiment, a pair of daughter photons was produced in an entangled state, in which the sum of their two energies was equal to the sharp energy of their parent photon, in the nonlinear optical process of spontaneous parametric down-conversion. The wavepacket of one daughter photon collapsed upon a measurement-at-a-distance of the other daughter's energy, in such a way that the total energy of the two-photon system was conserved. Heisenberg's energy-time uncertainty principle was also demonstrated to hold in this experiment.

1. Introduction

In this Symposium in honor of Heisenberg's Centennial, it is appropriate to begin by recalling the fact that in the spring of 1929, during his lectures at the University of Chicago, Heisenberg introduced the important concept of the "collapse of the wavepacket" into quantum theory [1]. This idea, which he referred to as the "reduction of the wavepacket", was closely related to the idea of the "collapse of the wavefunction," which was introduced into the standard Copenhagen interpretation of quantum mechanics in connection with the probabilistic interpretation of the wavefunction due to Born [2]. In the context of a remark concerning the spreading of the wavepacket of an electron, Heisenberg stated the following [1]:

> In relation to these considerations, one other idealized experiment (due to Einstein) may be considered. We imagine a photon which is represented by a wave packet built up out of Maxwell waves [3]. It will thus have a certain spatial extension and also a certain range of frequency. By reflection at a semi-transparent mirror, it is possible to decompose it into two parts, a reflected and a transmitted packet. There is then a definite probability for finding the photon either in one part or in the other part of the divided wave packet. After a sufficient time the two parts will be separated by any distance desired; now if an experiment yields the result that the photon is, say, in the reflected part of the packet, then the probability of finding the photon in the other part of the the packet immediately becomes zero. The experiment at the position of the reflected packet thus exerts a kind of action (reduction of the wave packet) at the distant point occupied by the transmitted packet, and one sees that this action is propagated with a velocity greater than that of light. However, it is also obvious that this kind of action can never be utilized for the transmission of signals so that it is not in conflict with the postulates of the theory of relativity.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

At the heart of Heisenberg's (and, earlier, Einstein's) conception of the "collapse of the wavepacket", was the *indivisibility* of the individual quantum, here of the light quantum – the photon – at the beam splitter [4]. Ultimately, it was the indivisibility of the photon that enforced the collapse of the wavepacket, whenever a detection of the particle occurred at one of the two exit ports of the beam splitter (for example, whenever a "click" occurred at a counter placed at the reflection port); because the photon was *indivisible*, it had *either* to be reflected, *or* to be transmitted at the beam splitter, but not both. Remarkably, a non-detection by a 100%-efficient detector will also collapse the state, so that the photon is definitely in the other arm [5].

Heisenberg believed that the reduction or collapse of the wavepacket was indeed a physical *action*, and not just a convenient *fiction* which was useful in the interpretation of quantum phenomena but which had no physical reality attached to it. It is an irony of history that although Heisenberg attributed this idea to Einstein, it was later totally rejected by Einstein along with all of its consequences, as unexplained, "spooky actions-at-a-distance (spukhafte Fernwirkungen)". The "collapse" idea, and its later developments, culminated in Einstein's ultimate rejection of quantum theory as being an incomplete theory of the physical world.

In contrast, von Neumann embraced the idea, and further sharpened it by introducing the notion of "projection of rays in Hilbert space" upon measurement, in the last two chapters entitled 'Measurement and Reversibility' and 'The Measuring Process' of his book *Mathematical Foundations of Quantum Mechanics* [6]. The physics of the "collapse postulate" introduced by Heisenberg was thereby tied to the mathematics of the "von Neumann projection postulate." In this way, the Copenhagen interpretation of quantum mechanics, or at least one version of it, was completed.

Here we shall review an experiment performed at Berkeley on another aspect of the "collapse of the wavepacket", as viewed from the standard Copenhagen viewpoint. This experiment involved a study of the correlated behaviors of two photon wavepackets in an entangled state of energy. The two entangled photon wavepackets were produced in the process of spontaneous parametric down-conversion, in which a parent photon from an ultraviolet (UV) laser beam was split inside a nonlinear crystal into two daughter photons, conserving energy and momentum during this process. One can view this quantum nonlinear optical process as being entirely analogous to a radioactive decay process in nuclear physics, in which a parent nucleus decays into two daughter nuclei. We shall see that a measurement of the energy of one daughter photon has an instantaneous collapse-like action-at-distance upon the behavior of the other daughter photon.

2. The Entangled Photon-pair Light Source: Spontaneous Parametric Down-conversion

We produced pairs of energy-entangled photon wave packets by means of spontaneous parametric down-conversion, also known as "parametric fluorescence", in an optical crystal with a nonvanishing $\chi^{(2)}$ nonlinear optical susceptibility [7]. In our experiment, we employed a crystal of potassium dihydrogen phosphate (KDP) [8]. The lack of inversion symmetry in crystals such as KDP breaks the usual parity-conservation selection rule, so that it is not forbidden for one photon to decay into two photons inside the crystal.

There are many ways in which a single, monoenergetic, parent photon (conventionally called the "pump" photon, originating in our experiment from an argon ion UV laser operating at $\lambda = 351$ nm) can decay into two daughter photons (conventionally called the "signal" and the "idler" photons), while distributing its energy between the two in a continuous fashion. There is therefore no reason why the daughter photons would necessarily be monoenergetic. In fact, as a result of the normal dispersion in the linear refractive index of

Fig. 1. Color photograph (an end-on view) of the spontaneous parametric down-conversion from an ultraviolet ($\lambda = 351$ nm) laser beam which has traversed a KDP crystal. (Some of the UV laser light, which has leaked through an UV rejection filter, caused the irregular splotches near the center). By means of Pinhole 1 and Pinhole 2, correlated pairs of photons emitted near the same red wavelength ($\lambda = 702$ nm) were selected out for coincidence detection. This picture illustrates yet another striking aspect of the "collapse" idea, which is different from the one discussed in the text. Although the angular distribution of the probability amplitude for the emission of the photon pair starts off as an azimuthally isotropic ring around the center, once a "click" is registered by a detector placed behind one of the pinholes, there is the sudden onset of a type of "spontaneous symmetry breaking," in which the probability for detection of the other photon suddenly vanishes everywhere around the ring, except at the other pinhole, where the probability of finding it suddenly becomes unity. (Adapted from A. Migdall and A. V. Sergienko's original photo).

the crystal, it turns out that the conservation of energy and momentum in the two-photon decay process results in the production of a rainbow of conical emissions of photon pairs with a wide spectrum of colors, which is shown in Fig. 1. Two photons on diametrically opposite sides of the rainbow are emitted in a pairwise fashion, conserving energy and momentum in the emission process.

By means of two pinholes, we selected out of the rainbow for further study two tightly correlated, entangled photons, which were emitted around the same red wavelength (i.e., $\lambda = 702$ nm, at twice the wavelength of the pump photon). For millimeter-scale pinholes, which span a few percent of the full visible spectrum, the resulting photon wavepackets typically have subpicosecond widths [9]. Thus, two photon counters placed behind these two pinholes would register tightly correlated, coincident "clicks".

The KDP crystal which we used was 10 cm long, cut such that its c-axis was 50.3° to the normal of its input face; the UV laser beam was normally incident on the KDP crystal face, with a vertical linear polarization. The correlated signal and idler photon beams were both horizontally polarized. The two pinholes were placed at +1.5° and at −1.5° with respect to the UV laser beam, so that degenerate pairs of photons centered at a wavelength of 702 nm were selected for study. Thus, in this parametric fluorescence process, a single parent photon with a sharp spectrum from the UV laser was spontaneously converted inside the crystal into a pair of daughter photons with broad, conjugate spectra centered at half the parent's UV energy.

The photon pair-production process due to parametric down-conversion produces the following entangled state [10]:

$$|\psi\rangle = \int dE_1\, A(E_1)\, |1\rangle_{E_1}\, |1\rangle_{E_2=E_0-E_1}\,, \tag{1}$$

where E_0 is the energy of the parent UV photon, E_1 is the energy of the first member of the photon pair (the "idler" daughter photon), $E_2 = E_0 - E_1$ is the energy of the second member of the photon pair (the "signal" daughter photon), and $A(E_1)$ is the probability amplitude for the emission of the pair. The first (second) photon is in the one-photon Fock state $|1\rangle_{E_1}$ ($|1\rangle_{E_2}$). Energy must be conserved in the pair-emission process, and this is indicated by the energy subscript of the one-photon Fock state for the second photon $|1\rangle_{E_2=E_0-E_1}$. The *integral* over the product states $|1\rangle_{E_1}\,|1\rangle_{E_2=E_0-E_1}$ indicates that the total state $|\psi\rangle$ is the *superposition* of product states. Hence, this state exhibits mathematical *nonfactorizability*, the meaning of which is physical *nonseparability*: It is an entangled state of energy. Therefore, the results of measurements of the energy of the first photon will be tightly correlated with the outcome of measurements of the energy of the second photon, even when the two photons are arbitrarily far away from each other. There result Einstein-Podolsky-Rosen effects, which are nonclassical and nonlocal [11, 12].

3. Apparatus for the Detection of Entangled Photon Pairs: Michelson Interferometry, Spectral Filtering, and Coincidence Counting

In Fig. 2, we show a schematic of the apparatus. Entangled photons, labeled "signal" and "idler", were produced in the KDP crystal. The upper beam of idler photons was transmitted through the "remote" filter F1 to the detector D1, which was a cooled RCA C31034A-02 photomultiplier tube. Horizontally polarized signal photons in the lower beam entered a Michelson interferometer, inside one arm of which was placed a pair of zero-order quarter-wave plates Q1 and Q2. The fast axis of Q1 was fixed at 45° to the horizontal, while the fast axis of Q2 was slowly rotated by a computer-controlled stepping motor, in order to scan for fringes. After leaving the Michelson, the signal photon impinged on a second beamsplitter B2, where it was transmitted to detector D2 through filter F2, or reflected to detector D3 through filter F3. Filters F2 and F3 were identical: They both had a broad bandwidth of 10 nm centered at $\lambda = 702$ nm. Detectors D2 and D3 were identical to D1.

Coincidences between detectors D1 and D2 and between D1 and D3 were detected by feeding their outputs into constant fraction discriminators and coincidence detectors after appropriate delay lines. We used EGG C102B coincidence detectors with coincidence window resolutions of 1.0 ns and 2.5 ns, respectively. Also, triple coincidences between D1, D2, and D3 were detected by feeding the outputs of the two coincidence counters into a third coincidence detector (a Tektronix 11302 oscilloscope used in a counter mode). The various count rates were stored on computer every second.

The two quarter-wave plates Q1 and Q2 inside the Michelson generated a geometrical (Pancharatnam-Berry) phase, which was proportional to the angle between the fast axes of

Fig. 2. Schematic of apparatus to demonstrate another aspect of Heisenberg's "collapse of the wavepacket", in which a sharp measurement of the energy of one member of an entangled state results in a collapse in the width of the wavepacket of the other member. A photon pair selected by means of pinholes shown in Figure 1 (not shown here), is emitted from the KDP crystal in parametric down-conversion of a parent photon from the UV laser. One member of the pair (the "idler") is sent through the "remote" interference filter F1 before detection by a photomultiplier D1. The other member of the pair (the "signal") is sent through a Michelson interferometer, whose optical path length difference is scanned by means of two quarter-wave plates Q1 and Q2. The beam splitter B2, filters F2 and F3, photomultipliers D2 and D3, coincidence gates N_{12} and N_{13}, all serve to select out only those photons which are members of entangled pairs, for observation. The triple coincidence gate N_{123} serves to check that only single-photon Fock states are detected, so that no classical explanation of the results would be possible. (From Ref. [8])

the two plates. Here, one should view the use of the quarter-wave plates as simply a convenient method for scanning the phase difference of the Michelson interferometer. For the details concerning the geometrical phase, see [8] and [11].

We took data both inside and outside the white-light fringe region of the Michelson (i.e., where the two arms have nearly equal optical path lengths), where the usual interference in singles detection occurs. We report here only on data taken outside this region, where the optical path length difference was set at a fixed value much greater than the coherence length (or wavepacket width) of the signal photons determined by the filters F2 and F3. Hence, the fringe visibility seen by detectors D2 and D3 in singles detection was essentially zero.

4. Theory

We present here a simplified analysis of this experiment. For a detailed treatment based on Glauber's correlation functions, see [11]. The entangled state of light after passage of the signal photon through the Michelson is given by

$$|\psi\rangle_{\text{out}} = \frac{1}{2} \int dE_1\, A(E_1)\, |1\rangle_{E_1}\, |1\rangle_{E_2=E_0-E_1} (1 + \exp\left(i\phi(E_0 - E_1)\right), \tag{2}$$

where $\phi(E_0 - E_1)$ is the total phase shift of the signal photon inside the Michelson, including the Pancharatnam-Berry phase; the factor of $1/2 = (1/\sqrt{2})^2$ comes from the two interactions with the Michelson 50/50 beam splitter B1. The coincidence count rates N_{12} between the pair of detectors D1 and D2 (and N_{13} between the pair D1 and D3) are proportional to the probabilities of finding at the same instant t (more precisely, within the detection time, typically 1 ns) one photon at detector D1 and also one photon at detector D2 (and for N_{13}, one photon at detector D1 and also one photon at detector D3). When a *narrowband* filter F1 centered at energy E_{1_0} is placed in front of the detector D1, N_{12} becomes proportional to

$$\left|\psi'_{\text{out}}(\boldsymbol{r}_1, \boldsymbol{r}_2, t)\right|^2 = \left|\langle \boldsymbol{r}_1, \boldsymbol{r}_2, t \mid \psi\rangle'_{\text{out}}\right|^2 \propto 1 + \cos\phi\,, \tag{3}$$

where $\boldsymbol{r}_1$ is the position of detector D1, $\boldsymbol{r}_2$ is the position of detector D2, and the prime denotes the output state after a von-Neumann projection onto the eigenstate associated with the sharp energy E_{1_0} upon measurement. Therefore, the phase ϕ is determined at the sharp energy $E_0 - E_{1_0}$. In order to conserve total energy, the energy bandwidth of the collapsed signal photon wavepacket must depend on the bandwidth of the filter F1 in front of D1, through which it did not pass. Therefore, the visibility of the signal photon fringes seen in coincidences should depend critically on the bandwidth of this *remote* filter. For a narrowband F1, this fringe visibility should be high, provided that the optical path length difference of the Michelson does not exceed the coherence length of the *collapsed* wavepacket (recall that due to the energy-time uncertainty principle, collapsing to a *narrower* energy spread actually leads to *longer* wavepacket). It should be emphasized that the width of the collapsed signal photon wavepacket is therefore determined by the *remote* filter F1, through which this signal photon has apparently never passed! If, however, a sufficiently broadband remote filter F1 is used instead, such that the optical path length difference of the Michelson is much greater than the coherence length of the collapsed wavepacket, then the coincidence fringes should disappear.

5. Results

In Fig. 3, we show data which confirm these predictions. In the lower trace (squares) we display the coincidence count rate between detectors D1 and D3, as a function of the angle θ between the fast axes of the wave plates Q1 and Q2, when the remote filter F1 was quite narrowband (i.e., with a bandwidth of 0.86 nm). The calculated coherence length of the collapsed signal photon wavepacket (570 microns) was greater than the optical path length difference at which the Michelson was set (220 microns). The observed visibility of the coincidence fringes was quite high, viz., 60% $\pm$ 5%, indicating that the collapse of the signal photon wavepacket had indeed occurred.

In the upper trace (triangles) we display the coincidence count rate versus θ when a broadband remote filter F1 (i.e., one with a bandwidth of 10 nm) was substituted for the narrowband one. The coherence length of the collapsed signal photon wavepacket in this case should have been only 50 microns, which is shorter than the 220 micron optical path length difference at which the Michelson was set. The observed coincidence fringes have now indeed disappeared, indicating that the collapse of the signal photon wavepacket (this time to a *shorter* temporal width) has again occurred.

In addition to the above coincidences-counting data, we also took singles-counting data at detector D3 (where are not shown here), at the same settings of the optical path length of the Michelson for the traces shown in Fig. 3. We observed no visible fringes in the singles-counting data during the scan of θ, with a visibility less than 2%. This indicates that we were well outside of the white-light fringe of the Michelson. More importantly, this also demonstrates

Fig. 3. Data demonstrating the phenomenon of the "collapse of the wavepacket." The visibility of the fringes from the Michelson interferometer seen in coincidence detection for the signal photon (see Fig. 2), depends on the bandwidth of the *remote* filter F1, *through which it has apparently never passed*. In the lower trace, F1 is narrowband, which results in the collapsed signal photon wavepacket having a long coherence length, and thus in the observed high-visibility fringes. However, when the remote, narrowband filter F1 is replaced by a broadband one, the collapsed signal photon wavepacket now has a short coherence length (shorter than the optical path length difference of the Michelson), so that the fringes disappear, as is evident in the upper trace. (Data from Ref. [8])

that only those photons which are detected in *coincidence* with their twins which have passed through the narrowband filter F1, exhibit the observed phenomenon of the collapse of the wavepacket upon detection. In other words, only *entangled* pairs of photons detected by the coincidences-counting method show this kind of "collapse" behavior.

Heisenberg's energy-time uncertainty principle was also demonstrated during the course of this experiment [13]. The width Δt_2 of the collapsed signal photon wavepacket, which was measured by means of the Michelson, satisfied the inequality

$$\Delta E_2 \, \Delta t_2 \geq \hbar/2 \,, \tag{4}$$

where the energy width ΔE_2 of the collapsed signal photon wavepacket, was determined by the measured energy width ΔE_1 of the idler photon, in order to conserve total energy. Hence, the energy width ΔE_2 of the signal photon, which enters into the Heisenberg uncertainty relation (4), was actually the width ΔE_1 of the remote filter F1, through which this signal photon did not pass.

6. Discussion

Any classical electromagnetic field explanation of these results can be ruled out. We followed here the earlier experiment of Grangier, Roger, and Aspect [14], in which they showed that one can rule out any classical-wave explanation of the action of an optical

beam splitter, by means of triple coincidence measurements in a setup similar to that shown in Fig. 2. Let us define the parameter

$$a = \frac{N_{123} N_1}{N_{12} N_{13}} , \tag{5}$$

where N_{123} is the count rate of triple coincidences between detectors D1, D2, and D3, N_{12} is the count rate of double coincidences between detectors D1 and D2, N_{13} is the count rate of double coincidences between detectors D1 and D3, and N_1 is the count rate of singles detection by D1 alone. From Schwartz's inequality, it can be shown that for any classical-wave theory for electromagnetic wavepackets, the inequality $a \geq 1$ must hold [14]. By contrast, quantum theory predicts that $a = 0$.

The physical meaning of the inequality $a \geq 1$ is that any classical wavepacket divides itself smoothly at a beam splitter, and this always results in triple coincidences after the beam splitter B2, in conjunction with a semi-classical theory of the photoelectric effect. By contrast, a single photon is indivisible, so that it cannot divide itself at the beamsplitter B2, but rather must exit through only one or the other of the two exit ports of the beam splitter, and this results in zero triple coincidences (except for a small background arising from multiple pair events).

For coincidence-detection efficiencies η less than unity, this inequality becomes $a \geq \eta$. We calibrated our triple-coincidence counting system by replacing the photon-pair light source by an attenuated light bulb, and measured $\eta = 70\% \pm 7\%$. During the data run of Fig. 3 (lower trace), we measured values of a shown at the vertical arrows. The average value of a is 0.08 ± 0.04, which violates by more than thirteen standard deviations the predictions based on any classical-wave theory of electromagnetism.

It is therefore incorrect to explain these results in terms of a stochastic-ensemble model of classical electromagnetic waves, along with a semi-classical theory of photoelectric detection [15]. Pairs of classical waves with conjugately correlated, but random, frequencies could conceivably yield the observed interference patterns, but they would also yield many more triple coincidences than were observed.

One might be tempted to explain our results simply in terms of conditioning on the detection of the idler photons to post-select signal photons of pre-existing definite energies. And in fact, such a local realistic model can account for the results of this experiment with no need to invoke a nonlocal collapse. However, in light of the observed violations of Bell's inequalities based on energy-time variables [12], it is incorrect to interpret these results in terms of a statistical ensemble of signal and idler photons which possess definite, but unknown, energies before measurement (i.e., before filtering and detection). Physical observables, such as energy and momentum, cannot be viewed as local, realistic properties which are carried by the photon during its flight to the detectors, before they are actually measured.

We have chosen to interpret our experiment in terms of the "collapse" idea. However, it should be stressed that this is but one interpretation of quantum measurement. Other interpretations exist which could possibly also explain our results. They include the Bohm-trajectory picture [16], the many-worlds interpretation [17], the advanced-wave or transactional model by Cramer [18] (and related ideas by Klyshko [19]), the "consistent-histories" quantum-cosmology picture [20], and others.

An obvious follow-up experiment (but one which has not yet been performed) is a version of Wheeler's "delayed-choice" experiment [21], in which one could increase, as much as one desires, the distance from the source to the filter F1 and the detector D1 on the "remote" side of the apparatus, as compared to the distance to detectors D2 and D3, etc., on the "near" side. The arbitrary choice of whether filter F1 should be broadband or narrowband could then be delayed by the experimenter until well after "clicks" had already

been irreversibly registered in detectors D2 or D3. In this way, we can be sure that the signal wavepacket on the near side of the apparatus could not have known, well in advance of the experimenter's arbitrary and delayed choice of F1 on the remote side of the apparatus, whether to have collapsed to a broad or to a narrow wavepacket.

However, there is no paradox here since the determination of coincident events can only be made *after* the records of the "clicks" at both near and remote detectors are brought together and compared, and only then does the appearance or disappearance of interference fringes in coincidences become apparent. The bringing together of these records requires the propagation of signals through classical channels with appropriate delays, such as the post-detection coaxial delay lines leading to the coincidence gates shown in Fig. 2. Classical channels propagate signals with discontinuous fronts, such as "clicks," at a speed limited by *c*. Hence there is no conflict with the postulates of the theory of relativity.

It is therefore incorrect to say that the experimenter's arbitrary choice of the filter F1 on the remote side of the apparatus somehow *caused* the collapse of the signal photon wavepacket on the near side. Only nonlocal, instantaneous, uncaused correlations-at-a-distance are predicted by quantum theory. Clearly, the collapse phenomenon is nonlocal and noncausal in nature.

In conclusion, we have demonstrated that the *nonlocal* collapse of the wavefunction or wavepacket in the Copenhagen interpretation of quantum theory, which was introduced by Heisenberg in 1929, leads to a self-consistent description of our experimental results. Whether or not fringes in coincidence detection show up in a Michelson interferometer on the near side of the apparatus, depends on the arbitrary choice by the experimenter of the remote filter F1 through which the photon on the near side has evidently never passed. This collapse phenomenon, however, is clearly *noncausal*, as a "delayed-choice" extension of our experiment would show.

Acknowledgments

This work was supported in part by the ONR. We thank Charles H. Townes for a careful reading of this manuscript. RYC would like to thank John C. Garrison for helpful discussions, and Aephraim M. Steinberg for pointing out to me once again Ref. [1].

References

[1] *The Physical Principles of the Quantum Theory*, W. Heisenberg (University of Chicago Press, Chicago, 1930), reprinted by Dover Publications, p. 39.

[2] M. Born, Z. f. Physik **37** (1926) 863; see also the commentary by L. Rosenfeld in Quantum Theory and Measurement, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, 1983), p. 50.

[3] Heisenberg added the following footnote here: "For a single photon the configuration space has only three dimensions; the Schrödinger equation of a photon can thus be regarded as formally identical with the Maxwell equations".

[4] It is interesting to note here that much earlier, Newton, in connection with his experimental studies of the optical beam splitter (or what Heisenberg called a "semi-transparent mirror"), and in particular, of the phenomenon of Newton's rings, struggled with the problem of how to reconcile his concept of an indivisible "corpuscle" (i.e., a particle) of light, with the fact that these corpuscles were *coherently* divided into transmitted and reflected parts at air-glass interfaces, in such a way that these parts could later interfere with each other, and thus produce Newton's rings.

[5] R. H. Dicke, Am. J. Phys. **49** (1981) 925;
A. C. Elitzur and L. Vaidman, Found. Phys. **23** (1993) 987;
P. G. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M. A. Kasevich, Phys. Rev. Lett. **74** (1995) 4763.

[6] *Mathemathical Foundations of Quantum Mechanics*, J. VON NEUMANN (Princeton University Press, Princeton, 1955), p. 347 ff.
[7] S. E. HARRIS, M. K. OSHMAN, and R. L. BYER, Phys. Rev. Lett. **18** (1967) 732;
D. MAGDE and H. MAHR, Phys. Rev. Lett. **18** (1967) 905;
D. C. BURNHAM and D. L. WEINBERG, Phys. Rev. Lett. **25** (1970) 84.
[8] P. G. KWIAT and R. Y. CHIAO, Phys. Rev. Lett. **66** (1991) 588.
[9] Strictly speaking, because we have a *continuous* (CW) pump, before detecting one of the down-conversion photons we cannot describe them as wavepackets. It is only *after* detecting one photon that its sister photon is in a proper wavepacket.
[10] We omit here the predominant contribution of the vacuum, corresponding to events where the pump photon did *not* down convert; such events are never measured in a photon counting experiment such as ours. We also omit the entirely negligible contribution from multiple-pair events (i.e., in which more than one pair was produced within the detection time).
[11] P. G. KWIAT, Ph. D. Thesis, U. C. Berkeley (1993, unpublished but available at <http://wwwlib.umi.com/dxweb/details?doc_no=2717843>).
[12] J. D. FRANSON, Phys. Rev. Lett. **62** (1989) 2205;
J. BRENDEL, E. MOHLER, and W. MARTIENSSEN, Europhys. Lett. **20** (1992) 575;
P. G. KWIAT, A. M. STEINBERG, and R. Y. CHIAO, Phys. Rev. A **47** R2472;
S. AERTS, P. KWIAT, J.-Å. LARSSON, and M. ŻUKOWSKI, Phys. Rev. Lett. **83** (1999) 2872.
[13] R. Y. CHIAO, P. G. KWIAT, and A. M. STEINBERG, in the *Workshop on Squeezed States and Uncertainty Relations*, D. HAN et al., eds., NASA Conference Publication 3135, NASA, Washington, DC, 1991, p. 61.
[14] P. GRANGIER, G. ROGER, and A. ASPECT, Europhys. Lett. **1** (1986) 173.
[15] J. F. CLAUSER, Phys. Rev. D **9** (1974) 853.
[16] D. BOHM, Phys. Rev. **85** (1951) 166; **85** (1952) 180;
D. BOHM, B. J. HILEY, and P. N. KALOYEROU, Phys. Rep. **144** (1987) 321.
[17] H. EVERITT, Rev. Mod. Phys. **29** (1957) 454;
H. EVERITT, *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1973);
L. VAIDMAN 'The Many-Worlds Interpretation of Quantum Mechanics,' to appear in The Stanford Encyclopedia of Philosophy, available at <http://www.tau.ac.il/126vaidman/mwi/mw2.html>.
[18] J. G. CRAMER, Rev. Mod. Phys. **58** (1986) 647; Int. J. Th. Phys. **27** (1988) 227; 'The Plane of the Present and the New Transactional Paradigm of Time' in *Time and the Instant*, Robin Durie, ed., (Clinamen Press, Manchester, 2000), Chap. 8.
[19] D. N. KLYSKO, Phys. Lett. A **132** (1988) 299.
[20] M. GELL-MANN and J. B. HARTLE, 'Quantum mechanics in light of quantum cosmology,' in *Complexity, Entropy and the Physics of Information*, W. H. Zurek, ed. (Addison-Wesley, Reading, 1990), p. 425; Phys. Rev. D **47** (1993) 3345.
[21] J. A. WHEELER, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, 1983), p. 182.

Fortschr. Phys. **50** (2002) 5–7, 624–629

The Hamiltonian Mass and Asymptotically Anti-de Sitter Space-times

PIOTR T. CHRUŚCIEL

Département de Mathématiques, Faculté des Sciences
Parc de Grandmont, F-37200 Tours, France
chrusciel@univ-tours.fr

Abstract

We give a Hamiltonian definition of mass for asymptotically hyperboloidal Riemannian manifolds, or for spacelike hypersurfaces in space-times with metrics which are asymptotic to the anti-de Sitter one.

1. Introduction

In classical mechanics the energy is most conveniently defined, up to a constant, via Hamilton's equations of motion,

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}\,, \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}\,, \tag{1}$$

or, equivalently,

$$-dH = \frac{dp_i}{dt}\; dq^i - \frac{dq^i}{dt}\; dp_i\,. \tag{2}$$

The textbook generalization of (2) to the theory of a set of fields φ^A on Minkowski space-time is,

$$-\delta H = \int\limits_{\{x^0=\mathrm{const}\}} \frac{\partial \pi_A}{\partial x^0}\; \delta\varphi^A - \frac{\partial \varphi^A}{\partial x^0}\; \delta\pi_A\,, \tag{3}$$

the symbol δ denoting a variation of fields. If the field equations arise from a Lagrange function $\mathcal{L}(\varphi^A, \varphi^A{}_{,\mu})$, then a Hamiltonian is given by the formula

$$H = \int\limits_{\{x^0=\mathrm{const}\}} \pi_A\; \frac{\partial \varphi^A}{\partial x^0} - \mathcal{L}\,, \tag{4}$$

with π_A related to the field φ^A through the equation

$$\pi_A = \frac{\partial \mathcal{L}}{\partial \varphi^A{}_{,0}}\,. \tag{5}$$

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

On every path connected component of the phase space any other Hamiltonian differs from H given by (4) by a constant.

The generalization of those standard facts to geometric field theories, due to Kijowski and Tulczyjew [14], is perhaps somewhat less familiar: here the hypersurface $\{x^0 = \text{const.}\}$ is replaced by an arbitrary hypersurface $\mathcal{S}$ in the space-time manifold $\mathcal{M}$, the field momentum π_A is replaced by a collection of momenta $\pi_A{}^\mu$ which, again for a Lagrangian theory of first order, are related to the field as

$$\pi_A{}^\mu = \frac{\partial \mathcal{L}}{\partial \varphi^A{}_{,\mu}} \,, \tag{6}$$

while the partial time derivative $\partial/\partial x^0$ is, typically,[1)] replaced in (3−4) by the Lie derivative $\mathcal{L}_X$ along the flow of some chosen vector field X. Eq. (4−5) become [14]

$$-\,\delta H = \int_{\mathcal{S}} \mathcal{L}_X \pi_A \, \delta\varphi^A - \mathcal{L}_X \varphi^A \, \delta\pi_A + \int_{\partial\mathcal{S}} X^{[\mu} \pi_A^{\nu]} \, \delta\varphi^A \, dS_{\mu\nu} \,, \tag{7}$$

$$H = \int_{\mathcal{S}} (\pi_A{}^\mu \mathcal{L}_X \varphi^A - \mathcal{L} X^\mu) \, dS_\mu \,, \tag{8}$$

and to obtain a Hamiltonian dynamical system one needs to handle the boundary terms appearing in (7), e.g. by imposing boundary conditions on the fields. Specializing to vacuum general relativity, and using as a field variable the metric density

$$\mathfrak{g}^{\mu\nu} := \frac{1}{16\pi} \sqrt{-\det g} \; g^{\mu\nu} \,, \tag{9}$$

one is then led to the following equations [12, 13] (cf. also [4, 6])

$$-\,\delta H = \int_{\mathcal{S}} (\mathcal{L}_X p^\lambda{}_{\mu\nu} \, \delta\mathfrak{g}^{\mu\nu} - \mathcal{L}_X \mathfrak{g}^{\mu\nu} \, \delta p^\lambda{}_{\mu\nu}) \, dS_\lambda + \int_{\partial\mathcal{S}} X^{[\mu} p^{\nu]}{}_{\alpha\beta} \, \delta\mathfrak{g}^{\alpha\beta} \, dS_{\mu\nu} \,, \tag{10}$$

$$H(X, \mathcal{S}) = \int_{\mathcal{S}} (p^\mu_{\alpha\beta} \mathcal{L}_X \mathfrak{g}^{\alpha\beta} - X^\mu \mathcal{L}) \, dS_\mu \,. \tag{11}$$

There is actually a problem here, related to the fact that there is no invariant Lagrangian depending upon the metric and its first derivatives only. This can be taken care of by introducing a background metric b, and removing from the usual Hilbert Lagrangian R a complete divergence:

$$\mathfrak{g}^{\mu\nu} R_{\mu\nu} - 2\Lambda \, \frac{\sqrt{-\det g_{\mu\nu}}}{16\pi} = -\partial_\alpha (\mathfrak{g}^{\mu\nu} p^\alpha_{\mu\nu}) + \mathcal{L} \,,$$

where $\mathcal{L}$ depends now upon the physical metric, its first derivatives, as well as upon the background metric and its derivatives up to order two. Here

$$p^\alpha_{\mu\nu} := (B^\alpha_{\mu\nu} - \delta^\alpha_{(\mu} B^\kappa_{\nu)\kappa}) - (\Gamma^\alpha_{\mu\nu} - \delta^\alpha_{(\mu} \Gamma^\kappa_{\nu)\kappa}) \,, \tag{12}$$

[1)] More general situations can also be considered, as described in [6].

with $B^{\alpha}_{\mu\nu}$ – the Levi-Civita connection of the metric b. Assuming X to be a Killing vector field of the background b, somewhat lengthy calculations [4, 7] lead from (11) to[2)]

$$H(X, \mathcal{S}, b) = \frac{1}{2} \int_{\partial\mathcal{S}} \mathsf{U}^{\alpha\beta}\, dS_{\alpha\beta}\,, \tag{13}$$

$$\mathsf{U}^{\nu\lambda} = \mathsf{U}^{\nu\lambda}{}_{\beta} X^{\beta} + \frac{1}{8\pi} \left(\sqrt{|\det g_{\rho\sigma}|}\, g^{\alpha[\nu} - \sqrt{|\det b_{\rho\sigma}|}\, b^{\alpha[\nu}\right) \delta^{\lambda]}_{\beta} \tilde{\nabla}_{\alpha} X^{\beta}\,, \tag{14}$$

$$\mathsf{U}^{\nu\lambda}{}_{\beta} = \frac{2\,|\det b_{\mu\nu}|}{16\pi\sqrt{|\det g_{\rho\sigma}|}}\, g_{\beta\gamma} \tilde{\nabla}_{\kappa}(e^2 g^{\gamma[\nu} g^{\lambda]\kappa})\,, \tag{15}$$

$$e = \sqrt{|\det g_{\rho\sigma}|}/\sqrt{|\det b_{\mu\nu}|}\,. \tag{16}$$

In (13) we have added b to the list of arguments of H to emphasize its potential dependence upon the background b.

2. Spacelike Hypersurfaces in Asymptotically Anti-de Sitter Space-times

We consider from now on a strictly negative cosmological constant; see [9] for an alternative Hamiltonian treatment of that case. Let $\mathcal{S}$ be an n-dimensional spacelike hypersurface in a $n+1$-dimensional Lorentzian space-time $(\mathcal{M}, g)$. Suppose that $\mathcal{M}$ contains an open set $\mathcal{U}$ which is covered by a finite number of coordinate charts (t, r, v^A), with $r \in [R, \infty)$, and with (v^A) – local coordinates on some compact $n-1$ dimensional manifold ${}^{n-1}M$, such that $\mathcal{S} \cap \mathcal{U} = \{t = 0\}$. Assume that the metric g approaches a background metric b of the form

$$b = -a^{-2}(r)\, dt^2 + a^2(r)\, dr^2 + r^2 h\,, \qquad h = h_{AB}(v^C)\, dv^A\, dv^B\,, \tag{17}$$

with $a(r) = 1/\sqrt{r^2/\ell^2 + k}$, where $k = 0, \pm 1$, h is a Riemannian Einstein metric on ${}^{n-1}M$ with Ricci scalar $n(n-1)\,k$, and ℓ is a strictly positive constant related to the cosmological constant Λ by the formula $2\Lambda = -n(n-1)/\ell^2$. For example, if h is the standard round metric on S^2 and $k = 1$, then b is the anti-de Sitter metric. It seems that the most convenient way to make the approach rates precise is to introduce an orthonormal frame for b,

$$e_0 = a(r)\, \partial_t\,, \qquad e_1 = \frac{1}{a(r)}\, \partial_r\,, \qquad e_A = \frac{1}{r}\, f_A\,, \tag{18}$$

with f_A – an h-orthonormal frame on $({}^{n-1}M, h)$, so that $b_{ab} := b(e_a, e_b) = \eta_{ab}$ – the usual Minkowski matrix diag $(-1, +1, \ldots, +1)$. We then require that the frame components g_{ab} of g with respect to the frame (18) satisfy

$$e^{ab} = O(r^{-\beta})\,, \qquad e_a(e^{bc}) = O(r^{-\beta})\,, \qquad b_{ab} e^{ab} = O(r^{-\gamma})\,, \tag{19}$$

[2)] The integral over $\partial\mathcal{S}$ should be understood by a limiting process, as the limit as R tends to infinity of integrals over the sets $t = 0$, $r = R$. $dS_{\alpha\beta}$ is defined as $\frac{\partial}{\partial x^\alpha} \lrcorner \frac{\partial}{\partial x^\beta} \lrcorner dx^0 \wedge \ldots \wedge dx^n$, with $\lrcorner$ denoting contraction; g stands for the space-time metric unless explicitly indicated otherwise. Square brackets denote antisymmetrization with an appropriate numerical factor (1/2 for two indices), and $\tilde{\nabla}$ denotes covariant differentiation *with respect to the background metric b*. The summation convention is used throughout. We use Greek indices for coordinate components and lower-case latin indices for the tetrad ones; upper-case Latin indices run from 2 to n and are associated to frames on ${}^{n-1}M$.

where $e^{ab} = g^{ab} - b^{ab}$, with

$$\beta > n/2\,, \qquad \gamma > n\,. \tag{20}$$

(The $n+1$ dimensional generalizations of the Kottler metrics (sometimes referred to as "Schwarzschild-anti de Sitter" metrics) are of the form (17) with

$$a(r) = 1/\sqrt{r^2/\ell^2 + k - 2\eta/r}$$

for a constant η, and thus satisfy (19) with $\beta = n$, and with $\gamma = 2n$.) One can check (cf. [7]) that we have the following asymptotic behaviour of the frame components of the b-Killing vector fields,

$$X^a = O(r)\,, \qquad \tilde{\nabla}_a X^b = O(r)\,.$$

Assuming that $\mathcal{L}_X p^\lambda{}_{\mu\nu}$ and $\mathcal{L}_X \mathfrak{g}^{\mu\nu}$ have the same asymptotic behaviour as $\delta p^\lambda{}_{\mu\nu}$ and $\delta \mathfrak{g}^{\mu\nu}$ (which is equivalent to requiring that the dynamics preserves the phase space), it is then easily seen that under the asymptotic conditions (19−20) the volume integrals appearing in (10−11) are convergent, the undesirable boundary integral in the variational formula (10) vanishes, so that the integrals (13) do indeed provide Hamiltonians on the space of fields satisfying (19−20). (Assuming (19−20) and $X = \partial_t$, the numerical value of the integral (13) coincides with that of an expression proposed by Abbott and Deser [1]). This singles out the charges (13) amongst various alternative expressions because Hamiltonians are uniquely defined, up to the addition of a constant, on each path connected component of the phase space. The key advantage of the Hamiltonian approach is precisely this uniqueness property, which does not seem to have a counterpart in the Noether charge analysis [15] (cf., however [11, 16]), or in Hamilton-Jacobi type arguments [3].

To define the integrals (13) we have fixed a model background metric b, as well as an orthonormal frame as in (18); this last equation required the corresponding coordinate system (t, r, v^A) as in (17). Hence, the background structure necessary for our analysis consists of a *background metric* and a *background coordinate system*. This leads to a *potential coordinate dependence* of the integrals (13): let g be any metric such that its frame components g^{ab} tend to η^{ab} as r tends to infinity, in such a way that the integrals $H(\mathcal{S}, X, b)$ given by (13) converge. Consider another coordinate system $(\hat{t}, \hat{r}, \hat{v}^A)$ with the associated background metric $\hat{b}$:

$$\hat{b} = -a^{-2}(\hat{r})\, d\hat{t}^{\,2} + a^2(\hat{r})\, d\hat{r}^2 + \hat{r}^2 \hat{h}\,, \qquad \hat{h} = h_{AB}(\hat{v}^C)\, d\hat{v}^A\, d\hat{v}^B\,,$$

together with an associated frame $\hat{e}^a$,

$$\hat{e}_0 = a(\hat{r})\, \partial_{\hat{t}}\,, \qquad \hat{e}_1 = \frac{1}{a(\hat{r})}\, \partial_{\hat{r}}\,, \qquad \hat{e}_A = \frac{1}{\hat{r}}\, \hat{f}_A\,, \tag{21}$$

and suppose that in the new hatted coordinates the integrals defining the Hamiltonians $H(\hat{\mathcal{S}}, \hat{X}, \hat{b})$ converge again. An obvious way of obtaining such coordinate systems is to make a coordinate transformation

$$t \to \hat{t} = t + \delta t\,, \qquad r \to \hat{r} = r + \delta r\,, \qquad v^A \to \hat{v}^A = v^A + \delta v^A\,, \tag{22}$$

with $(\delta t,\, \delta r,\, \delta v^A)$ decaying sufficiently fast:

$$\begin{aligned} \hat{t} &= t + O(r^{-1-\beta})\,, & e_a(\hat{t}) &= \ell\delta_a^0 + O(r^{-1-\beta})\,, \\ \hat{r} &= r + O(r^{1-\beta})\,, & e_a(\hat{r}) &= \frac{\delta_a^1}{\ell} + O(r^{1-\beta})\,, \\ \hat{v}^A &= v^A + O(r^{-1-\beta})\,, & e_a(\hat{v}^A) &= \delta_a^A + O(r^{-1-\beta})\,, \end{aligned} \tag{23}$$

and with analogous conditions on second derivatives; this guarantees that the hatted analogue of Eqs. (19) and (20) will also hold. In [7] the following is proved:

- All backgrounds satisfying the requirements above and preserving $\mathcal{S}$ (so that $\hat{t} = t$) differ from each other by a coordinate transformation of the form (23). Equivalently, coordinate transformations compatible with our fall-off conditions are compositions of (23) with an isometry of the background. (This is the most difficult part of the work in [7]).
- Under the coordinate transformations (23) the integrals (13) remain unchanged:

  $$H(\mathcal{S},\, X,\, b) = H(\hat{\mathcal{S}},\, \hat{X},\, \hat{b})\,.$$

  Here, if $X = X^\mu(t,\, r,\, v^A)\,\partial_\mu$, then the vector field $\hat{X}$ is defined using the same functions X^μ of the *hatted* variables.
- The conditions (20) are optimal[3)], in the sense that allowing $\beta = n/2$ leads to a background-dependent numerical value of the Hamiltonian.
- For some topologies of ${}^{n-1}M$, isometries of b lead to interesting, non-trivial transformation properties of the mass integrals $H(\mathcal{S},\, X,\, b)$, which have to be accounted for when defining a single number called mass. More precisely, if ${}^{n-1}M$ is negatively curved, a geometric invariant is obtained by setting

  $$m = H(\mathcal{S},\, \partial_t,\, b)\,. \tag{24}$$

If ${}^{n-1}M$ is a flat torus, then any choice of normalization of the volume of ${}^{n-1}M$ leads again to an invariant via (24). If ${}^{n-1}M = S^{n-1}$, then the group G of isometries of b preserving $\{t = 0\}$ is the Lorentz group $O(n,\, 1)$, which acts on the space $\mathcal{K}^\perp$ of b-Killing vectors normal to $\{t = 0\}$ through its usual defining representation, in particular $\mathcal{K}^\perp$ is equipped in a natural way with a G-invariant Lorentzian scalar product $\eta^{(\mu)(\nu)}$. Choosing a basis $X_{(\mu)}$ of $\mathcal{K}^\perp$ and setting

$$m_{(\mu)} = H(\mathcal{S},\, X_{(\mu)},\, b)\,, \tag{25}$$

the invariant mass is obtained by calculating the Lorentzian norm of $m_{(\mu)}$:

$$m^2 := |\eta^{(\mu)(\nu)} m_{(\mu)} m_{(\nu)}|\,. \tag{26}$$

3. The Mass of Asymptotically Hyperboloidal Riemannian Manifolds

In the asymptotically flat case the mass is an object that can be defined purely in Riemannian terms [2], i.e., without making any reference to a space-time, and this remains true in the asymptotically hyperboloidal case. The situation is somewhat more delicate here, because the transcription of the notion of a *space-time background Killing vector field* to a purely Riemannian setting requires more care. The Riemannian information carried by space-time Killing

[3)] Strictly speaking, it is the Riemannian counterpart of (20) that is optimal, see [5].

vector fields of the form $X = Ve_0$, where e_0 is a unit normal to the hypersurface $\mathcal{S}$, is encoded in the function V, which for vacuum backgrounds satisfies the set of equations

$$\Delta_b V + \lambda V = 0\,, \tag{27}$$

$$\tilde{D}_i \tilde{D}_j V = V(\mathrm{Ric}\,(b)_{ij} - \lambda b_{ij})\,, \tag{28}$$

where $\tilde{D}$ is the Levi-Civita covariant derivative of b and λ is a constant. We can forget now that $\mathcal{S}$ is a hypersurface in some space-time, and consider an n-dimensional Riemannian manifold $(\mathcal{S}, g)$ together with the set, denoted by $\mathcal{N}_b$, of solutions of (27−28); we shall assume that $\mathcal{N}_b \neq \emptyset$. If one imposes boundary conditions in the spirit of (18−20) on the Riemannian metric g, *except that the condition there on the space-time trace $b_{ab}g^{ab}$ is not needed any more*, then well defined global geometric invariants can be extracted − in a way similar to that discussed at the end of the previous section − from the integrals

$$H(V, b) := \lim_{R\to\infty} \int_{r=R} \mathsf{U}^i(V)\, dS_i\,, \tag{29}$$

where $V \in \mathcal{N}_b$ and [5]

$$\mathsf{U}^i(V) := 2\sqrt{\det g}\,\left(V g^{i[k} g^{j]l} \tilde{D}_j g_{kl} + D^{[i} V g^{j]k}(g_{jk} - b_{jk})\right)\,. \tag{30}$$

If ${}^{n-1}M$ is an $(n-1)$-dimensional sphere, and if the manifold $\mathcal{S}$ admits a spin structure, then a positive energy theorem holds [5, 8, 17, 18]; this isn't true anymore for general ${}^{n-1}M$'s, cf., e.g., [10].

References

[1] L. F. ABBOTT and S. DESER, Nucl. Phys. **B195** (1982) 76.
[2] R. BARTNIK, Comm. Pure Appl. Math. **39** (1986) 661.
[3] J. D. BROWN and J. W. YORK, Jr., Phys. Rev. **D47** (1993) 1407.
[4] P. T. CHRUŚCIEL, Ann. Inst. H. Poincaré **42** (1985) 267.
[5] P. T. CHRUŚCIEL and M. HERZLICH, *The mass of asymptotically hyperbolic Riemannian manifolds*, (2001) [dg-ga/0110035].
[6] P. T. CHRUŚCIEL, J. JEZIERSKI, and J. KIJOWSKI, *Hamiltonian field theory in the radiating regime*, Lect. Notes in Physics, vol. m70, Springer, Berlin, Heidelberg, New York, 2001, URL http://www.phys.univ-tours.fr/~piotr/papers/hamiltonian_structure.
[7] P. T. CHRUŚCIEL and G. NAGY, *The mass of spacelike hypersurfaces in asymptotically anti-de Sitter space-times*, (2001) [gr-qc/0110014].
[8] G. W. GIBBONS, S. W. HAWKING, G. T. HOROWITZ, and M. J. PERRY, Commun. Math. Phys. **88** (1983) 295.
[9] M. HENNEAUX and C. TEITELBOIM, Commun. Math Phys. **98** (1985) 391.
[10] G. T. HOROWITZ and R. C. MYERS, Phys. Rev. **D59** (1999) 026005.
[11] B. JULIA and S. SILVA, Class. Quantum Grav. **17** (2000) 4733.
[12] J. KIJOWSKI, *Unconstrained degrees of freedom of gravitational field and the positivity of gravitational energy*, Gravitation, geometry and relativistic physics (Aussois, 1984), Springer, Berlin, 1984, pp. 40−50.
[13] J. KIJOWSKI, Gen. Rel. Grav. **29** (1997) 307.
[14] J. KIJOWSKI and W. M. TULCZYJEW, *A symplectic framework for field theories*, Lecture Notes in Physics, vol. 107, Springer, New York, Heidelberg, Berlin, 1979.
[15] A. TRAUTMAN, *Conservation laws in general relativity*, Gravitation: an introduction to current research (L. Witten, ed.), Wiley, 1962.
[16] R. M. WALD and A. ZOUPAS, Phys. Rev. **D61** (2000) 084027 [gr-qc/9911095].
[17] X. WANG, *Mass for asymptotically hyperbolic manifolds*, Ph.D. thesis, Stanford, 2000.
[18] X. ZHANG, *A definition of total energy-momenta and the positive mass theorem on asymptotically hyperbolic 3 manifolds I*, (2001), preprint.

Fortschr. Phys. **50** (2002) 5–7, 630–635

The Bohr-Heisenberg Correspondence Principle Viewed from Phase Space

JENS PEDER DAHL

Chemical Physics, Department of Chemistry, Technical University of Denmark
DTU 207, DK-2800 Kgs. Lyngby, Denmark
and Abteilung für Quantenphysik, Universität Ulm, D-89069 Ulm, Germany
jpd@kemi.dtu.dk

Abstract

Phase-space representations play an increasingly important role in several branches of physics. Here, we review the author's studies of the Bohr-Heisenberg correspondence principle within the Weyl-Wigner phase-space representation. The analysis leads to refined correspondence rules that can be successfully used far away from the classical limit considered by Bohr and Heisenberg.

1. Introduction

The Bohr-Heisenberg Correspondence Principle played an important role in the setting-up of quantum mechanics. It was formulated by Niels Bohr within the framework of early quantum mechanics. It played an important role in Heisenberg's construction of the new quantum mechanics. Niels Bohr discussed the principle at length in 1918–1923 [1]. Heisenberg describes his own use of the principle in his 1929 Chicago lectures [2]. Heisenberg writes:

> *The correspondence principle, which is due to Bohr, postulates a detailed analogy between the quantum theory and the classical theory appropriate to the mental picture employed. This analogy does not merely serve as a guide to the discovery of formal laws; its special value is that it furnishes the interpretation of the laws that are found in terms of the mental pictures used.*

It is in the spirit of this comment by Heisenberg that we shall be referring to the correspondence principle below, that is to say, as a principle that allows us to use different mental pictures to view the same physical situation, with one of these mental pictures being rooted in classical mechanics. The mental picture rooted in classical mechanics will, however, be derived from quantum mechanics, using the Weyl-Wigner representation. The assertion is that this representation is ideally suited to discuss the correspondence principle.

We shall use the Weyl-Wigner representation to derive two specific results for simple systems, namely

- The quantum-mechanical transition frequency corresponds to a multiple of the classical dynamical frequency.
- The quantum-mechanical transition probability corresponds to a classical Fourier coefficient.

These results match the two central aspects of the correspondence principle as discussed by Bohr and Heisenberg. In concluding, we shall also draw attention to the fact that one

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

obtains two different mental pictures of the hydrogen-atom ground state, depending on whether one adapts the Schrödinger picture or the Weyl-Wigner picture.

2. The Weyl-Wigner Representation

The Weyl-Wigner representation is the simplest possible phase-space representation of quantum mechanics. To introduce it, it is sufficient to consider a single spinless particle, restricted to moving in one spatial dimension q, with p the corresponding momentum. The generalization to higher dimensions and several particles is straightforward. The basic relations of the Weyl-Wigner representation are the following [3−7]:

States and transitions between states are described by Wigner functions,

$$\begin{aligned} f_{ij}(q,p) &= \frac{1}{2\pi\hbar}\int dy \left\langle q+\frac{y}{2}\,\middle|\,\psi_j\right\rangle \left\langle \psi_i\,\middle|\,q-\frac{y}{2}\right\rangle e^{-ipy/\hbar} \\ &= \frac{1}{2\pi\hbar}\int dy\, \psi_i^*\left(q-\frac{y}{2}\right)\psi_j\left(q+\frac{y}{2}\right) e^{-ipy/\hbar}\,, \end{aligned} \tag{1}$$

with $\psi_i(q)$ and $\psi_j(q)$ being the normalized position-space wavefunctions associated with the states $|\psi_i\rangle$ and $|\psi_j\rangle$, respectively. Hilbert-space operators $\hat{a}$ are represented by their Weyl transforms, which are dynamical functions $a(q,p)$ in phase space,

$$a(q,p) = \int dy \left\langle q+\frac{y}{2}\right| \hat{a} \left|q-\frac{y}{2}\right\rangle e^{-ipy/\hbar}\,. \tag{2}$$

For $\hat{a} = F(\hat{q}) + G(\hat{p})$, the corresponding phase-space function is $a(q,p) = F(q) + G(p)$. For other operators, the correspondence becomes more complicated.

We now have the relations

$$\int\int dq\, dp\, f_{ij}(q,p) = \langle \psi_i \mid \psi_j\rangle \tag{3}$$

and

$$\langle \psi_i|\,\hat{a}\,|\psi_j\rangle = \int\int dq\, dp\, f_{ij}(q,p)\, a(q,p)\,. \tag{4}$$

The expectation value $\langle \psi_i|\,\hat{a}\,|\psi_i\rangle$ is, in particular, obtained by averaging the dynamical function $a(q,p)$ with the normalized distribution function $f_{ii}(q,p)$. The distribution function is always real but, in general, it takes on both positive and negative values.

3. The Dynamical Equations

In the following, we shall confine our attention to stationary states. Hence, we assume that $|\psi_i\rangle$ and $|\psi_j\rangle$ are eigenstates of the Hamiltonian $\hat{H}$ with eigenvalues E_i and E_j, respectively. We also assume that the Hamiltonian, and hence its Weyl transform, may be written

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q})\,, \qquad H(q,p) = \frac{p^2}{2m} + V(q)\,. \tag{5}$$

With $|\psi_i\rangle$ and $|\psi_j\rangle$ being stationary states, the phase-space function $f_{ij}(q, p)$ satisfies the following two coupled equations [8−10], which we denote the dynamical equations [10],

$$\cos\left[\frac{\hbar}{2}\left(\frac{\partial}{\partial q_1}\frac{\partial}{\partial p_2}-\frac{\partial}{\partial p_1}\frac{\partial}{\partial q_2}\right)\right]H(q, p)\, f_{ij}(q, p)=\frac{1}{2}\,(E_i+E_j)\, f_{ij}(q, p)\,, \tag{6}$$

$$\sin\left[\frac{\hbar}{2}\left(\frac{\partial}{\partial q_1}\frac{\partial}{\partial p_2}-\frac{\partial}{\partial p_1}\frac{\partial}{\partial q_2}\right)\right]H(q, p)\, f_{ij}(q, p)=\frac{1}{2i}\,(E_j-E_i)\, f_{ij}(q, p)\,. \tag{7}$$

Here, the subscript 1 on a differential operator indicates that this operator only acts on the first function in the product $H(q, p)\, f_{ij}(q, p)$. Similarly, the subscript 2 is used with operators that only act on the second function in the product.

The dynamical equations in phase space replace the time-independent Schrödinger equation in Hilbert space. As shown elsewhere [10−13], they may be solved to directly give the phase-space functions $f_{ij}(q, p)$. Hence, the phase-space representation may be considered a representation in its own right.

To make contact with the correspondence principle, we make a crude semiclassical approximation [11] by neglecting second and higher order terms in $\hbar$ when expanding the cosine and sine operators in the dynamical equations. With the Hamiltonian (5) this leads to the equations

$$\left[\frac{p^2}{2m}+V(q)-\frac{1}{2}\,(E_i+E_j)\right]f_{ij}(q, p)=0\,, \tag{8}$$

$$\left(-\frac{p}{m}\frac{\partial}{\partial q}+\frac{dV}{dq}\frac{\partial}{\partial p}\right)f_{ij}(q, p)=\frac{1}{i\hbar}\,(E_j-E_i)\, f_{ij}(q, p)\,. \tag{9}$$

Next, we perform a transformation to new canonical variables $\sigma = p^2/2m + V(q)$ and $\tau = \tau(q, p)$, with the new momentum σ being identical with the Hamiltonian $H(q, p)$ and the new coordinate being the dynamical time [10]. To visualize this transformation, we imagine phase space filled with a fluid of virtual particles moving in accordance with the laws of classical mechanics. $\tau(q, p)$ is then the time at which a chosen virtual particle on the trajectory through (q, p) is found at that point. The velocity of this particle is $(\dot{q}, \dot{p}) = (p/m, -dV/dq)$, where the dot implies differentiation with respect to τ. We get, therefore, that

$$\frac{\partial}{\partial\tau}=\frac{p}{m}\frac{\partial}{\partial q}-\frac{dV}{dq}\frac{\partial}{\partial p}\,. \tag{10}$$

Hence, Eqs. (8) and (9) become

$$\left[\sigma-\frac{1}{2}\,(E_i+E_j)\right]f_{ij}(\tau, \sigma)=0\,, \tag{11}$$

$$\frac{\partial}{\partial\tau}f_{ij}(\tau, \sigma)=\frac{i}{\hbar}\,(E_j-E_i)\, f_{ij}(\tau, \sigma)\,. \tag{12}$$

These equations have the solution

$$f_{ij}(\tau, \sigma)=N_{ij}\delta\left[\sigma-\frac{1}{2}\,(E_i+E_j)\right]e^{i(E_j-E_i)\tau/\hbar}\,, \tag{13}$$

where N_{ij} is a normalisation constant. To evaluate it, one notes that $dq\,dp = d\sigma\,d\tau$ since the transition to the variables σ and τ is a canonical transformation.

Thus, the Wigner function associated with a stationary state, with energy E_i, is independent of τ, and equal to a δ-function on the phase-space trajectory defined by the energy value E_i. The Wigner function associated with a transition between stationary states, with energies E_i and E_j, is restricted to the phase-space trajectory defined by the mean energy $\frac{1}{2}(E_i + E_j)$. The phase of the function varies with τ along the trajectory.

4. Refined Correspondence Principle

To proceed, we limit ourselves to bound states. This implies that the classical phase-space trajectories become closed curves, and hence that each virtual particle executes a periodic motion. Each trajectory has, accordingly, an angular frequency ω and a period $T = 2\pi/\omega$ associated with it. Adding T to the dynamical time amounts to transforming each phase-space point on the trajectory into itself. A single-valued function, defined on the trajectory, must therefore be invariant under the transformation $\tau \to \tau + T$.

Applying this condition to the function (13) gives the quantization condition

$$E_j - E_i = s\omega\hbar\,, \qquad s = \pm 1, \pm 2, \ldots, \tag{14}$$

where ω is the angular frequency associated with the trajectory for which $\sigma = \frac{1}{2}(E_i + E_j)$. This condition expresses a deep relation between classical and quantum-mechanical structure. The classical part is the frequency ω, the quantum-mechanical part the allowed energy difference $E_j - E_i$. The two parts are connected through Planck's constant. The allowed energy difference divided by $\hbar$ is also the angular frequency of the electromagnetic radiation which the system can emit or absorb. Hence, Eq. (14) states that this frequency must be a multiple of the classical frequency of motion. But this is exactly the relation between the two types of frequency that Bohr required to be fulfilled for large i and j, and for small s, when he formulated the correspondence principle. The strength of the present discussion is that it uniquely associates the frequency of motion with the phase-space trajectory defined by the mean energy $\frac{1}{2}(E_i + E_j)$. It is a refined correspondence rule.

We have now derived the first result mentioned in the introduction. To derive the second result, we begin by evaluating the normalisation constant in Eq. (13), by putting $i = j$ and integrating over σ and τ. This gives $N_{ij} = 1/T$. Hence, Eq. (13) becomes

$$f_{ij}(\tau, \sigma) = (1/T)\,\delta\left[\sigma - \frac{1}{2}\,(E_i + E_j)\right] e^{i(E_j - E_i)\tau/\hbar}\,. \tag{15}$$

Equation (4) for the transition matrix element $\langle\psi_i|\,\hat{a}\,|\psi_j\rangle$ may therefore be written

$$\langle\psi_i|\,\hat{a}\,|\psi_j\rangle = \frac{1}{T}\int_0^T d\tau\; a(\tau, \bar{\sigma})\, e^{is\omega\tau}\,, \qquad \bar{\sigma} = \frac{1}{2}\,(E_i + E_j)\,, \tag{16}$$

where $a(\tau, \bar{\sigma})$ is the Weyl transform of the operator $\hat{a}$, evaluated on the trajectory defined by the mean energy $\frac{1}{2}(E_i + E_j)$. We may express the function $a(\tau, \bar{\sigma})$ by the Fourier series

$$a(\tau, \bar{\sigma}) = \sum_{k=-\infty}^{\infty} a_k(\bar{\sigma})\, e^{-ik\omega\tau}\,. \tag{17}$$

Inserting this series into the expression (16), while observing Eq. (14), gives

$$\langle \psi_i | \, \hat{a} \, | \psi_j \rangle = a_s(\bar{\sigma}) \, , \qquad \bar{\sigma} = \frac{1}{2} \, (E_i + E_j) \, . \tag{18}$$

This, then, is the second result mentioned in the introduction.

With $\hat{a} = \hat{q}$, Eq. (18) is the condition which Bohr required to be fulfilled for large i and j, and small s. Again, our result contains a refined correspondence rule, because the trajectory defining $a_s(\bar{\sigma})$ is uniquely specified.

5. The Harmonic Oscillator

We shall now apply the rules of the preceding section to the harmonic oscillator, and hence provide a justification for talking about a *refined* correspondence principle. Our Hamiltonian and its Weyl transform have the well-known form

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2} \, m\omega^2 \hat{q}^2 \, , \qquad H = \frac{p^2}{2m} + \frac{1}{2} \, m\omega^2 q^2 \tag{19}$$

and the allowed energies are

$$E_n = \left(n + \frac{1}{2} \right) \hbar\omega \, , \qquad n = 0, 1, 2, \ldots \tag{20}$$

The classical phase-space trajectories are ellipses, and each virtual particle executes a periodic motion with the same angular frequency ω.

The energy expression (20) shows that the distance between any two energy levels is a multiple of $\hbar\omega$. Hence, the relation (14) is exactly fulfilled for all combinations of i, j and s.

The classical phase-space trajectory associated with the energy $\bar{\sigma}$ is determined by the equations

$$q = \sqrt{\frac{2\bar{\sigma}}{m\omega^2}} \, \cos \omega\tau \, , \qquad p = -\sqrt{2m\bar{\sigma}} \, \sin \omega\tau \, . \tag{21}$$

It is obvious that the only non-vanishing Fourier components of q are q_1 and q_{-1}, and likewise, that the only non-vanishing Fourier components of p are p_1 and p_{-1}. We have, in fact, that

$$q_1(\bar{\sigma}) = q_{-1}(\bar{\sigma}) = \sqrt{\frac{\bar{\sigma}}{2m\omega^2}} \, , \qquad p_1(\bar{\sigma}) = -p_{-1}(\bar{\sigma}) = i \, \sqrt{\frac{m\bar{\sigma}}{2}} \, . \tag{22}$$

The value of $\bar{\sigma}$ that goes with the matrix elements $\langle \psi_n | \, \hat{q} \, | \psi_{n+s} \rangle$ and $\langle \psi_n | \hat{p} \, | \psi_{n+s} \rangle$ is

$$\bar{\sigma} = \frac{1}{2} \, (E_n + E_{n+s}) = \left(n + \frac{1}{2} + \frac{s}{2} \right) \hbar\omega \, . \tag{23}$$

For $q_1(\bar{\sigma})$ and $p_1(\bar{\sigma})$ we must put $s = 1$, and for $q_{-1}(\bar{\sigma})$ and $p_{-1}(\bar{\sigma})$ we must put $s = -1$. The expressions (18) and (22) give accordingly

$$\langle \psi_n | \, \hat{q} \, | \psi_{n+1} \rangle = q_1(\bar{\sigma}) = \sqrt{\frac{(n+1)\,\hbar}{2m\omega}} \, , \qquad \langle \psi_n | \, \hat{q} \, | \psi_{n-1} \rangle = q_{-1}(\bar{\sigma}) = \sqrt{\frac{n\hbar}{2m\omega}} \, , \tag{24}$$

and

$$\langle \psi_n | \hat{p} | \psi_{n+1} \rangle = p_1(\bar{\sigma}) = i \sqrt{\frac{(n+1)\, m\hbar\omega}{2}} \,,$$
$$\langle \psi_n | \hat{p} | \psi_{n-1} \rangle = p_{-1}(\bar{\sigma}) = -i \sqrt{\frac{nm\hbar\omega}{2}} \,. \tag{25}$$

These expressions are identical with the exact quantum-mechanical ones. Hence, also the relation (18) is exactly fulfilled for all combinations of states, with $\hat{a}$ equal to $\hat{q}$ or $\hat{p}$.

6. Discussion

The above results show that the refined correspondence rules (14) and (18) do, in fact have a range of applicability that goes beyond the range specified by Bohr and Heisenberg. Admittedly, we have considered the most favorable example of all, namely, the harmonic oscillator and the operators $\hat{q}$ and $\hat{p}$. For other systems and other operators, minor deviations from the exact results do occur. But as an actual analysis for the Morse oscillator shows [11], the relations (14) and (18) remain surprisingly accurate.

In closing the present paper on the connection between classical mechanics and quantum mechanics, we would like to draw attention to some earlier work [14], in which we bring the fact that the angular momentum in the ground-state Bohr orbit for the hydrogen atom is non-zero, into accordance with the fact that the angular momentum in the Schrödinger picture is zero. Likewise, we would like to mention some recent work [15] where, among other things, we show that the kinetic energy in the ground state of the hydrogen atom is purely radial or purely angular, depending on whether one adapts the Schrödinger picture or the Weyl-Wigner picture.

Acknowledgments

It is a pleasure to acknowledge the support of the Danish Natural Science Research Council, the Alexander von Humboldt Foundation, and the great hospitality enjoyed at the Abteilung für Quantenphysik.

References

[1] NIELS BOHR Collected Works Vol. 3: The Correspondence Principle (1918−1923), ed. J. Rud Nielsen (North Holland 1976).
[2] W. HEISENBERG, *The Physical Principles of the Quantum Theory* (Dover 1949).
[3] E. WIGNER, Phys. Rev. **40** (1931) 749.
[4] H. WEYL, *The Theory of Groups and Quantum Mechanics* (Methuen 1931).
[5] H. J. GROENEWOLD, Physica **12** (1946) 405.
[6] J. E. MOYAL, Proc. Cambridge Phil. Soc. **45** (1949) 99.
[7] M. HILLERY, R. F. O'CONNELL, M. O. SCULLY, and E. P. WIGNER, Phys. Rep. **106** (1984) 121.
[8] D. B. FAIRLIE, Proc. Camb. Phil. Soc. **60** (1964) 581.
[9] W. KUNDT, Z. Naturforsch. a **22** (1967) 1333.
[10] J. P. DAHL, in *Energy Storage and Redistribution in Molecules*, ed. J. HINZE (Plenum 1983) 557.
[11] J. P. DAHL, in *Semiclassical Descriptions of Atomic and Nuclear Collisions*, ed. J. Bang and J. de Boer (Elsevier 1985) 379.
[12] M. HUG, C. MENKE, and W. P. SCHLEICH, Phys. Rev. A **57** (1998) 3188 and 3206.
[13] W. P. SCHLEICH, *Quantum Optics in Phase Space* (Wiley-VCH 2001).
[14] J. P. DAHL and M. SPRINGBORG, Mol. Phys. **47** (1982) 1001.
[15] J. P. DAHL and W. P. SCHLEICH, accepted Phys. Rev. A. Preprint arXiv quant-ph/0110134.

Fortschr. Phys. **50** (2002) 5–7, 636–641

The Classical Atom: Stabilization of Electronic Trojan Wavepackets

DAVID FARRELLY[1)], ERNESTINE A. LEE[2)], and T. UZER[3)]

[1)] Department of Chemistry, Utah State University, Logan, Utah 84322, USA
[2)] Incyte Genomics, 3160 Porter Drive, Palo Alto, California 94304, USA
[3)] Center for Nonlinear Sciences and School of Physics, Georgia Institute of Technology Atlanta, Georgia 30332-0430, USA

Abstract

This article demonstrates that coherent states in Rydberg atoms can be produced and stabilized by combining a circularly polarized microwave field with a static, perpendicular magnetic field. These electronic wavepackets owe their stability to atomic analogs of the Lagrange equilibria which confine Jupiter's Trojan asteroids. While these "Trojan" wavepackets may slowly decay due to tunneling, a more significant source of dispersion will arise if the tails of the wavepacket penetrate appreciably into the non-linear or chaotic parts of phase space. In the laboratory frame, if these dispersive factors can be minimized – and this may be accomplished using magnetic fields – the electronic wavepacket will travel along a circular Kepler orbit while remaining localized radially and angularly for a finite – but large – number of Kepler periods. In this sense, the system is a classical Bohr atom.

1. Introduction

The gravitational and Coulombic potentials are functionally identical and a one-electron atom is, therefore, governed by the same Hamiltonian as is the Kepler problem. The classical equations of motion for the Kepler problem are easy to solve: in the Solar System, if only the Earth and the Sun were considered, then the Earth and Sun would endlessly orbit their center of mass, never varying from an elliptical orbit. The presence of the Moon destroys the integrability of the system and the resulting motions become extremely complicated. When the number of bodies is increased beyond three – by including the other planets – the problem becomes even more of a nightmare. In fact, despite an enormous expenditure of effort, no case beyond the two-body problem has been solved completely. This complexity has had its felicitous side, however, in that many of the important developments in applied mathematics and physics, from the publication of Sir Isaac Newton's *Principia* in 1687 to Max Born's *Mechanics of the Atom* in 1925 [1], and down to today, have been fueled by progress made in studies of the n-body problem [2]. In particular, the "Old Quantum Mechanics" as described in Born's book includes many concepts and methodologies that were carried over intact from celestial mechanics. Although celestial mechanics and atomic physics temporarily parted ways after the discovery of wave mechanics the synergism between the two fields left a lasting impression: the notion of a "classical" or "Bohr" atom in which a localized electron revolves around the nucleus in a Kepler orbit continues to persist not only among members of the general public but also among atomic physicists. Indeed, the goal of making a classical atom accounts for a notable fraction of the vast body of contemporary experimental and theoretical research on Rydberg atoms [3, 4]. However, the three-body problem itself, arguably the *raison d'être* of celestial mechanics for several centuries, has received relatively less attention from atomic physicists because it does not have a direct quantum counterpart. Three quantum particles cannot

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

mutually attract one another and at the same time interact through a purely Coulombic force law. Nevertheless, in the last five years it has been demonstrated by several research groups that quantum analogs of a particular limit of the three-body problem, *the restricted three-body problem* [5], not only exist but contain dramatically new physics. Based on ideas that can be traced directly back to Schrödinger, a definition of a classical atom that is consistent with quantum mechanics can be stated as follows;

1. The wavepacket representing the electron neither spreads nor disperses as its center moves along a Kepler orbit about the nucleus.
2. The spatial extent of the packet is small compared to the radius of the orbit.
3. The classical orbit along which the center of the wavepacket moves is confined to a single plane in space.

The close analogy between a Rydberg atom in a circularly polarized microwave field and the restricted three-body problem extends to the existence of atomic stable equilibrium points that are directly analogous to the Lagrangian equilibrium points in celestial mechanics. This led Bialynicki-Birula, et al. [6] to expect that wavepackets launched from the atomic equilibrium point analogous to the stable Lagrange points L_4 and L_5 would orbit the nucleus without spreading. The term "Trojan" wavepacket is appropriate for such states. While the analogy between Rydberg atoms and planetary systems turned out to be useful, the finite size of Planck's constant imposes an absolute scale on the atomic problem which soon leads to dispersion. The atomic analogs of these points are stable only over a limited range of parameters, and placing a finite-size minimum uncertainty wavepacket at such an equilibrium point may require unrealistically large principal quantum numbers to guarantee stability. In other words linear stability is not enough to guarantee the actual stability of a particular wavepacket.

The present authors invented a different method in which a homogeneous magnetic field is used in addition to the circularly polarized microwave field to continuously stabilize the packet as it moves around the nucleus; in essence, the states are robust coherent states of an atom dressed by both microwave and magnetic fields. Examination of all of these systems in a rotating frame reveals the key role played by Coriolis forces.

2. Trojan Wavepackets

The problem of a hydrogen atom in a circularly or elliptically polarized microwave field has received considerable attention, mainly from the standpoint of how the polarization of the field affects ionization [3]. In the dipole approximation and atomic units ($\hbar = m_e = e = 1$) the Hamiltonian for a hydrogen atom subjected to a circularly polarized microwave field and a magnetic field perpendicular to the plane of polarization is

$$H = \frac{\tilde{\boldsymbol{p}}^2}{2} - \frac{1}{r} - \frac{\omega_c}{2}\,(xp_y - yp_x) + \frac{\omega_c^2}{8}\,(x^2 + y^2) + F(x\cos\omega_f t + y\sin\omega_f t)\,. \tag{1}$$

The magnetic field is taken to lie along the positive z-direction, $\omega_c = eB/m_e c$ is the cyclotron frequency, while ω_f is the microwave field frequency and F is its strength. In a frame rotating with the field frequency ω_f the Hamiltonian becomes

$$K = \frac{\tilde{\boldsymbol{p}}^2}{2} - \frac{1}{r} - \left(\omega_f + \frac{\omega_c}{2}\right)(xp_y - yp_x) + Fx + \frac{\omega_c^2}{8}\,(x^2 + y^2)\,, \tag{2}$$

where K is called the Jacobi constant in analogy with the restricted three-body problem [2] and the coordinates are now interpreted as being in the rotating frame.

The equilibria may be found by examination of curves of zero velocity used by Hill in his studies of the motions of the Moon, see Ref. [2]. These curves are obtained most easily by the simple short cut: re-write the Hamiltonian in terms of the velocities using Hamilton's equations of motion and then, since the Hamiltonian is a quadratic form in the velocities, setting them to zero which gives;

$$\Upsilon = -\frac{1}{r} + Fx - \frac{\omega_f\,(\omega_f + \omega_c)}{2}\,(x^2 + y^2)\,. \tag{3}$$

The equilibrium points of this zero-velocity surface (ZVS) are found to lie along the x axis and provides the analogy with the Lagrange equilibrium points L_4 and L_5 in the restricted-three body problem. This configuration of equilibria occurs whether or not a magnetic field is added. Bialynicki-Birula et al. [6] originally studied the case in which only a circularly polarized microwave field was used.

In scaled coordinates the Hamiltonian becomes

$$\mathcal{K} = \frac{\tilde{\boldsymbol{p}}^2}{2} - \frac{1}{r} - \,(x\,p_y - y\,p_x) + \epsilon x + \frac{\omega_s^2}{8}\,(x^2 + y^2) \tag{4}$$

which shows that the *classical* dynamics depends only on the three parameters, $\mathcal{K}$, ω_s, and ϵ. Stability of the equilibrium can now be expressed in terms of the two parameters ω_s and ϵ. For $\omega_c \neq 0$ the curve separating stable from unstable motion consists of an upper branch ($\epsilon^>$) and a lower branch ($\epsilon^<$) given by

$$\epsilon^{><} = \frac{4 - 5\,\omega_s^2 \pm \left(2 \mp \frac{1}{2}\,\omega_s^2\right)\sqrt{4 - 9\,\omega_s^2}}{2^{\frac{4}{3}}\omega_s^{\frac{2}{3}}\,(2 \pm \sqrt{4 - 9\,\omega_s^2})^{\frac{2}{3}}}\,, \tag{5}$$

where the upper (lower) sign is taken throughout for $\epsilon^>$ ($\epsilon^<$). These two functions become equal to each other exactly at $\omega_s = 2/3$ which is the rightmost point in Fig. 4. It was something of a surprise to learn that at this point the planar limit of the problem reduces to the rigorously integrable case discovered by Raković and Chu [7]

$$K = \frac{(p_x^2 + p_y^2)}{2} - \frac{1}{r} - \,(xp_y - yp_x) + (2/3)^{\frac{4}{3}}x + \frac{1}{18}\,(x^2 + y^2)\,. \tag{6}$$

2.1. Coherent states

The next step is to expand the Hamiltonian at the equilibrium point using methods developed in celestial mechanics [8] and nuclear physics [9]. Full details and an exhaustive set of references are provided in Refs. [10, 11].

A harmonic approximation to the Hamiltonian to describe librations around the equilibrium point shows the motion in the z (or ζ)-direction to be stable, harmonic, and uncoupled from the planar motion. Therefore, we ignore this degree of freedom.

After a rotation in phase space

$$\xi' = A\xi + Bp_\eta\,, \tag{7}$$

$$\eta' = A\eta + Bp_\xi\,, \tag{8}$$

$$p'_\xi = p_\xi + C\eta\,, \tag{9}$$

$$p'_\eta = p_\eta + C\xi \tag{10}$$

with $A - BC = 1$ (to preserve the commutation relations between coordinates and momenta) H can be reduced to the following locally separable form [9]

$$\mathcal{K} = \frac{1}{2m_\xi} p_\xi'^2 + \frac{1}{2} m_\xi \Omega_\xi^2 \xi'^2 + \frac{1}{2m_\eta} p_\eta'^2 + \frac{1}{2} m_\eta \Omega_\eta^2 \eta'^2 , \tag{11}$$

where we will assume (without loss of generality) that $\omega_\eta > \omega_\xi$. The locally harmonic frequencies

$$\Omega_\eta^2 = \omega^2 \left(\frac{1}{4} \omega_s{}^2 + 1 - \frac{1}{2} q - \frac{1}{2} \sqrt{9q^2 - 8q + 4\omega_s{}^2} \right) , \tag{12}$$

$$\Omega_\xi^2 = \omega^2 \left(\frac{1}{4} \omega_s{}^2 + 1 - \frac{1}{2} q + \frac{1}{2} \sqrt{9q^2 - 8q + 4\omega_s{}^2} \right) \tag{13}$$

and the "masses"

$$m_\xi = \frac{\sqrt{9q^2 - 8q + 4\omega^2{}_s}}{\left(2 + \frac{3}{2} q\right) + \frac{1}{2} \sqrt{9q^2 - 8q + 4\omega_s{}^2}} , \tag{14}$$

$$m_\eta = \frac{-\sqrt{9q^2 - 8q + 4\omega_s{}^2}}{\left(2 - \frac{3}{2} q\right) - \frac{1}{2} \sqrt{9q^2 - 8q + 4\omega_s{}^2}} \tag{15}$$

are here given in terms of the dimensionless quantity

$$q = \frac{1}{\omega^2 x_0^3} . \tag{16}$$

The "masses" are not physical masses and can be either positive or negative; there is no bound motion if both masses are negative. In the case of motion at a maximum, the masses have opposite signs. Based on the zero-velocity surface, this case occurs when $\omega^2 > \omega_\xi{}^2,\ \omega_\eta{}^2$. When both masses are positive, stable motion at a minimum is indicated. In order to cover the two possibilities for stable motion, we define the index

$$\Lambda = \frac{m_\xi m_\eta}{|m_\xi m_\eta|} \tag{17}$$

which is 1 for a minimum and -1 for a maximum.

The energy eigenvalues are given by

$$E = \frac{m_\xi}{|m_\xi|} \left(n_\xi + \frac{1}{2} \right) \hbar \, |\Omega_\xi| + \frac{m_\eta}{|m_\eta|} \left(n_\eta + \frac{1}{2} \right) \hbar \, |\Omega_\eta| . \tag{18}$$

The magnitude of the ground state energy is defined in terms of an average frequency Ω through

$$E = \frac{\hbar}{2} \left(|\Omega_\xi| + \Lambda \, |\Omega_\eta| \right) = \hbar \Omega \tag{19}$$

with this frequency explicitly being given by

$$\Omega = \frac{|\omega|}{2}\sqrt{2 - q + \frac{1}{2}\omega_s^2 + 2\Lambda s(q,\,\omega_s)}\,, \tag{20}$$

where

$$s(q,\,\omega_s) = \sqrt{\left(1 + 2q - \frac{1}{4}{\omega_s}^2\right)\left(1 - q - \frac{1}{4}\omega_s^2\right)}\,. \tag{21}$$

2.2. The initial wavepacket

The ground-state wavefunction of the (local) electronic Hamiltonian becomes

$$\Psi_{000}(\xi,\,\eta,\,\zeta) = N\psi_C(\xi,\,\eta)\,\exp\left(-\frac{\omega_\zeta}{2}\,\zeta^2\right), \tag{22}$$

where $\psi_C(\xi,\,\eta)$ is the normalized ground state wavefunction of the cranked oscillator

$$\psi_C(\xi,\,\eta) = \left(\frac{\alpha\beta}{\pi^2}\right)^{\frac{1}{4}}\exp\left(-\frac{\alpha}{2}\,\xi^2 - \frac{\beta}{2}\,\eta^2 - i\gamma\xi\eta\right). \tag{23}$$

The parameters α, β, γ are given by,

$$\alpha = \left(\frac{\omega}{3q\hbar}\right)\left(1 + 2q - \frac{1}{4}{\omega_s}^2 + \Lambda s\right)\sqrt{2 - q + \frac{1}{2}\,\omega_s^2 + 2\Lambda s(q,\,\omega_s)}\,, \tag{24}$$

$$\beta = \left(\frac{\omega}{3q\hbar}\right)\left(q - 1 + \frac{1}{4}\,{\omega_s}^2 - \Lambda s\right)\sqrt{2 - q + \frac{1}{2}\omega_s^2 + 2\Lambda s(q,\,\omega_s)}\,, \tag{25}$$

$$\gamma = \left(\frac{\omega}{3q\hbar}\right)\left(2 + q - \frac{1}{2}{\omega_s}^2 + 2\Lambda s\right). \tag{26}$$

2.3. Phase factors

The transformation to barycentric synodical coordinates requires two shifts, one in coordinate and another in momentum, to reach the equilibrium point from the center of mass. These shifts introduce all-important phase factors. The quantum mechanical consequence of these shifts can be described by a translation operator

$$\mathcal{T}(x_0) = \exp\left(-ix_0 p_\xi/\hbar\right) \tag{27}$$

and a boost-like operator

$$\mathcal{B}(\omega x_0) = \exp\left(i\omega x_0\eta/\hbar\right). \tag{28}$$

If $|C\rangle$ is the ket which is represented by $\psi_C(\xi,\,\eta)$, then the ket $|I\rangle$ that we need to use as the initial state in the barycentric coordinates is related to it by

$$|I\rangle = \mathcal{T}(x_0)\,\mathcal{B}(\omega x_0)\,|C\rangle \tag{29}$$

and therefore the wavefunctions are related by

$$\psi_C(x, y) = \left(\frac{\alpha\beta}{\pi^2}\right)^{\frac{1}{4}} \exp\left(i\nu x_0 y\right) \exp\left[-\frac{\alpha}{2}\left(x - x_0\right)^2 - \frac{\beta}{2} y^2 - i\gamma(x - x_0)\, y\right], \quad (30)$$

where $\nu = \omega/\hbar$.

Numerical integration of the quantum motion in [11] demonstrates explicitly that these wavepackets are stable to dispersion, and are localized on a scale small compared to the size of the atom.

3. Conclusions

This article has demonstrated that coherent states in Rydberg atoms can be produced by a combination of circularly polarized microwave and magnetic fields. Such states neither spread nor disperse as they execute their revolutions around the nucleus, although a Trojan wavepacket will slowly decay due to tunneling. A more significant source of dispersion will arise if the tails of the wavepacket penetrate appreciably into the non-linear or chaotic parts of phase space. In the laboratory frame, if these dispersive factors can be minimized, the electronic wavepacket will travel along a circular Kepler orbit while remaining localized radially and angularly for a finite (but possibly very large) number of Kepler periods. An important point in our study is that the stability of such a packet can be enhanced considerably by using a magnetic field in addition to the circularly polarized field. Indeed, the initial wavepacket and field choices displayed remarkable localization with the addition of a static magnetic field whereas the absence of this field can lead to rapid delocalization.

The applications of Hill's methods and the importance of Lagrange equilibria in quantum physics are not exhausted [2]. Applications have also been made to Rydberg molecules (important in zero electron-kinetic energy or ZEKE spectroscopy), to the stability of Bose-Einstein condensates in the time-averaged orbiting potential (TOP) trap, and to quantum dots [12]. Without doubt the legacies of Lagrange and Hill as well as other giants of celestial mechanics continue to influence the direction of many important areas of physics and astronomy.

References

[1] M. BORN, *Mechanics of the Atom* (republished by F. Ungar, New York, 1960). Tranlsated by J. W. Fisher.
[2] J. M. A. DANBY, *Fundamentals of Celestial Mechanics*, 2nd edition (Willman-Bell, Richmond, Virginia, 1992). pp. 384–385.
[3] T. GALLAGHER, *Rydberg Atoms* (Cambridge University Press, Cambridge, 1994).
[4] J. P. CONNERADE, *Highly Excited Atoms* (Cambridge University Press, Cambridge, 1998).
[5] V. SZEBEHELY, *Theory of Orbits: The Restricted Problem of Three Bodies* (Academic, New York, 1967).
[6] I. BIALYNICKI-BIRULA, M. KALIŃSKI, and J. H. EBERLY, Phys. Rev. Lett. **73** (1994) 1777.
[7] M. J. RAKOVIĆ and S. I. CHU, Phys. Rev. A **50** (1994) 5077.
[8] A. DEPRIT, Astron. J. **71** (1966) 77.
[9] D. GLAS, U. MOSEL, and P. G. ZINT, Z. Phys. A **285** (1978) 83.
[10] E. LEE, A.F. BRUNELLO, and D. FARRELLY, Phys. Rev. A **55** (1997) 2203.
[11] C. CERJAN, E. LEE, D. FARRELLY, and T. UZER **55** (1997) 2222.
[12] E. LEE, *Theoretical Investigations of the Electronic Dynamics of Rydberg Atoms and Molecules and Quantum Dots in the Presence of Nonconserved Paramagnetic Interactions*. Doctoral Dissertation, Utah State University, 1997 (unpublished).

Fortschr. Phys. **50** (2002) 5–7, 642–645

Heisenberg's Equations in Laser Theory. A Historical Overview

HERMANN HAKEN

Institute for Theoretical Physics, Center of Synergetics
Pfaffenwaldring 57/4, D-70550 Stuttgart, Germany
haken@theo1.physik.uni-stuttgart.de

Abstract

Heisenberg's equations were used to formulate equations of motion for the operators of the electro-magnetic field and its interaction with the electrons of atoms, where second quantization is used. In order to take into account the effect of reservoirs, and to secure quantum mechanical consistency, damping terms and stochastic forces must be added. The averaged equations yield those of semiclassical laser theory. The solution of fully quantum mechanical equations lead to my prediction that below laser threshold the light source emits uncorrelated wave tracks, while above it laser light consists of a stabilized classical amplitude, on which small quantum fluctuations and phase diffusion are superimposed.

Heisenberg's foundation of quantum mechanics is what was called at that time *matrix mechanics*. In his approach, Heisenberg assigned to physical observables, such as position and momentum, quantities that, as Born observed, had the properties of matrices. These matrices obey the famous Heisenberg equations

$$\frac{dM}{dt} = \frac{i}{\hbar}\,[H, M]\,, \tag{1}$$

where the brackets denote the commutator and H is the Hamiltonian. In the sixties of the last century, I applied his equations to laser theory, where I used the representation in which the matrices M appear as operators in the sense of second quantization including quantum electrodynamics. In the laser, a device that produces coherent light, the atoms or molecules within a cavity interact with the electromagnetic field. The electric field strength E as a function of space-coordinate x and time t can be decomposed into so-called cavity modes

$$E(x, t) = \sum_k (b_k + b_k^+)\, N_k \sin(kx)\,, \tag{2}$$

where b_k, b_k^+ are the annihilation and creation operators, respectively, of a photon of the type k. N_k is a numerical constant and k the wave vector. The Hamiltonian of the field reads

$$H_{\text{field}} = \sum_k \hbar\omega_k b_k^+ b_k\,, \tag{3}$$

where ω_k is the frequency of the cavity mode k. The Hamiltonian of the atoms can be written in the form

$$H_{\text{atoms}} = \sum_{\mu\ell} W_\ell a_{\mu\ell}^+ a_{\mu\ell}\,, \tag{4}$$

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

where $a_{\mu\ell}$, $a^+_{\mu\ell}$ are the annihilation and creation operators, respectively, of an electron in state ℓ and of atom μ. W_ℓ is the energy of that state. Using the rotating wave approximation and confining our analysis to two-level atoms, the interaction Hamiltonian can be written in the form

$$H_{\text{int}} = \hbar \sum_\mu g_\mu b^+ a^+_{1\mu} a_{2\mu} + \text{h.c.}\,, \tag{5}$$

where g_μ is proportional to the optical dipole matrix element. The application of Heisenberg's equations requires a generalization that had been considered in particular by Senitzky and Schwinger with respect to the field operator b. Namely, in the context of the laser, we have to take into account the damping of cavity modes, which in the frame of equation of the form (1) must be taken into account by a term $-\kappa b$. This term, however, violates the quantum mechanical commutation relations between the operators b and b^+. In order to remedy this deficiency, fluctuating forces F have to be added to the Heisenberg Eqs. (1). Confining our analysis to a single mode, the thus generalized Heisenberg equations read

$$\frac{d}{dt}\, b = -i\omega b + ig \sum_\mu a^+_{1\mu} a_{2\mu} - \kappa b + F(t)\,. \tag{6}$$

The annihilation and creation operators of the electrons become dynamical variables by themselves so that also Heisenberg equations must be formulated for them. In the present case, they acquire the form

$$\frac{d}{dt}\,(a^+_{1\mu} a_{2\mu}) = i(W_1 - W_2)/\hbar (a^+_{1\mu} a_{2\mu}) + igb(a^+_{2\mu} a_{2\mu} - a^+_{1\mu} a^+_{1\mu}) - \gamma a^+_{1\mu} a_{2\mu} + \Gamma_\mu(t)\,. \tag{7}$$

These equations can be interpreted as equations for the dipole moment operator that appears on the left-hand side of (7).

Finally, equations for the so-called inversion can be derived, where we indicate them only by dots

$$\frac{d}{dt}\,(a^+_{2\mu} a_{2\mu} - a^+_{1\mu} a_{1\mu}) = \ldots \tag{8}$$

When we take quantum mechanical averages over the Eqs. (6), (7), (8) with mean values

$$\langle b\rangle\,, \qquad \langle a^+_{1\mu} a_{2\mu}\rangle\,, \qquad \langle F\rangle = 0\,, \qquad \langle \Gamma\rangle = 0\,, \tag{9}$$

and factorize the expectation values appropriately, we obtain the semiclassical laser equations of Haken, Sauermann, as well as Lamb, whereby the latter started from a density matrix formulation. These equations have found wide-spread applications. They allow the determination of the laser threshold; single mode as well as multimode laser action can be treated, line shifts determined, phenomena, such as frequency locking studied, and, as I could show, they can lead to laser light chaos. The original unaveraged Heisenberg equations provide us, however, with much more information especially about quantum fluctuations. In order to extract that information, I eliminated the atomic variables a^+, a from the Heisenberg equations (including damping and fluctuating forces), which leads to equations of the form

$$\dot{b}(t) = (-i\omega - \kappa)\, b - gbb^+ b + F(t)\,. \tag{10}$$

This elimination procedure was to become the slaving principle of synergetics, but that would be another story. The Eq. (10) are equations that refer to the field operators alone and their most interesting part is given by the nonlinear term containing three operators. In order to solve these equations above laser threshold, I made the hypothesis

$$b(t) = r(t)\, T \exp\left(\int^{t} \chi(\tau)\, d\tau\right) \tag{11}$$

for the operator b, whereby T is a time-ordering operator and χ a phase operator. Making the hypothesis

$$r(t) = (r_0 + \rho) \tag{12}$$

and assuming that ρ is a small quantity, I could show that r_0 is a c-number. In other words, above laser threshold, the laser field becomes classical. When jointly with Graham I extended this approach to a continuum of laser modes, we observed the condensation into a single mode quite in analogy to Bose-Einstein condensation. We, as well as M. Scully, established a close analogy between laser light close to its threshold and a second order phase transition. By means of the operators b and b^+, I could calculate correlation functions

$$\langle b^+(t)\, b(t')\rangle \propto e^{-\Gamma(t-t')}\,, \tag{13}$$

which allow one to determine the linewidth Γ. Furthermore, in order to make contact with the Hanbury Brown-Twiss experiment, higher order correlation functions were also calculated

$$\langle b^+(t)\, b^+(t')\, b(t)\, b(t')\rangle\,. \tag{14}$$

My prediction was that the statistical behavior of laser light below and above threshold changes dramatically. Whereas below laser threshold light is composed of incoherent wave tracks that become longer and longer the more the threshold is approached, above threshold laser light behaves like an amplitude stabilized wave with a practically classical amplitude on which phase diffusion is superimposed. While (10) can be interpreted as a quantum mechanical Langevin equation, Risken interpreted it as a classical Langevin equation that allowed him to establish the corresponding Fokker-Planck equation. His solution showed again the dramatic change of laser light below and above threshold and also was able to cover the range at the laser threshold. Later this approach could be put on a solid basis by deriving quantum mechanically a generalized Fokker-Planck equation. Parallel to this approach the density matrix formalism was applied by Scully and Lamb, by Weidlich as well as others.

This is a brief sketch of the early days of the quantum theory of the laser, whereby I stressed the approach on Heisenberg's equations. In my opinion they have the advantage that one can interpret the individual terms in a straight forward manner. These equations can also be used as the starting point for quantum optics, whereby one may use an effective Hamiltonian, in which the direct interaction between the individual modes including that of the pump field are considered, but not the role of the medium. For instance, in the case of second harmonic generation, the Hamiltonian reads

$$H = 2\hbar\omega b_1^+ b_1 + \hbar\omega b_2^+ b_2 + g b_2^{+2} b_1 + g b_1^+ b_2^2\,. \tag{15}$$

It describes how the pump field with operators b_1^+, b_1 leads to a generation of signal and idler photons with index 2, whereby in the present case signal and idler coincoide.

The following references stress the contributions to laser theory by what had been called the Stuttgart School. These books contain also an outline of the approaches of the above mentioned other authors.

References

[1] H. HAKEN, *Laser Theory*, Encyclopedia of Physics, Vol. **XXV/2c**, Springer, Berlin (1970), 2nd corr. ed. (1984).

[2] H. HAKEN, *Light*, Vol. **2**, *Laser Light Dynamics*, North Holland Physics Publishing (Elsevier Science Publishers B.V.) (1985).

Fortschr. Phys. **50** (2002) 5–7, 646–651

Quantum Mechanics from a Heisenberg-type Equality

MICHAEL J. W. HALL[1)] and MARCEL REGINATTO[2)]

[1)] Theoretical Physics, RSPSE, Australian National University, Canberra ACT 0200 Australia
michael.hall@anu.edu.au
[2)] Physikalisch-Technische Bundesanstalt, Bundesallee 100, D-38116 Braunschweig, Germany

Abstract

The usual Heisenberg uncertainty relation, $\Delta X\, \Delta P \geq \hbar/2$, may be replaced by an exact equality for suitably chosen measures of position and momentum uncertainty, which is valid for all wavefunctions. This *exact* uncertainty relation, $\delta X\, \Delta P_{\rm nc} \equiv \hbar/2$, can be generalised to other pairs of conjugate observables such as photon number and phase, and is sufficiently strong to provide the basis for moving from classical mechanics to quantum mechanics. In particular, the assumption of a nonclassical momentum fluctuation, having a strength which scales inversely with uncertainty in position, leads from the classical equations of motion to the Schrödinger equation.

1. An Exact Uncertainty Relation

The well known Heisenberg inequality $\Delta X\, \Delta P \geq \hbar/2$ has a fundamental significance for the interpretation of quantum theory – as argued by Heisenberg himself, it provides "that measure of freedom from the limitations of classical concepts which is necessary for a consistent description of atomic processes" [1]. It might be asked whether this "measure of freedom" from classical concepts can be formulated more precisely. The answer is, surprisingly, yes – and, as a consequence, the Heisenberg inequality can be replaced by an exact *equality*, valid for all pure states.

To obtain this equality, note that for a classical system, the position and momentum observables can be measured simultaneously, to an arbitrary accuracy. For a quantum system we therefore define the *classical component* of the momentum to be that observable which is *closest* to the momentum observable, under the constraint of being comeasurable with the position of the system. More formally, for the case that we have maximal knowledge about the system, i.e., we know its wavefunction $\psi(x)$, the classical component $P^{\psi}_{\rm cl}$ of the momentum is defined by the properties

$$[X, P^{\psi}_{\rm cl}] = 0\,, \qquad \langle (P - P^{\psi}_{\rm cl})^2 \rangle_{\psi} = \text{minimum}\,. \tag{1}$$

The unique solution to (1) is $P^{\psi}_{\rm cl} = \int dx\, |x\rangle\, \langle x|\, P^{\psi}_{\rm cl}(x)$, where [2]

$$P^{\psi}_{\rm cl}(x) = \frac{\hbar}{2i}\, [\psi'(x)/\psi(x) - \psi'^{*}(x)/\psi^{*}(x)]\,. \tag{2}$$

Thus $P^{\psi}_{\rm cl}(x)$ provides the best possible estimate of momentum for state $\psi(x)$ consistent with position measurement result $X = x$. It may also be recognised as the momentum current which appears in the quantum continuity equation for the probability density $|\psi(x)|^2$, as well as the momentum associated with the system in the de Broglie-Bohm interpretation of quantum mechanics [3].

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Having a classical momentum component, it is natural to define the corresponding *non-classical* component of the momentum, $P^{\psi}_{\rm nc}$, via the decomposition

$$P = P^{\psi}_{\rm cl} + P^{\psi}_{\rm nc} \,. \tag{3}$$

The average of $P^{\psi}_{\rm nc}$ is zero for state $\psi(x)$, and hence P may be thought of as comprising a nonclassical fluctuation about a classical average. It is the nonclassical component which is responsible for the commutation relation $[X, P] = i\hbar$. One has the related decomposition [2]

$$(\Delta P)^2 = (\Delta P^{\psi}_{\rm cl})^2 + (\Delta P^{\psi}_{\rm nc})^2 \tag{4}$$

of the momentum variance into classical and nonclassical components, and there is a similar decomposition of the kinetic energy.

The magnitude of the nonclassical momentum fluctuation, $\Delta P^{\psi}_{\rm nc}$, provides a natural measure for that "degree of freedom from the limitations of classical concepts" referred to by Heisenberg [1]. Note that this magnitude can be operationally determined from the statistics of X and P, via Eqs. (2) and (4).

It is remarkable that $\Delta P^{\psi}_{\rm nc}$ satisfies an *exact* uncertainty relation [2]

$$\delta X \, \Delta P^{\psi}_{\rm nc} \equiv \hbar/2 \,, \tag{5}$$

where δX denotes a classical measure of position uncertainty, called the "Fisher length", defined via [4]

$$(\delta X)^{-2} = \int_{-\infty}^{\infty} dx \; p(x) \, [(d/dx) \ln p(x)]^2$$

for probability density $p(x)$. There is thus a *precise* connection between the statistics of complementary observables.

The exact uncertainty relation (5) is far stronger than the usual Heisenberg inequality (the latter follows immediately via (4) and the Cramer-Rao inequality $\Delta X \geq \delta X$ of classical statistical estimation theory). For example, suppose that at some time the wavefunction $\psi(x)$ is confined to some interval. Then, since $\ln p(x)$ changes from $-\infty$ to a finite value in any neighbourhood containing an endpoint of the interval, the Fisher length δX vanishes. The exact uncertainty relation (5) thus immediately implies that $\Delta P^{\psi}_{\rm nc}$, and hence ΔP, is unbounded for any such confined wavefunction.

Exact uncertainty relations may be generalised and/or applied to, for example, density operators, higher dimensions, energy bounds, photon number and phase, and entanglement [2]. A conjugate relation, $\delta P \, \Delta X^{\psi}_{\rm nc} \equiv \hbar/2$, may also be derived. In the following section the very *existence* of exact uncertainty relations is used as a basis for deriving the Schrödinger equation.

2. QM from an Exact Uncertainty Principle

Landau and Lifschitz wrote, referring to the Heisenberg uncertainty principle, that "this principle in itself does not suffice as a basis on which to construct a new mechanics of particles" [5]. However, the existence of exact uncertainty relations for quantum systems raises anew the question of whether the uncertainty principle, at the conceptual core of the standard Copenhagen interpretation of quantum theory, can be put in a form strong enough to provide an axiomatic means for moving from classical to quantum equations of motion. The corresponding *exact* uncertainty principle would thus be on a par with alternative derivations based on the principle of superposition, C^*-algebras, quantum logics, etc.

It has recently been shown that an "exact uncertainty principle" does indeed exist, where the Schrödinger equation may be derived via the postulate that classical systems are subjected to random momentum fluctuations of a strength inversely proportional to uncertainty in position [6]. This new approach is summarised below (though using a different method of proof than in [6]).

Now, in any axiomatic-type construction of quantum mechanics one must first choose a classical starting point, to be generalised or modified appropriately. The starting point here is a statistical one – the classical motion of an ensemble of particles – and indeed most of the assumptions to be made below will be seen to have a statistical character.

Consider then a classical ensemble of n-dimensional particles of mass m moving under a potential V. The motion may be described via the Hamilton-Jacobi and continuity equations

$$\frac{\partial s}{\partial t} + \frac{1}{2m}\,|\nabla s|^2 + V = 0\,, \qquad \frac{\partial p}{\partial t} + \nabla\cdot\left[p\,\frac{\nabla s}{m}\right] = 0\,,$$

respectively, for the "momentum potential" s and the position probability density p. These equations follow from the variational principle $\delta L = 0$ with Lagrangian

$$L = \int dt\, d^n x p \left[\frac{\partial s}{\partial t} + \frac{1}{2m}\,|\nabla s|^2 + V\right], \tag{6}$$

under fixed endpoint variation with respect to p and s. This Lagrangian is therefore chosen as our classical starting point.

It is now assumed that the classical Lagrangian (6) must be modified, due to the existence of random momentum fluctations. The nature of these fluctuations is not important to the argument – they may be postulated to model experimental evidence that the momentum variance is sometimes greater than the expected $\int dx\, p\,|\nabla s - \langle\nabla s\rangle|^2$; or they may be regarded as a device to make the system irreducibly statistical (since such fluctuations imply the velocity relation $\boldsymbol{v} = m^{-1}\nabla s$ no longer holds, and hence cannot be integrated to give corresponding trajectories). The assumption is simply that *the momentum associated with position x is given by*

$$\boldsymbol{P} = \nabla s + \boldsymbol{N}\,,$$

where the fluctuation term $\boldsymbol{N}$ vanishes on the average at each point x. The physical meaning of s thus changes to being an *average* momentum potential.

It follows that the average kinetic energy $\langle|\nabla s|^2\rangle/(2m)$ appearing in (6) should be replaced by $\langle|\nabla s + \boldsymbol{N}|^2\rangle/(2m)$ (where $\langle\,\rangle$ now denotes the average over fluctuations *and* position), giving rise to the modified Lagrangian

$$L' = L + (2m)^{-1}\int dt\,\langle\boldsymbol{N}\cdot\boldsymbol{N}\rangle = L + (2m)^{-1}\int dt\,(\Delta N)^2\,, \tag{7}$$

where $\Delta N = \langle\boldsymbol{N}\cdot\boldsymbol{N}\rangle^{1/2}$ is a measure of the strength of the fluctuations. Note that the additional term in (7) corresponds to the kinetic energy of the fluctuations, and so is positive.

The additional term is specified uniquely, up to a multiplicative constant, by the following three assumptions:

(1) Action principle: L' is a scalar Lagrangian with respect to the fields p and s, where the variational principle $\delta L' = 0$ yields causal equations of motion. Thus

$$(\Delta N)^2 = \int d^n x\, p f(p,\, \nabla p,\, \partial p/\partial t,\, s,\, \nabla s,\, \partial s/\partial t,\, \boldsymbol{x},\, t)$$

for some scalar function f.

(2) Additivity: If the system comprises two independent non-interacting subsystems 1 and 2, with $p = p_1 p_2$, then the Lagrangian decomposes into additive subsystem contributions. Thus

$$f = f_1 + f_2 \quad \text{for} \quad p = p_1 p_2 \,,$$

where subscripts denote corresponding subsystem quantities. Note this is equivalent to the statistical assumption that the corresponding subsystem fluctuations are linearly uncorrelated, i.e., $\langle \boldsymbol{N}_1 \cdot \boldsymbol{N}_2 \rangle = 0$.

(3) Exact uncertainty principle: The strength of the momentum fluctuation at any given time is determined by, and scales inversely with, the uncertainty in position at that time. Thus

$$\Delta N \to k\, \Delta N \quad \text{for} \quad \boldsymbol{x} \to \boldsymbol{x}/k \,,$$

and moreover, since position uncertainty is entirely characterised by the probability density p at a given time, the function f cannot depend on s, nor explicitly on t, nor on the time-derivative of p.

We now have the following theorem [6]:

Theorem: *The above three assumptions of an action principle, additivity, and an exact uncertainty principle imply that*

$$(\Delta N)^2 = C \int d^n x\, p\, |\nabla \ln p|^2 \,,$$

where C is a positive universal constant.

A (new) proof of the above theorem is outlined below. Here its main consequence is noted. In particular, it follows directly via (7) that the equations of motion for p and s, corresponding to the variational principle $\delta L' = 0$, can be expressed as the single complex equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V \psi \,,$$

where one defines $\hbar := 2\sqrt{C}$ and $\psi := \sqrt{p}\, e^{is/\hbar}$. Thus the above postulates yield equations of motion equivalent to the Schrödinger equation. It can be shown that the mapping from fields p and s to the wavefunction ψ arises naturally from seeking canonical transformations which map to "normal modes" of the system [6]. It is remarkable that a linear equation results from the assumptions used.

3. Proof of Theorem

To see how the above theorem follows, note first that the action principle and exact uncertainty principle imply that the scalar function f depends only on p, ∇p, and $\boldsymbol{x}$. It can therefore be written in the form

$$f = g(u, v, w, r^2) \,,$$

where

$$u = \ln p \,, \qquad v = (\boldsymbol{x} \cdot \nabla p)/p \,, \qquad w = |\nabla p|^2/p^2 \,, \qquad r^2 = \boldsymbol{x} \cdot \boldsymbol{x} \,.$$

For $p = p_1 p_2$ one finds $u = u_1 + u_2$, $v = v_1 + v_2$, and $w = w_1 + w_2$, and hence the additivity assumption implies that

$$g(u_1 + u_2,\, v_1 + v_2,\, w_1 + w_2,\, r_1^2 + r_2^2) = g(u_1,\, v_1,\, w_1,\, r_1^2) + g(u_2,\, v_2,\, w_2,\, r_2^2)\,.$$

Thus g must be linear in u, v, w and r^2, yielding

$$f = A \ln p + B\,\frac{\boldsymbol{x}\cdot\nabla p}{p} + C\,\frac{|\nabla p|^2}{p^2} + D\boldsymbol{x}\cdot\boldsymbol{x}\,, \tag{8}$$

where A, B, C and D are universal constants. Note that the first term corresponds to an entropic potential in the Lagrangian L'.

Finally, noting that $p(x) \to k^n p(kx)$ under the transformation $\boldsymbol{x} \to \boldsymbol{x}/k$, the exact uncertainty principle forces $A = B = D = 0$, and the theorem is proved (where the positivity of C follows from the positivity of $(\Delta N)^2$).

4. Conclusions

It has been shown that the Heisenberg uncertainty relation may be upgraded to an exact uncertainty relation, and that a corresponding *exact* uncertainty principle may be used as the single nonclassical element necessary for obtaining the Schrödinger equation. Thus the uncertainty principle is able to provide not only a conceptual underpinning of quantum mechanics, but an axiomatic underpinning as well. The above approach immediately generalises to include electromagnetic potentials, and work on generalisations to systems with spin and to quantum fields is in progress.

The form chosen for the exact uncertainty principle, that classical systems are subject to random momentum fluctuations of a strength inversely proportional to uncertainty in position, is of course motivated by the momentum decomposition (3) and exact uncertainty relation (5) holding for quantum systems. Thus, not surprisingly, there are a number of connections between the latter uncertainty relation and the Theorem of Section 2. For example, for a 1-dimensional system, the Theorem immediately implies the uncertainty relation

$$\delta X\, \Delta N = \sqrt{C} = \hbar/2$$

for the momentum fluctations, which may be compared to (5).

The exact uncertainty principle has a type of "nonlocality" built into it: the form of ΔN specified by the Theorem implies that a change in the position probability density arising from actions on one subsystem (e.g., a position measurement), will typically influence the behaviour of a second subsystem correlated with the first. This nonlocality corresponds to quantum entanglement, and has been analysed to some extent via exact uncertainty relations in [2].

It is worth noting that the approach here, based on exact uncertainty, is rather different from other approaches which assign physical meaning to fields p and s related to the wavefunction. For example, in the de Broglie-Bohm approach [3], there are *no* momentum fluctuations, and the classical equations of motion for p and s are instead modified by adding a mass-dependent "quantum potential", Q, to the classical potential term in the Hamilton-Jacobi equation. The form of this quantum potential is left unexplained, and is interpreted as arising from the influence of an associated wave acting on the system. In contrast, in the exact uncertainty approach ∇s is an *average* momentum, the form of an additional kinetic energy term arising from random momentum fluctuations is *derived*, and no associated

wave is assumed. The formal connection between the two approaches is the relation

$$\delta(L' - L) = \int dt\, d^n x Q\, \delta p\,.$$

Finally, the exact uncertainty approach is also very different from the stochastic mechanics approach [7]. The latter postulates the existence of a classical stochastic process in configuration space, with a drift velocity assumed to be the gradient of some scalar, and defines an associated time-symmetric "mean acceleration" $\boldsymbol{a}$ in terms of averages over both the stochastic process and a corresponding time-reversed process, which is postulated to obey Newton's law $m\boldsymbol{a} = -\nabla V$. In contrast, the exact uncertainty approach does not rely on a classical model of fluctuations, nor on a new definition of acceleration, nor on properties of stochastic processes running backwards in time. The formal connections between the approaches are

$$\nabla s = m\boldsymbol{u}\,, \qquad (\Delta N)^2 = m^2 \langle \boldsymbol{v} \cdot \boldsymbol{v} \rangle\,,$$

where $\boldsymbol{u} + \boldsymbol{v}$ and $\boldsymbol{u} - \boldsymbol{v}$ are the drift velocities of the forward-in-time and backward-in-time processes respectively. It should be noted that $\langle \boldsymbol{u} \cdot \boldsymbol{v} \rangle \neq 0$, and hence one cannot identify $m\boldsymbol{v}$ with the random momentum fluctuation $\boldsymbol{N}$.

References

[1] W. HEISENBERG, *The Physical Principles of the Quantum Theory* (Dover, USA, 1930), page 4.
[2] M. J. W. HALL, Phys. Rev. A **64** (2001) 052103 [quant-ph/0107149].
[3] D. BOHM, Phys. Rev. **85** (1952) 143, 187.
[4] R. A. FISHER, Proc. Cambridge Philos. Soc. **22** (1925) 700.
[5] L. D. LANDAU and E. M. LIFSCHITZ, *Quantum Mechanics*, 3rd edition (Pergamon, Oxford, 1977), page 2.
[6] M. J. W. HALL and M. REGINATTO [quant-ph/0102069].
[7] E. NELSON, Phys. Rev. **150** (1966) 1079;
E. NELSON, *Dynamical Theories of Brownian Motion* (Princeton University Press, USA, 1967).

Fortschr. Phys. **50** (2002) 5–7, 652–656

Entanglers: Beam Splitters and Thermal Fields

MYUNG SHIK KIM

Sir David Bates Building, School of Mathematics and Physics
Queen's University, Belfast, BT7 1NN, United Kingdom
m.s.kim@qub.ac.uk

Abstract

Entanglement is an important ingredient for quantum information processing. We discuss some sources of entanglement, namely a beam splitter and a thermal field. For the generation of entangled continuous-variable states, we consider a beam splitter and find some conditions for input fields to see entanglement in the output. While a beam splitter is a unitary device to generate an entangled state for a bipartite continuous-variable system, a thermal field is shown to mediate entanglement of two qubits.

1. Introduction

An odd aspect of quantum mechanics is contained in the Heisenberg uncertainty principle [1]. The trial of the uncertainty principle for perfectly correlated particles made Einstein, Podolsky and Rosen in doubt about quantum mechanics [2]. A pair of quantum-mechanically correlated particles are said to be in an entangled state. Mathematically, a bipartite system in an entangled state is not separable, in other words, the system is not possible to be represented by a convex sum of direct products:

$$\hat{\rho} \neq \sum P_i \hat{\rho}_1(i) \otimes \hat{\rho}_2(i)\,, \tag{1}$$

where $\hat{\rho}$, $\hat{\rho}_1$ and $\hat{\rho}_2$ are density operators for the total system and the composite systems 1 and 2, respectively. It has been shown that the positivity of the partial transposition of a density operator is necessary and sufficient for its separability [3, 4].

In quantum algorithm, an entangled state is generated from two initial qubits in $|0\rangle\,|0\rangle$ by a unitary operation composed of a controlled-NOT and a Hadamard operation. After a qubit (control bit) passes through the Hadarmad gate, the control bit becomes in $(|0\rangle - |1\rangle)/\sqrt{2}$, the controlled-NOT operation then flips the target bit only when the control bit is in $|1\rangle$. Thus the two qubits turn into

$$|\Phi_-\rangle = \frac{1}{\sqrt{2}}\,(|0\rangle\,|0\rangle - |1\rangle\,|1\rangle)\,. \tag{2}$$

Experimentally, Type-II or type-I [5] parametric down-conversion may produce polarisation entangled states. By putting one photon state into a beam splitter, an entangled state is generated in the two output ports [6]. The beam splitter may also be used to produce a continuous-variable entangled state [7]. By a non-degenerate optical parametric amplifier a two-mode squeezed state, which is a continuous-variable entangled state, is generated [8]. Entanglement transformation at absorbing and amplifying four-port devices was studied [9].

In this paper, we consider a beam splitter as a linear unitary device to produce entangled Gaussian continuous-variable states. We are also interested in entangling two qubits via

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

mutual unitary interaction with a thermal field which is a chaotic field defined in infinite-dimensional Hilbert space.

2. Beam Splitter as an Entangler

We denote two input modes of a lossless beam splitter by a and b and two output modes by c and d. The output mode c is due to transmission of mode a and reflection of mode b. The output-field annihilation operators are given by

$$\hat{c} = \hat{B}\hat{a}\hat{B}^\dagger\,, \qquad \hat{d} = \hat{B}\hat{b}\hat{B}^\dagger \tag{3}$$

where the beam splitter operator is [10]

$$\hat{B}(\theta,\,\phi) = \exp\left[\frac{\theta}{2}\,(\hat{a}^\dagger\hat{b}\,e^{i\phi} - \hat{a}\hat{b}^\dagger\,e^{-i\phi})\right] \tag{4}$$

with the amplitude reflection and transmission coefficients $t = \cos(\theta/2)$, $r = \sin(\theta/2)$. The beam splitter gives the phase difference ϕ between the reflected and transmitted fields.

In this paper, we are interested in the action of a beam splitter for Gaussian fields [11]. A general pure Gaussian field is a squeezed coherent state, $\hat{S}(\zeta)\,\hat{D}(\alpha)\,|0\rangle$ where $\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$ is the displacement operator and the squeezing operator

$$\hat{S}(s) = \exp\left(\frac{s}{2}\,\hat{a}^2 - \frac{s}{2}\,\hat{a}^{\dagger 2}\right) \tag{5}$$

with the real squeezing parameter. It is straightforward to show that displacing the input fields does not change entanglement of the output fields [7]. It can also be shown that, without losing generality, we take the input squeezing parameter to be real while keeping ϕ of the beam splitter variable because the relative phase ϕ gives the effect of the rotation of the squeezing angle for the input fields.

Consider the output field from a 50:50 beam splitter for two squeezed input fields:

$$\begin{aligned}\hat{B}(\pi/4,\,\phi)\,\hat{S}_a(s_1)\,\hat{S}_b(s_2)\,|0,\,0\rangle = {} & \hat{S}_a\left(\frac{1}{2}\,(s_1 + s_2\,e^{2i\phi})\right)\hat{S}_b\left(\frac{1}{2}\,(s_1\,e^{-2i\phi} + s_2)\right) \\ & \times\hat{S}_{ab}\left(\frac{1}{2}\,(s_1\,e^{i\phi} - s_2\,e^{-i\phi})\right)|0,\,0\rangle\,,\end{aligned} \tag{6}$$

where $\hat{S}_{ab}(\zeta) = \exp(-\zeta\hat{a}\hat{b} + \zeta^*\hat{a}^\dagger\hat{b}^\dagger)$ is the two-mode squeezing operator. The single-mode squeezing operators $\hat{S}_a$ and $\hat{S}_b$ in the right-hand side of Eq. (6) do not contribute toward entanglement of the output state because they can be cancelled by local unitary operations. Thus only the two-mode squeezing operator $\hat{S}_{ab}$ determines the entanglement of the output state as only it represents a joint action on both pairs of the bipartite system. For a given squeezing, s_1 and s_2, when $\phi = \pi/2$, the output state is maximally entangled. When $\phi = 0$, entanglement is minimized. In fact, if $s_1 = s_2$ we completely lose entanglement for $\phi = 0$. We notice that *a two-mode squeezed state is produced from a single-mode squeezed state by an action of a beam splitter and local unitary operations* [12].

When the input fields are mixed, the output fields from a beam splitter are also mixed. A general mixed continuous-variable state is not easy to deal with because of its complicated nature. However, for a Gaussian two-mode state, the separability condition has been studied extensively [13−15]. The separability of a Gaussian state is discussed with quasi-probability

functions and their characteristic functions in phase space. It has been shown that if a two-mode Gaussian state is represented by a positive well-behaved P-function $P(\alpha, \beta)$, the state is separable [14, 15]. It is straightforward to prove *a sufficient condition* for separability of the output state from a beam splitter: when two *classical* Gaussian input fields are incident on a beam splitter, the output state is always separable [7]. It follows that for creating a Gaussian entangled state with a help of a beam splitter, it is necessary that the input exhibits nonclassical behavior.

We have already seen that two nonclassical input fields do not necessarily bring about entanglement in the output state as two squeezed state inputs may not give entangled outputs in the beam splitter. We investigate the entanglement of the output state when two Gaussian mixed states are incident on a beam splitter. The necessary and sufficient criterion for the separability of a Gaussian mixed state has been studied using the Weyl characteristic function $C^{(w)}(\zeta, \eta)$ [13–15].

Consider two thermal states of the same average photon number $\bar{n}$. The squeezed thermal field is represented by the following characteristic function:

$$C^{(w)}(\zeta) = \exp\left[-\frac{1}{2}\,(2\bar{n}+1)\,e^{2s}\,\zeta_r^2 - \frac{1}{2}\,(2\bar{n}+1)\,e^{-2s}\,\zeta_i^2\right]. \tag{7}$$

The squeezed thermal state is said to be nonclassical when one of the quadrature variables has its variance smaller than the vacuum limit; the squeezed thermal state of (7) is nonclassical when $(2\bar{n}+1)\,e^{-2s} - 1 < 0$. Here, $s > 0$ is assumed without loss of generality. For the maximum entanglement of the squeezed input, we consider a 50:50 beam splitter, in which case we find that the output state is entangled when the two identical input squeezed thermal fields become nonclassical.

As another example, we take a squeezed thermal state incident on one input port and vacuum on to the other port. After a little algebra, we confirm our earlier finding that the nonclassicality of the input state provides the entanglement criterion for the output state. When a squeezed thermal state and vacuum are incident on a beam splitter, the output state is entangled only if the squeezed thermal state is nonclassical.

Differently from the earlier cases, let us consider that one of the input states is always nonclassical while the other is always classical: one input is the squeezed vacuum and the other is the thermal state. In this case, it can be shown that the separability criterion coincides with the nonclassicality condition for one of the output field.

3. Thermal Field as an Entangler

A thermal field, which is emitted by a source in thermal equilibrium, is the field about which we have minimal information. It arises frequently in problems involving the coupling of a system to its environment which is in thermodynamic equilibrium. Recently, Bose et al. [16] showed that entanglement can always arise in the interaction of a single qubit in a pure state with an arbitrarily large system in any mixed state and illustrated this using a model of the interaction of a two-level atom with a thermal field. Using this model, they studied the possibility of entangling a qubit with a large system defined in an infinite dimensional Hilbert space. The entanglement between the system and the thermal field reduces the system to a mixed state when the field variables are traced over. Here, we investigate a possibility for a thermal field to induce entanglement between qubits [17].

We consider a quantum register composed of two two-level atoms interacting with a single-mode thermal field. We will investigate the entanglement between the two atoms, which are initially separable. This simple interaction model of a quantum register with its environment allows us to understand how the memory of the environment affects the state of a quantum register.

When a quantum system of two qubits prepared in $\hat{\rho}_s$ interacts with an environment represented by the density operator $\hat{\rho}_E$, the system and environment evolve for a finite time, governed by the unitary time evolution operator $\hat{U}(t)$. The density operator for the system and environment at time t is

$$\hat{\rho}(t) = \hat{U}(t)\,(\hat{\rho}_E \otimes \hat{\rho}_s)\,\hat{U}^\dagger(t)\,. \tag{8}$$

After performing a partial trace over environment variables, we find the final density matrix $\hat{\rho}_s(t)$ of the quantum system in the following Kraus representation

$$\hat{\rho}_s(t) = \mathrm{Tr}_E\,\hat{\rho}(t) = \sum_\mu \hat{K}_\mu \hat{\rho}_s \hat{K}_\mu^\dagger\,, \tag{9}$$

where the Kraus operators $\hat{K}_\mu$'s satisfy the property: $\sum_\mu \hat{K}_\mu^\dagger \hat{K}_\mu = 1_s$. Unitary evolution of the quantum system is a special case in which there is only one non-zero term in the operator sum (9). If there are two or more terms, the pure initial state becomes mixed. For the mixed thermal environment, not only entanglement but also classical correlation of the system and environment can make the system evolve from the initial pure state into a mixed one. The mutual information between the system and the environment is non-zero if the information for the total system is not the same as the total sum of information for each subsystem. A non-zero mutual information is due to classical correlation and/or entanglement [18]. Thus there being non-zero mutual information between the system and environment is a sufficient condition for the initial pure system to have evolved into a mixed state. The action of $\hat{K}_\mu$ projects the system into a pure state of $\hat{K}_\mu \hat{\rho} \hat{K}_\mu^\dagger$ when the initial state is pure.

For simplicity, we consider the identical two-level atoms 1 and 2 which are coupled to a single-mode thermal field with the same coupling constant γ. The ground and excited states for the atom i $(i = 1, 2)$ are, respectively, denoted by $|g\rangle_i$ and $|e\rangle_i$. The cavity mode is assumed to be resonant with the atomic transition frequency. Under the rotating wave approximation, the Hamiltonian in the interaction picture is

$$\hat{H}_I = \hbar\gamma \sum_{i=1,2} (\hat{a}\hat{\sigma}_i^+ + \hat{a}^\dagger \hat{\sigma}_i^-)\,, \tag{10}$$

where the atomic transition operators are $\hat{\sigma}_i^- = |g\rangle_i \langle e|$ and $\hat{\sigma}_i^+ = |e\rangle_i \langle g|$.

The density operator for the combined atom-field system follows a unitary time evolution generated by the evolution operator, $\hat{U}(t) = \exp(-i\hat{H}t/\hbar)$ (see [19]). For a two qubit system described by the density operator $\hat{\rho}$, a measure of entanglement can be defined in terms of the negative eigenvalues of the partial transposition [18]:

$$\mathcal{E} = -2 \sum_i \mu_i^-\,, \tag{11}$$

where μ_i^- are the negative eigenvalues of the partial transposition of $\hat{\rho}$. When $\mathcal{E} = 0$ the two qubits are separable [3, 4] and $\mathcal{E} = 1$ indicates maximum entanglement between them.

If the atoms are initially in their excited states, we find

$$\hat{\rho}_s(t) = A_e\,|ee\rangle\,\langle ee| + 2A_s\,|s\rangle\,\langle s| + A_g\,|gg\rangle\,\langle gg|\,, \tag{12}$$

where the coefficients A_e, A_s and A_g depend on the interaction time and the initial temperature of the thermal field. After a little algebra, it is found that the atoms are never entangled during the evolution when the atoms are initially in their excited states. If the atoms are initially in their ground states, the time-dependent density operator for the two atoms appears in the same form as Eq. (12) but with different coefficients A_e, A_s and A_g. In this

case, the value of A_s is relatively larger so that we see entanglement between two atoms at some times of their mutual interaction with the thermal field. More interestingly, we find that atoms are always entangled by a thermal field when one of the atoms is prepared in its excited state and the other in the ground state.

4. Remarks

We have demonstrated that two atoms can become entangled through their interaction with a highly chaotic system depending on the initial preparation of the atoms. When a beam splitter is used to entangle two Gaussian continuous-variable fields, we have proved that if both the Gaussian input fields are classical, it is not possible to create entanglement in the output of the beam splitter. From here it automatically follows that nonclassicality is a necessary condition for the entanglement. That is the nonclassicality of individual inputs can be traded for quantum entanglement of the output of the beam splitter.

Acknowledgment

We thank Profs. Knight and Bužek, Dr. Lee and Mr. Son for stimulating discussions. We acknowledge for support of the UK Engineering and Physical Sciences Research Council through the grant GR/R33304.

References

[1] W. Heisenberg, Z. Phys. **43** (1927) 172.
[2] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47** (1935) 777.
[3] A. Peres, Phys. Rev. Lett. **77** (1996) 1413.
[4] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. **A223** (1996) 1.
[5] K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger, Phys. Rev. Lett. **76** (1996) 4656;
A. G. White, D. F. V. James, P. H. Eberhard, and P. G. Kwiat, Phys. Rev. Lett. **83** (1999) 3103.
[6] S. M. Tan, D. F. Walls, and M. J. Collett, Phys. Rev. Lett. **66** (1991) 252.
[7] M. S. Kim, W. Son, V. Bužek, and P. L. Knight, LANL e-print, quant-ph/010636 (2001)
[8] S. Braunstein and H. J. Kimble, Phys. Rev. Lett. **80** (1998) 869;
A. Furusawa, J. L. Serensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik, Science **282** (1999) 706.
[9] S. Scheel, L. Knöll, T. Opartný, and D.-G. Welsch, Phys. Rev. A **62** (2000) 043803.
[10] R. A. Campos, B. E. A. Saleh, and M. C. Teich, Phys. Rev. A **40** (1989) 1371.
[11] G. S. Agarwal, J. Mod. Opt. **34** (1987) 909;
H. Huang and G. S. Agarwal, Phys. Rev. A **49** (1994) 52.
[12] M. S. Kim and B. C. Sanders, Phys. Rev. A **53** (1996) 3694;
S. Scheel and D.-G. Welsch, Phys. Rev. A **64** (2001) 063811;
Ch. Silberhorn, P. K. Lam, O. Weiss, F. König, N. Korolkova, and G. Leuchs, Phys. Rev. Lett. **86** (2001) 4267.
[13] R. Simon, Phys. Rev. Lett. **84** (2000) 2726.
[14] L.-M. Duan, G. Riedke, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **84** (2000) 2722.
[15] J. Lee, M. S. Kim, and H. Jeong, Phys. Rev. A **62** (2000) 032305.
[16] S. Bose, I. Fuentes-Guridi, P. L. Knight, and V. Vedral, Phys. Rev. Lett. **87** (2001) 050401.
[17] M. S. Kim, J. Lee, D. Ahn, and P. L. Knight, LANL e-print, quant-ph/0109052 (2001)
[18] J. Lee and M. S. Kim, Phys. Rev. Lett. **84** (2000) 4236;
J. Lee, M. S. Kim, Y. J. Park, and S. Lee, J. Mod. Opt. **47** (2000) 2151.
[19] M. S. Kim and G. S. Agarwal, J. Mod. Opt. **46** (1999) 2111.

Fortschr. Phys. **50** (2002) 5–7, 657–663

Enhancing Otto-mobile Efficiency via Addition of a Quantum Carnot Cycle

TOMÁŠ OPATRNÝ and MARLAN O. SCULLY

Department of Physics, Texas A & M University, College Station, TX 77843, USA

Abstract

It was shown recently that one can improve the efficiency of the Otto cycle by taking advantage of the internal degrees of freedom of an ideal gas [M. O. Scully, "The Quantum Afterburner", Phys. Rev. Lett., to be published]. Here we discuss the limiting improvement of the efficiency by considering reversible cycles with both internal and external degrees of freedom.

1. Introduction

Quantum thermodynamics, a field opened by Planck, lies at the heart of the old quantum theory. The new quantum mechanics, i.e., matrix and wave mechanics as developed by Heisenberg and Schrödinger, is currently being applied to thermodynamics. This new take on quantum thermodynamics is yielding interesting new insights. For example, the quantum heat engine idea by Scovil and Schulz-DuBois [1] and various extensions of this work [2] have contributed to better understanding of the thermodynamic concepts from the quantum mechanical point of view. A recent example is the demonstration of the Otto cycle efficiency improvement by extracting useful laser work from the internal degrees of freedom [3, 4]. Here we discuss the limiting case of such an improvement, illustrating it by a simple model as in Fig. 1.

The Otto cycle consists of two adiabatic and two isochoric processes (see Fig. 2) and its efficiency η_O is determined by the volume span

$$\eta_O = 1 - \left(\frac{V_1}{V_2}\right)^{\frac{\gamma-1}{\gamma}} \tag{1}$$

(γ being the gas constant), or equivalently, by the temperature span of the hotter adiabatic process

$$\eta_O = 1 - \frac{T_2}{T_1}\,. \tag{2}$$

Interestingly, the efficiency does not depend on the temperatures of the colder adiabatic process, in particular, on the coldest temperature of the cycle T_3. It was suggested recently [3, 4] to take advantage of the internal degrees of freedom of the gas particles (which are considered to be completely decoupled from the center of mass (COM) degrees of freedom) to improve the Otto cycle efficiency. A combined system of masers and lasers working between the hottest and coldest temperatures of the cycle can extract useful laser work, thus increasing the overall efficiency.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Fig. 1. Combined Otto and Carnot cycle. The Otto cycle works with the COM degrees of freedom and produces work $W^{(\mathrm{Otto})}$, while the Carnot cycle works with the internal magnetic levels of the medium. A permanent magnet producing field $\boldsymbol{B}$ is attracted to the paramagnetic medium. The magnet moves towards the medium when the medium is cold and the attractive force is large, and it moves in the opposite direction when the medium is hot and the attractive force is small. As a result, net work $W^{(\mathrm{Carnot})}$ is produced.

We will show that a simple Carnot cycle operating on the internal degrees of freedom between the hottest and coldest temperatures can increase the Otto cycle efficiency by the maximum amount. In particular, if each of the gas particles has N internal states decoupled from the COM and capable of a separate thermodynamic cycle, then the combined cycle efficiency can approach the limiting value

$$\eta_{\max} = \frac{T_1 - T_2 + \dfrac{T_1 - T_3}{T_2 - T_3}\, T_2(\gamma - 1)\ln N}{T_1 + \dfrac{T_1}{T_2 - T_3}\, T_2(\gamma - 1)\ln N}\,. \tag{3}$$

An example (see Fig. 1) of such an improved cycle is discussed.

2. Basic Considerations of the Otto Cycle

Let us consider the cycle as in Fig. 2. The heat received by the working fluid from the hot reservoir during the isochoric heating with the smaller volume V_1 is

$$Q_1^{(\mathrm{Otto})} = C_V(T_1 - T_4)\,, \tag{4}$$

whereas the heat transmitted to the cold reservoir during the isochoric cooling with the bigger volume V_2 is

$$Q_2^{(\mathrm{Otto})} = C_V(T_2 - T_3)\,. \tag{5}$$

The net work produced by the working fluid is the difference between the heat input and output, $W^{(\mathrm{Otto})} = Q_1^{(\mathrm{Otto})} - Q_2^{(\mathrm{Otto})}$. Since from the adiabatic processes we have

$$\frac{T_2}{T_1} = \left(\frac{V_1}{V_2}\right)^{\frac{\gamma-1}{\gamma}} = \frac{T_3}{T_4}\,, \tag{6}$$

Fig. 2. *TS* and *PV* diagrams for an ideal Otto cycle

we can write for the work

$$W^{(\mathrm{Otto})} = C_V(T_1 - T_4)\left(1 - \frac{T_2}{T_1}\right). \tag{7}$$

Using the definition of the heat-engine efficiency

$$\eta_O = \frac{W^{(\mathrm{Otto})}}{Q_1^{(\mathrm{Otto})}}, \tag{8}$$

we obtain the expressions (1) and (2). As can be seen, lowering the coldest temperature T_3 does not improve the Otto cycle efficiency, although it helps in increasing the net work $W^{(\mathrm{Otto})}$.

3. Carnot Cycle with an *N* Level System

Let us consider first a two-level system, the generalization to an N level system being straightforward. Let us assume that one can change the energy difference of the two levels by an external parameter (e.g., if the levels refer to magnetic momentum states, their energies are determined by the external magnetic field). The Carnot cycle then looks as follows (see Fig. 3):

State 1: the energy difference of the levels is ϵ_1, the temperature is T_1. The entropy of the state is $S_1 = S(\epsilon_1, T_1)$, where the entropy of a two-level system is given by

$$S(\epsilon, T) = k\left[\ln\left(1 + e^{-x}\right) + \frac{x}{e^x + 1}\right], \tag{9}$$

$$x \equiv \frac{\epsilon}{kT}, \tag{10}$$

where k is the Boltzmann constant.

Fig. 3. A scheme of a Carnot cycle with a two-level system

Isothermal heating: the energy difference is being decreased while heat flows from the hot reservoir to the system. The heat given to the system is

$$Q_1^{(\text{Carnot})} = T_1[S(\epsilon_2, T_1) - S(\epsilon_1, T_1)]\,. \tag{11}$$

State 2: the energy difference is ϵ_2, the temperature remains T_1. The entropy is $S_2 = S(\epsilon_2, T_1)$.

Adiabatic decrease of the energy difference: the population of the levels remains constant, and because the energy difference is decreasing, the temperature decreases, too.

State 3: the energy difference is ϵ_3, the temperature is T_3 (taken to be equal to the lowest temperature of the Otto cycle). The entropy is $S_3 = S(\epsilon_3, T_3) = S_2$.

Isothermal cooling: the energy difference increases and the temperature is kept low by contact with the reservoir. The heat extracted from the system is

$$Q_2^{(\text{Carnot})} = T_3[S(\epsilon_3, T_3) - S(\epsilon_4, T_4)]. \tag{12}$$

State 4: the energy difference is $\epsilon 4$ and the temperature is T_3. $\epsilon 4$ is chosen such that the entropy is $S_4 = S(\epsilon_4, T_3) = S_1$.

Adiabatic increase of the energy difference: the system returns to the initial state 1.

The net work performed by the system is

$$W^{(\text{Carnot})} = Q_1^{(\text{Carnot})} - Q_2^{(\text{Carnot})}. \tag{13}$$

If the temperatures T_1 and T_3 are fixed, the maximum work is obtained if the system energy differences are chosen such that

$$\epsilon_1,\, \epsilon_4 \gg kT_1 \rightarrow S(\epsilon_1, T_1)\,, \qquad S(\epsilon_4, T_3) \ll k\,, \tag{14}$$

and

$$\epsilon_2,\, \epsilon_3 \ll kT_1 \rightarrow S(\epsilon_2, T_1)\,, \qquad S(\epsilon_3, T_3) \approx k \ln 2\,. \tag{15}$$

One then obtains

$$Q_1^{(\text{Carnot})} \approx kT_1 \ln 2\,, \tag{16}$$

$$Q_2^{(\text{Carnot})} \approx kT_3 \ln 2\,, \tag{17}$$

from which the maximum extractable work is

$$W_{\max}^{(\text{Carnot})} = k(T_1 - T_3) \ln 2\,. \tag{18}$$

It is not difficult to generalize the considerations to an N level system, for which

$$W_{\max}^{(\text{Carnot})} = k(T_1 - T_3) \ln N\,. \tag{19}$$

A Carnot engine working with internal degrees of freedom can be realized using, e.g., magnetic levels of a paramagnetic substance in an external magnetic field (see the TS and mB

Fig. 4. *TS* and *mB* diagrams for a Carnot cycle with a two-level magnetic system. The quantities m and B refer to the magnetic moment and the external magnetic field, respectively. Note that during an adiabatic process the mean magnetic moment of a two-level system remains constant.

diagrams in Fig. 4). Similar cycles are used in cooling devices [5, 6] to reach temperatures about 1 K. Magnetic Carnot cycles with mechanical output work were studied, e.g., in [7]. In our model the Carnot cycle produces mechanical work by moving a permanent magnet in vicinity of the paramagnetic working fluid (see Fig. 1). The magnet moves towards the fluid when the fluid is cold and the magnetic moments of its particles are mostly aligned with the magnetic field. Thus the magnetization of the fluid is large and useful work is produced by moving the magnet in the direction of the large attractive force. The magnet moves in the opposite direction when the fluid is hot and many magnetic moments point against the external field due to the thermal motion. The magnetization of the fluid is then small so that the waste work necessary for moving the magnet against the force is smaller than the useful work. Alternately, the magnetic Carnot cycle can produce electric work, e.g., if the working fluid is used as a core of a coil in an LC oscillator. Periodic connection of the paramagnetic fluid to the heat and cold reservoirs then induce parametric oscillations in the circuit.

4. Combined Cycle

Let us consider a combined cycle with the COM degrees of freedom used to produce mechanical work in an Otto cycle, and the internal levels are used separately in a Carnot cycle working between the maximum an minimum temperatures T_1 and T_3. The individual steps of the cycle are (cf. [3]):

- $(1 \to 2)$ The hot gas expands adiabatically, doing useful work. COM temperature drops from T_1 to T_2, internal temperature remains T_1 (same as in [3]).
- $(2 \to 3)$ Heat is extracted from the COM at constant volume, reaching the lowest temperature T_3 (same as in [3]).
- $(3 \to 3' \to 4)$ Adiabatic change of the internal energy-level difference from ϵ_2 to ϵ_3 (see Fig. 3), the internal temperature drops from T_1 to T_3 (e.g., adiabatic paramagnetic cooling). In contact with the cold reservoir, isothermal change of the internal energy difference to ϵ_4 (different from [3]).
- $(4 \to 5)$ Adiabatic compression of the COM volume, increasing the COM temperature to T_4. The internal temperature remains T_3 (same as in [3]).
- $(5 \to 6)$ Isochoric heating of the COM to the initial temperature T_1. The internal temperature remains T_3 (same as in [3]).
- $(6 \to 6' \to 1)$ Adiabatic change of the internal energy-level difference from ϵ_4 to ϵ_1, the internal temperature increases to T_1. In contact with the hot reservoir, isothermal change of the internal energy difference to ϵ_2 (different from [3]).

The efficiency of the cycle is

$$\eta = \frac{W^{(\mathrm{Otto})} + W^{(\mathrm{Carnot})}}{Q_1^{(\mathrm{Otto})} + Q_1^{(\mathrm{Carnot})}} \,. \tag{20}$$

Using the expressions for the Otto cycle (4) and (7), taking into account that for the ideal gas

$$C_V = \frac{k}{\gamma - 1} \,, \tag{21}$$

and assuming the limiting case for the Carnot cycle, Eq. (19), one obtains the final formula (3).

As an example, let us assume a single-atom ideal gas with $\gamma = 5/3$. Let us assume that the volume is compressed such that $V_1/V_2 = 1/10$, and the lowest temperature is $T_3 = T_1/5$. One finds for the Otto cycle that

$$Q_1^{(\mathrm{Otto})} \approx 0.741 kT_1 \,, \tag{22}$$

$$W^{(\mathrm{Otto})} \approx 0.450 kT_1 \,, \tag{23}$$

$$\eta_O \approx 60.2\% \,. \tag{24}$$

A two-level system Carnot cycle has in the limiting case

$$Q_1^{(\mathrm{Carnot})} \approx 0.693 kT_1 \,, \tag{25}$$

$$W^{(\mathrm{Carnot})} \approx 0.554 kT_1 \,, \tag{26}$$

$$\eta_C \approx 80\% \,. \tag{27}$$

The efficiency of the combined system is then, on using (20)

$$\eta \approx 69.7\% \,, \tag{28}$$

which is a substantial improvement of (24).

5. Conclusion

The improvement of the Otto cycle efficiency is based on running an additional thermodynamic cycle between the hottest and coldest temperatures, with working fluid represented by the internal degrees of freedom of the gas particles. In [3] this additional cycle was realized by lasers and masers. The cycle of [3] contained irreversible steps (e.g., an initially hot internal state at T_1 is being thermalized with a cold maser cavity at T_3, etc.) which suggests that other improvements are possible. On the other hand, the cycle presented here is completely reversible. Thus, given the temperatures T_1, T_2 and T_3, and the number N of internal states participating in the additional cycle, the formula (3) represents the maximum possible efficiency of the complete cycle.

References

[1] H. E. D. Scovil and E. O. Schulz-DuBois, *Three-level masers as heat engines*, Phys. Rev. Lett. **2** (1959) 262.

[2] See for example E. Geva and R. Kosloff, *On the classical limit of quantum thermodynamics in finite-time*, J. Chem. Phys. **97** (1992) 4398;

S. LLOYD, *Quantum-mechanical Maxwell's demon*, Phys. Rev. A **56** (1997) 3374;
T. FELDMANN and R. KOSLOFF, *Performance of discrete heat engines and heat pumps in finite time*, Phys. Rev. **E61** (2000) 4774;
R. KOSLOFF, E. GEVA, and J. M. GORDON, *Quantum refrigerators in quest of the absolute zero*, J. Appl. Phys. **87** (2000) 8093 and references therein.

[3] M. O. SCULLY, *The quantum afterburner*, Phys. Rev. Lett., to be published.

[4] M. O. SCULLY, A. B. MATSKO, N. NAYAK, Y. V. ROSTOVTSEV, and M. S. ZUBAIRY, *The quantum Otto-mobile: improving engine efficiency via lasing off exhaust gassess*, in preparation.

[5] T. NUMAZAWA, H. KIMURA, M. SATO, and H. MAEDA, *Carnot magnetic refrigerator operating between 1.4-K and 10-K*, Cryogenics **33** (1993) 547;
A. BEZAGUET, J. CASASCUBILLOS, P. LEBRUN, R. LOSSERANDMADOUX, M. MARQUET, M. SCHMIDTRICKER, and P. SEYFERT, *Design and construction of a static magnetic refrigerator operating between 1.8-K and 4.5-K*, ibid **34** (1994) 227, Suppl. ICEC.

[6] Z. J. YAN and J. C. CHEN, *The characteristics of polytropic magnetic refrigeration cycles*, J. Appl. Phys. **70** (1991) 1911;
L. X. CHEN and Z. J. YAN, *Main characteristics of a Brayton refrigeration cycle of paramagnetic salt*, ibid **75** (1994) 1249.

[7] H. TOFTLUND, *A rotary Curie-point magnetic engine — a simple demonstration of a Carnot-cycle device*, Am. J. Phys. **55** (1987) 48;
G. BARNES, *The 2 cycles of the rotary Curie-point heat engine*, ibid **57** (1989) 223;
P. G. MATTOCKS, *A 0.7-mW magnetic heat engine*, ibid **58** (1990) 545.

Fortschr. Phys. **50** (2002) 5–7, 664–669

Atom Optics – from de Broglie Waves to Heisenberg Ferromagnets

HAN PU, CHRIS SEARCH, WEIPING ZHANG, and PIERRE MEYSTRE

Optical Sciences Center, The University of Arizona, Tucson, AZ 85721, USA

Abstract

We review some of the key developments that lead to the field of atom optics, and discuss how it has recently began to make connections with Werner Heisenberg's trailblazing work on magnetism.

1. Introduction

At the September 10, 1923 session of the French Academy of Sciences, Louis Perrin presented a note by Prince Louis de Broglie entitled "Radiations – Ondes et Quanta" [1]. This was a remarkable paper drawing on the analogy between "atoms of light" – les atomes de lumière – and electrons. De Broglie believed that an "atom of light" should be considered as a moving object with a very small mass ($< 10^{-50}$ g) and with a speed nearly equal to c, although slightly less." By analogy, he found that it was almost necessary to suppose ("il est presque nécessaire de supposer") that there is, associated with the electron motion, a fictitious wave at what is now known as the de Broglie wavelength. Amazingly, this somewhat flawed argumentation lead, together with other developments in atomic and optical physics, to one of the greatest scientific revolutions, the invention of quantum mechanics, carried out independently by Heisenberg [2] and by Schrödinger [3] in 1925.

The optics of atomic de Broglie waves, or *atom optics*, is now thriving and rapidly maturing [4]. Linear, nonlinear, quantum and integrated atom optics are witnessing exciting developments and may soon lead to practical applications, such as e.g. sensors of unprecedented accuracy and sensitivity.

Experimental atom optics can be traced back to the seminal experiments by O. Stern [5], who demonstrated the reflection and diffraction of atoms from metallic and crystalline surfaces, and to R. Frisch [6], who measured the deflection of atoms as a result of the absorption and spontaneous emission of light. These pioneering experiments were truly heroic, due to the difficulties associated with the extremely short room-temperature atomic de Broglie wavelength $\Lambda_{\rm th}$, of the order of a fraction of an angstrom. The situation changed drastically following the invention of laser cooling [7], which allows one to readily bring atomic samples down to temperatures of the order of 10^{-6} kelvin. Since $\Lambda_{\rm th} \propto 1/\sqrt{T}$, it is easily seen that such low temperatures lead to very large atomic de Broglie wavelengths, of the order of microns or longer.

Even lower temperatures can be achieved, most notably using forced evaporative cooling [8, 9], which can lead to temperatures of the order of a few nanokelvin. At this point, even atomic samples of modest densities ($10^{13} - 10^{14}$ cm^{-3}) suffer the effects of quantum statistics and must be described as undistinguishable particles. In the case of bosonic atoms, this can lead under appropriate conditions to the Bose-Einstein condensation of the sample, as predicted by Einstein [10] in 1924 following the work of Bose [11], and first experimentally verified in an atomic vapor in 1995 [12, 13].

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Atom optics bridges many subfields of physics, from atomic and optical physics to statistical mechanics and to condensed matter physics. It also bears on information science, metrology, nonlinear science, and even relativity. It borrows from and contributes to these fields in numerous ways. In the context of this meeting celebrating Werner Heisenberg, we find it appropriate to discuss one specific aspect of atom optics, the close analogy between the dynamics of Bose-Einstein condensates on optical lattices [14−18], and models of magnetism in condensed matter physics [20−23], since Heisenberg was of course the father of the quantum theory of magnetism.

A particularly interesting geometry in the context of the present paper involves quantum-degenerate Bose gases trapped in optical lattices. They were first used to demonstrate "mode-locked" atom lasers [14]. It was also predicted [15] and demonstrated [16] that they undergo a Mott insulator phase transition as the depth of the lattice wells is increased. We also mention their application in studies of quantum chaos [17−19].

In a parallel development, recent experiments have established that the ground state of optically trapped ^{87}Rb spinor Bose condensates must be ferromagnetic at $T = 0$ [24, 25]. What this means is that when located in optical lattices deep enough for the individual sites to be independent, the "mini-condensates" at each lattice site behave as individual ferromagnets. In the absence of external magnetic fields and of site-to-site interactions, these mesoscopic magnets point in random directions. However, magnetic and optical dipole-dipole interactions can couple neighboring sites, and as a result it can be expected that under appropriate conditions the mini-condensates will arrange themselves in ferromagnetic or antiferromagnetic configurations, that spin waves can be launched in the lattice, etc.

We recall that ferromagnetism remained mysterious until Heisenberg identified the crucial role played by the exchange interaction. In particular, the magnetic dipole-dipole interaction was too weak to explain the observed properties of ferromagnets. The traditional starting point for the study of magnetism in solids is now the Heisenberg Hamiltonian [23] $H_{\rm spin} = -\sum J_{ij}\, \boldsymbol{S}_i \cdot \boldsymbol{S}_j$, where $\boldsymbol{S}_i$ is the spin operator for ith electron, the J_{ij} are exchange coupling constants. They arise from the combined effects of the direct Coulomb interaction among electrons and the Pauli exclusion principle.

That it has now become possible to "turn back the clock" and study such effects in systems where a "classical" interaction between sites is dominant, is an immediate consequence of the collective behavior of the atoms in each of the mini-condensates: because of the Bose enhancement resulting from the presence of coherent ensembles of N_i atoms − typically several thousands in one-dimensional lattices − in the "mini-condensate" at each lattice site i; each of the mini-condensates acts as a mesoscopic ferromagnet N_i times stronger than that associated with an individual atom. The magnetic dipole-dipole interaction can therefore become significant despite the large distance between sites. Hence it becomes possible to carry out detailed static and dynamical studies of magnetism in one- to three-dimensional condensate lattices [26, 27]. A particularly attractive feature is that since Bose-condensed atomic systems are weakly interacting, a detailed microscopic description of these systems is possible, in contrast to typical condensed-matter situations.

2. Ferromagnetism on a Bosonic Lattice

Our starting point is the Hamiltonian H describing an $F = 1$ spinor condensate at zero temperature trapped in an optical lattice along the z-axis, subject to a magnetic dipole-dipole interaction $H_{\rm dd}$. More precisely, we include the long-range magnetic dipole-dipole interaction between different lattice sites, but neglect it within each site, assuming that it is much weaker than the s-wave ground-state collisions. We further assume that the optical lattice potential is deep enough that there is no spatial overlap between the condensates at different lattice sites. Finally, the atoms are coupled to an external magnetic field $\boldsymbol{B}_{\rm ext}$ [28−30].

For a one-dimensional lattice of mini-condensates with equal number of atoms, and assuming that all Zeeman sublevels of the atoms share the same spatial wave function, the total Hamiltonian of the system takes the form [30]

$$H = \sum_i \left[\lambda'_a \boldsymbol{S}_i^2 + \sum_{j \neq i} \lambda_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j - 3 \sum_{j \neq i} \lambda_{ij} S_i^z S_j^z - \gamma_B \boldsymbol{S}_i \cdot \boldsymbol{B}_{\text{ext}} \right] . \tag{1}$$

In this expression, the coefficient λ'_a account for spin-changing collisions within the individual mini-condensates, λ_{ij} accounts for the magnetic dipole-dipole interaction between condensates, and γ_B is the gyromagnetic ratio. μ_B being the Bohr magneton and g_F the Landé factor. In the case of ^{87}Rb the individual spinor condensates at the lattice sites are ferromagnetic, $\lambda'_a < 0$.

In contrast to the usual situation in ferromagnetism, where the temperature plays an essential role and the ferromagnetic transition occurs at the Curie temperature T_C, the present analysis is at zero temperature, $T = 0$. Nonetheless, effects similar to those of a finite temperature analysis can be simulated by an appropriate choice of the applied magnetic field $\boldsymbol{B}_{\text{ext}}(\boldsymbol{r})$. Specifically, we consider a situation where $\boldsymbol{B}_{\text{ext}}(\boldsymbol{r})$ consists of two contributions, $\boldsymbol{B}_{\text{ext}} = B_z \hat{\boldsymbol{z}} + B_\rho \hat{\boldsymbol{\rho}}$, where $\rho = \sqrt{x^2 + y^2}$ is the radial coordinate. Here, B_z is a controlled, external applied field that we take along the z-axis, and B_ρ is an effective "stray" field, that we take along the transverse direction. It accounts for uncontrolled aspects of the experimental environment, such as e.g. environmental magnetic fluctuations. We will see shortly that this transverse field plays a role analogous to temperature in conventional ferromagnetism.

In the case of an infinite lattice, the Hamiltonian of a generic site i reads

$$h_i = \lambda'_a \boldsymbol{S}^2 - \gamma_B \boldsymbol{S} \cdot \left[\left(B_z + 2 \sum_{j \neq i} \lambda_{ij} S_j^z \right) \hat{\boldsymbol{z}} + \left(B_\rho - \sum_{j \neq i} \lambda_{ij} S_j^\rho \right) \hat{\boldsymbol{\rho}} \right] . \tag{2}$$

We determine its ground state in the Weiss molecular potential approximation [21, 22], which approximates this Hamiltonian by

$$h_{\text{mf}} = \lambda'_a \boldsymbol{S}^2 - \gamma_B \boldsymbol{S} \cdot \boldsymbol{B}_{\text{eff}} , \tag{3}$$

where we have introduced the effective magnetic field

$$\boldsymbol{B}_{\text{eff}} = (B_z + 2\Lambda m_z)\, \hat{\boldsymbol{z}} + (B_\rho - \Lambda m_\rho)\, \hat{\boldsymbol{\rho}} \tag{4}$$

with $\Lambda = N \sum_{j \neq i} \lambda_{ij}$ and $m_\alpha = \langle S_j^\alpha \rangle / N$.

For ferromagnetic mini-condensates, the ground state corresponds to the situation where the condensate at site i is aligned along $\boldsymbol{B}_{\text{eff}}$ and S takes its maximum possible value N. That is, the ground state of the mean-field Hamiltonian (3) is simply

$$|GS\rangle = |N, N\rangle_{\boldsymbol{B}_{\text{eff}}} , \tag{5}$$

where the first number denotes the total angular momentum and the second its component along the direction of $\boldsymbol{B}_{\text{eff}}$.

Equation (5) illustrates clearly the essential element brought about by the presence of mini-condensates at the individual lattice sites: in contrast to the situation for an incoherent atomic sample hwere the enhancement factor would be $\sqrt{N}$, the coherent behavior of all N atoms within one site results in a ground-state magnetic dipole moment equal to N times that of an individual atom. It is this property that leads to a significant magnetic

dipole-dipole interaction even for lattice points separated by hundreds of nanometers. As such, the situation at hand is in stark contrast with usual ferromagnetism, where the magnetic interaction is negligible compared to exchange and the use of fermions is essential [22].

From the ground state (5) and for $B_z = 0$, the components $m_\alpha = \frac{1}{N}\langle GS| S_i^\alpha |GS\rangle = \cos\theta_\alpha$ of the magnetization, where θ_α is the angle between $\boldsymbol{B}_{\rm eff}$ and the α-axis, are

$$m_z = \frac{2\Lambda m_z}{\sqrt{(2\Lambda m_z)^2 + (B_\rho - \Lambda m_\rho)^2}}\,, \qquad m_\rho = \frac{B_\rho - \Lambda m_\rho}{\sqrt{(2\Lambda m_z)^2 + (B_\rho - \Lambda m_\rho)^2}}\,. \tag{6}$$

For $B_\rho \geq 3\Lambda$, the only solutions are $m_z = 0$ and $m_\rho = 1$. That is, the lattice of mini-condensates is magnetically polarized along the transverse magnetic field. For $B_\rho < 3\Lambda$, in contrast, there are two coexisting sets of solutions: (i) $m_z = 0$ and $m_\rho = 1$; and (ii) $m_z = \pm\sqrt{1 - (B_\rho/3\Lambda)^2}$ and $m_\rho = B_\rho/3\Lambda$. However, it is easily seen that the state associated with the latter solution has the lower energy. Hence it corresponds to the true ground state, while solution (i) represents an unstable equilibrium.

We have, then, the following situation: As B_ρ is reduced below a critical value 3Λ, the condensate lattice ceases to be polarized along its direction. A phase transition occurs, characterized by a spontaneous magnetization with a finite m_z along the z-direction.

From the preceding discussion, one might expect that the site-to-site coupling provide by the magnetic dipole-dipole interaction conceive that the magnetic dipole-dipole interaction will lead to the excitation of spin waves.[1)] This, however, is not quite the case. Simple order-of-magnitude estimates indicate that this interaction alone is not sufficient to excite spin waves in a lattice of Bose-Einstein condensate under presently achievable experimental conditions. One way out of this difficulty is to add an external laser field to couple the mini-condensates on the various sites via the electric dipole-dipole interaction. With this added interaction, the Hamiltonian (1) becomes [32, 33]

$$H = \sum_i \left[\lambda'_a \hat{\boldsymbol{S}}_i^2 - \gamma_B \hat{\boldsymbol{S}}_i \cdot \boldsymbol{B} - \sum_{j\neq i} J_{ij}^z \hat{S}_i^z \hat{S}_j^z - \sum_{j\neq i} J_{ij}(\hat{S}_i^{(-)}\hat{S}_j^{(+)} + \hat{S}_i^{(+)}\hat{S}_j^{(-)})\right]. \tag{7}$$

The explicit form of the coupling coefficients for the case of linearly polarized optical fields is given in Ref. [27].

From the Hamiltonian (7), we can derive the Heisenberg equations of motion for the spin excitations as

$$i\hbar\,\frac{\partial \hat{S}_q^{(-)}}{\partial t} = (\omega_0 + \Delta\omega_q)\,\hat{S}_q^{(-)} - \sum_{j\neq q} \chi_{qj}\hat{S}_j^{(-)}\,, \tag{8}$$

where we invoked the mean-field approximation to replace the spin operator $\hat{S}_q^z$ by its ground state expectation value. The frequencies $\omega_0 = -\gamma_B B$ and $\Delta\omega_q = 2\sum_{j\neq q} J_{qj}^z N_j \hbar$ describe the precessing of the qth spin caused by the external magnetic field and the static magnetic dipolar interaction. The site-to-site spin coupling coefficients $\chi_{qj} = 2J_{qj}N_q\hbar$ determine the propagation of the spin waves.

[1)] We recall that in analogy with phonons, which are normal modes of motion of atoms displaced from their equilibrium position in a crystal, spin waves are normal modes of spin excitation in materials with an ordered magnetic structure. These waves and their quanta of excitation, the magnons, were first discussed by F. Bloch, a student of Heisenberg, in 1930 [31].

In the long-wavelength limit, Eq. (8) leads to an effective Schrödinger equation [27]

$$i\,\frac{\partial S(y,\,t)}{\partial t} = \left[-\frac{\beta_1}{2}\,\frac{\partial^2}{\partial y^2} - \beta_0 + \omega(y)\right] S(y,\,t)\,, \tag{9}$$

where we have introduced the continuous limit quantities $\hat{S}_q^{(-)} \to S(y,\,t)$, $\chi_{qj} \to \chi(y-y')$, and $\omega_0 + \Delta\omega_1 \to \omega(y)$, and we have defined $\beta_n = (2/\lambda_L)\int d\eta\,\chi(\eta)\,\eta^{2n}$ for $n = 0, 1$. Eq. (9) describes the motion of "waves" caused by spin excitations in the x-y plane. The magnon dispersion relation is presented in Ref. [27], which also discusses possible techniques to experimentally establish the existence of spin waves in the lattice of mini-condensates.

3. Outlook

While the possibility of atom optics was of course obvious from de Broglie's work, it took three quarters of a century before it became somewhat practical, and it is only now that we begin to get a glimpse at its future promise.

While device applications such as rotation sensors and improved clocks are rather evident, the true technological potential of this emerging field remains unclear. In fundamental science, though, the situation is less murky. Atom optics now has a profound impact on our understanding of problems ranging from statistical and manybody physics to atomic physics and to condensed matter physics. In particular, we have shown in this note how it opens up the way to the study of magnetism in situations under exquisite control, both theoretically and experimentally.

Discussions with Brian Anderson on the experimental aspects of this problem are gratefully acknowledged. This work is supported in part by the US Office of Naval Research under Contract No. 14-91-J1205, by the National Science Foundation under Grant No. PHY98-01099, by the US Army Research Office, by NASA, and by the Joint Services Optics Program.

References

[1] L. de Broglie, Comp. Rend. Ac. Sci. **177** (1923) 507.
[2] W. Heisenberg, Z. Physik **33** (1925) 879;
M. Born, W. Heisenberg, and P. Jordan, Z. Physik **35** (1926) 557.
[3] E. Schrödinger, Ann. Phys. **79** (1926) 361; **79** (1926) 489; **79** (1926) 734.
[4] see e.g. P. Meystre, *Atom Optics*, Springer Verlag, New York (2001).
[5] O. Stern, Naturwiss. **17** (1929) 391.
[6] R. Frisch, Z. Physik **86** (1933) 42.
[7] see e.g. C. Cohen-Tannoudji, in *Fundamental Systems in Quantum Optics*, edited by J. Dalibard, J. M. Raimond, and J. Zinn-Justin, North-Holland, Amsterdam (1992), p. 3;
H. J. Metcalf and P. van der Straten, *Laser Cooling and Trapping*, Springer Verlag, New York (1999).
[8] D. E. Pritchard, K. Helmerson, and A. G. Martin, in *Atomic Physics 11*, edited by S. Haroche, J. C. Gay, and G. Grynberg, World Scientific, Singapore (1989).
[9] T. W. Hijmans, O. J. Luiten, I. D. Setija, and J. T. M. Walraven, J. Opt. Soc. Am. **6** (1989) 2235.
[10] A. Einstein, Sitzungsber. Preuss. Akad. Wiss. **1924** (1924) 261; **1925** (1925) 3.
[11] S. N. Bose, Z. Physik **26** (1924) 178.
[12] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. Cornell, Science **269** (1995) 198.

[13] K. B. DAVIS, M. O. MEWES, M. R. ANDREWS, N. J. VAN DRUTEN, D. S. DURFEE, D. M. KURN, and W. KETTERLE, Phys. Rev. Lett. **75** (1995) 3969.
[14] B. P. ANDERSON and M. A. Kasevich, Science **281** (1998) 1686.
[15] D. JAKSCH et al., Phys. Rev. Lett. **81** (1998) 3108.
[16] T. W. HÄNSCH, Private communication (2001).
[17] M. RAIZEN, private communication (2001).
[18] R. CARRETERO-GONZALEZ and K. PROMISLOW, e-print cond-mat/0105600, (2001).
[19] W. K. HENSINGER et al., Nature (London) **412** (2001) 52.
[20] W. HEISENBERG, Z. Physik **49** (1926) 619.
[21] C. KITTEL, *Introduction to Solid State Physics*, Wiley, New York (1996).
[22] C. KITTEL, *Quantum Theory of Solids*, second revised printing, Wiley, New York (1987).
[23] K. HUANG, *Statistical Mechanics* (John Wiley & Sons, New York, 1987).
[24] D. HEINZEN, private communication (2001).
[25] J. P. BURKE, Jr. and J. L. BOHN, Phys. Rev. A **59** (1999) 1303;
N. N. CLAUSEN, J. L. BOHN, and C. H. GREENE, Phys. Rev. **A64** (2001) 053602.
[26] H. PU, W. ZHANG, and P. MEYSTRE, Phys. Rev. Lett. **87** (2001) 140405.
[27] W. ZHANG, H. PU, C. SEARCH, and P. MEYSTRE, preprint cond-mat/0109415 (2001).
[28] T.-L. HO, Phys. Rev. Lett. **81** (1998) 742.
[29] T. OHMI and K. MACHIDA, J. Phys. Soc. Japan. **67** (1998) 1882.
[30] C. K. LAW, H. PU, and N. P. BIGELOW, Phys. Rev. Lett. **81** (1998) 5257.
[31] F. BLOCH, Z. Physik **61** (1930) 206.
[32] WEIPING ZHANG and D. F. WALLS, Phys. Rev. A **49** (1994) 3799.
[33] WEIPING ZHANG and D. F. WALLS, Phys. Rev. A **57** (1998) 1248.

Fortschr. Phys. **50** (2002) 5–7, 670–674

Quantum Field Theory and Gravitation

On Tachyon Condensation in String Theory: Worldsheet Computations

OLEG ANDREEV[1)]

Humboldt-Universität zu Berlin, Institut für Physik, Invalidenstr. 110
D-10115 Berlin, Germany
andreev@physik.hu-berlin.de

Abstract

I briefly review some recent results of computations of the tachyon potentials by worldsheet methods.

1. Introduction

It is a big problem to better understand the vacuum structure of string/M theory. Even in the case of bosonic string theory which contains the tachyon near its perturbative vacuum, it turned out to be not a simple deal to find a stable vacuum. Note that hopes for the existence of such a vacuum have their origins already in the seventies [1]. However, for open strings a real breakthrough has been achieved in recent years when the open string tachyons were related with annihilation or decay of D-branes via the process of tachyon condensation [3] (see [3] for a review and a list of references).

In studying the phenomenon of tachyon condensation, string field theory methods are the most appropriate ones (see, e.g., [4] for recent developments). However, in practice it turns out to be very hard to deal with them. This was a motivation for finding simpler methods, for example toy models [5], that can allow one to gain some intuition on the physics of the phenomenon. Recently, it has been realized that the background independent open string field theory (BIOSFT) of Witten [6, 7] is a powerful tool to attack the problem. In particular, it allows one to compute the tree level effective tachyon potentials [8–10] which amusingly coincide with the potentials found from the toy models based on the exactly solvable Schrödinger problem [5]. In fact, it is also possible to take a more intuitive approach (see, e.g., [11]) which generalizes the ordinary sigma model approach to string theory (see [12] for a review).

2. Tree Level

The BIOSFT is based on the Batalin-Vilkovisky formalism whose master equation provides the effective action of the theory. In bosonic string theory, the tree level action of the tachyon field living on an unstable p-brane is given by [6, 8, 9]

$$S = \tau_p \int d^{p+1}x\, e^{-T}(1 + T + \alpha'\, \partial_i T\, \partial_i T + \ldots)\,, \tag{1}$$

[1)] On leave from Landau Institute for Theoretical Physics, Moscow.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

where the dots stand for an infinite number of higher derivative terms. These terms can in principle be considered as a result of integration over the other open string modes. τ_p stands for the p-brane tension.

The tachyon potential is

$$V(T) = (1+T)\, e^{-T}\,. \tag{2}$$

It has two extrema. The standard perturbative vacuum corresponds to the first one, $T = 0$. A new vacuum to which the tachyon condenses is related to the second extremum, $T = +\infty$.

In superstring theory, the tree level action of unstable D-brane is [10]

$$\hat{S} = \hat{\tau}_p \int d^{p+1}x\, e^{-T^2}(1 + 4\ln 2\alpha'\, \partial_i T\, \partial_i T + \ldots)\,, \tag{3}$$

where the dots again stand for an infinite number of higher derivative terms. Note that the odd and even p's are referred to type IIA and IIB, respectively. In other words, these unstable D-branes are simply the D-branes with the "wrong" p's.

The tachyon potential is now symmetric under reflection $T \to -T$ and given by

$$\hat{V}(T) = e^{-T^2}\,. \tag{4}$$

It has three extrema. As in the bosonic case, the standard perturbative vacuum corresponds to the first one, $T = 0$. A new vacuum to which the tachyon condenses is related to the rest, $T = \pm\infty$.

I finish the discussion of the tree level results with a few comments:

(i) It is easy to incorporate the abelian gauge field. To the leading order in α', the actions are given by [11, 13, 14]

$$S = \tau_p \int d^{p+1}x\, \sqrt{\det(1+F)}\, e^{-T}\, (1 + T + \alpha' G^{ij}\, \partial_i T\, \partial_j T + \ldots) \tag{5}$$

and

$$\hat{S} = \hat{\tau}_p \int d^{p+1}x\, \sqrt{\det(1+F)}\, e^{-T^2}\, (1 + 4\ln 2\alpha' G^{ij}\, \partial_i T\, \partial_j T + \ldots)\,, \tag{6}$$

where $G^{ij} = [(1-F^2)^{-1}]^{ij}$.

(ii) It is also easy to incorporate a constant B field that slightly modifies the closed string background. As a result, unstable branes are now described by noncommutative field theories. Their actions have the same form as the actions (1) and (3) but the product of the tachyon field is taken to be the star product [14, 15]. To be more precise, the star product is defined as $T(x) * T(x) = e^{\frac{i}{2}\theta^{ij}\frac{\partial}{\partial y^i}\frac{\partial}{\partial z^j}}\, T(y)\, T(z)|_{y=z=x}$, where θ is the antisymmetric matrix defined in terms of the B field.

(iii) In superstring theory, there is another class of unstable systems than the D-branes with the "wrong" p's. These are the brane-antibrane ($D\bar{D}$) systems. It turns out to be difficult even at the leading order in α' to compute their couplings to NS-NS backgrounds in the form that includes all powers of the fields like (6). However, what one can do now is to compute the couplings to a constant R-R background [16]. The result is a simple generalization of the well-known couplings of the D-branes to the R-R potentials. The only novelty is that the field strength of the gauge field is replaced to a curvature associated with the superconnection $\mathcal{A}$ which is defined as a 2×2 matrix whose elements expressed in terms of the tachyon and gauge fields.

3. One-Loop

A deeper understanding of the process of tachyon condensation requires information about quantum corrections to the tree level potentials (2) and (4). Unfortunately, the quantum master equation which is a proper machinery for doing so in the framework of the BIOSFT is still missing. Thus one is forced to use more intuitive approaches [17−19].

For example, Tassilo Ott and I [19] have recently used the consistency with the field theory results as a basic principle in studying one-loop corrections to the potentials. In the bosonic case, we have found the one-loop correction to the potential (2) in the following form

$$V_{1-\text{loop}}(T) = -\frac{1}{2}\int_0^\infty \frac{dl}{l}\,(8\pi^2\alpha' l)^{-\frac{p+1}{2}}\,e^{-\pi l T}\,\eta(il)^{-24}\,, \tag{7}$$

where η is the Dedekind eta function.

A detail analysis of (7) shows that the one-loop correction develops an imaginary part. There are in fact two contributions to it. The first one is due to the open string tachyon, while the second one comes from massive open string modes or, equivalently, from the closed string tachyon. The imaginary part of $V_{1-\text{loop}}(T)$ signals that the theory is perturbatively unstable. A physically meaningful quantity to compute in this case is the decay rate per unit volume of an initial unstable state. Explicitly, it is defined as $\Gamma = 2\,\text{Im}\,V_{1-\text{loop}}$.

In the supersymmetric case, the one-loop correction to the potential (4) takes the form

$$\hat{V}_{1-\text{loop}}(T) = -\frac{1}{2}\int_0^\infty \frac{dl}{l}\,(8\pi^2\alpha' l)^{-\frac{p+1}{2}}\,e^{-\pi l T^2}\,\eta(il)^{-8}[f_{NS}^8(l) - f_R^8(l)]\,, \tag{8}$$

where $f_{\text{NS}}(l) = e^{\frac{\pi l}{24}}\prod_{n=1}^\infty\,(1+e^{-\pi l(2n-1)})$ and $f_R(l) = \sqrt{2}\,e^{-\frac{\pi l}{12}}\prod_{n=1}^\infty\,(1+e^{-2\pi l n})$.

As in the bosonic case, $\hat{V}_{1-\text{loop}}$ develops an imaginary part. This time the open string tachyon only contributes. Indeed, there is no the closed string tachyon because of the corresponding cancellation between the NS and R sectors. The decay rate associated with this perturbative instability is found to be

$$\Gamma = \pi(8\pi\alpha')^{-(p+1)/2}\,\frac{(1-T^2)^{\frac{p+1}{2}}}{\Gamma(1+(p+1)/2)}\,H(1-T^2)\,, \tag{9}$$

where H stands for the Heaviside step function.

I conclude the discussion of the one-loop corrections by making a couple of remarks:

(i) We have found the one-loop corrections by computing the open string partition functions in the presence of a constant tachyon background. Doing the standard modular transformation, it is easy to rewrite the partition functions via closed string modes. The result might seem curious. Indeed, it reveals the transverse directions to the brane and looks like closed string modes appearing from one brane, propagating a distance R[2)], and then disappearing on the other brane. In other words, the description in terms of closed strings assumes that the original brane splits into two parallel branes separated with a distance depending on the vev of the tachyon field.

It is of some interest to notice that there is an analogy with field theory models that undergo spontaneous symmetry breaking [20] that helps to understand the effect of brane

[2)] $R = \sqrt{2\pi^2\alpha' T}$ for bosonic string and $R = \sqrt{2\pi^2\alpha'}\,|T|$ for superstring.

splitting. Indeed, given an effective potential, for example $V_{\text{eff}} \sim \left(T^2 - T_0^2\right)^2$, one can study the dynamics of spontaneous symmetry breaking by setting an initial narrow distribution of the scalar field at the maximum of the potential ($T = 0$). Then, by quantum fluctuations it spreads out and shows two maxima. To reveal such a picture for the branes, let me note two key facts: The first is that in terms of closed strings T transforms into R (a distance into a transverse direction to a brane). The second is that a brane itself can be considered as a lump solution of the higher dimensional worldvolume theory that allows one to say that this solution plays a role of the initial distribution in field theory. Finally, let me note that such a description of the decay shows that these lumps interact by the closed string exchanges.

(ii) The instabilities I have discussed so far are perturbative ones. The form of the potential in the bosonic case (see Eq. (2)) assumes that the new vacuum $T = +\infty$ is also unstable but now in the nonperturbative sense because of the tunneling through the potential barrier. It is well-known that in field theory one can compute decay probabilities of such unstable vacua by the instanton methods (see, e.g., [21]). Recently, I have computed the exponential factor in the decay probability of the vacuum $T = +\infty$ [22]. It is simply given by

$$w \sim e^{-\tau_p Z_D} , \tag{10}$$

where Z_D is the standard partition function of open bosonic string on the unit disk with the Dirichlet boundary conditions.

Acknowledgment

I thank the Alexander von Humboldt Foundation for inviting me to the Heisenberg Symposium. A part of the work reported here was done in collaboration with Tassilo Ott. This work was supported in part by DFG under grant No. DO 447/3-1 and the European Commission RTN Programme HPRN-CT-2000-00131.

References

[1] K. BARDAKCI, Nucl. Phys. **B68** (1974) 331;
K. BARDAKCI and M. B. HALPERN, Phys. Rev. **D10** (1974) 4230; Nucl. Phys. **B96** (1975) 285;
K. BARDAKCI, Nucl. Phys. **B133** (1978) 297;
V. A. KOSTELECKY and S. SAMUEL, Nucl. Phys. **B336** (1990) 263.

[2] A. SEN, Int. Jour. Mod. Phys. **A14** (1999) 4061 [hep-th/9902105]; JHEP **9912** (1999) 027 [hep-th/9911116].

[3] K. OHMORI, *A Review on Tachyon Condensation in Open String Field Theories*, [hep-th/0102085].

[4] http://online.itp.ucsb.edu/online/mtheory. This is a link to talks of D. GROSS, A. SEN, W. TAYLOR, and B. ZWIEBACH delivered at Avatars of M-Theory, Santa Barbara, California.

[5] J. A. MINAHAN and B. ZWIEBACH, JHEP **0009** (2000) 029 [hep-th/0008231]; JHEP **0103** (2001) 038 [hep-th/0009246].

[6] E. WITTEN, Phys. Rev. **D46** (1992) 5467 [hep-th/9208027]; Phys. Rev. **D47** (1993) 3405 [hep-th/9210065];
S. L. SHATASHVILI, Phys. Lett. **B311** (1993) 83 [hep-th/9303143]; *On the Problems with Background Independence in String Theory*, [hep-th/9311177].

[7] M. MARINO, JHEP **0106** (2001) 059 [hep-th/0103089];
V. NIARCHOS and N. PREZAS, Nucl. Phys. **B619** (2001) 51 [hep-th/0103102].

[8] A. A. GERASIMOV and S. SHATASHVILI, JHEP **0010** (2000) 034 [hep-th/0009103].

[9] D. KUTASOV, M. MARINO, and G. MOORE, JHEP **0010** (2000) 045 [hep-th/0009148].

[10] D. KUTASOV, M. MARINO, and G. MOORE, *Remarks on Tachyon Condensation in Superstring Theory*, [hep-th/0010108].

[11] A. A. Tseytlin, J. Math. Phys. **42** (2001) 2854 [hep-th/0011033].
[12] A. A. Tseytlin, Int. J. Mod. Phys. **A4** (1989) 1257.
[13] K. Li and E. Witten, Phys. Rev. **D48** (1993) 853 [hep-th/9303067].
[14] O. Andreev, Nucl. Phys. **B598** (2001) 151 [hep-th/0010218].
[15] L. Cornalba, Phys. Lett. **B504** (2001) 55 [hep-th/0010021].
K. Okuyama, Phys. Lett. **B499** (2001) 167 [hep-th/0010028].
[16] P. Kraus and F. Larsen, Phys. Rev. **D63** (2001) 106004 [hep-th/0012198];
T. Takayanagi, S. Terashima, and T. Uesugi, JHEP **0103** (2001) 019 [hep-th/0012210];
M. Alishahiha, H. Ita, and Y. Oz, Phys. Lett. **B503** (2001) 181 [hep-th/0012222].
[17] T. Suyama, *Tachyon Condensation and Spectrum of Strings on D-branes*, [hep-th/0102192];
R. Rashkov, K. S. Viswanathan, and Y. Yang, *Background Independent Open String Field Theory with Constant B Field On the Annulus*, [hep-th/0101207];
K. Bardakci and A. Konechny, *Tachyon condensation in boundary string field theory at one loop*, [hep-th/0105098];
B. Craps, P. Kraus, and F. Larsen, JHEP **0106** (2001) 062 [hep-th/0105227].
[18] K. S. Viswanathan and Y. Yang, Phys. Rev. **D64** (2001) 106007 [hep-th/0104099];
M. Alishahiha, Phys. Lett. **B510** (2001) 285 [hep-th/0104164];
G. Arutyunov, A. Pankiewicz, and B. Stefánski, jr, JHEP **0106** (2001) 49 [hep-th/0105238];
T. Lee, K. S. Viswanathan, and Y. Yang, *Boundary String Field Theory at One-Loop* [hep-th/0109032].
[19] O. Andreev and T. Ott, *On One-Loop Approximation to Tachyon Potentials*, [hep-th/0109187].
[20] G. Felder, J. Garcia-Bellido, P. Greene, L. Kofman, A. D. Linde, and I. Tkachev, Phys. Rev. Lett. **87** (2001) 011601 [hep-th/ hep-th/0106179].
[21] M. B. Voloshin, I. Yu. Kobzarev, and L. B. Okun', Yad. Fiz. **20** (1974) 1229;
S. Coleman, Phys. Rev. **D15** (1977) 2929;
C. G. Callan and S. Coleman, Phys. Rev. **D16** (1977) 1762.
[22] O. Andreev, *A Note on Nonperturbative Instability in String Theory* [hep-th/0112088].

Fortschr. Phys. **50** (2002) 5–7, 675–679

Duality in the Low Dimensional Field Theory

IVAN ANDRIĆ and DANIJEL JURMAN

Theoretical Physics Division, Rudjer Bosković Institute
P. O. Box 180, HR-10002 Zagreb, Croatia
andric@thphys.irb.hr
jurman@thphys.irb.hr

Abstract

A strong-weak coupling duality appears in the quantum mechanical many body system with the interacting potential proportional to the pair-wise inverse-squared distance. In the large N limit the field theory formulation possess the soliton solution in non-perturbative "Bogomolny-Prasad-Sommerfeld" limit (BPS). In this limit we establish duality between original and soliton fields and the dual relation among coupling constants.

1. Introduction

Duality is an important generalization of symmetry for studying relations between different theories. It is already present in Maxwell equations and is the powerful property of the spin models. In higher dimensions there is a web of dualities between several theories. With more degrees of freedom duality is enlarged.

In this short review we would like to show the appearance of duality in the systems with N particles on the line [1]. We shall formulate the field theory [2] and we shall show the existance of solitons in the theory [3, 4]. Due to symmetry $SU(1,1)$ (isomorphic to $\mathcal{O}(2,1)$) in this model we can construct parent Hamiltonian, [5, 6] which exhibits duality properties between original fields and solitons. There are moving solitons in the model but we shall only consider ground states due to existance of the BPS limit in which solitons are static. Here we find an interesting property of this construction, the hierarchy. This is the possibility that solitons condense and form new solitons of the next hierarchy.

2. Solitons in the Collective-field Approach

Let us mention briefly the few properties of the Calogero Hamiltonian [1]

$$\mathcal{H} = \frac{1}{2} \sum^{N} p_i^2 + \frac{1}{2} \sum_{i \neq j} \frac{\lambda(\lambda - 1)}{(x_i - x_j)^2} , \tag{1}$$

where the coupling constant $\lambda(\lambda - 1)$ is factorized by dimensionless parameter λ which will determine statistical properties of the particles. The interaction is singular when particles are approaching each other, so the wave function ought to have a prefactor $\prod_{iej} (x_i - x_j)^{\lambda}$ which is vanishing for coincidence particles. In order to preserve scaling and translational invariance of the model a proper treatment of the center-of-mass degrees of freedom are important. In the field theory formulation with center-of-mass variables the terms of order $1/N$ may appear, but in large N limit they are vanishing. Model (1) is exactly solvable and

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

integrable, both classically and quantum-mechanicaly [7]. We still lack a local canonical field-theoretical formulation [8] of (1) due to singular character of (1). But, in the large N limit a collective-field theory [9] is established [2], and it is used throughout the paper.

The idea of collective field approach is that on symmetric spaces Hamiltonian can be expressed entirely in terms of the density of particles $\rho(x) = \sum_i \delta(x - x_i)$ and the current $\rho(x)\, \partial_x \pi(x) = \sum_i \delta(x - x_i)\, p_i$. In the large N limit $\partial_x \pi(x)$ is a canonically conjugate to $\rho(x)$

$$[\rho(x),\, \partial_y \pi(y)] = i\, \partial_y \delta(x - y) \tag{2}$$

and the (1) can be written as the collective-field Hamiltonian given in the Schrödinger representation:

$$\begin{aligned}\mathcal{H} = \frac{1}{2} \int dx\, \rho(x)\, (\partial_x \pi)^2 + \frac{1}{2} \int dx\, \rho(x) \left(\frac{\lambda - 1}{2}\, \frac{\partial_x \rho}{\rho} + \lambda \int P\, \frac{\rho(y)}{x - y}\, dy \right)^2 \\ - \mu \int dx\, \rho(x) \ldots .\end{aligned} \tag{3}$$

Here, we have implemented the Lagrange multiplier μ which defines the energy scale of the problem. The dots stays for the terms which are not contributing in the leading order in N. To find the ground-state energy of our system, we assume that the corresponding collective-field configuration is static and has a vanishing momentum $\partial_x \pi$ therefore in the $1/N$ expansion is given by effective potential

$$V_{\text{eff}} = \frac{1}{2} \int dx\, \rho(x) \left(\frac{\lambda - 1}{2}\, \frac{\partial_x \rho(x)}{\rho(x)} + \lambda \int P\, dy\, \frac{\rho(y)}{x - y} \right)^2 . \tag{4}$$

The potential is positive semidefinite and therefore its contribution to the ground-state energy vanishes if there exist a positive solution of the first-order differential BPS type equation which we get when the expression inside brackets is zero. The solution is constant density configuration. There is a soliton solution which comes from appropriate prefactor in front of the wave function describing the quasihole. Corresponding solution satisfies extended BPS equation

$$\frac{\lambda - 1}{2}\, \frac{\partial_x \rho}{\rho} + \frac{1 - \lambda}{x} + \lambda \int P\, \frac{\rho(y)}{x - y}\, dy = 0 \tag{5}$$

with the solution which describes the hole in the condensate ρ_0:

$$\rho_s(x) = \rho_0\, \frac{x^2}{x^2 + b^2}\, , \qquad \rho_0 b \pi = \frac{1 - \lambda}{\lambda} \tag{6}$$

and with the "charge"

$$\int dx\, (\rho_s(x) - \rho_0) = \frac{\lambda - 1}{\lambda}\, . \tag{7}$$

Afterwards, in the context of duality we shall show existance of the M solitons solutions. Besides BPS type solutions there are moving solitons as solutions of equation of motion, which are also part of duality scheme, but they will be not considered here.

3. Duality

We shall show the existance of the duality in the collective field model for particles interacting with inverse-distance squared law by constucting the parent Hamiltonian for dual fields. (in [6] a corresponding master Lagrangian method was introduced).

To construct the parent Hamiltonian we shall use spectrum generating algebra $SU(1,1)$ (isomorphic to $\mathcal{O}(2,1)$) which is a symmetry of the action [10]. In this approach Hamiltonian [10] (up to similarity transformation) is one of the generators

$$T_+(\rho,\,\lambda) = -J^{1/2}\mathcal{H}J^{-1/2} = -\frac{1}{2}\int \rho(x)\,(\partial_x\pi)^2_{dx} - \frac{1}{2}\int \omega(\rho)\,\pi\,dx\,, \tag{8}$$

where ω is determining Jacobian $J(\rho)$:

$$\omega(\rho) = (\lambda-1)\,\partial_x^2\rho(x) + 2\lambda\,\partial_x\rho(x)\int P\,\frac{\rho(y)}{x-y}\,dy = \partial_x\left(\rho(x)\,\partial_x\,\frac{\partial\ln J}{\delta\rho(x)}\right). \tag{9}$$

Owing to the special invariance we introduce the generator

$$T_+(\rho,\,\lambda) = \frac{1}{2}\int x^2\rho(x)\,dx \tag{10}$$

and from SU(1,1) the algebra commutator

$$[T_+,\,T_-] = -2T_0 \tag{11}$$

we obtain scale invariance generator

$$T_0(\rho,\,\lambda) = -\frac{1}{2}\left[i\int x\rho(x)\,\partial_x\pi(x)\,dx + E_0\right] \tag{12}$$

with

$$E_0 = \lambda\,\frac{N(N-1)}{2} + \frac{N}{2}\,. \tag{13}$$

We can verify also

$$[T_0,\,T_\pm] = \pm T_\pm\,. \tag{14}$$

After having established the representation of SU(1,1) algebra, we show that the solutions for wave functionals which are eigen-functionals of (3) are determined assuming the zero-energy eigen-functions are known:

$$T_+(\rho,\,\lambda)\,P_m(\rho) = 0\,, \tag{15}$$

$$T_0(\rho,\,\lambda)\,P_m(\rho) = \mu_m P_m(\rho)\,, \qquad \mu_m = -\frac{1}{2}\,(m+E_0)\,. \tag{16}$$

The functionals $P_m(\rho)$ describes ground state and in order to obtain Calogero solution [1] it must be functional polynomial in $\rho(x)$.

Ansatz for the nonzero energy eigenstates in terms of generators can be written

$$\begin{aligned}\psi(T_-,\,T_0,\,T_+) &= \sum_{p,l,n} c_{pln}T_-^pT_0^lT_+^nP_m(\rho)\\ &= \psi(T_-,\,T_0)\,P_m(\rho) = \psi_m(T_-)\,P_m(\rho)\,.\end{aligned} \tag{17}$$

$\psi_m(T_-)$ is determined from eigenvalue equation

$$-T_+\psi_m(T_-)\ P_m(\rho) = E\psi_m(T_-)\ P_m(\rho)\,. \tag{18}$$

From (11) and (14) we can derive the formula

$$[T_+,\, f(T_-)] = T_- f''(T_-) - 2f'(T_-)\ T_0\,, \tag{19}$$

and from (18) and (19) for $\psi(T_-)$ we get the Bessel function solution:

$$\psi(T_-)\ P_m(\rho) \sim T_-^{-m-E_0}\,\mathcal{Z}_{m+E_0}(2\sqrt{ET_-})\ P_m(\rho)\,. \tag{20}$$

The prefactor in the wave functional can be *guessed ?* from solitons in collective-field-approach. We shall take (with appropriate ϵ regularization)

$$V^\kappa(x-z) = \exp\frac{\kappa}{2}\iint dx\ \rho(x)\ [\ln(x-z+i\epsilon) + \ln(x-z-i\epsilon)]\ m(z)\ dz\,. \tag{21}$$

Here we have introduced a new field $m(z)$ describing the density of the solitons. The duality between fields ρ, $\partial_x\pi_\rho$, m, $\partial_z\pi_m$ can be displayed by the following relations

$$i\,\partial_x\pi_\rho\, V^\kappa = \kappa\int\frac{m(z)\ dz}{x-z}\ V^\kappa\,, \tag{22}$$

$$i\,\partial_x\pi_m\, V^\kappa = \kappa\int\frac{\rho(x)\ dx}{2-x}\ V^\kappa\,, \tag{23}$$

$$\begin{aligned} T_+(\rho,\,\lambda)V^\kappa = \Bigg[&-\frac{\lambda}{\kappa}\ T_+\left(m,\,\frac{\kappa^2}{\lambda}\right) + \frac{(\lambda+\kappa)\ (\kappa-1)}{4}\iint P\ \frac{(\partial_x\rho(x)\ m(z) - \rho(x)\ \partial_z m(z))}{x-z} \\ &+\frac{\kappa\pi^2}{2}\int\rho(x)\ m(x)\ [\lambda\rho(x)+\kappa m(x)] + \text{ const}\Bigg]\ V^\kappa \end{aligned} \tag{24}$$

and similar duality relations for $T_0(\rho,\,\lambda)$ and $T_0(m,\,\frac{\kappa^2}{\lambda})$ denotes an operator with the same functional dependence on $m(x)$ as the $T_{\pm,0}$ on $\rho(x)$ with the coupling constant λ changed into κ^2/λ. Here we have manifest strong/weak coupling duality if we interchange the fields λ goes to $\frac{\kappa^2}{\lambda}$. Duality (24) is crucial. T_+ is Hamiltonian (up to similarity transformation) so this is duality relations for Hamiltonians. Let us define new $SU(1,1)$ generators. Adding the generators for particles and solitons

$$\begin{aligned} \mathcal{T}_+ = {} & T_+(\rho,\,\lambda) + \frac{\lambda}{\kappa}\ T_+\left(m,\,\frac{\kappa^2}{\lambda}\right) \\ & -\frac{(\kappa+\lambda)\ (\kappa-1)}{4}\iint[\partial_x\rho(x)\ m(z) - \rho(x)\ \partial_z m(z)]\ \frac{P}{x-z} \\ & -\frac{\kappa\pi^2}{2}\int\rho(x)\ m(z)\ [\lambda\rho(x)+\kappa m(x)] + \text{ const}\,, \end{aligned}$$

$$\mathcal{T}_0 = T_0(\rho,\,\lambda) + T_0\left(m,\,\frac{\kappa^2}{\lambda}\right),$$

$$\mathcal{T}_- = T_-(\rho,\,\lambda) + \frac{\kappa}{\lambda}\ T_-\left(m,\,\frac{\kappa^2}{\lambda}\right),$$

it can easily be checked that above generators satisfies the SU(1,1) conformal algebra. In terms of new generators duality relations turns out to be exactly the sufficient condition for solving the eigenvalue problem (15), (16), (18):

$$\mathcal{T}_+ V^\kappa = 0\,,$$
$$\mathcal{T}_0 V^\kappa = -\frac{(N+M)\,(\kappa+1)-2}{4}\,V^\kappa\,.$$

Now we can continue similary as in (18) and (19). We interpret the operator $\mathcal{T}_+$ as the parent Hamiltonian [5, 6]. After performing a similarity transformation of $\mathcal{T}_+$ we obtain [11]:

$$\begin{aligned}\mathcal{H}^P = {} & \frac{1}{2}\int \rho(x)\left\{\frac{\lambda-1}{2}\,\frac{\partial_x\rho(x)}{\rho(x)} + \lambda\int P\,\frac{\rho(y)}{x-y}\,dy + \kappa\int P\,\frac{m(z)}{x-2}\,dz\right\}^2 \\ & + \frac{\lambda}{2\kappa}\int m(z)\left\{\left(\frac{\kappa^2}{\lambda}-1\right)\frac{\partial_z m(z)}{m(z)} + \frac{\kappa^2}{\lambda}\int P\,\frac{m(y)\,dz}{z-y} + \kappa\int P\,\frac{\rho(y)\,dy}{z-y}\right\} \\ & + \frac{1}{2}\int \rho(x)\,(\partial_x\pi_\rho)^2 + \frac{\lambda}{2\kappa}\int m(z)\,(\partial_z\pi_m)^2 + \ldots\,.\end{aligned}$$

In BPS limit we can find solutions for m and ρ. We can look for finite number of solitons. Then there is a solution for $m(z) = \sum_\alpha \delta(z - z_\alpha)$ and for $\rho(x)$ we get microscopic description of solitons as given by (6). For two solitons solution is $\rho_{2s}(x) = \rho_0\,\frac{(x-x_0)^2\,(x+x_0)^2}{(x^2-b^2)\,(x^2-b^{*2})}$ with the charge $\int(\rho(x)-\rho_0) = 2\,\frac{\lambda-1}{\lambda}$. In dual description m describes solitons as particles and ρ is giving a microscopic description in terms of fields [12].

Acknowledgement

One of the authors (I. Andrić) would like to thank the organizers of the Symposium for the invitation and for creating a memorable and stimulating atmosphere at the meeting.

References

[1] F. CALOGERO, J. Math. Phys. **10** (1969) 2191, 2197; J. Math. Phys. **12** (1971) 419.
[2] I. ANDRIĆ, A. JEVICKI, and H. LEVINE, Nucl. Phys. **B215** (1983) 307.
[3] I. ANDRIĆ, V. BARDEK, and L. JONKE, Phys. Lett. **B357** (1995) 374.
[4] A. JEVICKI, Nucl. Phys. **B376** (1992) 75;
A. P. POLYCHRONAKOS, Phys. Rev. Lett. **74** (1995) 5153.
[5] I. ANDRIĆ and L. JONKE, *Duality and quasiparticles in the Calogero-Sutherland model: Some exact results*, hep-th 0010033, to be published in Phys. Rev. A.
[6] S. DESER and R. JACKIW, Phys. Lett. **B139** (1984) 2366.
[7] G. BARUCCHI and T. REGGE, Math. Phys. **18** (1977) 1149;
M. A. OLSHANETSKY and A. M. PERELOMOV, Phys. Rep. **94** (1983) 313.
[8] G. AMELINO-CAMELIA, Mod. Phys. Lett. **A11** (1996) 1861.
[9] A. JEVICKI and B. SAKITA, Nucl. Phys. **B165** (1980) 511.
[10] V. DE ALFARO, S. FUBINI, and G. FURLAN, Nuovo Cimento **A34** (1976) 569.
[11] I. ANDRIĆ and D. JURMAN, unpublished.
[12] B. JULIA, this proceedings.

Fortschr. Phys. **50** (2002) 5–7, 680–685

Monopoles in Space-time Noncommutative Born-Infeld Theory

PAOLO ASCHIERI

Ludwig-Maximilians-Universität München, Sektion Physik
Theresienstr. 37, D-80333 München, Germany
aschieri@theorie.physik.uni-muenchen.de

Abstract

We transform static solutions of space-noncommutative Dirac-Born-Infeld theory (DBI) into static solutions of space-time noncommutative DBI. Via Seiberg-Witten map we match this symmetry transformation with a corresponding symmetry of commutative DBI. This allows to: 1) study new BPS type magnetic monopoles, with constant electric and magnetic background and describe them both in the commutative and in the noncommutative setting; 2) relate by S-duality space-noncommutative magnetic monopoles to *space*-noncommutative electric monopoles.

1. Introduction

Dirichlet branes effective actions can be described by noncommutative gauge theories, the noncommutativity arising from a nonzero constant background NS B field, see [1] and references therein. In fact, when $B \neq 0$, the effective physics on the D-brane can be described both by a commutative gauge theory $\mathcal{L}(F+B)$ and by a noncommutative one $\hat{\mathcal{L}}(\hat{F})$, where

$$\hat{F}_{\mu\nu} = \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu - i[\hat{A}_\mu \underset{\star}{,} \hat{A}_\nu] \tag{1}$$

and $\star$ is the Moyal star product; on coordinates $[x^\mu \underset{\star}{,} x^\nu] = x^\mu \star x^\nu - x^\nu \star x^\nu = i\Theta^{\mu\nu}$. The noncommutativity parameter Θ depends on B and on the metric on the D-brane. The commutative/noncommutative descriptions are related by Seiberg-Witten map (SW map) [1]. Initially space-noncommutativity has been considered ($\Theta^{ij} \neq 0$ i.e. $B_{ij} \neq 0$) then theories also with time noncommutativity ($\Theta^{0i} \neq 0$ i.e $B_{0i} \neq 0$) have been studied [2]. It turns out that unitarity of noncommutative Yang-Mills theory (NCYM) holds only if Θ is space-like or light-like i.e. only if the electric and magnetic components of Θ (or B) are perpendicular and if the electric component is not bigger in magnitude than the magnetic component. These are precisely the NCYM theories that can be obtained from open strings in the decoupling limit $\alpha' \to 0$ [3]. In this talk we consider these two kinds of space-time noncommutativity.

In Section 2 we show that for any space-noncommutative static solution we can turn on time-noncommutativity and obtain a static solution with space-time noncommutativity. This holds in particular for solutions of noncommutative Dirac-Born-Infeld-theory (NCDBI). Via SW map we obtain the nontrivial action of this symmetry in the corresponding commutative DBI theory and show that it is a rotation (boost) between the time component A_0 of the gauge potential and the worldvolume D-brane coordinates x^i. This boost, first studied in [4], is similar to the target space rotation that relates the linear monopole to the nonlinear monopole [4, 5]. In Section 3 we study BPS solutions of NCDBI and DBI theories with

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

both electric and magnetic background. In [6, 7] solutions to the BPS equations of noncommutative electromagnetism (NCEM), and of NCDBI, with just space-noncommutativity are found. The solution in [6] describes a smeared monopole connected with a string-like flux tube and is interpreted as a D1-string ending on a D3-brane with constant magnetic field background. We see that this solution remains a BPS solution also when we turn on time-noncommutativity. The corresponding commutative configuration is also found: It is a new BPS configuration. It describes a monopole plus string in a background that is both electric and magnetic. The monopole has the fundamental magneton charge and the string tension is that of a D1-string, this strongly suggests that we have a D1-D3 brane system with both electric and magnetic background. The D1-string tension is also matched with the string tension of the corresponding space-time noncommutative BPS monopole.

Finally in Section 4 we address the issue of duality rotations in NCDBI and NCEM [8, 9]. At the field theory level duality is present only if Θ is light-like i.e. the magnetic and electric component of Θ are perpendicular and equal in magnitude [10, 11]. For Θ space-like as shown in [8] we do not have noncommutative gauge theory self-duality and the S-dual of space-like NCEM is a noncommutative open string theory decoupled from closed strings. A main point here is that under S-duality the magnetic background is mapped into an electric one, and this background does not lead to a field theory in the $\alpha' \to 1$ limit. Since Θ must be light-like, it may seem that duality rotations have a very restricted range of application. This is not the case because the symmetry we present in Sections 2 and 3 changes the background too. It turns out that it is possible to compose duality rotations with this symmetry, we are thus able to consider duality rotations with background fixed and arbitrary. In particular we briefly discuss the S-dual of the D1-string D3-brane configuration of [6]; it describes an electric monopole plus string in a magnetic background, possibly a fundamental string ending on a D3-brane in the presence of a constant magnetic field.

2. Gauge Theory with Space-Time Noncommutativity

We use the following notations: Θ is a generic constant noncommutativity tensor, we have $[x^\mu \star x^\nu] = i\Theta^{\mu\nu}$; θ is just a space-noncommutativity tensor θ^{ij}, $\theta^{0i} = 0$; θ^ε is a space-time noncommutativity tensor obtained from θ adding electric components, $\theta^{\varepsilon^{ij}} = \theta^{ij}$, $\theta^{\varepsilon^{0i}}$. In three vector notation the electric and magnetic components of θ, respectively θ^ε, are $(0, \boldsymbol{\theta})$ and $(\boldsymbol{\varepsilon}, \boldsymbol{\theta})$. The background fields corresponding to Θ, θ, θ^ε are B, b, b^ε.

Consider a noncommutative Lagrangian $\hat{\mathcal{L}}^\theta = \hat{\mathcal{L}}(\hat{F}, \hat{\phi}, G, \star_\theta)$ where $\hat{\phi}$ are scalar fields and G is the metric. The equations of motion (EOM) for $\hat{\mathcal{L}}^\theta$ read

$$f_\alpha(\hat{F}, \hat{\phi}, G, \star_\theta) = 0\,, \tag{2}$$

where f_α are functions of the noncommutative fields and their derivatives. We notice that a static solution of the $\hat{\mathcal{L}}^\theta$ EOM (2) is also a static solution of (2) with θ^ε instead of θ, i.e. it is a static solution of the $\hat{\mathcal{L}}^{\theta^\varepsilon}$ EOM. Indeed the star products $\star_\theta$ and $\star_{\theta^\varepsilon}$ act in the same way on time independent fields. Moreover the energy and charges of the solution are invariant. A similar θ-θ^ε symmetry property holds if the fields are independent from a coordinate x^μ (not necessarily t). This θ-θ^ε symmetry property of static solutions can be used to construct moving solutions of a space-noncommutative theory $\hat{\mathcal{L}}^\theta$ from static solutions of the same theory $\hat{\mathcal{L}}^\theta$. Indeed, given θ, if we turn on an electric component such that $\boldsymbol{\varepsilon} \perp \boldsymbol{\theta}$ and $|\boldsymbol{\varepsilon}| < |\boldsymbol{\theta}|$, then with a Lorentz boost we can transform this new θ^ε into a space-like θ' proportional to the initial θ. Rescaling $\theta' \to \theta$ we thus obtain a solution (moving with constant velocity) of the space-noncommutative Lagrangian $\hat{\mathcal{L}}^\theta$.

We now use SW map and study how the θ-θ^ε symmetry acts in the commutative theory. We have to consider the two SW maps SW^θ and $\mathrm{SW}^{\theta^\varepsilon}$. In general a static solution $\hat{\phi}$, $\hat{A}_\mu$ is mapped by SW^θ and $\mathrm{SW}^{\theta^\varepsilon}$ into two different commutative solutions, however if $\hat{A}_0 = 0$ then $\mathrm{SW}^\theta = \mathrm{SW}^{\theta^\varepsilon}$. This can be seen from the index structure of SW map. In general we have

$$\begin{aligned} A_\mu &= \hat{A}_\mu + \sum_{n>s} \left(\Theta^{(n)}\, \partial^{(n+s)} \hat{A}^{(n-s)}\, \hat{A}\right)_\mu , \\ \phi &= \hat{\phi} + \sum_{n>s} \Theta^{(n)}\, \partial^{(n+s)} \hat{A}^{(n-s)}\, \hat{\phi} , \end{aligned} \tag{3}$$

where the number of times n, $n+s$, $n-s$ that Θ, ∂, $\hat{A}$ appear is dictated by dimensional analysis. In (3) we do not specify which ∂ acts on which $\hat{A}$ and we do not specify the coefficients of each addend. Because of the index structure we notice that Θ^{0i} never enters (3) if $\hat{\phi}, \hat{A}$ are time independent and $\hat{A}_0 = 0$. The commutative fields ϕ, A_i corresponding to $\hat{\phi}$, $\hat{A}_i$ $(i \neq 0)$ are solution of both $\mathcal{L}^\theta$ and $\mathcal{L}^{\theta^\varepsilon}$. Here $\mathcal{L}^\theta$ and $\mathcal{L}^{\theta^\varepsilon}$ are the commutative Lagrangians corresponding to $\hat{\mathcal{L}}^\theta$ and $\hat{\mathcal{L}}^\theta \varepsilon$ via SW map. In the case of the DBI Lagrangian with a scalar field ϕ, $\mathcal{L}^\theta$ and $\mathcal{L}^{\theta^\varepsilon}$ reads

$$\mathcal{L}_{\mathrm{DBI}}(F+b,\, \phi,\, g,\, g_s) = \frac{-1}{\alpha'^2 g_s} \sqrt{-\det\left(g + \alpha'(F+b) + \alpha'^2\, \partial\phi_a\, \partial\phi^a\right)} \tag{4}$$

and $\mathcal{L}_{\mathrm{DBI}}(F+b^\varepsilon,\, \phi,\, g^\varepsilon,\, g_s)$. We should write g_s^ε instead of g_s in this last expression, however we can rescale G_s and thus impose the invariance of the closed string coupling constant g_s. The relation between closed and open string parameters is given by (see [1]):

$$(g + \alpha' B)^{-1} = G^{-1} + \Theta/\alpha' , \qquad G_s = g_s \sqrt{\det G \det\left(g + \alpha' B\right)^{-1}} .$$

In order to have a more explicit formulation of the θ-θ^ε symmetry, from now on we set the noncommutative open string metric $G = \eta = \mathrm{diag}\,(-1, 1, 1, 1)$ and we consider (rigid) coordinate transformations $x \to x''$ and $x \to x'$ that respectively orthonormalize the closed string metrics g^ε and g, while preserving time independence of the transformed ϕ, A fields. The result is that if ϕ', $\mathcal{A}'_i$, $\mathcal{A}'_0 = 0$ is a static solution of $\mathcal{L}_{\mathrm{DBI}}$ then ϕ'', $\mathcal{A}''_\mu$ is a new static solution [11]. Here $\mathcal{A}'$ is the gauge potential of $\mathcal{F}' = F' + b'$ in the x' basis and

$$\phi''(x'') = \phi'(x') , \qquad \mathcal{A}''_i(x'') = \Upsilon^j_i \mathcal{A}'_j(x') , \qquad \mathcal{A}''_0(x'') = e'' x''_2 \tag{5}$$

with $x'' = \Upsilon^{-1} x'$; the nonvanishing components of Υ are

$$\Upsilon_{00} = \Upsilon_{22} = \sqrt{1 - \alpha'^2 e''^2} , \qquad \Upsilon_{11} = \Upsilon_{33} = 1 , \qquad \Upsilon_{01} = \alpha'^2 e''\, b'$$

and we have turned on time-noncommutativity just in the x''_2 direction.

Which is the symmetry of commutative $\mathcal{L}_{\mathrm{DBI}}$ that underlies this family of solutions? We split $\mathcal{F}'$ in its electric field $\mathcal{E}'$ and magnetic induction $\mathcal{B}'$ components. We then consider the Legendre transformation of $\mathcal{L}_{\mathrm{DBI}}$

$$\check{H}(\mathcal{E}',\, \mathcal{H}',\, \phi') = \frac{1}{g_s}\, \mathcal{B}' \cdot \mathcal{H}' + \mathcal{L}_{\mathrm{DBI}} , \quad \text{where} \quad \mathcal{H}'_i = -g_s\, \frac{\partial \mathcal{L}_{\mathrm{DBI}}}{\partial \mathcal{B}'^i} . \tag{6}$$

For time independent fields we have that the EOM imply $\mathcal{H}' = -\nabla'\chi'$ and $\mathcal{E}' = -\nabla'\psi'$, $(\psi' = -\mathcal{A}'_0)$. As shown in [4] it follows that $\check{H}(\psi',\, \chi',\, \phi')$ is the action of a space-like 3-brane immersed in a target space of coordinates $X'^A = \{\alpha'\psi',\, \alpha'\chi',\, \alpha'\phi',\, x'^i\}$ and metric $\eta = \mathrm{diag}\,(-1, -1, 1, 1, 1, 1)$

$$\int d^3x'\, \check{H} = \frac{-1}{g_s \alpha'^2} \int d^3x' \sqrt{\det\left(\eta_{AB}\, \frac{\partial X'^A}{\partial x'^i}\, \frac{\partial X'^B}{\partial x'^j}\right)} . \tag{7}$$

It is the $SO(2,4)$ symmetry [4] of this static gauge action that is relevant in our context: Consider the Lorentz transformation $Y^A = \Lambda^A_B X'^B$ (where $X'^i = x'^i$) and express Y^A as $Y^A = Y^A(y^i)$ (where $Y^i = y^i$) so that we are still in static gauge; the action (7) is invariant under $X'^A(x'^i) \to Y^A(y^i)$. In particular a boost in the $\alpha'\psi'$, x'_2 plane with velocity $\beta = -\alpha'\, e''$ gives (5) [11].

3. BPS Solutions for (NC)DBI with a Scalar Field

A BPS solution of DBI theory in the x-reference system with metric g and background $b = \theta/(\alpha'^2 + \theta^2)$, $\theta = -\theta^{12}$, all others $\theta^{\mu\nu} = 0$ is given by

$$\phi = \frac{-\theta}{\alpha'^2}\, x^3 - \frac{1}{2r}\,, \qquad \mathcal{B}_i = -\partial_i \phi\,; \qquad r^2 = g_{ij}x^i x^j = (x^3)^2 + \frac{(x^1)^2 + (x^2)^2}{1 + \theta^2/\alpha'^2}\,, \tag{8}$$

where $\frac{1}{2r} = \frac{q_m}{4\pi r}$ with q_m the magneton charge $q_m = 2\pi$. This solution describes a D1-string ending on a D3-brane. Because of the magnetic backgound field b on the brane the string is not perpendicular to the brane, in (8) the string is vertical and the brane is tilted w.r.t. the horizontal direction. The magnetic force acting on the end of the D1-string is compensated by the tension of the D1-string.

A BPS solution of NCDBI with space-like noncommutativity given by $\theta^{12} = -\theta$, (all others $\theta^{\mu\nu} = 0$) has been studied in [6]. It is a static solution with $\hat{A}_0 = 0$. This noncommutative BPS solution is caracterized by a noncommutative string tension and a magneton charge. It is expected to correspond, via SW map (and the target space rotation relating the linear monopole to the nonlinear one [5]), to (8). A first evidence is the correspondence between the tension of the noncommutative string and that of the D1-string [6], then in [7] the spectrum of small fluctuations around (a limit of) this solution is studied and found in agreement with the expectations from string theory.

We now discuss what happens to these commutative/noncommutative BPS solutions when we apply the θ-θ^ε symmetry. The result [11] is that we obtain two new BPS solutions that describe a D1-string ending on a D3-brane with both an electric and magnetic background. The noncommutative one is obtained simply writing θ^ε instead of θ (and rescaling the open string coupling constant $G_s \to G_s^\varepsilon$ since we keep g_s invariant), it still satisfies the noncommutative BPS equations $\hat{\mathcal{B}}_j = -\hat{D}_j\hat{\phi}$ with $\hat{D}_j\hat{\phi} = \partial_j\hat{\phi} - i[\hat{A}_j \stackrel{\star}{,} \hat{\phi}]$. The commutative solution is most easily written in the orthonormal reference system x'':

$$\begin{aligned} &\mathcal{B}''_2 = -\gamma\, \partial''_2\phi''\,, \\ &\mathcal{B}''_q = -\gamma^{-1}\, \partial''_q\phi''\,, \qquad q = 1,3 \\ &\mathcal{E}''_2 = e''\,, \qquad \mathcal{E}''_q = 0 \end{aligned} \tag{9}$$

with $\gamma^{-1} = \sqrt{1 - \alpha'^2\, e''^2}$ and

$$\phi'' = -\gamma b'' x''^3 - \frac{1}{2R}\,, \qquad R^2 \equiv (x''^1)^2 + \gamma^{-2}(x''^2)^2 + (x''^3)^2\,. \tag{10}$$

Eq. (9) are obtained from the (x'-reference system) BPS equations $\mathcal{B}'_i = -\partial_i\phi'$ via the boost Υ, cf. (5). The nonvanishing components of $b''_{\mu\nu}$ and $\theta^{\mu\nu}$ are $e'' = -b''_{02} = \frac{-\varepsilon}{\alpha'\sqrt{\alpha'^2+\theta^2}}$, $b'' = b''_{12} = \frac{\theta}{\alpha'^2}\sqrt{1 - \frac{\varepsilon^2}{\alpha'^2+\theta^2}}$, $\theta = -\theta^{12}$, $\varepsilon = \theta^{\varepsilon^{02}}$. A solution of (9) has an

energy

$$\Sigma'' = \Sigma' + \frac{1}{g_s} \int d^3x''\ \mathcal{E}''_i \mathcal{D}''^{\,i}\,, \tag{11}$$

where Σ' is the energy of the corresponding solution of $\mathcal{B}'_i = \partial_i \phi'$. We see that the energy is of BPS type, indeed it is the sum of the old BPS energy Σ' plus the topological charge $Z''_e = \frac{1}{g_s} \int d^3x'' \mathcal{E}''_i \mathcal{D}''^{\,i} = -\frac{1}{g_s} \int d^3x''\ \partial_i \psi''\ \mathcal{D}''^{\,i}$. The explicit value of $\mathcal{D}''^{\,i} \equiv g_s \frac{\partial \mathcal{L}_{\rm DBI}}{\partial \mathcal{E}''_i}$ is $\mathcal{D}''^{\,2} = e''\gamma,\ \mathcal{D}''^{\,1} = \mathcal{D}''^{\,3} = 0$. We can also write $\Sigma'' = \frac{1}{\alpha'^2 g_s} \int d^3x'' + |Z''_m| + \frac{\gamma - 1}{g_s} \int d^3x''$, and recognize the brane tension, the topological charge $Z''_m = \int d^3x''\ \partial_i \phi''\ \mathcal{B}''^{\,i}$ and the energy of just the electric field e'' in DBI theory. We also have that the magnetic charge and the string tension associated to solution (9), (10) are those of a D1-string as we expect from a BPS state. Notice that the shape of the funnel representing this D1-string is no more symmetric in the $x''_1,\ x''_2$ directions. A section determined by $\phi'' = $ const, $x''_3 = 0$ is an ellipsis in the $x''_1,\ x''_2$ plane. The ratio between the ellipsis axes is given by γ. One can project the D1-string on the D3-brane and consider the tension of this projected string: It matches the tension associated to the corresponding noncommutative BPS solution.

4. Dual String-brane Configuration

If we duality rotate the D1-D3 brane configuration (9, 10) we obtain a soliton solution that describes a fundamental string ending on a D3-brane with electric and magnetic background. Under a $\pi/2$ duality rotation we have (we set $2\pi = 1$, recall also $g''_{\mu\nu} = \eta_{\mu\nu}$)

$$g^D_s = \frac{1}{g_s}\,, \qquad g''^D_{\mu\nu} = \frac{1}{g_s}\ \eta_{\mu\nu}\,, \qquad \phi''_D = \left(\frac{1}{g_s}\right)^{\frac{1}{2}} \phi'' \tag{12}$$

the dual of solution (9), (10) is given by (12) and

$$\mathcal{A}''^D_0 = -\left(\frac{1}{g_s}\right)^{\frac{1}{2}} \phi''_D\,, \qquad \mathcal{A}''^D_1 = -\frac{1}{g_s}\ \gamma e'' x''^3\,, \qquad \mathcal{A}''^D_2 = \mathcal{A}''^D_3 = 0\,. \tag{13}$$

Is there a noncommutative field theory description of the F1-D3 system? Since NCDBI and its $\alpha' \to 1$ limit, NCEM, admit duality rotations only if θ is light-like, it seems that we have a F1-D3 system only if we consider a light-like background. This light-like condition may appear a strong constraint. Actually, using the θ-θ^ε symmetry we are *not* bound to consider only this restrictive case of light-like background. Indeed for any space-noncommutative static solution we can turn on time-noncommutativity and obtain a static solution with light-like noncommutativity. We can then apply a duality rotation, switch off the time-noncommutativity and thus obtain a new solution of the original pure space-noncommutative theory.

In particular in order to obtain the duality rotated configuration of the one described in [6], we consider the corresponding commutative configuration (8), that in the x' orthonormal frame reads $\mathcal{B}'_i = -\partial'_i \phi',\ \phi' = -\frac{1}{\alpha'}\theta x'^3 - \frac{1}{2r'},\ r'^2 \equiv (x'^1)^2 + (x'^2)^2 + (x'^3)^2$. In order to have a light-like background we turn on a constant electric field keeping here fixed G_s besides the open string metric $G = \eta$ (therefore here $g_s \to g^\varepsilon_s \neq g_s$). We then obtain (9) and (10) with $e'' = -b''$ (i.e. $\theta = \varepsilon$). Next we duality rotate this solution and obtain (12) and (13) with $g^{\varepsilon D}_s,\ g^\varepsilon_s$ instead of $g^D_s,\ g_s$. Finally we go back to the original x-reference system (the $x \to x''$ coordinate transformation commutes with duality rotations), we apply SW map

and arrive at the noncommutative fields $\hat{A}^D$, $\hat{\phi}_D$ with open string coulping constant, metric and light-like noncommutativity given by

$$G_s^D = \frac{1}{G_s}\,, \qquad G_{\mu\nu}^D = \frac{1}{G_s}\,\eta_{\mu\nu}\,, \qquad \theta_D^{\varepsilon\mu\nu} = \frac{1}{2}\,G_s\,\epsilon^{\mu\nu\rho\sigma}\,\theta_{\rho\sigma}^{\varepsilon}\,. \tag{14}$$

The noncommutative fields $\hat{A}^D$, $\hat{\phi}_D$ correspond to a fundamental string ending on a D3-brane with light-like background. These fields solve also $\hat{\mathcal{L}}_{\mathrm{DBI}}(\hat{F}, \hat{\phi}, G^D, G_s^D, \star_{\theta_D})$ where θ_D ($\theta_D^{13} = G_s\theta$, all others $\theta_D^{\mu\nu} = 0$) is just the space part of θ_D^{ε}; they describe an electric monopole with a string attached. Since the noncommutative string tension and charges are invariant under $\theta_D^{\varepsilon} \to \theta_D$, the $\hat{A}^D$, $\hat{\phi}_D$ fields are a good candidate to describe an F-string ending on a D3-brane with constant magnetic background.

Acknowledgements

I wish to thank the organizers and Alexander von Humboldt-Stiftung for the stimulating atmosphere of the conference. This work has been supported by a Marie Curie Fellowship of the European Community programme IHP under contract number MCFI-2000-01982.

References

[1] N. SEIBERG and E. WITTEN, JHEP**9909** (1999) 032 [hep-th/9908142].
[2] N. SEIBERG, L. SUSSKIND, and N. TOUMBAS, JHEP**0006** (2000) 044 [hep-th/0005015].
J. L. BARBON and E. RABINOVICI, Phys. Lett. B **486** (2000) 202 [hep-th/0005073].
J. GOMIS and T. MEHEN, Nucl. Phys. B **591** (2000) 265 [hep-th/0005129].
[3] O. AHARONY, J. GOMIS, and T. MEHEN, JHEP**0009** (2000) 023 [hep-th/0006236].
[4] G. W. GIBBONS, Nucl. Phys. B **514** (1998) 603 [hep-th/9709027].
[5] K. HASHIMOTO and T. HIRAYAMA, Nucl. Phys. B **587** (2000) 207 [hep-th/0002090].
S. MORIYAMA, Phys. Lett. **B485** (2000) 278 [hep-th/0003231].
S. MORIYAMA, JHEP **0008** (2000) 014 [hep-th/0006056].
K. HASHIMOTO, T. HIRAYAMA, and S. MORIYAMA, JHEP **0011** (2000) 014 [hep-th/0010026].
[6] D. J. GROSS and N. A. NEKRASOV, JHEP **0007** (2000) 034 [hep-th/0005204].
[7] D. J. GROSS and N. A. NEKRASOV, JHEP **0010** (2000) 021 [hep-th/0007204].
[8] R. GOPAKUMAR, J. MALDACENA, S. MINWALLA, and A. STROMINGER, JHEP**0006** (2000) 036 [hep-th/0005048].
[9] O. J. GANOR, G. RAJESH, and S. SETHI, Phys. Rev. D **62** (2000) 125008 [hep-th/0005046].
S. REY and R. VON UNGE, Phys. Lett. B **499** (2001) 215 [hep-th/0007089].
J. X. LU, S. ROY, and H. SINGH, Nucl. Phys. B **595** (2001) 298 [hep-th/0007168].
[10] P. ASCHIERI, Mod. Phys. Lett. A **16** (2001) 163 [hep-th/0103150].
[11] P. ASCHIERI, Nucl. Phys. **617** (2001) 321 [hep-th/0106281].

Fortschr. Phys. **50** (2002) 5–7, 686

Canonical Quantization, Twistors and Relativistic Wave Equations

IWO BIAŁYNICKI-BIRULA

Center for Theoretical Physics, Al. Lotników 32/46, 02-668 Warsaw, Poland
and Institute of Theoretical Physics, Warsaw University
birula@cft.edu.pl

Full version of this paper will be published in the special issue of Fortschritte der Physik devoted to Quantum Challenges

Abstract

Relativistic wave equations describing massless particles (Maxwell equations, Weyl equation) have never been derived directly from classical physics by a straightforward quantization procedure. However, it turns out that there is a road that leads from classical dynamics of a fictitious point particle living in a four-dimensional configuration space to relativistic wave equations. The procedure that accomplishes that is closely related to the twistor formalism of Penrose, but it offers a new interpretation of twistors within the standard phase-space formulation of quantum mechanics. The wave functions that arise in this approach may be subjected to all conformal transformations. In the special case of space-time translations, one obtains a somewhat different version of the Penrose transform. Having all the formal tools of canonical quantization at our disposal, one may introduce standard structures of this theory: position-momentum duality realized by the Fourier transformation, uncertainty relations, Schrödinger and Heisenberg pictures, Wigner functions, etc. The description of the quantum states in terms of the Wigner function defined on the eight-dimensional phase-space seems to be especially promising. Wigner functions possess very simple transformation properties under the full conformal group and this leads in a natural way to a tomography of twistor Wigner functions based on a suitably generalized Radon transform.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Fortschr. Phys. **50** (2002) 5–7, 687–693

Geometry of Non-supersymmetric Strings

ANAMARÍA FONT

Dept. de Física, Fac. de Ciencias, Univ. Central de Venezuela
Apdo. 20513, Caracas 1020-A, Venezuela
afont@fisica.ciens.ucv.ve

Abstract

We analyze the backreaction of dilaton tadpoles on the background of non-supersymmetric strings. In a prototype example we find explicit solutions for the warped metric and the dilaton profile and discuss their consequences. Our analysis confirms a process of spontaneous compactification in non-supersymmetric strings.

1. Introduction

In open string models containing lower dimensional D-branes, extra large transversal directions can give rise to a large Planck scale while leaving the string scale essentially a free parameter. Indeed, the string scale can be lowered to the TeV range so that supersymmetry is not necessarily needed to protect the gauge hierarchy. It is thus natural to explore non-supersymmetric string models. Of particular interest are models with brane supersymmetry breaking [1, 2], where the tree level bulk still preserves some supersymmetry while supersymmetry is broken on the D-branes. All these models have a non-vanishing dilaton tadpole, implying that the string equations of motion are not satisfied by a factorized metric and a constant dilaton. In particular, for the open string models the dilaton tadpoles appear at disc level and their backreaction should be taken into account to come closer to the true quantum vacuum of the theory. In the deformed background the tadpole has disappeared [3]. To find the true perturbative quantum vacuum, one would need to solve the string equations of motion to all string loop levels. However, for non-supersymmetric strings this is far beyond the reach of computational power.

In this work we wish to examine whether taking the tadpole backreaction into account, it is still possible to disentangle the Planck scale from the string scale. This is a non-trivial issue as the leading order relations

$$M_{Pl}^2 \sim \frac{M_s^8 V_d V_{6-d}}{g_s^2}\,, \qquad \frac{1}{g_{YM}^2} \sim \frac{M_s^d V_d}{g_s} \tag{1}$$

for the four dimensional Planck mass and the gauge couplings get modified. In (1) V_d denotes the volume longitudinal to the D-branes and V_{6-d} the volume transversal to the D-branes.

In Ref. [4], the effective equations of motion for the string background in the non-supersymmetric USp(32) Sugimoto model [1] in ten dimensions were solved. In the resulting warped metric, the four-dimensional Planck scale and gauge coupling take the values

$$M_{Pl}^2 \sim M_s^{\frac{17}{2}} V_5 R_c^{\frac{3}{2}}\,, \qquad \frac{1}{g_4^2} \sim M_s^{\frac{11}{2}} V_5 R_c^{\frac{1}{2}}\,, \tag{2}$$

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

with R_c denoting the effective size of the spontaneously compactified direction x_9. With gauge coupling of order one, the relations (2) imply $M_{Pl}^2 \sim M_s^3 R_c$, so that even with spacetime filling $\overline{\text{D9}}$-branes a large extra dimension scenario is possible (at least in the next-to-leading order approximation).

We extend the analysis of [4] to the case of D-branes allowing transversal directions. One simple such situation is provided by a T-dual version of the USp(32) model that contains $\overline{\text{D8}}$-branes as we will explain in Section 2. In Section 3 we present our solution that leads to finite couplings and in Section 4 we discuss the results.

2. A Model with $\overline{\text{D8}}$-branes

We begin with the prototype open string model featuring brane supersymmetry breaking, namely the USp(32) Sugimoto model [1] which is a non-supersymmetric version of the Type I string. Whereas the supersymmetric Type I string contains orientifold planes of negative tension and RR charge, the Sugimoto model contains orientifold planes of positive tension and RR charge. This modification does not change the Klein bottle amplitude at all, implying that at closed string tree level the bulk still preserves supersymmetry. However, in order to cancel the dangerous RR tadpole one has to introduce 32 anti-D9-branes, which of course also have positive tension. Thus, even though the RR charge is cancelled, the background contains positive tension branes generating a non-vanishing dilaton tadpole. Moreover, the Möbius amplitude is non-vanishing, so that the model explicitly breaks supersymmetry. Note that there does not exist any way of cancelling the RR tadpole by a supersymmetric configuration of D-branes.

To obtain lower dimensional branes we first compactify the tenth dimension, denoted y, in a circle of length $2L$. We next perform a T-duality along y, i.e. we combine the world sheet parity with the reflection $y \to -y$. In this way we find a model with two positively charged O8-planes located at the two fixed points of the reflection. We cancel the RR charge locally by putting 16 anti-D8 branes on each fixed point, chosen to be at $y = L/2,\ 3L/2$. In one loop, $e^{0\Phi}$, order, this appears to be a stable configuration. Since the Klein-bottle and the annulus amplitude vanish, the leading order force between two O8-planes, respectively two $\overline{\text{D8}}$-branes vanish. Only the Möbius amplitude is non-vanishing and leads to an attractive force between an O8-plane and a $\overline{\text{D8}}$-brane.

As in the original USp(32) model, we are left with non-zero dilaton tadpoles due to the positive tension localized at the two fixed points. In string frame the effective action for the metric and the dilaton is

$$\begin{aligned} S_S = \frac{M_S^8}{2} \int d^{10}x\, \sqrt{-G}\, e^{-2\Phi}\, [R + 4(\partial\Phi)^2] \\ - 32T \int d^{10}x\, \sqrt{-g}\, e^{-\Phi} \left[\delta\left(y - \frac{L}{2}\right) + \delta\left(y - \frac{3L}{2}\right)\right], \end{aligned} \tag{3}$$

where $g_{ab} = \delta_a^M\, \delta_b^N G_{MN}$ denotes the 9-dimensional metric induced on the branes. We take M, N to run over all spacetime and a, b over the longitudinal coordinates. Note that the brane tension $32T$ is the same on both fixed points and that due to the cancelled RR-Flux we can set the RR nine-form to zero.

To study a more general set-up, we will compactify the string theory on a $(8 - D)$ dimensional torus of volume V_{8-D}. Thus, in string frame we split the metric as

$$ds_{10,S}^2 = ds_{D+2,S}^2 + \sum_{m,n=D+3}^{10} \delta_{mn}\, dx^m\, dx^n\,. \tag{4}$$

We transform the resulting effective action via $G_E = e^{-\frac{4}{D}\Phi}\, G_S$ into Einstein frame to obtain

$$S_E = \frac{M_S^8\, V_{8-D}}{2} \int d^{D+2}x\, \sqrt{-G}\, \left[R - \frac{4}{D}\, (\partial\Phi)^2 \right] - 32 T V_{8-D} \int d^{D+2}x\, \sqrt{-g}\, e^{\frac{D+2}{D}\Phi}\, \Delta(y)\,, \tag{5}$$

where $\Delta(y) = \delta(y - \frac{L}{2}) + \delta(y - \frac{3L}{2})$.

3. Solutions

In this section we will construct solutions of the equations of motion resulting from the action (4). These are

$$\begin{aligned} E_{MN} &= -\frac{4}{D} \left(\frac{1}{2}\, G_{MN} G^{PQ}\, \partial_P \Phi\, \partial_Q \Phi - \partial_M \Phi\, \partial_N \Phi \right) - \lambda g_{ab} \delta^a_M \delta^b_N \sqrt{\frac{g}{G}}\, e^{\frac{D+2}{D}\Phi}\, \Delta(y)\,, \\ \partial_M(\sqrt{-G}\;\, G^{MN}\, \partial_N \Phi) &= \frac{D+2}{4}\, \lambda\, \sqrt{-g}\, e^{\frac{D+2}{D}\Phi}\, \Delta(y)\,, \end{aligned} \tag{6}$$

where $E_{MN} = R_{MN} - \frac{1}{2} G_{MN} R$ and we have introduced $\lambda = 32T/M_s^8$.

We can show that there does not exist a solution to these equations with $(D+1)$ dimensional Poincaré invariance. Therefore, the best we can try is to look for solutions with D-dimensional Poincaré invariance, for which we make the following warped ansatz

$$ds^2_{D+2} = e^{2M(r,y)}\, \eta_{\mu\nu}\, dx^\mu\, dx^\nu + e^{2N(r,y)}\, (dy)^2 + e^{2P(r,y)}\, (dr)^2\,. \tag{7}$$

By redefining the coordinates r and y we are free to choose $N = P$. We also assume that Φ depends only on r and y. Note that this ansatz assumes that the cosmological constant in the effective D-dimensional theory vanishes.

As usual, one first constructs solutions of Eq. (6) in the bulk and then imposes the jump conditions due to the branes at $y = L/2,\ 3L/2$. In order to satisfy these jump conditions we are led to choose warp factors that depend separately on r and y and likewise for Φ. More precisely, we take

$$M(y,\, r) = A(y) + X(r)\,, \qquad P(y,\, r) = C(y) + Y(r)\ , \qquad \Phi(y,\, r) = \varphi(y) + \chi(r)\,. \tag{8}$$

The jump conditions implied by the delta functions will thus require discontinuous φ', A' and C', where we use primes and dots for derivatives with respect to y and r respectively. Moreover, we impose $\dot{Y} = -\frac{D+2}{D}\, \dot{\chi}$, so that there is no r dependence on the right hand side of the jump conditions.

Inserting the above ansatz into (6) we arrive at the equations for the warp factors and the dilaton that can be found in the original work [5]. Altogether we have found two consistent solutions, labelled I and II. In solution I, the only r dependence is in $X \sim r$ and this implies divergent couplings [5]. Instead, in solution II, $\dot{X} = \alpha\dot{\chi}$, with α constant. Then, using variable separation, the bulk equations reduce to

$$\begin{aligned} A'' + D\,(A')^2 &= -D\,K^2\,, \qquad & \varphi'' + DA'\varphi' &= -DK^2/\alpha \\ C'' + DA'C' &= -D\mu K^2\,, \qquad & \ddot{\chi} + D\dot{X}\dot{\chi} &= DK^2/\alpha\,, \end{aligned} \tag{9}$$

where K and μ are constants. There are further relations

$$\begin{aligned}
&2DA'C' - (D-1)\,A'' - \frac{4}{D}\,(\varphi')^2 = -DK^2\left(\mu + 1 + \frac{D+2}{\alpha D}\right),\\
&2D\dot{X}\dot{Y} - (D-1)\,\ddot{X} - \frac{4}{D}\,(\dot{\chi})^2 = -DK^2\left(\mu - 1 + \frac{D+2}{\alpha D}\right),\\
&\alpha D^2(A' - C') + [(D+2)\,DA' + 4\varphi'] = 0\,.
\end{aligned} \tag{10}$$

The solutions to these bulk equations turn out to be

$$\begin{aligned}
A(y) &= \frac{1}{D}\log|\sin[DKy + 2\theta]|\,,\\
C_\pm(y) &= \frac{\mu}{D}\log|\sin[DKy + 2\theta]| \pm \sqrt{\frac{8}{D}}\,\frac{1}{\alpha D}\log\left|\tan\left[\frac{D}{2}Ky + \theta\right]\right|,\\
\varphi_\pm(y) &= \frac{1}{\alpha D}\log|\sin[DKy + 2\theta]| \pm \sqrt{\frac{D}{2}}\,\log\left|\tan\left[\frac{D}{2}Ky + \theta\right]\right|,\\
\chi(r) &= \frac{1}{\alpha D}\log|\cosh[DKr + \beta]| + \Phi_0\,.
\end{aligned} \tag{11}$$

Substituting into Eq. (10) determines the constants α and μ to be

$$\mu = \frac{D+1}{2} + \frac{2}{\alpha^2 D^2}\,, \qquad \alpha_\pm = \frac{(D+2) \pm \sqrt{(D+8)\,D}}{D(D-1)}\,. \tag{12}$$

An integration constant in $C_\pm$ was absorbed in a redefinition of K. The solution with α_- leads to diverging couplings so that we take α_+. To solve the matching relations we choose A of the form

$$A(y) = \begin{cases} \dfrac{1}{D}\log|\sin[DKy]|\,, & 0 \le y \le \dfrac{L}{2}\\[2ex] \dfrac{1}{D}\log|\sin[DK(y-L)]|\,, & \dfrac{L}{2} \le y \le L\,. \end{cases} \tag{13}$$

C and φ have a similar form. Remarkably, the three jump conditions turn out to be compatible and lead to

$$\cos\left[\frac{D}{2}KL\right] = \mp\sqrt{\frac{8}{D+8}}\,, \qquad e^{\frac{D+2}{D}\Phi_0} = \frac{K}{\kappa(D)\lambda}\,, \tag{14}$$

where the sign corresponds to the free sign in $C_\pm$ and $\Phi_\pm$. We choose $\frac{D}{2}KL < \pi/2$ so that the metric, as well as the dilaton, have singularities only at $y = 0$, L. Thus, we choose Φ_- and C_- in (11). The numerical coefficients $\kappa(D)$ can be found in Table 1. Consistently, the right hand side of the first equation in (14) is smaller than one.

By computing the Ricci scalar in string frame we find divergences at $y = 0, L$ implying naked singularities in the internal space. However, the dilaton also diverges at these points leading to infinite string coupling at the singularity. Thus, our next-to-leading order treatment of the string loop expansion breaks down and one might hope that higher loop or non-perturbative effects cure this singularity. After all it is not too surprising that we find

these singularities in the solution. Roughly speaking, developing these singularities is the way gravity can handle a configuration of sources (two positive tensions) that for RR-fields (two positive RR-charges) would be inconsistent. There are also singularities for $r \to \pm\infty$, where the string coupling diverges as well.

In the resulting metric with D-dimensional Poincaré invariance the coordinate r turns out to be compact. Indeed, given the warp factor $Y_s = Y + \frac{2}{D}\,\chi$ in the string metric, the effective size for this coordinate is

$$\rho = \int_{-\infty}^{\infty} dr\, e^{Y_s} = \epsilon(D)\, L\,, \tag{15}$$

where the numerical coefficients can be found in Table 1. In contrast to the tree level result, the sizes of the transverse and longitudinal non-flat directions are correlated.

4. Discussion

Even though at string tree level we compactified only the direction y, the backreaction of the dilaton tadpoles forced us to spontaneously compactify another direction r. Thus, we do not get an effective theory with $(D+1)$ dimensional Poincaré invariance. The best we can hope for is an effective theory with D dimensional Poincaré symmetry. To determine whether the solution found in the previous section lead to gravity and gauge interactions really confined to the D-dimensional space-time we compute the D-dimensional Planck mass and gauge couplings. After transforming to the Einstein frame these quantities are given by

$$\begin{aligned} M_{Pl}^{D-2} &= M_s^8 V_{8-D} \int_0^{2L} dy \int_{-\infty}^{\infty} dr\, e^{(D-2)A+B+C+(D-2)X+Y+Z}\,, \\ \frac{1}{g_D^2} &= M_s^5 V_{8-D} \int_0^{2L} dy \int_{-\infty}^{\infty} dr\, e^{\frac{(D-6)}{D}\Phi}\, e^{(D-4)A+C+(D-4)X+Z}\, \Delta(y)\,. \end{aligned} \tag{16}$$

In solution I, the integral in r diverges for both quantities. In solution II we find that M_{Pl} is finite provided we choose α_+. More concretely, by evaluating the integrals numerically we obtain

$$M_{Pl}^{D-2} = \gamma(D)\, \frac{M_s^8 V_{8-D}}{K^2}\,, \tag{17}$$

where the numerical coefficients are given in Table 1. On the other hand, for the Yang-Mills coupling we obtain

$$\frac{1}{g_D^2} = \delta(D)\, \frac{M_s^5 V_{8-D}}{K}\, e^{\frac{D-6}{D}\Phi_0}\,. \tag{18}$$

As can be seen from Table 1, the coefficient $\delta(D)$ diverges for $D = 8, 9$ and is finite only for $D \leq 6$. Thus, we only get a bona fide effective theory with at most six-dimensional Poincaré symmetry. This is in contrast to supersymmetric vacua, where the number of flat directions is a free parameter. We conclude, that in non-supersymmetric theories the number

Table 1
Numerical coefficients

D	$\kappa(D)$	$\epsilon(D)$	$\gamma(D)$	$\delta(D)$
2	1.41	19.73	0.74	0.14
3	1.04	10.98	0.72	0.33
4	0.85	8.07	0.70	0.67
5	0.73	6.60	0.68	1.44
6	0.64	5.72	0.67	5.87
7	0.58	5.12	0.66	∞
8	0.53	4.70	0.65	∞

of flat non-compact directions is not a free parameter, but can be restricted by the dynamics. This hints to an appealing dynamical mechanism to explain why we live in four dimensions.

Finally, let us see whether the solution admits to disentangle the Planck and the string scale. After a further toroidal compactification on $T^{(D-4)}$ to four flat dimensions, we obtain the following relations for the four dimensional scales

$$M_{Pl}^2 \sim M_s^8 V_{8-D} W_{D-4} L^2 , \qquad \frac{1}{g_4^2} \sim M_s^{\frac{4D+16}{D+2}} V_{8-D} W_{D-4} L^{\frac{8}{D+2}} , \tag{19}$$

where W_{D-4} is the volume of $T^{(D-4)}$. Note that these relations differ from the tree level results (1). Choosing the gauge coupling of order one implies

$$M_{Pl}^2 \sim M_s^{\frac{4D}{D+2}} L^{\frac{2D-4}{D+2}} , \tag{20}$$

showing that M_s is a free parameter as long as we choose the radius L large enough. We conclude, that large extra dimension scenarios are possible even when the next to leading order quantum corrections to the background are taken into account. Inserting numerical values into (20) and choosing $M_s = 1$ TeV gives the rough estimates of the internal dimensions shown in Table 2. Thus, in agreement with the naive tree-level result (1), in order to obtain phenomenologically acceptable sizes one has to apply more T-dualities to get D-branes with more transversal directions. However, extrapolating the results presented in [4] and in this work, it is a non-trivial question whether the critical dimension for such solutions would be larger than three.

Table 2
Large extra dimensions

D	L	$(V_{8-D} W_{D-4})^{\frac{1}{4}}$
6	10^{14} m	10^{-27} m
5	10^{19} m	10^{-30} m
4	10^{30} m	10^{-35} m

Acknowledgements

This work was done in collaboration with Ralph Blumenhagen whom I thank for sharing his ideas and insights. I am deeply grateful to the Alexander von Humboldt Foundation and to my host Dieter Lüst for providing the opportunity to carry out this research at the Humboldt Universität zu Berlin.

References

[1] S. SUGIMOTO, Prog. Theor. Phys. **102** (1999) 685 [hep-th/9905159].
[2] I. ANTONIADIS, E. DUDAS, and A. SAGNOTTI, Phys. Lett. **B464** (1999) 38 [hep-th/9908023]; C. ANGELANTONJ, I. ANTONIADIS, G. D'APPOLLONIO, E. DUDAS, and A. SAGNOTTI, Nucl. Phys. **B572** (2000) 36 [hep-th/9911081].
[3] W. FISCHLER and L. SUSSKIND, Phys. Lett. **B171** (1986) 383; Phys. Lett. **B173** (1986) 262.
[4] E. DUDAS and J. MOURAD, Phys. Lett. **B486** (2000) 172 [hep-th/0004165].
[5] R. BLUMENHAGEN and A. FONT, Nucl. Phys. **B599** (2001) 241 [hep-th/0011269].

Fortschr. Phys. **50** (2002) 5–7, 694–698

Observing Quanta on a Cosmic Scale

CRAIG J. HOGAN

Astronomy and Physics Departments, University of Washington
Seattle, WA 98195-1580, USA

Abstract

Our entire galaxy, like all others, originated as a fairly smooth patch of binding energy, which in turn originated as a single quantum perturbation of the inflaton field on a subatomic scale during inflation. The best preserved relic of these perturbations is the anisotropy of the microwave background radiation, which on the largest scales preserves a faithful image of the primordial quantum fields. It is possible that close study of these perturbations might reveal signs of discreteness caused by spacetime quantization.

1. Cosmic Background Anisotropy: Images of Inflaton Quanta on the Sky

At the beginning of the first session of the Humboldt Foundation's symposium "Werner Heisenberg: 100 years Works and Impact", an eminent particle theorist chose to begin the conference with a talk not on the grand tradition of Heisenberg quantum mechanics we had gathered to celebrate, but on the much less elegant topic of astrophysical cosmology. This choice made perfect sense: at this moment in the development of science, the most fundamental advances in understanding the innermost world of small things are coming from studies of big things: The newest physics is coming from astrophysics.

The last few years have seen spectacular discoveries of this kind. Astrophysical experiments have determined that neutrinos have mass, demonstrating the first new extension of Standard Model phenomena in many years. The mass-energy of the universe appears to be dominated by a new form of "Dark Energy" with a sufficiently large repulsive gravitational effect to accelerate the cosmic expansion of the universe, as detected in measurements of distant supernovae, and which defies all attempts at a deep theoretical understanding. Another great advance, the detailed measurement of the pattern of primordial anistropy in the cosmic microwave background, has now started to nail down basic parameters of the universe with some precision. We now know for example that the large scale geometry of the universe is nearly flat, or to put it another way, the entire cosmic hypersphere appears to be at least ten times larger in linear scale than the piece accessible to our telescopes. Better data soon to come will probably drive that lower limit up to a factor of a hundred. Real data are confirming the expectation, from inflation theory, that the universe is much larger than what we can ever see.

From the point of view of Heisenberg, perhaps most amazing newly discovered phenomena are the perturbations that create the anisotropy. Hot and cold patches in the cosmic background radiation correspond to small fractional perturbations in gravitational potential, on a vast scale, with coherent patches stretching across the entire visible universe. The perturbations are also ultimately responsible for causing all of the structure in the universe, the superclusters of galaxies, and the galaxies, stars and planets within them. They are most remarkable however because they are a quantum phenomenon: the pattern of hot and cold

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

patches scientists map on the sky today is a faithful image of the configuration of quantum field fluctuations that occurred in a tiny patch of space, far smaller than an atom, long ago during cosmic inflation. Roughly speaking, each hot and cold patch originated as a single quantum of the inflaton field, which was subsequently enormously magnified by the cosmic expansion acting as a coherent, nearly noise-free amplifier.

This audacious application of quantum mechanics is generally accepted as the correct explanation of the origin of these perturbations because a well-controlled model [1–8] makes definite predictions that agree well with the data [9–14]. The basic framework of the theory consists of free relativistic quantum fields in a classical curved spacetime: that is, Einstein's classical theory of gravity is used to compute the spacetime background on which the quantum fields propagate. The fields affect the spacetime through their energy-momentum tensor, as described by the Einstein field equations, but the spacetime itself is not quantized. The fields approximate a collection of quantized harmonic oscillators in an expanding background, with perturbations due to the zero-point field amplitudes in their ground state. Each mode expands with the background universe, and at a certain point its fluctuations are "frozen in" as its oscillation rate becomes smaller than the expansion rate, a process that can most accurately be descibed as a phase wrapping or state squeezing [15–20]. The fluctuation becomes a perturbation in the spacetime metric [2], which continues to expand by a huge factor as the universe exponentially increases in size during inflation. The final superposition of all the modes creates metric perturbations described as a continuous field with random Gaussian statistics, and a nearly (though not exactly) scale-free spectrum. These predictions are confirmed by the current data.

Thus the structure of the universe on the largest scales is directly connected to quantum processes on the smallest scales. The statistics of the fluctuations, such as the amplitude and slope of their spectrum, are determined by certain combinations of parameters of the field Lagrangian. The data on anisotropy provide by far the most precise data we have on the structure of fundamental forces on such small scales (albeit, the information is limited in scope to a few special combinations of parameters).

Indeed, the theory extrapolates the basic theoretical framework tens of orders of magnitude from any other experimental data, which leads to some healthy scepticism about whether we ought to believe it at all. In some situations (in particular, if inflation occurs close enough to the Planck scale to produce detectable and separable tensor perturbations due to graviton fluctuations, in addition to the scalar modes that lead to galaxies), we may be able to implement a more detailed test, for example, a comparison of the relative amplitude of scalar and tensor mode perturbations with the slope of the spectrum [21, 22]. However, for plenty of choices of parameters the field theory framework predicts only unobservable departures from scale-invariant, random Gaussian noise.

2. Holographic Information Bound and Quantum Discreteness of Anisotropy

It has always been acknowledged that standard inflation theory is only itself an approximation, because it does not include quantization of spacetime itself – and for good reason, since there is not yet a widely accepted theory of quantum gravity. On the other hand, there are now some definite quantitative results in quantum gravity: the dimension of the Hilbert space is known or bounded, giving bounds on the total entropy and information accessible to all configurations of all fields, including the quantum degrees of freedom of the spacetime itself. Instead of having an arbitrary zero point, entropy can now be defined in absolute terms. In other words, there are absolute limits on how many different things can happen within the confines of any given region.

This in turn imposes bounds on everything during inflation, including the behavior of free quantum fields. Their modes are not truly and fundamentally independent as assumed in the standard picture, nor do they have an infinite Hilbert space. This is a radical departure from the foundations of field theory, but it is required if we wish to include spacetime as a quantum object in the fundamental theory, in particular one that obeys unitary evolution without fundamental loss of information.

Consider for example the thought experiment of a black hole that forms, then evaporates via Hawking radiation [23–28]. If the whole process is unitary [29–32], then the states of radiated particles depend in detail on how the hole was formed. Indeed, running things backwards in time, carefully assembled[1)] incoming particles would form a small black hole that grows and then disassembles by throwing out any particular macroscopic objects – TV sets, whatever – that went into the hole. This is only possible if the spacetime metric encodes at a fundamental level the information equivalent to the radiated entropy. The entropy is known from thermodynamic arguments to be one quarter of the area of the event horizon in Planck units. Since a black hole is the highest entropy state attainable by any amount of mass/energy, one is led to the "Holographic Principle" [33–35]: for any physical system, the total entropy S within any surface is bounded by one quarter of the area A of the surface in Planck units (adopting $\hbar = c = G = 1$). This is an "absolute" entropy; the dimension N of the Hilbert space is given by e^S, and the total number of distinguishable quantum states available to the system is given by a binary number with $n = S/\ln 2 = A/4 \ln 2$ digits [36].

A holographic bound can also be derived for the universe as a whole [37–42]. Because the cosmological solutions do not have the asymptotically flat infinity of the black hole solutions for defining particle states, the bound has a somewhat different operational meaning, referring to an "observable entropy". The cosmological version [43, 44] of the holographic principle is: the observable entropy of any universe cannot exceed $S_{\max} = 3\pi/\Lambda$, where Λ is the cosmological constant in Planck units. In a de Sitter universe, as in a black hole, this corresponds to one quarter of the area of the event horizon in Planck units, but the bound is conjectured to hold for any spacetime, even Friedmann-Robertson-Walker universes with matter as well as Λ. The inflationary part of our spacetime closely resembles a piece of a de Sitter universe, so there is a bound on the observable entropy of all quantum fields during inflation: $S_{\max} = \pi/H^2$, where H is the expansion rate during inflation. That means that all the modes of all the fields have the same size Hilbert space as a system of $n = \pi/H^2 \ln 2$ binary spins.

The value of H during inflation is not known, but the field theory predicts that graviton fluctuations are produced with amplitude $\approx H$, leading to tensor perturbations of the metric and anisotropy with amplitude $\delta T/T \approx H$; since the observations with $\delta T/T \approx 10^{-5}$ are dominated by scalar perturbations, we currently have a bound around $H \leq 10^{-5}$. (This bound will improve with the advent of experiments with better sensitivity to polarization that can separate tensor and scalar components [45]). In round numbers then, the universe during inflation has a Hilbert space equivalent to at least 10^{10} spins. That is unquestionably a large number, but then again it is much smaller than the infinite Hilbert space of the field theory description. It is also much smaller than the Hilbert space of other astrophysical systems, such as stellar-mass black holes.

The most interesting question is, can we detect any observable effect of the finite Hilbert space on the fluctuations? We derive some hope from the fact the the microwave background experiments have such a high precision, and that the effects of various complicated astrophysical foregrounds have been successfully removed to reveal the simple, cleanly modeled perturbations in the cosmic last scattering surface.

[1)] To do this right with the known CP-violating, CPT-invariant fields, one would have to take the final states and parity reverse them before running them backwards in time.

To get one very rough estimate of the possible amplitude of the effect, consider the following crude model. Pay attention only to one set of modes, those five independent components that contribute to the observed quadropole moment of the anisotropy. These perturbations are created by (and are amplified images of) quantum fluctuations around the time that the current Hubble length passed through the inflationary event horizon; the amplitude of the observed temperature perturbation traces in detail a quantity which is proportional to the amplitude of the quantum field. Imagine that these modes contain all of the information allowed by the cosmological holographic bound, and further that they are literally composed of n pixels on the sky, each of which represents a binary spin. For the maximal value $H \approx 10^{-5}$ (the maximum allowed by the tensor-mode limits), there would be of order 10^{10} "black-and-white" pixels, corresponding to a pixel scale of only a few arc seconds.[2)] This level of discreteness would not be observable in practical terms, for two reasons: the last scattering surface is much thicker than this (which smears out the contribution of many pixels due to optical transfer and acoustic effects at a redshift of about 1000), and there are many other modes on a smaller scale superimposed.

On the other hand, at least one important assumption in this estimate is likely to be wrong by orders of magnitude: the modes on the horizon scale probably carry only a small fraction of the holographic information bound. At any given time, most of the information is in much smaller wavelength modes. A toy model of spacetime discreteness [46] suggests that the Hilbert space of the horizon-scale quantum perturbations is equivalent to at most only about 10^5 binary spins. (This number is determined not the tensor-mode limit on H, but by the inverse of the observed scalar perturbation amplitude). This in principle may produce observable discreteness, since the all-sky cosmic background anisotropy includes about 10^4 independent samplings of inflationary fields (determined by the angular size of the horizon at last scattering).

Now the idea that there are literally binary pixels is also silly, and was used here only for illustrative convenience. In fact we have no clear idea of the actual character of the holographic eigenstates projected onto the inflaton perturbations; we have only the counting argument to guide us. In [46], a toy model assumed that H and δT come in discrete levels, but this again was only for calculational convenience. In the absence of a more concrete theory of the nature of the holographic modes, it makes sense to consider a variety of tests on the data to seek departures from the field-theory prediction of a continuous, random Gaussian field. The most straightforward check will be to test whether the amplitudes of the harmonics $A_{\ell m}$ come in discrete values rather than being selected from a random continuous Gaussian distribution. Although this test could already be performed fairly reliably using the all-sky COBE satellite data [47, 48], it will become much more powerful using the all-sky data from the MAP (and later PLANCK) satellites, which will have much finer angular resolution capable of cleanly resolving modes far below the horizon scale at recombination (a level already reached in ground- and balloon-based experiments with limited sky coverage). The high quality of this data will let us search for true quantum-gravity effects.

Acknowledgments

I am grateful for useful conversations with J. Bardeen, R. Bousso, and R. Brandenberger. This work was supported by NSF grant AST-0098557 at the University of Washington.

[2)] This level of graininess in the everyday world would be just below the level of detectability with the naked eye.

References

[1] A. A. Starobinsky, JETP Lett. **30** (1979) 682 [Pisma Zh. Eksp. Teor. Fiz. **30** (1979) 719].
[2] J. M. Bardeen, Phys Rev D **22** (1980) 1882.
[3] S. W. Hawking, Phys. Lett. B **115** (1982) 295.
[4] A. H. Guth and S. Y. Pi, Phys. Rev. Lett. **49** (1982) 1110.
[5] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys. Rev. D **28** (1983) 679.
[6] A. A. Starobinsky, Phys. Lett. B **117** (1982) 175.
[7] J. J. Halliwell and S. W. Hawking, Phys. Rev. D **31** (1985) 1777.
[8] L. P. Grishchuk, Class. Quant. Grav. **10** (1993) 2449.
[9] G. F. Smoot et al., Astrophys. J. **396** (1992) L1.
[10] P. de Bernardis et al., Nature **404** (2000) 955–959.
[11] C. Pryke et al., Ap. J. in press, astro-ph/0104490 (2001).
[12] N. W. Halverson et al., Angular astro-ph/0104489 (2001).
[13] S. Hanany et al., Astrophys. J. **545** (2000) L5.
[14] J. R. Bond and R. G. Crittenden, "CMB Analysis", in Proc. NATO ASI "Structure Formation in the Universe", eds. R. G. Crittenden and N. G. Turok, astro-ph/0108204, (Kluwer, 2001).
[15] L. P. Grishchuk and Y. V. Sidorov, Class. Quant. Grav. **6** (1989) L161.
[16] L. P. Grishchuk and Y. V. Sidorov, Phys. Rev. D **42** (1990) 3413.
[17] A. Albrecht, P. Ferreira, M. Joyce, and T. Prokopec, Phys. Rev. D **50** (1994) 4807.
[18] J. Lesgourgues, D. Polarski, and A. A. Starobinsky, Nucl. Phys. B **497** (1997) 479.
[19] D. Polarski and A. A. Starobinsky, Class. Quant. Grav. **13** (1996) 377.
[20] C. Kiefer, D. Polarski, and A. A. Starobinsky, Int. J. Mod. Phys. D **7** (1998) 455–462.
[21] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rept. **215** (1992) 203.
[22] D. H. Lyth and A. Riotto, Phys. Rept. **314** (1999) 1–146.
[23] J. D. Bekenstein, Lett. Nuovo Cim. **4** (1972) 737.
[24] J. D. Bekenstein, Phys. Rev. D **7** (1973) 2333.
[25] J. D. Bekenstein, Phys. Rev. D **9** (1974) 3292.
[26] S. W. Hawking, Commun. Math. Phys. **43** (1975) 199.
[27] S. W. Hawking, Phys. Rev. D **14** (1976) 2460.
[28] J. D. Bekenstein, Phys. Rev. D **23** (1981) 287.
[29] T. Banks, L. Susskind, and M. E. Peskin, Nucl. Phys. B **244** (1984) 125.
[30] G. 't Hooft, Nucl. Phys. B **256** (1985) 727.
[31] L. Susskind, L. Thorlacius, and J. Uglum, Phys. Rev. D **48** (1993) 3743.
[32] C. R. Stephens, G. 't Hooft, and B. F. Whiting, Class. Quant. Grav. **11** (1994) 621–648.
[33] G. 't Hooft, *Dimensional Reduction in Quantum Gravity*, essay in honor of Abdus Salam, gr-qc/9310026 (1993).
[34] L. Susskind, J. Math. Phys. **36** (1995) 6377.
[35] D. Bigatti and L. Susskind, *TASI lectures on the holographic principle*, arXiv:hep-th/0002044 (1999).
[36] J. D. Bekenstein, *Quantum Information and Quantum Black Holes*, Lectures delivered at NATO Advanced Study Institute: International School of Cosmology and Gravitation: 17th Course of the International School of Cosmology and Gravitation, Erice, May 1–11, 2001, to appear in Advances in the Interplay between Quantum and Gravity Physics, ed. V. de Sabbata (Kluwer, Dordrecht 2002).
[37] L. Susskind and E. Witten, *The Holographic Bound in Anti-de Sitter Space*, hep-th/9805114 (1998).
[38] W. Fischler and L. Susskind, *Holography and Cosmology*, hep-th/9806039 (1998).
[39] R. Bousso, JHEP 9907 (1999) 004.
[40] R. Bousso, JHEP 9906 (1999) 028.
[41] R. Bousso, Class. Quant. Grav. **17** (2000) 997.
[42] R. Bousso, Phys. Rev. D **60** (1999) 063503.
[43] R. Bousso, *Positive Vacuum Energy and the N-bound*, JHEP in press, hep-th/0010252 (2001).
[44] E. E. Flanagan, D. Marolf, and R. M. Wald, Phys. Rev. D **62** (2000) 084035.
[45] M. Kamionkowski and A. Kosowski, Ann. Rev. Nucl. Part. Sci. **49** (1999) 77.
[46] C. Hogan, *Holographic Discreteness of Inflationary Perturbations*, Phys. Rev. D. (submitted).
[47] C. L. Bennett et al., Astrophys. J. **464** (1996) L1.
[48] K. M. Gorski et al., Astrophys. J. **464** (1996) L11.

Fortschr. Phys. **50** (2002) 5–7, 699–704

Supersymmetric Configurations in Gauged Supergravity

SWAPNA MAHAPATRA[1)] and HARVENDRA SINGH[2)]

[1)] Physics Department, Utkal University, Bhubaneswar-751004, India
swapna@iopb.res.in
[2)] Fachgruppe Theoretische Physik, Martin-Luther-Universität Halle-Wittenberg
D-06120 Halle, Germany
singh@physik.uni-halle.de

Abstract

We discuss supersymmetric configurations in gauged supergravity theories with particular emphasis on Euclidean $D = 4$, $N = 4$ gauged $SU(2) \times SU(1,1)$ supergravity.

1. Introduction

Supergravity backgrounds play an important role in the analysis of the string theory. Apart from fully supersymmetric string vacua, solutions with partial supersymmetry have also played an important role in verifying various duality relations. The conjectured duality between supersymmetry on anti de Sitter (AdS) background and superconformal field theory (CFT) on the AdS boundary [1] as well as the domain-wall/QFT correspondence [2] has led to renewed interest in the study of gauged supergravity theories. These theories can be derived through a compactification of ten or eleven dimensional supergravity theories on space-times involving AdS subspaces, namely, $AdS_{p+2} \times S^{D-2-p}$, which are near horizon geometries of extremal solutions in supergravity theories. Gauged supergravity theories can be constructed when a subgroup of the R-symmetry group or the automorphism group of the symmetry algebra is gauged by the vector fields in the graviton supermultiplet. Other methods include gauging the isometries of the vector as well as hypermultiplet moduli spaces. The procedure of gauging does not change the particle spectrum or number of supersymmetries in the theory. However, gauging introduces new terms proportional to the square of the gauge coupling constant in the action. For $N \leq 3$, there are no scalars and gauging induces a cosmological constant. In the presence of scalar fields (for $N \geq 4$), gauging induces a scalar potential $V(\Phi)$, thereby changing the properties of the ground state. The induced potential $V(\Phi)$ is unbounded from below and it may or may not have critical points. So it is important to understand the nature of the ground state. The idea is to look for maximally supersymmetric solutions or solutions preserving some amount of supersymmetry which corresponds to stable configurations [3]. The supersymmetric ground states are often the AdS or domain wall type solutions. For obtaining supersymmetric configurations which are invariant under all or some of the supersymmetry transformations, one needs to check if the background possesses killing spinors. So one has to check for the existence of nontrivial spinor parameters for which the fermion supersymmetry variations vanish. Such nontrivial spinor parameters are called supersymmetric killing spinors. In this talk, we shall mainly focus on the Euclidean version of $N = 4$, $D = 4$ gauged supergravity theories and shall obtain various supersymmetric configurations.

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

2. Euclidean $N = 4$, $D = 4$ Supergravity Theories

In four dimensions, there are two versions of $N = 4$ supergravity theories, one with a global $SO(4)$ symmetry [4] and the other one with a global $SU(4)$ symmetry [5]. The equations of motion of the two formulations are equivalent by using field redefinition and duality transformations. However, the gauged models corresponding to the respective local internal symmetries are inequivalent. The $N = 4$, $SO(4)$ gauged supergravity [6] has one coupling constant. The theory has four-dimensional AdS space as a stable ground state. On the otherhand, $N = 4$ gauged $SU(2) \times SU(2)$ (subgroup of $SU(4)$) supergravity has two gauge coupling constants g_1 and g_2 and the scalar potential does not have any critical point. However, the scalars may be stabilized by turning on background gauge fields. The familiar examples are the Freedman- Schwarz (FS) electro-vac solution [7] with abelian electric fields where the space-time geometry is $AdS_2 \times R^2$ and the axio-vac solution with the space-time geometry being $AdS_3 \times R^1$. The above backgrounds are supersymmetric as they admit killing spinors. Other supersymmetric configurations in the FS model include domain-wall solutions preserving one-half of the original supersymmetry [8], nonabelian magnetic solitons [9] and BPS black holes [10]. The maximal $N = 8$ gauged supergravity in four dimensions has also been constructed [11].

Now we go over to the discussion of Euclidean Freedman-Schwarz (EFS) model which has been obtained by Volkov [12]. In the EFS model, the four-dimensional gravity multiplet contains the graviton E^m_μ, 4 majorana spin-$\frac{3}{2}$ gravitino Ψ^I_μ $(I = 1, \ldots, 4)$, gauge fields A^a_μ $(a = 1, 2, 3)$ belonging to $SU(2)$ with gauge coupling constant g_1, three pseudovector gauge fields $\dot{A}^a_\mu$ $(a = 1, 2, 3)$ belonging to $SU(1, 1)$ with gauge coupling constant g_2, four majorana spin-$\frac{1}{2}$ fields χ^I $(I = 1, \ldots, 4)$, the axion $\boldsymbol{a}$ and the dilaton Φ. Here greek indices $\mu, \nu, \ldots$ refer to base space indices and latin indices $m, n, \ldots$ refer to tangent space indices. Let us note that the EFS and FS models are not related by analytic continuations. Whereas, FS model can be embedded into $N = 1$ supergravity in $D = 10$ as an $S^3 \times S^3$ compactification with the group manifold being $SU(2) \times SU(2)$ [13], EFS model can be embedded into $N = 1$ supergravity in $D = 10$ as an $S^3 \times AdS_3$ compactification with the group manifold being $SU(2) \times SU(1, 1)$. In the later case, the internal space has signature $(+, +, +, +, +, -)$ and the corresponding four-dimensional theory becomes Euclidean. In both the theories, the solutions are not asymptotically flat due to the presence of the dilaton potential $V(\Phi)$, which is given by,

$$FS\ model\colon V(\Phi) = -\frac{1}{8}\,(g_1^2 + g_2^2)\,e^{-2\Phi}\,, \tag{1}$$

$$EFS\ model\colon V(\Phi) = -\frac{1}{8}\,(g_1^2 - g_2^2)\,e^{-2\Phi}\,. \tag{2}$$

As the scalar curvature of S^3 is positive and that of AdS_3 is negative, the dilaton potential in the corresponding four-dimensional theory becomes proportional to $g_1^2 - g_2^2$, where g_1 and g_2 are the gauge coupling constants corresponding to $SU(2)$ and $SU(1, 1)$ respectively. Since the potential is proportional to the square of the difference of the gauge couplings, one can consider a variety of cases, where the potential can be positive, negative or zero. The bosonic part of the ten dimensional action corresponding to $N = 1$ supergravity is given by,

$$S_{10} = \int \sqrt{-\hat{g}}\, d^{10}\hat{x}\, \left(\frac{1}{4}\,\hat{\boldsymbol{R}} - \frac{1}{2}\,\partial_M\hat{\boldsymbol{\Phi}}\,\partial^M\hat{\boldsymbol{\Phi}} - \frac{1}{12}\,e^{-2\hat{\boldsymbol{\Phi}}}\,\hat{\boldsymbol{H}}_{MNP}\hat{\boldsymbol{H}}^{MNP}\right), \tag{3}$$

where $\hat{\boldsymbol{R}}$ is the curvature scalar in $D = 10$ with $\hat{g}_{MN}(M, N, \ldots = 1, \ldots, 10)$ being the metric, $\hat{\boldsymbol{H}}_{MNP}$ is the three form anti-symmetric field strength and $\hat{\boldsymbol{\Phi}}$ is the dilaton. The idea is

to find a parametrization of $\hat{g}_{MN}$, $\hat{H}_{MNP}$, $\hat{\Phi}$ in terms of the four-dimensional variables which reduces the action and equations of motion to a consistent theory in four-dimensions.

After dimensional reduction, the four-dimensional Euclidean FS action involving the metric, dilaton, axion and gauge fields is given by,

$$S_4 = \int \sqrt{g}\, d^4x \left[\frac{R}{4} - \frac{1}{2}\, \partial_\mu \Phi\, \partial^\mu \Phi + \frac{1}{2}\, e^{-4\Phi}\, \partial_\mu \boldsymbol{a}\, \partial^\mu \boldsymbol{a} - \frac{1}{4}\, e^{2\Phi}\, (\eta^{(1)}_{ab} F^a_{\mu\nu} F^{b\mu\nu} + \eta^{(2)}_{ab} \dot{F}^a_{\mu\nu} \dot{F}^{b\mu\nu}) \right.$$
$$\left. - \frac{1}{2}\, \boldsymbol{a}\, (\eta^{(1)}_{ab} \star F^a_{\mu\nu} F^{b\mu\nu} + \eta^{(2)}_{ab} \star \dot{F}^a_{\mu\nu} \dot{F}^{b\mu\nu}) + \frac{1}{8}\, (g_1^2 - g_2^2)\, e^{-2\Phi} \right], \tag{4}$$

where, $\mu, \nu = 1, \ldots, 4$, $\eta^{(1)}_{ab} = \mathrm{diag}\,(1,1,1)$ and $\eta^{(2)}_{ab} = \mathrm{diag}\,(1,1,-1)$ are the cartan metrics corresponding to $SU(2)$ and $SU(1,1)$ respectively. The dual field strengths in the four dimensional Euclidean theory are defined as,

$$\star F^a_{\mu\nu} = \frac{1}{2}\, \sqrt{g}\, \epsilon_{\mu\nu\lambda\rho} F^{a\lambda\rho}\,, \qquad \star \dot{F}^a_{\mu\nu} = \frac{1}{2}\, \sqrt{g}\, \epsilon_{\mu\nu\lambda\rho} \dot{F}^{a\lambda\rho}\,. \tag{5}$$

The corresponding four dimensional equations of motion for metric, dilaton, axion and gauge fields are respectively given by,

$$R_{\mu\nu} - 2\, \partial_\mu \Phi\, \partial_\nu \Phi + 2\, e^{-4\Phi}\, \partial_\mu \boldsymbol{a}\, \partial_\nu \boldsymbol{a}$$
$$- 2\, e^{2\Phi} \left[\eta^{(1)}_{ab} \left(F^a_{\mu\lambda} F^{b\lambda}_{\nu} - \frac{1}{4}\, g_{\mu\nu} F^a_{\lambda\rho} F^{b\lambda\rho} \right) + \eta^{(2)}_{ab} \left(\dot{F}^a_{\mu\lambda} \dot{F}^{b\lambda}_{\nu} - \frac{1}{4}\, g_{\mu\nu} \dot{F}^a_{\lambda\rho} \dot{F}^{b\lambda\rho} \right) \right]$$
$$- 2 g_{\mu\nu} U(\Phi) = 0\,, \tag{6}$$

$$\nabla_\mu \nabla^\mu \Phi - 2\, e^{-4\Phi}\, \partial_\mu \boldsymbol{a}\, \partial^\mu \boldsymbol{a} - \frac{1}{2}\, e^{2\Phi}\, [\eta^{(1)}_{ab} F^a_{\mu\nu} F^{b\mu\nu} + \eta^{(2)}_{ab} \dot{F}^a_{\mu\nu} \dot{F}^{b\mu\nu}] + 2U(\Phi) = 0\,, \tag{7}$$

$$\nabla_\mu (e^{-4\Phi} \nabla^\mu \boldsymbol{a}) + \frac{1}{2}\, [\eta^{(1)}_{ab} \star F^a_{\mu\nu} F^{b\mu\nu} + \eta^{(2)}_{ab} \star \dot{F}^a_{\mu\nu} \dot{F}^{b\mu\nu}] = 0\,, \tag{8}$$

$$\nabla_\rho (e^{2\Phi} F^{a\rho\mu}) + g_1\, e^{2\Phi} f^a_{bc} A^b_\rho F^{c\rho\mu} - 2 \star F^{a\mu\rho}\, \partial_\rho \boldsymbol{a} = 0\,, \tag{9}$$

$$\nabla_\rho (e^{2\Phi} \dot{F}^{a\rho\mu}) + g_2\, e^{2\Phi} \dot{f}^a_{bc} \dot{A}^b_\rho \dot{F}^{c\rho\mu} - 2 \star \dot{F}^{a\mu\rho} \partial_\rho \boldsymbol{a} = 0\,, \tag{10}$$

where the structure constants are,

$$f^c_{ab} = \eta^{(1)cd} \epsilon_{dab}; \qquad \dot{f}^c_{ab} = \eta^{(2)cd} \epsilon_{dab}\,. \tag{11}$$

In the fermionic sector, after consistently reducing the ten dimensional spinors, the 4-dimensional supersymmetry variations are obtained as,

$$\delta\chi = \left(\frac{1}{\sqrt{2}}\, \gamma^\mu\, \partial_\mu \Phi - \frac{1}{\sqrt{2}}\, e^{-2\Phi}\, \gamma_5 \gamma^\mu\, \partial_\mu \boldsymbol{a} \right) \epsilon$$
$$+ \frac{1}{2}\, e^{\Phi} \left(\frac{1}{2}\, \eta^{(1)}_{ab} \gamma^\alpha \gamma^\beta F^a_{\alpha\beta} \alpha^b - \frac{1}{2}\, \gamma_5 \eta^{(2)}_{ab} \gamma^\alpha \gamma^\beta \dot{F}^a_{\alpha\beta} \dot{\alpha}^b \right) \epsilon + \frac{1}{4}\, e^{-\Phi}\, (g_1 - g_2 \gamma_5)\, \epsilon\,,$$

$$\delta\Psi_\mu = \left(\partial_\mu + \frac{1}{4}\, \omega^{\alpha\beta}_\mu \gamma_\alpha \gamma_\beta - \frac{g_1}{2}\, \eta^{(1)}_{ab} \alpha^a A^b_\mu + \frac{g_2}{2}\, \eta^{(2)}_{ab} \dot{\alpha}^a \dot{A}^b_\mu + \frac{1}{2}\, e^{-2\Phi}\, \gamma_5\, \partial_\mu \boldsymbol{a} \right) \epsilon$$
$$+ \frac{1}{2\sqrt{2}}\, e^{\Phi} (\eta^{(1)}_{ab} F^a_{\lambda\nu} \alpha^b + \gamma_5 \eta^{(2)}_{ab} \dot{F}^a_{\lambda\nu} \dot{\alpha}^b)\, \gamma^\lambda \gamma^\nu \gamma_\mu \epsilon + \frac{1}{4\sqrt{2}}\, e^{-\Phi}\, (g_1 + g_2 \gamma_5)\, \gamma_\mu \epsilon\,, \tag{12}$$

where ϵ is the supersymmetry transformation parameter, γ^α are the four dimensional tangent space gamma matrices, $\omega_\mu^{\alpha\beta}$ are the spin connections and α^a and $\dot{\alpha}^a$ are the 4×4 matrices which generate the Lie algebra of the group $SU(2)$ and $SU(1,1)$ respectively.

One can construct many interesting stable new vacua in the EFS model which are supersymmetric and are consistent with the four-dimensional background equations of motion.

3. Supersymmetric Configurations

Here we consider some examples where the nonzero gauge fields are abelian.

3.1. Euclidean domain walls

We consider the four dimensional Euclidean domain wall obtained by analytically continuing the Lorenztian domain walls [8] with the field configurations,

$$\begin{aligned} &ds^2 = U(y)\,(dt^2 + dx_1^2 + dx_2^2) + U^{-1}(y)\,dy^2\,, \\ &\Phi = \frac{1}{2}\,\ln U(y)\,, \qquad U(y) = m\,|y - y_0|\,, \\ &A_\mu^a = 0\,, \qquad \dot{A}_\mu^a = 0\,, \qquad \boldsymbol{a} = 0\,. \end{aligned} \tag{13}$$

This background is singular at $y = y_0$. For $g_1 \neq 0$ (here g_1 and m must be related) and $g_2 = 0$, one finds that the fermionic variations vanish provided the supersymmetry parameters satisfy,

$$\epsilon = -\gamma_3\epsilon\,, \qquad \epsilon = U(y)^{\frac{1}{4}}\,\epsilon_0\,, \tag{14}$$

where ϵ_0 is a constant spinor. These conditions break half of the supersymmetry thereby implying existence of nontrivial killing spinors preserving $N = 2$ supersymmetry for pure dilatonic Euclidean domain wall background.

3.2. $E^2 \times AdS_2$

We consider constant dilaton and nonzero abelian gauge fields. The field configurations are given by,

$$\begin{aligned} &ds^2 = d\psi^2 + d\chi^2 + \frac{1}{B}\left(r^2\,dt^2 + \frac{dr^2}{r^2}\right), \\ &\dot{F}^a = \delta^{\hat{a}3} Q\,dt \wedge dr\,, \qquad \Phi = \Phi_0 = \text{constant}\,, \\ &F = 0\,, \qquad \boldsymbol{a} = \text{constant}\,, \qquad g_2 = 0\,. \end{aligned} \tag{15}$$

The geometry corresponds to $E^2 \times AdS_2$, where $\frac{1}{\sqrt{B}}$ corresponds to the radius of the AdS space. The above background fields are consistent with the equations of motion (for $Qg_1 = 1$). One gets the following equations for the ψ, χ, t and r components for the killing spinor:

$$\partial_\psi\epsilon = 0\,, \qquad \partial_\chi\epsilon = 0\,, \tag{16}$$

$$\partial_t\epsilon + \frac{r}{2}\,\gamma_2\gamma_3\epsilon + \frac{1}{2}\,r\gamma_2\epsilon = 0\,, \qquad \partial_r\epsilon + \frac{1}{2r}\,\gamma_3\epsilon = 0\,. \tag{17}$$

The solution to these equations is

$$\epsilon = \frac{1}{2}\, r^{\frac{1}{2}}\big(\gamma_5\gamma_2\gamma_3\dot{\alpha}^3 + 1\big)\, \epsilon_- + \frac{1}{2}\, [r^{-\frac{1}{2}} - r^{\frac{1}{2}}\gamma_2 t]\, (\gamma_5\gamma_2\gamma_3\dot{\alpha}^3 + 1)\, \epsilon_+ \,, \tag{18}$$

where $\epsilon_\mp$ are constant spinors such that $\gamma_3\epsilon_\mp = \mp\epsilon_\mp$. Note that $\big[\gamma_5\gamma_2\gamma_3\dot{\alpha}^3 + 1\big]$ acts like a projector and $\dot{\alpha}^3$ is along the noncompact direction in $SU(1,1)$. This solution is the analog of the electro-vac solution in FS model.

We have also constructed the analog of pure axionic gravity solutions where the geometry corresponds to $E^1 \times S^3$. The field configuration is given by,

$$\begin{aligned} &ds^2 = d\psi^2 + \frac{1}{Q}\,(d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta\, d\phi^2))\,, \\ &\boldsymbol{a} = \pm\sqrt{Q}\,\psi; \qquad \Phi = 0\,, \\ &A = 0; \qquad \dot{A} = 0; \qquad g_1 = 0\,. \end{aligned} \tag{19}$$

The complete solution of the killing spinor equation is given by,

$$\epsilon = e^{-\frac{1}{2}\chi\gamma_5\gamma_1}\, e^{-\frac{1}{2}\theta\gamma_2\gamma_1}\, e^{-\frac{1}{2}\phi\gamma_3\gamma_2}\left[\frac{\gamma_0 + 1}{2}\right]\epsilon_0\,. \tag{20}$$

3.3. Gravitational instantons

It has been noted in [12] that with vanishing dilaton, axion, gauge fields and for $g_1 = g_2$, the flat gravitational instantons are vacua of EFS model. Here we show that even in the presence of (anti)self-dual gauge fields, the Eguchi-Hanson instanton [14] satisfying the flat space Einstein equations is a consistent background of the EFS model preserving certain fraction of the supersymmetry. The field configuration is given by,

$$\begin{aligned} &ds^2 = \frac{dr^2}{1 - \dfrac{a^4}{r^4}} + \frac{r^2}{4}\,(d\theta^2 + \sin^2\theta\, d\phi^2) + \frac{r^2}{4}\left(1 - \frac{a^4}{r^4}\right)(d\psi + \cos\theta\, d\phi)^2\,, \\ &F = \frac{2}{r^4}\,(e^3 \wedge e^0 + e^1 \wedge e^2) = \dot{\boldsymbol{F}}\,, \\ &\Phi = 0; \qquad \boldsymbol{a} = 0; \qquad g_1 = g_2\,. \end{aligned} \tag{21}$$

With the above choice of background fields, the supersymmetry variations give the projector conditions,

$$(1 - \gamma_5)\,\epsilon = 0\,, \qquad (\alpha^3 + \dot{\alpha}^3)\,\epsilon = 0\,. \tag{22}$$

With these projectors, the killing spinor equations look really simple, namely,

$$\partial_\mu\epsilon = 0\,. \tag{23}$$

So the killing spinors are independent of r, θ, ϕ, ψ. Because of the twin supersymmetric conditions, the solution preserves $\frac{1}{4}$ of the supersymmetry. However, once the gauge field backgrounds are switched off the second condition in Eq. (22) drops out and the pure gravitational instanton background becomes half supersymmetric.

4. Summary

To summarise, scalar potentials of gauged supergravity theories provide an interesting testing ground for studying supersymmetric configurations. In particular, the scalar potential in $D = 4,\ N = 4,\ SU(2) \times SU(1,1)$ gauged supergravity theory can be positive, negative or zero depending on the value of the gauge coupling constants. We have shown the existence of stable supersymmetric configurations like domain wall, electro-vac, axio-vac type solutions and flat gravitational instanton solutions in the above Euclidean Freedman-Schwarz model. Our findings of a large class of stable vacua for the EFS model makes the theory more interesting and worth exploring further.

Acknowledgements

SM is grateful to the Humboldt Foundation for invitation to participate in the Heisenberg Symposium held in an excellent atmosphere in Bamberg, Germany. Work of HS is also supported by the Humboldt Foundation. SM would like to thank Institute of Physics, Bhubaneswar for extending the computer facilities.

References

[1] J. Maldacena, Adv. Theor. Math. Phys. **2** (1998) 231 [hep-th/9711200].
[2] H. J. Boonstra, K. Skenderis, and P. K. Townsend, JHEP **9901** (1999) 003 [hep-th/9807137].
[3] E. Witten, Comm. Math. Phys. **80** (1981) 381;
G. W. Gibbons, C. M. Hull, and N. P. Warner, Nucl. Phys. **B218** (1983) 173.
[4] A. Das, Phys. Rev. **D15** (1977), 2805;
E. Cremmer and J. Scherk, Nucl. Phys. **B127** (1977) 259.
[5] E. Cremmer, J. Scherk, and S. Ferrara, Phys. Lett. **B74** (1978) 61.
[6] A. Das, M. Fischler, and M. Rocek, Phys. Rev. **D16** (1977) 3427.
[7] D. Freedman and J. H. Schwarz, Nucl. Phys. **B137** (1978) 333;
D. Freedman and G. W. Gibbons, Nucl. Phys. **B233** (1984) 24.
[8] H. Singh, Phys. Lett. **B429** (1998) 304 [hep-th/9801038], Phys. Lett. **B444** (1998) 327 [hep-th/9808181];
P. M. Cowdall, Class. Quant. Grav. **15** (1998) 2937 [hep-th/9710214];
P. Cowdall and P. K. Townsend, Phys. Lett. **B429** (1998) 281, erratum-ibid. **B434** (1998) 458 [hep-th/9801165].
[9] A. H. Chamseddine and M. S. Volkov, Phys. Rev. Lett. **79** (1997) 3343 [hep-th/9707176].
[10] D. Klemm, Nucl. Phys. **B545** (1999) 545 [hep-th/9810090].
[11] B. de Wit and H. Nicolai, Nucl. Phys. **B208** (1982) 323.
[12] M. S. Volkov, Nucl. Phys. **B566** (2000) 173 [hep-th/9910116].
[13] A. H. Chamseddine and M. S. Volkov, Phys. Rev. **D57** (1998) 6242 [hep-th/9711181].
[14] T. Eguchi and A. J. Hanson, Ann. Phys. **120** (1979) 82.

Fortschr. Phys. **50** (2002) 5–7, 705–710

Anomalies and Schwinger Terms in NCG Field Theory Models

JOUKO A. MICKELSSON

Mathematical Physics, Royal Institute of Technology, SCFAB
SE-10691, Stockholm, Sweden

Abstract

This talk is based on the paper [1], which contains references to the original literature. We study the quantization of chiral fermions coupled to generalized Dirac operators arising in NCG Yang-Mills theory. The cocycles describing chiral symmetry breaking are calculated. In particular, we introduce a generalized locality principle for the cocycles. Local cocycles are by definition expressions which can be written as generalized traces of operator commutators. In the case of pseudodifferential operators, these traces lead in fact to integrals of ordinary local de Rham forms. As an application of the general ideas we take the case of noncommutative tori.

1. Introduction

Anomalies in quantum field theory arise because certain symmetries (gauge and coordinate transformations) of a classical system cannot be preserved in quantization.

The right mathematical arena for the study of quantum field theory anomalies is the study of families of Dirac operators, parametrized by external classical fields, and their properties in relations to symmetry transformations. It is often more practical (and normally also sufficient) to study the properties of the QFT observables with respect to infinitesimal gauge and coordinate transformations. This leads in a natural way to a computation of cocycles for the Lie algebra of infinitesimal gauge (coordinate) transformations, with coefficients in a module consisting of functions of the external fields (vector potentials and metrics).

Here we want to discuss the Lie algebra cohomology of 'gauge' transformations in noncommutative geometry models.

Instead of a standard Dirac operator one considers self-adjoint unbounded Fredholm operators D with the property that $1/|D|^p$ is 'almost' trace-class, meaning that after removing a potential logarithmic divergence the trace is converging. After fixing such an operator we study generalized Dirac operators $D_A = D + A$ where A is any hermitean bounded perturbation. We stress that our considerations are very general: The family of Dirac operators could for example arise from a coupling of vector potentials through a star product (generalized Moyal brackets). All what is needed is that the L_p estimate mentioned above is valid. This is known to hold in a class of star product quantizations defined by a constant antisymmetric θ matrix; we shall comment an example of this in the end of the talk.

Our starting point for a construction of a NCG field theory model is the Connes' spectral triple $(D, \int, \mathcal{B})$, where D is the Dirac operator acting in a Hilbert space H, $\mathcal{B}$ is an associative algebra of operators in H such that $[D, b]$ is bounded for all $b \in \mathcal{B}$. In addition, there is an 'integration', or a generalized trace map, from operators ('p-forms') of the type $\omega = b_0[\epsilon, b_1]\,[\epsilon, b_2]\ldots[\epsilon, b_p]$, or alternatively $|D|^{-p}\, b_0[D, b_1]\ldots[D, b_p]$, to complex numbers. Here $\epsilon = D/|D|$ is the sign of the Dirac operator. In some cases (but not always) one can prove an equality between the integrals of the two alternative expressions. The general-

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

ized vector potentials are then linear combinations of 1-forms $A = b_0[D, b_1]$, or sometimes $a = b_0[\epsilon, b_1]$.

We study the BRST double complex based on de Rham forms on the space of vector potentials A with values in the space of the above operator p-forms ω, with $p = 0, 1, 2, \ldots$

One of the central results concerns the question of locality of the BRST cocycles. In abstract NCG models it is not a priori clear what this could mean because in the operator approach there is in general no space-time manifold to define local fields. However, in the ordinary space-time geometric setup many of the basic quantities can be written as 'traces' of commutators of pseudodifferential operators. For example, this holds for nonabelian Schwinger terms and the gauge anomaly in the path integral formulation. It turns out that these traces are in fact integrals of local differential forms. We adopt this as our starting point: We prove that in the NCG field theory models there is a 'local' anomaly formula expressing the BRST cocycles as traces of commutators of nontrace-class operators.

2. The General Setup for NCG Descent Equations and Anomalies

Let D_0 be an unbounded selfadjoint operator in a complex Hilbert space H such that $|D_0|^{-1} \in L_{p+}$, that is, $|D_0|^{-p} \in L_{1+}$ for some $p \geq 1$. Here L_{1+} is the Dixmier ideal in the algebra of bounded operators in H. A positive operator T is in L_{1+} if it is compact and

$$\frac{1}{\log N} \sum_{k=1}^{N} \lambda_k$$

has a finite limit, where $\lambda_1 \geq \lambda_2 \geq \ldots$ are the eigenvalues of T. We also assume for simplicity that D_0 is invertible. However, the following discussion can be easily generalized to the case when D_0 is a finite rank perturbation of an invertible operator.

We shall work with bounded perturbations of the 'free Dirac operator' D_0, denote $D_A = D_0 + A$ where A is a bounded selfadjoint operator in H such that $[|D_0|, A]$ is bounded. We shall denote F_A a smoothed sign operator associated to D_A. The technical complication is that the map $A \mapsto D_A/|D_A|$ is not continuous when D_A has zero modes. Instead, we can take a smooth function $f : \mathsf{R} \to \mathsf{R}_+$ such that $f(x) - |x|$ approaches zero faster than any power of x as $|x| \mapsto \infty$ and $f(x) \geq m$ for some positive constant, and we define $F_A = D_A/f(D_A)$. For example, take $f(x) = +\sqrt{x^2 + e^{-x^2}}$. Then $A \mapsto F_A$ is norm continuous. If D_A is the classical Dirac operator associated to a vector potential A on a compact manifold then the difference $D_A/|D_A| - F_A$ is an infinitely smoothing pseudodifferential operator and in particular a trace class operator.

The reason why F_A is important in QFT is that the sign operator $D_A/|D_A|$ defines the vacuum in the Fock space representation of the canonical anticommutation relations algebra CAR. It is also known that two CAR representations are equivalent if and only if the difference of the sign operators is Hilbert-Schmidt. For this reason F_A is a good parameter for the different Fock spaces.

As an operator function $D_A/f(D_A)$ of D_A, the operator F_A satisfies $g^{-1}F_A g = F_{A^g}$, with $A^g = g^{-1}Ag + g^{-1}[D_0, g]$ for a unitary transformation g such that $[D_0, g]$ is bounded. Denote by $\mathcal{B}$ the algebra of bounded operators b in H such that $[D_0, b]$ and $[|D_0|, b]$ are bounded. Then all the operators $A = b_0[D_0, b_1]$, for $b_i \in \mathcal{B}$, satisfy the condition $[|D_0|, A]$ is bounded. We denote by U_{p+} the group of unitary elements in $\mathcal{B}$. Any element $g \in U_{p+}$ satisfies $[\epsilon, g] \in L_{p+}$ where $\epsilon = D_0/|D_0|$.

The 'infinitesimal version' of the gauge transformation $A \mapsto A^g$ in terms of the parameter a is

$$\delta_X a = [a, X] + [\epsilon, X] \quad \text{for} \quad X \in \boldsymbol{u}_{p+} = \mathrm{Lie}\,(U_{p+})\,. \tag{2.1}$$

Let us recall the basic definitions in NCG differential calculus for Fredholm modules. The differentials of order n are linear combinations of operators of the type $b_0[\epsilon, b_1]\dots[\epsilon, b_n]$ where $b_i \in \mathcal{B}$. One denotes $db = [\epsilon, b]$ for $b \in \mathcal{B}$. If $\phi \in \Omega^n$ is a differential of order n then $d\phi = \epsilon\phi + (-1)^{n+1}\,\phi\epsilon$. This gives a map $d : \Omega^n \to \Omega^{n+1}$ with $d^2 = 0$. The cohomology of this complex is trivial.

The coboundary operator associated to infinitesimal gauge transformations is denoted by δ. We work with cochains of order k, $\tau \in \Omega_k$, consisting of functions $\tau(a; X_1, \dots, X_k)$ of $a \in \Omega^1$ and of Lie algebra elements $X_i \in \boldsymbol{u}_{p+}$, linear in each X_i and totally antisymmetric in the arguments X_i. The standard Lie algebra coboundary operator is defined by

$$(\delta\tau_n)\,(a; X_1, \dots, X_{n+1}) = \sum_i\,(-1)^{i-1}\,\delta_{X_i}\tau(a;\, X_1, \dots, \hat{X}_i, \dots, X_{n+1}) + \sum_{i<j}\,(-1)^{i+j}\,\tau(a;\, [X_i, X_j], \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_{n+1})\,, \tag{2.2}$$

where the hat means that the corresponding argument is deleted and δ_{X_i} is the Lie derivative acting on functions of a, the action on the argument being given by (2.1). We remind that the multilinear forms τ on a Lie algebra can be interpreted as left invariant differential forms on the corresponding Lie group (and vice versa) through the standard identification of a Lie algebra as left invariant vector fields.

There is a (d, δ) double complex consisting of δ forms taking values in the d complex Ω^*. In order to guarantee that $d\delta + \delta d = 0$ in addition to $d^2 = 0 = \delta^2$ we have to modify the sign conventions in the definition of δ in the standard way (a relative minus sign on odd order d forms as compared to even order forms).

In the standard discussion of anomalies in quantum field theory one constructs cocycles $c_{n,k}$ in the (Ω_*, δ) complex by integrating de Rham forms $\omega_{n,k}$ in Ω^n_* over a compact manifold of dimension n. In the NCG setting integration of forms is replaced by applying an appropriate trace functional to the operator valued forms.

The translation from the classical to NCG setting is straight-forward. The de Rham exterior derivation is replaced by the operation d described above (to the forms after the symbol '$\int$'). All the formal manipulations are done exactly in the same way as in the classical BRST complex. Especially, one can define NCG Chern forms F^n with $F = da + a^2$, and by descent equations one arrives at the forms $\omega_{n,k}$ in the (d, δ) double complex,

$$\delta(\omega_{2n-1-k,k}) + d(\omega_{2n-1,k+1}) = (\dots)\,, \qquad k = 0, \dots, 2n-1\,,$$

where the dots denote sums of commutators.

Of course, the nontriviality of the cohomology classes depends on what is meant by the trace (and the definition of the trace is intertwined by the choice of $(H, D_0, \mathcal{B})$).

3. 'Local' NCG Anomalies and Schwinger Terms

In the case of the classical BRST complex all the cocycles $c_{[i,j]}$ are given in terms of differential forms which are differential polynomials in variables a, v. The Fredholm module cocycles involve terms like $[\epsilon, v]$, $\epsilon a + a\epsilon$, and therefore are nonlocal in nature; when evaluated using the symbol calculus of pseudodifferential operators they contain terms of arbitrary high order in the partial derivatives.

However, even in the case of the Fredholm module cocycles (for classical vector potentials and gauge transformations) the locality is preserved in a certain sense. Namely, it turns out that the cocycles $c_{[i,j]}$ are equivalent (in the BRST cohomology) to cocycles $c'_{[i,j]}$ which can be written as renormalized traces of commutators of PSDO's.

The trace of a commutator depends only on the term in the asymptotic expansion of a PSDO which has order equal to $-\dim M$, and for this reason one needs to take into account only a finite number of derivatives of the symbols (since each differentiation in a homogeneous term decreases the order by one). In this sense the trace of a commutator is a local expression. In a more general setup, beyond the PSDO calculus, we take this as a **definition of locality:** cocycles which are traces of commutators in the algebra are called local.

We set up the following assumptions. There is a complex linear functional TR on the algebra generated by D_0, $|D_0|$ and $\mathcal{B}$ such that 1) it is equal to the ordinary trace for trace class operators, 2) it has the property that $\mathrm{TR}[A, B] = 0$ when $AB, BA \in L_{1+}$, and 3) $\mathrm{TR}\,[\epsilon, W] = 0$ (odd case), $\mathrm{TR}\,[\Gamma\epsilon, W] = 0$ (even case), for bounded operators W, with $\epsilon = D_0/|D_0|$. In the even case there is an operator Γ anticommuting with D_0, commuting with $\mathcal{B}$, and $\Gamma^2 = 1$.

Example 1: Set $p = 1$ and consider the cocycle $c(X, Y) = \mathrm{tr}_C\, X[\epsilon, Y]$, for $X, Y \in \mathcal{B}$. Here $\mathrm{tr}_C\, X \equiv \frac{1}{2}\mathrm{tr}\,(X + \epsilon X \epsilon)$. We can write

$$\begin{aligned} c_{1,2}(X, Y) &= \frac{1}{4}\ \mathrm{tr}\,\epsilon[\epsilon, X]\ [\epsilon, Y] = \mathrm{tr}_C\, X[\epsilon, Y] \\ &= \frac{1}{2}\ \mathrm{TR}\, X[\epsilon, Y] = \frac{1}{2}\ \mathrm{TR}\,[X\epsilon, Y] - \frac{1}{2}\ \mathrm{TR}\,\epsilon[X, Y]\,. \end{aligned}$$

The last term is the coboundary of the cochain $\theta(X) = \frac{1}{2}\mathrm{TR}\,\epsilon X$ and therefore the class of $c_{1,2}$ is given by

$$c_{\mathrm{loc}}(X, Y) = \frac{1}{2}\ \mathrm{TR}\,[X\epsilon, Y]\,.$$

One can also check by a direct computation that c_{loc} is a cocycle.

$$\begin{aligned} 2(\delta c_{\mathrm{loc}})(X, Y, Z) &= \mathrm{TR}\,\{[[X, Y]\ \epsilon, Z] + \mathrm{cycl.}\} \\ &= \mathrm{TR}\,\{[[X, Y],\ [\epsilon, Z]] + \mathrm{cycl.}\} \\ &= \mathrm{TR}\,[\epsilon,\ [[X, Y],\ Z]] + \mathrm{TR}\,[Y,\ [[\epsilon, X], Z]] - \mathrm{TR}\,[X,\ [[\epsilon, Y], Z]] + \mathrm{cycl.} \end{aligned}$$

By 3) the first term on the right vanishes and by 2) the second and third term vanish. In the case when X, Y are multiplication operators, by smooth functions, on the circle S^1 the local cocycle becomes the central term in an affine Lie algebra,

$$c_{\mathrm{loc}}(X, Y) = \frac{1}{2\pi i} \int_{S^1} \mathrm{tr}\, X\, dY\,,$$

where the trace under the integral sign is a finite-dimensional matrix trace for the matrix valued functions X, Y.

The cocycle $c_{1,2}$ (or c_{loc}) arises in canonical quantization of fermions. A '1-particle operator' X is mapped to an operator $\hat{X}$ in the Fock space; the commutation relations are modified by the cocycle c,

$$[\hat{X}, \hat{Y}] = \widehat{[X, Y]} + c(X, Y)\,.$$

Example 2: Here we consider the problem arising from quantization of gauge currents in three space dimensions. Typically, $[\epsilon, X]$ is not Hilbert-Schmidt but it belongs to the ideal

$L_{3+} \subset L_4$. For this reason the expression for the 2-cocycle in the previous example does not converge for 3-dimensional gauge currents. Instead, one has to introduce a renormalization of the 2-cocycle,

$$c_{3,2}(a;\, X,\, Y) = \frac{1}{8}\, \mathrm{tr}_C\, a[[\epsilon,\, X],\, [\epsilon,\, Y]]\,,$$

with $a = F_A - \epsilon$. One can check by a direct calculation that this is a cocycle in the sense that

$$c_{3,2}(a;\, [X,\, Y],\, Z) + \delta_X c_{3,2}(a;\, Y,\, Z) + \text{ cyclic perm. of } X,\, Y,\, Z = 0\,.$$

Let next $\eta(a;\, X) = \frac{1}{8}\mathrm{TR}\, \epsilon a[\epsilon,\, X]$. By a direct calculation one can check that $c_{3,2} = \delta\eta + c_{\mathrm{loc}}$ where now

$$8c_{\mathrm{loc}}(a;\, X,\, Y) = \mathrm{TR}\,[Y,\, \epsilon a[\epsilon,\, X]] - \mathrm{TR}\,[X,\, \epsilon a[\epsilon,\, Y]] + 2\,\mathrm{TR}\,[X\epsilon,\, Y] - 2\,\mathrm{TR}\,[Y\epsilon,\, X]$$

is explicitly a generalized trace of a sum of commutators.

In this example the canonical quantization of gauge currents is ill-defined even after normal ordering, precisely because $[\epsilon,\, X]$ is not Hilbert-Schmidt. However, there is an operator theoretic interpretation for second quantized $\hat{X}$, $\hat{Y}$. These are now generators for unitary transformations between Fock spaces carrying inequivalent representations of the CAR algebra. Geometrically, there is a bundle of Fock spaces parametrized by the external field a and the gauge transformations act as unitary maps between the fibers.

Example 3: The gauge anomaly in two space-time dimensions. Here we are in the even Fredholm module case. Then $c_{2,1}(a;\, X) = \mathrm{tr}_C\, \Gamma a[\epsilon,\, X]$ is a cocycle,

$$\delta_X c_{2,1}(a;\, Y) - \delta_Y c_{2,1}(a;\, X) - c_{2,1}(a;\, [X,\, Y]) = 0\,,$$

using the fact that $\mathrm{tr}_C[\epsilon,\, \cdot] = 0$. In this case

$$c_{2,1}(a;\, X) = \mathrm{TR}\,[\Gamma a\epsilon,\, X] + (\delta\eta)\,(a)\,,$$

where $\eta(a) = \mathrm{TR}\,\Gamma a\epsilon$. In BRST notation, for an even module, $c_{\mathrm{loc}} = \mathrm{TR}\,[a\epsilon,\, s(v)]$. Inserting $\epsilon = D_0/|D_0|$, $D_0 = -i\sum_{k=1}^{2} \gamma^k \frac{\partial}{\partial x_k}$, $D_A = D_0 + \sum \gamma^k A_k(x)$, and $a = D_A/|D_A| - \epsilon$ one obtains the standard formula for the nonabelian gauge anomaly in two space-time dimensions,

$$c_{\mathrm{loc}} = \frac{1}{4\pi}\int \mathrm{tr}\,(A_1\, \partial_2 X - A_2\, \partial_1 X)\,,$$

for a smooth infinitesimal gauge transformation X of compact support.

In the general case we have the following result:

Theorem: *Let $n = 1, 2, \ldots$ and $k = 0, 1, 2, \ldots$ and let the cocycle $c_{2n-1-k,k}$ be computed from the descent equations (starting from Chern classes). Then for even k the cohomology class $[c_{2n-1-k,k}]$ is represented by a cocycle $c'_{2n-1-k,k}$ which is a generalized trace of a sum of commutators; each commutator is a polynomial in the variables a, da, X, $dX \ldots$ In the case of odd k one has to add a term proportional to a generalized trace of $\omega_{0,k}$.*

The general ideas above can be applied to the case of a noncommutative torus. For example, in three dimensions the Schwinger term for the gauge algebra (example 2 above)

gives a simple formula

$$\omega_{3,2} = \frac{i\pi}{6}\ \tau(A[[D_0,\ X],\ [D_0,\ Y]]) = \frac{i\pi}{3}\ \epsilon^{ijk}\tau(A_i(\delta_j(X)\ \delta_k(Y) - \delta_j(Y)\ \delta_k(X)))\,.$$

Here δ_i (with $i = 1, 2, 3$) are the standard derivations on a noncommutative torus, τ is a trace functional, D_0 is the free Dirac operator, and A_i are the components of the vector potential.

References

[1] E. Langmann, J. Mickelsson, and S. Rydh, Anomalies and Schwinger terms in NCG field theory models, J. Math. Phys. **42** (2001), 10, 4779–4801. hep-th/0103006.

Fortschr. Phys. **50** (2002) 5–7, 711–716

On Gravitational Interaction of Fermions

YURI N. OBUKHOV

Department of Theoretical Physics, Moscow State University, 117234 Moscow, Russia

Abstract

We discuss some aspects of the gravitational interaction of the relativistic quantum particles with spin 1/2. The exact Foldy-Wouthuysen transformation is constructed for the Dirac particle coupled to the static spacetime metric. The quasi-relativistic limit of the theory is then analyzed. Using the analogous method, we obtain the exact Cini-Touschek transformation and discuss the ultra-relativistic limit of the fermion theory. We show that the Foldy-Wouthuysen transformation is not uniquely defined, and the corresponding ambiguity is deeply rooted in the relativistic quantum theory.

1. Introduction

The recent development of experimental technique, in particular of the neutron interferometric methods [1], has provided the first direct tests of the interaction of the quantum spinless particles with the classical gravitational field. There is a little doubt that a further technological progress (using the polarized neutrons, atomic interferometers, etc) will soon make it possible to measure the higher order gravitational and inertial effects of the quantum particle *with* spin. Theoretical studies of the relativistic quantum theory in a curved spacetime have predicted a number of interesting manifestations of the spin-gravity coupling for the Dirac fermion, see [2, 9, 12, 13], e.g. In most cases, the various approximate schemes were used for the case of the weak gravitational field. Here the exact results for an arbitrary static spacetime geometry are reported.

A massive quantum particle with spin $1/2$ is described by the relativistic Dirac theory. In the curved spacetime, the fermion wave function – 4-spinor field ψ – satisfies the covariant Dirac equation

$$(i\hbar\gamma^{\alpha}D_{\alpha} - mc)\ \psi = 0\,. \tag{1}$$

The spinor covariant derivative is defined by

$$D_{\alpha} = h^{i}_{\alpha}D_{i}\,, \qquad D_{i} := \partial_{i} + \frac{i}{4}\ \hat{\sigma}_{\alpha\beta}\Gamma_{i}{}^{\alpha\beta}\,, \tag{2}$$

which shows that the gravitational and inertial effects are encoded in the coframe (vierbein) and the Lorentz connection coefficients $h_{i}{}^{\alpha}$, $\Gamma_{i}{}^{\alpha\beta} = -\Gamma_{i}{}^{\beta\alpha}$. We use the Greek alphabet for the indices which label the components with respect to a local Lorentz frame $e_{\alpha} = h^{i}_{\alpha}\partial_{i}$, whereas the Latin indices refer to the local spacetime coordinates x^{i}. For the Dirac matrices, the conventions of [3] are used. In particular, we have $\beta = \gamma^{0}, \vec{\alpha} = \beta\vec{\gamma}$, $\gamma_{5} = -i\gamma^{\hat{0}}\gamma^{\hat{1}}\gamma^{\hat{2}}\gamma^{\hat{3}}$, $\hat{\sigma}^{\alpha\beta} = i\gamma^{[\alpha}\gamma^{\beta]}$. The spin matrix is defined by $\vec{\Sigma} = i\vec{\gamma}\times\vec{\gamma}/2 = -\gamma_{5}\vec{\alpha}$.

Let us consider the metric of the static spacetime

$$ds^{2} = V^{2}(dx^{0})^{2} - W^{2}\,(d\vec{x}\cdot d\vec{x})\,. \tag{3}$$

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

Here $x^0 = ct$, and $V = V(\vec{x})$, $W = W(\vec{x})$ are arbitrary functions of $\vec{x}$. Many important particular cases belong to this family: (i) the flat Minkowski spacetime in *accelerated frame* corresponds to the choice $V = 1 + (\vec{a} \cdot \vec{x})/c^2$, $W = 1$, (ii) *Schwarzschild* spacetime in the isotropic coordinates with $r := \sqrt{\vec{x} \cdot \vec{x}}$ is obtained for $V = \left(1 - \frac{GM}{2c^2 r}\right)\left(1 + \frac{GM}{2c^2 r}\right)^{-1}$, $W = \left(1 + \frac{GM}{2c^2 r}\right)^2$, (iii) *de Sitter* spacetime in static frame is recovered for $V = \frac{1+r^2/\ell^2}{1-r^2/\ell^2}$, $W = 2/(1 - r^2/\ell^2)$, where ℓ is the constant curvature radius (with the curvature 2-form $R^{\alpha\beta} = 1/\ell^2 \vartheta^\alpha \wedge \vartheta^\beta$), (iv) the *product* spacetime $R \times S^3$ (fermion on a sphere) arises when $V = 1$, $W = (1 + r^2/(4L^2))^{-1}$ with L radius of the sphere S^3.

Using (3), one can bring the Dirac Eq. (1) to the Schrödinger form

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{\mathcal{H}} \psi \tag{4}$$

with the Hamilton operator

$$\hat{\mathcal{H}} = \beta m c^2 V + \frac{c}{2} \left[(\vec{\alpha} \cdot \vec{p}) \, \mathcal{F} + \mathcal{F} (\vec{\alpha} \cdot \vec{p}) \right]. \tag{5}$$

Here we introduced $\mathcal{F} = V/W$.

2. Foldy-Wouthuysen Transformation

In order to reveal the true physical content of the theory and to find its correct interpretation, one needs to perform the Foldy-Wouthuysen (FW) transformation [4]. Technically, this yields the representation in which the quantum states with positive and negative energy become uncoupled.

We use here the approach of Eriksen [5] to construct the exact Foldy-Wouthuysen transformation. The energy sign operator (Pauli) is defined by $\hat{\Lambda} = \hat{\mathcal{H}}/\sqrt{\hat{\mathcal{H}}^2}$. It is Hermitian, unitary, and idempotent: $\hat{\Lambda}^2 = \hat{\Lambda}^\dagger \hat{\Lambda} = 1$. The unitary operator U which maps the Dirac representation to the FW-representation

$$\psi \to \psi^F = U \psi \,, \tag{6}$$

should satisfy the condition

$$U \hat{\Lambda} U^\dagger = \beta \,. \tag{7}$$

Remarkably, for our case, the *exact* FW-transformation exists. Consider the operator

$$J := i \gamma_5 \beta = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \tag{8}$$

It is Hermitian, $J^\dagger = J$, unitary, and idempotent: $JJ^\dagger = J^2 = 1$, and it anticommutes both with the Hamiltonian, and with β:

$$J \hat{\mathcal{H}} + \hat{\mathcal{H}} J = 0 \,, \qquad J\beta + \beta J = 0 \,. \tag{9}$$

Then the FW transformation (7) is realized by

$$U = \frac{1}{2} \, (1 + \beta J) \, (1 + J \hat{\Lambda}) \tag{10}$$

and the corresponding FW Hamiltonian reads

$$\hat{\mathcal{H}}^F = U\hat{\mathcal{H}}U^\dagger = \left[\sqrt{\hat{\mathcal{H}}^2}\right]\beta + \left\{\sqrt{\hat{\mathcal{H}}^2}\right\}J\,. \tag{11}$$

As usual, even and odd parts of an operator Q are defined by $[Q] := \frac{1}{2}(Q + \beta Q\beta)$ and $\{Q\} := \frac{1}{2}(Q - \beta Q\beta)$. Explicitly, we have for the square of (5)

$$\hat{\mathcal{H}}^2 = m^2c^4V^2 + \mathcal{F}c^2p^2\mathcal{F} + \frac{\hbar^2c^2}{2}\left(\mathcal{F}\vec{\nabla}\vec{f} - \vec{f}^{\,2}/2\right) + \hbar c^2\,\mathcal{F}\vec{\Sigma}\cdot(\vec{f}\times\vec{p} + Jmc\,\vec{\phi}) \tag{12}$$

which contains the *odd* piece (the last term). Here: $\vec{\phi} := \vec{\nabla}V$, $\vec{f} := \vec{\nabla}\mathcal{F}$.

The obtained FW Hamiltonian is *exact*. However, for most practical purposes, it is sufficient to consider the non-relativistic limit, and use the quasi-relativistic wave functions, treating all the interaction terms as perturbations. The quasi-relativistic approximation is straightforwardly obtained by assuming that mc^2 term is dominating, and expanding $\sqrt{\hat{\mathcal{H}}^2}$ in powers of $1/(mc^2)$. However, the massless case is covered only by the exact result.

Expanding $\sqrt{\hat{\mathcal{H}}^2}$ in powers of $1/(mc^2)$, we find finally the quasi-relativistic Hamiltonian:

$$\begin{aligned}\hat{\mathcal{H}}^F \approx\ & \underbrace{\beta mc^2V}_{\text{COW}} + \underbrace{\frac{1}{4m}\beta\left(\frac{1}{W}p^2\mathcal{F} + \mathcal{F}p^2\frac{1}{W}\right)}_{\text{kinetic rel. red shift}} + \underbrace{\frac{\hbar^2}{4mW}\beta\,(\triangle\mathcal{F})}_{\text{grav. Darwin term}} - \frac{\hbar^2}{8mV}\beta\vec{f}^{\,2} \\ & + \underbrace{\frac{\hbar}{4m}\beta\vec{\Sigma}\cdot\left(\frac{1}{W}\vec{f}\times\vec{p} + \vec{f}\times\vec{p}\,\frac{1}{W}\right)}_{\text{grav/inert. spin-orbit. term}} + \frac{\hbar c}{2W}\,(\vec{\Sigma}\cdot\vec{\phi})\,.\end{aligned} \tag{13}$$

The first two terms describe the familiar effects already measured experimentally for spinless particles (Colella-Overhauser-Werner and Bonse-Wroblewski, [1]). The first term in the second line represents the new inertial/gravitational spin-orbital momentum effects, cf. [12, 13]. The "gravitational Darwin" term admits a physical interpretation similar to usual electromagnetic Darwin term, reflecting the zitterbewegung fluctuation of the fermion's position with the mean square $\langle(\delta r)^2\rangle \sim \hbar^2/(mc)^2$.

It is interesting to observe the emergence of the spin-gravitational moment coupling which is described by the last term in (13). Such interaction was predicted, in a phenomenological approach, by Kobzarev and Okun [6] and was discussed by Peres [7], see also the recent reviews [8]. The presence of this term demonstrates the validity of the equivalence principle for the Dirac fermions [9].

3. Cini-Touschek Transformation

In the study of the high-energy neutrino effects in the gravitational field of a massive compact object (see [10], e.g.), it is convenient to use a different representation which is directly related to the ultra-relativistic rather than to the quasi-relativistic limit. The new representation is in this sense complementary to the FW picture.

The corresponding limit (when $mc^2 \ll c\,|p|$) is achieved with the help of the Cini-Touschek (CT) transformation [11]. We can construct the exact CT-transformation for a

fermion in the static metric (3) using the scheme similar to the above FW case. To begin with, we observe that the operator

$$\hat{\mathcal{P}} = \frac{\vec{\alpha} \cdot \vec{p}}{|p|} \tag{14}$$

is Hermitian, unitary, and idempotent, $\hat{\mathcal{P}}^2 = 1$. It is proportional to the chirality operator $\hat{\chi} = \vec{\Sigma} \cdot \vec{p}/|p| = -\gamma_5 \hat{\mathcal{P}}$. Evidently, we have

$$\hat{\mathcal{P}} J + J \hat{\mathcal{P}} = 0\,. \tag{15}$$

In complete analogy to (7), the CT-transformation is determined by the unitary operator U which satisfies

$$U \hat{\Lambda} U^{\dagger} = \hat{\mathcal{P}}\,. \tag{16}$$

Thus, technically, we need to replace $\beta \leftrightarrow \hat{\mathcal{P}}$ everywhere in the above derivations. The explicit CT-operator then reads

$$U = \frac{1}{2}\left(1 + \hat{\mathcal{P}} J\right)\left(1 + J \hat{\Lambda}\right) \tag{17}$$

and the Cini-Touschek Hamiltonian is

$$\hat{\mathcal{H}}^{CT} = \left[\sqrt{\hat{\mathcal{H}}^2}\,\right]^{\mathcal{P}} \hat{\mathcal{P}} + \left\{\sqrt{\hat{\mathcal{H}}^2}\right\}^{\mathcal{P}} J\,. \tag{18}$$

Here, the "$\hat{\mathcal{P}}$-odd/even" parts of an operator are defined by the same token as the usual "β-odd/even" parts. As a simple application, we check that for the free particle (18) yields $\hat{\mathcal{H}}^{CT} = E\hat{\mathcal{P}} \approx c\vec{\alpha} \cdot \vec{p}$ which is the correct ultra-relativistic Hamiltonian.

4. Ambiguities

The presence of the spin-gravitational moment in the quasi-relativistic FW Hamiltonian requires some comments.

FW transformation is defined with a certain ambiguity. Let us consider the unitary transformation $U' = e^{iS}$ with

$$S = \frac{\beta}{mc}\left\{b(x)\,(\vec{\Sigma}\vec{p}\,) + (\vec{\Sigma}\vec{p}\,)\;b(x)\right\}, \tag{19}$$

where $b(x)$ is an arbitrary function of the spatial coordinates. The spaces of quantum states with positive and negative energies are invariant under the action of this operator. For the Hamiltonian of the unitary equivalent representation, we find, in a perturbative manner:

$$\begin{aligned}\hat{\mathcal{H}}' = U'\hat{\mathcal{H}}^F U'^{\dagger} = \hat{\mathcal{H}}^F + 2\hbar c b(\vec{\Sigma}\vec{\phi}) + \frac{i\hbar}{m}\beta[b\,(\vec{\Sigma}\vec{p}\,),\,(1/W + 2b)\,(\vec{\Sigma}\vec{\phi})] \\ + \frac{\hbar^2}{2m}\,(1/W + 2b)\,\beta[(\vec{\Sigma}\vec{\nabla}b),\,(\vec{\Sigma}\vec{\phi})] + \mathcal{O}(1/m^2)\,.\end{aligned} \tag{20}$$

Using (13), we find that the choice $2b = -\frac{1}{2W}$ yields the approximate form of the FW Hamiltonian reported by Fischbach et al. [12] and by Hehl and Ni [13].

The mentioned ambiguity is deeply rooted in the relativistic quantum theory. The FW representation is often treated merely as a rigorous method to derive the quasi-relativistic limit of the Dirac equation (see [14], e.g.), refining the "non-rigorous" derivation based on the direct elimination of the so-called small components of the 4-spinor wave function. And indeed, one can straightforwardly verify that the direct derivation of the Pauli equation suffers from the same ambiguity: Recall that after eliminating the small components, the remaining 2-spinor wave function should be properly normalized [15]. The corresponding normalization operator is not uniquely defined and this yields the transformation of the type (20) of the quasi-relativistic Hamiltonian.

Furthermore, one can easily find the relevant ambiguities of that kind in the full (relativistic) Dirac theory. For example, the Hamiltonian of the free particle $\hat{\mathcal{H}} = c(\vec{\alpha}\vec{p}) + \beta mc^2$ is invariant under the unitary transformation of the wave function described by

$$U = \sqrt{\frac{E + mc^2}{2E}} \left(1 + \frac{ic}{E + mc^2} (\vec{\Sigma}\vec{p})\, \hat{\Lambda}\right). \tag{21}$$

Operators of position, spin and energy (Hamiltonian) can have different form in the unitary equivalent representations, and one should properly determine them in order to analyze the physical effects. Certainly, the observable quantities measured in experiment do not depend on the choice of representation.

5. Discussion and Conclusion

Approximate scheme (see Bjorken-Drell [3], e.g.) was developed for the case of electromagnetic coupling. As it is well known, the idea is to remove, order by order in $1/m$, odd terms from the Hamiltonian $\hat{\mathcal{H}} = \hat{\mathcal{H}}_1 = \beta mc^2 + \mathcal{E} + \mathcal{O}$. A unitary transformation $\psi_2 = U_{21}\psi_1$, with $U_{21} = e^{iS}$, yields (in the time-independent case) the perturbative construction of the new Hamiltonian

$$\hat{\mathcal{H}}_2 = U_{21}\hat{\mathcal{H}}_1 U_{21}^{\dagger} = \hat{\mathcal{H}}_1 + [iS, \hat{\mathcal{H}}_1] + \frac{1}{2}\,[iS, [iS, \hat{\mathcal{H}}_1]] + \frac{1}{3!}\,[iS, [iS, [iS, \hat{\mathcal{H}}_1]]] + \ldots \tag{22}$$

In electrodynamics, the odd $\mathcal{O}$ and even $\mathcal{E}$ parts do not depend on the mass m. Instead, they are proportional to *electromagnetic* charge e, and that fact makes the approximate scheme working. For example, choosing at the first step $S = -i\beta\mathcal{O}/2m$, we remove the original odd part and find $\hat{\mathcal{H}}_2 = \beta mc^2 + \mathcal{E}' + \mathcal{O}'$ where the new odd part

$$\mathcal{O}' = \frac{\beta}{2m}\,[\mathcal{O}, \mathcal{E}] - \frac{\mathcal{O}^3}{3m^2} \tag{23}$$

has a higher order in $1/m$ than the new even part. However, for the gravitational/inertial case, $\mathcal{E}$ is proportional to the *gravitational/inertial* charge m. As a result, the new odd term in $\hat{\mathcal{H}}_2$ is of order m^0. The same happens at every step of the approximate scheme: the "remaining" even terms have the same order in $1/m$ as the "removed" odd terms. This makes the issue of the convergence of the approximate scheme problematic. In our approach, this deficiency is avoided by using the *exact* FW transformation.

Here we have demonstrated how to obtain the exact FW and CT transformations in the covariant Dirac theory. The detailed discussion of the corresponding applications to the specific quasi-relativistic and ultra-relativistic physical problems will be presented elsewhere.

Acknowledgment

I would like to thank the Alexander von Humboldt Foundation for the invitation and support.

References

[1] R. Colella, A. W. Overhauser, and S. A. Werner, Phys. Rev. Lett. **34** (1975) 1472;
U. Bonse and T. Wroblewski, Phys. Rev. Lett. **51** (1983) 1401.
[2] C. G. de Oliveira and J. Tiomno, Nuovo Cim. **24** (1962) 672;
J. Audretsch and G. Schäfer, Gen. Rel. Grav. **9** (1978) 243;
Y. Q. Cai and G. Papini, Phys. Rev. Lett. **66** (1991) 1259; Phys. Rev. Lett. **68** (1992) 3811;
D. Singh and G. Papini, Nuovo Cim. **B115** (2000) 223;
J. C. Huang, Ann. d. Phys. **3** (1994) 53;
K. Varjú and L. H. Ryder, Phys. Lett. **A250** (1998) 263;
L. Ryder, J. Phys. A: Math. Gen. **A31** (1998) 2465;
K. Varjú and L. H. Ryder, Phys. Rev. **D62** (2000) 024016.
[3] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill: San Francisco, 1964).
[4] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78** (1950) 29.
[5] E. Eriksen and M. Kolsrud, Nuovo Cim. **18** (1960) 1;
A. G. Nikitin, J. Phys. A: Math. Gen. **A31** (1998) 3297.
[6] I. Yu. Kobzarev and L. B. Okun, Sov. Phys. JETP **16** (1963) 1343 [ZhETF **43** (1962) 1904 (in Russian)]; see also J. Leitner and S. Okubo, Phys. Rev. **136** (1964) 1542.
[7] A. Peres, Phys. Rev. **D18** (1978) 2739.
[8] B. Mashhoon, Class. Quantum Grav. **17** (2000) 2399. L. H. Ryder and B. Mashhoon, to be published in the Proceedings of the Ninth Marcel Grossmann Meeting (Rome, 2000); e-Print Archive: gr-qc/0102101.
[9] Yu. N. Obukhov, Phys. Rev. Lett. **86** (2001) 192.
[10] K. Konno and M. Kasai, Prog. Theor. Phys. **100** (1998) 1145.
[11] M. Cini and B. Touschek, Nuovo Cim. **7** (1958) 422;
S. K. Bose, A. Gamba, and E. C. G. Sudarshan, Phys. Rev. **113** (1959) 1661.
[12] E. Fischbach, B. S. Freeman, and W.-K. Cheng, Phys. Rev. **D23** (1981) 2157.
[13] F. W. Hehl and W.-T. Ni, Phys. Rev. **D42** (1990) 2045.
[14] M. E. Rose, *Relativistic electron theory* (John Wiley & Sons: New York, 1961).
[15] A. I. Akhiezer and V. B. Berestetsky, *Quantum electrodynamics* (Consultants Bureau: Oak Ridge) [Translation form the Russian edition: *Kvantovaya elektrodinamika* (GITTL: Moscow, 1953)], see Sec. 14.

Fortschr. Phys. **50** (2002) 5–7, 717–723

New Einstein-Hilbert Type Action for Unity of Nature

KAZUNARI SHIMA

Laboratory of Physics, Saitama Institute of Technology, Okabe-machi
Saitama 369-0293, Japan
shima@sit.ac.jp

Abstract

A new Einstein-Hilbert type (SGM) action is obtained by performing the Einstein gravity analogue geometrical arguments in high symmetric (SGM) spacetime. All elementary particles except graviton are regarded as the eigenstates of $SO(10)$ super-Poincaré algebra (SPA) and composed of the fundamental fermion "superons" of nonlinear supersymmetry (NL SUSY). Some phenomenological implications and the linearlization of the action are discussed briefly.

1. Introduction

The standard model (SM) is established as a unified model for the electroewak interaction. Nevertheless, it is very unsatisfactory in many aspects, e.g. it can not explain the particle quantum numbers (Q_e, I, Y, *color*, i.e. $1 \times 2 \times 3$ *gauge structure*), the three-generations structure and contains more than 28 arbitrary parameters (in the case of neutrino oscillations) even disregarding the mass generation mechanism for neutrino. The simple and beautiful extension to $SU(5)$ GUT has serious difficulties, e.g. the life time of proton, etc and is excluded so far. The SM and GUT equiptted naively with supersymmetry (SUSY) have improved the situations, e.g. the unification of the gauge couplings at about 10^{17}, relatively stable proton (now threatened by experiments), etc., but they posess more than 100 arbitrary parameters and less predictive powers. However SUSY [1] is an essential notion to unify various topological and non-topological charges and gives a natural framework to unify spacetime and matter leading to the birth of supergravity(SUGRA). Unfortunately the maximally extended $SO(8)$ SUGRA is too small to accommoate all observed particles as elementary fields. The straightforward extension to SO(N) SUGRA with $N > 9$ has a difficulty due to so called the no-go theorem on the massless elementary high spin (>2) (gauge) field. The massive high-spin is another.

Furthermore, we think that from the viewpoint of simplicity and beauty of nature it is interesting to attempt the accommodation of all observed particles in a single irreducible representation of a certain algebra (group) especially in the case of high symmetric spacetime having a certain boundary (i.e. a boundary condition) and the dynamics are described by the spontaneous breakdown of the high symmetry of spacetime by itself, which is encoded in the nonliner realization of the geometrical arguments of spacetime. Also the no-go theorem does not exclude the possibility that the fundamental action, if it exists, posesses the high-spin degrees of freedom not as the elementary fields but as some *composite* eigenstates of a certain symmetry (algebra) of the fundamental action. In this talk we would like to present a model along this scenario.

2. Superon-Graviton Model (SGM) – Phenomenology

Among all single irreducible representations of all SO(N) extended super-Poincaré(SP) symmetries, the massless irreducible representations of SO(10) SP algebra(SPA) is the only one

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

that accommodates minimally all observed particles including the graviton [2]. 10 generators $Q^N(N = 1, 2, \ldots, 10)$ of $SO(10)$ SPA are the fundamental represemtations of $SO(10)$ internal symmetry and decomposed $\underline{10} = \underline{5} + \underline{5}^*$ with respect to $SU(5)$ following $SO(10) \supset SU(5)$. For the massless case the little algebra of $SO(10)$ SPA for the supercharges in the light-cone frame $P_\mu = \epsilon(1, 0, 0, 1)$ becomes after a suitable rescaling

$$\{Q^M_\alpha, Q^N_\beta\} = \{\bar{Q}^M_{\dot\alpha}, \bar{Q}^N_{\dot\beta}\} = 0, \quad \{Q^M_\alpha, \bar{Q}^N_{\dot\beta}\} = \delta_{\alpha 1}\delta_{\dot\beta \dot 1}\delta^{MN}, \tag{1}$$

where $\alpha, \beta = 1, 2$ and $M, N = 1, 2, \ldots 5$. By identifying the graviton with the Clifford vacuum $|\Omega\rangle$ ($SO(10)$ singlet) satisfying $Q^M_\alpha |\Omega\rangle = 0$. and performing the ordinary procedures we obtain $2 \cdot 2^{10}$ dimensional irreducible representation of the little algebra (1) of $SO(10)$ SPA as follows [2]: $[\underline{1}(+2), \underline{10}(+\frac{3}{2}), \underline{45}(+1), \underline{120}(+\frac{1}{2}), \underline{210}(0), \underline{252}(-\frac{1}{2}), \underline{210}(-1), \underline{120}(-\frac{3}{2}), \underline{45}(-2), \underline{10}(-\frac{5}{2}), \underline{1}(-3)] + [$CPT-conjugate, where $\underline{d}(\lambda)$ represents $SO(10)$ dimension $\underline{d}$ and the helicity λ. By noting that the helicities of these states are automatically determined by $SO(10)$ SPA in the light-cone and that Q^M_1 and $\bar{Q}^M_1$ satisfy the algebra of the annihilation and the creation operators for the massless spin $\frac{1}{2}$ particle, we speculate boldly that these massless states spanned upon the Clifford vacuum $|\ \Omega(\pm 2)\rangle$ are the massless (gravitational) eigenstates of spacetime and matter with $SO(10)$ SP symmetric structure, which are composed of the fundamental massless object Q^N, *superon* with spin $\frac{1}{2}$. Because they correspond merely to all possible nontrivial combinations of the multiplications of the spinor charges (i.e. generators) of $SO(10)$ SP algebra(clustering by a universal force?). Therefore we regard $\underline{5} + \underline{5}^*$ as *a superon-quintet* and *an antisuperon-quintet*. The speculation is dicussed later. To survey the physical implications of superon model for matter we assign tentatively the following SM quantum numbers to superons and adopt the following symbols.

$$\begin{aligned}
\underline{10} &= \underline{5} + \underline{5}^* \\
&= [Q_a(a = 1, 2, 3),\ Q_m(m = 4, 5)] + [Q^*_a(a = 1, 2, 3), Q^*_m(m = 4, 5)]\,, \\
&= \left[\left(\underline{3}, \underline{1}; -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}\right), (\underline{1}, \underline{2}; 1, 0)\right] + \left[\left(\underline{3}^*, \underline{1}; \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right), (\underline{1}, \underline{2}^*; -1, 0)\right],
\end{aligned} \tag{2}$$

where we have specified ($SU(3)$, $SU(2)$; electric charges) and $a = 1, 2, 3$ and $m = 4, 5$ represent the color and electroweak components of superons respectively. Interestingly our model needs only five superons which have the same quantum numbers as the fundamental matter multiplet $\underline{5}$ of $SU(5)$ GUT and satisfy the Gell-Mann-Nishijima relation

$$Q_e = I_z + \frac{1}{2}\ (B - L)\,. \tag{3}$$

Accordingly all $2 \cdot 2^{10}$ states are specified uniquely with respect to ($SU(3)$, $SU(2)$; electric charges). Here we suppose drastically that a field theory of SGM exists and all unnecessary (for SM) higher helicity states become massive in $SU(3) \times SU(2) \times U(1)$ invariant way by eating the lower helicity states corresponding to the superHiggs mechanism and/or to the diagonalizations of the mass terms of the high-spin fields via [$SO(10)$ SPA upon the Clifford vacuum] $\rightarrow$ [$SU(3) \times SU(2) \times U(1)$] $\rightarrow$ [$SU(3) \times U(1)$]. We have carried out the recombinations of the states and found surprisingly that all the massless states necessary for the SM with three generations of quarks and leotons appear in the surviving massless states (therefore, no sterile neutrinos). Among predicted new particles one lepton-type electroweak-doublet (ν_Γ, Γ^-) with spin $\frac{3}{2}$ with the mass of the electroweak scale $(\leq Tev)$ and doubly charged leptons $(> Tev)$ are color singlets and can be observed directly.

As for the assignments of observed particles, we take for simplicity the following left-right symmetric assignment for quarks and leptons by using the conjugate representations naively, i.e. $(\nu_l, l^-)_R = (\bar{\nu}_l, l^+)_L$, etc. [3]. Furthermore as for the generation assignments we assume simply that the states with more (color-) superons turn to acquiring larger masses in the low energy and no a priori mixings among generations. The surviving massless states identified with SM(GUT) are as follows.

For three generations of leptons $[(\nu_e, e), (\nu_\mu, \mu), (\nu_\tau, \tau)]$, we take

$$[(Q_m \varepsilon_{ln} Q_l^* Q_n^*), (Q_m \varepsilon_{ln} Q_l^* Q_n^* Q_a Q_a^*), (Q_a Q_a^* Q_b Q_b^* Q_m^*)] \tag{4}$$

and the conjugate states respectively.

For three generations of quarks $[(u, d), (c, s), (t, b)]$, we have *uniquely*

$$[(\varepsilon_{abc} Q_b^* Q_c^* Q_m^*), (\varepsilon_{abc} Q_b^* Q_c^* Q_l \varepsilon_{mn} Q_m^* Q_n^*), (\varepsilon_{abc} Q_a^* Q_b^* Q_c^* Q_d Q_m^*)] \tag{5}$$

and the conjugate states respectively. For $\mathrm{SU}(2) \times U(1)$ gauge bosons $[W^+, Z, \gamma, W^-]$, $\mathrm{SU}(3)$ color-octet gluons $[G^a (a = 1, 2, \ldots, 8)]$, [SU(2) Higgs Boson], $[(X, Y)]$ leptoquark bosons in GUTs, and a color- and $SU(2)$-singlet neutral gauge boson from $\underline{3} \times \underline{3}^*$ (which we call simply S boson to represent the singlet) we have $[Q_4 Q_5^*, \frac{1}{\sqrt{2}} (Q_4 Q_4^* \pm Q_5 Q_5^*), Q_5 Q_4^*]$, $[Q_1 Q_3^*, Q_2 Q_3^*, -Q_1 Q_2^*, \frac{1}{\sqrt{2}} (Q_1 Q_1^* - Q_2 Q_2^*), Q_2 Q_1^*, \frac{1}{\sqrt{6}} (2 Q_3 Q_3^* - Q_2 Q_2^* - Q_1 Q_1^*), -Q_3 Q_2^*, Q_3 Q_1^*]$, $[\varepsilon_{abc} Q_a Q_b Q_c Q_m]$, $[Q_a^* Q_m]$ and $Q_a Q_a^*$, (and their conjugates) respectively. Now in order to see the potential of superon-graviton model (SGM) as a composite model of matter we try to interpret the Feynman diagrams of SM(GUT) in terms of the superon pictures of all particles in SM(GUT), i.e. a single line of a particle in the Feynman diagrams of SM(GUT) is replaced by multiple lines of superons constituting the particle under two assumptions at the vertex; (i) the analogue of the OZI-rule of the quark model and (ii) the superon number consevation. We find many remarkable results, e.g. in SM, naturalness of the mixing of K^0-$\overline{K^0}$, D^0-$\overline{D^0}$ and B^0-$\overline{B^0}$, no CKM-like mixings among the lepton generations, $\nu_e \leftrightarrow \nu_\mu \leftrightarrow \nu_\tau$ transitions beyond SM, strong CP-violation, small Yukawa couplings and no $\mu \to e + \gamma$ despite compositeness, etc. and in (SUSY)GUT, no dangerous diagrams for proton decay (without R-parity by hand), etc. [2, 3]. SGM may be the most economic model.

3. Fundamental Action of SGM

The supercharges Q of Volkov-Akulov(V-A) model [4] of the nonlinear SUSY(NL SUSY) is given by the supercurrents

$$J^\mu(x) = \frac{1}{i}\ \sigma^\mu \psi(x) - \kappa \left\{ \text{the higher order terms of } \kappa,\ \psi(\mathrm{x}) \text{ and } \partial\psi(x) \right\}. \tag{6}$$

(10) means the field-current identity between the elementary N-G spinor field $\psi(x)$ and the supercurrent, which justifies our bold assumption that the generator (supercharge) Q^N $(N = 1, 2, \ldots 10)$ of $SO(10)$ SPA in the light-cone frame represents the fundamental massless particle, *superon with spin* $\frac{1}{2}$. Therefore the fundamental theory of SGM for spacetime and matter at (above) the Planck scale is $SO(10)$ NL SUSY in the curved spacetime (corresponding to the Clifford vacuum $|\Omega(\pm 2)\rangle\rangle$). We extend the arguments of V-A to high symmetric curved SGM spacetime, where NL SUSY SL(2C) degrees of freedom (i.e. the coset space coordinates representing N-G fermions) $\psi(x)$ in addition to Lorentz $SO(3,1)$ coordinates x^a are embedded at every curved spacetime point with GL(4R) invariance. By defining a new tetrad $w^a{}_\mu(x)$, $w_a{}^\mu(x)$ and a new metric tensor $s^{\mu\nu}(x) \equiv w_a{}^\mu(x)\, w^{a\nu}(x)$ in SGM

spacetime we obtain the following Einstein-Hilbert (E-H) type Lagrangian as the fundamental theory of SGM for spacetime and matter [3].

$$L = -\frac{c^3}{16\pi G}\,|w|\,(\Omega + \Lambda)\,, \tag{7}$$

$$|w| = \det w^a{}_\mu = \det\left(e^a{}_\mu + t^a{}_\mu\right), \qquad t^a{}_\mu = \frac{\kappa}{2i}\sum_{j=1}^{10}\left(\bar{\psi}^j\gamma^a\,\partial_\mu\psi^j - \partial_\mu\bar{\psi}^j\,\gamma^a\psi^j\right), \tag{8}$$

where $i = 1, 2, \ldots, 10$, κ is a fundamental volume of four dimensional spacetime, $e^a{}_\mu(x)$ is the vierbein of Einstein general relativity theory (EGRT) and Λ is a cosmological constant related to the superon-vacuum coupling constant. Ω is a new scalar curvature analogous to the Ricci scalar curvature R of EGRT. The explicit expression of Ω is obtained by just replacing $e^a{}_\mu(x)$ by $w^a{}_\mu(x)$ in Ricci scalar R. The action (7) is invariant at least under GL(4R), local Lorentz, global $SO(10)$ and the following new (NL) SUSY transformation

$$\delta\psi^i(x) = \zeta^i + i\kappa(\bar{\xi}^j\gamma^\rho\psi^j(x))\,\partial_\rho\psi^i(x)\,, \qquad \delta e^a{}_\mu(x) = i\kappa(\bar{\xi}^j\gamma^\rho\psi^j(x))\,\partial_{[\rho}e^a{}_{\mu]}(x)\,, \tag{9}$$

where $\zeta^i, (i = 1, \ldots 10)$ is a constant spinor and $\partial_{[\rho}e^a{}_{\mu]}(x) = \partial_\rho e^a{}_\mu - \partial_\mu e^a{}_\rho$. These results can be understood intuitively by observing that $w^a{}_\mu(x) = e^a{}_\mu(x) + t^a{}_\mu(x)$ defined by $\omega^a = w^a{}_\mu\,dx^\mu$, where ω^a is the NL SUSY invariant differential forms of V-A [4] and $w^a{}_\mu(x)$ and $s^{\mu\nu}(x) \equiv w_a{}^\mu(x)\,w^{a\nu}(x)$ are formally a new vierbein and a new metric tensor in SGM spacetime. In fact, it is not difficult to show the same behaviors of $w_a{}^\mu(x)$ and $s^{\mu\nu}(x)$ as those of $e_a{}^\mu(x)$ and $g^{\mu\nu}(x)$, i.e., $w_a{}^\mu(x)$ and $s^{\mu\nu}(x)$ are invertible, $w_a{}^\mu w_{b\mu} = \eta_{ab}$, $s_{\mu\nu}w_a{}^\mu w_b{}^\mu = \eta_{ab}, \ldots$ etc. and the following GL(4R) transformations of $w^a{}_\mu(x)$ and $s_{\mu\nu}(x)$ under (9) with the field dependent parameters

$$\delta_\zeta w^a{}_\mu = \xi^\nu\,\partial_\nu w^a{}_\mu + \partial_\mu\xi^\nu w^a{}_\nu\,, \qquad \delta_\zeta s_{\mu\nu} = \xi^\kappa\,\partial_\kappa s_{\mu\nu} + \partial_\mu\xi^\kappa s_{\kappa\nu} + \partial_\nu\xi^\kappa s_{\mu\kappa}\,, \tag{10}$$

where $\xi^\rho = i\kappa(\bar{\xi}^j\gamma^\rho\psi^j(x))$. Therefore the similar arguments to EGRT in Riemann space can be carried out straightforwadly by using $s^{\mu\nu}(x)$ (or $w^a{}_\mu(x)$) in stead of $g^{\mu\nu}(x)$ (or $e^a{}_\mu(x)$), which leads to (7) manifestly invariant at least under the above mentioned symmetries, which are isomorphic to $SO(10)$ SP. The commutators of two new supersymmetry transformations on $\psi(x)$ and $e^a{}_\mu(x)$ are the general coordinate transformations

$$[\delta_{\zeta_1}, \delta_{\zeta_2}]\,\psi = \Xi^\mu\,\partial_\mu\psi\,, \qquad [\delta_{\zeta_1}, \delta_{\zeta_2}]\,e^a{}_\mu = \Xi^\rho\,\partial_\rho e^a{}_\mu + e^a{}_\rho\,\partial_\mu\Xi^\rho\,, \tag{11}$$

where Ξ^μ is defined by $\Xi^\mu = 2ia(\bar{\zeta}_2\gamma^\mu\zeta_1) - \xi_1^\rho\xi_2^\sigma e_a{}^\mu(\partial_{[\rho}e^a{}_{\sigma]})$, which form a closed algebra.

In addition, to embed simply the local Lorentz invariance we follow EGRT formally and require that the new vierbein $w^a{}_\mu(x)$ should also have formally a local Lorentz transformation, i.e.,

$$\delta_L w^a{}_\mu = \epsilon^a{}_b w^b{}_\mu \tag{12}$$

with the local Lorentz transformation parameter $\epsilon_{ab}(x) = (1/2)\,\epsilon_{[ab]}(x)$. Interestingly, we find that the following generalized local Lorentz transformations on ψ and $e^a{}_\mu$

$$\delta_L\psi(x) = -\frac{i}{2}\,\epsilon_{ab}\sigma^{ab}\psi\,, \qquad \delta_L e^a{}_\mu(x) = \epsilon^a{}_b e^b{}_\mu + \frac{\kappa}{4}\,\varepsilon^{abcd}\bar{\psi}\gamma_5\gamma_d\psi(\partial_\mu\epsilon_{bc}) \tag{13}$$

are compatible with (12). [Note that the Eq. (13) reduces to the familiar form of the Lorentz transformations if the global transformations are considered, e.g., $\delta_L g_{\mu\nu} = 0$.] Also the

local Lorentz transformation on $e^a{}_\mu(x)$ forms a closed algebra.

$$[\delta_{L_1},\, \delta_{L_2}]\; e^a{}_\mu = \beta^a{}_b e^b{}_\mu + \frac{\kappa}{4}\; \varepsilon^{abcd}\bar{\psi}\gamma_5\gamma_d\psi(\partial_\mu\beta_{bc})\,, \tag{14}$$

where $\beta_{ab} = -\beta_{ba}$ is defined by $\beta_{ab} = \epsilon_{2ac}\epsilon_1{}^c{}_b - \epsilon_{2bc}\epsilon_1{}^c{}_a$. These arguments show that SGM action (7) is invariant at least under [5]

$$[\text{global NL SUSY}] \otimes [\text{local GL}(4,\text{R})] \otimes [\text{local Lorentz}] \otimes [\text{global SO(N)}]\,. \tag{15}$$

$SO(10)$ SP for spacetime and matter is the (isomorphic) case with $N = 10$.

4. Toward Low Energy Theory of SGM

For deriving the low energy behavior of the SGM action it is often useful to linearize such a high nonlinear theory and obtain a low energy effective theory which is renormalizable. Toward the linearization of the SGM we investigate the linearization of V-A model in detail. The linearization of V-A model is investigated [6, 7] and proved that $N = 1$ V-A model of NL SUSY is equivalent to $N = 1$ scalar supermultiplet action of L SUSY which is renormalizable. The general arguments on the constraints which gives the relations between the linear and the nonlinear realizations of global SUSY have been established [6]. Following the general arguments we show explicitly that non-renormalizable $N = 1$ V-A model is equivalent to a renormalizable total action of a U(1) gauge supermultiplet of the linear SUSY [8] with the Fayet-Iliopoulos (F-I) D term indicating a spontaneous SUSY breaking [9]. Remarkably we find that the magnitude of F-I D term (vacuum value) is determined to reproduce the correct sign of V-A action. For $N = 1$, a $U(1)$ gauge field constructed explicitly in terms of N-G fermion fields is an axial vector for $N = 1$.

An $N = 1$ $U(1)$ gauge supermultiplet is given by a real superfield [1]

$$\begin{aligned} V(x,\,\theta,\,\bar{\theta}) = {} & C + i\theta\chi - i\bar{\theta}\bar{\chi} + \frac{1}{2}\, i\theta^2(M + iN) - \frac{1}{2}\, i\bar{\theta}^2(M - iN) - \theta\sigma^m\bar{\theta}v_m \\ & + i\theta^2\bar{\theta}\left(\bar{\lambda} + \frac{1}{2}i\bar{\sigma}^m\,\partial_m\chi\right) - i\bar{\theta}^2\theta\left(\lambda + \frac{1}{2}\, i\sigma^m\,\partial_m\bar{\chi}\right) \\ & + \frac{1}{2}\,\theta^2\bar{\theta}^2\left(D + \frac{1}{2}\,\Box C\right), \end{aligned} \tag{16}$$

where $C(x)$, $M(x)$, $N(x)$, $D(x)$ are real scalar fields, $\chi_\alpha(x)$, $\lambda_\alpha(x)$ and $\bar{\chi}_{\dot{\alpha}}(x)$, $\bar{\lambda}_{\dot{\alpha}}(x)$ are Weyl spinors and their complex conjugates, and $v_m(x)$ is a real vector field. We adopt the notations in Ref. [1]. Following Ref. [6], we define the superfield $\tilde{V}(x,\theta,\bar{\theta})$ by

$$\tilde{V}(x,\,\theta,\,\bar{\theta}) = V(x',\,\theta',\,\bar{\theta}')\,, \tag{17}$$

$$x'^m = x^m + i\kappa(\zeta(x)\;\sigma^m\bar{\theta} - \theta\sigma^m\bar{\zeta}(x))\,, \qquad \theta' = \theta - \kappa\zeta(x)\,, \qquad \bar{\theta}' = \bar{\theta} - \kappa\bar{\zeta}(x)\,. \tag{18}$$

$\tilde{V}$ may be expanded as (16) in component fields $\{\tilde{\phi}_i(x)\} = \{\tilde{C}(x),\, \tilde{\chi}(x),\, \bar{\tilde{\chi}}(x), \ldots\}$, which can be expressed by C, χ, $\bar{\chi}, \ldots$ and ζ, $\bar{\zeta}$ by using the relation (17). κ is now defined with the dimension $(length)^2$. They have the supertransformations of the form

$$\delta\tilde{\phi}_i = -i\kappa\big(\zeta\sigma^m\bar{\epsilon} - \epsilon\sigma^m\bar{\zeta}\big)\;\partial_m\tilde{\phi}_i\,. \tag{19}$$

Therefore, a condition $\tilde{\phi}_i(x) =$ constant is invariant under supertransformations. As we are only interested in the sector which only depends on the N-G fields, we eliminate other degrees of freedom than the N-G fields by imposing SUSY invariant constraints

$$\tilde{C} = \tilde{\chi} = \tilde{M} = \tilde{N} = \tilde{v}_m = \tilde{\lambda} = 0\,, \qquad \tilde{D} = \frac{1}{\kappa}\,. \tag{20}$$

Solving these constraints we find that the original component fields C, χ, $\bar{\chi}$, ... can be expressed by the N-G fields ζ, $\bar{\zeta}$. Among them, the leading terms in the expansion of the fields v_m, λ, $\bar{\lambda}$ and D, which contain gauge invariant degrees of freedom, in κ are

$$\begin{aligned} v_m &= \kappa\zeta\sigma_m\bar{\zeta} + \cdot\,, \qquad \lambda = i\zeta - \frac{1}{2}\,\kappa^2\zeta\big(\zeta\sigma^m\,\partial_m\bar{\zeta} - \partial_m\zeta\,\sigma^m\bar{\zeta}\big) + \cdot\,, \\ D &= \frac{1}{\kappa} + i\kappa\big(\zeta\sigma^m\,\partial_m\bar{\zeta} - \partial_m\zeta\,\sigma^m\bar{\zeta}\big) + \cdot\,, \end{aligned} \tag{21}$$

where $\cdot$ are higher order terms in κ. Our discussion so far does not depend on a particular form of the action. We now consider a free action of a U(1) gauge supermultiplet of L SUSY with a Fayet-Iliopoulos D term. In component fields we have

$$S = \int d^4x\left[-\frac{1}{4}\,v_{mn}v^{mn} - i\lambda\sigma^m\,\partial_m\bar{\lambda} + \frac{1}{2}\,D^2 - \frac{1}{\kappa}\,D\right]. \tag{22}$$

The last term proportional to κ^{-1} is the Fayet-Iliopoulos D term. The field equation for D gives $D = \frac{1}{\kappa} \neq 0$ in accordance with Eq. (20), which indicates the spontaneous breakdown of supersymmetry. We substitute Eq. (21) into the action (22) and obtain an action for the N-G fields ζ, $\bar{\zeta}$ which is exactly $N = 1$ V-A action

$$S = -\frac{1}{2\kappa^2}\int d^4x\,\det\left[\delta^n_m + i\kappa^2(\zeta\sigma^n\,\partial_m\bar{\zeta} - \partial_m\zeta\,\sigma^n\bar{\zeta})\right]. \tag{23}$$

For $N = 1$, U(1) gauge field becomes $v_m \sim \kappa\bar{\zeta}\gamma_m\gamma_5\zeta + \ldots$ in the four-component spinor notation, which is an axial vector. These are very suggestive and favourable to SGM.

5. Discussion

SGM action in SGM spacetime is a nontrivial generalization of E-H action in Riemann spacetime despite the liner relation $w^a{}_\mu = e^a{}_\mu + t^a{}_\mu$. In fact, by the redefinition $e^a{}_\mu \to e^a{}_\mu - t^a{}_\mu$ the inverse $w_a{}^\mu$ does not reduce to $e_a{}^\mu$, i.e. interestingly the higher order nonlinear terms in $t^\mu{}_a(\neq t_a{}^\mu)$ in the inverse $w_a{}^\mu$ can not be eliminated. Because $t^a{}_\mu$ is not a metric. Such a redefinition breaks the metric properties of $w^a{}_\mu$ and $w_a{}^\mu$. Note that SGM action (7) posesses two inequivalent flat spaces, i.e. SGM-flat $w^a{}_\mu \to \delta^a{}_\mu$ and Riemann-flat $e^a{}_\mu \to \delta^a{}_\mu$. The expansion of (7) in terms of $e^a{}_\mu$ and $t^a{}_\mu$ is a spontaneous breakdown of spacetime (7) due to the degeneracy (9) from SGM to Riemann connecting with Riemann-flat spacetime [10]. SGM (and V-A model with $N > 1$) posesses rich structures and the potential for defining completely the renormalizable (broken SUSY) models of the local field theory containing the massive high spin field by the lineralization in the curved spacetime. SGM for spin $\frac{3}{2}$ N-G fermion [11] and SGM with the extra dimensions to be compactified are also in the same scope. SGM cosmology is open.

References

[1] J. WESS and J. BAGGER, *Supersymmetry and Supergravity* (Princeton Univ. Press, 1992).
[2] K. SHIMA, Z. Phys. **C18** (1983) 25.
K. SHIMA, European. Phys. J. **C7** (1999) 341.
[3] K. SHIMA, Phys. Lett. **B501** (2001) 237.
[4] D. V. VOLKOV and V. P. AKULOV, Phys. Lett. **B46** (1973) 109.
[5] K. SHIMA and M. TSUDA, Phys. Lett. **B507** (2001) 260.
[6] E. A. IVANOV and A. A. KAPUSTNIKOV, J. Phys. **A11** (1978) 2375.
[7] M. RŐCEK, Phys. Rev. Lett. **41** (1978) 451.
[8] J. WESS and B. ZUMINO, Nucl. Phys. **B78** (1974) 1.
[9] K. SHIMA, Y. TANII, and M. TSUDA, To appear in Phys. Lett. **B** (2002). hep-th/0110102.
[10] K. SHIMA and M. TSUDA, hep-th/0109042.
[11] K. SHIMA and M. TSUDA, Phys. Lett. **B521** (2001) 67.

Fortschr. Phys. **50** (2002) 5–7, 724–728

The Heisenberg Algebra and Spin

DMITRI SOROKIN[1,2)]

[1)] Institute for Theoretical Physics, NSC KIPT, Akademicheskaya Street 1
61108 Kharkov, Ukraine
[2)] INFN, Sezione Di Padova Via F. Marzolo, 8, I-35131 Padova, Italy
sorokin@pd.infn.it

Abstract

Examples of the use of the Heisenberg algebra for the description of relativistic particles and fields of arbitrary spin in three- and four-dimensional space-time are considered.

1. Introduction

The Heisenberg algebra is well known to be the mathematical expression for the one of fundamental principles of Quantum Mechanics, the uncertainty principle, descovered by W. Heisenberg in 1927 [1]. The uncertainty relation

$$[x, p] = xp - pq = i\hbar \tag{1}$$

reads that the spatial coordinate x and the momentum p of a quantum particle along the same axis cannot simultaneously have definite values. In this contribution I would like to demonstrate that the Heisenberg algebra is not only the basis for the description of quantum mechanical properties of particles in space and time, but is also suitable for the description of their spins. We shall consider two examples of the use of the Heisenberg algebra for the description of spin. In a three-dimensional space-time ($D = 2 + 1$) it describes particles with fractional quantum statistics and fractional spin $\frac{1}{4}$ and $\frac{3}{4}$ [2, 3, 4]. These particles are called semions (or quartions) and are a special class of anyons [5], planar quasiparticles assumed to be responsible for the fractional quantum Hall effect and High-T_C superconductivity (see e.g. [6]). The second example will be the description of higher-spin particles and fields in four-dimensional relativistic field theory [7, 8].

2. Group-theoretical Description of Semions [3]

Dynamical properties of relativistic particles are described by the momentum P_m and the orbital momentum $M_{mn} = \frac{1}{2}\,(x_m P_n - P_m x_n)$ or, in $D = 2 + 1$, by $M_l = \epsilon_{lmn} M^{mn}$ $(l, m, n = 0, 1, 2)$. P_m and M_m form the $D = 2 + 1$ Poincare algebra

$$[P_m, P_n] = 0\,, \quad [M_m, P_n] = i\epsilon_{mnl} P^l\,, \quad [M_m, M_n] = i\epsilon_{mnl} M^l\,. \tag{2}$$

The orbital momenta M_m themselves form the Lorentz algebra $SO(1, 2)$. To describe spin degrees of freedom one should introduce a spin momentum S_m satisfying the same commutation relations as M_m

$$[S_m, S_n] = i\epsilon_{mnl} S^l\,,$$

100 Years Werner Heisenberg: Works and Impact. Edited by D. Papenfuß, D. Lüst, W.P. Schleich
Copyright © 2002 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32065-3

$J_m = M_m + S_m$ being the total angular momentum of the particle (or field). We now show that the spin of the semions is described by spinning variables q and ρ satisfying the Heisenberg algebra $[q, \rho] = i$ (for simplicity we put $\hbar = 1$). For the description to be manifestly $SO(1,2)$ invariant let us combine q and ρ into a two–component hermitian spinor operator

$$\lambda_\alpha = (q, \rho)\,, \quad (\alpha = 1,2) \in SL(2, R) \sim SO(1,2)\,. \tag{3}$$

Then the Heisenberg algebra takes the form

$$[\lambda_\alpha, \lambda_\beta] = i\epsilon_{\alpha\beta}\,, \qquad \epsilon_{12} = -\epsilon_{21} = 1\,. \tag{4}$$

The spin momentum is composed of λ_α as follows

$$S_m = \lambda_\alpha \gamma_m^{\alpha\beta} \lambda_\beta\,, \quad \text{or} \quad S_{\alpha\beta} = \gamma_{\alpha\beta}^m S_m = \frac{1}{4}\,(\lambda_\alpha\lambda_\beta + \lambda_\beta\lambda_\alpha)\,, \tag{5}$$

where $\gamma_m^{\alpha\beta}$ are $D = 2 + 1$ Dirac matrices. The Casimir operator $S_m S^m$ of this $SL(2,R)$ representation can be computed to have the following eigenvalue

$$\langle S_m S^m \rangle = s_0(s_0 - 1) = \frac{3}{16}, \tag{6}$$

from which we conclude that $s_0 = \frac{1}{4}$ or $s_0 = \frac{3}{4}$. This means that the spins of quantum states described by the operator (5) are $\frac{1}{4}$ and $\frac{3}{4}$. These are the fractional spins of the semions. The classical dynamics of relativistic particles of mass m whose quantization gives the semionic states with spin $\frac{1}{4}$ and $\frac{3}{4}$ is described by the action [3]

$$S = \int d\tau \left[p_m \dot{x}^m + e(\tau)\,(p_m p^m + m^2) + c(\tau)\left(\frac{1}{4}\,p_m \lambda \gamma^m \lambda - s_0 m\right) + \lambda^\alpha \dot{\lambda}_\alpha \right], \tag{7}$$

where $e(\tau)$ and $c(\tau)$ are Lagrange multipliers. The quantization of the model based on the action (7) results in the following wave equation for semion states [2, 3]

$$(P_m \gamma_{\alpha\beta}^m - im\epsilon_{\alpha\beta})\,\lambda^\beta \Phi(x, \lambda) = 0 \tag{8}$$

which resembles the Dirac equation but where λ_α is the spinorial operator generating the Heisenberg algebra (4). This equation singles out a state with the 'helicity' $\frac{1}{m} P_m S^m = \frac{1}{4}$ as we shall now show. The integrability condition of (8) is the Klein-Gordon equation for the field of a mass m

$$(P_m P^m + m^2)\;\Phi(x, \lambda) = 0\,. \tag{9}$$

Therefore we can choose the rest frame $P_m = (m, 0, 0)$ in which Eq. (8) reduces to

$$(\lambda^\alpha - i\epsilon^{\alpha\beta}\lambda^\beta)\;\Phi(x, \lambda) = 0 \quad \Rightarrow \quad \frac{1}{\sqrt{2}}\,(q - ip)\;\Phi(x, \lambda) = a\Phi(x, \lambda) = 0\,, \tag{10}$$

where $a = \frac{1}{\sqrt{2}}\,(q - ip)$ can be regarded as an annihilation operator, the creation operator being naturally $a^+ = \frac{1}{\sqrt{2}}\,(q + ip)$. Thus the semion wave function can be represented as the wave function of a one-dimensional harmonic oscillator which is the sum of states with

even and odd powers of a^+. It forms an irreducible representation of the Heisenberg-Weil group generated by the Heisenberg algebra

$$\Phi = \Phi_{\frac{1}{4}} + \Phi_{\frac{3}{4}}\,, \tag{11}$$

where

$$\Phi_{\frac{1}{4}} = \sum_{n=0}^{\infty} f^{2n}(x)\ (a^+)^{2n}\ |0\rangle \quad \text{and} \quad \Phi_{\frac{3}{4}} = \sum_{n=0}^{\infty} f^{2n+1}(x)\ (a^+)^{2n+1}\ |0\rangle$$

take values in the discrete $1/4$ and $3/4$ representation of the $SL(2,R)$ group, respectively, the helicity of an nth state being $s = \frac{1}{2}(n+\frac{1}{2})$. The semion wave Eq. (8), requires that all the states in the infinite tower (11) are zero except for the lowest $s = \frac{1}{4}$ helicity state $|0\rangle$ which is annihilated by a. To include into consideration the semion with the spin $\frac{3}{4}$ it is natural to consider it as the superpartner of the $\frac{1}{4}$ semion (because the spins of these two differ by $1/2$). Supersymmetric models of semions have been constructed in [3, 4]. The anyons of arbitrary fractional spin can be described (see e.g. [9] for details) by the deformed Heisenberg algebra [10]

$$[q,\,\rho] = i(1+\nu R)\,, \qquad R^2 = 1\,, \qquad Rq + qR = R\rho + \rho R = 0\,,$$

where ν is a numerical deformation parameter and R is a reflection operator.

3. Higher Spin Superparticles in $D = 4$ [8]

In the previous example we considered λ_α to be hermitian (real) spinors in the fundamental representation of $SL(2,R)$. Consider now λ_α to be a complex (Weyl) spinor $(\lambda_\alpha)^* = \bar{\lambda}_{\dot\alpha}$ transforming under a fundamental representation of $SL(2,C) \sim SO(1,3)$, the Lorentz group in four-dimensional space-time. The Heisenberg algebra becomes

$$[\lambda_\alpha,\,\lambda_\beta] = i\epsilon_{\alpha\beta}\,, \qquad [\bar{\lambda}_{\dot\alpha},\,\bar{\lambda}_{\dot\beta}] = i\epsilon_{\dot\alpha\dot\beta}\,, \qquad [\lambda_\alpha,\,\bar{\lambda}_{\dot\beta}] = 0\,, \tag{12}$$

Combine λ_α and $\bar{\lambda}_{\dot\alpha}$ into a Majorana bi-spinor $\lambda_{\hat\alpha} = \frac{1}{\sqrt{r}}\ (\lambda_\alpha,\,\bar{\lambda}^{\dot\alpha})$, where r will be interpreted as an anti-de-Sitter space radius, and $\hat\alpha = 1,\ldots,4$. The Heisenberg algebra takes the form

$$[\lambda_{\hat\alpha},\,\lambda_{\hat\beta}] = \frac{i}{r}\ C_{\hat\alpha\hat\beta} = \begin{pmatrix} \epsilon_{\alpha\beta} & 0 \\ 0 & \epsilon^{\dot\alpha\dot\beta} \end{pmatrix}. \tag{13}$$

Actually, $\lambda_{\hat\alpha}$ takes values in the fundamental representation of a larger group $Sp(4) \sim SO(2,3) \supset SL(2,C)$ which is the isometry group of an AdS_4 space of a radius r. The generators of $Sp(4)$ can be realized as belinears of $\lambda_{\hat\alpha}$

$$M_{\hat\alpha\hat\beta} = \frac{1}{4}\ (\lambda_{\hat\alpha}\lambda_{\hat\beta} + \lambda_{\hat\beta}\lambda_{\hat\alpha}) = -2\gamma^a_{\hat\alpha\hat\beta}P^a + \gamma^{ab}_{\hat\alpha\hat\beta}Z_{ab}\,, \tag{14}$$

where $\gamma^{ab} = \frac{1}{2}(\gamma^a\gamma^b - \gamma^b\gamma^a)$, P_a are the generators of the boosts in AdS_4 and Z_{ab} are the generators of the AdS_4 holonomy group $SO(1,3)$ $(a,b = 0,1,2,3)$. The $Sp(4)$ algebra is

$$[M_{\hat\alpha\hat\beta},\,M_{\hat\gamma\hat\delta}] = -\frac{4i}{r}\ (C_{\hat\gamma(\alpha}M_{\hat\beta)\hat\delta} + C_{\hat\delta(\alpha}M_{\hat\beta)\hat\gamma})\,. \tag{15}$$

In the limit $r \to \infty$ the AdS_4 space becomes flat, and the generators P_a and Z_{ab} become commutative, the former playing the role of the momentum generators in the flat space M_4. From the relations (13) and (14) we find that at $r \to \infty$

$$P_a = \lambda\gamma_a\lambda \quad \Rightarrow \quad P_aP^a \equiv 0 \tag{16}$$

due to $D = 4$ Dirac matrix identities. We thus conclude that P_a are momenta of massless states in M_4. To understand the algebraic and physical meaning of the commuting generators Z_{ab} let us consider a supersymmetric extension of the above construction. For this we introduce a Grassmann-odd scalar χ with the following properties

$$\chi^2 = \frac{1}{2}, \qquad \chi\lambda_{\hat\alpha} - \lambda_{\hat\alpha}\chi = 0 \tag{17}$$

and construct a Grassmann-odd supercharge $Q_{\hat\alpha} = \chi\lambda_{\hat\alpha}$ whose anticommutator is

$$\{Q_{\hat\alpha}, Q_{\hat\beta}\} = M_{\hat\alpha\hat\beta} = -2\gamma^a_{\hat\alpha\hat\beta}P^a + \gamma^{ab}_{\hat\alpha\hat\beta}Z_{ab}\,. \tag{18}$$

We thus get the superalgebra $OSp(1\,|\,4)$[1)] which at $r \to \infty$ reduces to the $N = 1$, $D = 4$ super Poincare algebra with the tensorial central charge

$$Z_{ab} = \lambda\gamma_{ab}\lambda\,. \tag{19}$$

We now construct a superparticle model where Z_{ab} plays the role of momenta associated with 'internal' tensorial coordinates y^{ab} describing its spinning degrees of freedom. The massless superparticle is assumed to move in $N = 1$, $D = 4$ superspace parametrized by bosonic coordinates x^a and fermionic coordinates $\theta_{\hat\alpha}$ and enlarged with the tensorial coordinates y^{ab}. The model is described by the following action [7]

$$S = \int d\tau\,[\lambda\gamma_a\lambda(\dot x^a + i\theta\gamma\dot\theta) + \lambda\gamma_{ab}\lambda(\dot y^{ab} + i\theta\gamma^{ab}\theta)]\,, \tag{20}$$

which implies that (16) and (19) are momenta conjugate to x^a and y^{ab}, respectively. The supercharge is $Q_{\hat\alpha} = \pi_{\hat\alpha} + i(\theta^{\hat\beta}\lambda_{\hat\beta})\,\lambda_{\hat\alpha}$, where $\pi_{\hat\alpha}$ is the momentum conjugate to $\theta^{\hat\alpha}$ and $\theta^{\hat\beta}\lambda_{\hat\beta} = \chi$ is associated with the classical counterpart of the Grassmann odd scalar operator (17). As it has been shown in [8] the dynamical constraints and the quantization of the system results in the quantum spectrum which consists of massless states of arbitrary integer and half integer spin described in the momentum representation by the following wave function

$$\Phi = \sum_{n=0}^{\infty} \Phi^{\hat\alpha_1\ldots\hat\alpha_n}(p_a)\lambda_{\hat\alpha_1}\ldots\lambda_{\hat\alpha_n} + \chi\sum_{n=0}^{\infty}\Psi^{\hat\alpha_1\ldots\hat\alpha_n}(p_a)\,\lambda_{\hat\alpha_1}\ldots\lambda_{\hat\alpha_n}\,. \tag{21}$$

The states with an even n have integer spins $n = 2s_{\text{integer}}$ and the states with an odd n have half integer spins $n = 2(k+\frac{1}{2}) = 2s_{\frac{1}{2}\text{integer}}$. The model is the classical counterpart of the quantum field theory of higher spins developed by M. Vasiliev (see [11] and references therein). One more particular feature of this model is that it describes exotic BPS states which preserve $3/4$ of $N = 1$, $D = 4$ supersymmetry [7], while usually superparticles (and in general superbranes) preserve one half or less target space supersymmetry. By now this model (and its higher dimensional analogs) is the only known example of such BPS config-

[1)] Note that $(\lambda_{\hat\alpha}, \chi)$ transform under the fundamental (singleton) representation of the $OSp(1\,|\,4)$ supergroup.

urations. In conclusion we have demonstrated that the Heisenberg algebra has proved to be not only a basic ingredient of Quantum Mechanics but also very useful for the construction of models of particles and fields of arbitrary (fractional and higher) spins.

Acknowledgements

The author is grateful to the Alexander von Humboldt Foundation for the invitation and financial support to attend the Werner Heisenberg Symposium and to visit the Humboldt University of Berlin. This work was partially supported by the Grant N 383 of the Ukrainian State Fund for Fundamental Research, by the European Commission TMR Programme HPRN-CT-2000-00131 to which the author is associated with the University of Padua and by the INTAS Research Project N 2000-254.

References

[1] W. Heisenberg, Z. Phys. **43** (1927) 172.
[2] D. V. Volkov, JETP Lett. **49** (1989) 473.
[3] D. P. Sorokin and D. V. Volkov, Nucl. Phys. B **409** (1993) 547.
[4] D. P. Sorokin and D. V. Volkov, JETP Lett. **57** (1993) 343; Int. J. Mod. Phys. A **9** (1994) 1555.
[5] M. G. Laidlow and C. M. de Wit, Phys. Rev. **D3** (1971) 1375;
J. Leinaas and J. Mirheim, Nuovo Cimento **B37** (1971) 1;
F. Wilczek, Phys. Rev. Lett. **48** (1982) 1144; **49** (1982) 1;
R. Mackenzie and F. Wilczek, Int. J. Mod. Phys. **A3** (1988) 2827;
R. Iengo and K. Lechner, Phys. Rep. **213** (1992) 179.
[6] L. B. Laughlin, Science **242** (1988) 525. F. Wilczek, Fractional statistics and anyon superconductivity, World Scientific, Singapore, 1990.
[7] I. Bandos and J. Lukierski, Mod. Phys. Lett. **14** (1999) 1257.
[8] I. Bandos, J. Lukierski, and D. Sorokin, Phys. Rev. D **61** (2000) 045002; hep-th/9912264.
[9] M. S. Plyushchay, Phys. Lett. B **320** (1994) 91; Nucl. Phys. B **491** (1997) 619; Mod. Phys. Lett. A **12** (1997) 1153; Ann. Phys. **245** (1996) 339. S. M. Klishevich, M. S. Plyushchay, and M. Rausch de Traubenberg, Nucl. Phys. B **616** (2001) 419.
[10] Y. Ohnuki and S. Kamefuchi, *Quantum Field Theory and Parastatistics* (University Press of Tokyo, 1982).
M. Vasiliev, JETP Lett. **50** (1989) 344; Int. J. Mod. Phys. **A6** (1991) 1115.
[11] M. A. Vasiliev, hep-th/0106149; hep-th/0111119; Nucl. Phys. B **616** (2001) 106.

D. Papenfuß, D. Lüst, W. P. Schleich (eds.)

100 Years Werner Heisenberg

Works and Impact

The volume contains 53 selected papers presented at an international conference held in Bamberg/Germany to commemorate Werner Heisenberg's 100th birthday. The meeting was organized by the Alexander von Humboldt-Foundation in Bonn/Germany. Heisenberg was the president of the Foundation from 1953 to 1975.

Werner Heisenberg (1901–1976) is the undisputed leader in the new field of Quantum Physics. The conference emphasized his many major contributions to physics and at the same time illuminated the latest developments in Quantum Physics. Several plenary talks tried to put Heisenberg in the proper historical perspective and three parallel sessions were dedicated to Elementary Particles, Quantum Physics, and Quantum Field Theory and Gravitation.

ISBN 3-527-40392-2

9 783527 403929

www.wiley-vch.de

WILEY-VCH