

Publishers' page

Publishers' page

Publishers' page

Publishers' page



Hagen Kleinert

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

Preface

The subject of this volume embodies a wide variety of research topics whose theoretical methods are based on path integrals, or more generally *Fluctuating Paths and Fields* – a heading which reflects best the numerous research areas of Hagen Kleinert to whom this *Festschrift* is devoted. The selection of authors and articles is guided by the difficult task to present a cross section of his research activities. With a few prominent exceptions, the contributors to this volume are his former and present students, postdocs, co-workers, or colleagues who have enjoyed working with him. With their articles they wish to celebrate his 60th birthday on June 15, 2001. We are especially grateful to the Physics Nobel Laureate Gerard 't Hooft for delivering the birthday colloquium at the Freie Universität Berlin.^a

The first part of the book contains recent advances in the path-integral description of quantum physics, covering and supplementing a subset of the topics discussed in Hagen Kleinert's widespread textbook on *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*. Among Kleinert's many contributions to this field, the best known is his path-integral solution of the three-dimensional hydrogen atom based on non-holonomic space-time transformations in 1979. The clue to the successful treatment was his deep knowledge of the relation between three-dimensional Coulomb and four-dimensional oscillator systems which was established in the context of celestial mechanics by Kustaanheimo and Stiefel. In standard quantum mechanics, this relation is rooted in the existence of the dynamical group $O(4, 2)$ of the hydrogen atom which Kleinert discovered in collaboration with the late Asim O. Barut during the years 1965 to 1967, when working on his Ph.D. thesis at the University of Colorado at Boulder. The main ideas for the path-integral solution are contained in a paper of 1979, written with his postdoc Ismail H. Duru. A fully consistent mathematical formulation, however, turned

^aThe abstract of this lecture is printed on p. xiii.

out to be quite subtle. In particular, it required an understanding of the path integration measure in curved space-time with torsion. The complete solution of the problem was found by Kleinert in the late eighties. Generalizing the theory allowed him to calculate path integrals of all systems for which the Schrödinger equation can be solved analytically. A full account of the theory can be found in his textbook on path integrals.

The impetus to tackle this fundamental problem of atomic physics goes back to Kleinert's discussions with Richard P. Feynman during a sabbatical at Caltech in 1973/74. As described in the Feynman biography by John and Mary Gribbin, "... Feynman had stopped teaching path integrals at a less advanced level, because he never derived a complete path integral description of the hydrogen atom, and was embarrassed by this failure. ... Kleinert ... not only solved the problem (much to Feynman's delight), but wrote a major textbook on the path integral approach, re-establishing path integrals as a research tool, not only conceptually useful but now capable of solving problems ..."^b Kleinert was hired to Pasadena by Murray Gell-Mann whom he had met before from 1968 to 1972 during various extended visits at CERN in Geneva, which was and continues to be one of the main centers for meeting important people (... such as his wife Annemarie). At that time, Gell-Mann was interested in a relation between current and constituent quarks derived by Kleinert and his colleagues at CERN, two of them, Franco Buccella and Carlos A. Savoy, having written articles for this *Festschrift*. In Pasadena, Kleinert shared office with Yuval Ne'eman whose pioneering work on $SU(3)$ had paved the way to Gell-Mann's development of a quark field theory and his successful prediction of the Ω^- particle. During that time Kleinert found a theoretical explanation for the algebra of Regge residues proposed by Ne'eman in collaboration with Cabibbo and Horwitz. This led to a deep friendship with Ne'eman, and even in his period as a Minister of Science and Technology in Israel, he found time to visit Kleinert in Berlin for discussing physics (see his contribution in this *Festschrift*).

The articles in Part I cover topics ranging from rigorous definitions of functional integrals over applications to group spaces, quasi-classical approximations, and semi-classical dynamics to numerical Monte Carlo evaluations of path integrals, which Kleinert, while working mainly analytically, often enjoyed performing on his personal computer at home. One of the editors

^bJ. Gribbin and M. Gribbin, *Richard Feynman: A Life in Science* (Viking Penguin Books, London, 1997), pp. 215-216.

(W.J.) took part in a race of two good-old Atari computers against each other – one at the institute and the other at Kleinert’s home, the latter PC, of course, turning out to be newer and thus faster . . .

Part II of this volume collects articles on quantum field theory which has always been the backbone of Hagen Kleinert’s research. In fact, this was one of the reasons why Werner Theis, then one of the senior professors of theoretical physics at the Freie Universität Berlin, asked the 27 years-old Kleinert to join the faculty as an Associate Professor. After turning down two other offers for a full professorship elsewhere, he became a Full Professor in Berlin in 1976. Since then Kleinert, who grew up in Hannover and received his undergraduate education at the Technische Universität, became a real “Berliner”, refusing two further attractive offers from other universities in Germany and abroad. Kleinert’s early applications of field theoretic methods include the introduction of composite fields via path-integral transformations in models of elementary particle physics which he called “hadronization of quark theories” and, by analogy, in models of superconductors and superfluid Helium. Field theoretic methods involving gauge fields play a crucial role in Kleinert’s comprehensive work on defect-driven phase transitions which is collected in his two extensive monographs on *Gauge Fields in Condensed Matter* – Vol. I: *Superflow and Vortex Lines*, Vol. II: *Stresses and Defects*. In this context non-holonomic mappings are again of central importance. The articles in Part II discuss gauge theories, in particular quantum electrodynamics and quantum chromodynamics, and in some more detail ϕ^4 -field theories, where Kleinert and his collaborators contributed most over the recent years.

Variational perturbation theory and the resulting highly efficient scheme for resumming divergent perturbation series are discussed in Part III. This technique has its roots in a variational treatment of quantum-statistical partition functions in the path-integral representation developed by Kleinert in joint work with Feynman during his later sabbaticals at the Universities of California at Berkeley in 1979/80 and at Santa Barbara in 1982/83. After a considerable delay due to Feynman’s illness, the manuscript was finally finished and submitted for publication in *Physical Review* during Kleinert’s sabbatical stay at San Diego in 1985/86. This last period of intense scientific interactions with Feynman is also witnessed by the famous photographs of Feynman’s last office blackboards.^c The joint theory is an improvement of a 20 years-old less powerful approach found by Feynman himself. It is an amus-

^c *Feynman’s Office: The Last Blackboards*, in *Physics Today* **42**, 88 (February 1989).

ing fact of scientific history that, after such a long time, the same idea for an improved version was worked out independently also by Riccardo Giachetti and Valerio Tognetti in Florence in 1985/86. While subsequently the Italian group was mainly interested in applications to solid state physics, Kleinert pushed the method to higher orders and made it applicable to quantum field theory. Finally he arrived at an extremely powerful formulation of variational perturbation theory which may be viewed as a highly efficient resummation procedure for divergent, asymptotic perturbation series. Besides impressive applications in quantum mechanics and atomic physics, this has recently led to a successful treatment of ϕ^4 -field theories in the strong-coupling limit. The results are collected in a monograph on *Critical Properties of ϕ^4 -Theories*, written in collaboration with Verena Schulte-Frohlinde, who also contributes here.

The articles in Part IV deal with phase transitions and critical phenomena, many of them centering around explicit calculations of critical exponents of ϕ^4 -theories with and without anisotropies or quenched, random disorder. They have important applications to phase transitions in magnets and the λ -transition in liquid ^4He . Others are devoted to the related Ginzburg-Landau model and phase transitions in superconductors and superfluid ^3He where topological defects play a crucial role. Here, Kleinert's development of a disorder field theory dual to Ginzburg-Landau's order field theory has led to his discovery of the tricritical point in superconductors in 1982. By studying the limit of strongly bound electron pairs as the origin of high- T_c superconductivity he pointed out that pairing and phase decoherence may occur at different temperatures. In ^3He he found, together with Kazumi Maki who has written an article in this part, an interesting helical texture whose existence was subsequently confirmed experimentally. To explain the blue phase of cholesteric liquid crystals, Kleinert and Maki furthermore investigated icosahedral symmetries. Such structures were later discovered in sputtered aluminum and are now called quasicrystals. In between their joint calculations Maki especially enjoyed singing Italian operas with Kleinert, who knows a great number of them by heart.

Part V comprises articles on topological defects and their role in phase transitions as well as on fluctuating strings and membranes. Defects play an important role in crystal melting, superfluidity, superconductivity, cosmology, and elementary particle physics as described in Kleinert's above-mentioned monographs. Several new developments are discussed there. As far as strings and membranes are concerned the key-point of Kleinert's research emphasized

the structural similarity between real membranes and the surfaces swept out by color-electric strings which hold quarks together. This led to the by now famous Polyakov-Kleinert action for curvature dominated membranes and strings. Many field theoretic techniques could be successfully transferred between these physically quite different fields. In particular, a fundamental constant which rules the so-called Helfrich pressure exerted by fluctuating membranes was thus found analytically for the first time in good agreement with earlier Monte Carlo simulations.

The final Part VI deals with questions related to gravitation, cosmology, and astrophysics. Here Kleinert's research interests focused mainly around extending Einstein's gravitational theory in such a way that also the torsion of space-time can be taken into account. In close analogy to forces between defects in real crystals he developed a replacement and extension of Einstein's equivalence principle in the form of a non-holonomic mapping principle which permits deriving laws of Nature in geometries with curvature and torsion from those in flat space-time. With using non-holonomic mappings, Kleinert closed the circle between gravity, the theory of dislocations and disclinations, and his early path-integral solution of the hydrogen atom. In gravity, he also obtained interesting consequences for the early universe by deriving a gravitational action which differs from Einstein's at high curvature as a result of quantum fluctuations of elementary fields.

The comprehensive work of Hagen Kleinert is witnessed by his four books and more than 300 research papers, all accessible on his internet homepage

<http://www.physik.fu-berlin.de/~kleinert>.

His work cannot be fully reflected by the selected contributions in this *Festschrift*. It contains 65 articles arranged in each part in topical order and further linked together through an extensive index. We can only hope that in this way at least a flavor of Kleinert's diverse research interests and achievements becomes visible. We refrain from reproducing his current list of publications as it continues to grow rapidly and would be outdated at the time of printing.

We are grateful to all authors for their immediate agreement to contribute an article. Many of them emphasize Kleinert's stimulating influence on their research. His Italian-, Spanish-, French-, and certainly English-speaking visitors all have an easy time discussing with him since he converses with them in their own language. The number of former and present students, postdocs, collaborators, and friends of Hagen Kleinert is so large that we were unable

to ask all of them to contribute to this book. We apologize to all those we have missed to contact. Financial support from the HSP III program of the German Ministry of Education and Research (BMBF) is appreciated. We thank Sieglinde Endrias for reading the manuscript and spotting many printing errors, Renate Schmidt and Lea Voigt for secretarial help as well as Ms. E.H. Chionh for the pleasant negotiations with the publishing company. Finally, we acknowledge the constant interest and valuable advice of Dr. Annemarie Kleinert. She was very helpful in the editing process and brought several former collaborators of her husband to our attention.

Leipzig, Berlin, and Potsdam, April 2001

Wolfhard Janke
Axel Pelster
Hans-Jürgen Schmidt
Michael Bachmann

Editors

Special Lecture

**How Does God Throw Dice? Speculations About
Quantum Mechanics at the Planck Scale.**

of Nobel Laureate

Gerard 't Hooft

*Spinoza Institute, Utrecht University,
Leuvenlaan 4, Postbus 80.195, 3508 TD Utrecht, Netherlands
E-mail: g.thooft@phys.uu.nl*

in honor of

Hagen Kleinert

on the occasion of his 60th birthday on June 15th, 2001
at the Freie Universität Berlin

Attempts to arrive at consistent theories combining quantum mechanics with general relativity not only require new concepts of space, time, and matter, such as the ideas that lead to superstring theory, D-brane theory, and M-theory, but they may also require a reconsideration of what quantum mechanics itself really is about. Although completely deterministic scenarios appear to be ruled out by the Bell inequalities, it is nevertheless worth-while to investigate a set-up where we start with a deterministic theory and add to this the notion of information loss. Models proposed so-far all show deficiencies of some sort which make them unrealistic for describing the real world, but they do show how chaotic phenomena in a deterministic theory might be found to lie at the basis of the quantum nature of our world.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

Contents

Part I: Path Integrals and Quantum Mechanics

Semiclassical Quantum Mechanics: A Path-Integral Approach	3
<i>Barry R. Holstein</i>	
Forward-Backward Semiclassical Dynamics	15
<i>Nancy Makri</i>	
The Feynman Path Goes Monte Carlo	29
<i>Tilman Sauer</i>	
Coordinate Independence of Path Integrals in One-Dimensional Target Space	43
<i>Alexander Chervyakov</i>	
Perturbatively Defined Path Integral in Phase Space	57
<i>Michael Bachmann</i>	
The Use of Non-Relativistic Path Integrals in Field Theories	69
<i>Ismail H. Duru</i>	
Path Integration and Coherent States for the 5D Hydrogen Atom	73
<i>Nuri Ünal</i>	
The Aharonov-Bohm Effect in Four Dimensions	83
<i>De-Hone Lin</i>	
Path Integral for Separable Hamiltonians of Liouville-Type	95
<i>Kazuo Fujikawa</i>	
Conjecture on the Reality of Spectra of Non-Hermitian Hamiltonians	105
<i>Carl M. Bender, Stefan Boettcher, and Peter N. Meisinger</i>	

Time-Transformation Approach to q-Deformed Objects	117
<i>Akira Inomata</i>	
Feynman Integral on a Group	129
<i>Zbigniew Haba</i>	
Characterizing Volume Forms	139
<i>Pierre Cartier, Marcus Berg, Cécile DeWitt-Morette, and Alexander Wurm</i>	
Vassiliev Invariants and Functional Integration	157
<i>Louis H. Kauffman</i>	
 Part II: Quantum Field Theory	
Higher Algebraic Geometrization Emerging from Noncommutativity	173
<i>Yuval Ne'eman</i>	
Dynamical Fermion Masses Under the Influence of Kaluza-Klein Fermions in Randall-Sundrum Background	185
<i>Hiroyuki Abe, Tomohiro Inagaki, and Taizo Muta</i>	
The Ultraviolet Fixed-Point in Quantum Electrodynamics – Adler Conjecture: Is QED Trivial?	197
<i>Raghunath Acharya</i>	
From Z Operator to $SO(10)$, Neutrino Oscillations, and Fermi-Dirac Functions for Quark Parton Distributions	201
<i>Franco Buccella</i>	
A Tilt at Constituent Quarks	207
<i>Carlos A. Savoy</i>	
The Breaking of Isospin and the ρ-ω System	219
<i>Harald Fritzsch</i>	
Analytic Confinement and Regge Spectrum	225
<i>Garij V. Efimov</i>	

Recursive Graphical Construction of Tadpole-Free Feynman Diagrams and Their Weights in ϕ^4-Theory	235
<i>Axel Pelster and Konstantin Glaum</i>	

Recursive Construction of Feynman Graphs in Spontaneously Broken $O(N)$-Symmetric ϕ^4-Theory	247
<i>Boris Kastening</i>	

Critical Behavior of Correlation Functions and Asymptotic Expansions of Feynman Amplitudes	259
<i>Adolfo P.C. Malbouisson</i>	

Gauge Symmetry and Neural Networks	271
<i>Tetsuo Matsui</i>	

Part III: Variational Perturbation Theory

Note on the Path-Integral Variational Approach in Many-Body Theory	283
<i>Jozef T. Devreese</i>	

Variational Perturbation Theory: A Powerful Method for Deriving Strong-Coupling Expansions	301
<i>Wolfhard Janke</i>	

Variational Perturbation Theory for the Ground-State Wave Function	315
<i>Axel Pelster and Florian Weißbach</i>	

Uniformly Suitable Estimation for Thermodynamic Values	327
<i>Ilya D. Feranchuk and Alexei Ivanov</i>	

Path-Integral and Perturbation Methods for Debye-Waller Factors Observed by Extended X-Ray-Absorption Fine Structure Spectroscopy	337
<i>Toshihiko Yokoyama</i>	

The Highest-Derivative Version of Variational Perturbation Theory	347

Bodo Hamprecht and Axel Pelster

Fast-Convergent Resummation Algorithm and Critical Exponents of ϕ^4 -Theory in Three Dimensions 365

Florian Jasch

Critical Exponent α of Superfluid Helium from Variational Strong-Coupling Theory 377

Bruno van den Bossche

Five-Loop Expansion of the ϕ^4 -Theory and Critical Exponents from Strong-Coupling Theory 387

Verena Schulte-Frohlinde

Part IV: Phase Transitions and Critical Phenomena

Nonasymptotic Critical Behavior from Field Theory 401

Claude Bagnuls and Claude Bervillier

Nonanalyticity of the Beta-Function and Systematic Errors in Field-Theoretic Calculations of Critical Quantities 413

Michele Caselle, Andrea Pelissetto, and Ettore Vicari

A Remark on the Numerical Validation of Triviality for Scalar Field Theories by High-Temperature Expansions 425

Paolo Butera and Marcattilio Comi

Multiloop ϕ^4 -Theory at Criticality 433

John A. Gracey

Phase Ordering Dynamics of ϕ^4 -Theory with Hamiltonian Equations of Motion 445

Bo Zheng, Volkard Linke, and Steffen Trimper

Phase Transition in the Random Anisotropy Model 457

Maksym Dudka, Rheinhard Folk, and Yuriy Holovatch

Gaps and Magnetization Plateaus in Low-Dimensional Quantum Spin

Systems	469
<i>Karl-Heinz Mütter</i>	
Effective Free Energy of Ginzburg-Landau Model	485
<i>Adriaan M.J. Schakel</i>	
Scaling and Duality in the Superconducting Phase Transition	497
<i>Flavio S. Nogueira</i>	
Thermal Fluctuations in the Gross-Neveu Model with $U(1)$-Symmetry at Small N	507
<i>Egor Babaev</i>	
FLEX-Theory for High-T_c Superconductivity Due to Spin Fluctuations	517
<i>Dirk Manske and Karl H. Bennemann</i>	
From Superfluid ^3He to Triplet Superconductor Sr_2RuO_4	533
<i>Kazumi Maki and Hyekyung Won</i>	
 Part V: Topological Defects, Strings, and Membranes	
Description of Vorticity by Grassmann Variables and an Extension to Supersymmetry	553
<i>Roman Jackiw</i>	
Non-Equilibrium Worldline Duality in Condensed Matter	565
<i>Ray J. Rivers</i>	
Field Theories and the Problem of Topological Entanglement in Polymer Physics	577
<i>Franco Ferrari</i>	
The Role of Topological Excitations at Second-Order Transitions	589
<i>Luís M.A. Bettencourt</i>	
Topological Singularities, Defect Formation, and Phase Transitions in Quantum Field Theory	601

Giuseppe Vitiello

Confinement in the Ensembles of Monopoles 613

Dmitri Antonov

Nambu-Goto String without Tachyons Between a Heavy and a Light Quark 625

Gaetano Lambiase and Vladimir V. Nesterenko

Deconfinement of Quarks in the Nambu-Goto String with Massive Ends 635

Vladimir V. Nesterenko and Gaetano Lambiase

Principles of Non-Local Field Theories and Their Application to Polymerized Membranes 645

Kay J. Wiese

Random Paths and Surfaces with Rigidity 663

Bergfinnur Durhuus

Part VI: Gravitation, Cosmology, and Astrophysics

String, Scalar Field, and Torsion Interactions 677

Richard T. Hammond

Metric and Connection: Kinematic and Dynamic Solutions of the Space Problem 685

Horst-Heino von Borzeszkowski and Hans-Jürgen Treder

Asymptotic Freedom in Curvature-Saturated Gravity 697

Salvatore Capozziello, Gaetano Lambiase, and Hans-Jürgen Schmidt

What Can Ising Spins Teach Us about Quantum Gravity? 707

Christian Holm

Gravitational Excitons – Fluctuating Particles from Extra Dimensions 721

Uwe Günther and Alexander Zhuk

Path Integrals in Quantum Cosmology	729
<i>Claus Kiefer</i>	
Varying Light Velocity as a Solution to the Problems in Cosmology	741
<i>John W. Moffat</i>	
Renormalization Group Method and Inhomogeneous Universe	759
<i>Yasuada Nambu</i>	
Analogies, New Paradigms, and Observational Data as Growing Factors of Relativistic Astrophysics	771
<i>Remo J. Ruffini</i>	
Index	799

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

Part I

Path Integrals and Quantum Mechanics

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

SEMICLASSICAL QUANTUM MECHANICS: A PATH-INTEGRAL APPROACH

B.R. HOLSTEIN

Department of Physics, University of Massachusetts, Amherst, MA 01003, USA

E-mail: holstein@physics.umass.edu

It is well known that one can approach problems in electrodynamics at either the macroscopic or microscopic level. However, we have less knowledge about the same approach in the case of semiclassical quantum mechanics. This case will be discussed here.

1 Introduction

When dealing with simple optical phenomena such as transmission and reflection, one can proceed in two quite different but totally equivalent ways. One is macroscopic and utilizes the solution of the Maxwell equations subject to appropriate boundary conditions. The second is microscopic. It is based on the multiple scattering series and follows a simple ray through all possible paths to its final destination. This procedure offers a very visual and a much clearer intuitive picture for what is taking place. The equivalence of macroscopic and microscopic techniques in this case is well known. Less familiar to many physicists is that an equivalent microscopic approach can be used in the case of semiclassical quantum mechanics. This will be discussed here. Our results are implicit in the work of other authors [1], so the presentation will be more of a didactic nature.

2 Semiclassical Methods

It is possible to treat quantum mechanics via macroscopic and/or microscopic methods, at least in the situation that the action divided by $\hbar$ is large – the

so-called semiclassical limit. The macroscopic procedure is well known as the WKB technique, the basic idea of which is that for a slowly varying potential $V(x)$, a stationary state solution can be written as [2]

$$\psi_{\text{WKB}}(x, t) \sim \sqrt{\frac{m}{k(x)}} \exp \left\{ i \left(\int_{x_0}^x k(x') dx' - Et \right) \right\}, \quad (1)$$

with the local wave number

$$k(x) = \sqrt{2m(E - V(x))}. \quad (2)$$

This approximation is valid when refraction dominates over reflection, or equivalently when the change in potential energy over a distance of the de Broglie wavelength is much smaller than the local kinetic energy, i.e.

$$\lambda \frac{|dV/dx|}{E - V(x)} \ll 1. \quad (3)$$

This approximation clearly breaks down at a classical turning point $x = a$, where $E = V(a)$. The usual procedure to deal with this breakdown is to use the WKB form of the wave function except in the immediate vicinity of the turning point, wherein a linear approximation to the potential is used in order to match the form of the wave function. This technique is a standard one and a prototypical problem treated via WKB is that of barrier penetration, i.e. the case of a particle of mass m and energy E incident on a potential barrier $V(x)$ with maximum height such that $V_{\text{max}} > E$. The results of this procedure are well known and yield reflection and transmission formulae [3]

$$R = 1 - T = \frac{e^{-2\sigma}}{(1 + \frac{1}{4}e^{-2\sigma})^2}, \quad (4)$$

with the WKB penetration factor

$$\sigma = \int_a^b dx \sqrt{2m(V(x) - E)}, \quad (5)$$

and a, b being the classical turning points. The derivation is straightforward, but using the connection formulae, derived from the asymptotic behavior of Airy functions, renders this procedure somewhat in the realm of black magic, so that one does not really develop a "feel" for the physics.

As an alternative approach, we present below a microscopic version of the semiclassical approximation which is based upon the Feynman path integral. The results obtained in this fashion are totally equivalent to those found via

WKB in most cases. However, the path integral viewpoint generates a more intuitive picture of the physics of the process under consideration and offers ways to treat some aspects, such as the superbarrier penetration [4], which lie outside the simple WKB domain of applicability.

2.1 Quadratic Approximation to the Propagator

In order to understand the semiclassical procedure, consider the path-integral form of the propagator

$$D_F(x_2, t; x_1, 0) = \int \mathcal{D}[x(t)] \exp \left\{ \frac{i}{\hbar} S[x(t), \dot{x}(t)] \right\}. \quad (6)$$

Now write

$$x(t) = x_{\text{cl}}(t) + \delta x(t), \quad (7)$$

where $x_{\text{cl}}(t)$ is a solution of Hamilton's equation $\delta S[x_{\text{cl}}(t)] = 0$, and expand the action in terms of $\delta x(t)$. The terms linear in $\delta x, \delta \dot{x}$ vanish after an integration by parts and use of the classical equations of motion, yielding

$$D_F(x_2, t; x_1, 0) \simeq \exp \left\{ \frac{i}{\hbar} S[x_{\text{cl}}(t)] \right\} \int \mathcal{D}[\delta x(t)] \times \exp \left\{ \frac{i}{\hbar} \int_0^t dt' \left(\frac{1}{2} m (\delta \dot{x}(t'))^2 - \frac{1}{2} V''(x_{\text{cl}}(t')) (\delta x(t'))^2 \right) \right\} \quad (8)$$

as an approximate representation of the propagator. The exponential phase factor is just the classical action, while the additional multiplicative term is the closed path propagator for a particle of mass m moving under the influence of the time-dependent potential $x^2 V''(x_{\text{cl}}(t))/2$.

Eq. (8) is exact for a quadratic potential, but for a general potential the semiclassical approximation involves dropping terms of higher order than quadratic in δx , since one can show that the contribution of the quadratic order terms to the path integral is of lower order in $\hbar$ than that from cubic or higher order terms. Indeed, only the classical path survives in the limit $\hbar \rightarrow 0$, and deviations $\delta x(t)$ from this path are only important if the change which they induce in $S/\hbar$ is of order $\sim \pi$ or less, i.e. $\delta^2 S[x_{\text{cl}}]/\hbar \leq \pi$, since larger deviations imply a rapid oscillation of the integrand and result in a nearly complete cancellation. We conclude that $\delta x \sim \hbar^{1/2}$ so that $S[x_{\text{cl}}]/\hbar \sim \mathcal{O}(\hbar^{-1})$ and $\delta^2 S[x_{\text{cl}}]/\hbar \sim \mathcal{O}(1)$. However, $\delta^3 S[x_{\text{cl}}]/\hbar \sim \mathcal{O}(\hbar^{1/2})$, implying

that cubic and higher order terms in δx can be neglected in the semiclassical limit $\hbar \rightarrow 0$.

The quadratic path integration over $\delta x(t)$ may be performed via a clever analyticity argument due to Coleman [5], or by other means, yielding the so-called Van Vleck determinant (hereafter setting $\hbar = 1$)

$$\begin{aligned} & \int \mathcal{D}[\delta x(t)] \exp \left\{ i \int_0^t dt' \left(\frac{1}{2} m (\delta \dot{x}(t'))^2 - \frac{1}{2} (\delta x(t'))^2 V''(x_{\text{cl}}(t')) \right) \right\} \\ &= \left(\frac{m}{2\pi i \dot{x}_{\text{cl}}(0) \dot{x}_{\text{cl}}(t)} \int_{x_1}^{x_2} dx \dot{x}_{\text{cl}}^{-3}(x) \right)^{\frac{1}{2}}. \end{aligned} \quad (9)$$

2.2 The WKB Propagator

We can now demonstrate that this quadratic approximation to the path integral is, in general, completely equivalent to the WKB approximation [6]. In order to make this connection, consider the propagator written in terms of WKB solutions of Eq. (1):

$$D(x_2, t; x_1, 0) = \frac{m}{2\pi} \int dE e^{-iEt} \psi_{\text{WKB}}(x_2, E) \psi_{\text{WKB}}^*(x_1, E), \quad x_2 > x_1, \quad (10)$$

where

$$\psi_{\text{WKB}}(x, E) = \frac{1}{\sqrt{k(x)}} \exp \left[i \int_a^x dx' k(x') \right], \quad (11)$$

with the local wave number (2). The energy integration in Eq. (10) can be performed via the stationary phase approximation, wherein one approximates an integral of the form

$$J = \int_{-\infty}^{\infty} dE e^{if(E)} g(E) \simeq \sqrt{\frac{2\pi i}{f''(E_0)}} g(E_0) e^{if(E_0)}, \quad (12)$$

with E_0 defined by the condition $f'(E_0) = 0$. In our case

$$f(E) = \int_{x_1}^{x_2} dx' \sqrt{2m(E - V(x'))} - Et, \quad (13)$$

and the stationary phase energy E_0 is found by requiring

$$0 = \frac{\partial f(E)}{\partial E} = \int_{x_1}^{x_2} dx' \sqrt{\frac{m}{2(E_0 - V(x'))}} - t. \quad (14)$$

Since $\sqrt{2(E - V(x))/m} = \dot{x}_{\text{cl}}$ is the classical velocity, we find $t = \int_{x_1}^{x_2} dx' / \dot{x}_{\text{cl}}(x')$ as the defining equation for E_0 , i.e. $E_0 = E_{\text{cl}}$, where E_{cl} is the classical energy for the trajectory $x_{\text{cl}}(t)$ connecting $(x_1, 0)$ and (x_2, t) . Also, since

$$f''(E_{\text{cl}}) = -\frac{1}{2} \int_{x_1}^{x_2} dx' \left(\frac{m}{2(E_{\text{cl}} - V(x'))^3} \right)^{\frac{1}{2}} = -\frac{1}{m} \int_{x_1}^{x_2} dx \dot{x}_{\text{cl}}^{-3}(x) \quad , \quad (15)$$

the WKB propagator becomes

$$D(x_2, t; x_1, 0) = \left(\frac{m}{2\pi i \dot{x}_{\text{cl}}(t) \dot{x}_{\text{cl}}(0) \int_{x_1}^{x_2} dx \dot{x}_{\text{cl}}^{-3}(x)} \right)^{1/2} \times \exp \left(i \int_{x_1}^{x_2} dx k(x) - i E_{\text{cl}} t \right) \quad , \quad (16)$$

which is identical in form to the quadratic approximation to the path integral, since the prefactor is obviously the same, while the classical action can be written as

$$\begin{aligned} S[x_{\text{cl}}(t)] &= \int_0^t dt' \left(\frac{1}{2} m \dot{x}_{\text{cl}}^2(t') - V(x_{\text{cl}}(t')) \right) = \int_0^t dt' (m \dot{x}_{\text{cl}}^2(t') - E_{\text{cl}}) \\ &= \int_{x_1}^{x_2} dx m \dot{x}_{\text{cl}}(x) - E_{\text{cl}} t = \int_{x_1}^{x_2} dx \sqrt{2m(E_{\text{cl}} - V(x))} - E_{\text{cl}} t. \end{aligned} \quad (17)$$

We conclude that any problem which can be treated via WKB methods can equally well be handled using path-integral techniques in the quadratic approximation.

2.3 Barrier Penetration: Semiclassical Approach

An example of such a problem is that of barrier penetration. We shall demonstrate here how the same results found via WKB can be derived via semiclassical path-integral methods. In order to do so, it is necessary to generalize the idea of a classical path. In the case of barrier penetration a conventional classical trajectory connecting points x_1, x_2 far to the left or right of the barrier, respectively, and with classical energy *less* than the height of the barrier, does not exist since the particle must be completely reflected. Nevertheless, there *do* exist trajectories connecting the two space-time points $(x_1, 0)$ and

(x_2, t) which satisfy

$$m\ddot{x}_{\text{cl}} = -\left.\frac{dV}{dx}\right|_{x_{\text{cl}}}, \quad (18)$$

provided we allow the time t to become complex, and such paths can be chosen to have a real classical energy $E < V_{\text{max}}$. In fact, for given x_1, x_2, E , there exists a denumerably *infinite* set of such paths which can be labeled by their propagation "times" with $t^{(n)}$, $n = 0, 1, 2, \dots$:

$$t^{(n)} = \int_{x_1}^a + \int_b^{x_2} dx \sqrt{\frac{m}{2(E - V(x))}} - i(2n+1) \int_a^b dx \sqrt{\frac{m}{2(V(x) - E)}}. \quad (19)$$

The propagation from x_1 to x_2 can be considered as occurring in three successive steps. First the particle travels from the initial point x_1 to the left-hand classical turning point a in the real time interval

$$\Delta t_a = \int_{x_1}^a dx \sqrt{\frac{m}{2(E - V(x))}}, \quad (20)$$

and then makes $2n + 1$ traversals of the interval between a and b in pure *imaginary* time

$$\Delta t_i^{(n)} = -i(2n+1) \int_a^b dx \sqrt{\frac{m}{2(V(x) - E)}}. \quad (21)$$

Finally, the particle propagates from the right-hand turning point b to x_2 in the real time interval

$$\Delta t_b = \int_b^{x_2} dx \sqrt{\frac{m}{2(E - V(x))}}. \quad (22)$$

The total temporal interval is then $t^{(n)} = \Delta t_a + \Delta t_i^{(n)} + \Delta t_b$, as given in Eq. (19).

Physically, the different values $t^{(n)}$ ($n = 0, 1, 2, \dots$) correspond to the possibility of internal reflections from the walls inside the barrier. This interpretation follows from writing Newton's equations in terms of imaginary time $\tau = it$

$$m \frac{d^2 x}{d\tau^2} = -m \frac{d^2 x}{dt^2} = \left.\frac{dV(x)}{dx}\right|_{x=x(\tau)}. \quad (23)$$

Equation (23) describes the motion of a classical particle of mass m moving inside a potential well $-V(x)$. The particle can travel from turning point a to turning point b in an infinite number of ways, corresponding to the direct path plus an *arbitrary* number ($n = 1, 2, \dots$) of complete loops from a to b and back again to a . Thus the "trajectory" labeled by $t^{(0)}$ represents a path which enters the forbidden region at point a , "propagates" directly to turning point b and exits. The path labeled by $t^{(1)}$ passes into the barrier at point a , "propagates" to b where it is reflected, returns to a where it is reflected again, and finally "propagates" to point b where it exits the barrier. Similarly, paths labeled by $t^{(n)}$ correspond to "trajectories" with $2n + 1$ traversals of the forbidden region before final exit at point b .

Obviously these "classical" paths cannot be directly utilized to produce a semiclassical propagator. They become relevant, however, if the propagator is analytically continued through the introduction of its Fourier transform^a

$$D(x_2, x_1; E) = \int_0^\infty dt e^{iEt} D(x_2, t; x_1, 0) . \quad (24)$$

In fact, it is this energy-space form of the propagator which is relevant for calculating the barrier reflection/transmission properties, since wavepackets are constructed to have fixed energy.

The form of $D(x_2, t; x_1, 0)$ to be used here is given by

$$D(x_2, t; x_1, 0) \cong \left(\frac{1}{2\pi i k(x_2)k(x_1) \int_{x_1}^{x_2} dx k^{-3}(x)} \right)^{1/2} \times \exp \left(i \int_{x_1}^{x_2} k(x) dx - i\tilde{E}t \right) . \quad (25)$$

The integral $\int_{x_1}^{x_2} k(x) dx$ is to be understood as a line integral that can loop around the edges of the barrier $2n$ times ($n = 0, 1, 2, \dots$). In the immediate vicinity of a turning point, the quadratic approximation to the propagator is no longer valid, and an additional factor $\lambda = i/2$ is produced each time one passes through such a point, as shown by careful study of the turning point region [7]. Thus for the "classical" path labeled by integer n , we interpret

$$\int_{x_1}^{x_2} dx k(x) = \int_{x_1}^a + \int_b^{x_2} dx k(x) + i(2n + 1) \int_a^b \kappa(x) dx - i2n \ln \lambda, \quad (26)$$

^aThe lower limit of integration is set by the feature that the Feynman propagator $D(x_2, t; x_1, 0)$ is defined to vanish for $t < 0$ by the $i\epsilon$ prescription.

where $\kappa(x) = \sqrt{2m(V(x) - E)}$ and where the energy $\tilde{E}(t^{(n)})$ is defined implicitly in terms of the time $t^{(n)}$ via

$$t^{(n)} = \int_{x_1}^a + \int_b^{x_2} dx \sqrt{\frac{m}{2(\tilde{E} - V(x))}} + i(2n+1) \int_a^b dx \sqrt{\frac{m}{2(V(x) - \tilde{E})}}. \quad (27)$$

Here the turning points a and b are functions of $\tilde{E}(x_2, x_1, t^{(n)})$, but λ is independent of energy (and time).

The Fourier transform is performed via the stationary phase approximation. Writing $D(x_2, t; x_1, 0) = \rho(t) e^{i\phi(t)}$ (where we have suppressed the dependence upon x_2, x_1) we have

$$D(x_2, x_1; E) = \int_0^\infty dt e^{i(Et + \phi(t))} \rho(t), \quad (28)$$

and the stationary phase point $\bar{t}$ is determined via

$$\begin{aligned} 0 &= \frac{\partial}{\partial t} (Et + \phi(t)) = \frac{\partial}{\partial t} \left(Et + \int_{x_1}^{x_2} k(x) dx - \tilde{E}t \right) \\ &= E - \tilde{E}(t) - t \frac{\partial \tilde{E}}{\partial t} + \int_{x_1}^{x_2} dx \frac{\partial k}{\partial \tilde{E}} \frac{\partial \tilde{E}}{\partial t}. \end{aligned} \quad (29)$$

However, $t - \int_{x_1}^{x_2} dx \partial k / \partial \tilde{E} = 0$, since this is the equation which defined $\tilde{E}(t)$ in the first place. Thus the stationary phase point $\bar{t}$ is defined via $E = \tilde{E}(\bar{t})$. Finally, using

$$\left. \frac{\partial^2}{\partial t^2} (Et + \phi(t)) \right|_{t=\bar{t}} = - \left. \frac{\partial \tilde{E}}{\partial t} \right|_{t=\bar{t}} = - \frac{1}{\int_{x_1}^{x_2} dx \frac{\partial^2 k}{\partial \tilde{E}^2}} = \left(m^2 \int_{x_1}^{x_2} dx k^{-3}(x) \right)^{-1}, \quad (30)$$

the stationary phase approximation yields

$$D(x_2, x_1; E) = \left(\frac{m^2}{k(x_2)k(x_1)} \right)^{1/2} \exp \left\{ i \int_{x_1}^{x_2} k(x) dx \right\}, \quad (31)$$

where the integration is understood in the sense of Eq. (26).

As the $\bar{t}^{(n)}$ are complex, the stationary phase procedure requires that the path from $t = 0$ to $t = +\infty$ is deformed into the complex plane in such a way that it passes through *each* $\bar{t}^{(n)}$ with $n = 0, 1, 2, \dots$. We then find

$$D(x_2, x_1; E) = \sum_{n=0}^{\infty} D^{(n)}(x_2, x_1; E)$$

$$\begin{aligned}
&= \left(\frac{m^2}{k(x_2)k(x_1)} \right)^{1/2} \exp \left(i \int_{x_1}^a + i \int_b^{x_2} k(x) dx \right) \\
&\quad \times \sum_{n=0}^{\infty} \lambda^{2n} \exp \left(-(2n+1) \int_a^b \kappa(x) dx \right) . \quad (32)
\end{aligned}$$

The contribution from paths involving $n \geq 1$ are exponentially suppressed and of questionable accuracy, as the corrections to the semiclassical approximation to the path integral could well be larger. Nevertheless, it is important to include the effects of these “interior bounce” solutions since only then we have a unitary and fully consistent picture of the transmission and reflection process which conserves probability. Thus calculating the propagator for transmission, as given above, we can perform the summation over n , yielding

$$D(x_2, x_1; E) = \left(\frac{m^2}{k(x_2)k(x_1)} \right)^{1/2} \exp \left(i \int_{x_1}^a + i \int_b^{x_2} k(x) dx \right) \frac{e^{-\sigma}}{1 - \lambda^2 e^{-2\sigma}}, \quad (33)$$

where $\sigma \equiv \int_a^b \kappa(x) dx$.

We see that this path-integral approach to the semiclassical approximation allows a very appealing and graphical picture of the transmission process. Instead of the sum over *all possible* paths required in calculating the complete path integral, the semiclassical approximation utilizes a sum over *all “classical”* paths connecting the initial and final points. (Here “classical” is used in the sense defined above, wherein analytic continuation into the complex time plane is permitted.) The form of the propagator can be found by using the simple rules:^b

- i) Propagation from x_1 to x_2 in a classically allowed, forbidden region produces a factor

$$\exp \left\{ i \int_{x_1}^{x_2} k(x) dx \right\} , \quad \exp \left\{ - \int_{x_1}^{x_2} \kappa(x) dx \right\} , \quad (34)$$

respectively;

- ii) Reflection from a classical turning point within a classically forbidden region yields a factor $\lambda = i/2$, as discussed in Ref. [7].

^bFor completeness, it should be noted that there exist corresponding phases $\exp(\pm i\pi/4)$ which arise upon *transmission* through a barrier. However, these do not play a role when the transmission or reflection probabilities are determined and therefore, for simplicity, will be omitted here.

In order to deal with the corresponding reflection coefficient, we require one additional rule

- iii) Reflection from a classical turning point within a classically allowed region yields a factor $\eta = -i$, according to Ref. [7].

The propagator for the reflection process may then be constructed by taking x_1, x_2 both to the left of the barrier and including all possible "classical" trajectories:

- i) Propagation from x_1 to the left-hand wall at $x = a$ followed by a reflection and propagation back to x_2 ;
- ii) Propagation from x_1 to the left-hand wall at $x = a$ followed by transmission into the barrier, reflection from the right-hand wall, transmission back to $x = a$ and then propagation from $x = a$ to x_2 , etc.

The total contribution to the propagator is then

$$D(x_2, x_1; E) = \left(\frac{m^2}{k(x_2)k(x_1)} \right)^{1/2} \exp \left(i \int_{x_1}^a + i \int_a^{x_2} k(x) dx \right) f \quad (35)$$

with

$$f = \eta + \sum_{n=0}^{\infty} \lambda^{2n+1} e^{-(2n+2)\sigma} = \eta + \frac{\lambda e^{-2\sigma}}{1 - \lambda^2 e^{-2\sigma}}, \quad (36)$$

where σ is defined in Eq. (5). The connection with the transmission and reflection coefficients

$$R = |r(E)|^2 \quad \text{and} \quad T = |t(E)|^2 \quad (37)$$

can now be made via the identifications

$$\begin{aligned} D(x_2, x_1; E) &= \left(\frac{m^2}{k(x_2)k(x_1)} \right)^{1/2} t(E), & x_1 \ll a, \quad b \ll x_2, \\ D(x_2, x_1; E) &= \left(\frac{m^2}{k(x_2)k(x_1)} \right)^{1/2} r(E), & x_1, x_2 \ll a, \end{aligned} \quad (38)$$

and we find, using Eqs. (32) and (35),

$$T = |t(E)|^2 = \frac{e^{-2\sigma}}{|1 - \lambda^2 e^{-2\sigma}|^2} = \frac{e^{-2\sigma}}{(1 + \frac{1}{4}e^{-2\sigma})^2},$$

$$R = |r(E)|^2 = \left| \eta + \frac{\lambda e^{-2\sigma}}{1 - \lambda^2 e^{-2\sigma}} \right|^2 = \left(\frac{1 - \frac{1}{4}e^{-2\sigma}}{1 + \frac{1}{4}e^{-2\sigma}} \right)^2 \quad (39)$$

in complete agreement with the corresponding WKB expressions, Eq. (5).

3 Conclusions

We have seen how the problem of barrier penetration can be handled equally well via macroscopic (WKB) or microscopic (semiclassical path integral) techniques. Similarly, it is straightforward to understand how other classic problems usually treated via WKB, such as α decay [8], can be derived microscopically. The combination of macroscopic and microscopic approaches can provide an improved understanding of the problem and its solution.

Acknowledgments

This work was supported in part by the National Science Foundation.

References

- [1] R. Dashen, B. Hasslacher, and A. Neveu, *Phys. Rev. D* **10**, 4114, 4130 (1974); R. Rajaraman, *Phys. Rep. C* **21**, 227 (1975); H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [2] G. Wentzel, *Z. Phys.* **39**, 518 (1926); H.A. Kramers, *Z. Phys.* **39**, 828 (1926); L. Brillouin, *J. Phys.* **7**, 353 (1926).
- [3] See, e.g. E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1998).
- [4] B.R. Holstein, *Am. J. Phys.* **52**, 321 (1984).
- [5] S. Coleman, *Aspects of Symmetry* (Cambridge Univ. Press, New York, 1985).
- [6] B.R. Holstein and A.R. Swift, *Am. J. Phys.* **50**, 829 (1982); R.D. Carlitz and D.A. Nicole, *Ann. Phys. (N.Y.)* **164**, 411 (1985).
- [7] See, e.g. D.W. McLaughlin, *J. Math. Phys.* **13**, 1099 (1972); M.V. Berry and K.E. Mount, *Rep. Prog. Phys.* **35**, 315 (1972); J.B. Keller, *Ann. Phys. (N.Y.)*, **4**, 180 (1958).
- [8] B.R. Holstein, *Am. J. Phys.* **64**, 1061 (1996).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

FORWARD-BACKWARD SEMICLASSICAL DYNAMICS

N. MAKRI

*Departments of Chemistry and Physics, University of Illinois,
601 S. Goodwin Avenue, Urbana, Illinois 61801, USA
E-mail: nancy@makri.scs.uiuc.edu*

Recent developments in semiclassical methodology for time-dependent properties of multi-dimensional systems are reviewed. Combining the time evolution operator and its adjoint into a single forward-backward propagator provides a natural smoothing of the semiclassical integrand, making the latter suitable for integration by Monte Carlo methods.

1 Introduction

Computing the time evolution of quantum systems composed of many coupled degrees of freedom presents serious challenges. Currently, simulations of time-dependent properties in polyatomic molecules, clusters, or condensed phases employ a variety of approximations, which are largely built around variational, quantum classical or semiclassical ideas. While approximate methods have proven extremely valuable in many situations, their regime of applicability is limited. Perhaps more importantly, checking the accuracy of the results is usually impractical, and thus reliability becomes a significant concern.

Numerical solution of the Schrödinger equation requires storage of the multi-dimensional wave function, which occupies a volume that grows exponentially with the number of particles. This limitation is circumvented in the path-integral formulation, where quantum mechanical amplitudes are expressed as sums over paths [1,2]. To reproduce quantum interference, the amplitude along each path must be a complex-valued phase. Because of the rapid oscillations of the latter, numerical evaluation of the real-time path

integral by means of stochastic methods is plagued by severe numerical instabilities [3,4]. At the same time, explicit summation over all paths is far from feasible, as their number (if space and time are discretized) increases exponentially with the number of degrees of freedom and the total propagation time. Recently, exact evaluation of the path integral by virtue of an iterative algorithm became feasible in cases of a low-dimensional system coupled to a dissipative bath (for a review see Ref. [5]).

2 Semiclassical Dynamics with Forward-Backward Trajectories

The conventional time-dependent semiclassical approximation can be derived from the path-integral expression for the coordinate propagator [1,2],

$$\langle \mathbf{x}_2 | U(t_2, t_1) | \mathbf{x}_1 \rangle = \int \mathcal{D}\mathbf{x}(t) e^{iS[\mathbf{x}(t)]/\hbar}, \quad (1)$$

where U is the time evolution operator that propagates from time t_1 to t_2 . The last equation is a functional integral over all paths $\mathbf{x}(t)$ with endpoints $\mathbf{x}(t_1) = \mathbf{x}_1$ and $\mathbf{x}(t_2) = \mathbf{x}_2$, and the amplitude contributed by each path is the classical action in units of Planck's constant. Taking the limit $\hbar \rightarrow 0$ and evaluating the functional integral by the stationary phase method (i.e. restricting the sum in Eq. (1) to paths that satisfy the Euler-Lagrange equations as well as quadratic fluctuations around them) leads to the Van Vleck expression for the semiclassical propagator [6,7]:

$$\langle \mathbf{x}_2 | U(t_2, t_1) | \mathbf{x}_1 \rangle_{\text{SC}} = \sum_{\text{classical paths } \mathbf{x}(t)} D_{VV}(\mathbf{x}_2, \mathbf{x}_1) e^{iS[\mathbf{x}(t)]/\hbar} e^{-i\mu\pi/2}. \quad (2)$$

Here the prefactor D_{VV} is a determinant given by one of the elements of the stability matrix, and μ is the Maslov index which supplies the proper phase by keeping track of focusing characteristics of the classical trajectories. The sum in Eq. (2) arises because there are in general multiple solutions to the boundary value problem specified by the endpoint constraints imposed on the classical paths. An excellent presentation of the time-dependent semiclassical approximation and its relation to the path integral can be found in the books of Kleinert [8] and of Schulman [9].

Equation (2) is formulated in terms of the classical solution(s) to a double-ended boundary value problem and thus is impractical. In addition, the Van Vleck prefactor diverges at caustics, and its rapidly growing rate in the vicinity of such points gives rise to numerical instabilities. Miller has shown that

both of these problems can be overcome by switching to an initial value representation [10], provided that one is interested in an observable rather than the propagator itself. This is achieved by changing the integration variable associated with the final position of a trajectory to that specifying its initial momentum and absorbing the Van Vleck determinant and the Jacobian into a single prefactor. The most favorable representation is obtained in terms of coherent states, which introduce a natural weight factor for sampling the initial conditions of classical trajectories. Inserting overcomplete sets of coherent states and applying repeatedly the stationary phase approximation, Herman and Kluk showed [11,12] that the semiclassical propagator can also be brought into the form

$$\langle \mathbf{x}_2 | U(t_2, t_1) | \mathbf{x}_1 \rangle_{\text{SC}} = \int d\mathbf{x}_0 \int d\mathbf{p}_0 \langle \mathbf{x}_2 | g(\mathbf{x}_f, \mathbf{p}_f) \rangle D_{HK}(\mathbf{x}_0, \mathbf{p}_0) \times e^{iS(\mathbf{x}_0, \mathbf{p}_0)/\hbar} \langle g(\mathbf{x}_0, \mathbf{p}_0) | \mathbf{x}_1 \rangle. \quad (3)$$

Here $\mathbf{x}_f, \mathbf{p}_f$ stand for the final coordinate and momentum of a classical trajectory with initial conditions $\mathbf{x}_0$ and $\mathbf{p}_0$, and the coherent state wave functions are defined by the usual relation

$$\langle \mathbf{x} | g(\mathbf{x}_0, \mathbf{p}_0) \rangle = \sqrt{\det \frac{2\gamma}{\pi}} \exp \left[-(\mathbf{x} - \mathbf{x}_0) \cdot \gamma \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{i}{\hbar} \mathbf{p}_0 \cdot (\mathbf{x} - \mathbf{x}_0) \right]. \quad (4)$$

The semiclassical propagator is capable of reproducing semiquantitatively all aspects of quantum mechanics [10,13,14]. In spite of these appealing features, rigorous semiclassical theory has found little practical utility in the past. As in the case of the real time path integral, evaluation of ensemble-averaged observables by means of the semiclassical propagator is plagued by the problem of severe phase cancellation, which now occurs due to interfering classical trajectories. Persistent efforts in the last few years have led to some methodological advances, which have enabled calculations in several models and small molecular systems [11,12,15-34].

Systems of many atoms are usually characterized in terms of ensemble-averaged quantities, such as correlation functions or the reduced density matrix. In these, the dynamical quantity of interest (usually an operator that depends on a few degrees of freedom of direct interest, the "system"), propagated to the desired time, is traced over all unprobed coordinates (those of the "bath") subject to a distribution function that characterizes the particular

ensemble. Such expressions have the general structure

$$\text{Tr} (\rho_0 \cdots U_{\text{back}}^{-1}(t, 0) \cdots U_{\text{for}}(t, 0)) , \quad (5)$$

where ρ_0 is the equilibrium density matrix and U_{for} , U_{back} are time evolution operators corresponding to the dynamics generated by the same or two distinct Hamiltonians, that may also be explicitly time-dependent. This structure can be exploited to reduce the severity of the sign problem.

The conventional semiclassical procedure for evaluating Eq. (5) is to approximate each of the two propagators by the semiclassical (Van Vleck or coherent state) expression and attempt to perform the resulting multi-dimensional integrals by Monte Carlo methods. In the coordinate representation, this procedure produces an expression of the form

$$\int d\mathbf{x}_0 \int d\mathbf{x}_f \int d\mathbf{x}_t e^{-iS_{\text{back}}(\mathbf{x}_t, \mathbf{x}_f)/\hbar} \cdots e^{iS_{\text{for}}(\mathbf{x}_0, \mathbf{x}_t)/\hbar} \cdots . \quad (6)$$

According to this, one integrates trajectories from $(\mathbf{x}_0, 0)$ to $(\mathbf{x}_t, t)$ and also from $(\mathbf{x}_t, t)$ to $(\mathbf{x}_f, 0)$. With the exception of very short times, each of the action integrals in the above equation is large relative to $\hbar$ and thus the integrand is highly oscillatory.

Imagine, however, that one finds a way to combine the two time evolution operators in Eq. (5) into a single propagator. Various ways of doing this are discussed in the subsections that follow, but notice that the presence of the operator B in the correlation function prevents the combined exponentials from producing the identity (and thus eliminating all dynamics). Applying the semiclassical approximation to the new forward-backward propagator would lead to an expression with a *single* semiclassical amplitude [25]:

$$\int d\mathbf{x}_0 \int d\mathbf{x}_f \cdots e^{iS(\mathbf{x}_0, \mathbf{x}_f)/\hbar} \cdots . \quad (7)$$

Here the classical trajectories run from $(\mathbf{x}_0, 0)$ to $(\mathbf{x}_t, t)$ and subsequently backwards in time to $(\mathbf{x}_f, 0)$, and S is the combined forward-backward action.

The pairing of the two evolution operators has accomplished two tasks. First, the midpoint integral has been eliminated. In fact, Eq. (7) can be viewed as the result of evaluating the midpoint integral of Eq. (6) by the stationary phase method. Second, and more important, the action entering Eq. (7) is the *net* result of forward and backward propagation. If the trajectory simply reversed its steps during the backward part, the overall action would be equal to zero. As will become clear in the following, the presence of

an operator between the two propagators alters somewhat the course of the backward trajectory; nevertheless, the net forward-backward action generally remains small. The significance of this fact is obvious: the rapid oscillations of the integrand have been eliminated and Monte Carlo methods should be applicable. Phrased differently, the advantage of forward-backward semiclassical dynamics (FBS) is that the phase cancellation which plagues the multi-dimensional integrals of Eq. (6) is shifted naturally to a cancellation between forward and backward actions.

The smoothing achieved by combining the forward and backward propagation steps are illustrated in Figs. 1 and 2 which show the semiclassical integrand as a function of the trajectory initial conditions for a harmonic potential after (a) propagation in the forward direction only (Fig. 1) and (b) combined forward-backward propagation (Fig. 2). In multi-dimensional space, the oscillatory character of the integrand depicted in Fig. 1 is responsible for the failure of Monte Carlo methods. On the other hand, the smooth forward-backward integrand of Fig. 2 is perfectly suitable for stochastic sampling.

The numerical advantages achieved by combining the forward and backward evolution operators into a single semiclassical step are not without consequences. Consider two classical trajectories satisfying the boundary conditions of the forward and backward propagators in Eq. (6). As no constraint is imposed on these trajectories, they may join with different slopes (and thus different momenta) at the midpoint $\mathbf{x}_t$. Now imagine evaluating the midpoint integral by the stationary phase method in order to arrive at the FBS expression, Eq. (7). The stationary phase condition is [26]

$$\left. \frac{\partial S_{\text{for}}(\mathbf{x}_0, \mathbf{x}_t)}{\partial \mathbf{x}_t} \right|_{\mathbf{x}_0} - \left. \frac{\partial S_{\text{back}}(\mathbf{x}_t, \mathbf{x}_f)}{\partial \mathbf{x}_t} \right|_{\mathbf{x}_f} = \mathbf{p}_t^{\text{for}} - \mathbf{p}_t^{\text{back}} = 0, \quad (8)$$

where $\mathbf{p}_t^{\text{for}}$ is the final momentum of the forward trajectory and $\mathbf{p}_t^{\text{back}}$ is the initial momentum of the backward trajectory. Thus, the forward and backward trajectories in FBS are no longer independent but are constrained to maintain momentum continuity at all times. This restriction can lead to loss of interference.

It is thus apparent that the FBS approach should be implemented with caution. Depending on the nature of the problem at hand and the desired degree of accuracy it may be advantageous to treat all of just some of the degrees of freedom by FBS. Various situations that emerge are described

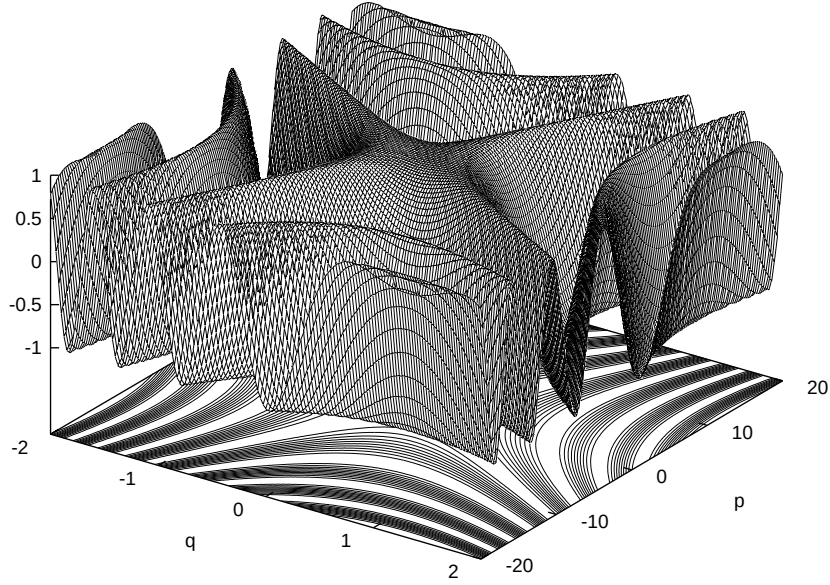


Figure 1. The real part of the semiclassical propagator as a function of initial position and momentum for a one-dimensional harmonic oscillator driven by a time-dependent force arising from its coupling to a quantum system for the action arising from a path in the conventional forward time direction. The total propagation time is $\omega t = 1$.

below.

Perhaps the most straightforward and unambiguous use of FBSD is in the evaluation of influence functionals [35] in the context of the path-integral representation of time-dependent quantum mechanics. Consider, for example, the reduced density matrix for a system described by the Hamiltonian $H_0(r, p_r)$ in contact with an n -dimensional environment of coordinates $\mathbf{x}, \mathbf{p}$ whose force field and interaction with the solute are given by the Hamiltonian $H_b(\mathbf{x}, \mathbf{p}, r)$. Using the path-integral representation for each of the two propagators, the reduced density matrix can be brought in the form

$$\langle r'' | \rho(t) | r' \rangle = \int dr_0^+ \int dr_0^- \int \mathcal{D}r^+ \int \mathcal{D}r^- e^{iS_0[r^+]/\hbar} e^{-iS_0[r^-]/\hbar} F[r^+, r^-], \quad (9)$$

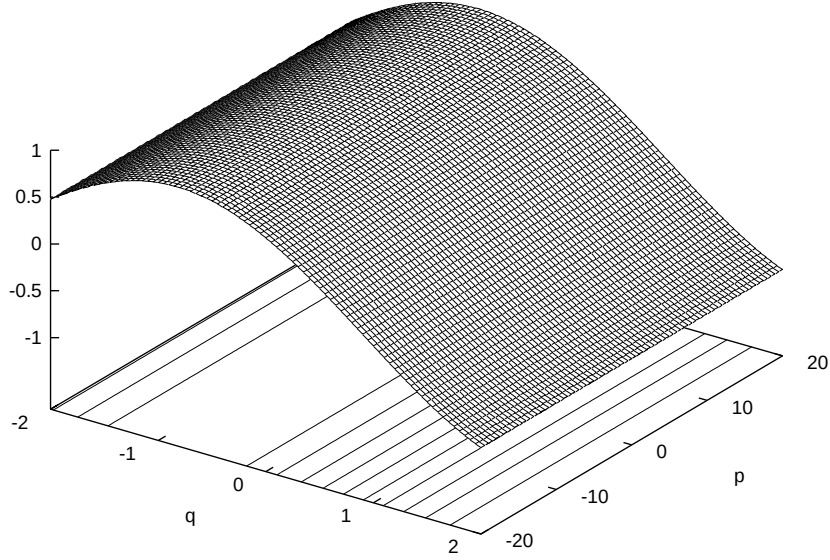


Figure 2. Same as Fig. 1, but for combined action along the forward-backward time contour for an arbitrary choice of distinct forward and backward paths.

where

$$F[r^+, r^-] = \text{Tr}_b (U_{\text{for}}(t, 0) \rho_0 U_{\text{back}}^{-1}(t, 0)) \quad (10)$$

is the influence functional that contains all the effects of the environment on the dynamics of the particle of interest. Here $U_{\text{for}}(t, 0)$ and $U_{\text{back}}(t, 0)$ are operators that propagate the bath in the presence of a potential that is time-dependent by virtue of the time-parameterization of the system coordinate:

$$H_{\text{for}}(t') = H_b(\mathbf{x}, \mathbf{p}, r^+(t')), \quad H_{\text{back}}(t') = H_b(\mathbf{x}, \mathbf{p}, r^-(t')). \quad (11)$$

Noting that the trace operation allows a cyclic rearrangement of the two arguments brings the influence functional in a form suitable for FBSD. Its coherent state representation becomes [25]

$$F[r^+, r^-] = \int d\mathbf{x}_0 \int d\mathbf{p}_0 D_{HK}(\mathbf{x}_0, \mathbf{p}_0) e^{iS(\mathbf{x}_0, \mathbf{p}_0)/\hbar}$$

$$\times \langle g(\mathbf{x}_0, \mathbf{p}_0) | \rho_0 | g(\mathbf{x}_f, \mathbf{p}_f) \rangle. \quad (12)$$

Consider now the evaluation of general correlation functions of the type

$$C(t) = \text{Tr} \left(\rho_0 A e^{iHt/\hbar} B e^{-iHt/\hbar} \right), \quad (13)$$

where A and B are general operators. In this case, the forward-backward structure of the integrand is disrupted by the presence of these operators and formulating a semiclassical forward-backward scheme is neither straightforward nor unique.

The most straightforward treatments involve bringing the operator B into an exponential form and applying FBSD to the resulting evolution operator. Miller and coworkers have proposed the use of the Weyl identity,

$$B_W(\mathbf{x}, \mathbf{p}) = \int d\mathbf{p}' \int d\mathbf{x}' W_{\rho_0 B}(\mathbf{x}', \mathbf{p}') e^{i\mathbf{p}' \cdot \mathbf{x}/\hbar} e^{-i\mathbf{p} \cdot \mathbf{x}'/\hbar}, \quad (14)$$

where $W_{\rho_0 B}$ is the Wigner transform [36] of the operator $\rho_0 B$. The correlation function is brought in the form [29,30]

$$C(t) = \int d\mathbf{x}_0 \int d\mathbf{p}_0 \int d\mathbf{x}' \int d\mathbf{p}' W_{\rho_0 A}(\mathbf{x}_0, \mathbf{p}_0, \mathbf{x}_f, \mathbf{p}_f) B(\mathbf{x}_t, \mathbf{p}_t) \\ \times D_W(\mathbf{x}_0, \mathbf{p}_0, \mathbf{x}_t, \mathbf{p}_t) e^{iS(\mathbf{x}_0, \mathbf{p}_0, \mathbf{x}', \mathbf{p}')/\hbar}, \quad (15)$$

where the phase space variables with initial conditions $\mathbf{x}_0, \mathbf{p}_0$ and the action follow the classical equations of motion forward and backward in time to the final values $\mathbf{x}_f, \mathbf{p}_f$, changing discontinuously at the time t by the amounts

$$\delta\mathbf{x} = \mathbf{x}', \quad \delta\mathbf{p} = \mathbf{p}'. \quad (16)$$

A very practical scheme arises by using a derivative identity to convert the operators to an exponential form and evaluating the resulting expression by FBSD. The correlation function is written as

$$C(t) = -i \frac{\partial}{\partial \mu} \text{Tr} \left(\rho_0 A e^{iHt/\hbar} e^{i\mu B} e^{-iHt/\hbar} \right) \Big|_{\mu=0}. \quad (17)$$

By applying the semiclassical approximation to the product of exponentials in this expression and manipulating the resulting equation, Shao and Makri have shown in Refs. [33,34] that Eq. (17) leads to the expression

$$C(t) = -i\hbar^{-n} \frac{\partial}{\partial \mu} \int d\mathbf{x}_0 \int d\mathbf{p}_0 \exp \left(\frac{i}{\hbar} S(\mathbf{x}_0, \mathbf{p}_0) \right)$$

$$\times \langle g(\mathbf{x}_0, \mathbf{p}_0) | \rho_0 A | g(\mathbf{x}_f, \mathbf{p}_f) \rangle |_{\mu=0}, \quad (18)$$

where the coordinates and action increment at time t by the infinitesimal amounts to

$$\delta \mathbf{x}_t = -\hbar \mu \frac{\partial B(\mathbf{x}_t, \mathbf{p}_t)}{\partial \mathbf{p}_t}, \quad \delta \mathbf{p}_t = \frac{1}{2} \hbar \mu \frac{\partial B(\mathbf{x}_t, \mathbf{p}_t)}{\partial \mathbf{x}_t}, \quad \delta S = \hbar \mu B(\mathbf{x}_t). \quad (19)$$

Note that Eq. (18) contains no prefactor, and that the momentum jump in this expression is exactly one half of that dictated by Hamilton's equations for the effective potential governing the dynamics. This way, the semiclassical prefactor is accounted for by the combined contribution of the action and initial density. The absence of a prefactor from Eq. (18) leads to nearly linear scaling of the required computer time with the number of degrees of freedom. Miller and coworkers have shown that Eq. (18) can also be expressed in a form involving only forward propagation.

Equation (18) is the rigorous stationary phase limit of Eq. (17). However, because the stationary phase approximation with respect to the midpoint integral is applied here to *all* degrees of freedom, it is clear that the interference arising from distinct forward and backward trajectories will be missed, and thus Eq. (18) is not as accurate as semiclassical expressions employing two separate propagators. In fact, the overall behavior of Eq. (18) is similar to that of the Wigner approximation [37,38]

$$C(t) = \hbar^{-n} \int d\mathbf{x}_0 \int d\mathbf{p}_0 W_{\rho_0 A}(\mathbf{x}_0, \mathbf{p}_0) B(\mathbf{x}_t, \mathbf{p}_t), \quad (20)$$

which involves a quasiclassical trajectory average with the Wigner function replacing the classical density. Here W is the Wigner transform of the operator ρ_0 . Sun and Miller have shown that the Wigner quasiclassical method can also be obtained from the full semiclassical expression by linearizing the action difference of the forward and backward trajectories [21].

As a test, Shao and Makri applied Eq. (18) to a model of an initially displaced quartic oscillator coupled to a bath of 30 harmonic oscillators [33] and compared the resulting average position to exact solutions [39]. As seen in Fig. 3, for zero coupling the prefactor-free expression reproduces the first several oscillations of the average position semiquantitatively but fails to capture the rephasing of the wavepacket. When a small amount of dissipation is included, the average position decays irreversibly, and both prefactor-free expressions follow closely the exact results.

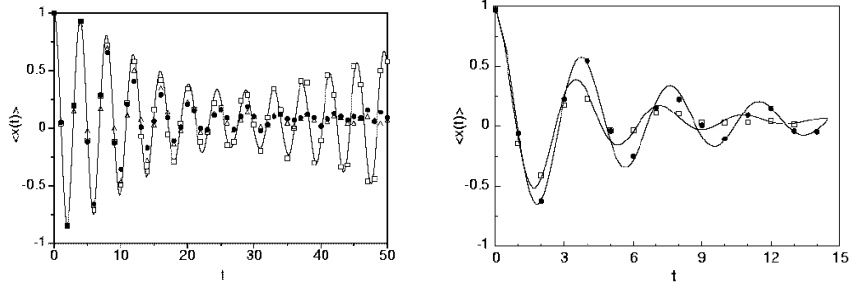


Figure 3. Average position for a one-dimensional quartic oscillator coupled to a bath of 30 harmonic degrees of freedom at zero temperature. Solid line: exact results. Solid circles and hollow squares: FBSD without prefactor, Eq. (18). Hollow triangles: linearized (Wigner) approximation; *left*: no system-bath coupling, *right*: weak and moderate system-bath coupling.

3 Discussion

Shao and Makri used the derivative version of the FBSD formulation summarized in equations (18) and (19) and calculated correlation functions of various normal modes in clusters of two and four water molecules at zero temperature. The largest of these clusters has 30 active degrees of freedom and the calculation involves a 60-dimensional integral which was evaluated with only 2,500 sampling points per integration variable, i.e. a total of 150,000 trajectories. Representative results are shown in Fig. 4 for the water tetramer [34]. The large imaginary parts of these correlation functions is of purely quantum mechanical origin. It is concluded that the prefactor-free FBSD is capable of describing the vibrational dynamics of sizable systems with numerical effort that is comparable to that required in conventional molecular dynamics, while capturing important quantum mechanical effects.

To improve the accuracy of FBSD, one should express the correlation function in a form that includes some explicit interference between forward and backward trajectories. One way to do this is to apply the stationary phase technique to the coordinates of the bath, but evaluate the midpoint integral with respect to the system coordinate numerically [27]. Another possibility is to start from the full semiclassical expression but introduce damping factors that bias sampling around stationary phase paths, with adjustable weights [40]. Both of these approaches lead to improved results where important quantum features can be captured accurately, but at considerably higher

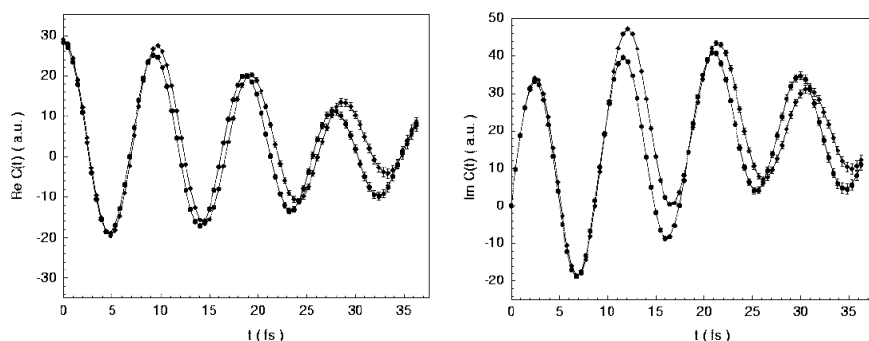


Figure 4. Real and imaginary parts of the correlation function for the acceptor OH stretching vibrations in the water tetramer as obtained by the derivative version of the FBSD procedure.

computational cost.

In summary, the FBSD approach and its numerous variants have tremendously enhanced the feasibility of semiclassical calculations in chemical systems. However, the most practical of these schemes are accurate only for short times and/or when dephasing effects due to multi-dimensional environments quench quantum coherence fairly rapidly. The more accurate versions of FBSD remain computationally expensive and even prohibitive in some cases. Devising semiclassical approximations that are capable of capturing quantum interference effects, while being routinely applicable to large systems, remains the focus of intense efforts in the dynamics community.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Award No. NSF CHE-9877050.

References

- [1] R.P. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948).
- [2] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
- [3] J.D. Doll and D.L. Freeman, *Adv. Chem. Phys.* **73**, 289 (1988).
- [4] N. Makri, *Comp. Phys. Comm.* **63**, 389 (1991).
- [5] N. Makri, *J. Phys. Chem.* **102**, 4414 (1998).

- [6] J.H. Van Vleck, *Proc. Nat. Acad. U.S. Sci.* **14**, 178 (1928).
- [7] C. Morette, *Phys. Rev.* **81**, 848 (1952).
- [8] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [9] L.S. Schulman, *Techniques and Applications of Path Integration* (John Wiley and Sons, New York, 1981).
- [10] W.H. Miller, *Adv. Chem. Phys.* **25**, 69 (1974).
- [11] M.F. Herman and E. Kluk, *Chem. Phys.* **91**, 27 (1984).
- [12] E. Kluk, M.F. Herman, and H.L. Davis, *J. Chem. Phys.* **84**, 326 (1986).
- [13] W.H. Miller, *Adv. Chem. Phys.* **30**, 77 (1975).
- [14] M.S. Child, *Semiclassical Mechanics With Molecular Applications* (Clarendon, Oxford, 1991).
- [15] M.F. Herman, *J. Chem. Phys.* **87**, 4779 (1987); *Ann. Rev. Phys. Chem.* **45**, 83 (1994).
- [16] E.J. Heller, *J. Chem. Phys.* **94**, 2723 (1991).
- [17] M.A. Sepulveda, S. Tomsovic, and E.J. Heller, *Phys. Rev. Lett.* **69**, 402 (1992).
- [18] S. Tomsovic and E.J. Heller, *Phys. Rev. Lett.* **67**, 664 (1991).
- [19] A.R. Walton and D.E. Manolopoulos, *Mol. Phys.* **84**, 961 (1996).
- [20] M.L. Brewer, J.S. Hulme, and D.E. Manolopoulos, *J. Chem. Phys.* **106**, 4832 (1997).
- [21] H. Wang, X. Sun, and W.H. Miller, *J. Chem. Phys.* **108**, 9726 (1998).
- [22] X. Sun and W.H. Miller, *J. Chem. Phys.* **106**, 916 and 6346 (1997); **108**, 8870 (1998).
- [23] X. Sun, H. Wang, and W.H. Miller, *J. Chem. Phys.* **109**, 4190 and 7064 (1998).
- [24] V.S. Batista and W.H. Miller, *J. Chem. Phys.* **108**, 498 (1998).
- [25] N. Makri and K. Thompson, *Chem. Phys. Lett.* **291**, 101 (1998).
- [26] K. Thompson and N. Makri, *J. Chem. Phys.* **110**, 1343 (1999).
- [27] K. Thompson and N. Makri, *Phys. Rev. E* **59**, R4729 (1999).
- [28] O. Kuhn and N. Makri, *J. Phys. Chem.* **103**, 9487 (1999).
- [29] W.H. Miller, *Faraday Discuss.* **110**, 1 (1998).
- [30] X. Sun and W.H. Miller, *J. Chem. Phys.* **110**, 6635 (1999).
- [31] H. Wang, X. Song, D. Chandler, and W.H. Miller, *J. Chem. Phys.* **110**, 4828 (1999).
- [32] V.S. Batista, M.T. Zanni, J. Greenblatt, D.M. Neumark, and W.H. Miller, *J. Chem. Phys.* **110**, 3736 (1999).

- [33] J. Shao and N. Makri, *J. Phys. Chem.* **103**, 7753 (1999).
- [34] J. Shao and N. Makri, *J. Phys. Chem.* **103**, 9479 (1999).
- [35] R.P. Feynman and J.F.L. Vernon, *Ann. Phys. (NY)* **24**, 118 (1963).
- [36] E.J. Wigner, *Chem. Phys.* **5**, 720 (1937).
- [37] E.J. Heller, *J. Chem. Phys.* **65**, 1289 (1976).
- [38] R.C. Brown and E.J. Heller, *J. Chem. Phys.* **75**, 186 (1981).
- [39] N. Makri and D.E. Makarov, *J. Chem. Phys.* **102**, 4611 (1995).
- [40] M. Thoss, H. Wang, and W.H. Miller, *J. Chem. Phys.*, in press (2001).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

THE FEYNMAN PATH GOES MONTE CARLO

T. SAUER

*History and Philosophy of Science,
University of Berne, Sidlerstrasse 5, CH-3012 Berne, Switzerland
E-mail: tilman.sauer@philو.unibe.ch*

Path-integral Monte Carlo (PIMC) simulations have become an important tool for the investigation of the statistical mechanics of quantum systems. I discuss some of the history of applying the Monte Carlo method to non-relativistic quantum systems in path-integral representation. The principle feasibility of the method was well established by the early eighties, a number of algorithmic improvements have been introduced in the last two decades.

1 Introduction

Feynman's classic paper of 1948 presented a *Space-Time Approach to Non-Relativistic Quantum Mechanics* [1], or, in Hagen Kleinert's words [2], "an all-time global approach to the calculation of quantum mechanical amplitudes". Within the philosophy of this approach, we must find, as Kleinert often stressed, "all properties, including the Schrödinger wave functions, from the globally determined time displacement amplitude". The Feynman path, governed by the classical Lagrangian of the quantum system, is the very object we need to study if we want to establish a truly independent alternative to Schrödinger's equation. In avoiding the operator formalism, the sum over paths provides an independent conceptual link between quantum and classical mechanics.

Another attractive feature of the path-integral formulation of quantum mechanics is that it allows to establish a bridge between quantum mechanics and statistical mechanics. Technically, the oscillating exponential of the time-displacement amplitude turns into a positive Boltzmann weight if the paths

are expressed in imaginary time. The quantum mechanical propagator thus turns into the quantum statistical density matrix. It is this very feature which allows the application of methods of classical statistical mechanics, notably the Monte Carlo method, to quantum systems.

The Monte Carlo method came into being roughly around the same time as the Feynman path. Anecdotally, the idea of gaining insight into a complex phenomenon by making various trials and studying the proportions of the respective outcomes occurred to Stanislaw Ulam while playing solitaire during an illness in 1946 [3]. The immediate application was, of course, the problem of neutron diffusion studied in Los Alamos at that time. The name of the procedure first appeared in print in a classic paper by Metropolis and Ulam in 1949 [4], where the authors explicitly mentioned that the method they presented as a statistical approach to the study of integro-differential equations would sometimes be referred to as the Monte Carlo method. In classical statistical mechanics it quickly became a standard calculation tool.

2 Where's Monte Carlo?

The object of interest in Monte Carlo evaluations of Feynman's path integral is the quantum statistical partition function $\mathcal{Z}$ given, in operator language, by the trace of the density operator $\exp(-\beta\hat{H})$ of the canonical ensemble ($\beta = 1/k_B T$) associated with a Hamilton operator describing N particles of mass m_i moving under the influence of a potential V :

$$\hat{H} = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m_i} + V(\hat{r}_1, \dots, \hat{r}_N). \quad (1)$$

Expressed as a Feynman integral, the density matrix elements read

$$\langle \mathbf{r} | \exp(-\beta\hat{H}) | \mathbf{r}' \rangle = \int_{\mathbf{r}(0)=\mathbf{r}}^{\mathbf{r}(\hbar\beta)=\mathbf{r}'} \mathcal{D}\mathbf{r}(\tau) \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} L(\{\dot{\vec{r}}_i(\tau), \vec{r}_i(\tau)\}) d\tau \right\}, \quad (2)$$

where $\mathbf{r} \equiv \{\vec{r}_1, \dots, \vec{r}_N\}$ and L denotes the classical Lagrangian

$$L(\{\dot{\vec{r}}_i(\tau), \vec{r}_i(\tau)\}) = \sum_{i=1}^N \frac{m_i}{2} \dot{\vec{r}}_i^2 + V(\vec{r}_1, \dots, \vec{r}_N(\tau)) \quad (3)$$

expressed in imaginary time τ .^a The particles are assumed to be distinguishable. To evaluate the trace, we only need to set $\mathbf{r} = \mathbf{r}'$ and integrate over $\mathbf{r}$. To take into account Bose or Fermi statistics for indistinguishable particles, the partition function splits into a sum of the direct Boltzmann part and parts with permuted endpoints.

The right hand side of Eq. (2) is a path integral for the $3N$ functions $\mathbf{r}$. The idea of a Monte Carlo evaluation of this quantity is to sample these paths stochastically and to get (approximate) information about the quantum statistics of the system by averaging over the finite set of paths generated in the sampling process.

Monte Carlo data always come with error bars and, in general, the errors associated with numerical Monte Carlo data stem from two distinct sources. A *systematic* error of Monte Carlo evaluations of the path integral follows from the need to identify the paths by a finite amount of computer information. This can be done by discretizing the paths at some set of points in the interval $(0, \hbar\beta)$. For a single particle moving in one dimension, the simplest discrete time approximation for L time slices reads ($\epsilon = \hbar\beta/L$)

$$\langle x | \exp(-\beta\hat{H}) | x' \rangle = \lim_{L \rightarrow \infty} \frac{1}{A} \prod_{j=1}^{L-1} \left[\int \frac{dx_j}{A} \right] \exp \left\{ -\frac{1}{\hbar} \sum_{j=1}^L \left[\frac{m}{2} \frac{(x_j - x_{j-1})^2}{\epsilon} + \epsilon V(x_{j-1}) \right] \right\}, \quad (4)$$

where $A = (2\pi\hbar\epsilon/m)^{1/2}$, $x_0 = x$ and $x_L = x'$. Alternatively, one may expand the individual paths in terms of an orthogonal function basis, e.g. by the Fourier decomposition

$$x(\tau) = x + \frac{(x' - x)\tau}{\hbar\beta} + \sum_{k=1}^{\infty} a_k \sin \frac{k\pi\tau}{\hbar\beta}, \quad (5)$$

and express the density matrix as

$$\langle x | \exp(-\beta\hat{H}) | x' \rangle = \lim_{L' \rightarrow \infty} J \exp \left\{ -\frac{m}{2\hbar^2\beta} (x - x')^2 \right\}$$

^aThere have been attempts to apply the Monte Carlo method to path integrals also for real time. However, due to the oscillating exponential one then has to deal with problems of numerical cancellation, and it is much harder to obtain results of some numerical accuracy. Therefore, I shall here restrict myself to Monte Carlo working in imaginary time.

$$\times \int \prod_{k=1}^{L'} da_k \exp \left\{ -\frac{a_k^2}{2\sigma_k^2} \right\} \times \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} V(x(\tau)) d\tau \right\}, \quad (6)$$

where $\sigma_k = [2\hbar^2\beta/m(\pi k)^2]^{1/2}$ and J is the Jacobian of the transformation from the integral over all paths to the integral over all Fourier coefficients. A systematic error then arises from the loss of information by the finite number L of points x_i on the discretized time axis or by the finite number L' of Fourier components a_k that are taken into account in the Monte Carlo sampling of the paths.

The other error source of Monte Carlo data is the *statistical* error due to the finite number N_m of paths that form the sample used for evaluating the statistical averages. To make matters worse, the probability of configurations is, in general highly peaked, making an independent sampling of paths highly inefficient in most cases. The remedy is to introduce some way of "importance sampling" where configurations are generated according to their probability given by the exponential in Eq. (2). Statistical averages may then be computed as simple arithmetic means. A way to achieve this is by constructing Markov chains where transition probabilities between configurations are constructed that allow to generate a new configuration from a given one such that in the limit of infinitely many configurations the correct probability distribution of paths results. A very simple and universally applicable algorithm to set up such a Markov chain is the Metropolis algorithm introduced in 1953 [5]. Here a new configuration is obtained by looking at some configuration with only one variable changed and accepting or rejecting it for the sample on the basis of a simple rule that only depends on the respective energies of the two configurations. The advantages of importance sampling on the basis of Markov chains are obtained on the cost that, in general, successive configurations are not statistically independent but autocorrelated. The crucial quantity is the integrated autocorrelation time $\tau_{\mathcal{O}}^{\text{int}}$ of a quantity of interest $\mathcal{O} = \langle \overline{\mathcal{O}} \rangle$ with $\overline{\mathcal{O}} = (1/N_m) \sum_{i=1}^{N_m} \mathcal{O}_i$ and $\mathcal{O}_i$ computed for each path i in the sample. It enters the statistical error estimate $\Delta_{\mathcal{O}}$ for expectation values of $\mathcal{O}$ computed from a Monte Carlo sample of N_m autocorrelated configurations as

$$\Delta_{\overline{\mathcal{O}}} = \sqrt{\frac{\sigma_{\mathcal{O}_i}^2}{N_m}} \sqrt{2\tau_{\mathcal{O}}^{\text{int}}}, \quad (7)$$

where $\sigma_{\mathcal{O}_i}^2$ is the variance of $\mathcal{O}_i$.

With Monte Carlo generated samples of Feynman paths one can thus “measure” thermodynamic properties of quantum systems like the internal energy and the specific heat, but also gain more detailed information about correlation functions, probability distributions and the like. In the low-temperature limit, $\beta \rightarrow \infty$, quantum mechanical ground-state properties are recovered.

3 Blazing Trails

A pioneer in the application of the Monte Carlo method to physics problems, notably by applying it to the Ising model, was Lloyd D. Fosdick. He appears to have also been one of the first to consider the stochastic sampling of paths. In 1962, he considered the possibility of sampling paths [6] for what he called the conditional Wiener integral, i.e. the Wiener integral for fixed end points. As a toy example he investigated the expectation value of the functional $\exp \left[- \int_0^1 \int_0^1 \tau \tau' x(\tau) x(\tau') d\tau d\tau' \right]$ for a conditional Wiener process, and, more generally, for the quantity $\exp \left[- \int_0^\beta V(x(\tau)) d\tau \right]$, i.e. he considered direct computation of the partition function for a quantum particle moving in a potential V . He introduced a Fourier decomposition of the paths and generated these by direct Monte Carlo sampling of the Fourier components as Gaussian stochastic variables. He did some explicit sampling of his toy model to demonstrate the feasibility of the method, but a theoretical consideration of the one-dimensional harmonic oscillator was not considered worthwhile to be put on the computer, even though Fosdick, at the time, was at the University of Illinois and had access to the university's ILLIAC computer. His examples were primarily used to investigate the principle feasibility and possible accuracy obtainable by the method. Instead, Fosdick studied a pair of two identical particles and presented some numerical results for this problem. Continuation of the work on the two-particle problem together with a graduate student led to the publication of a paper on the Slater sum, i.e. the diagonal density matrix elements, for ^4He in 1966 [7], and on three-particle effects in the pair distribution function for ^4He in 1968 [8]. In the same year [9], Fosdick elaborated on the numerical method in a report on the Monte Carlo method in quantum statistics in the SIAM review. Instead of sampling the Fourier components he now used the discrete time approximation of the paths. Sampling of x_i at the discrete points was done using a trick that later would gain prominence in PIMC simulations in an algorithm called staging. The

idea was to express the discretized kinetic term in the relative probability density $p(x_i|x_{i-1}, x_{i+1}) = (1/2\pi\epsilon) \exp [-(x_i - x_{i-1})^2/2\epsilon - (x_{i+1} - x_i)^2/2\epsilon]$ as $-(x_i - \bar{x}_i)^2/2\sigma^2 + (x_{i-1} - x_{i+1})^2/4\epsilon$ with $\bar{x}_i = (x_{i-1} + x_{i+1})/2$ and to sample $(x_i - \bar{x}_i)/\sigma$ as an independent Gaussian variable. The procedure could be iterated recursively for all x_i and thus allowed to obtain statistically independent paths which were used to "measure" the potential energy term $\exp \left[- \int_0^\beta V(x(\tau)) d\tau \right]$.

In 1969, Lawande, Jensen, and Sahlin introduced Metropolis sampling of paths in discrete time, broken line approximation [10]. They investigated the ground-state wave functions of simple one-dimensional problems (harmonic oscillator, square well, and Morse potential) and, theoretically, also addressed the problem of extracting information about excited energies and of simulating many particle problems. In a follow-up paper [11] they presented studies of the Coulomb problem using Monte Carlo simulations of the path integral in polar coordinates. Not surprisingly, the singularity at the origin had to be avoided by artificial constraints and the authors admitted that a more rigorous justification of their procedure was called for. The path integral was later solved exactly by Duru and Kleinert in 1979 [12]. It became clear that there were fundamental problems with such singularities in time-sliced path integrals [2].

Little activity is recorded in the seventies, and I am only aware of a brief theoretical consideration of the possibility of Monte Carlo sampling of paths in a paper by Morita on the solution of the Bloch equation for many particle systems in terms of path integrals from 1973 [13]. The paper is cited in a later one by J.A. Barker published in 1979 [14] in which the one-dimensional particle in a box is considered, and numerical estimations of the ground-state energy and wave function are presented. The data were obtained by introducing image sources to take account of the boundary conditions of the box and using Metropolis sampling of the broken line approximation of the paths. Incidentally, the analytic solution of this problem, i.e. of the path integral for the particle in the box was given by Janke and Kleinert almost simultaneously [15]. Barker also computed distribution functions for the problem of two particles in a box.

Very much in the spirit of Lawande *et al.*, but possibly unaware of this work, Creutz and Freedman published a didactic paper on *a statistical approach to quantum mechanics* in 1981 [16]. They, too, performed Metropolis sampling of paths in the broken line approximation and studied the energies

and ground-state wave functions of the one-dimensional harmonic oscillator. The background of these authors were Monte Carlo studies of gauge field theories, and the paper was meant as an attempt to better understand the Monte Carlo method by applying it to simple one-degree-of-freedom Schrödinger problems. It still is a useful introduction to the basics of the technique, and in particular, it presents a brief primer on the theory of Markov chains underlying the Metropolis algorithm. To compute energies, they introduced an alternative estimator by invoking the virial theorem. They also studied double-well problems, presenting snap shot pictures of double-kink instanton paths. The problem of determining the energy level splitting was addressed by computing correlation functions.

The papers by Lawande *et al.* and by Creutz and Freedman appear to have been cited very rarely, possibly because they presented their work as being only of pedagogic value and not so much because the Monte Carlo method could be a useful method to obtain numerical results for Schrödinger problems which, in real life, should be handled by numerical methods more suitable in this simple case. These remarks also hold for work published a little later by Shuryak [17,18].

Fosdick's work from 1962 was done very much at the forefront of the technological possibilities of high-speed computing at the time. By the mid-eighties, path-integral simulations of simple quantum mechanical problems had become both conceptually and technically "easy". Indeed, the exposition by Creutz and Freedman was already written in an introductory, didactic manner, and in 1985 the simulation of the one-particle harmonic oscillator was explicitly proposed as an undergraduate project, to be handled on a Commodore CBM3032 microcomputer, in a paper published in the American Journal of Physics [19].

4 Speeding Up

The feasibility of evaluating the quantum statistical partition function of many-particle systems by Monte Carlo sampling of paths was well established in the early eighties and the method began to be applied to concrete problems, in particular in the chemical physics literature. It had also become clear that the method had severe restrictions if numerical accuracy was called for. In addition to the statistical error inherent to the Monte Carlo method, a systematic error was unavoidably introduced by the necessary discretization of the paths. Attempts to improve the accuracy by algorithmic improvements

to reduce both the systematic and the statistical errors were reported in subsequent years. The literature is abundant and rather than trying to review the field I shall only indicate some pertinent paths of development.

In Fourier PIMC methods, introduced in 1983 in the chemical physics context by Doll and Freeman [20,21], the systematic error arises from the fact that only a finite number of Fourier components are taken into account. Here the systematic error could be reduced by the method of partial averaging [22,23].

In discrete time approximations arising from the short-time propagator or, equivalently, the high-temperature Green's function, various attempts have been made to find more rapidly converging formulations. Among these are attempts to include higher terms in an expansion of the Wigner-Kirkwood type, i.e. an expansion in terms of $\hbar^2/2m$. Taking into account the first term of such an expansion would imply to replace the potential term $\epsilon V(x_{j-1})$ in (4) by [24-26]

$$\epsilon V(x_{j-1}) \rightarrow \frac{\epsilon}{x - x'} \int_x^{x'} dy V(y). \quad (8)$$

This improves the convergence of the density matrix (4) (from even less than $\mathcal{O}(1/L)$ [24] to $\mathcal{O}(1/L^2)$). For the full partition function, the convergence of the simple discretization scheme is already of order $\mathcal{O}(1/L^2)$: due to the cyclic property of the trace, the discretization $\epsilon V(x_{j-1})$ is then equivalent to a symmetrized potential term $\epsilon(V(x_{j-1}) + V(x_j))/2$. The convergence behavior of these formulations follows from the Trotter decomposition formula,

$$e^{-(A+B)} = \left[e^{-\frac{A}{L}} e^{-\frac{B}{L}} \right]^L + \mathcal{O}(1/L) = \left[e^{-\frac{A}{2L}} e^{-\frac{B}{L}} e^{-\frac{A}{2L}} \right]^L + \mathcal{O}(1/L^2), \quad (9)$$

valid for non-commuting operators A and B in a Banach space [27], identifying A with the kinetic energy $\beta \sum \hat{p}_i^2/2m_i$ and B with the potential energy $\beta V(\{\hat{x}_i\})$. More rapidly converging discretization schemes were investigated on the basis of higher-order decompositions. Unfortunately, a direct "fractal" decomposition [28] of the form

$$e^{-(A+B)} = \lim_{L \rightarrow \infty} \left[e^{\alpha_1 \frac{A}{L}} e^{\beta_1 \frac{B}{L}} e^{\alpha_2 \frac{A}{L}} e^{\beta_2 \frac{B}{L}} \dots \right]^L, \quad \sum \alpha_i = \sum \beta_i = 1 \quad (10)$$

inevitably leads to negative coefficients for higher decompositions [29] and is therefore not amenable to Monte Carlo sampling of paths [30]. Higher-order Trotter decomposition schemes involving commutators have proven to

be more successful [31-34]. In particular, a decomposition of the form

$$\mathcal{Z} = \lim_{L \rightarrow \infty} \text{Tr} \left[e^{-\frac{A}{2L}} e^{-\frac{B}{2L}} e^{-\frac{[[B,A],B]}{24L^3}} e^{-\frac{B}{2L}} e^{-\frac{A}{2L}} \right]^L, \quad (11)$$

derivable by making use of the cyclic property of the trace, is convergent of order $\mathcal{O}(1/L^4)$ and amounts to simply replacing the potential ϵV in (4) by an effective potential [32]

$$V_{\text{eff}} = V + \frac{(\beta\hbar)^2}{24mL^2} (V')^2. \quad (12)$$

Another problem for the numerical accuracy of PIMC simulations arises from the analog of the “critical slowing down” problem which is well known for local update algorithms at second-order phase transitions in the simulation of spin systems and lattice field theory. Since the correlations $\langle x_j x_{j+k} \rangle$ between variables x_j and x_{j+k} in the discrete time approximation only depend on the temperature and on the gaps between the energy levels and not, or at least not appreciably, on the discretization parameter ϵ , the correlation length ζ along the discretized time axis always diverges linearly with L when measured in units of the lattice spacing ϵ . Hence in the continuum limit of $\epsilon \rightarrow 0$ with β fixed or, equivalently, of $L \rightarrow \infty$ for local, importance sampling update algorithms, like the standard Metropolis algorithm, a slowing down occurs because paths generated in the Monte Carlo process become highly correlated. Since autocorrelation times diverge in simulations using the Metropolis algorithm as [35] $\tau_{\mathcal{O}}^{\text{int}} \propto L^z$ with $z \approx 2$, the computational effort (CPU time) to achieve comparable numerical accuracy in the continuum limit $L \rightarrow \infty$ diverges as $L \times L^z = L^{z+1}$.

To overcome this drawback, ad hoc algorithmic modifications like introducing collective moves of the path as a whole between local Metropolis updates were introduced then and again. One of the earliest more systematic and successful attempts to reduce autocorrelations between successive path configurations was introduced in 1984 by Pollock and Ceperly [36]. Rewriting the discretized path integral, their method essentially amounts to a recursive transformation of the variables x_i in such a way that the kinetic part of the energy can be taken care of by sampling direct Gaussian random variables and a Metropolis choice is made for the potential part. The recursive transformation can be done between some fixed points of the discretized paths, and the method has been applied in such a way that successively finer discretizations of the path were introduced between neighbouring points. Invoking the poly-

mer analog of the discretized path this method was christened the “staging” algorithm by Sprik, Klein, and Chandler in 1985 [37].

The staging algorithm decorrelates successive paths very effectively because the whole staging section of the path is essentially sampled independently. In 1993, another explicitly non-local update was applied to PIMC simulations [35,38] by transferring the so-called multigrid method known from the simulation of spin systems. Originating in the theory of numerical solutions of partial differential equations, the idea of the multigrid method is to introduce a hierarchy of successively coarser grids in order to take into account long wavelength fluctuations more effectively. Moving variables of the coarser grids then amounts to a collective move of neighbouring variables of the finer grids, and the formulation allows to give a recursive description of how to cycle most effectively through the various levels of the multigrid. Particularly successful is the so-called W-cycle. Both the staging algorithm and the multigrid W-cycle have been shown to beat the slowing down problem in the continuum limit completely by reducing the exponent z to $z \approx 0$ [39].

Another cause of severe correlations between paths arises if the probability density of configurations is sharply peaked with high maxima separated by regions of very low probability density. In the statistical mechanics of spin systems this is the case at a first-order phase transition. In PIMC simulations the problem arises for tunneling situations like, e.g. for a double-well potential with a high potential barrier between the two wells. In these cases, an unbiased probing of the configuration space becomes difficult because the system tends to get stuck around one of the probability maxima. A remedy to this problem is to simulate an auxiliary distribution that is flat between the maxima and to recover the correct Boltzmann distribution by an appropriate reweighting of the sample. The procedure is known under the name of umbrella sampling or multicanonical sampling. It was shown to reduce autocorrelations for PIMC simulations of a single particle in a one-dimensional double well, and it can also be combined with multigrid acceleration [40].

The statistical error associated with a Monte Carlo estimate of an observable $\mathcal{O}$ cannot only be minimized by reducing autocorrelation times $\tau_{\mathcal{O}}^{\text{int}}$. If the observable can be measured with two different estimators U_i that yield the same mean $U_i^{(L)} = \langle U_i \rangle$ with $\mathcal{O} = \lim_{L \rightarrow \infty} U_i^{(L)}$, the estimator with the smaller variance $\sigma_{U_i}^2$ is to be preferred. Straightforward differentiation of the discretized path integral (4) leads to an estimator of the energy that explicitly

measures the kinetic and potential parts of the energy by

$$U_k = \frac{L}{2\beta} - \frac{m}{2L} \sum_{j=1}^L \left(\frac{x_j - x_{j-1}}{\epsilon} \right)^2 + \frac{1}{L} \sum_{i=1}^L V(x_i). \quad (13)$$

The variance of this so-called “kinetic” energy estimator diverges with L . Another estimator can be derived by invoking the path-integral analog of the virial theorem

$$\frac{L}{2\beta} - \frac{m}{2} \left\langle \left(\frac{x_j - x_{j-1}}{\epsilon} \right)^2 \right\rangle = \frac{1}{2} \langle x_j V'(x_j) \rangle, \quad (14)$$

and the variance of the “virial” estimator

$$U_v = \frac{1}{2L} \sum_{i=1}^L x_i V'(x_i) + \frac{1}{L} \sum_{i=1}^L V(x_i) \quad (15)$$

does not depend on L . In the early eighties, investigations of the “kinetic” and the “virial” estimators focussed on their variances [32,41,42]. Some years later, it was pointed out [43] that a correct assessment of the accuracy also has to take into account the autocorrelations, and it was demonstrated that, for a standard Metropolis simulation of the harmonic oscillator, the allegedly less successful “kinetic” estimator gave smaller errors than the “virial” estimator. In 1989 it was shown [44] that conclusions about the accuracy also depend on the particular Monte Carlo update algorithm at hand, since modifications of the update scheme such as inclusion of collective moves of the whole path affect the autocorrelations of the two estimators in a different way. A careful comparison of the two estimators which disentangles the various factors involved was given only quite recently [45]. Here it was also shown that a further reduction of the error may be achieved by a proper combination of both estimators without extra cost.

5 Concluding Remarks

The application of the Monte Carlo method to quantum systems is not restricted to direct sampling of Feynman paths, but this method has attractive features. It is not only conceptually suggestive but also allows for algorithmic improvements that help to make the method useful even when the problems at hand require considerable numerical accuracy. However, algorithmic improvements like the ones alluded to above have been proposed and tested

mainly for simple one-particle systems. On the other hand, the power of the Monte Carlo method is, of course, most welcome in those cases where analytical methods fail. For more complicated systems, however, evaluating the algorithms and controlling the numerical accuracy is also more difficult. Only recently, a comparison of the efficiency of Fourier- and discrete time-path integral Monte Carlo for a cluster of 22 hydrogen molecules was presented [46] and debated [47,48]. Nevertheless, path-integral Monte Carlo simulations have become an essential tool for the treatment of strongly interacting quantum systems, like, e.g. the theory of condensed helium [49].

Acknowledgments

I wish to thank Wolfhard Janke for an instructive and enjoyable collaboration.

References

- [1] R.P. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948).
- [2] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics* (World Scientific, Singapore, 1990).
- [3] W. Aspray, in *The History of Modern Mathematics. Vol. II: Institutions and Applications*, Eds. D.E. Rowe and J. McCleary (Academic Press, Boston, 1989), p. 312.
- [4] N. Metropolis and S. Ulam, *J. Am. Stat. Ass.* **44**, 1949 (335).
- [5] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [6] L.D. Fosdick, *J. Math. Phys.* **3**, 1251 (1962).
- [7] L.D. Fosdick and H.F. Jordan, *Phys. Rev.* **143**, 58 (1966).
- [8] H.F. Jordan and L.D. Fosdick, *Phys. Rev.* **171**, 129 (1968).
- [9] L.D. Fosdick, *SIAM Review* **10**, 315 (1968).
- [10] S.V. Lawande, C.A. Jensen, and H.L. Sahlin, *J. Comp. Phys.* **3**, 416 (1969).
- [11] S.V. Lawande, C.A. Jensen, and H.L. Sahlin, *J. Comp. Phys.* **4**, 451 (1969).
- [12] I.H. Duru and H. Kleinert *Phys. Lett. B* **84**, 185 (1979).
- [13] T. Morita, *J. Phys. Soc. Jpn.* **35**, 980 (1973).
- [14] J.A. Barker, *J. Chem. Phys.* **70**, 2914 (1979).
- [15] W. Janke and H. Kleinert, *Lett. Nuovo Cim.* **25**, 297 (1979).
- [16] M. Creutz and B. Freedman, *Ann. Phys. (N.Y.)* **132**, 427 (1981).

- [17] E.V. Shuryak and O.V. Zhirov, *Nucl. Phys. B* **242**, 393 (1984).
- [18] E.V. Shuryak, *Sov. Phys. Usp.* **27**, 448 (1984).
- [19] P.K. MacKeown, *Am. J. Phys.* **53**, 880 (1985).
- [20] J.D. Doll and D.L. Freeman, *J. Chem. Phys.* **80**, 2239 (1984).
- [21] D.L. Freeman and J.D. Doll, *J. Chem. Phys.* **80**, 5709 (1984).
- [22] J.D. Doll, R.D. Coalson, and D.L. Freeman, *Phys. Rev. Lett.* **55**, 1 (1985).
- [23] R.D. Coalson, D.L. Freeman, and J.D. Doll, *J. Chem. Phys.* **85**, 4567 (1986).
- [24] N. Makri and W.H. Miller, *Chem. Phys. Lett.* **151**, 1 (1988).
- [25] N. Makri and W.H. Miller, *J. Chem. Phys.* **90**, 904 (1989).
- [26] I. Bender, D. Gromes, and U. Marquard, *Nucl. Phys. B* **346**, 593 (1990).
- [27] M. Suzuki, *J. Math. Phys.* **26**, 601 (1985).
- [28] M. Suzuki, *Phys. Lett. A* **146**, 319 (1990).
- [29] M. Suzuki, *J. Math. Phys.* **32**, 400 (1991).
- [30] W. Janke and T. Sauer, *Phys. Lett. A* **165**, 199 (1992).
- [31] H. De Raedt and B. De Raedt, *Phys. Rev. A* **28**, 3575 (1983).
- [32] M. Takahashi and M. Imada, *J. Phys. Soc. Jpn.* **53**, 963, 3765 (1984).
- [33] X.-P. Li and J.Q. Broughton, *J. Chem. Phys.* **86**, 5094 (1987).
- [34] H. Kono, A. Takasaka, and S.H. Lin, *J. Chem. Phys.* **88**, 6390 (1988).
- [35] W. Janke and T. Sauer in *Path Integrals from meV to MeV: Tutzing '92*, Eds. H. Grabert, A. Inomata, L.S. Schulman, and U. Weiss (World Scientific, Singapore, 1993).
- [36] E.L. Pollock and D.M. Ceperley, *Phys. Rev. B* **30**, 2555 (1984).
- [37] M. Sprik, M.L. Klein, and D. Chandler, *Phys. Rev. B* **31**, 4234 (1985).
- [38] W. Janke and T. Sauer, *Chem. Phys. Lett.* **201**, 499 (1993).
- [39] W. Janke and T. Sauer, *Chem. Phys. Lett.* **263**, 488 (1996).
- [40] W. Janke and T. Sauer, *Phys. Rev. E* **49**, 3475 (1994).
- [41] M.F. Herman, E.J. Bruskin, and B.J. Berne, *J. Chem. Phys.* **76**, 5150 (1982).
- [42] M. Parrinello and A. Rahman, *J. Chem. Phys.* **80**, 860 (1984).
- [43] A. Giansanti and G. Jacucci, *J. Chem. Phys.* **89**, 7454 (1988).
- [44] J.S. Cao and B.J. Berne, *J. Chem. Phys.* **91**, 6359 (1989).
- [45] W. Janke and T. Sauer, *J. Chem. Phys.* **107**, 5821 (1997).
- [46] C. Chakravarty, M.C. Gordillo, and D.M. Ceperley, *J. Chem. Phys.* **109**, 2123 (1998).
- [47] J.D. Doll and D.L. Freeman, *J. Chem. Phys.* **111**, 7685 (1999).

- [48] C. Chakravarty, M.C. Gordillo, and D.M. Ceperley, *J. Chem. Phys.* **111**, 7687 (1999).
- [49] D.M. Ceperley, *Rev. Mod. Phys.* **67**, 279 (1995).

COORDINATE INDEPENDENCE OF PATH INTEGRALS IN ONE-DIMENSIONAL TARGET SPACE

A. CHERVYAKOV

*Institut für Theoretische Physik, Freie Universität Berlin,
Arnimallee 14, D-14195 Berlin, Germany
E-mail: chervyak@physik.fu-berlin.de*

The recent perturbative formulation of the quantum mechanical path integral in a curved space allows one to calculate the Feynman diagrams involving multiple temporal integrals over products of distributions. To test the consistency of this formulation the exactly solvable path integral for the particle in a flat space is transformed into the curvilinear form of a nonlinear sigma model where the only solution is the perturbation expansion. By an explicit three-loop calculation we show that the perturbatively defined path integral in new curvilinear coordinates reproduces the ground-state energy of the original system thus ensuring the coordinate independence.

1 Introduction

In honor of Professor Kleinert's birthday I would like to review some of the most interesting results we found in a series of recent papers written in a common collaboration [1–3].

By an arbitrary coordinate transformation from flat to curved space the exactly solvable path integral for the particle in a flat space can be turned into the curvilinear form of a nonlinear sigma model where the only solution is the perturbation expansion. The formal perturbative definition of quantum mechanical path integrals in curvilinear coordinates, however, poses problems. To exhibit the difficulties, consider the associated partition function calculated for periodic paths on the imaginary-time axis τ :

$$Z = \int \mathcal{D}q(\tau) \sqrt{g} e^{-\mathcal{A}[q]}, \quad (1)$$

where $\mathcal{A}[q]$ is the Euclidean action with the general form

$$\mathcal{A}[q] = \int d\tau \left[\frac{1}{2} g_{\mu\nu}(q(\tau)) \dot{q}^\mu(\tau) \dot{q}^\nu(\tau) + V(q(\tau)) \right]. \quad (2)$$

The dots denote τ -derivatives, $g_{\mu\nu}(q)$ is a metric, and $g = \det g_{\mu\nu}$ its determinant. The path integral may formally be defined perturbatively as follows: The metric $g_{\mu\nu}(q)$ is expanded around some point q_0^μ in powers of the deviation $\delta q^\mu \equiv q^\mu - q_0^\mu$. The same thing is done with the potential $V(q)$. After this, the action $\mathcal{A}[q]$ is separated into a free part $\mathcal{A}_0[q_0; \delta q] \equiv g_{\mu\nu}(q_0) \dot{q}^\mu \dot{q}^\nu / 2 + \omega^2 \delta q^\mu \delta q^\nu / 2$, and an interacting part $\mathcal{A}_{\text{int}}[q_0; \delta q] \equiv \mathcal{A}[q] - \mathcal{A}_0[q_0; \delta q]$. The square root in the path integral (1) is taken into the exponent and expanded likewise, defining an effective action $\mathcal{A}_{\sqrt{g}} = -\delta(0) \log[g(q_0 + \delta q)/g(q_0)]/2$, where $\delta(0)$ is the δ -function at the origin. It represents formally the inverse infinitesimal lattice spacing on the time axis, and is equal to $\delta(0) \equiv \int dp/(2\pi)$.

An expansion of Z in powers of the interaction leads to a sum of loop diagrams. Problems arise from the fact that there are interaction terms involving $\dot{q}^2 q^n$ which lead to Feynman integrals over products of distributions. There are now three types of lines representing the correlation functions:

$$\Delta(\tau, \tau') \equiv \langle q(\tau) q(\tau') \rangle = \text{———}, \quad (3)$$

$$\partial_\tau \Delta(\tau, \tau') \equiv \langle \dot{q}(\tau) q(\tau') \rangle = \text{-----}, \quad (4)$$

$$\partial_\tau \partial_{\tau'} \Delta(\tau, \tau') \equiv \langle \dot{q}(\tau) \dot{q}(\tau') \rangle = \text{-----}. \quad (5)$$

The right-hand sides define the line symbols to be used in Feynman diagrams for the interaction terms.

Explicitly, the first correlation function reads

$$\Delta(\tau, \tau') = \frac{1}{2\omega} e^{-\omega|\tau - \tau'|}. \quad (6)$$

The second correlation function (4) has a discontinuity

$$\partial_\tau \Delta(\tau, \tau') = -\frac{1}{2} \epsilon(\tau - \tau') e^{-\omega|\tau - \tau'|}, \quad (7)$$

where

$$\epsilon(\tau - \tau') \equiv -1 + 2 \int_{-\infty}^{\tau} d\tau'' \delta(\tau'' - \tau') \quad (8)$$

is a distribution which vanishes at the origin and is equal to ± 1 for positive and negative arguments, respectively. The third correlation function (5)

contains a δ -function:

$$\partial_\tau \partial_{\tau'} \Delta(\tau, \tau') = \delta(\tau - \tau') - \frac{\omega}{2} e^{-\omega|\tau - \tau'|}. \quad (9)$$

Recently, the temporal integrals over products of such distributions were defined uniquely and independently of the choice of coordinates [1–3]. In Ref. [1] it was shown that Feynman integrals in *momentum space* can be uniquely defined as $\epsilon \rightarrow 0$ -limits of $(1 - \epsilon)$ -dimensional integrals via an analytic continuation à la 't Hooft and Veltman [4]. This definition makes path integrals coordinate-independent. In Ref. [2] the rules for the calculation of temporal integrals over products of distributions were formulated directly for Feynman integrals in the $(1 - \epsilon)$ -dimensional time space. In Ref. [3] these integrals were expressed uniquely via integrals over products of nonsingular correlation functions $\Delta(\tau, \tau')$, plus integrals over pure products of δ -functions. The rules defining the integrals over distributional products were set up directly in one dimension by the requirement of coordinate invariance of the path integral (1). This approach has the important advantage to make superfluous the somewhat tedious analytic continuation to $1 - \epsilon$ dimensions. In fact, it does not require specifying any regularization scheme. In addition, it gives a foundation of a new and general mathematics of extending the theory of distributions from a linear space to products.

In this note, using the results of the previous work [1–3], we test the coordinate independence of perturbatively defined path integrals by an explicit three-loop calculation of the ground-state energy.

2 Model System

We want to test the coordinate independence of the exactly solvable path integral of a point particle of unit mass in a harmonic potential $\omega^2 x^2/2$, over a large imaginary-time interval β ,

$$Z_\omega = \int \mathcal{D}x(\tau) e^{-\mathcal{A}_\omega[x]} = e^{-\text{Tr} \log(-\partial^2 + \omega^2)} = e^{-\beta\omega/2}. \quad (10)$$

The action is

$$\mathcal{A}_\omega = \frac{1}{2} \int d\tau [\dot{x}^2(\tau) + \omega^2 x^2(\tau)]. \quad (11)$$

Here and in the following, the target space is assumed to be one-dimensional for simplicity.

A coordinate transformation turns (10) into a path integral of the type (1) with a singular perturbation expansion.

From our work in Refs. [1,2] we know that all terms in this expansion vanish in dimensional regularization. Here we shall use the identities for integrals over products of distributions found in Ref. [3] to perform the perturbative calculation directly in one dimension.

For simplicity we assume the coordinate transformation to preserve the symmetry $x \leftrightarrow -x$ of the initial oscillator, such that its power series expansion starts out like $x(\tau) = f(q(\tau)) = q - gq^3/3 + g^2aq^5/5 - \dots$, where g is a smallness parameter and a an extra parameter. We shall see that the identities are independent of a , such that a will merely serve to check the calculations. The transformation changes the partition function (10) into

$$Z = \int \mathcal{D}q(\tau) e^{-\mathcal{A}_J[q]} e^{-\mathcal{A}[q]}, \quad (12)$$

where $\mathcal{A}[q]$ is the transformed action, whereas $\mathcal{A}_J[q]$ represents an effective action coming from the Jacobian of the coordinate transformation:

$$\mathcal{A}_J[q] = -\delta(0) \int d\tau \log \frac{\delta f(q(\tau))}{\delta q(\tau)}. \quad (13)$$

Of course, the quantum mechanics is the UV-finite one-dimensional quantum field theory. Up to the three-loop level, this was shown in Refs. [1,2] via an analytic continuation to $1 - \epsilon$ dimensions in the $\epsilon \rightarrow 0$ -limit. Thus, working in one dimension directly, we can ignore the intermediate linear and quadratic divergences arising from the loop diagrams. Instead, we want to trace out explicitly the cancellation of these divergences against those coming from the Jacobian action (13).

The transformed action is decomposed into a free part

$$\mathcal{A}_\omega[q] = \frac{1}{2} \int d\tau [\dot{q}^2(\tau) + \omega^2 q^2(\tau)], \quad (14)$$

and an interacting part, which reads to second order in g :

$$\begin{aligned} \mathcal{A}_{\text{int}}[q] = \frac{1}{2} \int d\tau \left\{ -g \left[2\dot{q}^2(\tau)q^2(\tau) + \frac{2\omega^2}{3}q^4(\tau) \right] \right. \\ \left. + g^2 \left[(1+2a)\dot{q}^2(\tau)q^4(\tau) + \omega^2 \left(\frac{1}{9} + \frac{2a}{5} \right) q^6(\tau) \right] \right\}. \end{aligned} \quad (15)$$

To the same order in g , the Jacobian action (13) is

$$\mathcal{A}_J[q] = -\delta(0) \int d\tau \left[-gq^2(\tau) + g^2 \left(a - \frac{1}{2} \right) q^4(\tau) \right]. \quad (16)$$

For $g = 0$, the transformed partition function (12) coincides with (10). When expanding Z of Eq. (12) in powers of g , we obtain sums of Feynman diagrams contributing to each order g^n , which must vanish to ensure coordinate invariance. By considering only connected Feynman diagrams, we are dealing directly with the ground-state energy.

3 Free Energy Density

The graphical expansion for the ground-state energy will be carried here only up to three loops. At any order g^n , there exist different types of Feynman diagrams with $L = n + 1, n$, and $n - 1$ number of loops coming from the interaction terms (15) and (16), respectively. The diagrams are composed of the three types of lines in Eqs. (3)–(5), and of new interaction vertices for each power of g . The diagrams coming from the Jacobian action (16) are easily recognized by an accompanying power of $\delta(0)$.

At first order in g , there exist only three diagrams, two originated from the interaction (15), one from the Jacobian action (16):

$$-g \text{ (diagram 1)} - g\omega^2 \text{ (diagram 2)} + g\delta(0) \text{ (diagram 3)}. \quad (17)$$

At order g^2 , we distinguish several contributions. First there are two three-loop local diagrams coming from the interaction (15), and one two-loop local diagram from the Jacobian action (16):

$$g^2 \left[3 \left(\frac{1}{2} + a \right) \text{ (diagram 4)} + 15\omega^2 \left(\frac{1}{18} + \frac{a}{5} \right) \text{ (diagram 5)} - 3 \left(a - \frac{1}{2} \right) \delta(0) \text{ (diagram 6)} \right]. \quad (18)$$

We call a diagram local if it involves no temporal time integral. The Jacobian action (16) contributes further the nonlocal diagrams:

$$-\frac{g^2}{2!} \left\{ 2\delta^2(0) \text{ (diagram 7)} - 4\delta(0) [\text{diagram 8} + \text{diagram 9} + 2\omega^4 \text{ (diagram 10)}] \right\}. \quad (19)$$

The remaining diagrams come from the interaction (15) only. They are either of the three-bubble type, or of the watermelon type, each with all possible

combinations of the three-line types (3)–(5): The sum of all three-bubbles diagrams is

$$-\frac{g^2}{2!} \left[4 \text{ (diagram 1)} + 2 \text{ (diagram 2)} + 2 \text{ (diagram 3)} + 8\omega^2 \text{ (diagram 4)} + 8\omega^2 \text{ (diagram 5)} + 8\omega^4 \text{ (diagram 6)} \right], \quad (20)$$

while the watermelon-like diagrams contribute

$$-\frac{g^2}{2!} 4 \left[\text{diagram 7} + 4 \text{ (diagram 8)} + \text{diagram 9} + 4\omega^2 \text{ (diagram 10)} + \frac{2}{3}\omega^4 \text{ (diagram 11)} \right]. \quad (21)$$

Since the equal-time expectation value $\langle \dot{q}(\tau) q(\tau) \rangle$ vanishes according to Eq. (7), there are a number of trivially vanishing diagrams, which have been omitted.

In previous papers [1,2], all integrals were calculated individually in $D = 1 - \varepsilon$ dimensions, taking the limit $\varepsilon \rightarrow 0$ at the end. The results for the integrals ensured that the sum of all Feynman diagrams contributing to each order g^n vanishes.

4 Integrals over Distributions

Here we shall follow the developments of Ref. [3] for calculating temporal integrals over products of distributions directly in one dimension. These integrals can be expressed formally via integrals over products of nonsingular correlation functions $\Delta(\tau, \tau')$, plus integrals containing pure products of δ - and ϵ -functions.

First, we shall calculate directly some basic integrals. Most simply, we obtain the relations

$$\int d\tau \Delta^2(\tau) = \frac{1}{2\omega^2} \Delta(0), \quad (22)$$

$$\int d\tau \dot{\Delta}^2(\tau) = \frac{1}{2} \Delta(0). \quad (23)$$

Therefore we have

$$\int d\tau \left[\dot{\Delta}^2(\tau) + \omega^2 \Delta^2(\tau) \right] = \Delta(0). \quad (24)$$

Two other elementary integrals read

$$\int d\tau \Delta^4(\tau) = \frac{\Delta^3(0)}{4\omega^2}, \quad (25)$$

$$\int d\tau \dot{\Delta}^2(\tau) \Delta^2(\tau) = \frac{1}{4} \Delta^3(0). \quad (26)$$

Note that these relations can easily be obtained with the help of partial integration and the inhomogeneous field equation satisfied by the correlation function

$$\ddot{\Delta}(\tau) = - \int dk \frac{k^2}{k^2 + \omega^2} e^{ik\tau} = -\delta(\tau) + \omega^2 \Delta(\tau). \quad (27)$$

To derive, for example, Eq. (24) we have first to integrate the first term partially,

$$\int d\tau \dot{\Delta}^2(\tau) = - \int d\tau \Delta(\tau) \ddot{\Delta}(\tau), \quad (28)$$

with no boundary term due to the exponential vanishing at infinity of all functions involved. Using then the field equation (27) and the definition of the δ -function

$$\int d\tau f(\tau) \delta(\tau) = f(0), \quad (29)$$

we obtain (24).

The same is valid for Eq. (26): by a partial integration, the left-hand side becomes

$$\int d\tau \dot{\Delta}^2(\tau) \Delta^2(\tau) = -\frac{1}{3} \int d\tau \ddot{\Delta}(\tau) \Delta^3(\tau). \quad (30)$$

Applying again the field equation (27) and Eq. (25), we reduce the right-hand side to

$$- \int d\tau \ddot{\Delta}(\tau) \Delta^3(\tau) = \Delta^3(0) - \omega^2 \int d\tau \Delta^4(\tau) = \frac{3}{4} \Delta^3(0). \quad (31)$$

Substituting this into Eq. (30), yields (26).

Consider now more singular integrals over $\ddot{\Delta}^2(\tau)$, $\ddot{\Delta}(\tau) \dot{\Delta}^2(\tau) \Delta(\tau)$ and $\ddot{\Delta}^2(\tau) \Delta^2(\tau)$ involving pure products of δ - and ϵ -functions. As a first step in calculating these integrals we replace $\ddot{\Delta}(\tau)$ by the left-hand side of Eq. (27). This reduces the integral involving $\ddot{\Delta}^2(\tau)$ to the relation

$$\int d\tau \left[\ddot{\Delta}^2(\tau) + 2\omega^2 \dot{\Delta}^2(\tau) + \omega^4 \Delta^2(\tau) \right] = \int d\tau \delta^2(\tau). \quad (32)$$

Combining the same tool with Eq. (26), we rewrite the integral over $\ddot{\Delta}(\tau)\dot{\Delta}^2(\tau)\Delta(\tau)$ in the following form

$$\int d\tau \ddot{\Delta}(\tau)\dot{\Delta}^2(\tau)\Delta(\tau) = -\frac{1}{8\omega} \int d\tau \epsilon^2(\tau)\delta(\tau) + \frac{1}{4}\omega^2\Delta^3(0). \quad (33)$$

Let us turn now to the integral over $\ddot{\Delta}^2(\tau)\Delta^2(\tau)$. This can be expressed in terms of the regular integral

$$\int d\tau \dot{\Delta}^4(\tau) = \frac{1}{4}\omega^2\Delta^3(0) \quad (34)$$

as follows: Invoking once more the field equation (27) and Eq. (34), we find

$$\int d\tau \ddot{\Delta}^2(\tau)\Delta^2(\tau) = \int d\tau \Delta^2(\tau)\delta^2(\tau) - \frac{7}{4}\omega^2\Delta^3(0). \quad (35)$$

As in standard quantum field theory, we should define now the integrals over distributional products of δ - and ϵ -functions, occurring in Eqs. (32), (33) and (35). Before coming to this, however, we note that the partial integration rule for integrals of this kind is not generally allowed in the purely one-dimensional approach. To illustrate the problem involved with partial integration, consider the integral

$$\int d\tau \epsilon^2(\tau)\delta(\tau)e^{-\lambda|\tau|}, \quad (36)$$

with an arbitrary parameter λ . Applying partial integration reduces this to the completely regular form

$$3 \int d\tau \epsilon^2(\tau)\delta(\tau)e^{-\lambda|\tau|} = \frac{\lambda}{2} \int d\tau \epsilon^4(\tau)e^{-\lambda|\tau|} = 1, \quad (37)$$

for any value of λ . Taking Eq. (34) into account, it is easy to see that the result (37) satisfies the identity

$$\begin{aligned} \int d\tau \Delta(\tau)\dot{\Delta}^2(\tau)\ddot{\Delta}(\tau) &= -\frac{1}{3} \int d\tau \dot{\Delta}^4(\tau) \\ &= -\frac{1}{12}\omega^2\Delta^3(0) = -\frac{1}{24\omega} + \frac{1}{4}\omega^2\Delta^3(0) \end{aligned} \quad (38)$$

derived partially from Eq. (33).

On the other hand, the integral (36) can be transformed by the partial integration into another form

$$\begin{aligned} \int d\tau \epsilon^2(\tau) \delta(\tau) e^{-\lambda|\tau|} &= \int d\tau \epsilon(\tau) \delta(\tau) \frac{d}{d\tau} \left[-\frac{1}{\lambda} e^{-\lambda|\tau|} \right] \\ &= \frac{1}{\lambda} \left[2 \int d\tau \delta^2(\tau) e^{-\lambda|\tau|} + \int d\tau \epsilon(\tau) \dot{\delta}(\tau) e^{-\lambda|\tau|} \right], \end{aligned} \quad (39)$$

where the role of the parameter λ becomes crucial. Integrating by parts now the left-hand side of Eq. (32) and using Eq. (39), we obtain

$$\begin{aligned} \int d\tau \left[\ddot{\Delta}^2(\tau) + 2\omega^2 \dot{\Delta}^2(\tau) + \omega^4 \Delta^2(\tau) \right] &= \\ &= \int d\tau \delta^2(\tau) - \frac{1}{2\omega} \int d\tau \epsilon^2(\tau) \delta(\tau) + \frac{1}{2\omega}. \end{aligned} \quad (40)$$

Contrary to the result (37), it follows from Eqs. (32) and (40) that

$$\int d\tau \epsilon^2(\tau) \delta(\tau) e^{-\lambda|\tau|} = 1. \quad (41)$$

Therefore, in integrals involving singular products of distributions, partial integration is bound to fail since it leaves no room to specify the integral over $\epsilon^2 \delta$ consistently. Abandoning the partial integration, we shall define the following rules for an integral over products of two δ -functions in Eqs. (32) and (35)

$$\int d\tau f(\tau) \delta^2(\tau) = f(0) \delta(0), \quad (42)$$

and for the integral over a product of two ϵ -functions with one δ -function in Eq. (33):

$$\int d\tau f(\tau) \epsilon^2(\tau) \delta(\tau) = \frac{1}{4} f(0), \quad (43)$$

for any smooth test function $f(\tau)$.

The obtained relations have reduced all integrals over singular products of correlation functions to regular integrals plus integrals containing $\delta^2(\tau)$ and $\epsilon^2(\tau) \delta(\tau)$. For the last, we have the integration rules (42) and (43) which determine completely the right-hand sides of relations (32), (33) and (35). We are now going to show that these rules hold the required coordinate independence of the path integral up to a three-loop approximation.

5 Reparametrization Invariance

To first order in g , the sum of Feynman diagrams (17) vanishes:

$$\text{Diagram 1} + \omega^2 \text{Diagram 2} - \delta(0) \text{Diagram 3} = 0. \quad (44)$$

The analytic form of this relation is

$$\left[-\ddot{\Delta}(0) + \omega^2 \Delta(0) - \delta(0) \right] \Delta(0) = 0, \quad (45)$$

and the vanishing is a direct consequence of the field equation (27) for the correlation function at the origin.

At order g^2 , the same equation reduces the sum of all local diagrams in (18) to a finite result plus a term proportional to $\delta(0)$:

$$\begin{aligned} & \left[-3 \left(\frac{1}{2} + a \right) \ddot{\Delta}(0) + 15 \left(\frac{1}{18} + \frac{a}{5} \right) \omega^2 \Delta(0) \right. \\ & \left. - 3 \left(a - \frac{1}{2} \right) \delta(0) \right] \Delta^2(0) = \left[3\delta(0) - \frac{2}{3} \omega^2 \Delta(0) \right] \Delta^2(0). \end{aligned} \quad (46)$$

Representing the right-hand side diagrammatically, we obtain the identity

$$\Sigma(18) = 3\delta(0) \text{Diagram 4} - \frac{2}{3} \omega^2 \text{Diagram 5}, \quad (47)$$

where $\Sigma(18)$ denotes the sum of all diagrams in Eq. (18). Using the identity (24) together with the field equation (27), we reduce the sum (19) of all one- and two-loop bubbles diagrams to terms involving $\delta(0)$ and $\delta^2(0)$:

$$\begin{aligned} & -\frac{1}{2!} \left\{ 2\delta^2(0) \int d\tau \Delta^2(\tau) \right. \\ & \quad \left. - 4\delta(0) \int d\tau \left[\Delta(0) \dot{\Delta}^2(\tau) - \ddot{\Delta}(0) \Delta^2(\tau) + 2\omega^2 \Delta(0) \Delta^2(\tau) \right] \right\} \\ & = 2\Delta^2(0) \delta(0) + \delta^2(0) \int d\tau \Delta^2(\tau). \end{aligned} \quad (48)$$

Hence we find the diagrammatic identity

$$-\frac{1}{2!} \Sigma(19) = 2\delta(0) \text{Diagram 6} + \delta^2(0) \text{Diagram 7}. \quad (49)$$

Now, the terms accompanying $\delta^2(0)$ turn out to be cancelled by similar terms coming from the sum of all three-loop bubbles diagrams in (20). In fact, the identities (24) and (32) lead to

$$\begin{aligned}
& -\frac{1}{2!} \int d\tau \left[-4\Delta(0)\ddot{\Delta}(0)\dot{\Delta}^2(\tau) + 2\Delta^2(0)\ddot{\Delta}^2(\tau) \right. \\
& \quad + 2\ddot{\Delta}^2(0)\Delta^2(\tau) + 8\omega^2\Delta^2(0)\dot{\Delta}^2(\tau) \\
& \quad \left. - 8\omega^2\Delta(0)\ddot{\Delta}(0)\Delta^2(\tau) + 8\omega^4\Delta^2(0)\Delta^2(\tau) \right] \\
& = - \left[\int d\tau \delta^2(\tau) + 2\delta(0) \right] \Delta^2(0) - \delta^2(0) \int d\tau \Delta^2(\tau). \quad (50)
\end{aligned}$$

Thus, using the rule (42), we find the diagrammatical sum for all bubbles diagrams:

$$-\frac{1}{2!} \Sigma(19) - \frac{1}{2!} \Sigma(20) = -\delta(0) \bigcirc \bigcirc . \quad (51)$$

Finally, the relations (25), (26), (33), (34), and (35) reduce the sum (21) of all watermelon-like diagrams to a finite contribution plus integrals involving $\delta^2(\tau)$ and $\epsilon^2(\tau)\delta(\tau)$:

$$\begin{aligned}
& -\frac{4}{2!} \int d\tau \left[\Delta^2(\tau)\ddot{\Delta}^2(\tau) + 4\Delta(\tau)\dot{\Delta}^2(\tau)\ddot{\Delta}(\tau) \right. \\
& \quad \left. + \dot{\Delta}^4(\tau) + 4\omega^2\Delta^2(\tau)\dot{\Delta}^2(\tau) + \frac{2}{3}\omega^4\Delta^4(\tau) \right] \\
& = -2 \int d\tau \Delta^2(\tau)\delta^2(\tau) + \frac{1}{\omega} \int d\tau \epsilon^2(\tau)\delta(\tau) - \frac{4}{3}\omega^2\Delta^3(0). \quad (52)
\end{aligned}$$

Combining these with all local diagrams (47), we easily verify that all finite contributions cancel each other, thus leading, with the help of Eq. (43), to the diagrammatic identity

$$\Sigma(18) - \frac{4}{2!} \Sigma(21) = \delta(0) \bigcirc \bigcirc . \quad (53)$$

The remaining singular terms in Eqs. (51) and (53) are summing up to zero, as required by the coordinate invariance of perturbatively defined path integrals.

The procedure can easily be continued to higher-loop diagrams to obtain integrals over any desired products of singular correlation functions, and over products of δ -functions. At no place we have to specify the value of $\delta(0)$ and the regularization scheme.

6 Summary

In this note, using the simple rules for relating singular to regular Feynman integrals found in Ref. [3], we discuss the reparametrization invariance of the perturbatively defined path integral for a particle in one-dimensional target space by performing an explicit three-loop calculation of the ground-state energy. The unique rules supply both kinds of cancellation: the loop divergences as well as the finite contributions, thus providing the UV-finiteness simultaneously with the coordinate independence of the quantum mechanical path integrals. This approach has an important advantage allowing us to avoid the explicit calculation of dimensionally regularized integrals over products of distributions. The procedure is *independent* of regularization prescriptions, using only the fact that regularized integrals can be integrated by parts. The results are, of course, perfectly compatible with those derived before in Refs. [1,2] by dimensional regularization.

Considerations of this paper are more general than those of Ref. [5] where two model Hamiltonians related to some transformation were found. It was shown that they lead to the same ground-state energy.

To prove the coordinate independence of perturbatively defined path integrals in one-dimensional target space there is no need for extra compensating noncovariant terms found necessary in the treatment of phase-space path integrals in Refs. [6–8]. A perturbative definition of the phase-space path integral was recently developed in Ref. [9].

Acknowledgments

I wish to thank Professor Hagen Kleinert for many interesting and stimulating discussions, as well as for his invaluable collaboration at various stages of this work.

I am grateful for financial support from the German university support program HSP III-Potsdam.

References

- [1] H. Kleinert and A. Chervyakov, *Phys. Lett. B* **464**, 257 (1999), eprint: hep-th/9906156.
- [2] H. Kleinert and A. Chervyakov, *Phys. Lett. B* **477**, 373 (2000), eprint: quant-ph/9912056.

Coordinate Independence of Path Integrals in One-Dimensional Target Space 55

- [3] H. Kleinert and A. Chervyakov, eprint: quant-ph/0002067.
- [4] G. 't Hooft and M. Veltman, *Nucl. Phys. B* **44**, 189 (1972).
- [5] T.D. Lee and C.N. Yang, *Phys. Rev.* **128**, 885 (1962).
- [6] J.L. Gervais and A. Jevicki, *Nucl. Phys. B* **459**, 631 (1996).
- [7] J. de Boer, B. Peeters, K. Skenderis, and P. van Nieuwenhuizen, *Nucl. Phys. B* **459**, 631 (1996), Appendix A.
- [8] P. Salomonson, *Nucl. Phys. B* **121**, 433 (1977).
- [9] M. Bachmann, *Perturbatively Defined Path Integral in Phase Space*, this volume, p. 57.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

PERTURBATIVELY DEFINED PATH INTEGRAL IN PHASE SPACE

M. BACHMANN

*Freie Universität Berlin, Institut für Theoretische Physik,
Arnimallee 14, 14195 Berlin, Germany
E-mail: mbach@physik.fu-berlin.de*

Introducing a perturbative definition, phase space path integrals can be calculated without slicing. This leads to a short-time expansion of the quantum mechanical path amplitude or a high-temperature expansion of the unnormalized density matrix. We use the proposed formalism to calculate the effective classical Hamiltonian for the harmonic oscillator.

1 Introduction

Path integrals are usually evaluated by time-slicing [1], since the continuum definition is mathematically problematic. This becomes obvious for physical systems with nontrivial metric where reparametrization invariance has been a problem for many years [2]. It was solved recently by Kleinert and Chervyakov who defined path integrals perturbatively in *configuration space* [3] using dimensional regularization methods developed in the quantum theory of non-Abelian gauge fields [4]. They found rules for calculating integrals over products of distributions which establish a unique procedure for a perturbative evaluation of path integrals which fully respects parameterization invariance. The path integral of any system is expanded around the exactly known solution for the free particle in powers of the coupling constant of the potential.

In this article we want to present an extension of this definition to *phase space* and derive a short-time expansion of the Hamiltonian quantum mechanical time evolution amplitude. In Euclidean space, the density matrix is obtained as a high-temperature expansion. By simple resummation, this

series can be turned into an expansion in powers of the coupling constant of the potential described above. As will be shown in the sequel, the knowledge of an exactly known nontrivial path integral as that of a free particle is *not* required. Thus, the perturbative definition presented here is completely general. It reproduces the expansion around the free particle by a simple resummation.

The method is then applied to calculate the effective classical Hamiltonian of the harmonic oscillator $H_{\omega,\text{eff}}(p_0, x_0)$ by exactly summing up the perturbation series. The quantity is related to the quantum statistical partition function via the classically looking phase space integral

$$Z_\omega = \int \frac{dx_0 dp_0}{2\pi\hbar} \exp \{-\beta H_{\omega,\text{eff}}(p_0, x_0)\}, \quad (1)$$

where $\beta = 1/k_B T$ is the inverse thermal energy.

2 Perturbative Definition of the Path Integral for the Density Matrix

After slicing the interval $[0, \hbar\beta]$ into $N+1$ pieces of width $\varepsilon = \hbar\beta/(N+1)$, the unnormalized density matrix can be expressed in the continuum limit as [1]

$$\begin{aligned} \tilde{\rho}(x_b, x_a) = & \lim_{N \rightarrow \infty} \prod_{n=1}^N \left[\int_{-\infty}^{\infty} dx_n \right] \prod_{n=1}^{N+1} \left[\int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{ip_n(x_n - x_{n-1})/\hbar} \right] \\ & \times \exp \left\{ -\varepsilon \sum_{n=1}^{N+1} H(p_n, x_n)/\hbar \right\}, \end{aligned} \quad (2)$$

where $x_a = x_0$ and $x_b = x_{N+1}$ are the fixed end points of the path. Upon expanding the last exponential in powers of $\varepsilon/\hbar$, we recognize that the zeroth-order contribution to the density matrix (2) is an *infinite product of δ -functions* due to the identity

$$\int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} e^{ip_n(x_n - x_{n-1})/\hbar} = \delta(x_n - x_{n-1}). \quad (3)$$

This infinite product simply reduces to

$$\lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} dx_N \cdots dx_1 \delta(x_{N+1} - x_N) \cdots \delta(x_2 - x_1) \delta(x_1 - x_0) = \delta(x_b - x_a), \quad (4)$$

which is the unperturbed contribution to the unnormalized density matrix (2) obtained here without solving a nontrivial path integral. Thus, the phase space path integral for the unnormalized density matrix (2) can be perturbatively defined as

$$\begin{aligned} \tilde{\varrho}(x_b, x_a) = & \delta(x_b - x_a) + \sum_{n=1}^{\infty} \frac{(-1)^n}{\hbar^n n!} \int_0^{\hbar\beta} d\tau_1 \cdots \int_0^{\hbar\beta} d\tau_n \\ & \times \langle H(p(\tau_1), x(\tau_1)) \cdots H(p(\tau_n), x(\tau_n)) \rangle_0 \end{aligned} \quad (5)$$

with expectation values

$$\langle \cdots \rangle_0 = \lim_{N \rightarrow \infty} \prod_{n=1}^N \left[\int_{-\infty}^{\infty} dx_n \right] \prod_{n=1}^{N+1} \left[\int_{-\infty}^{\infty} \frac{dp_n}{2\pi\hbar} \cdots e^{ip_n(x_n - x_{n-1})/\hbar} \right]. \quad (6)$$

These expectation values may be reexpressed by Feynman diagrams. This is possible for polynomial as well as nonpolynomial functions of momentum and position [5].

3 Restricted Partition Function and Two-Point Correlations

The trace over the unnormalized density matrix (5) of our unperturbed system with vanishing Hamiltonian $H(p, x) = 0$ leads to a diverging partition function. In addition, the classical partition function diverges with the phase space volume. The regularization of these divergences is possible by excluding, from the phase space path integral, the zero frequency fluctuations x_0 and p_0 of the Fourier decomposition of the periodic path $x(\tau)$ and momentum $p(\tau)$, respectively [1,6]. At the end, we shall calculate the quantum statistical partition function of any system from the classical phase space integral

$$Z = \int \frac{dx_0 dp_0}{2\pi\hbar} Z^{p_0 x_0}. \quad (7)$$

The restricted partition function in the integrand contains the Boltzmann factor of the effective classical Hamiltonian defined by the path integral:

$$\begin{aligned} Z^{p_0 x_0} \equiv & \exp \{ -\beta H_{\text{eff}}(p_0, x_0) \} = 2\pi\hbar \oint \mathcal{D}x \mathcal{D}p \delta(x_0 - \bar{x}) \delta(p_0 - \bar{p}) \\ & \times \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau \left[-i(p(\tau) - p_0) \frac{\partial}{\partial \tau} (x(\tau) - x_0) + H(p(\tau), x(\tau)) \right] \right\} \end{aligned} \quad (8)$$

with the measure

$$\oint \mathcal{D}x \mathcal{D}p = \lim_{N \rightarrow \infty} \prod_{n=1}^{N+1} \left[\int_{-\infty}^{\infty} \frac{dx_n dp_n}{2\pi\hbar} \right]. \quad (9)$$

The quantities $\bar{x}$ and $\bar{p}$ are the temporal mean values $\bar{x} = \int_0^{\hbar\beta} d\tau x(\tau)/\hbar\beta$ and $\bar{p} = \int_0^{\hbar\beta} d\tau p(\tau)/\hbar\beta$.

As illustrated above, the unperturbed system may have a vanishing Hamiltonian H . The calculation of the restricted partition function for this system to be denoted by $Z_0^{p_0 x_0}$, is *trivial* as for its density matrix in Eq. (4). A cancellation of δ -functions yields directly $Z_0^{p_0 x_0} = 1$.

Let us now concentrate on the correlation functions of position- and momentum-dependent quantities. For this purpose it is convenient to introduce the generating functional

$$\begin{aligned} Z_0^{p_0 x_0}[j, v] &= 2\pi\hbar \oint \mathcal{D}x \mathcal{D}p \delta(x_0 - \bar{x}) \delta(p_0 - \bar{p}) \\ &\times \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau \left[-i\tilde{p}(\tau) \frac{\partial}{\partial \tau} \tilde{x}(\tau) + j(\tau) \tilde{x}(\tau) + v(\tau) \tilde{p}(\tau) \right] \right\}, \end{aligned} \quad (10)$$

with abbreviations $\tilde{x}(\tau) = x(\tau) - x_0$ and $\tilde{p}(\tau) = p(\tau) - p_0$. The action in the exponent contains only the trivial Euclidean eikonal $S = -i \int d\tau \tilde{p} \partial \tilde{x} / \partial \tau$. The calculation yields

$$Z_0^{p_0 x_0}[j, v] = \exp \left\{ \frac{1}{\hbar^2} \int_0^{\hbar\beta} d\tau \int_0^{\hbar\beta} d\tau' j(\tau) G^{p_0 x_0}(\tau, \tau') v(\tau') \right\}, \quad (11)$$

where the periodic Green function is

$$\begin{aligned} G^{p_0 x_0}(\tau, \tau') &= -\frac{i}{2\beta} \{2(\tau - \tau') - \hbar\beta [\Theta(\tau - \tau') - \Theta(\tau' - \tau)]\} \\ &= \frac{2i}{\beta} \sum_{m=1}^{\infty} \frac{\sin \omega_m(\tau - \tau')}{\omega_m}. \end{aligned} \quad (12)$$

In the last line the Fourier decomposition is given with respect to Matsubara frequencies $\omega_m = 2\pi m/\hbar\beta$, omitting the zero mode. Observe the antisymmetry $G^{p_0 x_0}(\tau, \tau') = -G^{p_0 x_0}(\tau', \tau)$.

These Green functions possess an interesting scaling property: Substituting $\bar{\tau} \equiv \tau/\beta$, they become *independent* of β :

$$G^{p_0 x_0}(\bar{\tau}, \bar{\tau}') = -\frac{i}{2} \{2(\bar{\tau} - \bar{\tau}') - \hbar [\Theta(\bar{\tau} - \bar{\tau}') - \Theta(\bar{\tau}' - \bar{\tau})]\}. \quad (13)$$

Introducing expectation values as

$$\langle \cdots \rangle_0^{p_0 x_0} = 2\pi\hbar \oint \mathcal{D}x \mathcal{D}p \delta(x_0 - \bar{x}) \delta(p_0 - \bar{p}) \cdots \exp \left\{ \frac{1}{\hbar} \int_0^{\hbar\beta} d\tau i\tilde{p}(\tau) \frac{\partial}{\partial \tau} \tilde{x}(\tau) \right\}, \quad (14)$$

the two-point functions are obtained from the functional (10) by performing appropriate functional derivatives with respect to $j(\tau)$ and $v(\tau)$, respectively:

$$\langle \tilde{x}(\tau) \tilde{x}(\tau') \rangle_0^{p_0 x_0} = 0, \quad (15)$$

$$\langle \tilde{x}(\tau) \tilde{p}(\tau') \rangle_0^{p_0 x_0} = G^{p_0 x_0}(\tau, \tau'), \quad (16)$$

$$\langle \tilde{p}(\tau) \tilde{p}(\tau') \rangle_0^{p_0 x_0} = 0. \quad (17)$$

The off-diagonal nature of the trivial action in (14) entails that only *mixed* position-momentum correlations do not vanish.

4 Perturbative Expansion for Effective Classical Hamiltonian

Expanding the restricted partition function (8) in powers of the Hamiltonian,

$$Z^{p_0 x_0} = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{\hbar^n n!} \int_0^{\hbar\beta} d\tau_1 \cdots \int_0^{\hbar\beta} d\tau_n \times \langle H(p(\tau_1), x(\tau_1)) \cdots H(p(\tau_n), x(\tau_n)) \rangle_0^{p_0 x_0}, \quad (18)$$

rewriting this into a cumulant expansion, and utilizing the relation (8) between restricted partition function and effective classical Hamiltonian, we obtain

$$H_{\text{eff}}(p_0, x_0) = \frac{1}{\beta} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\hbar^n n!} \int_0^{\hbar\beta} d\tau_1 \cdots \int_0^{\hbar\beta} d\tau_n \times \langle H(p(\tau_1), x(\tau_1)) \cdots H(p(\tau_n), x(\tau_n)) \rangle_{0,c}^{p_0 x_0}. \quad (19)$$

Using Wick's rule, all correlation functions can be expressed in terms of products of two-point functions. Since only mixed two-point functions (12) can lead to nonvanishing contributions to the effective classical Hamiltonian, we

use the rescaled version (13) of the Green function. The scaling transformation gives a factor β from each of the n integral measures. Thus the expansion (19) is a *high-temperature* expansion of the effective classical Hamiltonian:

$$H_{\text{eff}}(p_0, x_0) = \sum_{n=1}^{\infty} \beta^{n-1} \frac{(-1)^{n+1}}{\hbar^n n!} \int_0^{\hbar} d\bar{\tau}_1 \cdots \int_0^{\hbar} d\bar{\tau}_n \times \langle H(p(\bar{\tau}_1), x(\bar{\tau}_1)) \cdots H(p(\bar{\tau}_n), x(\bar{\tau}_n)) \rangle_{0,c}^{p_0 x_0}. \quad (20)$$

For the following considerations it is useful to assume the Hamilton function to be of standard form:

$$H(p(\bar{\tau}), x(\bar{\tau})) = p^2(\bar{\tau})/2M + gV(x(\bar{\tau})). \quad (21)$$

We have introduced the coupling constant g of the potential. Defining the functionals

$$a[p] = \int_0^{\hbar} d\bar{\tau} p^2(\bar{\tau})/2M, \quad b[x] = \int_0^{\hbar} d\bar{\tau} V(x(\bar{\tau})), \quad (22)$$

the high-temperature expansion (20) is expressed as

$$H_{\text{eff}}(p_0, x_0) = \sum_{n=1}^{\infty} \beta^{n-1} \frac{(-1)^{n+1}}{n! \hbar^n} \sum_{k=0}^n g^k \binom{n}{k} \langle a^{n-k}[p] b^k[x] \rangle_{0,c}^{p_0 x_0}. \quad (23)$$

In the sequel we point out how this high-temperature expansion is connected with an expansion in powers of the coupling constant g of the potential.

5 High-Temperature Versus Weak-Coupling Expansion

Having shown that the perturbative expansion around a vanishing Hamiltonian leads to a perturbative series in powers of the inverse temperature in a natural manner, we now elaborate on its relation to more customary perturbative expansions in powers of the coupling constant g of the potential. Changing the order of summation in Eq. (23), we obtain

$$H_{\text{eff}}(p_0, x_0) = \sum_{k=0}^{\infty} g^k \sum_{n=0}^{\infty} \beta^{n+k-1} \binom{n+k}{k} \frac{(-1)^{n+k+1}}{(n+k)! \hbar^{n+k}} \langle a^n[p] b^k[x] \rangle_{0,c}^{p_0 x_0} + \frac{1}{\beta}, \quad (24)$$

which is rewritten after explicitly evaluating the contributions for $n = 0$ and $k = 0$:

$$\begin{aligned}
 H_{\text{eff}}(p_0, x_0) &= \frac{p_0^2}{2M} + gV(x_0) + \frac{1}{\beta} \sum_{k=1}^{\infty} g^k \sum_{n=1}^{\infty} \frac{(-1)^{n+k+1}}{n!k!\hbar^{n+k}(2M)^n} \\
 &\quad \times \int_0^{\hbar\beta} d\tau_1 \cdots \int_0^{\hbar\beta} d\tau_k \int_0^{\hbar\beta} d\tau_{k+1} \cdots \int_0^{\hbar\beta} d\tau_{k+n} \\
 &\quad \times \langle V(x(\tau_1)) \cdots V(x(\tau_k)) p^2(\tau_{k+1}) \cdots p^2(\tau_{k+n}) \rangle_{0,c}^{p_0, x_0}. \quad (25)
 \end{aligned}$$

In this expression, we have inverted the scaling transformation in (13) and used the expectation values

$$\int_0^{\hbar\beta} d\tau \langle p^2(\tau) \rangle_{0,c}^{p_0, x_0} = \hbar\beta p_0^2, \quad \int_0^{\hbar\beta} d\tau \langle V(x(\tau)) \rangle_{0,c}^{p_0, x_0} = \hbar\beta V(x_0). \quad (26)$$

All higher-order expectations of functions which only depend on x or p are zero, due to the vanishing of expectations of functions of $\tilde{x}$ or $\tilde{p}$ in a Wick expansion into products of two-point functions (15) and (17). All other possible contributions are disconnected.

We now observe that the expansion (25) is equal to a perturbation expansion around a free-particle theory

$$\begin{aligned}
 H_{\text{eff}}(p_0, x_0) &= \frac{p_0^2}{2M} + gV(x_0) + \frac{1}{\beta} \sum_{k=1}^{\infty} g^k \frac{(-1)^{k+1}}{k!\hbar^k} \int_0^{\hbar\beta} d\tau_1 \cdots \int_0^{\hbar\beta} d\tau_k \\
 &\quad \times \langle V(x(\tau_1)) \cdots V(x(\tau_k)) \rangle_{\text{free}, c}^{x_0}, \quad (27)
 \end{aligned}$$

in which cumulants are formed from expectation values

$$\begin{aligned}
 \langle \cdots \rangle_{\text{free}}^{x_0} &= 2\pi\hbar \oint \mathcal{D}x \mathcal{D}p \delta(x_0 - \bar{x}) \delta(p_0 - \bar{p}) \cdots \\
 &\quad \times \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau \left[-i(p(\tau) - p_0) \frac{\partial}{\partial \tau} (x(\tau) - x_0) + \frac{1}{2M} (p(\tau) - p_0)^2 \right] \right\}. \quad (28)
 \end{aligned}$$

6 Effective Classical Hamiltonian of the Harmonic Oscillator

We now calculate the effective classical Hamiltonian for the harmonic oscillator,

$$H_\omega(p, x) = \frac{p^2}{2M} + \frac{1}{2}M\omega^2 x^2, \quad (29)$$

by an exact resummation of the high-temperature expansion (23). For systematically expressing the terms of this expansion, it is useful to introduce the following Feynman rules:

$$\bar{\tau}_1 \text{ --- } \bar{\tau}_2 \equiv \langle p(\bar{\tau}_1)p(\bar{\tau}_2) \rangle_0^{p_0 x_0} = p_0^2, \quad (30)$$

$$\bar{\tau}_1 \text{ --- } \bar{\tau}_2 \equiv \langle x(\bar{\tau}_1)x(\bar{\tau}_2) \rangle_0^{p_0 x_0} = x_0^2, \quad (31)$$

$$\bar{\tau}_1 \text{ --- } \bar{\tau}_2 \equiv \langle x(\bar{\tau}_1)p(\bar{\tau}_2) \rangle_0^{p_0 x_0} = G^{p_0 x_0}(\bar{\tau}_1, \bar{\tau}_2) + x_0 p_0, \quad (32)$$

$$\bar{\tau}_1 \text{ --- } \bar{\tau}_2 \equiv \langle p(\bar{\tau}_1)x(\bar{\tau}_2) \rangle_0^{p_0 x_0} = -G^{p_0 x_0}(\bar{\tau}_1, \bar{\tau}_2) + p_0 x_0, \quad (33)$$

$$\bar{\tau} \text{ --- } \star \equiv \langle p(\bar{\tau}) \rangle_0^{p_0 x_0} = p_0, \quad (34)$$

$$\bar{\tau} \text{ --- } \star \equiv \langle x(\bar{\tau}) \rangle_0^{p_0 x_0} = x_0, \quad (35)$$

$$\bullet \equiv \int_0^{\hbar} d\bar{\tau}, \quad (36)$$

where the current-like expectations in (34) and (35) arise from $\langle \tilde{p}(\bar{\tau}) \rangle_0^{p_0 x_0} = 0$ and $\langle \tilde{x}(\bar{\tau}) \rangle_0^{p_0 x_0} = 0$, respectively. In order to simplify the calculation of the expectation values in the high-temperature expansion of the effective classical Hamiltonian (20), we also define operational subgraphs

$$\text{---}\bullet\text{---} \equiv \frac{1}{2M\hbar} \int_0^{\hbar} d\bar{\tau} p^2(\bar{\tau}), \quad (37)$$

$$\text{---}\bullet\text{---} \equiv \frac{1}{2\hbar} M\omega^2 \int_0^{\hbar} d\bar{\tau} x^2(\bar{\tau}), \quad (38)$$

which are useful for the systematic construction of Feynman diagrams. These diagrams are composed by attaching the legs of such subgraphs to one another or by connecting legs with suitable currents. Note that only combinations of different types of subgraphs lead to non-vanishing contributions, since the connection of subgraphs of the same type,

$$\text{---}\bullet\text{---}\bullet\text{---}, \quad \text{---}\bullet\text{---}\bullet\text{---}, \quad (39)$$

creates a new subgraph which contains a propagator (30) or (31), respectively. These propagators are, however, independent of τ , such that the τ -integrals related to the vertices in these subgraphs are trivial. Thus, there does not really exist a connection between these vertices, and the propagators (30) and (31) can be expressed by the currents (34) and (35):

$$\overline{\tau}_1 \text{ --- } \overline{\tau}_2 = \overline{\tau}_1 \text{ --- } \star \text{ --- } \overline{\tau}_2, \quad (40)$$

$$\overline{\tau}_1 \text{ --- } \overline{\tau}_2 = \overline{\tau}_1 \text{ --- } \star \text{ --- } \overline{\tau}_2. \quad (41)$$

As a consequence, connected diagrams for $n > 1$ containing propagators of type (30) or (31) *must* break up into disconnected parts. Analytically, this is seen by considering for example

$$\langle x(\overline{\tau}_1)x(\overline{\tau}_2) \rangle_0^{p_0 x_0} = \langle \tilde{x}(\overline{\tau}_1)\tilde{x}(\overline{\tau}_2) \rangle_0^{p_0 x_0} + \langle x(\overline{\tau}_1) \rangle_0^{p_0 x_0} \langle x(\overline{\tau}_2) \rangle_0^{p_0 x_0}. \quad (42)$$

The first term on the right-hand side vanishes due to Eq. (15), while the second simply yields x_0^2 , which proves Eq. (33). This means that only Feynman diagrams which consist of a mixture of subgraphs (37) and (38) contribute to the effective classical Hamiltonian. To illustrate this, we discuss the first and second order of expansion (20) in more detail.

The Feynman diagrams of the first-order contribution to the effective classical Hamiltonian are simply constructed from the subgraphs

$$\begin{aligned} H_{\omega, \text{eff}}^{(1)}(p_0, x_0) &\propto \text{---} \text{---} \text{---} + \text{---} \bullet \text{---} \\ &= \frac{1}{2M\hbar} \text{---} \text{---} \text{---} + \frac{1}{2\hbar} M\omega^2 \text{---} \bullet \text{---} = \frac{1}{2M\hbar} \text{---} \star \text{---} \text{---} \star + \frac{1}{2\hbar} M\omega^2 \text{---} \star \text{---} \bullet \text{---} \star \\ &= \frac{p_0^2}{2M} + \frac{1}{2} M\omega^2 x_0^2, \end{aligned} \quad (43)$$

where we have used the identities (40) and (41) in the second expression of the second line. Note that the first-order term (43) obviously reproduces the classical Hamiltonian. This is the consequence of the high-temperature expansion (23), since only the first-order contribution is nonzero in the limit $\beta = 1/k_B T \rightarrow 0$. The second-order contribution reads

$$\begin{aligned} H_{\omega, \text{eff}}^{(2)}(p_0, x_0) &\propto (\text{---} \text{---} \text{---} + \text{---} \bullet \text{---}) (\text{---} \text{---} \text{---} + \text{---} \bullet \text{---}) \\ &= -\frac{\omega^2}{8\hbar^2\beta} \left(8 \text{---} \star \text{---} \bullet \text{---} \star + 4 \text{---} \bullet \text{---} \bullet \text{---} \right). \end{aligned} \quad (44)$$

The chain diagram is zero, while the loop diagram has the value $-\hbar^4\zeta(2)/2\pi^2$,

where

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z} \quad (45)$$

is the ζ -function. Thus we obtain

$$H_{\omega,\text{eff}}^{(2)}(p_0, x_0) = \beta \hbar^2 \omega^2 \zeta(2)/4\pi^2.$$

This second-order contribution (44) shows the characteristic types of Feynman diagrams in each order $n > 1$ of the expansion (20) for the harmonic oscillator: chain and loop diagrams. In order to calculate the n th-order contribution, we must evaluate these diagrams in a more general way. By constructing Feynman diagrams from the product of n sums of subgraphs,

$$H_{\omega, \text{eff}}^{(n)}(p_0, x_0) \propto \underbrace{(\text{---}\bullet\text{---}\text{---} + \text{---}\bullet\text{---}) (\text{---}\text{---}\bullet\text{---} + \text{---}\bullet\text{---}) \cdots (\text{---}\text{---}\bullet\text{---} + \text{---}\bullet\text{---})}_{n \text{ times}}, \quad (46)$$

it turns out that only the following chain and loop diagrams contribute:

$$\begin{array}{l}
\star \cdots \bullet \cdots \blacktriangle \cdots \bullet \cdots \star, \\
\star \cdots \blacktriangle \cdots \bullet \cdots \blacktriangle \cdots \bullet \cdots \star, \\
\star \cdots \blacktriangle \cdots \bullet \cdots \blacktriangle \cdots \bullet \cdots \star,
\end{array}
\quad
\begin{array}{c}
\text{Diagram of a circle with 12 vertices and directed edges forming a cycle.}
\end{array}
\quad (47)$$

The evaluation of the chain diagrams is easily done and yields zero. An explicit calculation in Fourier space shows that there occur Kronecker- δ 's $\delta_{m,0}$. Since the Matsubara sum of the Green function (12) does not contain the zero mode $m = 0$, all chain diagrams are zero.

A determination of the values of loop diagrams is more involved. It is obvious that loop diagrams can only be constructed in *even* order ($n = 2, 4, 6, \dots$), since for a loop diagram with mixed propagators (32) or (33) pairs of different subgraphs (37) and (38) are necessary. Thus we find the result that *odd* orders of expansion (23) *vanish*, and only loop diagrams for $n = 2, 4, 6, \dots$ must be calculated. Evaluating loop diagrams of n th order in Fourier space is straightforward and entails

$$= 2(-1)^k \left(\frac{\hbar^2}{2\pi} \right)^{2k} \zeta(2k), \quad (48)$$

where $k = n/2$. The high-temperature expansion for the effective classical Hamiltonian of the harmonic oscillator can then be written as

$$H_{\omega,\text{eff}}(p_0, x_0) = \frac{p_0^2}{2M} + \frac{1}{2}M\omega^2 x_0^2 + \sum_{k=1}^{\infty} \beta^{2k-1} \frac{(-1)^{k+1}}{k} \left(\frac{\hbar\omega}{2\pi}\right)^{2k} \zeta(2k). \quad (49)$$

Substituting the ζ -function by its definition (45) and exchanging the summations, the last term in Eq. (49) can be expressed as a logarithm

$$\sum_{k=1}^{\infty} \beta^{2k-1} \frac{(-1)^{k+1}}{k} \left(\frac{\hbar\omega}{2\pi}\right)^{2k} \zeta(2k) = \frac{1}{\beta} \ln \left(\prod_{n=1}^{\infty} \left[1 + \frac{\hbar^2 \beta^2 \omega^2}{4\pi^2 n^2} \right] \right). \quad (50)$$

Applying the relation

$$\frac{1}{z} \sinh z = \prod_{n=1}^{\infty} \left(1 + \frac{z^2}{n^2 \pi^2} \right), \quad (51)$$

we find the more familiar form of the effective classical Hamiltonian for a harmonic oscillator

$$H_{\omega,\text{eff}}(p_0, x_0) = \frac{p_0^2}{2M} + \frac{1}{2}M\omega^2 x_0^2 - \frac{1}{\beta} \ln \frac{\hbar\omega\beta}{2 \sinh(\hbar\omega\beta/2)}. \quad (52)$$

When performing the x_0 - and p_0 -integrations in Eq. (1), we obtain the well-known form of the partition function of the harmonic oscillator $Z_{\omega} = 1/2 \sinh(\hbar\omega\beta/2)$.

7 Summary

We have used a perturbative definition of the path integral in phase space representation which produces an effective classical Hamiltonian for the harmonic oscillator. Our procedure represents an alternative way to evaluate path integrals: The unperturbed system is trivial and the calculation of appropriate Feynman diagrams is simple. As a further advantage, the perturbative expansion for the effective classical Hamiltonian is identical to the high-temperature expansion used frequently in statistical mechanics.

Acknowledgments

I deeply thank Professor H. Kleinert for many exciting discussions, useful hints, and expert advice in the course of my Ph.D. research. I am also in-

debted to Dr. A. Pelster and Dr. A. Chervyakov for the exchange of experiences with respect to the perturbatively defined path integral. Finally, I am grateful to the Studienstiftung des deutschen Volkes for financial support.

References

- [1] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [2] J. de Boer, B. Peeters, K. Skenderis, and P. van Nieuwenhuizen, *Nucl. Phys. B* **446**, 211 (1995); **459**, 631 (1996); F. Bastianielli, K. Schalm, and P. van Nieuwenhuizen, *Phys. Rev. D* **58**, 44002 (1998).
- [3] H. Kleinert and A. Chervyakov, *Phys. Lett. A* **273**, 1 (2000); *Phys. Lett. B* **464**, 257 (1999); **477**, 373 (2000); eprint: quant-ph/000206. See also: A. Chervyakov, *Coordinate Independence of Path Integrals in One-Dimensional Target Space*, this volume, p. 43.
- [4] G. 't Hooft and M. Veltman, *Nucl. Phys. B* **44**, 189 (1972).
- [5] H. Kleinert, A. Pelster, and M. Bachmann, *Phys. Rev. E* **60**, 2510 (1999).
- [6] M. Bachmann, H. Kleinert, and A. Pelster, *Phys. Rev. A* **62**, 52509 (2000); *Phys. Lett. A* **279**, 23 (2001).

THE USE OF NON-RELATIVISTIC PATH INTEGRALS IN FIELD THEORIES

I.H. DURU

Trakya University, Mathematics Department, P.O. Box 126, Edirne, Turkey;
Feza Gursey Institute, P.O. Box 6, 81220, Cengelkoy, Istanbul, Turkey
E-mail: duru@gursey.gov.tr

We list a few examples on the application of non-relativistic path integrals to field theoretical problems.

In 1979, I had the fortune to collaborate with Hagen Kleinert. It was the time, when we solved the path integral for the hydrogen atom [1]. Before that, only exact solutions of the path integrals for the free particle and systems with quadratic [2] and inverse quadratic [3] potentials were known. After we found the procedure to calculate the path integral for the H -atom [1], however, almost all of the exactly solvable quantum mechanics text book problems were also solved by path integrals [4]. The basic methods are derived from the H -atom solution: Point canonical transformations, time parameter changes and introduction of auxiliary dynamics in additional space dimensions, reflections from the boundaries. Only the employment of the underlying symmetry group spaces [5] does not have its roots in the techniques used in H -atom path integral.

Solving the quantum mechanical problems by path integrals which have already been solved by the conventional Schrödinger method has its own interest. However apart from its obvious beauty, the path-integration method is suitable for discovering the connections between seemingly different potentials in the Schrödinger picture. It also gives the wave function normalization quite automatically.

The purpose of this note is to mention the use of the exact path-integral solutions for the non-relativistic quantum mechanical point particle motions in some relativistic field theoretical problems.

Particle pair creations in given classical environments are good examples. As a specific case let us consider the Robertson-Walker cosmologies with the metric

$$ds^2 = -dt^2 + a^2(t)dx^2, \quad (1)$$

where the expansion factor $a(t)$ is a given function of time.

The amplitude A_{ij} for a particle created by the gravitational field at time t in states $f_i(x)$ and $f_j(x)$ is given by

$$A_{ij} = -A_0 \int d^3k_a d^3k_b \int d\sigma^\mu(x_a) d\sigma^\nu(x_b) \times f_i^*(\vec{k}_a, x_a) \overleftrightarrow{\partial}_\mu^a f_j^*(\vec{k}_b, x_b) \overleftrightarrow{\partial}_\nu^b K(x_b, x_a). \quad (2)$$

Here A_0 is the no-particle production amplitude, and $d\sigma^\mu$ is the element of the constant t hypersurface. $K(x_b, x_a)$ is the propagator for a scalar particle of mass μ to go from the space-time point x_b to x_a .

Developing the propagator $K(x_b, x_a)$ by path integrals is important, for it avoids the determination of the initial particle states which is problematic because of the cosmological singularity [6].

The propagator $K(x_b, x_a)$ is given by [1]

$$K(x_b, x_a) = \int_0^\infty dW e^{-i\mu^2 W} F(W, x_b, x_a), \quad (3)$$

where W is the total parameter time. $F(W, x_b, x_a)$ is the Green function which can be expressed as a Hamiltonian path integral as

$$F(W, x_b, x_a) = \int \mathcal{D}^4x \mathcal{D}^4k \exp \left\{ i \int_0^W d\omega \left[k_t \dot{t} + \vec{k} \cdot \dot{\vec{x}} + k_t^2 - \frac{\vec{k}^2}{a^2(t)} \right] \right\}, \quad (4)$$

which is formally the same as the non-relativistic particle motion taking place in "space" $(t, \vec{x})$ in the "time" ω . After performing the trivial space integrations over $\mathcal{D}^3x \mathcal{D}^3k$ one arrives at

$$F(W, x_b, x_a) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k} \cdot (\vec{x}_a - \vec{x}_b)} \int \mathcal{D}t \mathcal{D}k_t \times \exp \left\{ i \int_0^W d\omega \left[k_t \dot{t} + k_t^2 - \frac{\vec{k}^2}{a^2(t)} \right] \right\}. \quad (5)$$

The path integral in the above expression is equivalent to the one written for the "mass" $= -1/2$, taking place in the $(1+1)$ -dimensional "space-time" (t, ω) ,

under the influence of the potential $\vec{k}^2/a^2(t)$, where $\vec{k}^2$ is constant through the path integration. Thus for example for the inflationary Universe with $a(t) = e^{-Ht}$, the problem at hand is equivalent to the exponential potential whose exact solution is known from the Morse potential. For the radiation-dominated Universe with $a(t) \sim \sqrt{t}$ on the other hand, the potential is of the one-dimensional Kepler type that is again exactly solvable.

The second type of field theoretical example I would like to mention briefly is the decay of the vacuum in an external magnetic field. In this case the propagator one needs is expressible as the following type of phase-space path integral

$$G(x_b, x_a) = \int_0^\infty dW e^{-i\mu^2 W} \int \mathcal{D}^4 x \mathcal{D}^4 k \exp \left[i \int_0^W d\omega (k_t \dot{t} + \vec{k} \cdot \dot{\vec{x}} + \vec{k}_t^2 - \vec{k}^2 + eA^\nu \dot{x}_\nu) \right], \quad (6)$$

with A^ν being the external electromagnetic potential. The path integral in the above Green function is again formally the same as the one written for a non-relativistic particle motion taking place in $(3+1)$ -dimensional "space" and in "time" ω .

For specific cases one gets even simpler forms: For example for constant electric field $\vec{E} = E\hat{x}$ in time dependent gauge, i.e. with $A_\mu = (0, Et, 0, 0)$, one ends up formally with a $(1+1)$ -dimensional quadratic potential [7]. For the magnetic monopole field the path integral one has in hand is of Pöschl-Teller type in the angular coordinates and the inverse square potential in radial coordinates [8].

References

- [1] I.H. Duru and H. Kleinert, *Phys. Lett. B* **84**, 185 (1979); *Fortschr. Phys.* **30**, 401 (1982).
- [2] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (Mc Graw-Hill, New York, 1965).
- [3] D. Peak and A. Inomata, *J. Math. Phys.* **10**, 1422 (1969).
- [4] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [5] See for example: I.H. Duru, *Phys. Rev. D* **30**, 212 (1984); M.S. Marinov and M.V. Terentyev, *Fortschr. Phys.* **27**, 511 (1979); J.S. Dowker, *Ann.*

Phys. (N.Y.) **62**, 361 (1971).

[6] D.M. Chitre and J.B. Hartle, *Phys. Rev. D* **16**, 251 (1977).

[7] A.O. Barut and I.H. Duru, *Phys. Rev. D* **41**, 1312 (1990).

[8] I.H. Duru, *J. Phys. A: Math. Gen.* **28**, 5883 (1995).

PATH INTEGRATION AND COHERENT STATES FOR THE 5D HYDROGEN ATOM

N. ÜNAL

Akdeniz University, Physics Department, P.K. 510 Antalya, 07058, Turkey

E-mail: nunal@pascal.sci.akdeniz.edu.tr

We transform the five-dimensional hydrogen atom into eight oscillators in a parametric time by adding three new degrees of freedom and by using the Hurwitz transformation. The path integration is then performed over the holomorphic coordinates of the oscillators. The end-point integrations over the extra three degrees of freedom give the coherent states of the five-dimensional hydrogen atom.

1 Introduction

In 1926, Schrödinger constructed the coherent states for the harmonic oscillator [1] and later some authors searched to obtain the coherent states for the hydrogen atom [2-7].

In 1979, Duru and Kleinert performed the path integration for the hydrogen atom [8]. They made use of the $SO(4)$ dynamical symmetry of the system and transformed the three-dimensional Kepler problem into four harmonic oscillators in a new parametric time by using the Kustaanheimo-Stiefel transformation [9].

Recently, we discussed the path integral for the hydrogen atom in holomorphic coordinates a_i and $a_i^\dagger$ instead of the physical coordinates $\vec{x}$ and $\vec{p}$ [10]. Since the coherent states are the eigenstates of the lowering operator a_i , we showed that the kernel between the holomorphic coordinates governs the time evolution of the coherent states. We also discussed the time evolution of the expectation values of $\langle \vec{x} \rangle$ and showed that they satisfy the Kepler laws and that the dispersions also oscillate with the frequency of the elliptic orbit.

The transformation $x_1 + ix_2 = (u_1 + iu_2)^2$ maps the two-dimensional Kepler problem into two harmonic oscillators [11]. The Kustaanheimo-Stiefel transformation maps the three-dimensional Kepler problem into four oscillators and this transformation has been used to derive the path-integral solution of the Kepler problem. Cornish discussed the complex form of the Kustaanheimo-Stiefel transformation and showed that it can be written as $x = u^2$ by using quaternions, without using the nonholonomic constraints for the Kepler problem [12].

The aim of this study is to generalize the path-integral method over the holomorphic coordinates of the 5-dimensional hydrogen atom. In order to generalize the path integration methods for the three-dimensional hydrogen atom we will use the Hurwitz transformation which maps the 5-dimensional Kepler problem into 8 harmonic oscillators [13]. In Section 2 we discuss the mapping between the 5-dimensional hydrogen atom and 8 harmonic oscillators and derive the Lagrangian of the system in holomorphic coordinates. In Section 3 we solve the path integral over the holomorphic coordinates, and derive the coherent states of the five-dimensional hydrogen atom in spherical coordinates.

2 Classical Kepler Problem in 5D

The action of the 5-dimensional Kepler problem is

$$A = \int_{t_a}^{t_b} dt \left[\vec{p} \cdot \frac{d\vec{x}}{dt} - \left(\frac{\vec{p} \cdot \vec{p}}{2m} - \frac{k}{r} \right) \right], \quad (1)$$

where $k = e^2$, and $\vec{p}$, $\vec{x}$, and r are

$$\vec{p} = (p_1, p_2, p_3, p_4, p_5), \quad \vec{x} = (x_1, x_2, x_3, x_4, x_5), \quad r = [\vec{x} \cdot \vec{x}]^{\frac{1}{2}}.$$

This system has 15 conserved quantities: 10 for the angular momentum $\epsilon_{ijk}x_jp_k$ and 5 for the Runge-Lenz vector A_i . To use this dynamical symmetry we add new degrees of freedom to the free-particle Lagrangian. Then, Eq. (1) becomes

$$A = \int_{t_a}^{t_b} dt \left[p_A \cdot \frac{dx_A}{dt} - \left(\frac{p_A \cdot p_A}{2m} - \frac{k}{r} \right) \right], \quad (2)$$

where

$$x_A = (\vec{x}, x_6, x_7, x_8), \quad p_A = (\vec{p}, p_6, p_7, p_8).$$

We introduce a new parametric time λ and the dimensionless variables X_A and P_A according to

$$\frac{dt}{d\lambda} = r, \quad (3)$$

$$X_A = (2m |p_0|)^{\frac{1}{2}} x_A, \quad (4)$$

$$P_A = (2m |p_0|)^{-\frac{1}{2}} p_A. \quad (5)$$

Then the action becomes

$$A = \int_{\lambda_a}^{\lambda_b} d\lambda \left[P_A \frac{dx_A}{d\lambda} + (-p_0) \frac{dt}{d\lambda} - \frac{(2m |p_0|)^{\frac{1}{2}}}{2m} R(P_A P_A + 1) + k \right].$$

The complex form of the Hurwitz transformation is

$$\begin{aligned} dX_a &= \frac{1}{\sqrt{2}} d(X_1 + iX_2) = d\xi_A^* \xi_C + \xi_A^* d\xi_C - d\xi_B \xi_D^* - \xi_B d\xi_D^*, \\ dX_b &= \frac{1}{\sqrt{2}} d(X_3 + iX_4) = d\xi_A^* \xi_D + \xi_A^* d\xi_D + d\xi_B \xi_C^* + \xi_B d\xi_C^*, \\ dX_c &= \frac{1}{\sqrt{2}} d(X_5 + iX_6) = \xi_A d\xi_A^* + \xi_B d\xi_B^* - \xi_C^* d\xi_C - \xi_D^* d\xi_D, \\ dX_d &= \frac{1}{\sqrt{2}} d(X_7 + iX_8) = \xi_B d\xi_A + \xi_A d\xi_B - \xi_D d\xi_C - \xi_C d\xi_D, \end{aligned} \quad (6)$$

where

$$\begin{aligned} \xi_A &= \frac{1}{\sqrt{2}} (u_1 + iu_2), \quad \xi_B = \frac{1}{\sqrt{2}} (u_3 + iu_4), \\ \xi_C &= \frac{1}{\sqrt{2}} (u_5 + iu_6), \quad \xi_D = \frac{1}{\sqrt{2}} (u_7 + iu_8). \end{aligned}$$

In this transformation X_6 , X_7 , and X_8 are non-holonomic.

The transformations for the momenta are

$$\begin{aligned}
 P_{\xi_A} &= \sqrt{2} [P_{x_a}^* \xi_C^* + P_{x_b}^* \xi_D^* + P_{x_c}^* \xi_A^* + P_{x_d} \xi_B], \\
 P_{\xi_B} &= \sqrt{2} [-P_{x_a} \xi_D^* + P_{x_b} \xi_C^* + P_{x_c}^* \xi_B^* - P_{x_d} \xi_A], \\
 P_{\xi_C} &= \sqrt{2} [P_{x_a} \xi_A^* + P_{x_b}^* \xi_B^* - P_{x_c} \xi_C^* - P_{x_d} \xi_D], \\
 P_{\xi_D} &= \sqrt{2} [-P_{x_a}^* \xi_B^* + P_{x_b} \xi_A^* - P_{x_c} \xi_D^* + P_{x_d} \xi_C]. \quad (7)
 \end{aligned}$$

By using Eqs. (6) and (7) we write the action as

$$A = \int_{\lambda_a}^{\lambda_b} d\lambda \left[\left(P_A \cdot \frac{dX_A}{d\lambda} - \frac{dP_A}{d\lambda} \cdot X_A \right) - p_0 \frac{dt}{d\lambda} - \omega (P_\xi P_{\xi^\dagger} + \xi^\dagger \xi) + k \right], \quad (8)$$

where $\omega = (|p_0|/2m)^{1/2}$, and ξ and $P_{\xi^\dagger}$ are the following four-component complex spinors

$$\xi = \frac{1}{2} \begin{pmatrix} \xi_A \\ \xi_B \\ \xi_C \\ \xi_D \end{pmatrix}, \quad P_{\xi^\dagger} = \frac{1}{2} \begin{pmatrix} P_{\xi_A}^* \\ P_{\xi_B}^* \\ P_{\xi_C}^* \\ P_{\xi_D}^* \end{pmatrix}.$$

Here $\xi^\dagger$ and P_ξ are defined as

$$\xi^\dagger = (\xi^*)^T, \quad P_\xi = \left[(P_{\xi^\dagger})^* \right]^T.$$

The coherent states are the eigenstates of the lowering operators of the harmonic oscillators. For this reason we define the holomorphic coordinates of the eight oscillators:

$$a = \frac{1}{2} \begin{pmatrix} \xi^\dagger + ip_\xi \\ \xi + ip_{\xi^\dagger} \end{pmatrix}, \quad a^\dagger = \frac{1}{2} (\xi - ip_{\xi^\dagger}, \xi^\dagger - ip_\xi). \quad (9)$$

In Eq. (9) a and $a^\dagger$ represent eight-component spinors. Then the Lagrangian of the oscillators becomes

$$L = \frac{1}{2i} \left(\frac{da^\dagger}{d\lambda} a - a^\dagger \frac{da}{d\lambda} \right) - p_0 \frac{dt}{d\lambda} - \omega a^\dagger a + k, \quad (10)$$

where ω $a^\dagger a$ is the Hamiltonian of the eight oscillators in the parametric time λ . The value of the parametric energy is k .

3 Path Integration

The kernel of the eight oscillators between the holomorphic coordinates $a_a = a(\lambda_a)$ and $a_b^\dagger = a^\dagger(\lambda_b)$ is

$$K(a_b^\dagger, t_b; a_a, t_a) = \int \frac{D(-p_0) Dt}{2\pi} \int \frac{Da^\dagger Da}{(2\pi i)^8} e^{iA}, \quad (11)$$

where A is the action of the system and where we use $\hbar = 1$. The relation between the physical kernel $K(\vec{x}_b, t_b; \vec{x}_a, t_a)$ of the 5D Kepler problem and $K(a_b^\dagger, t_b; a_a, t_a)$ is

$$K(\vec{x}_b, t_b; \vec{x}_a, t_a) = \int du_6(b) du_7(b) du_8(b) \int d\lambda K(a_b^\dagger, t_b; a_a, t_a).$$

We evaluate $K(a_b^\dagger, t_b; a_a, t_a)$ by using the method given in Ref. [14]. The result is

$$K(a_b^\dagger, t_b; a_a, t_a) = \exp \left[a_b^\dagger e^{-i\omega(\lambda_b - \lambda_a)} a_a - i(4\omega - k)(\lambda_b - \lambda_a) \right]. \quad (12)$$

Since the Hurwitz transformation is double-valued, the kernel is symmetric with respect to the end points:

$$K(a_b^\dagger, t_b; a_a, t_a) = \frac{1}{2} \left\{ \exp \left[a_b^\dagger e^{-i\omega(\lambda_b - \lambda_a)} a_a \right] + \exp \left[-a_b^\dagger e^{i\omega(\lambda_b - \lambda_a)} a_a \right] \right\} \\ \times e^{-i(4\omega - k)(\lambda_b - \lambda_a)}.$$

$K(a^\dagger(\lambda); \alpha(\lambda_a))$ is the matrix element of the evolution operator between $a^\dagger(\lambda) = a^\dagger$, $\alpha(\lambda_a) = \alpha$. It gives the time evolution of the eigenstate of the lowering operator a in parametric time:

$$K(a^\dagger(\lambda); \alpha) = |a(\lambda)\rangle = \frac{1}{2} \left[\exp(a^\dagger e^{i\omega\lambda}) \alpha + \exp(-a^\dagger e^{-i\omega\lambda}) \alpha \right]. \quad (13)$$

We expand the exponential in Eq. (13) into the power series of $a_i^\dagger$:

$$|a(\lambda)\rangle = \sum_{n_1, \dots, n_8=0}^{\infty} \left(\frac{1 + (-1)^{n_1 + \dots + n_8}}{2} \right)$$

$$\times \frac{(\alpha_1)^{n_1}}{\sqrt{n_1!}} \dots \frac{(\alpha_8)^{n_8}}{\sqrt{n_8!}} e^{-i\omega(n_1+\dots+n_8)}, \quad (14)$$

where $|n_i\rangle$ is the energy eigenstate of the i th oscillator.

The coherent state of the five-dimensional Kepler problem is obtained by integrating over the final values of the additional free particle coordinates $X_6(b)$, $X_7(b)$, and $X_8(b)$. These integrals are performed in spherical coordinates. We represent the complex coordinates ξ_A , ξ_B , ξ_C , and ξ_D in terms of spherical harmonics as

$$\begin{aligned} \xi_A &= |u| \cos \frac{\Theta}{2} (D_{--} D'_{--} + D_{-+} D'_{+-}), \\ \xi_B &= |u| \cos \frac{\Theta}{2} (-D_{+-} D'_{--} - D_{++} D'_{+-}), \\ \xi_C &= |u| \sin \frac{\Theta}{2} (D_{++} D'_{--} - D_{-+} D'_{+-}), \\ \xi_D &= |u| \sin \frac{\Theta}{2} (D_{+-} D'_{--} - D_{--} D'_{+-}), \end{aligned}$$

where $D_{mm'}$ and $D'_{nn'}$ are

$$D_{mm'} = D_{m/2, m'/2}^{(\frac{1}{2})}(\phi, \theta, \psi), \quad D'_{nn'} = D_{n/2, n'/2}^{(\frac{1}{2})}(\alpha, \beta, \gamma).$$

Then we can write the $\langle \xi, \xi^\dagger | \alpha(\lambda) \rangle$ as

$$\langle \xi, \xi^\dagger | \alpha(\lambda) \rangle = e^{-|\xi|^2 + \sqrt{2}(\xi\alpha + \xi^\dagger\alpha^\dagger) - \alpha\cdot\alpha^\dagger}, \quad (15)$$

where α and $\alpha^\dagger$ are the complex spinors

$$\alpha = (\alpha_-, \beta_-, \gamma_-, \delta_-) = \left(\frac{\alpha_1 - i\alpha_2}{\sqrt{2}}, \frac{\alpha_3 - i\alpha_4}{\sqrt{2}}, \frac{\alpha_5 - i\alpha_6}{\sqrt{2}}, \frac{\alpha_7 - i\alpha_8}{\sqrt{2}} \right),$$

$$\alpha^\dagger = \begin{pmatrix} \alpha_+ \\ \beta_+ \\ \gamma_+ \\ \delta_+ \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_1 + i\alpha_2 \\ \alpha_3 + i\alpha_4 \\ \alpha_5 + i\alpha_6 \\ \alpha_7 + i\alpha_8 \end{pmatrix}.$$

The expressions $\alpha \cdot \xi$ and $\xi^\dagger \cdot \alpha^\dagger$ are

$$\begin{aligned} \alpha \cdot \xi &= |u| (A_- D'_{--} + B_- D'_{+-}), \\ \alpha^\dagger \cdot \xi^\dagger &= |u| (A_+ D'^*_{--} + B_+ D'^*_{+-}), \end{aligned}$$

with

$$\begin{aligned} A_- &= \left[(\alpha - D_{--} - \beta_- D_{+-}) \cos \frac{\Theta}{2} + (\gamma_- D_{++} + \delta_- D_{+-}) \sin \frac{\Theta}{2} \right], \\ B_- &= \left[(\alpha - D_{--} - \beta_- D_{+-}) \cos \frac{\Theta}{2} + (-\gamma_- D_{++} - \delta_- D_{--}) \sin \frac{\Theta}{2} \right], \\ A_+ &= \left[(\alpha_+ - D_{--}^* - \beta + D_{+-}^*) \cos \frac{\Theta}{2} + (-\gamma_+ D_{-+}^* + \delta_+ D_{+-}^*) \sin \frac{\Theta}{2} \right], \\ B_+ &= \left[(\alpha_+ D_{-+}^* - \beta_+ D_{++}^*) \cos \frac{\Theta}{2} + (-\gamma_+ D_{-+}^* - \delta_+ D_{--}^*) \sin \frac{\Theta}{2} \right]. \end{aligned}$$

We substitute $\alpha \cdot \xi$ and $\alpha^\dagger \cdot \xi^\dagger$ into Eq. (15) and expand the exponential into power series of D'_{--} , D'_{+-} , D'^{*}_{--} , and D'^{*}_{+-} :

$$\begin{aligned} \langle \xi, \xi^\dagger | \alpha \rangle &= e^{-|\xi|^2 - \alpha \cdot \alpha^\dagger} \sum_{n_1, \dots, n_4=0}^{\infty} |u|^{(n_1+n_2+n_3+n_4)} (D'_{--})^{n_1} (D'_{+-})^{n_2} \\ &\quad \times \frac{(A_-)^{n_1} (B_-)^{n_2} (A_+)^{n_3} (B_+)^{n_4}}{n_1! n_2! n_3! n_4!} (D'^{*}_{--})^{n_3} (D'^{*}_{+-})^{n_4}. \end{aligned}$$

Using the following identities

$$\begin{aligned} (D'_{--})^{n_1} &= D_{-n_1/2, -n_1/2}^{n_1/2}(\alpha, \beta, \gamma), \\ (D'_{+-})^{n_2} &= D_{+n_2/2, -n_2/2}^{n_2/2}(\alpha, \beta, \gamma), \end{aligned}$$

we combine $(D'_{--})^{n_1}$ and $(D'_{+-})^{n_2}$ as

$$\begin{aligned} (D'_{--})^{n_1} (D'_{+-})^{n_2} &= \left\langle \frac{n_1 + n_2}{2}, -\frac{n_1 - n_2}{2} \middle| \frac{n_1}{2}, -\frac{n_1}{2}; \frac{n_2}{2}, \frac{n_2}{2} \right\rangle \\ &\quad \times D_{-(n_1+n_2)/2, -(n_1-n_2)/2}^{(n_1+n_2)/2}(\alpha, \beta, \gamma). \end{aligned}$$

Then $\langle \xi, \xi^\dagger | \alpha \rangle_{\text{phys}}$ becomes

$$\begin{aligned} \langle \xi, \xi^\dagger | \alpha \rangle_{\text{phys}} &= e^{-|\xi|^2 - \alpha^\dagger \alpha} \sum_{\substack{n_1, n_2, \\ n_3, n_4=0}}^{\infty} \frac{|u|^{(n_1+n_2+n_3+n_4)}}{n_1! \dots n_4!} (A_-)^{n_1} (B_-)^{n_2} \\ &\quad \times (A_+)^{n_3} (B_+)^{n_4} \left\langle \frac{n_1 + n_2}{2}, -\frac{n_1 - n_2}{2} \middle| \frac{n_1}{2}, -\frac{n_1}{2}; \frac{n_2}{2}, \frac{n_2}{2} \right\rangle \\ &\quad \times \left\langle \frac{n_3}{2}, -\frac{n_3}{2}; \frac{n_4}{2}, \frac{n_4}{2} \middle| \frac{n_3 + n_4}{2}, -\frac{n_3 - n_4}{2} \right\rangle \frac{1}{32\pi^2} \int d\alpha d(\cos \beta) d\gamma \\ &\quad \times D_{-(n_3+n_4)/2, -(n_3-n_4)/2}^{*(n_3+n_4)/2}(\alpha, \beta, \gamma) D_{-(n_1+n_2)/2, -(n_1-n_2)/2}^{(n_1+n_2)/2}(\alpha, \beta, \gamma). \quad (16) \end{aligned}$$

We perform the α, β and γ integration. The result is

$$\begin{aligned} \langle \xi, \xi^\dagger | \alpha \rangle_{phys} &= e^{-|\xi|^2 - \alpha^\dagger \alpha} \sum_{n_1, n_2=0}^{\infty} |u|^{2(n_1+n_2)} \frac{(A_- A_+)^{n_1} (B_- B_+)^{n_2}}{(n_1! n_2!)^2} \\ &\times \left| \left\langle \frac{n_1 + n_2}{2}, -\frac{n_1 - n_2}{2} \middle| \frac{n_1}{2}, -\frac{n_1}{2}; \frac{n_2}{2}, \frac{n_2}{2} \right\rangle \right|^2, \end{aligned} \quad (17)$$

where $A_- A_+$ and $B_- B_+$ are linear functions of $D_{mm'}^1(\phi, \theta, \psi)$ and $\cos \Theta$. Then the final expression of the coherent states is a function of $D_{mm'}^j(\phi, \theta, \psi)$, $r^{(n_1+n_2)}$ and $\cos \Theta$.

4 Conclusion

We have used the mapping between the five-dimensional hydrogen atom and eight oscillators and quantized the latter via path integration over holomorphic coordinates. The kernel governs the parametric time evolution of the eigenstates of the lowering operators and corresponds to the coherent states for the harmonic oscillators. The contribution of the extra degrees of freedom is eliminated by integrating over their final points. We expressed the state functions of the oscillators in terms of the spherical harmonics $D_{mm'}^{1/2}(\phi, \theta, \psi)$ and $D_{nn'}^{1/2}(\alpha, \beta, \gamma)$ and performed the integrals over α, β and γ by using the orthogonality properties of the D functions. The final expression of the coherent states depends on five physical coordinates of the hydrogen atom.

The Kustaanheimo-Stiefel transformation can be written as $x = u^2$ between the 3-component x and 4-component u quaternions. One of the open problems is to write the Hurwitz transformation in the same way as the Kustaanheimo-Stiefel transformation in terms of the octonions without using the additional degrees of freedom. We used the path integration over the c -number holomorphic coordinates a and $a^\dagger$, instead of x and p . Another open problem is to define the path integrals over the quaternions or octonions.

Acknowledgments

This work is partly supported by the Akdeniz University (99.01.0105.03).

References

- [1] E. Schrödinger, *Naturwissenschaften* **14**, 664 (1926).

- [2] A. ten Wolde, L.D. Noordam, A. Lagendijk, and H.B. van Linden van den Heuvel, *Phys. Rev. Lett.* **61**, 2099 (1988).
- [3] M.M. Nieto and L.M. Simmons Jr., *Phys. Rev. Lett.* **41**, 207 (1978).
- [4] L.S. Brown, *Am. J. Phys.* **41**, 525 (1973).
- [5] D. Bhaumik, B. Dutta-Roy, and G. Ghosh, *J. Phys: Math. Gen.* **19**, 1355 (1986).
- [6] C.C. Gerry, *Phys. Rev. A* **33**, 6 (1988).
- [7] T. Toyoda and S. Wakayama, *Phys. Rev. A* **59**, 1021 (1999).
- [8] I.H. Duru and H. Kleinert, *Phys. Lett. B* **84**, 185 (1979) and *Fortschr. Phys.* **30**, 401 (1982).
- [9] P. Kustaanheimo and E. Stiefel, *J. Reine Angew. Math.* **218**, 204 (1965).
- [10] N. Ünal, preprint.
- [11] T. Levi-Civita, *Opere Matematiche* **2**, 411 (1956).
- [12] F.H.J. Cornish, *J. Phys A: Math. Gen.* **17**, 323 (1984).
- [13] A. Hurwitz, *Mathematische Werke*, Band II (Birkhäuser, Basel, 1933), p. 641.
- [14] N. Ünal, *Found. Phys.* **28**, 755 (1998).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

THE AHARONOV-BOHM EFFECT IN FOUR DIMENSIONS

D.H. LIN

*Institute of Electro-Physics, National Chiao Tung University,
Hsinchu 30043, Taiwan*

E-mail: dhlin@cc.nctu.edu.tw

The path-integral method is particularly useful for studying global questions of topology. We discuss the influence of the Aharonov-Bohm effect in four-dimensional spherical systems via the path integral. As a realization, we give the exact Green's function of the relativistic A-B-C system. The procedure can be generalized to any-dimensional systems.

1 Introduction

It is an honor and a pleasure to present an article at this celebration of Professor Dr. Hagen Kleinert's birthday. During the past several years, I have spent some time studying relativistic systems using path integrals. It was Kleinert who first introduced the space-time transformation in the path integral [2] to compute the hydrogen atom [2–4] leading to a vigorous development in this field. We want to show the consequences of the topological effects of the Aharonov-Bohm (A-B) effect in four-dimensional spherically symmetric systems. We study such systems, since several physical systems turn in momentum space into a dynamical problem of a three-dimensional surface in four dimensions, such as the hydrogen atom whose dynamics in momentum space takes place on a three-dimensional sphere in four-dimensional space [4].

2 The General Influence of the A-B Effect in Spherical Symmetric Systems

The path-integral representation for the Green's function of a relativistic particle moving in external electromagnetic fields in the four-dimensional space is given by [3,5,6]:

$$G(\mathbf{x}_b, \mathbf{x}_a; E) = \frac{i\hbar}{2mc} \int_0^\infty dS \int \mathcal{D}\rho(\lambda) \Phi[\rho(\lambda)] \int \mathcal{D}^4 x(\lambda) \times \exp \{ -A_E[\mathbf{x}, \dot{\mathbf{x}}] / \hbar \} \rho(0), \quad (1)$$

where the Euclidean action reads

$$A_E[\mathbf{x}, \dot{\mathbf{x}}] = \int_{\lambda_a}^{\lambda_b} d\lambda \left[\frac{m}{2\rho(\lambda)} \dot{\mathbf{x}}^2(\lambda) - i(e/c) \mathbf{A}(\mathbf{x}) \cdot \dot{\mathbf{x}}(\lambda) - \rho(\lambda) \frac{(E - V(\mathbf{x}))^2}{2mc^2} + \rho(\lambda) \frac{mc^2}{2} \right], \quad (2)$$

with S defined as

$$S = \int_{\lambda_a}^{\lambda_b} d\lambda \rho(\lambda). \quad (3)$$

Here $\rho(\lambda)$ is an arbitrary dimensionless fluctuating scale variable, $\rho(0)$ is the terminal point of the function $\rho(\lambda)$, and $\Phi[\rho(\lambda)]$ is some convenient gauge-fixing functional [3,5,6]. The only condition on $\Phi[\rho(\lambda)]$ is that

$$\int \mathcal{D}\rho(\lambda) \Phi[\rho(\lambda)] = 1. \quad (4)$$

The prefactor $\hbar/mc$ in Eq. (1) is the Compton wave length of a particle of mass m , and $\mathbf{A}(\mathbf{x})$, $V(\mathbf{x})$ stand for the vector and scalar potential of the system, respectively. The constant E is the system energy, and $\mathbf{x}$ is the spatial part of the $(4+1)$ -vector $x^\mu = (\mathbf{x}, \tau)$. The functional integral for $\mathbf{x}$ in the representation of Eq. (1) can be interpreted as the expectation value of the real functional $\exp \left\{ - \int_{\lambda_a}^{\lambda_b} d\lambda \beta \rho(\lambda) V(\mathbf{x}(\lambda)) / \hbar \right\}$ over the measure

$$K_0(\mathbf{x}_b, \mathbf{x}_a; \lambda_b - \lambda_a) = \int \mathcal{D}^4 x(\lambda) \exp \left\{ - \frac{1}{\hbar} \int_{\lambda_a}^{\lambda_b} d\lambda \left[\frac{m}{2\rho(\lambda)} \dot{\mathbf{x}}^2(\lambda) - i \frac{e}{c} \mathbf{A}(\mathbf{x}) \cdot \dot{\mathbf{x}}(\lambda) - \rho(\lambda) \frac{V^2(\mathbf{x})}{2mc^2} \right] \right\}. \quad (5)$$

With this interpretation, the entire Green's function turns into the following formula

$$G(\mathbf{x}_b, \mathbf{x}_a; E) = \frac{i\hbar}{2mc} \int_0^\infty dS \int \mathcal{D}\rho(\lambda) \Phi[\rho(\lambda)] e^{-\frac{1}{\hbar} \int_{\lambda_a}^{\lambda_b} d\lambda \rho(\lambda) \mathcal{E}} \\ \times \left\langle \exp \left\{ -\frac{1}{\hbar} \int_{\lambda_a}^{\lambda_b} d\lambda \beta \rho(\lambda) V(\mathbf{x}(\lambda)) \right\} \right\rangle \rho(0), \quad (6)$$

in which $\mathcal{E} = (m^2 c^4 - E^2)/2mc^2$ and $\beta = E/mc^2$, with the notation $\langle \star \rangle$ standing for the expectation value of the moment $\star$ over the measure $K_0(\mathbf{x}_b, \mathbf{x}_a; \lambda_b - \lambda_a)$. Equation (6) forms the basis for studying the relativistic potential problems via the Feynman-Kac type formula.

Expanding the potential $V(\mathbf{x})$ in Eq. (6) into a power series and interchanging the order of integration and summation, we have

$$G(\mathbf{x}_b, \mathbf{x}_a; E) = \frac{i\hbar}{2mc} \int_0^\infty dS \int \mathcal{D}\rho \Phi[\rho] e^{-\frac{1}{\hbar} \int_{\lambda_a}^{\lambda_b} d\lambda \rho(\lambda) \mathcal{E}} \\ \times \sum_{n=0}^\infty \frac{(-\beta/\hbar)^n}{n!} \left\langle \left(\int_{\lambda_a}^{\lambda_b} d\lambda \rho(\lambda) V(\mathbf{x}(\lambda)) \right)^n \right\rangle \rho(0). \quad (7)$$

We see that the calculation of the relativistic path integral now turns into the computation of the expectation value of moments Q^n ($Q = \int_{\lambda_a}^{\lambda_b} d\lambda \rho V(\mathbf{x})$) over the Feynman measure and their summation in accordance with the Feynman-Kac type formula. Ordering the λ as $\lambda_1 < \lambda_2 < \dots < \lambda_n < \lambda_b$ and denoting $\mathbf{x}(\lambda_i) = \mathbf{x}_i$, the perturbation series in Eq. (7) explicitly turns into [1]

$$\sum_{n=0}^\infty \frac{(-\beta/\hbar)^n}{n!} \left\langle \left(\int_{\lambda_a}^{\lambda_b} d\lambda \rho(\lambda) V(\mathbf{x}(\lambda)) \right)^n \right\rangle = K_0(\mathbf{x}_b, \mathbf{x}_a; \lambda_b - \lambda_a) \\ + \sum_{n=1}^\infty \left(-\frac{\beta}{\hbar} \right)^n \int_{\lambda_a}^{\lambda_b} d\lambda_n \int_{\lambda_a}^{\lambda_n} d\lambda_{n-1} \dots \int_{\lambda_a}^{\lambda_2} d\lambda_1 \\ \times \int \left[\prod_{j=0}^n K_0(\mathbf{x}_{j+1}, \mathbf{x}_j; \lambda_{j+1} - \lambda_j) \right] \prod_{i=1}^n \rho_i V(\mathbf{x}_i) d\mathbf{x}_i, \quad (8)$$

where $\lambda_0 = \lambda_a$, $\lambda_{n+1} = \lambda_b$, $\mathbf{x}_{n+1} = \mathbf{x}_b$, and $\mathbf{x}_0 = \mathbf{x}_a$. The merit of expanding the path integral into this form is that we can let $\Phi[\rho]$ be equal to the delta functional $\delta[\rho - 1]$ in order to fix the value of $\rho(\lambda)$ to unity, such that the integration over S in Eq. (7) leads to a Laplace transformation [7–9]. Because

of the convolution property of the Laplace transformation, we obtain

$$G(\mathbf{x}_b, \mathbf{x}_a; E) = \frac{i\hbar}{2mc} \left\{ G_0(\mathbf{x}_b, \mathbf{x}_a; \mathcal{E}) + \sum_{n=1}^{\infty} \left(-\frac{\beta}{\hbar} \right)^n \int \left[\prod_{j=0}^n G_0(\mathbf{x}_{j+1}, \mathbf{x}_j; \mathcal{E}) \right] \prod_{i=1}^n V(\mathbf{x}_i) d\mathbf{x}_i \right\}, \quad (9)$$

where $G_0(\mathbf{x}_b, \mathbf{x}_a; \mathcal{E})$ is the Laplace transformation of the pseudo propagator $K_0(\mathbf{x}_b, \mathbf{x}_a; \lambda_b - \lambda_a)$. For calculating the explicit form of the Green's function G_0 , the relations between the Cartesian coordinates (x_1, x_2, x_3, x_4) and the polar coordinates $(r, \vartheta, \theta, \varphi)$, with the range of validity $0 < r < \infty$, $0 < \vartheta, \theta < \pi$, and $0 < \varphi < 2\pi$, read

$$\begin{aligned} x_1 &= r \cos \vartheta, & x_2 &= r \sin \vartheta \cos \theta, & x_3 &= r \sin \vartheta \sin \theta \cos \varphi, \\ x_4 &= r \sin \vartheta \sin \theta \sin \varphi. \end{aligned} \quad (10)$$

With these transformations, the volume element is

$$d\mathbf{x} = r^3 \sin^2 \vartheta \sin \theta dr d\vartheta d\theta d\varphi. \quad (11)$$

We have the angular decomposition of the spherical symmetric systems

$$G_0(\mathbf{x}_{j+1}, \mathbf{x}_j; \mathcal{E}) = \sum_{s=0}^{\infty} \sum_{l=0}^s \sum_{k=-l}^l g_l^{(0)}(r_{j+1}, r_j; \mathcal{E}) Y_{slk}(\hat{\mathbf{x}}_{j+1}) Y_{slk}^*(\hat{\mathbf{x}}_j), \quad (12)$$

where $Y_{slk}(\hat{\mathbf{x}}_j) = Y_{slk}(\vartheta_j, \theta_j, \varphi_j)$ is the four-dimensional spherical harmonics. Using the orthogonality relations,

$$\int \sin^2 \vartheta \sin \theta dr d\vartheta d\theta d\varphi Y_{slk}(\hat{\mathbf{x}}) Y_{s'l'k'}^*(\hat{\mathbf{x}}) = \delta_{ss'} \delta_{ll'} \delta_{mm'}, \quad (13)$$

the intermediate angular part can be performed. It yields

$$G(\mathbf{x}_b, \mathbf{x}_a; E) = \frac{i\hbar}{2mc} \sum_{s=0}^{\infty} \sum_{l=0}^s \sum_{k=-l}^l G_l(r_b, r_a; \mathcal{E}) Y_{slk}(\hat{\mathbf{x}}_b) Y_{slk}^*(\hat{\mathbf{x}}_a), \quad (14)$$

with the pure radial Green's function

$$G_l(r_b, r_a; \mathcal{E}) = \sum_{n=0}^{\infty} \left(-\frac{\beta}{\hbar} \right)^n g_l^{(n)}(r_b, r_a; \mathcal{E}). \quad (15)$$

The radial integrations are given by

$$g_l^{(n)}(r_b, r_a; \mathcal{E}) = \int_0^\infty \cdots \int_0^\infty \left[\prod_{j=0}^n g_l^{(0)}(r_{j+1}, r_j; \mathcal{E}) \right] \prod_{i=1}^n r_i^3 V(r_i) dr_i. \quad (16)$$

Here the unperturbed radial Green's function $g_l^{(0)}$ has the path-integral representation

$$g_l^{(0)}(r_{j+1}, r_j; \mathcal{E}) = \frac{1}{r_{j+1} r_j} \int_0^\infty dS e^{-\mathcal{E}S/\hbar} \int \mathcal{D}r(\lambda) \times \exp \left\{ -\frac{1}{\hbar} \int_{\lambda_a}^{\lambda_b} d\lambda \left[\frac{m}{2} \dot{r}^2(\lambda) - i \frac{e}{c} \mathbf{A}(x) \cdot \dot{\mathbf{x}}(\lambda) - \frac{V^2(r)}{2mc^2} \right] \right\}. \quad (17)$$

It is now that we can consider the influence of the A-B effect for systems with spherical symmetry. The vector potential of the A-B effect is given by

$$\mathbf{A}(\mathbf{x}) = 2g \frac{-x_4 \hat{e}_3 + x_3 \hat{e}_4}{x_3^2 + x_4^2}, \quad (18)$$

where $\hat{e}_{3,4}$ stands for the unit vector along the x, y axis. Introducing the azimuthal angle around the A-B tube

$$\varphi(\mathbf{x}) = \arctan(x_4/x_3), \quad (19)$$

the interaction from the A-B effect, denoted in the action by the subscript mag, has the form

$$A_{\text{mag}} = i\hbar\mu_0 \int_0^S d\lambda \dot{\varphi}(\lambda), \quad (20)$$

where $\varphi(\lambda) = \varphi(\mathbf{x}(\lambda))$, $\dot{\varphi} = d\varphi/d\lambda$, and $\mu_0 = -2eg/\hbar c$ is a dimensionless number. The minus sign is a matter of convention. Since the particle orbits are present at all times, their worldlines in space-time can be considered as being closed at infinity, and the integral

$$k = \frac{1}{2\pi} \int_0^S d\lambda \dot{\varphi}(\lambda) \quad (21)$$

is a topological invariant with integer values of the winding number k . The magnetic interaction is therefore purely topological, its value being

$$A_{\text{mag}} = i\hbar\mu_0 2k\pi. \quad (22)$$

The influence of the A-B effect in the entire Green's function $G(\mathbf{x}_b, \mathbf{x}_a; \mathcal{E})$ can be considered by writing down the explicit form of the four-dimensional spherical harmonics

$$\begin{aligned} & \sum_{s=0}^{\infty} \sum_{l=0}^s \sum_{k=-l}^l Y_{slk}(\hat{\mathbf{x}}_b) Y_{slk}^*(\hat{\mathbf{x}}_a) \\ &= \sum_{s=0}^{\infty} \sum_{l=0}^s \sum_{k=-l}^l \frac{2^{2l+1}(s+1)\Gamma^2(l+1)\Gamma(s-l+1)}{\pi\Gamma(s+l+2)} (\sin\vartheta_b \sin\vartheta_a)^l \\ & \quad \times C_{s-l}^{l+1}(\cos\vartheta_b) C_{s-l}^{l+1}(\cos\vartheta_a) Y_{lk}(\theta_b, \varphi_b) Y_{lk}^*(\theta_a, \varphi_a), \end{aligned} \quad (23)$$

where C_n^λ and Y_{lk} are the Gegenbauer polynomials and the usual three-dimensional spherical harmonics. The orthogonality relations of $Y_{slk}(\hat{\mathbf{x}})$ can be easily checked by the formula:

$$\int_0^\pi d\vartheta \sin^{2\lambda} \vartheta C_m^\lambda(\cos\vartheta) C_n^\lambda(\cos\vartheta) = \frac{\pi 2^{1-2\lambda} \Gamma(n+2\lambda)}{n! (\lambda+n) \Gamma^2(\lambda)} \delta_{m,n}. \quad (24)$$

With the help of the following relation between the associated Legendre polynomial $P_\nu^\mu(x)$ and the Jacobi function $P_n^{(\alpha,\beta)}(x)$ [8],

$$P_l^k(\cos\theta) = (-1)^k \frac{\Gamma(1+k+l)}{\Gamma(1+l)} (\cos\theta/2 \sin\theta/2)^k P_{l-k}^{(k,k)}(\cos\theta), \quad (25)$$

the angular part of Eq. (23) turns into

$$\begin{aligned} & \sum_{s=0}^{\infty} \sum_{l=0}^s \sum_{k=-l}^l Y_{slk}(\hat{\mathbf{x}}_b) Y_{slk}^*(\hat{\mathbf{x}}_a) \\ &= \sum_{s=0}^{\infty} \sum_{l=0}^s \sum_{k=-l}^l \frac{2^{2l+1}(s+1)(2l+1)\Gamma(s-l+1)\Gamma(l+k+1)\Gamma(l-k+1)}{4\pi^2\Gamma(s+l+2)} \\ & \quad \times (\sin\vartheta_b \sin\vartheta_a)^l C_{s-l}^{l+1}(\cos\vartheta_b) C_{s-l}^{l+1}(\cos\vartheta_a) e^{ik(\varphi_b - \varphi_a)} \\ & \quad \times (\cos\theta_b/2 \cos\theta_a/2 \sin\theta_b/2 \sin\theta_a/2)^k P_{l-k}^{(k,k)}(\cos\theta_b) P_{l-k}^{(k,k)}(\cos\theta_a). \end{aligned} \quad (26)$$

To go further, let us change the variables s, l by defining $s-l = t$, $l-k = q$ into t, q . Thus the Green's function of Eq. (14) becomes

$$G(\mathbf{x}_b, \mathbf{x}_a; E) = \frac{i\hbar}{2mc} \sum_{t=0}^{\infty} \sum_{q=0}^{\infty} \sum_{k=-\infty}^{\infty} 2^{2(q+k)+1} G_{q+k}(r_b, r_a; \mathcal{E})$$

$$\begin{aligned}
& \times \frac{(t+q+k+1)[2(q+k)+1]\Gamma(q+2k+1)\Gamma(q+1)\Gamma(t+1)}{4\pi^2\Gamma(t+2(q+k)+2)} \\
& \times (\sin\vartheta_b \sin\vartheta_a)^{q+k} C_t^{q+k+1}(\cos\vartheta_b) C_t^{q+k+1}(\cos\vartheta_a) e^{ik(\varphi_b-\varphi_a)} \\
& \times (\cos\theta_b/2 \cos\theta_a/2 \sin\theta_b/2 \sin\theta_a/2)^k P_q^{(k,k)}(\cos\theta_b) P_q^{(k,k)}(\cos\theta_a). \quad (27)
\end{aligned}$$

Now, let us consider the A-B effect by invoking Poisson's summation formula

$$\sum_{k=-\infty}^{\infty} f(k) = \int_{-\infty}^{\infty} dy \sum_{n=-\infty}^{\infty} e^{2\pi n y i} f(y). \quad (28)$$

The entire Green's function $G_0(\mathbf{x}_b, \mathbf{x}_a; \mathcal{E})$ containing the A-B effect becomes

$$\begin{aligned}
G(\mathbf{x}_b, \mathbf{x}_a; E) &= \frac{i\hbar}{2mc} \sum_{t=0}^{\infty} \sum_{q=0}^{\infty} \int dz \sum_{k=-\infty}^{\infty} 2^{2(q+z)+1} G_{q+z}(r_b, r_a; \mathcal{E}) \\
& \times \frac{(t+q+z+1)[2(q+z)+1]\Gamma(q+2z+1)\Gamma(q+1)\Gamma(t+1)}{4\pi^2\Gamma(t+2(q+z)+2)} \\
& \times (\sin\vartheta_b \sin\vartheta_a)^{q+z} C_t^{q+z+1}(\cos\vartheta_b) C_t^{q+z+1}(\cos\vartheta_a) e^{i(z-\mu_0)(\varphi_b+2k\pi-\varphi_a)} \\
& \times (\cos\theta_b/2 \cos\theta_a/2 \sin\theta_b/2 \sin\theta_a/2)^z P_q^{(z,z)}(\cos\theta_b) P_q^{(z,z)}(\cos\theta_a). \quad (29)
\end{aligned}$$

The sum over all k in Eq. (29) forces z to be equal to μ_0 modulo an arbitrary integral number leading to

$$\begin{aligned}
G(\mathbf{x}_b, \mathbf{x}_a; E) &= \frac{i\hbar}{2mc} \sum_{t=0}^{\infty} \sum_{q=0}^{\infty} \sum_{k=-\infty}^{\infty} 2^{2(q+|k+\mu_0|)+1} G_{q+|k+\mu_0|}(r_b, r_a; \mathcal{E}) \\
& \times \frac{[2(q+|k+\mu_0|)+1]\Gamma(q+2|k+\mu_0|+1)\Gamma(q+1)\Gamma(t+1)}{4\pi^2\Gamma(t+2(q+|k+\mu_0|)+2)} \\
& \times (t+q+|k+\mu_0|+1) (\cos\theta_b/2 \cos\theta_a/2 \sin\theta_b/2 \sin\theta_a/2)^{|k+\mu_0|} \\
& \times (\sin\vartheta_b \sin\vartheta_a)^{q+|k+\mu_0|} C_t^{q+|k+\mu_0|+1}(\cos\vartheta_b) C_t^{q+|k+\mu_0|+1}(\cos\vartheta_a) \\
& \times P_q^{(|k+\mu_0|, |k+\mu_0|)}(\cos\theta_b) P_q^{(|k+\mu_0|, |k+\mu_0|)}(\cos\theta_a) e^{ik(\varphi_b-\varphi_a)}. \quad (30)
\end{aligned}$$

From this equation, we see that the influence of the A-B flux will cover any dimensions although it is just described by the two-dimensional coordinates. Its effects are not only in the energy spectra but also in the wave functions of higher-dimensional spaces. As a realization, we consider the pure Coulomb system moving in the fields of A-B in four dimensions. Since the potential of Coulomb is $-e^2/r$, the pure radial amplitude $G_{q+|k+\mu_0|}(r_b, r_a; \mathcal{E})$ has the

form

$$G_{q+|k+\beta_0|}(r_b, r_a; \mathcal{E}) = \frac{m}{\hbar} \frac{1}{r_b r_a} \sum_{n=0}^{\infty} \left(\frac{m \beta e^2}{\hbar^2} \right)^n g_{q+|k+\mu_0|}^{(n)}(r_b, r_a; \mathcal{E}), \quad (31)$$

with

$$g_{q+|k+\beta_0|}^{(n)}(r_b, r_a; \mathcal{E}) = \int_0^{\infty} \cdots \int_0^{\infty} \left[\prod_{j=0}^n g_{q+|k+\beta_0|}^{(0)}(r_{j+1}, r_j; \mathcal{E}) \right] \prod_{i=1}^n dr_i, \quad (32)$$

in which $g_{q+|k+\beta_0|}^{(0)}$ is given by [7]

$$g_{q+|k+\beta_0|}^{(0)}(r_{j+1}, r_j; \mathcal{E}) = \int_0^{\infty} \frac{dS}{S} e^{-\frac{\xi}{\hbar} S} e^{-m(r_{j+1}^2 + r_j^2)/2\hbar S} I_{\sqrt{[(q+|k+\mu_0|)+1]^2 - \alpha^2}} \left(\frac{m}{\hbar} \frac{r_{j+1} r_j}{S} \right). \quad (33)$$

To obtain the explicit result of $g_{q+|k+\beta_0|}^{(n)}$, we note that [7]

$$\begin{aligned} \int_0^{\infty} \frac{dS}{S} e^{-\frac{\xi}{\hbar} S} e^{-m(r_b^2 + r_a^2)/2\hbar S} I_{\rho} \left(\frac{m}{\hbar} \frac{r_b r_a}{S} \right) \\ = 2 \int_0^{\infty} dz \frac{1}{\sinh z} e^{-\kappa(r_b + r_a) \coth z} I_{2\rho} \left(\frac{2\kappa \sqrt{r_b r_a}}{\sinh z} \right) \end{aligned} \quad (34)$$

with $\kappa = \sqrt{m^2 c^4 - E^2}/\hbar c$. With the help of the integral formula

$$\int_0^{\infty} dr r e^{-r^2/a} I_{\nu}(\zeta r) I_{\nu}(\xi r) = \frac{a}{2} e^{a(\xi^2 + \zeta^2)/4} I_{\nu}(a\xi\zeta/2), \quad (35)$$

we obtain the result

$$\begin{aligned} g_{q+|k+\beta_0|}^{(1)}(r_b, r_a; \mathcal{E}) &= \int_0^{\infty} g_{q+|k+\beta_0|}^{(0)}(r_b, r; \mathcal{E}) g_{q+|k+\beta_0|}^{(0)}(r, r_a; \mathcal{E}) dr \\ &= \frac{2^2}{\kappa} \int_0^{\infty} z h(z) dz, \end{aligned} \quad (36)$$

where the function $h(z)$ is defined as

$$h(z) = \frac{1}{\sinh z} e^{-\kappa(r_b + r_a) \coth z} I_{2\sqrt{[(q+|k+\beta_0|)+1]^2 - \alpha^2}} \left(\frac{2\kappa \sqrt{r_b r_a}}{\sinh z} \right). \quad (37)$$

The expression for $g_{q+|k+\beta_0|}^{(n)}(r_b, r_a; \mathcal{E})$ can be obtained by induction with respect to n , and is given by

$$g_{q+|k+\beta_0|}^{(n)}(r_b, r_a; \mathcal{E}) = \frac{2^{n+1}}{n!} \frac{1}{\kappa^n} \int_0^\infty z^n h(z) dz. \quad (38)$$

Inserting this expression in Eq. (31), we obtain

$$G_{q+|k+\beta_0|}(r_b, r_a; \mathcal{E}) = \frac{m}{\hbar} \frac{2}{r_b r_a} \int_0^\infty dz e^{\left(\frac{2m\beta_0 e^2}{\hbar^2 \kappa}\right)z} \times \frac{1}{\sinh z} e^{-\kappa(r_b+r_a) \coth z} I_{2\sqrt{[(q+|k+\beta_0|)+1]^2 - \alpha^2}} \left(\frac{2\kappa\sqrt{r_b r_a}}{\sinh z} \right). \quad (39)$$

The integration can be done by using the formula

$$\int_0^\infty dy \frac{e^{2\nu y}}{\sinh y} \exp \left[-\frac{t}{2} (\zeta_a + \zeta_b) \coth y \right] I_\mu \left(\frac{t\sqrt{\zeta_b \zeta_a}}{\sinh y} \right) = \frac{\Gamma((1+\mu)/2 - \nu)}{t\sqrt{\zeta_b \zeta_a} \Gamma(\mu+1)} W_{\nu, \mu/2}(t\zeta_b) M_{\nu, \mu/2}(t\zeta_a), \quad (40)$$

where $M_{\mu, \nu}$ and $W_{\mu, \nu}$ are the Whittaker functions. We complete the integration and obtain

$$G_{q+|k+\mu_0|}(r_b, r_a; \mathcal{E}) = \frac{1}{(r_b r_a)^{3/2}} \frac{mc}{\sqrt{m^2 c^4 - E^2}} \times \frac{\Gamma \left(1/2 + \sqrt{[(q+|k+\mu_0|)+1]^2 - \alpha^2} - E\alpha/\sqrt{m^2 c^4 - E^2} \right)}{\Gamma \left(1 + 2\sqrt{[(q+|k+\mu_0|)+1]^2 - \alpha^2} \right)} \times W_{E\alpha/\sqrt{m^2 c^4 - E^2}, \sqrt{[(q+|k+\beta_0|)+1]^2 - \alpha^2}} \left(\frac{2}{\hbar c} \sqrt{m^2 c^4 - E^2} r_b \right) \times M_{E\alpha/\sqrt{m^2 c^4 - E^2}, \sqrt{[(q+|k+\beta_0|)+1]^2 - \alpha^2}} \left(\frac{2}{\hbar c} \sqrt{m^2 c^4 - E^2} r_a \right). \quad (41)$$

Inserting this in Eq. (30), we obtain the entire relativistic Green's function of the A-B-C system in four-dimensional space. The energy spectrum is determined by the poles of the Gamma function

$$1/2 + \sqrt{[(q+|k+\mu_0|)+1]^2 - \alpha^2} - E\alpha/\sqrt{m^2 c^4 - E^2} = -n_r, \quad (42)$$

where $n_r = 0, 1, 2, \dots$. It is easy to find the energy levels by expanding the parameter α in the lowest two orders:

$$E_{q,k} = mc^2 \left\{ 1 - \frac{\alpha^2}{2(n_r + q + |k + \mu_0| + 3/2)^2} - \frac{\alpha^4}{(n_r + q + |k + \mu_0| + 3/2)^3} \right. \\ \left. \times \left[\frac{1}{2q + 2|k + \mu_0| + 2} - \frac{3}{8} \frac{1}{(n_r + q + |k + \mu_0| + 3/2)} \right] + \mathcal{O}(\alpha^6) \right\}. \quad (43)$$

We see that, if the flux is quantized, the result is reduced to the pure relativistic Coulomb spectrum in four dimensions [7] by $4\pi g = 2\pi\hbar c/e \times \text{integer}$, leading to an integer-valued $|k + \mu_0|$. In comparison with the conventional path-integral treatment of the hydrogen atom, the merit of our approach is that the complicated multi-valued Kustaanheimo-Stiefel and the space-time transformation techniques are avoided. Furthermore, the calculation of the entire Green's function is just related to the unperturbed one.

3 Discussion

We have discussed the influence of the A-B effect on the single particle system moving in a spherical symmetric potential. We have shown that the influence of the flux on the energy spectra and wave functions will cover any space dimensions. This is not common for particles moving in magnetic fields. We remind the Landau levels of a particle moving in a homogeneous magnetic field $\mathbf{B}$ pointing along the direction of the third axis. The magnetic field can be described by the vector potential $\mathbf{A}(\mathbf{x}) = (0, Bx, 0)$ or a more complicated one by performing phase transformation of an appropriate gauge. For the exact propagator it turns out that the particle is free in the third axis and, of course, in higher dimensions. The reason lies in the topological effect of the A-B coupling to the angular momentum, leading to effects in any dimension. In recent years, topological effects and geometrical phases have penetrated many areas of science. We hope that our discussion contributes to our understanding of the role of topology in physics.

References

- [1] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw Hill, New York, 1965).

- [2] I.H. Duru and H. Kleinert, *Phys. Lett. B* **84**, 185 (1979). More details are explained in: H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [3] H. Kleinert, *Phys. Lett. A* **212**, 15 (1996).
- [4] H. Kleinert, *Phys. Lett. A* **252**, 277 (1999).
- [5] D.H. Lin, *J. Math. Phys.* **40**, 1246 (1999).
- [6] D.H. Lin, *J. Phys. A* **31**, 4785 (1998).
- [7] D.H. Lin, *J. Phys. A* **31**, 7577 (1998).
- [8] D.H. Lin, *J. Math. Phys.* **41**, 2723 (2000).
- [9] D.H. Lin, *J. Phys. A* **30**, 4365 (1997).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

PATH INTEGRAL FOR SEPARABLE HAMILTONIANS OF LIOUVILLE-TYPE

K. FUJIKAWA

Department of Physics, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

E-mail: fujikawa@www-hep.phys.s.u-tokyo.ac.jp

A general path-integral analysis of the separable Hamiltonian of Liouville-type is reviewed. The basic dynamical principle used is Jacobi's principle of least action for given energy which is reparametrization invariant, and thus the gauge freedom naturally appears. The choice of gauge in a path integral corresponds to the separation of variables in operator formalism. The gauge independence and the operator ordering are closely related. The path integral in this formulation sums over orbits in space instead of space-time. An exact path integral of the Green's function for the hydrogen atom in parabolic coordinates is illustrated as an example, which is also interpreted as a one-dimensional quantum gravity with a quantized cosmological constant.

1 Introduction

In 1979, Duru and Kleinert [1] showed an elegant path-integral method to exactly evaluate the Green's function for the hydrogen atom. Two basic ingredients in their method are the use of a re-scaled time variable and the so-called Kustaanheimo-Stiefel transformation [2] which reveals the $O(4)$ symmetry explicitly in the coordinate space. The physical meaning of the "re-scaled time variable" however remained somewhat unclear [3]. I have studied this issue by recognizing the procedure in Ref. [1] as a special case of the general treatment of the classically separable Hamiltonian of Liouville-type. The basic dynamical principle involved is then identified as Jacobi's principle of least action for given energy [4].

The path integral on the basis of Jacobi's principle of least action is basically static and analogous to geometrical optics, and one deals with a sum

over orbits in *space* instead of space-time. Another characteristic of the Jacobi principle of least action is that it is reparametrization invariant. The general technique of gauge theory is thus applicable to the evaluation of a path integral, and a suitable choice of gauge simplifies the problem such as the hydrogen atom. A simple trick in parabolic coordinates, which was used before in a different context by Ravndal and Toyoda [5], renders the hydrogen atom Hamiltonian, a separable form of Liouville-type.

The hydrogen atom path integral is also understood as a one-dimensional quantum gravity; an interesting aspect of this picture is that the cosmological constant is *quantized* [6].

2 Separable Hamiltonian of Liouville-Type

We start with a separable Hamiltonian

$$H = \frac{1}{V_1(q_1) + V_2(q_2)} \left\{ \frac{1}{2m} (p_1^2 + p_2^2) + U_1(q_1) + U_2(q_2) \right\}, \quad (1)$$

where the variables vary over $-\infty < q_1, q_2 < \infty$. A general Hamiltonian of Liouville-type is given by

$$H = \frac{1}{V_1(Q_1) + V_2(Q_2)} \times \left\{ \frac{1}{2mW_1(Q_1)} P_1^2 + \frac{1}{2mW_2(Q_2)} P_2^2 + U_1(Q_1) + U_2(Q_2) \right\}, \quad (2)$$

but after a canonical transformation

$$\begin{aligned} \frac{1}{\sqrt{W_1(Q_1)}} P_1 &= p_1, \quad \int_0^{Q_1} \sqrt{W_1(Q)} dQ = q_1, \\ \frac{1}{\sqrt{W_2(Q_2)}} P_2 &= p_2, \quad \int_0^{Q_2} \sqrt{W_2(Q)} dQ = q_2 \end{aligned} \quad (3)$$

and a suitable redefinition of V and U , we can write the Hamiltonian in the form of (1).

We may then solve the Schrödinger problem

$$E\psi = \frac{1}{V_1(q_1) + V_2(q_2)} \left\{ \frac{1}{2m} (\hat{p}_1^2 + \hat{p}_2^2) + U_1(q_1) + U_2(q_2) \right\} \psi, \quad (4)$$

where

$$\hat{p}_l = -i\hbar \frac{\partial}{\partial q_l} \quad (5)$$

for $l = 1, 2$, and where the volume element dV , which renders the Hamiltonian H in (4) Hermitian, is given by

$$dV = (V_1(q_1) + V_2(q_2))dq_1dq_2. \quad (6)$$

The classical Hamiltonian (1) does not completely specify the operator ordering in (4), and the simplest ordering is adopted here. A precise operator ordering needs to be fixed depending on each explicit example.

One may rewrite the above Schrödinger equation (4) as

$$\hat{H}_T \psi = 0 \quad (7)$$

with a *total Hamiltonian* defined by a specific gauge condition,

$$\hat{H}_T = \frac{1}{2m}(\hat{p}_1^2 + \hat{p}_2^2) + U_1(q_1) + U_2(q_2) - E(V_1(q_1) + V_2(q_2)). \quad (8)$$

The meaning of the total Hamiltonian is clarified later. A general procedure to deal with a completely separated operator $\hat{H}_T$ is to consider an evolution operator for a parameter τ defined by

$$\begin{aligned} & \langle q_{1b}, q_{2b} | e^{-i\hat{H}_T \tau / \hbar} | q_{1a}, q_{2a} \rangle \\ &= \left\langle q_{1b} \left| \exp \left[-\frac{i}{\hbar} \left(\frac{1}{2m} \hat{p}_1^2 + U_1(q_1) - EV_1(q_1) \right) \tau \right] \right| q_{1a} \right\rangle \\ & \quad \times \left\langle q_{2b} \left| \exp \left[-\frac{i}{\hbar} \left(\frac{1}{2m} \hat{p}_2^2 + U_2(q_2) - EV_2(q_2) \right) \tau \right] \right| q_{2a} \right\rangle \\ &= \int \mathcal{D}q_1 \mathcal{D}p_1 \exp \left[\frac{i}{\hbar} \int_0^\tau \left\{ p_1 \dot{q}_1 - \left(\frac{1}{2m} p_1^2 + U_1(q_1) - EV_1(q_1) \right) \right\} d\tau \right] \\ & \quad \times \int \mathcal{D}q_2 \mathcal{D}p_2 \exp \left[\frac{i}{\hbar} \int_0^\tau \left\{ p_2 \dot{q}_2 - \left(\frac{1}{2m} p_2^2 + U_2(q_2) - EV_2(q_2) \right) \right\} d\tau \right]. \quad (9) \end{aligned}$$

The parameter τ is arbitrary, and by integrating over τ from 0 to ∞ one obtains a physically meaningful quantity

$$\begin{aligned} & \left\langle q_{1b}, q_{2b} \left| \frac{\hbar}{\hat{H}_T} \right| q_{1a}, q_{2a} \right\rangle_{\text{semi-classical}} \\ &= i \int_0^\infty d\tau \frac{1}{\sqrt{2\pi i \hbar (\partial q_1(\tau) / \partial p_1(0))_{q_{1a}}}} \frac{1}{\sqrt{2\pi i \hbar (\partial q_2(\tau) / \partial p_2(0))_{q_{2a}}}} \end{aligned}$$

$$\times \exp \{ (i/\hbar) S_{\text{cl}}(q_{1b}, q_{1a}, \tau) + (i/\hbar) S_{\text{cl}}(q_{2b}, q_{2a}, \tau) \}. \quad (10)$$

Here, the result is written as a semi-classical approximation for the path integral [7], though in certain cases one may be able to perform an exact path integral in (9). The prefactor in (10) is written in terms of classical paths, for example,

$$q_{1\text{cl}}(\tau) = q_1(\tau; q_{1a}, p_1(0)). \quad (11)$$

This means that the classical paths, dictated by the total Hamiltonian $\hat{H}_T$, are expressed as functions of the initial positions and momenta. On the other hand, the classical action S_{cl} is expressed as a function of the initial position, the final position, and the elapsed "time" τ by eliminating, for example, the dependence on $p_1(0)$. Thus we obtain

$$S_{\text{cl}}(q_{1b}, q_{1a}, \tau) = \int_0^\tau \left\{ p_1 \dot{q}_1 - \left(\frac{1}{2m} p_1^2 + U_1(q_1) - EV_1(q_1) \right) \right\}_{\text{cl}} d\tau \quad (12)$$

with $q_1(\tau) = q_{1b}$. If one solves the Hamilton-Jacobi equation in the form of

$$S(q_{1b}, q_{1a}; \tau) = -A\tau + S(q_{1b}, q_{1a}; A), \quad (13)$$

one treats A as a dynamical variable and regards the above equation as a Legendre transformation defined by

$$\begin{aligned} \frac{\partial S(q_{1b}, q_{1a}; \tau)}{\partial \tau} &= -A, \\ \frac{\partial S(q_{1b}, q_{1a}; A)}{\partial A} &= \tau. \end{aligned} \quad (14)$$

The variable A is then eliminated. This may be regarded as a classical analogue of the uncertainty relation; if one specifies τ , the conjugate variable A becomes implicit. It is well known that the semi-classical approximation (10) is exact for a quadratic system.

We next note the relation for the quantity defined on the left-hand side of (10):

$$\begin{aligned} &\left\langle q_{1b}, q_{2b} \left| \frac{1}{\hat{H}_T} \right| q_{1a}, q_{2a} \right\rangle \\ &= \left\langle q_{1b}, q_{2b} \left| \frac{1}{\left(\frac{1}{\hat{V}_1(q_1) + \hat{V}_2(q_2)} \right) \hat{H}_T} \right| q_{1a}, q_{2a} \right\rangle \frac{1}{V_1(q_{1a}) + V_2(q_{2a})} \end{aligned}$$

$$\begin{aligned}
&= \left\langle q_{1b}, q_{2b} \left| \frac{1}{\hat{H} - E} \right| q_{1a}, q_{2a} \right\rangle \frac{1}{V_1(q_{1a}) + V_2(q_{2a})} \\
&= \frac{1}{H \left(q_{1b}, \frac{\hbar}{i} \frac{\partial}{\partial q_{1b}}, q_{2b}, \frac{\hbar}{i} \frac{\partial}{\partial q_{2b}} \right) - E} \\
&\quad \times \left\{ \frac{1}{\sqrt{V_1(q_{1b}) + V_2(q_{2b})}} \langle q_{1b}, q_{2b} | q_{1a}, q_{2a} \rangle \frac{1}{\sqrt{V_1(q_{1a}) + V_2(q_{2a})}} \right\},
\end{aligned} \tag{15}$$

by recalling $(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1}$. The state vectors in these relations are defined for the volume element $dq_1 dq_2$ as

$$\int dq_1 dq_2 |q_1, q_2\rangle \langle q_1, q_2| = 1. \tag{16}$$

Note that the definition of the δ -function in $\langle q'_1, q'_2 | q_1, q_2 \rangle = \delta(q'_1 - q_1) \delta(q'_2 - q_2)$ depends on the choice of the volume element in (16) and thus on the choice of H_T . The last expression in (15) is thus correctly defined for the original Hamiltonian H and the original state ψ in (4) with the volume element dV in (6), since we have the completeness relation from (16)

$$\int |q_1, q_2\rangle \frac{dV}{V_1(q_1) + V_2(q_2)} \langle q_1, q_2| = 1. \tag{17}$$

The left-hand side of (15) thus defines the correct Green's function for the original operator $(\hat{H} - E)^{-1}$ by noting the symmetry in q_a and q_b . One can then define the conventional evolution operator by

$$\begin{aligned}
&\left\langle q_{1b}, q_{2b} \left| e^{-i\hat{H}(t_b - t_a)/\hbar} \right| q_{1a}, q_{2a} \right\rangle_{\text{conv}} \\
&= \frac{1}{2\pi i \hbar} \int_{-\infty}^{\infty} dE e^{-iE(t_b - t_a)/\hbar} \\
&\quad \times \left\langle q_{1b}, q_{2b} \left| \frac{\hbar}{\hat{H} - i\epsilon - E} \right| q_{1a}, q_{2a} \right\rangle \frac{1}{V_1(q_{1a}) + V_2(q_{2a})},
\end{aligned} \tag{18}$$

where ϵ is an infinitesimal positive number. The total Hamiltonian changes by a different choice of gauge condition in Jacobi's principle of least action to be explained below. Consequently, the volume element, which renders H_T Hermitian, generally depends on the choice of gauge. In the present case, one has the relation

$$\langle q_{1b}, q_{2b} | q_{1a}, q_{2a} \rangle_{\text{conv}}$$

$$\begin{aligned}
&= \langle q_{1b}, q_{2b} | q_{1a}, q_{2a} \rangle \frac{1}{V_1(q_{1a}) + V_2(q_{2a})} \\
&= \frac{1}{\sqrt{V_1(q_{1b}) + V_2(q_{2b})}} \langle q_{1b}, q_{2b} | q_{1a}, q_{2a} \rangle \frac{1}{\sqrt{V_1(q_{1a}) + V_2(q_{2a})}}. \quad (19)
\end{aligned}$$

3 Jacobi's Principle of Least Action

The meaning of the total Hamiltonian H_T (8) becomes transparent if one starts with Jacobi's principle of least action for a given E ,

$$\begin{aligned}
S &= \int_0^\tau d\tau L \quad (20) \\
&= \int_0^\tau d\tau \sqrt{2m[E(V_1(q_1) + V_2(q_2)) - (U_1(q_1) + U_2(q_2))](\dot{q}_1^2 + \dot{q}_2^2)} \\
&= \int \sqrt{2m[E(V_1(q_1) + V_2(q_2)) - (U_1(q_1) + U_2(q_2))][(dq_1)^2 + (dq_2)^2]},
\end{aligned}$$

which is reparametrization invariant. One then defines the momenta conjugate to coordinates

$$\begin{aligned}
p_l &= \frac{\partial L}{\partial \dot{q}_l} \quad (21) \\
&= \sqrt{2m[E(V_1(q_1) + V_2(q_2)) - (U_1(q_1) + U_2(q_2))]} \times \frac{\dot{q}_l}{\sqrt{(\dot{q}_1^2 + \dot{q}_2^2)}}
\end{aligned}$$

and obtains a vanishing Hamiltonian as a result of reparametrization invariance. Furthermore we obtain a first class constraint ϕ as the generator of reparametrization gauge symmetry,

$$\begin{aligned}
H &= p_l \dot{q}_l - L = 0, \\
\phi(q_l, p_l) &= \frac{1}{V_1(q_1) + V_2(q_2)} \left\{ \frac{1}{2m} (p_1^2 + p_2^2) + U_1(q_1) + U_2(q_2) \right\} - E \\
&\simeq 0. \quad (22)
\end{aligned}$$

Following Dirac [8], one may then define a *total Hamiltonian*

$$H_T = H + \alpha(q_l, p_l) \phi(q_l, p_l) = \alpha(q_l, p_l) \phi(q_l, p_l) \simeq 0, \quad (23)$$

where an arbitrary function $\alpha(q_l, p_l)$ specifies a choice of *gauge* or a choice of the arbitrary parameter τ in (20), which parameterizes the orbit for a given

E. The quantum theory is defined, up to an operator ordering, by

$$i\hbar \frac{\partial}{\partial \tau} \psi = \hat{H}_T \psi \quad (24)$$

with the physical state condition

$$\hat{\alpha}(q_l, p_l) \hat{\phi}(q_l, p_l) \psi_{phy} = 0. \quad (25)$$

The choice of the specific gauge $\alpha(q_l, p_l) = V_1(q_1) + V_2(q_2)$ gives rise to (7) and the choice $\alpha(q_l, p_l) = 1$ gives the conventional static Schrödinger equation (4), since ψ appearing in these equations are physical states.

The basic dynamical principle involved is thus identified as the Jacobi principle of least action, which is analogous to geometrical optics, and the formula for an evolution operator (9) dictated by (24) provides a basis for the path-integral approach to a general separable Hamiltonian of Liouville-type. The path integral in (9) deals with a sum over orbits in space instead of space-time, and the notion of re-scaled time does not explicitly appear in the present approach. The evolution operator (9) essentially generates a *gauge transformation*.

4 Hydrogen Atom

It is known that the Schrödinger equation for the hydrogen atom can be written in parabolic coordinates as

$$\begin{aligned} \hat{H}_T \psi &= 0, \\ (\hat{p}_\varphi - \hat{p}_{\varphi'}) \psi &= 0, \end{aligned} \quad (26)$$

with the Hamiltonian

$$\begin{aligned} \hat{H}_T &= \frac{1}{2m} \left[\hat{p}_u^2 + \frac{1}{u^2} \hat{p}_\varphi^2 + \hat{p}_v^2 + \frac{1}{v^2} \hat{p}_{\varphi'}^2 \right] + \frac{m\omega^2}{2} (u^2 + v^2) - e^2 \\ &= \frac{1}{2m} \vec{p}_u^2 + \frac{m\omega^2}{2} \vec{u}^2 + \frac{1}{2m} \vec{p}_v^2 + \frac{m\omega^2}{2} \vec{v}^2 - e^2, \end{aligned} \quad (27)$$

where we define

$$\begin{aligned} \vec{u} &= (u_1, u_2) = (u \cos \varphi, u \sin \varphi), \\ \vec{p}_u^2 &= \hat{p}_u^2 + \frac{1}{u^2} \hat{p}_\varphi^2, \\ \vec{v} &= (v_1, v_2) = (v \cos \varphi', v \sin \varphi'), \end{aligned}$$

$$\vec{p}_v^2 = \hat{p}_v^2 + \frac{1}{v^2} \hat{p}_{\varphi'}^2. \quad (28)$$

The subsidiary condition (26) renders a quadratic Hamiltonian of Liouville-type for the hydrogen atom without recourse to the Kustaanheimo-Stiefel transformation [2]. This introduction of auxiliary variables (28) has been discussed by Ravndal and Toyoda [5].

The general procedure is thus applicable to the hydrogen atom, and we recover the classical exact result of Duru and Kleinert [1].

5 Quantum Gravity with a Quantized Cosmological Constant

An alternative way to see the physical meaning of the parameter τ is to study the one-dimensional quantum gravity coupled to matter variables $\vec{x}$, defined by

$$\int \frac{\mathcal{D}\vec{x} \mathcal{D}h}{\text{gauge volume}} \exp \left\{ \frac{i}{\hbar} \int_0^\tau Lh d\tau \right\}, \quad (29)$$

with

$$L = \frac{m}{2\hbar^2} \left(\frac{d\vec{x}}{d\tau} \right)^2 - V(r) + E, \quad (30)$$

where h stands for the einbein, a one-dimensional analogue of the vierbein h_μ^a , and $h = \sqrt{g}$ in one dimension. If one uses the solution of the equation of motion for h defined by the Lagrangian $\mathcal{L} = Lh$, the action in (29) is reduced to the one appearing in Jacobi's principle of least action. An interesting aspect of this interpretation is that one has a toy model of quantum gravity with a quantized cosmological constant [6].

6 Discussion

A method for exactly solving the Green's function for the hydrogen atom via path integrals [1] opened a new avenue for the path-integral treatment of a general separable Hamiltonian of Liouville-type. This new point of view, which has been shown to be based on Jacobi's principle of least action, provides a more flexible framework of path integrals to deal with a wider class of problems of physical interest. Jacobi's principle of least action, besides being reparametrization invariant, gives a geometrical picture of particle orbits in a

curved space deformed by the potential. On the other hand, the fundamental space-time picture of the conventional Feynman path integral, which is associated with Hamilton's principle of stationary action, is lost.

I learned the path integral of the hydrogen atom through a seminar given by C. Bernido, who provided me with references to the work of Kleinert. Though I have spent much time on the path integral in relativistic quantum field theory, I wrote only two papers on path integrals of non-relativistic quantum mechanics. By publishing these papers [4,6], I was embarrassed to find an enormous difference of *culture* between physicists in relativistic and non-relativistic fields regarding path integrals; the major issue in the latter field of non-relativistic path integral is *operator ordering*, which is not fundamental in relativistic problems mainly due to the strong constraints provided by the symmetry principles such as Lorentz and gauge invariance. An over-emphasis on the operator ordering problem does not appear to be healthy.

I believe that a more fruitful field will open if the researchers in these two realms of path integrals work together. Professor H. Kleinert will certainly be one of the leaders in such an attempt.

References

- [1] I.H. Duru and H. Kleinert, *Phys. Lett. B* **84**, 185 (1979).
- [2] P. Kustaanheimo and E. Stiefel, *J. Reine Angew. Math.* **218**, 204 (1965).
- [3] As for complete references to the related problem, see H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [4] K. Fujikawa, *Nucl. Phys. B* **484**, 495 (1997).
- [5] F. Ravndal and T. Toyoda, *Nucl. Phys. B* **3**, 312 (1967).
- [6] K. Fujikawa, *Prog. Theor. Phys.* **96**, 863 (1996).
- [7] W. Pauli, *Ausgewählte Kapitel aus der Feldquantisierung*, Lecture Notes (Zürich, 1951); P. Choquard, *Helv. Phys. Acta* **28**, 89 (1955).
- [8] P.A.M. Dirac, *Lectures on Quantum Field Theory* (Yeshiva Univ., New York, 1966).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

CONJECTURE ON THE REALITY OF SPECTRA OF NON-HERMITIAN HAMILTONIANS

C.M. BENDER

Department of Physics, Washington University, St. Louis, MO 63130, USA

E-mail: cmb@howdy.wustl.edu

S. BOETTCHER

Department of Physics, Emory University, Atlanta, GA 30322, USA

E-mail: sboettc@emory.edu

P.N. MEISINGER

Department of Physics, Washington University, St. Louis, MO 63130, USA

E-mail: pnm@howdy.wustl.edu

Ordinarily, one demands a Hamiltonian to satisfy $H^\dagger = H$, where $\dagger$ represents the mathematical operation of complex conjugation and matrix transposition. This conventional hermiticity condition ensures that H has a real spectrum. Recently, it was discovered that replacing this mathematical condition by the weaker and more physical requirement $H^\ddagger = H$, where $\ddagger$ represents combined parity reflection and time reversal $\mathcal{PT}$, one obtains new classes of complex Hamiltonians whose spectra are often real and positive. However, the condition of $\mathcal{PT}$ symmetry alone is not strong enough to *guarantee* that the spectrum of H is real. In this paper evidence is presented to support the conjecture that the spectrum of the complex Hamiltonian $H = p^2 + V(x)(ix)^\epsilon$ ($\epsilon > 0$) will be real, positive, and discrete if (1) $V(x)$ is real when x is real; (2) the spectrum of $H = p^2 + V(x)$ is positive and discrete; (3) the function $V(x)$ is even [$V(x) = V(-x)$]; (4) the function $V(x)$ is an *entire* function of complex x . We believe that if these conditions are satisfied then the complex deformation $H = p^2 + V(x)(ix)^\epsilon$ of $H = p^2 + V(x)$ will have a real, positive, and discrete spectrum.

1 Complex Hamiltonians Having Real Spectra

In a recent letter [1], a class of complex quantum mechanical Hamiltonians

$$H = p^2 + x^2(ix)^\epsilon \quad (\epsilon \text{ real}) \quad (1)$$

was investigated. Despite the lack of hermiticity, the spectrum of H is real and positive for all $\epsilon \geq 0$. As shown in Fig. 1 (and Fig. 1 of Ref. [1]) the spectrum is discrete and each of the energy levels increases with increasing ϵ . The reality of the spectrum is in part a consequence of $\mathcal{PT}$ invariance.

Here, the operators $\mathcal{P}$ and $\mathcal{T}$ represent parity reflection and time reversal. These operators are defined by their action on the position and momentum operators x and p :

$$\begin{aligned} \mathcal{P} : x &\rightarrow -x, & p &\rightarrow -p, \\ \mathcal{T} : x &\rightarrow x, & p &\rightarrow -p, & i &\rightarrow -i. \end{aligned} \quad (2)$$

When the operators x and p are real, the canonical commutation relation $[x, p] = i$ is invariant under both parity reflection and time reversal. This commutation relation remains invariant under $\mathcal{P}$ and $\mathcal{T}$ even if x and p are complex provided that the above transformations hold [2].

The spectrum of the Hamiltonian (1) is obtained by solving the corresponding Schrödinger equation

$$-\psi''(x) + [x^2(ix)^\epsilon - E]\psi(x) = 0 \quad (3)$$

subject to appropriate boundary conditions imposed in the complex- x plane. These boundary conditions are described in Ref. 1.

Note from Fig. 1 that the spectrum of the Hamiltonian (1) exhibits two regions: When $\epsilon \geq 0$, the energy spectrum of H is real and positive. However, a transition occurs at $\epsilon = 0$. As ϵ goes below 0, the eigenvalues as functions of ϵ pair off and become complex, starting with the highest-energy eigenvalues. For negative ϵ the spectrum contains an infinite number of complex eigenvalues and a finite number of real, positive eigenvalues. As ϵ decreases, there are fewer and fewer real eigenvalues and below approximately $\epsilon = -0.57793$ only one real energy remains. This energy then begins to increase with decreasing ϵ and becomes infinite as ϵ approaches -1 .

In summary, the theory defined by Eq. (1) exhibits two phases, an unbroken-symmetry phase with a purely real energy spectrum when $\epsilon \geq 0$ and a spontaneously-broken-symmetry phase with a partly real and partly

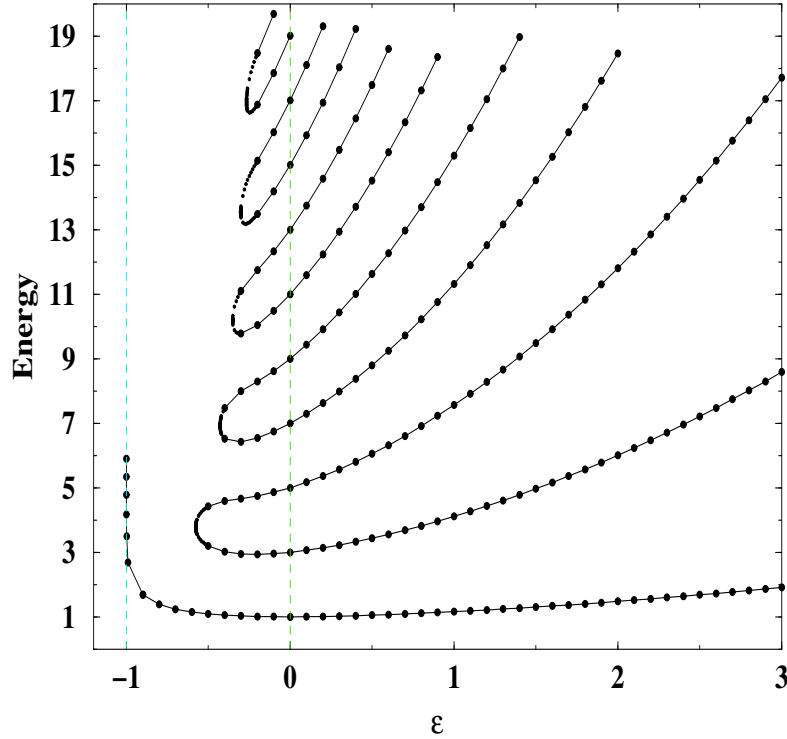


Figure 1. Energy levels of the Hamiltonian $H = p^2 + x^2(ix)^\epsilon$ as a function of the parameter ϵ . When $\epsilon \geq 0$, the spectrum is real and positive and the energy levels rise with increasing ϵ . The lower bound of this region, $\epsilon = 0$, corresponds to the harmonic oscillator, whose energy levels are $E_n = 2n + 1$. When $-1 < \epsilon < 0$, there are a finite number of real positive eigenvalues and an infinite number of complex conjugate pairs of eigenvalues. As ϵ decreases from 0 to -1 , the number of real eigenvalues decreases; when $\epsilon \leq -0.57793$, the only real eigenvalue is the ground-state energy. As ϵ approaches -1^+ , the ground-state energy diverges. For $\epsilon \leq -1$ there are no real eigenvalues.

complex spectrum when $-1 < \epsilon < 0$, the transition at $\epsilon = 0$ can be seen in both the quantum mechanical system and the underlying classical system [2]. We have numerically verified that the eigenfunctions of H in Eq. (1) are also eigenfunctions of the operator $\mathcal{PT}$ when $\epsilon \geq 0$ [2]. However, when $\epsilon < 0$, the $\mathcal{PT}$ symmetry of the Hamiltonian is spontaneously broken; even though $\mathcal{PT}$ commutes with H , the eigenfunctions of H are *not* all simultaneously eigenfunctions of $\mathcal{PT}$. For these eigenfunctions of H the energies are complex.

Similar qualitative features are exhibited by complex deformations of real

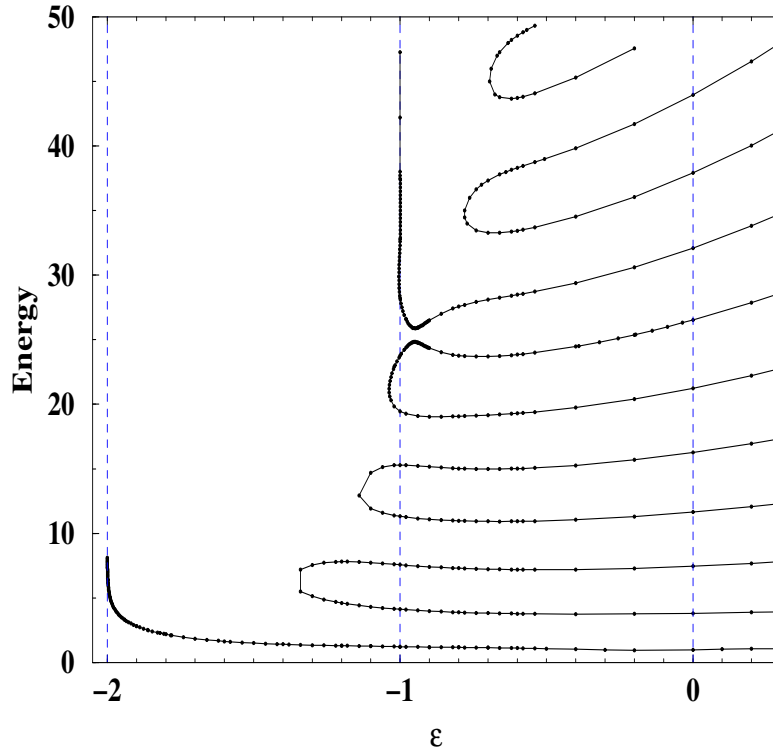


Figure 2. Energy levels of the Hamiltonian $H = p^2 + x^4(ix)^\epsilon$ as a function of the parameter ϵ . This figure is similar to Fig. 1, but now there are four regions. When $\epsilon \geq 0$, the spectrum is real and positive and it rises monotonically with increasing ϵ . Below $\epsilon = 0$ there are complex eigenvalues (except at $\epsilon = -1$).

Hamiltonians other than the harmonic oscillator. The class of Hamiltonians

$$H = p^2 + x^{2N}(ix)^\epsilon \quad (4)$$

have the same qualitative properties as H in Eq. (1). As ϵ decreases below 0, all of these theories exhibit a transition from an unbroken $\mathcal{PT}$ -symmetric regime to a regime in which $\mathcal{PT}$ symmetry is spontaneously broken. Each of these $\mathcal{PT}$ -symmetric theories may be viewed as analytic continuations of conventional Hermitian theories from real to complex phase space [2].

Consider, for example, the spectrum of an $x^4(ix)^\epsilon$ theory ($N = 2$), which is displayed in Fig. 2. This figure resembles Fig. 1 for the case $N = 1$ except that now there are four regions: When $\epsilon \geq 0$, the spectrum is discrete, real, and positive and it rises monotonically with increasing ϵ . The lower

bound $\epsilon = 0$ of this $\mathcal{PT}$ -symmetric region corresponds to the pure quartic anharmonic oscillator, whose Hamiltonian is $H = p^2 + x^4$. When $-1 < \epsilon < 0$, $\mathcal{PT}$ symmetry is spontaneously broken. There are a finite number of real positive eigenvalues and an infinite number of complex conjugate pairs of eigenvalues; as a function of ϵ the eigenvalues pinch off in pairs and move into the complex plane. As ϵ approaches -1 from above only eight real eigenvalues remain. Just as ϵ reaches -1 the entire spectrum reemerges from the complex plane and becomes real. Note that at $\epsilon = -1$ the entire spectrum agrees with the entire spectrum in Fig. 1 at $\epsilon = 1$. This reemergence is difficult to see in this figure. Just below $\epsilon = -1$, the eigenvalues once again begin to pinch off and disappear in pairs into the complex plane. However, this pairing is different from the pairing in the region $-1 < \epsilon < 0$. Above $\epsilon = -1$ the lower member of a pinching pair is even and the upper member is odd (that is, E_8 and E_9 combine, E_{10} and E_{11} combine, and so on); below $\epsilon = -1$ this pattern reverses (that is, E_7 combines with E_8 , E_9 combines with E_{10} , and so on). As ϵ decreases from -1 to -2 , the number of real eigenvalues continues to decrease until the only real eigenvalue is the ground-state energy. Then, as ϵ approaches -2^+ , the ground-state energy diverges logarithmically. For $\epsilon \leq -2$ there are no real eigenvalues.

The spectrum for $H = p^2 + x^6(ix)^\epsilon$ ($N = 3$) is displayed in Fig. 3. This figure resembles Fig. 2 for the case $N = 2$. However, now there are transitions at both $\epsilon = -1$ and $\epsilon = -2$. The spectrum is discrete, real, and positive when $\epsilon \geq 0$ and the energy levels rise monotonically with increasing ϵ . The lower bound $\epsilon = 0$ of this $\mathcal{PT}$ -symmetric region corresponds to the sextic anharmonic oscillator, whose Hamiltonian is $H = p^2 + x^6$. The other four regions are $-1 < \epsilon < 0$, $-2 < \epsilon < -1$, $-3 < \epsilon < -2$, and $\epsilon < -3$. The $\mathcal{PT}$ symmetry is spontaneously broken when ϵ is negative, and the number of real eigenvalues decreases as ϵ becomes more negative. However, at the boundaries $\epsilon = -1$, -2 there is a complete real positive spectrum. When $\epsilon = -1$, the eigenspectrum is identical to the eigenspectrum in Fig. 2 at $\epsilon = 1$. For $\epsilon \leq -3$ there are no real eigenvalues.

While there is as yet no proof that the spectrum of H in Eq. (1) is real, it is possible to gain insight regarding the reality and positivity of the spectrum of a $\mathcal{PT}$ -invariant Hamiltonian H by calculating the spectral zeta function. An exact calculation of the zeta function was done for the case $\epsilon = 3$ by Mezincescu [3] and this work was generalized to arbitrary $\epsilon > 0$ by Bender and Wang [4]. Other work has been done by Delabaere *et al.* [5]

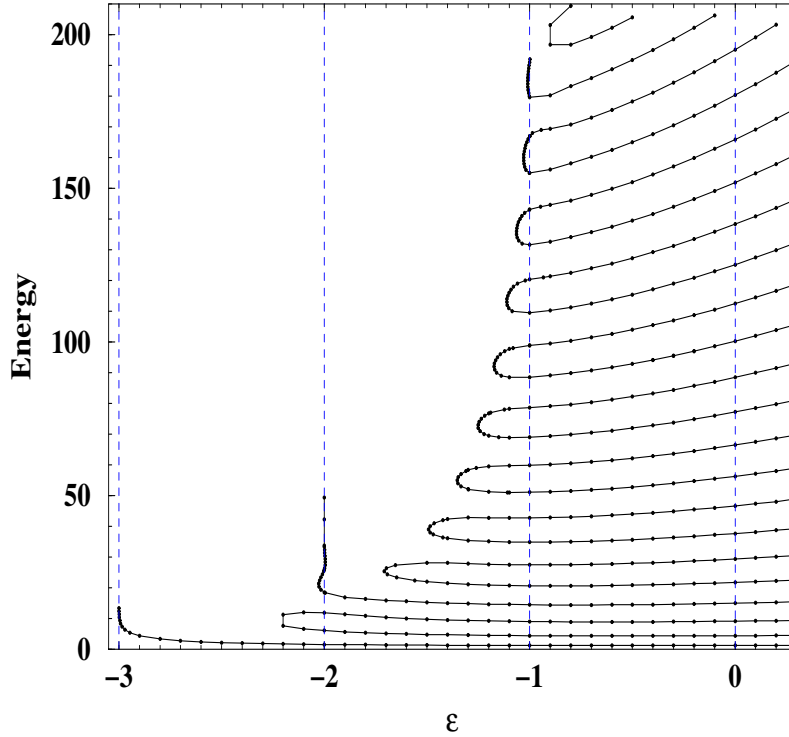


Figure 3. Energy levels of the Hamiltonian $H = p^2 + x^6(ix)^\epsilon$ as a function of the parameter ϵ . This figure is similar to Fig. 2, but now there are five regions. When $\epsilon \geq 0$, the spectrum is real and positive and it rises monotonically with increasing ϵ . Below $\epsilon = 0$ the spectrum is complex except when $\epsilon = -1$ and $\epsilon = -2$.

2 Complex Deformations of Nonanalytic Potentials

$\mathcal{PT}$ symmetry alone is not sufficient to guarantee that the spectrum of a Hamiltonian is real. We conjecture that analyticity of the underlying real potential is also required. To illustrate this, we study complex deformations of some nonanalytic potentials. In our discussion so far we have considered complex deformations of the potentials x^{2N} , which are entire functions of x . Let us now consider deformations of the *nonanalytic* potentials $|x|^P$ (P real). We will see that the eigenvalues of the potentials $|x|^P(ix)^\epsilon$ are real only when $\epsilon = 0$ (and sometimes at other isolated values of ϵ). Thus, it appears that deforming a nonanalytic potential destroys a crucial property of the theory; namely, that the spectrum be real.

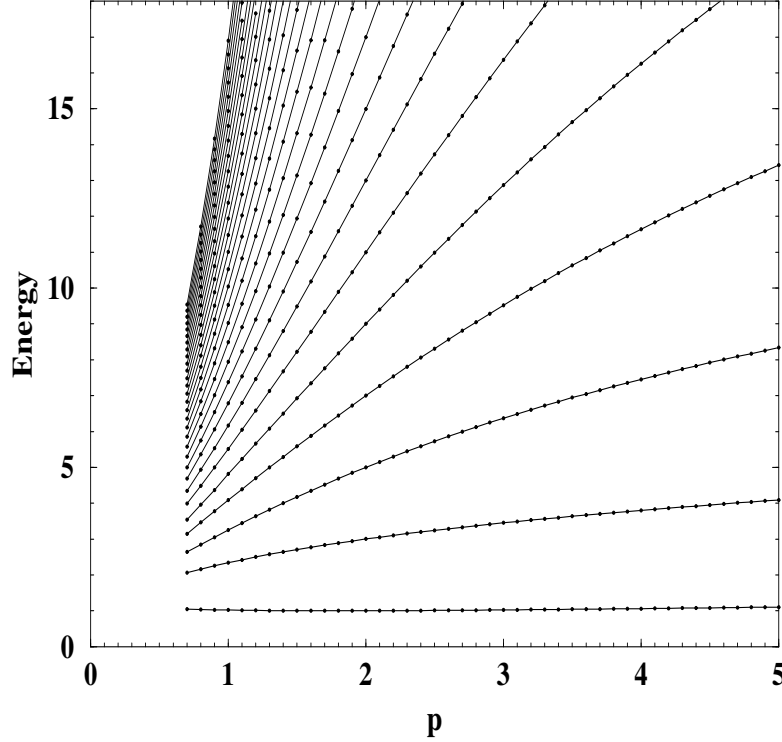


Figure 4. Energy levels of the Hamiltonian $H = p^2 + |x|^P$ as a function of the parameter P . This figure is similar to Fig. 1, but the eigenvalues do not pinch off and go into the complex plane because H is Hermitian (instead, the spectrum becomes denser as P approaches 0).

We begin by examining the spectrum of the $|x|^P$ potential. The eigenvalues of this potential are displayed as a function of P in Fig. 4 (see Ref. [6]). The spectrum of this potential is similar to that of the $x^2(ix)^\epsilon$ potential for positive ϵ (see Fig. 1). The difference between the spectra of these two potentials becomes apparent when ϵ is large: As $\epsilon \rightarrow \infty$, the spectrum of $|x|^{2+\epsilon}$ approaches that of the square-well potential $[E_n = (n+1)^2\pi^2/4]$, while the energies of the $x^2(ix)^\epsilon$ potential diverge [7].

WKB theory gives an excellent approximation to the spectrum of both real and complex potentials and thus provides an interesting comparison. For the $x^2(ix)^\epsilon$ potential, when $\epsilon \geq 0$, the WKB calculation must be performed in the complex plane [8]. The turning points $x_\pm$ are the roots of $E - x^2(ix)^\epsilon = 0$

that *analytically continue* off the real axis as ϵ moves away from zero:

$$x_- = E^{\frac{1}{2+\epsilon}} e^{i\pi \frac{4+3\epsilon}{4+2\epsilon}}, \quad x_+ = E^{\frac{1}{2+\epsilon}} e^{-i\pi \frac{\epsilon}{4+2\epsilon}}. \quad (5)$$

These turning points lie in the lower-half (upper-half) x -plane when $\epsilon > 0$ ($\epsilon < 0$).

The WKB phase-integral quantization condition to leading order is $(n + 1/2)\pi = \int_{x_-}^{x_+} dx \sqrt{E - x^2(ix)^\epsilon}$. It is crucial that this integral follows a path along which the *integral is real*. When $\epsilon > 0$, this path lies entirely in the lower-half x -plane and when $\epsilon = 0$ the path lies on the real axis. But, when $\epsilon < 0$ the path is in the upper-half x -plane; it crosses the cut on the positive imaginary axis and thus is *not a continuous path joining the turning points*. Hence, WKB fails when $\epsilon < 0$.

When $\epsilon \geq 0$, we deform the phase-integral contour so that it follows the rays from x_- to 0 and from 0 to x_+ :

$$\left(n + \frac{1}{2}\right)\pi = 2 \sin[\pi/(2 + \epsilon)] E^{\frac{4+\epsilon}{4+2\epsilon}} \int_0^1 ds \sqrt{1 - s^{2+\epsilon}}.$$

We then solve for E_n :

$$E_n \sim \left[\frac{\Gamma\left(\frac{8+3\epsilon}{4+2\epsilon}\right) \sqrt{\pi}(n + 1/2)}{\sin\left(\frac{\pi}{2+\epsilon}\right) \Gamma\left(\frac{3+\epsilon}{2+\epsilon}\right)} \right]^{\frac{4+2\epsilon}{4+\epsilon}} \quad (n \rightarrow \infty). \quad (6)$$

To perform a higher-order WKB calculation we replace the phase integral by a *closed contour* that encircles the path connecting the two turning points (see Ref. 9). With $Q(x) = x^2(ix)^\epsilon - E$, the next-to-leading-order WKB quantization condition is

$$\left(n + \frac{1}{2}\right)\pi = \frac{1}{2i} \oint_C dx \sqrt{Q(x)} + \frac{1}{2i} \oint_C dx \frac{Q''(x)}{48Q(x)^{\frac{3}{2}}}, \quad (7)$$

where the contour C encircles the turning points x_+ and x_- in a counter-clockwise direction. Assuming that n is large, we obtain

$$E_n \sim \left[\frac{\Gamma\left(\frac{8+3\epsilon}{4+2\epsilon}\right) \sqrt{\pi}(n + 1/2)}{\sin\left(\frac{\pi}{2+\epsilon}\right) \Gamma\left(\frac{3+\epsilon}{2+\epsilon}\right)} \right]^{\frac{4+2\epsilon}{4+\epsilon}} \left[1 + \frac{(2 + \epsilon)(1 + \epsilon) \sin\left(\frac{2\pi}{2+\epsilon}\right)}{6\pi \left(n + \frac{1}{2}\right)^2 (4 + \epsilon)^2} \right]. \quad (8)$$

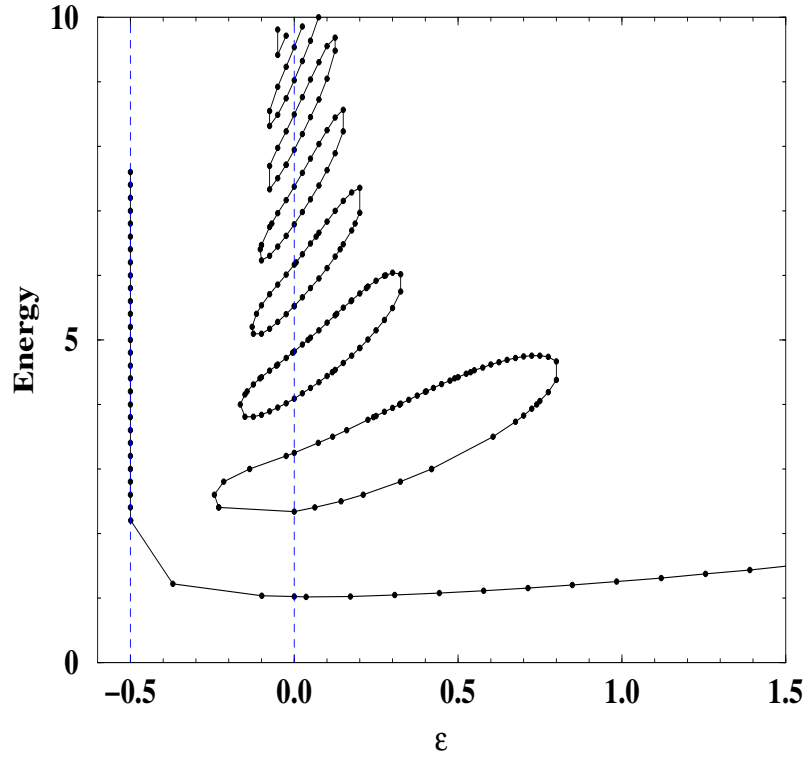


Figure 5. Energy levels of the Hamiltonian $H = p^2 + |x|(ix)^\epsilon$ as a function of the parameter ϵ . The spectrum is entirely real only when $\epsilon = 0$.

This is the next-to-leading-order WKB result for the energy. The corresponding WKB result for the $|x|^{2+\epsilon}$ ($\epsilon > -2$) potential is quite similar:

$$E_n \sim \left[\frac{\Gamma\left(\frac{8+3\epsilon}{4+2\epsilon}\right) \sqrt{\pi}(n+1/2)}{\Gamma\left(\frac{3+\epsilon}{2+\epsilon}\right)} \right]^{\frac{4+2\epsilon}{4+\epsilon}} \left[1 + \frac{(2+\epsilon)(1+\epsilon) \cot\left(\frac{\pi}{2+\epsilon}\right)}{3\pi \left(n + \frac{1}{2}\right)^2 (4+\epsilon)^2} \right]. \quad (9)$$

Now we perform a complex deformation of the $|x|^P$ potential. That is, we consider an $|x|^P(ix)^\epsilon$ potential. Of course, since $|x|$ is not an analytic function, we cannot define an analytic continuation of the Schrödinger eigenvalue problem

$$-\psi''(x) + |x|^P(ix)^\epsilon \psi(x) = E\psi(x) \quad (10)$$

into the complex x -plane. However, for sufficiently small ϵ we can allow x

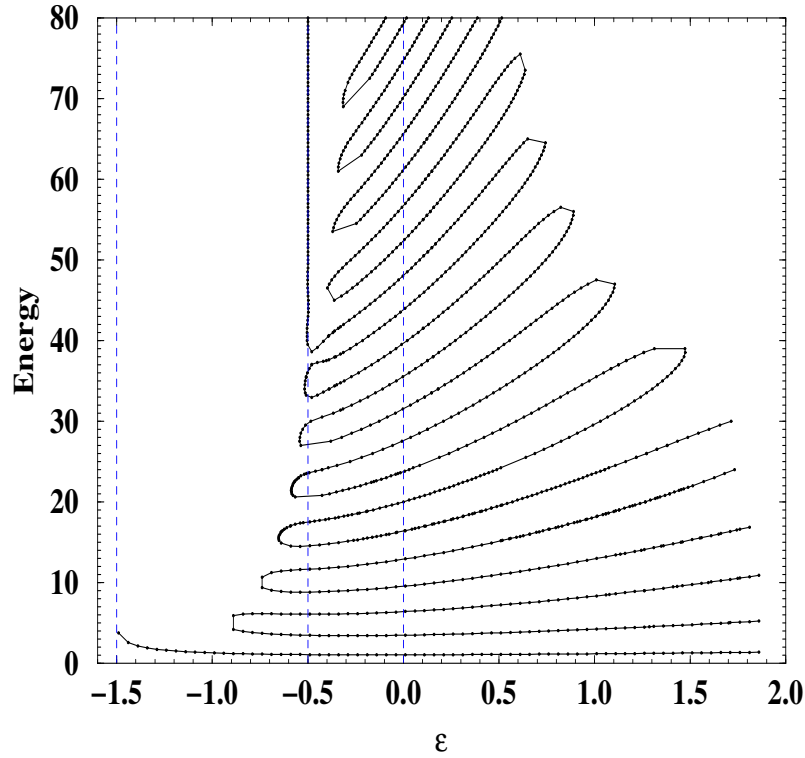


Figure 6. Energy levels as a function of the parameter ϵ for the Hamiltonian $H = p^2 + |x|^3(ix)^\epsilon$. The spectrum is real when $\epsilon = 0$ and $\epsilon = -0.5$.

to remain *real* and we can impose the boundary condition that $\psi(x) \rightarrow 0$ for $x \rightarrow \pm\infty$. Specifically, we have the condition for *all* P that if $|\epsilon| < 2$, then this boundary condition on the real x -axis may be consistently imposed to define the eigenspectrum.

Let us consider two cases: $P = 1$ (Fig. 5) and $P = 3$ (Fig. 6). Figure 5 is quite similar to Fig. 1, and Fig. 6 resembles Fig. 2. The key features of these figures are that (1) the lowest energy level diverges at $\epsilon = -P/2$, and that (2) the energy levels pinch off and move into the complex plane on *both* sides of $\epsilon = 0$. Thus, the spectrum is entirely real only when $\epsilon = 0$.

Acknowledgments

We are grateful to the U.S. Department of Energy for financial support.

References

- [1] C.M. Bender and S. Boettcher, *Phys. Rev. Lett.* **80**, 5243 (1998). In this reference the notation N represents the quantity $2 + \epsilon$.
- [2] C.M. Bender, S. Boettcher, and P.N. Meisinger, *J. Math. Phys.* **40**, 2201 (1999).
- [3] G.A. Mezincescu, *J. Phys. A: Math. Gen.* **33**, 4911 (2000). See also the references therein.
- [4] C.M. Bender and Q. Wang, submitted to *J. Phys. A: Math. Gen.*
- [5] See E. Delabaere and F. Pham, *Phys. Lett. A* **250**, 25 (1998) and **250**, 29 (1998); E. Delabaere and D.T. Trinh, Univ. of Nice preprint, 1999.
- [6] S. Boettcher and C.M. Bender, *J. Math. Phys.* **31**, 2579 (1990).
- [7] C.M. Bender, S. Boettcher, H.F. Jones, and V.M. Savage, *J. Phys. A: Math. Gen.* **32**, 1 (1999).
- [8] C.M. Bender and A. Turbiner, *Phys. Lett. A* **173**, 442 (1993).
- [9] C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978), Chap. 10.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

TIME-TRANSFORMATION APPROACH TO Q -DEFORMED OBJECTS

A. INOMATA

Department of Physics, SUNY-Albany, Albany, NY 12222, USA

E-mail: inomata@albany.edu

Time-transformations in path integrals are revisited. In particular the time-dependent coordinate transformation associated with the global time-transformation is discussed and applied for deriving the propagator of a q -deformed object.

1 Introduction

Duru and Kleinert [1] were the first to publish a paper on the time-transformation applied to a path integral. More specifically, they calculated a path integral for the hydrogen atom by using the time-transformation of Kustaanheimo and Stiefel [2] and succeeded for the first time in obtaining the Coulomb propagator from Feynman's path integral. Feynman asserted in his 1949 paper [3] by deriving Schrödinger's equation from the path integral that the path integral approach is equivalent to the standard canonical approach. However, the path integral remained incapable of solving the hydrogen atom problem until Duru and Kleinert made a breakthrough by bringing the time-transformation into the path-integral calculation.

After the success in the Coulomb problem, many authors have employed various time-transformations to solve other problems which can be solved by Schrödinger's equation but had not been solved by path integration [4]. Feynman's path-integral method, if time-transformations are appropriately used, is indeed capable of producing exact solutions for many integrable systems. In this manner, Feynman's approach is shown to be as powerful as the standard canonical approach in obtaining exact results. However, it is

as yet unclear whether the time-transformation technique is effectively applicable to other nontrivial problems. In this paper, we attempt to utilize the time-transformation technique in dealing with q -deformation.

Time-transformations, $t \rightarrow s$, pertinent to path integrals in nonrelativistic quantum mechanics, may be expressed in the form,

$$ds = F[\mathbf{r}(t), t] dt, \quad (1)$$

which may be classified into two types: global (holonomic) transformations and local (non-holonomic) transformations [5].

A holonomic time-transformation with $F = dg^{-1}/dt$ takes the time parameter t into a new "time" parameter s by

$$s = g^{-1}(t) \quad \text{or} \quad t = g(s). \quad (2)$$

In this case, the new parameter s is globally meaningful as the time parameter t is. An important example is the time-transformation considered by de Alfaro, Fubini, and Furlan [6], and by Jackiw [7] in combination with a time-dependent conformal coordinate transformation which will be discussed in Section 2.

A non-holonomic time-transformation for which Eq. (1) may be integrable along a certain path is generally significant only for an infinitesimal time interval. The time-transformation of Kustaanheimo and Stiefel [2], used for the path integral of the hydrogen atom, is a prototype of local time-transformations. The transformations used in carrying out path integration for exact results are mainly of the non-holonomic type.

In Section 2, we focus our attention on the global time-transformation in combination with a time-dependent conformal coordinate transformation. In Section 3 we use time-transformations to analyze q -deformed objects and derive the propagators of the q -deformed free particle and of the pulsed harmonic oscillator.

2 Time-Dependent Conformal Transformations

The time-dependent conformal transformation (TDCT) is an isotropic scale transformation of coordinates from $\mathbf{r}$ to $\mathbf{R}$:

$$\mathbf{r}(t) = \mathbf{R}(s) f(s), \quad (3)$$

associated with a global time-transformation

$$t = g(s). \quad (4)$$

In the above we assume that $g(s)$ and $f(s)$ are both well-behaved single-valued functions of s and mutually related by

$$f^2(s) = \overset{\circ}{g}(s), \quad (5)$$

where $\overset{\circ}{g} = dg/ds$.

To see how TDCT works in a path integral [8], we consider a particle of charge e and mass M under the influence of a scalar potential $V(\mathbf{r})$ and a vector potential $\mathbf{A}(\mathbf{r})$ such that $\mathbf{A}(\mathbf{r}) \cdot \mathbf{r} = 0$. Upon application of TDCT, the action integral for the system,

$$S(t'', t') = \int_{t'}^{t''} \left[\frac{1}{2} M \dot{\mathbf{r}}^2 + \frac{e}{c} \mathbf{A}(\mathbf{r}) \cdot \dot{\mathbf{r}} - V(\mathbf{r}) \right] dt, \quad (6)$$

transforms into

$$S(t'', t') = S_0(s'', s') + \tilde{S}(s'', s'), \quad (7)$$

where $s' = g^{-1}(t')$ and $s'' = g^{-1}(t'')$. The first term is a quantity depending only on the end point values,

$$S_0(s'', s') = \frac{1}{2} \left[M \overset{\circ}{\mathbf{R}}^2 \frac{\overset{\circ}{f}}{f} \right]_{s'}^{s''}, \quad (8)$$

which does not contribute to the equation of motion. The second term is the principal part of the new action,

$$\tilde{S}(s'', s') = \int_{s'}^{s''} \left[\frac{1}{2} M \overset{\circ}{\mathbf{R}}^2 + \frac{e}{c} \tilde{\mathbf{A}} \cdot \overset{\circ}{\mathbf{R}} - \tilde{V}(\mathbf{R}, s) \right] ds \quad (9)$$

with

$$\tilde{\mathbf{A}}(\mathbf{R}, s) = f(s) \mathbf{A}(f\mathbf{R}, g), \quad (10)$$

$$\tilde{V}(\mathbf{R}, s) = f^2(s) V(\mathbf{r}) + \frac{1}{2} M \omega^2(s) \mathbf{R}^2, \quad (11)$$

$$\omega^2(s) = (\overset{\circ\circ}{f} f - 2 \overset{\circ}{f}^2) / f^2. \quad (12)$$

If the original vector potential is given in the form $\mathbf{A}(\mathbf{r}) = \mathbf{a}(\theta, \phi)/r$ with $\mathbf{a} \cdot \mathbf{r} = 0$, the vector interaction term remains form-invariant: $\mathbf{A} \cdot d\mathbf{r} = \tilde{\mathbf{A}} \cdot d\mathbf{R}$. As is apparent from (11), TDCT, when applied to the scalar potential $V(\mathbf{r})$, generates an extra term which has the form of the time-dependent harmonic oscillator. Thus, for instance, if both the vector and scalar potentials are absent, TDCT converts the action of the free particle into that of a time-dependent harmonic oscillator. Or conversely, the inverse of TDCT reduces the harmonic oscillator to a free particle.

Next we take this transformation scheme to quantum mechanics via the path integral. The propagator corresponding to the action (6) is

$$K(\mathbf{r}'', t''; \mathbf{r}', t') = \int_{\mathbf{r}'=\mathbf{r}(t')}^{\mathbf{r}''=\mathbf{r}(t'')} e^{iS/\hbar} \mathcal{D}\mathbf{r}. \quad (13)$$

It can be shown explicitly by the time-sliced calculation [8] that TDCT changes this propagator into

$$K(\mathbf{r}'', t''; \mathbf{r}', t') = (f' f'')^{-3/2} \exp[iS_0(s'', s')/\hbar] \tilde{K}(\mathbf{R}'', s''; \mathbf{R}', s'), \quad (14)$$

where

$$\tilde{K}(\mathbf{R}'', s''; \mathbf{R}', s') = \int_{\mathbf{R}'=\mathbf{R}(s')}^{\mathbf{R}''=\mathbf{R}(s'')} e^{i\tilde{S}/\hbar} \mathcal{D}\mathbf{R}. \quad (15)$$

Therefore the transformed propagator can be calculated directly by substituting (3)-(5) as

$$\tilde{K}(\mathbf{R}'', s''; \mathbf{R}', s') = (f' f'')^{3/2} \exp[-iS_0(s'', s')/\hbar] K(\mathbf{r}'', t''; \mathbf{r}', t')|_{TDCT}. \quad (16)$$

As an example, let us take the Alfaro-Fubini-Furlen-Jackiw (AFFJ) transformation [6,7]:

$$\mathbf{r}(t) = \mathbf{R}(s) \sec(\omega s), \quad t = \frac{\tan(\omega s)}{\omega}. \quad (17)$$

Apparently $f(s) = \sec(\omega s)$ and $g(s) = \omega^{-1} \tan(\omega s)$ satisfy the condition (5). Under this transformation, the action integral for a free particle ($\mathbf{A} = 0$ and $V = 0$) takes the form $S_0(s'', s') + \tilde{S}(s'', s')$ with

$$S_0(s'', s') = \frac{1}{2} M [\mathbf{R}^2(s'') - \mathbf{R}^2(s')] \quad (18)$$

and

$$\tilde{S}(s'', s') = \int_{s'}^{s''} \left[\frac{1}{2} M \dot{\mathbf{R}}^2 - \frac{1}{2} M \omega^2 \mathbf{R}^2 \right] ds. \quad (19)$$

The last action is the one for the simple harmonic oscillator. Since S_0 does not contribute to the equation of motion, the AFFJ transformation (17) converts the free particle into a harmonic oscillator [9]. The direct substitution of (17) via the formula (16) into the free particle propagator,

$$K^{(\text{free})}(\mathbf{r}'', t''; \mathbf{r}', t') = \left[\frac{M}{2\pi i \hbar (t'' - t')} \right]^{3/2} \exp \left[\frac{iM(\mathbf{r}'' - \mathbf{r}')^2}{2\hbar(t'' - t')} \right], \quad (20)$$

leads to the propagator of the harmonic oscillator,

$$\begin{aligned} \tilde{K}^{(\text{osc})}(\mathbf{R}'', s''; \mathbf{R}', s') &= \left[\frac{M\omega}{2\pi i \hbar \sin[\omega(s'' - s')]} \right]^{3/2} \\ &\times \exp \left[\frac{iM\omega}{2\hbar \sin[\omega(s'' - s')]} \left\{ (\mathbf{R}'^2 + \mathbf{R}''^2) \cos[\omega(s'' - s')] - 2\mathbf{R}'\mathbf{R}'' \right\} \right]. \end{aligned} \quad (21)$$

In this process, we have utilized the following identities [9],

$$\frac{A \sec^2 \alpha + B \sec^2 \beta}{\tan \alpha - \tan \beta} = (A + B) \cot(\alpha - \beta) + A \tan \alpha - B \tan \beta, \quad (22)$$

$$\frac{\sec \alpha \sec \beta}{\tan \alpha - \tan \beta} = \csc(\alpha - \beta). \quad (23)$$

Other examples can be found in Refs. [8,10].

3 Global Time-Transformation in q -Deformation

In this section, we study global time-transformations related to q -deformation.

3.1 The q -Free Particle

Let us start with the Newtonian free particle in one dimension which obeys the difference equation

$$x(t+T) - 2x(t) + x(t-T) = 0, \quad (24)$$

where T is any finite-time period. Obviously its solution is given by

$$x(t) = at + b, \quad (25)$$

where a and b are constants.

Now we apply to the free particle the following transformation [11],

$$x(t) \rightarrow y(\tau) = x(\tau), \quad t \rightarrow \tau = q^{2t/T}, \quad (26)$$

where $q \neq 0$ and $q \neq 1$. The new variable,

$$y(\tau(t)) = a\tau + b = aq^{2t/T} + b, \quad (27)$$

does not satisfy (24) any longer but obeys the equation

$$q^{-1}y(q^2\tau) - (q + q^{-1})y(\tau) + qy(q^{-2}\tau) = 0. \quad (28)$$

While the difference equation (24) dictates the time-evolution of $x(t)$ under the discrete time-translation $t - T \rightarrow t \rightarrow t + T$, the new difference equation (28) stipulates the progression of $y(\tau)$ under the time-scaling $q^{-2}\tau \rightarrow \tau \rightarrow q^2\tau$.

In fact, the difference equation (28) is equivalent to the q -deformation counterpart of the force-free Newton equation:

$$\frac{D^2y(\tau)}{D\tau^2} = 0, \quad (29)$$

where $Dy(\tau)/D\tau$ is the symmetric Jackson q -derivative defined by [12]

$$\frac{Dy(\tau)}{D\tau} = \frac{y(q\tau) - y(q^{-1}\tau)}{(q - q^{-1})\tau}. \quad (30)$$

Thus the scaling factor q appearing in Eq. (26) turns out to be the same as the deformation parameter.

In this way, the time-transformation (26) converts the Newtonian free particle to the q -deformed object obeying Eq. (28) or Eq. (29) which we call the q -deformed free particle or the q -free particle in short.

In order for the new time parameter τ to be real for any t , the deformation parameter q must be a positive real number. In this case, the trajectory described by (27) is also real and continuous. Thus the q -free particle moving along a real continuous trajectory is characterized by a positive real parameter $q \in \mathbf{R}^+$. If, however, the object is allowed only to hop along a real discrete sequential trajectory $\{y(s(nT))\}$ associated with discrete periodic

time-translations $t = nT$ ($n \in \mathbf{N}$), then q may be negative. Thus, for the proper q -free particle and the hopping q -free particle we have $q \in \mathbf{R}/(0, 1)$.

If we let $x_n = x(t_0 + nT)$ ($n \in \mathbf{N}_0$), then Eq. (24) may be written as the recursion relation

$$x_{n+1} - 2x_n + x_{n-1} = 0. \quad (31)$$

Correspondingly, defining $\tau_n = q_{2n}\tau_0$ with $\tau_0 = q^{2t_0/T}$, we obtain the recursion relation obeyed by $y_n = y(\tau_n)$:

$$q^{-1}y_{n+1} - (q + q^{-1})y_n + qy_{n-1} = 0. \quad (32)$$

3.2 The q -Object

Next we apply the time-dependent coordinate transformation

$$y(\tau) \rightarrow Q(\tau) = \tau^{-1/2}y(\tau) \quad (33)$$

to the q -free particle. Under this scale transformation, the difference equation (28) becomes

$$Q(q^2\tau) - (q + q^{-1})Q(\tau) + Q(q^{-2}\tau) = 0. \quad (34)$$

As is evident from Eqs. (27) and (33), the solution for this equation is given by

$$Q(\tau) = a\tau^{1/2} + b\tau^{-1/2}, \quad (35)$$

which may be rewritten as

$$Q(\tau) = \left(A \frac{\tau^{1/2} - \tau^{-1/2}}{\tau^{1/2} + \tau^{-1/2}} + B \right) (\tau^{1/2} + \tau^{-1/2}), \quad (36)$$

where $A = (a - b)/2$ and $B = (a + b)/2$.

Then it becomes clear that $Q(\tau)$ may be related to the coordinate variable $x(t)$ of the Newtonian free particle via a time-dependent scale transformation $x(t) \rightarrow Q(s)$ associated with the time-transformation $t \rightarrow s \in \mathbf{R}$:

$$x(t) = 2[(c/T) \ln q]^{1/2} Q(s) (q^{s/T} + q^{-s/T})^{-1}, \quad t = c \frac{q^{s/T} - q^{-s/T}}{q^{s/T} + q^{-s/T}}, \quad (37)$$

where

$$Q(s) = Q(\tau(s)) = \left(A \frac{q^{s/T} - q^{-s/T}}{q^{s/T} + q^{-s/T}} + B \right) (q^{s/T} + q^{-s/T}), \quad (38)$$

and c is a constant to be chosen such that the condition (5) is met.

Although we have assumed $q \neq 1$ in Eq. (26), we may remove the assumption if we define $Q(s)$ by Eq. (37). It is obvious from Eq. (37) that $s \rightarrow t$ and $Q(s) \rightarrow x(t)$ in the limit $q \rightarrow 1$. The difference equation (34) also reduces to the equation for the free particle (24) when $q \rightarrow 1$. Furthermore, Eq. (34) indicates that the function $Q(s)$ may remain to be real-valued under the condition

$$q + q^{-1} \in \mathbf{R}. \quad (39)$$

This condition implies either $q \in \mathbf{R}$ or $q \in S^1$. In this regard, the object described by the real-valued variable $Q(s)$ is more general than the q -free particle. We refer to this generic object subjected to the condition (39) as the q -object.

Let $s_n = s_0 + nT$ and $Q_n = Q(s_n)$ where $n \in \mathbf{N}_0$. Then it is evident that Q_n satisfies the relation,

$$Q_{n+1} - (q + q^{-1})Q_n + Q_{n-1} = 0, \quad (40)$$

which is nothing but the recursion relation obeyed by the Chebyshev polynomials,

$$T_n(\cos \varphi) = \cos(n\varphi) \quad \text{and} \quad U_n(\cos \varphi) = \sin(n\varphi), \quad (41)$$

with $q = e^{-i\varphi}$ ($\varphi \in \mathbf{C}$). Thus the time-evolution of the q -object is a real Chebyshev process (with $q + q^{-1} \in \mathbf{R}$).

3.3 The Number Representation

The deformed harmonic oscillator of Biedenharn [13] and Macfarlane [14] is a well-known q -deformed object based on the commutator

$$\hat{a}\hat{a}^\dagger - q\hat{a}^\dagger\hat{a} = q^{-\hat{N}}. \quad (42)$$

In the Fock space $\mathcal{F} = \{|n\rangle : n \in \mathbf{N}_0\}$, the operators $\hat{N}$, $\hat{a}^\dagger\hat{a}$ and $\hat{a}\hat{a}^\dagger$ are diagonalized with the diagonal elements, n , u_n and u_{n+1} , respectively; $\hat{a}^\dagger$ and $\hat{a}$ shift the energy states as

$$\hat{a}^\dagger|n\rangle = \sqrt{u_{n+1}}|n+1\rangle, \quad \hat{a}|n\rangle = \sqrt{u_n}|n-1\rangle. \quad (43)$$

Here u_n is often called the structure function of the algebra (42).

By using Eq. (42) it is straightforward to show that the structure function u_n satisfies the recursion relation

$$u_{n+1} - (q + q^{-1})u_n + u_{n-1} = 0. \quad (44)$$

Surprisingly this recursion relation coincides with Eq. (40). However, we do not hastily conclude that the q -object is a q -deformed harmonic oscillator. The deformed oscillator is characterized by a Hamiltonian given as a function of $\hat{a}$ and $\hat{a}^\dagger$, whereas the q -object has nothing to do with a Hamiltonian. Here, associating $|n\rangle$ with the position of the q -object at a time $t = t_0 + nT$, we just point out that the discrete time-evolution of the q -object can be represented by the transition of the Fock states.

3.4 The p -Oscillator

The pulsed harmonic oscillator (p -oscillator) is defined here as a free particle which undergoes periodic pulses of Hooke's force $F(t) = -M\omega^2 X\delta(t/T - m)$, where $M\omega^2$ is Hooke's constant, T is the period of pulses and $m \in \mathbf{Z}$. The symmetrized action of this system for a time interval containing a single pulse is given by

$$S(t_m, t_{m-1}) = \frac{M}{2T}(X_m - X_{m-1})^2 - \frac{1}{4}M\omega^2 T(X_m^2 + X_{m-1}^2). \quad (45)$$

Calculating the canonical momenta by

$$P_m = \partial S / \partial X_m = (M/T)(X_m - X_{m-1}) - (M\omega^2 T/2)X_m, \quad (46)$$

$$P_{m-1} = -\partial S / \partial X_{m-1} = (M/T)(X_m - X_{m-1}) + (M\omega^2 T/2)X_{m-1},$$

we find the area-preserving linear map in phase space:

$$\begin{aligned} X_m &= X_{m-1} + (T/2M)(1 - \omega^2 T^2/4)^{-1}(P_m + P_{m-1}), \\ P_m &= P_{m-1} - (M\omega^2 T/2)(X_m + X_{m-1}). \end{aligned} \quad (47)$$

It is interesting to notice that both X_m and P_m satisfy the recursion relation

$$X_{m+1} - (2 - \omega^2 T^2)X_m + X_{m-1} = 0. \quad (48)$$

Obviously the time-evolution of the p -oscillator is also a Chebyshev process with

$$\varphi = \cos^{-1}(1 - \omega^2 T^2/2). \quad (49)$$

If $0 < \omega^2 T^2 < 4$, then $\varphi \in \mathbf{R}$ and the discrete trajectory $\{X_m\}$ oscillates sinusoidally. If $\omega^2 T^2 < 0$ or $4 < \omega^2 T^2$, then φ is a complex number, and the solutions are no longer physical and do not represent the p -oscillator. Therefore the proper p -oscillator is a q -object with $q \in S^1$, i.e. $q = e^{-i\varphi}$ ($\varphi \in \mathbf{R}$).

4 The Propagator for the q -Object

Here we utilize the time-dependent scale transformation (37) to derive the propagator for the q -object and discuss special cases.

4.1 The q -Object

Letting $c = iT/\varphi$, we put the transformation (37) into the form,

$$x_n = Q_n \sec(in \ln q), \quad t_n = (iT/\ln q) \tan(in \ln q). \quad (50)$$

Then, making use of the identities (22) and (23) once again, we can implement the transformation (50) to the one-dimensional free particle propagator for the time interval $t_n - t_0 = nT$,

$$K^{(\text{free})}(x_n, x_0; nT) = \left[\frac{M}{2\pi i \hbar (t_n - t_0)} \right]^{1/2} \exp \left[\frac{iM}{2\hbar} \frac{(x_n - x_0)^2}{t_n - t_0} \right], \quad (51)$$

to find the propagator for the q -object

$$\begin{aligned} \hat{K}^{(q\text{-object})}(Q_n, Q_0; nT) &= \left[\frac{M(q - q^{-1})}{2\pi i \hbar T (q^n - q^{-n})} \right]^{1/2} \\ &\times \exp \left[\frac{iM(q - q^{-1})}{4\hbar T (q^n - q^{-n})} \{ (Q_n^2 + Q_0^2)(q^n + q^{-n}) - 4Q_n Q_0 \} \right], \end{aligned} \quad (52)$$

where $q + q^{-1} \in \mathbf{R}$.

4.2 The p -Oscillator

For $q = e^{-i\varphi} \in S^1$ with $\varphi = \cos^{-1}(1 - \omega^2 T^2/2)$, we let $Q_n = X_n \in \mathbf{R}$ in Eq. (52). Then the propagator for the proper p -oscillator is given by

$$\hat{K}^{(p\text{-osc})}(X_n, X_0; nT) = \left[\frac{M U_1(\cos \varphi)}{2\pi i \hbar T U_n(\cos \varphi)} \right]^{1/2}$$

$$\times \exp \left[\frac{iM U_1(\cos \varphi)}{2\hbar T U_n(\cos \varphi)} \{ (X_n^2 + X_0^2) T_n(\cos \varphi) - 2X_n X_0 \} \right], \quad (53)$$

where $T_n(\cos \varphi)$ and $U_n(\cos \varphi)$ are the Chebyshev polynomials given in Eq. (41). The caustics of the propagator correspond to $q^n = e^{im\pi}$ ($m \in \mathbf{Z}$). In the limit $T \rightarrow 0$ ($q \rightarrow 1$) and $n \rightarrow \infty$, such that the total time interval nT remains constant, Eq. (53) approaches the propagator for the usual harmonic oscillator.

4.3 The q -Free Particle

The propagator for the T -evolution of the q -free particle is obtained by transforming X_n of Eq. (52) back to $y_n = q^n X_n$ via Eq. (50). Namely,

$$\begin{aligned} \hat{K}^{(q\text{-free})}(y_n, y_0; nT) &= \left[\frac{M(q - q^{-1})}{2\pi i \hbar T (q^n - q^{-n})} \right]^{1/2} \\ &\times \exp \left[\frac{iM(q - q^{-1})}{4\hbar T (q^n - q^{-n})} \{ (q^{-2n} y_n^2 + y_0^2)(q^n + q^{-n}) - 4q^{-n} y_n y_0 \} \right]. \end{aligned} \quad (54)$$

For the proper q -free particle, $q \in \mathbf{R}^+$. If we include the hopping type in the q -free particle, then q can take any real number excluding $q = 0$. In the limit $q \rightarrow 1$, Eq. (55) reduces to the propagator for the usual free particle.

5 Concluding Remarks

The global time-transformations have been used to discuss q -deformed objects. The propagator for the generic q -object is also obtained, which contains those of the pulsed harmonic oscillator and the q -free particle as special cases. Other aspects of the path integral for the q -object will be given elsewhere [15].

It is an interesting question why the AFFJ transformation (17) can generate the discrete energy spectrum for the harmonic oscillator out of the continuous spectrum of the free particle. By the q -analysis, it becomes clear that the AFFJ transformation in the q -version (50) is a sort of analytical continuation through q . As the bound and the continuous states of the Coulomb problem are related to the compact group $O(4)$ and the non-compact group $O(3, 1)$, respectively, the q -free particle with $q \in \mathbf{R}^+$ is analytically continued to the p -oscillator with $q \in S^1$. The generation of the discrete spectrum can be ascribed to the compactification of the space belonging to the deformation parameter q .

References

- [1] I.H. Duru and H. Kleinert, *Phys. Lett. B* **84**, 185 (1979); *Fortschr. Phys.* **30**, 401 (1982). For the time-sliced path integration, see R. Ho and A. Inomata, *Phys. Rev. Lett.* **48**, 231 (1982).
- [2] P. Kustaanheimo and E. Stiefel, *J. Rein. Angew. Math.* **218**, 204 (1965).
- [3] R.P. Feynman, *Rev. Mod. Phys.* **76**, 769 (1949).
- [4] See, for instance, H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995); A. Inomata, H. Kuratsuji, and C.C. Gerry, *Path Integrals and Coherent States of SU(2) and SU(1,1)* (World Scientific, Singapore, 1992); C. Grosche and F. Steiner, *Handbook of Feynman Path Integrals* (Springer-Verlag, Berlin, 1998).
- [5] A. Inomata, in *Path Summation: Achievements and Goals*, Eds. S. Lundviquist, A. Rangfani, V. Sayakanit, and L.S. Schulman (World Scientific, Singapore, 1988), p.114. For a unification of the two types, see A. Pelster and A. Wunderlin, *Z. Phys. B* **89**, 373 (1992).
- [6] V. de Alfaro, S. Fubini, and G. Furlan, *Nuovo Cim.* **34**, 569 (1976).
- [7] R. Jackiw, *Ann. Phys. (N.Y.)* **129**, 183 (1980).
- [8] J.M. Cai, P.Y. Cai, and A. Inomata, *Proceedings of International Symposium on Advanced Topics of Quantum Physics - Shanxi'92*, Eds. J.Q. Liang, M.L. Wang, S.N. Qiao, and D.C. Su (Science Press, Beijing, 1993), p. 75.
- [9] P.Y. Cai, A. Inomata, and P. Wang, *Phys. Lett. A* **91**, 331 (1982).
- [10] G. Junker and A. Inomata, *Phys. Lett. A* **110**, 195 (1985).
- [11] A. Dimakis and F. Müller-Hoissen, *Phys. Lett. B* **295**, 242 (1992).
- [12] See, e.g. M. Chainchian and A. Domichev, *Introduction to Quantum Groups* (World Scientific, Singapore, 1996), p. 154.
- [13] L.C. Biedenharn, *J. Phys. A* **22**, L873 (1989).
- [14] A.J. Macfarlane, *J. Phys. A* **22**, 4581 (1989).
- [15] A. Inomata, J.C. Kimball, and C. Pigorsch, SUNY-A preprint (2000).

FEYNMAN INTEGRAL ON A GROUP

Z. HABA

Institute of Theoretical Physics, University of Wrocław, Poland

E-mail: zhab@ift.uni.wroc.pl

If G is a semisimple Lie group and G^C its complexification then we define a stochastic process g_t with values in G^C such that $\langle f(g_t) \rangle$ is a solution of the Schrödinger equation on G . We consider a function V (a potential) continuous on G and meromorphic on G^C . We obtain the Schrödinger evolution in a potential by means of the Feynman formula. We discuss an equivalent construction of the Schrödinger evolution by means of stochastic equations which are defined as a stochastic perturbation of classical dynamics.

1 Introduction

A definition of the Feynman integral on a manifold encounters numerous difficulties. These difficulties arise because there seems to be no natural finite-dimensional approximation to the Feynman integral on a manifold (contrary to the flat case). It should be useful to apply a map of a manifold (locally) into a region in a flat space, where the path integral is well-defined. The method has been applied by mathematicians [1] in order to define the Brownian motion and the Wiener measure on a Riemannian manifold. Multivalued mappings have been introduced by Hagen Kleinert [2]. Such transformations allow us to define the Feynman integral on a manifold with curvature and torsion.

In this contribution we study the mapping method in order to define the Feynman integral on a semisimple Lie group in terms of the Brownian motion on R^d . The Feynman integral is understood here in a broad sense as defining the solution of the Schrödinger equation by means of a functional measure. In our earlier works [3,4] we have studied the Brownian motion

representation of the Feynman integral in a flat case. The manifolds were only briefly discussed. The semisimple Lie groups are particularly simple from our point of view because there is a natural map from a flat space (the algebra) onto the group. So, we can transform the Brownian motion from the algebra into a complex process on a complexification of the group. This process gives a realization of the Feynman integral. Subsequently, we consider a function V on a group (a potential) and derive the Feynman integral representation of the solution of the Schrödinger equation with the potential. We discuss also the semiclassical expansion and a Langevin equation on a group which can be applied to generate this semiclassical expansion.

2 Notation and Preliminaries

Let G be the d -dimensional Lie group, whose complexification is denoted by G^C . $\mathcal{G}$ is the Lie algebra of G (we denote its complexification by $\mathcal{G}^C$). The exponential map $\mathcal{E}$ is a diffeomorphism of a neighbourhood of 0 in $\mathcal{G}$ into a neighbourhood of 1 in G . It extends to a holomorphic map of a neighbourhood of $\mathcal{G}^C$ into a neighbourhood of G^C . We shall restrict ourselves to matrix groups (which give a faithful representation of a semisimple Lie group G). By τ^j we denote a basis of the Lie algebra $\mathcal{G}$ normalized by the Killing form C^{jk} . Then we define $\Upsilon = \sum_1^d y_j \tau^j$, where $y_j \in R$. When restricted to the matrix group the exponential map will also be denoted by $\mathcal{E}(\Upsilon) = \exp(\Upsilon)$, because in this case the exponential has a direct meaning as an exponential of a matrix (now $\text{Tr}(\tau^j \tau^k) = -C^{jk}$).

Let r_j ($j = 1, \dots, d$), be independent random Gaussian variables. The vector $\mathbf{r}$ can be considered as a random variable in a d -dimensional algebra $\mathcal{G}$ with the probability distribution

$$p(\mathbf{y}) = (2\pi)^{-\frac{d}{2}} \exp\left(-\frac{|\mathbf{y}|^2}{2}\right). \quad (1)$$

We wish to transform this random variable to the group by means of the exponential map. Let $\mathcal{R} = \sum_{k=1}^d r_j \tau^j \in \mathcal{G}$. We define the probability distribution on the semisimple Lie group G^C by the formula

$$E[F(\exp(z\mathcal{R}))] = \int_{R^d} d\mathbf{y} p(\mathbf{y}) F(\exp(z\Upsilon)). \quad (2)$$

For the Feynman integral we choose

$$z = \epsilon\lambda\sigma,$$

where ϵ is a real parameter, λ is defined by

$$\lambda = \frac{1}{\sqrt{2}}(1 + i),$$

and

$$\sigma = \sqrt{\frac{\hbar}{m}}$$

with Planck's constant $\hbar$ and mass m .

3 Discretization of the Feynman Integral

The Gaussian random variables are applied in order to define a finite-dimensional approximation to the Feynman integral. We define recursively a random walk on the group as a solution of the equation

$$g(n+1) = \exp(z\mathcal{R}(n))g(n), \quad (3)$$

where $n \geq 0$ and $g(0) = g$. Then, we take the limit $n \rightarrow \infty$, $\epsilon \rightarrow 0$ and $\epsilon^2 n \rightarrow t \geq 0$ in order to define a continuous process $g(t)$ with values in G^C . It has the following properties

- i) $g(0) = g$.
- ii) The increments

$$g(t_1)^{-1}g(t_2), \dots, g(t_{n-1})^{-1}g(t_n), \quad t_1 \leq t_2 \leq \dots \leq t_n \quad (4)$$

are mutually independent.

- iii) The probability distribution of $g(t_1)^{-1}g(t_2)$ depends only on $t_2 - t_1$.
- iv) $g(t)$ is the solution of the Stratonovitch stochastic equation [5]

$$dg(t) = zdb(t) \circ g(t), \quad (5)$$

where $b(s) = \sum_{k=1}^d \tau^k b_k(s)$ and $b_k(s)$ are independent Brownian motions.

Eq. (5) follows from Eq. (3). In fact, we have

$$\sum_{k=1}^n (g(k+1) - g(k)) g(k)^{-1} = \sum_{k=1}^n (\exp(z\mathcal{R}) - 1),$$

and the right-hand side tends to the Brownian motion as $\epsilon \rightarrow 0$ and $n \rightarrow \infty$ (because the sum of independent random variables tends to the Brownian motion).

When the probability distribution (2) of independent random variables is known then the distribution of group elements can be formally expressed by a change of variables $y \rightarrow g$,

$$\prod_k dy_k \exp(-y_k^2/2) = \prod_k dy_k \exp \left\{ -\frac{1}{2z^2} \text{Tr} \left[\ln \left(g(k+1) g(k)^{-1} \right) \right]^2 \right\}. \quad (6)$$

It is clear that the logarithm is only locally defined around the unit element of the group. Moreover, in order to replace the R^d integral by a group integral, we would need to calculate the Jacobian. Again we could do it only locally. Fortunately, we do not need to derive the corresponding formulas explicitly. We know that in general a linear continuous functional $\langle f(g_1, \dots, g_n) \rangle$ on a Cartesian product of groups determines a measure on the product of groups. This is the discrete version of the Feynman integral. We discuss still another definition in the next section.

4 Unitary Evolution in $L^2(dg)$

Instead of describing explicitly the discrete Feynman measure we consider the discrete time evolution of the group element. The probability distribution (2) of $g(n)$ determines the operator

$$(K_{\epsilon, \sigma} F)(g) = E[F(g(\mathcal{R})g)], \quad (7)$$

where we denoted $g(\mathcal{R}) = \exp(\epsilon \lambda \sigma \mathcal{R})$. Clearly,

$$(K_{\epsilon, \sigma}^n F)(g) = E[F(g(\mathcal{R}(n)) \dots g(\mathcal{R}(1))g)], \quad (8)$$

where $\mathcal{R}(k)$ numerates independent random variables. From Eqs. (2), (3), and (7) we obtain the kernel of $K_{\epsilon, \sigma}$

$$(K_{\epsilon, \sigma} F)(g) = \int d\mathbf{y} p(\mathbf{y}) F(g(\Upsilon)g) \equiv \int d\mathbf{y} K_{\epsilon, \sigma}(g, \mathbf{y}) F(g(\Upsilon)). \quad (9)$$

Eq. (7) has a meaning only for analytic functions F . However, using the (distributional) kernel (9) we can extend the definition of $K_{\epsilon, \sigma}$ to any $F \in L^1(dg)$. From Eq. (8) we have

$$(K_{\epsilon, \sigma}^n F)(g) = \int d\mathbf{y}_1 \dots d\mathbf{y}_n p(\mathbf{y}_1) \dots p(\mathbf{y}_n) F(g(\Upsilon_n) \dots g(\Upsilon_1)g). \quad (10)$$

We define the generator A_σ of $K_{\epsilon,\sigma}$ as a limit in $L^2(d\Upsilon)$,

$$A_\sigma = \lim_{\epsilon \rightarrow 0} \epsilon^{-2} (K_{\epsilon,\sigma} - 1) \equiv \frac{i\hbar}{2m} \Delta_G. \quad (11)$$

Applying the Taylor formula on the Lie group we can compute the generator A_σ of $K_{\epsilon,\sigma}$ directly from Eq. (11). We obtain

$$A_\sigma = -\frac{i\hbar}{2} \sum_{k=1}^d C^{kl} X_k X_l, \quad (12)$$

where X_k is the basis of right-invariant vector fields on G corresponding to the basis $\tau \in \mathcal{G}$. Then Δ_G is the Laplace-Beltrami operator on the group.

The limit $\epsilon \rightarrow 0$ in Eq. (9) defines a continuous time semigroup. In fact, it follows from Eq. (11) that if we define

$$(U_t F)(g') \equiv F_t(g') = E[F(g(t)g')], \quad (13)$$

where $g(t)$ (with $g(0) = 1$) is the solution of Eq. (5), then U_t defines a semigroup in $L^2(dg)$ with the generator A_σ . F_t is the solution of the Schrödinger equation

$$\frac{d}{dt} F_t(g') = A_\sigma F_t(g') \quad (14)$$

with the initial condition F .

After a change of coordinates from the algebra to the group, Eq. (10) could be considered as another rigorous version of the discrete Feynman integral (a formal expression is described by Eq. (6)). Note however that the map $G \rightarrow R^d$ is multivalued and discontinuous (as the maps considered in Ref. [6]). Hence, the formula for a change of variables in Eq. (9) would be rather complicated. The Feynman integral on a manifold has been defined in some earlier papers [7–9] in terms of the short-time propagator. However, the short-time propagator must be derived first by other methods. Then, the continuum limit (13) of the products of short-time propagators does not define a measure. Finally, in order to obtain the generator (11) one must artificially subtract the $R/6$ term (where R is the scalar curvature) from the propagator. The $R/6$ term comes from a determinant of the map $R^d \rightarrow G$ (as was shown in Ref. [2]). Hence, its presence is explained by the method of non-holonomic equivalence principle [6].

5 Feynman Integral on a Group

Let V be a continuous function on G which is meromorphic on G^C . We are going to express a solution of the Schrödinger equation with the Hamiltonian $H = i\hbar A_\sigma + V$ by the Feynman integral. Assume that for a function F defined on G^C the function

$$\exp\left(-\frac{i}{\hbar} \int_0^t V(g(s)g') ds\right) F(g(t)g')$$

is integrable. Then

$$(U^V(t)F)(g) = E\left[\exp\left(-\frac{i}{\hbar} \int_0^t V(g(s)g) ds\right) F(g(t)g)\right] \quad (15)$$

satisfies the Schrödinger equation

$$\frac{d}{dt} F_t(g) = \left(A_\sigma - \frac{i}{\hbar} V(g)\right) F_t(g) \quad (16)$$

with the initial condition $\lim_{t \rightarrow 0} F_t(g) = F(g)$.

The semigroup property of the operator defined by the right-hand side of Eq. (15) follows from the Markov property of $g(t)$ and the additivity of $\int_0^t ds V$ as a function of t . The Markov property of $g(t)$ follows from its construction (3). Then, using the semigroup property it is sufficient to prove Eq. (16) at $t = 0$. At $t = 0$, Eq. (16) follows from Eq. (14) and the Leibniz rule of differentiation.

There remains the basic problem of showing the integrability of the function inside the square brackets in Eq. (15).

Our basic tool in the proof of integrability is the Jensen inequality

$$\begin{aligned} & \left| E\left[\exp\left(-\frac{i}{\hbar} \int_0^t V(g(s)g) ds\right) F(g(t)g)\right] \right| \\ & \leq \int_0^t \frac{ds}{t} E\left[\exp\left(\Im \frac{t}{\hbar} V(g(s)g)\right) |F(g(t)g)|\right] \\ & = \int_0^t \frac{ds}{t} dg' K_s(g, g') \exp\left(\Im \frac{t}{\hbar} V(g')\right) |F(g')|, \end{aligned} \quad (17)$$

where K_s is the kernel of U_s . We apply this inequality for a regularized V , e.g.

$$V_R(g(s)) = V(g(s)) \exp\left(-\int_0^s d\tau \text{Tr}(b(\tau))^2 / R\right),$$

which is a bounded function, then we take the limit $R \rightarrow \infty$. Hence, the inequality (17) holds true for V itself. Another method of a regularization of the Feynman formula comes from the Trotter product formula (U_t is defined in Eq. (13))

$$U_t^V F = \lim_{n \rightarrow \infty} \left[U_{t/n} \exp \left(-\frac{it}{n\hbar} V \right) \right]^n F,$$

where for a large n

$$U_{t/n} F(g) \approx \int d\mathbf{y} p(\mathbf{y}) F \left(\exp \left(\lambda \sigma \frac{t}{n} \Upsilon \right) g \right). \quad (18)$$

Hence

$$\begin{aligned} & \left| U_{t/n} \exp \left[-\frac{it}{n\hbar} V(g) \right] F(g) \right| \\ & \approx \int d\mathbf{y} p(\mathbf{y}) \left| \exp \left[\Im \frac{t}{n\hbar} V \left(\exp \left(\lambda \sigma \frac{t}{n} \Upsilon \right) g \right) \right] \right| \\ & \quad \times \left| F \left(\exp \left(\lambda \sigma \frac{t}{n} \Upsilon \right) g \right) \right|. \end{aligned} \quad (19)$$

At this stage we can apply the Jensen inequality (17).

As an example let us consider the group $SU(2)$. Each element of the group can be expressed in the form (where $\mathbf{r} = (y_1, \dots, y_d)$ and $r = |\mathbf{r}|$)

$$g = \exp(i\mathbf{s}\mathbf{r}) = I \cos r + i\mathbf{s}\mathbf{r} \sin r, \quad (20)$$

where $\mathbf{s}$ are the Pauli matrices satisfying

$$[s_j, s_k] = 2i\epsilon_{jkl}s_l.$$

As a typical meromorphic function on G^C we may consider

$$V(g) = \text{Tr}(g)(1 + (\text{Tr}(g))^2)^{-1}, \quad (21)$$

where $\text{Tr}(g) = 2 \cos r$. Now, we have

$$g = \exp(\lambda \sigma \Upsilon) = I \cos(\lambda \sigma r) + i\lambda \sigma \mathbf{s} \Upsilon \sin(\lambda \sigma r).$$

In order to prove that the integral (15) is finite for a small time we must show that the singularities of $\exp(\Im V(\exp(\lambda \sigma \Upsilon)g))$ are integrable. Let us take for simplicity $g = 1$ in Eq. (15). Then we have $\text{Tr}(\exp(\lambda \sigma \Upsilon)) = 2 \cos((1+i)A)$

for a certain real A . Hence, the denominator in Eq. (15) with the potential (21) is

$$1 + (\text{Tr}(\exp(\lambda\sigma\Upsilon)))^2 = 1 + \cos^2 A \cosh^2 A - \sin^2 A \sinh^2 A - \frac{i}{2} \sin 2A \sinh 2A. \quad (22)$$

We can see that this expression is never zero. Hence, the exponential factor is integrable. In general, the denominator can be zero but on a set of measure zero. Then, the integrability can still be established.

6 Semiclassical Expansion

The classical limit of Eq. (13) (without the potential) is determined by the Lagrangian

$$L = \frac{m}{2} \text{Tr} \left(\frac{dg}{dt} g^{-1} \right)^2. \quad (23)$$

As there is only the kinetic term in L , the time derivative of L vanishes when calculated on the classical solution. Hence, the classical equations read

$$\frac{dg}{dt} g^{-1} = v = \text{const.}$$

We can now construct the action and the semiclassical approximation which appears to be the exact solution of the quantum problem [9]. Hence, without the potential, the classical limit is trivial. It becomes non-trivial if the potential is present.

There are two approaches to the semiclassical expansion:

- i) We transform first the Schrödinger equation using a solution of the Hamilton-Jacobi equation, express the solution by a modified Feynman integral and subsequently estimate the remainder.
- ii) We make a shift of variables in the Feynman integral using the classical solution and then estimate the remainder.

The methods i)-ii) lead to the same result but they require different assumptions to justify the estimates. Let us begin with the first method. Assume that W_s is the solution (with the initial condition W) of the Hamilton-Jacobi equation on G with a scalar potential V :

$$\partial_s W_s + \frac{1}{2m} (\nabla W_s, \nabla W_s) + V = 0. \quad (24)$$

The scalar product in Eq. (24) is with respect to the Riemannian metric on G . Let us consider a solution of the Schrödinger equation with the initial condition $\psi = \exp(iW/\hbar)\phi$. Then, the solution of the Schrödinger equation can be expressed in the form $\psi_t = \exp(iW_t/\hbar)\phi_t$, where ϕ_t is the solution of the equation

$$\partial_s \phi_s = \frac{i\hbar}{2m} \Delta_G \phi_s - \frac{1}{m} (\nabla W_s, \nabla \phi_s) + \frac{1}{2m} (\Delta_G W_s) \phi_s. \quad (25)$$

On a formal level we may neglect the second-order differential operator in Eq. (25) in the limit $\hbar \rightarrow 0$. Then, the limit $\hbar \rightarrow 0$ of the solution of Eq. (25) is determined by the classical flow. In order to make the formal argument rigorous as well as to derive an expansion in $\hbar$ it is useful to express the solution of Eq. (25) by a Markov process

$$\phi_t(x) = E \left[\exp \left(\frac{1}{2m} \int_0^t (\Delta_G W_{t-s}(\zeta_s)) ds \right) \phi(\zeta_t(x)) \right], \quad (26)$$

where the process $\zeta_s(x)$ is defined for $0 \leq s \leq t$ as a solution of the Langevin equation (with the initial condition $x \in G$)

$$d\zeta^\mu(s) = -\frac{1}{m} g^{\mu\nu}(\zeta(s)) \partial_\nu W_{t-s}(\zeta(s)) ds + d\xi^\mu(s), \quad (27)$$

where $g^{\mu\nu}$ is the Riemannian metric on G and ξ denotes the Markov process generated by $i\hbar\Delta_G/2m$ defined in a coordinate independent way in Eq. (5). In the classical limit $\hbar \rightarrow 0$ the solution of Eq. (27) converges to the classical trajectory of a particle moving on a group manifold in the potential V .

Acknowledgments

This paper is dedicated to Hagen Kleinert on the occasion of his 60th birthday with thanks for his inspiring influence on our collaboration. My stay at the Freie Universität Berlin was made possible thanks to grant HSP III-Potsdam from the German university support program. I am grateful to Dr. Hans-Jürgen Schmidt for kindly integrating me into his joint project with Professor Kleinert.

References

- [1] D.K. Elworthy and A. Truman, *J. Math. Phys.* **22**, 2144 (1981).

-
- [2] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
 - [3] Z. Haba, *J. Phys. A* **27**, 6457 (1994).
 - [4] Z. Haba, *Feynman Integral and Random Dynamics in Quantum Physics* (Kluwer, Dordrecht, 1999).
 - [5] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North-Holland, Amsterdam, 1981).
 - [6] H. Kleinert, *Gen. Rel. Grav.* **32**, 769 (2000).
 - [7] B.S. DeWitt, *Rev. Mod. Phys.* **29**, 377 (1957).
 - [8] D.W. McLaughlin and L.S. Schulman, *J. Math. Phys.* **12**, 2520 (1971).
 - [9] J.S. Dowker, *Ann. Phys. (N.Y.)* **62**, 361 (1971).

CHARACTERIZING VOLUME FORMS

P. CARTIER

*Ecole Normale Supérieure,
45 rue d'Ulm F-75005 Paris, France*

M. BERG, C. DEWITT-MORETTE, AND A. WURM

*Department of Physics and Center for Relativity,
University of Texas, Austin, TX 78712, USA*

E-mail: alex@einstein.ph.utexas.edu

We present old and new results for characterizing volume forms. The Cartier/DeWitt-Morette group, which regularly shares questions and findings, hopes that Hagen Kleinert will enjoy a new approach for characterizing volume forms on Riemannian and symplectic manifolds, using integration by parts.

1 The Wiener measure

Defining volume forms on infinite-dimensional spaces is a key problem in the theory of functional integration. The first volume form used in functional integration has been the Wiener measure. From the several equivalent definitions of the Wiener measure, we choose one [1] which can easily be extended for use in Feynman integrals. We recall the Cameron-Martin and Malliavin formulae because they are, respectively, integrated and infinitesimal formulae for changes of variable of integration which can be imposed on volume forms other than the Wiener measure.

1.1 Definition

The Wiener measure γ on the space $\mathbb{W}$ of pointed continuous paths w on the time interval $T = [0, 1]$,

$$w : [0, 1] \longrightarrow \mathbb{R} , \qquad w(0) = 0,$$

can be characterized by the equation

$$\int_{\mathbb{W}} d\gamma(w) \exp(-i\langle w', w \rangle) = \exp\left(-\frac{1}{2} \int_T \int_T dw'(t) dw'(t') \inf(t, t')\right), \quad (1)$$

where w' is an element of the topological dual $\mathbb{W}'$ of $\mathbb{W}$, i.e. a bounded measure on the semi-open interval $T =]0, 1]$,

$$\langle w', w \rangle = \int_T dw'(t) w(t). \quad (2)$$

1.2 Cameron-Martin Formula

The Cameron-Martin formula can be written as

$$\int_{\mathbb{W}} d\gamma(w) F(w + \varphi) = \int_{\mathbb{W}} d\gamma(w) J(\varphi, w) F(w), \quad (3)$$

where $J(\varphi, w)$ is the Radon-Nikodym derivative

$$J(\varphi, w) = \frac{d\gamma(w - \varphi)}{d\gamma(w)}. \quad (4)$$

This formally obvious expression for $J(\varphi, w)$ has a not-so-obvious explicit expression. When $\varphi \in L^{2,1}$, i.e. when $\dot{\varphi}(t)$ is square integrable, and $\varphi(0) = 0$, then^a

$$J(\varphi, w) = \exp\left(-\frac{1}{2} \int_T dt \dot{\varphi}(t)^2 + \int_T dw(t) \dot{\varphi}(t)\right). \quad (5)$$

The meaning of the second term is subtle since w is not of bounded variation. If $\varphi \in C^2(0, 1)$ with the boundary conditions $\varphi(0) = \dot{\varphi}(1) = 0$, the second term can be integrated by parts, giving $-\int_T dt w(t) \ddot{\varphi}(t)$, and can then be extended by continuity in the $L^{2,1}$ norm for φ .

A heuristic proof of the Cameron-Martin formula makes the explicit expression (5) "obvious". Let us write formally

$$d\gamma(w) = \mathcal{D}w \exp(-\pi Q(w)), \quad (6)$$

^aThe boundary condition $\varphi(0) = 0$ is required in order that $w + \varphi$ belongs to $\mathbb{W}$ when w does, since $w(0) = 0$ for every element w of $\mathbb{W}$.

where $\mathcal{D}$ is a translation-invariant symbol^b

$$\mathcal{D}(w + \varphi) = \mathcal{D}w \quad (7)$$

and where Q follows from the definition (1) of the Wiener measure:

$$\int_{\mathbb{W}} \mathcal{D}w \exp(-\pi Q(w)) \exp(-2\pi i \langle w', w \rangle) = \exp(-\pi W(w')) \quad (8)$$

with

$$W(w') = 2\pi \int_T \int_T dw'(t) dw'(t') \inf(t, t'). \quad (9)$$

By analogy with the finite-dimensional case (43), the quadratic form Q on $\mathbb{W}$ is required to be the inverse of W on $\mathbb{W}'$ in the following sense. Represent W and Q as

$$W(w') = \langle w', Gw' \rangle \quad \text{and} \quad Q(w) = \langle Dw, w \rangle; \quad (10)$$

then Q is said to be the inverse of W if

$$DG = \mathbf{1}. \quad (11)$$

It follows from (9) and (10) that

$$Gw'(t) = 2\pi \int_T dw'(t') \inf(t, t') \quad (12)$$

and from (10) and (11) that

$$Q(w) = \frac{1}{2\pi} \int_T dt \left(\frac{dw(t)}{dt} \right)^2 = \frac{1}{2\pi} \int_T \frac{(dw(t))^2}{dt}. \quad (13)$$

The Cameron-Martin formula (3) is now the obvious statement

$$\int_{\mathbb{W}} \mathcal{D}w \exp[-\pi Q(w)] F(w + \varphi) = \int_{\mathbb{W}} \mathcal{D}w \exp[-\pi Q(w - \varphi)] F(w), \quad (14)$$

that is $J(\varphi, w) = \exp\{\pi[Q(w) - Q(w - \varphi)]\}$. We calculate

$$Q(w - \varphi) = \frac{1}{2\pi} \int_T dt \left(\frac{d}{dt}(w - \varphi)(t) \right)^2$$

^bThe symbol $\mathcal{D}$ is often used in physics where $\mathcal{D}w := \prod_t dw(t)$. Here it is defined by (8), which in finite dimensions reduces to (45).

$$= \frac{1}{2\pi} \left(Q(w) + \int_T dt \dot{\varphi}(t)^2 - 2 \int_T dw(t) \dot{\varphi}(t) \right). \quad (15)$$

Thus formula (5) follows immediately and this completes the heuristic demonstration.

1.3 An Analogy with the Dobrushin-Lanford-Ruelle Characterization of Gibbs States

What is missing in the heuristic proof of the Cameron-Martin formula to be rigorous? The difficulty is that the Brownian trajectories w are so rough that $Q(w)$ is infinite if calculated as the limit of the Riemann sums $\sum_{i=1}^N (\Delta w_i)^2 / \Delta t_i$, where

$$0 = t_0 < t_1 < \dots < t_N = 1, \quad \Delta t_i = t_i - t_{i-1}, \quad \Delta w_i = w(t_i) - w(t_{i-1}).$$

However, in the Cameron-Martin formula we need only the difference $Q(w) - Q(w - \varphi)$. The infinite part drops out in the difference provided φ is smooth enough, e.g. if φ is of class C^2 on $T = [0, 1]$. A similar situation occurs in statistical mechanics in the case of infinite volumes. For a configuration w , the formal Hamiltonian $H(w)$ may be infinite. But if a configuration w' is obtained by a local modification of w , by changing the states in finitely many sites, then $H(w) - H(w')$ is finite. This is the strategy underlying the Dobrushin-Lanford-Ruelle (DLR) characterization of Gibbs states [2].

1.4 Malliavin Formula

According to the Malliavin formula,

$$\int_{\mathbb{W}} d\gamma(w) D_\varphi F(w) = \int_{\mathbb{W}} d\gamma(w) A_\varphi(w) F(w), \quad (16)$$

where $D_\varphi F$ is the Gateaux differential of F in the φ -direction,

$$D_\varphi F(w) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [F(w + \epsilon \varphi) - F(w)] = \int_T dt \frac{\delta F(w)}{\delta w(t)} \varphi(t), \quad (17)$$

and where

$$\begin{aligned} A_\varphi(w) &:= \int_T dw(t) \dot{\varphi}(t) \\ &= - \int_T dt w(t) \ddot{\varphi}(t) \end{aligned} \quad (18)$$

for $\varphi \in C^2(0, 1)$, with $\varphi(0) = \varphi(1) = 0$. Again, the Malliavin formula is "obvious" if we use the formal expression (6) and integrate the left hand side of (16) by parts,

$$\int_{\mathbb{W}} \mathcal{D}w \exp(-\pi Q(w)) D_{\varphi} F(w) = - \int_{\mathbb{W}} \mathcal{D}w D_{\varphi}(\exp(-\pi Q(w))) F(w). \quad (19)$$

The Cameron-Martin formula and its infinitesimal form, the Malliavin formula, pave the way for defining a formal translation-invariant symbol " $\mathcal{D}$ ". In this paper, we propose an infinitesimal characterization of $\mathcal{D}$; namely, given an arbitrary functional U integrable by $\mathcal{D}w$, the translation invariance of $\mathcal{D}$ can be expressed by integrating by parts

$$\int_{\mathbb{W}} \mathcal{D}w \frac{\delta U}{\delta w(t)} = - \int_{\mathbb{W}} \frac{\delta}{\delta w(t)} \mathcal{D}w \cdot U = 0 \quad \forall U. \quad (20)$$

1.5 Some Lessons from the Malliavin Formula

- Let us write formally

$$d\gamma(w) = \mu(w) \mathcal{D}w, \quad (21)$$

where $\mu(w)$, often called "measure" in physics, is not necessarily $\exp(-\pi Q(w))$. Malliavin's formula reads

$$\int_{\mathbb{W}} \mu(w) \mathcal{D}w (D_{\varphi} F(w) - A_{\varphi}(w) F(w)) = 0, \quad (22)$$

which, by a formal integration by parts, becomes

$$\int_{\mathbb{W}} \mathcal{D}w (D_{\varphi} \mu(w) + A_{\varphi}(w) \mu(w)) F(w) = 0. \quad (23)$$

Since $F(w)$ is arbitrary, Malliavin's formula is equivalent to

$$D_{\varphi} \mu(w) + A_{\varphi}(w) \mu(w) = 0 \quad (24)$$

for all φ sufficiently regular, e.g. $\varphi \in C^2(0, 1)$. The "measure" $\mu(w)$ is, modulo a multiplicative constant, characterized by (24). If $\mu(w) = \exp(-\pi Q(w))$ as before, then

$$D_{\varphi} Q(w) = \frac{1}{\pi} A_{\varphi}(w). \quad (25)$$

According to (18), $A_\varphi(w)$ is bilinear in φ and w ; hence Q is quadratic in w and we recover formula (13).

- The above remark is a heuristic proof that the Wiener measure γ is characterized by Malliavin's formula. The proof can be made rigorous by choosing in (16)

$$F(w) = \exp(-2\pi i \langle w', w \rangle) , \quad w' \in \mathbb{W}' .$$

- Conversely the Cameron-Martin formula provides a rigorous proof of the Malliavin formula (16): replace φ by $\epsilon\varphi$ in (3) and take the derivative of both sides with respect to ϵ . At $\epsilon = 0$, one checks that

$$\left. \frac{d}{d\epsilon} J(\epsilon\varphi, w) \right|_{\epsilon=0} = A_\varphi(w) .$$

- The Malliavin formula can be used for realizing creation and annihilation operators on bosonic Fock spaces [3], thanks to the Wiener chaos isomorphism: Let $\mathcal{H} := L^{2,1}(T)$ be a one-particle (real) Hilbert space with scalar product $(\varphi_1|\varphi_2) = \int_T dt \dot{\varphi}_1(t) \dot{\varphi}_2(t)$. Let $\mathcal{F}(\mathcal{H})$ be the Fock space with vacuum Ω . The *Wiener chaos* is an isomorphism (see e.g. Ref. [4])

$$L^2(\mathbb{W}, d\gamma) \simeq \mathcal{F}(\mathcal{H}) . \quad (26)$$

With

$$A_\varphi(w) := \int_T dw(t) \dot{\varphi}(t),$$

one obtains

$$\int_{\mathbb{W}} d\gamma(w) A_{\varphi_1}(w) A_{\varphi_2}(w) = \int_T dt \dot{\varphi}_1(t) \dot{\varphi}_2(t). \quad (27)$$

It suffices to consider the case $\varphi_1 = \varphi_2 = \varphi$. This fundamental formula can then be established by a linear change of variable of integration in $\mathbb{W}$,

$$A_\varphi : w \longmapsto \int_T dw(t) \dot{\varphi}(t).$$

Equation (27) says that the map from $\mathcal{H}$ to $\mathcal{F}$,

$$\mathcal{H} := L^{2,1}(T) \longrightarrow \mathcal{F} := L^2(\mathbb{W}, d\gamma) \quad (28)$$

$$\text{by} \quad \varphi \longmapsto A_\varphi \quad (29)$$

is an isometry. The Malliavin formula (16), with $\varphi = \varphi_1$ and $F = A_{\varphi_2}$, gives another proof of (27),

$$(\varphi_1|\varphi_2) = \int_{\mathbb{W}} d\gamma(w) D_{\varphi_1}(A_{\varphi_2}(w)) = \int_{\mathbb{W}} d\gamma(w) A_{\varphi_1}(w) A_{\varphi_2}(w), \quad (30)$$

and the following commutation relations are obvious:

$$[D_{\varphi_1}, D_{\varphi_2}] = 0, \quad [A_{\varphi_1}, A_{\varphi_2}] = 0, \quad [D_{\varphi_1}, A_{\varphi_2}] = (\varphi_1|\varphi_2). \quad (31)$$

Therefore

$$a^\dagger(\varphi) := A_\varphi - D_\varphi, \quad a(\varphi) := D_\varphi \quad (32)$$

obey the bosonic commutation relation of creation and annihilation operators on $\mathcal{F}$, respectively,

$$[a(\varphi_1), a^\dagger(\varphi_2)] = (\varphi_1|\varphi_2) \cdot \mathbf{1},$$

other commutators vanishing. It can be proved that a and $a^\dagger$ are adjoint in the Hilbert space $L^2(\mathbb{W}, d\gamma)$ by integrating by parts the Malliavin formula (16) with $F = F_1 F_2$,

$$\int_{\mathbb{W}} d\gamma D_\varphi F_1 \cdot F_2 = \int_{\mathbb{W}} d\gamma F_1 \cdot (A_\varphi F_2 - D_\varphi F_2). \quad (33)$$

The vacuum $\Omega \in \mathcal{F}$ is the constant function equal to 1. If a functional F of the Brownian path w acts on $\mathcal{F}$ by multiplication, i.e.

$$F(w) : \Psi(w) \longmapsto F(w) \Psi(w),$$

then we derive the tautology

$$\langle \Omega | F | \Omega \rangle = \int_{\mathbb{W}} d\gamma(w) F(w). \quad (34)$$

The Wiener measure is therefore the spectral measure corresponding to the vacuum state Ω . The vacuum is characterized by

$$a(\varphi) \Omega = D_\varphi \Omega = 0, \quad \forall \varphi$$

or alternatively by

$$\frac{\delta \Omega(w)}{\delta w(t)} = 0, \quad \forall t \in T.$$

With the notation of (34), we write

$$0 = \langle F | a(\varphi) \Omega \rangle = \langle a^\dagger(\varphi) F | \Omega \rangle.$$

Hence the vacuum Ω is, up to a scalar, the unique state orthogonal to all $a^\dagger(\varphi)F$, that is to the functions $D_\varphi F - A_\varphi \cdot F$. This gives another interpretation to Malliavin's formula (16).

To show that $\{L^2(\mathbb{W}, d\gamma), \Omega, a(\varphi), a^\dagger(\varphi)\}$ is indeed a model of Fock space, it remains to check that the vectors $a^\dagger(\varphi_1) \dots a^\dagger(\varphi_n) \Omega$ make a total system for $L^2(\mathbb{W}, d\gamma)$, i.e. that the finite linear combinations of $A_{\varphi_1} \dots A_{\varphi_n}$ are dense in $L^2(\mathbb{W}, d\gamma)$. The Wiener chaos follows from the general theory of Fock space. The general theory includes not only the symmetric space considered here, but also the antisymmetric Fock space which we have not yet considered.

1.6 Feynman Volume Form

The Fourier transform of the Wiener measure (1) or (8) suggests a characterization of the Feynman volume form by its Fourier transform (see Refs. [5,6]). Let $s \in \{1, i\}$, then we can define $\mathcal{D}x$ by

$$\int_{\mathbb{X}} \mathcal{D}x \exp\left(-\frac{\pi}{s} Q(x)\right) \exp(-2\pi i \langle x', x \rangle) = \exp(-s\pi W(x')), \quad (35)$$

where $\mathbb{X}$ is the space of paths x and the quadratic form

$$Q(x) > 0 \text{ for } s = 1, \quad \text{Im } Q(x) > 0 \text{ for } s = i. \quad (36)$$

The case $s = 1$ corresponds to the Wiener measure, while the case $s = i$ corresponds to the Feynman sum over paths in quantum mechanics. Everything said before can be repeated with obvious changes, e.g. the Malliavin formula.

2 Volume Forms in Quantum Field Theory; Schwinger's Dynamical Principle

The functional integral representation of the Schwinger dynamical principle has led Bryce DeWitt to the introduction of a ubiquitous volume form in quantum field theory. According to Schwinger, the variation of the probability amplitude for a transition $\langle \text{out} | \text{in} \rangle$ is given by the variation of the action $\mathbf{S}$ of

the system^c:

$$\delta \langle \text{out} | \text{in} \rangle = \frac{i}{\hbar} \langle \text{out} | \delta \mathbf{S} | \text{in} \rangle, \quad (37)$$

where $\mathbf{S}$ is a functional of the field operators, which are globally designated by φ .

2.1 Evolution Equations for the Field Operators φ

The Schwinger variational principle gives evolution equations for the field operators φ different from the classical Euler-Lagrange equations

$$\frac{\delta S}{\delta \varphi} = 0. \quad (38)$$

The Schwinger-Dyson equations give the quantum evolution of polynomials of fields $F(\varphi)$ for a system with classical action S by the expectation value of a time ordered operator,^d

$$\langle \text{vac} | T \left(\frac{i}{\hbar} \frac{\delta S}{\delta \varphi} F(\varphi) + \frac{\delta F}{\delta \varphi} \right) | \text{vac} \rangle = 0. \quad (39)$$

2.2 Functional Integral Solution of the Schwinger Principle

To exploit the Schwinger variational principle (37), one varies an external source J added to the original action S . The new action is

$$S + \langle J, \varphi \rangle,$$

and the principle (37) now reads

$$\frac{\hbar}{i} \frac{\delta}{\delta J} \langle \text{out} | \text{in} \rangle = \langle \text{out} | \varphi | \text{in} \rangle.$$

Bryce DeWitt has constructed the following functional integral solution of this equation (for details, see pp. 4160-4164 and related references in Ref. [6]):

$$\langle \text{out} | \text{in} \rangle = \mathcal{N} \int_{\Phi(\text{in}, \text{out})} \mu(\varphi) \mathcal{D}\varphi \exp \left(\frac{i}{\hbar} (S(\varphi) + \langle J, \varphi \rangle) \right), \quad (40)$$

where

^cWe use boldface for operators on Fock space.

^dThe proof of Eq. (39) and some of its applications can be found e.g. in the textbook by Peskin and Schroeder [7], Section 9.6.

- $\mathcal{N}$ is a normalization constant,
- the domain of integration is defined by the *in* and *out* states,
- $\mathcal{D}\varphi$ is invariant under translations,
- $\mu(\varphi)$ is, to leading order, given by the advanced Green's function G^+ :

$$\mu(\varphi) = |\text{sdet } G^+(\varphi)|^{-1/2} + \dots, \quad (41)$$

where "sdet" is the superdeterminant. In (41) the advanced Green's function G^+ is the unique inverse of the leading non-constant term S'' of the expansion of S restricted to its domain of integration $\Phi(\text{in}, \text{out})$:

$$S(\varphi_0 + \delta\varphi) = S(\varphi_0) + \frac{1}{2} S''(\varphi_0) \delta\varphi \delta\varphi + \dots, \quad S'(\varphi_0) = 0. \quad (42)$$

Equations (35) and (40) define volume forms $\mu(\varphi)\mathcal{D}\varphi$ and $\mathcal{D}x$, respectively. In both cases $\mathcal{D}$ is a translation-invariant symbol. For comparing the structure of the two volume forms, we write the finite-dimensional version of (35) with $s = 1$:

$$\int_{\mathbb{R}^D} \mathcal{D}x \exp(-\pi Q_{\alpha\beta} x^\alpha x^\beta) \exp(-2\pi i x'_\alpha x^\alpha) = \exp(-\pi W^{\alpha\beta} x'_\alpha x'_\beta), \quad (43)$$

where

$$Q_{\alpha\beta} W^{\beta\gamma} = \delta_\alpha^\gamma, \quad (44)$$

$$\begin{aligned} \mathcal{D}x &= (\det Q_{\alpha\beta})^{1/2} dx^1 \dots dx^D \\ &= (\det W^{\alpha\beta})^{-1/2} dx^1 \dots dx^D. \end{aligned} \quad (45)$$

Hence G^+ is to field theory what W is to a Gaussian on $\mathbb{R}^D$, that is, the covariance matrix:

$$W^{\lambda\mu} = 2\pi \int_{\mathbb{R}^D} \mathcal{D}x \exp(-\pi Q_{\alpha\beta} x^\alpha x^\beta) x^\lambda x^\mu.$$

In (40) the term $(2\pi i/h)\langle J, \varphi \rangle$ corresponds to $-2\pi i\langle x', x \rangle$ in (35); they both mean that the left-hand side is a Fourier transform.

In finite dimensions it is easy to write explicitly $\mathcal{D}x$; in infinite dimensions it has meaning only in the context of the integral (35). However, this implicit definition is sufficient for computing functional integrals [5,6]. It is also easy to generalize it to cases other than Gaussians.

Using a generalized formulation of $\mathcal{D}x$, or using $\mu(\varphi)\mathcal{D}\varphi$ obtained from the solution of the Schwinger variational principle is a matter of choice and

is often dictated by the context. Once $\mu(\varphi)$, or at least its leading term, is defined, the following equation

$$\langle \text{out} | T(F(\varphi)) | \text{in} \rangle = \mathcal{N} \int \mu(\varphi) \mathcal{D}\varphi F(\varphi) \exp\left(\frac{i}{\hbar} S(\varphi)\right)$$

can be exploited in a variety of cases, i.e. in cases where $F(\varphi)$ is not simply $\exp(i\langle J, \varphi \rangle / \hbar)$.

3 Volume Forms in Differential Geometry

We shall use differential geometry for defining volume forms on finite-dimensional Riemannian and symplectic manifolds in a formulation which paves the way for the infinite-dimensional case. The knowledgeable reader for whom using the infinite limit of a finite volume element is, rightly, anathema, please bear with us. Finite-dimensional volume elements are useful in the following situations:

- A rule of thumb. A statement which is independent of the dimension of the space of interest has a chance to generalize to infinite-dimensional spaces; for example a Gaussian on $\mathbb{R}^D$ defined by (43) generalizes easily to (8).
- Infinite-dimensional spaces defined by a projective system of finite-dimensional spaces. This strategy was used in defining Feynman volume forms by their Fourier transforms [8,9].
- Differential calculus on Banach spaces, and differential geometry on Banach manifolds. They are natural generalizations of their finite dimensional counterparts. For this reason we propose a formula which defines volume elements by their Lie derivatives.

Let $\mathcal{L}_X$ be the Lie derivative with respect to a vector field X on a D -dimensional manifold M^D , either a (pseudo-)Riemannian manifold (M^D, g) or a symplectic manifold (M^{2N}, Ω) . The volume forms are, respectively,

$$\omega_g(x) = |\det g_{\alpha\beta}(x)|^{1/2} dx^1 \wedge \dots \wedge dx^D \quad \text{on } (M^D, g) \quad (46)$$

and

$$\omega_\Omega(x) = \frac{1}{N!} \Omega \wedge \dots \wedge \Omega \quad (N \text{ factors}) \quad \text{on } (M^{2N}, \Omega). \quad (47)$$

In canonical coordinates (p, q) ,

$$\Omega = \sum_{\alpha} dp_{\alpha} \wedge dq^{\alpha} \quad (48)$$

and

$$\omega_{\Omega} = dp_1 \wedge dq^1 \wedge \dots \wedge dp_N \wedge dq^N. \quad (49)$$

Surprisingly ω_g and ω_{Ω} satisfy equations of the same structure:

$$\mathcal{L}_X \omega_g = \frac{1}{2} \text{Tr} (g^{-1} \mathcal{L}_X g) \omega_g, \quad (50)$$

$$\mathcal{L}_X \omega_{\Omega} = \frac{1}{2} \text{Tr} (\Omega^{-1} \mathcal{L}_X \Omega) \omega_{\Omega}. \quad (51)$$

Riemannian and symplectic geometry are notoriously different (see e.g. McDuff [10]) and the analogies between them are not superficial. For instance, with Riemannian geometry on the left and symplectic geometry on the right

$\int ds$	$\int \Omega$	
geodesics	minimal surfaces	
$\mathcal{L}_X g = 0$ defines	$\mathcal{L}_X \Omega = 0$ defines	
Killing vector	Hamiltonian vector fields	(52)

Killing vector fields are few, Hamiltonian vector fields are many.

3.1 The General Case

Before proving Eqs. (50) and (51), we consider the more general equation

$$\mathcal{L}_X \omega = D(X) \cdot \omega \quad (53)$$

or its integrated formulation

$$\int_M (\mathcal{L}_X F) \omega = - \int_M F \mathcal{L}_X \omega = - \int_M F D(X) \cdot \omega, \quad (54)$$

where ω is a top form (a D -dimensional form on M^D) and $D(X)$ is a function on M depending on the vector field X on M .

- Properties of $D(X)$ dictated by properties of $\mathcal{L}_X$:

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X \Leftrightarrow D([X,Y]) = X(D(Y)) - Y(D(X)) \quad (55)$$

On a top form:

$$\mathcal{L}_{fX} = f\mathcal{L}_X + X(f) \Leftrightarrow D(fX) = fD(X) + X(f). \quad (56)$$

PROOF: On a top form ω , the Cartan formula $\mathcal{L}_X = di_X + i_X d$ yields

$$\mathcal{L}_X \omega = di_X \omega,$$

and since $i_{fX} \omega = i_X(f\omega)$, we have $\mathcal{L}_{fX} \omega = di_X(f\omega) = \mathcal{L}_X(f\omega)$. ■

- In coordinates,

$$\omega_\mu(x) = \mu(x) dx^1 \wedge \dots \wedge dx^D = \mu(x) d^D x. \quad (57)$$

By the Leibniz rule,

$$\mathcal{L}_X(\mu d^D x) = \mathcal{L}_X(\mu) d^D x + \mu \mathcal{L}_X(d^D x). \quad (58)$$

Because $d^D x$ is a top form on M^D ,

$$\mathcal{L}_X(d^D x) = d(i_X d^D x) = X^\alpha{}_{,\alpha} d^D x. \quad (59)$$

Finally, combining (57), (58) and (59),

$$\begin{aligned} \mathcal{L}_X \omega_\mu &= (X^\alpha \mu_{,\alpha} + \mu X^\alpha{}_{,\alpha}) d^D x \\ &= D(X) \cdot \omega_\mu, \end{aligned} \quad (60)$$

with

$$\begin{aligned} D(X) &= (X^\alpha \mu_{,\alpha} + \mu X^\alpha{}_{,\alpha}) \mu^{-1} \\ &= X^\alpha{}_{,\alpha} + X^\alpha (\log |\mu|)_{,\alpha}. \end{aligned} \quad (61)$$

3.2 The Riemannian Case (M, g)

Let $\omega_g(x) = \mu(x) d^D x$. We shall show that the basic equation (50) is satisfied if and only if $\mu(x) = \text{const } |\det g(x)|^{1/2}$. Indeed

$$(\mathcal{L}_X g)_{\alpha\beta} = X^\gamma g_{\alpha\beta,\gamma} + g_{\gamma\beta} X^\gamma{}_{,\alpha} + g_{\alpha\gamma} X^\gamma{}_{,\beta} \quad (62)$$

and

$$\text{Tr}(g^{-1} \mathcal{L}_X g) = g^{\beta\alpha} X^\gamma g_{\alpha\beta,\gamma} + 2X^\alpha{}_{,\alpha} \quad (63)$$

and, as already computed in Eq. (60),

$$\mathcal{L}_X(\mu(x) d^D x) = (X^\alpha \mu_{,\alpha} + \mu X^\alpha{}_{,\alpha}) d^D x. \quad (64)$$

Therefore the basic equation (50) is satisfied if, and only if

$$(X^\gamma \mu_{,\gamma} + \mu X^\alpha_{,\alpha}) \mu^{-1} = \frac{1}{2} (g^{\beta\alpha} X^\gamma g_{\alpha\beta,\gamma} + 2 X^\alpha_{,\alpha}), \quad (65)$$

i.e.

$$\frac{\mu_{,\gamma}}{\mu} = \frac{1}{2} g^{\beta\alpha} g_{\alpha\beta,\gamma} = \frac{1}{2} \partial_\gamma \ln |\det g|, \quad (66)$$

$$\mu(x) = \text{const. } |\det g(x)|^{1/2}. \quad (67)$$

The equation

$$\mathcal{L}_X \omega_g = \frac{1}{2} \text{Tr} (g^{-1} \mathcal{L}_X g) \omega_g \quad (68)$$

has, up to multiplication by a constant, a unique solution

$$\omega_g(x) = |\det g(x)|^{1/2} dx^1 \wedge \dots \wedge dx^D. \quad \blacksquare \quad (69)$$

We use the classical formula

$$\Gamma^\alpha_{\alpha\gamma} = \frac{1}{2} g^{\beta\alpha} g_{\alpha\beta,\gamma}, \quad (70)$$

with the Christoffel symbols $\Gamma^\alpha_{\beta\gamma}$. Hence (63) says

$$\frac{1}{2} \text{Tr} (g^{-1} \mathcal{L}_X g) = X^\alpha_{;\alpha} =: \text{Div}_g(X) \quad (71)$$

with the standard definition of the covariant divergence $X^\alpha_{;\alpha}$ of the vector field X , and we can write the basic equation (50) in the form

$$\mathcal{L}_X \omega_g = \text{Div}_g(X) \cdot \omega_g. \quad (72)$$

If X is a Killing vector field with respect to isometries, then $\mathcal{L}_X g = 0$, $\mathcal{L}_X \omega_g = 0$ and $X_{\alpha;\beta} + X_{\beta;\alpha} = 0$. Hence $X^\alpha_{;\alpha} = 0$ and (72) is satisfied.

3.3 The Symplectic Case (M^D, Ω) , $D = 2N$

The symplectic form Ω on M^{2N} is a closed 2-form of rank $D = 2N$.

$$\begin{aligned} \Omega &= \Omega_{AB} dx^A \wedge dx^B && \text{with } A < B, \Omega_{AB} = -\Omega_{BA}, d\Omega = 0 \\ &= \frac{1}{2} \Omega_{\alpha\beta} dx^\alpha \wedge dx^\beta && \text{no restriction on the order of } \alpha, \beta \\ &= \frac{1}{2} \Omega_{\alpha\beta} (dx^\alpha \otimes dx^\beta - dx^\beta \otimes dx^\alpha) \\ &= \Omega_{\alpha\beta} dx^\alpha \otimes dx^\beta, \end{aligned} \quad (73)$$

since $\Omega_{\alpha\beta} = -\Omega_{\beta\alpha}$.

REMARK: There are two different definitions of the exterior product, each with its concomitant definition of exterior derivative, e.g.

$$dx^1 \wedge dx^2 = dx^1 \otimes dx^2 - dx^2 \otimes dx^1, \quad (74)$$

$$\tilde{dx}^1 \tilde{\wedge} \tilde{dx}^2 = \frac{1}{2} \left(\tilde{dx}^1 \otimes \tilde{dx}^2 - \tilde{dx}^2 \otimes \tilde{dx}^1 \right). \quad (75)$$

With the second definition, Stokes' formula for a p -form θ reads $\int_M \tilde{d}\theta = (p+1) \int_{\partial M} \theta$; with the first one, it is simply $\int_M d\theta = \int_{\partial M} \theta$. We choose the first definition, namely

$$df^1 \wedge \dots \wedge df^p = \epsilon_{j_1 \dots j_p} df^{j_1} \otimes \dots \otimes df^{j_p},$$

and in particular

$$dx^1 \wedge \dots \wedge dx^D = \epsilon_{j_1 \dots j_D} dx^{j_1} \otimes \dots \otimes dx^{j_D},$$

where ϵ is totally antisymmetric. Since Ω is of rank $D = 2N$,

$$\Omega^{\wedge N} := \Omega \wedge \dots \wedge \Omega \quad (N \text{ factors})$$

is a nonzero top form on M^{2N} and the volume element

$$\omega_\Omega = \frac{1}{N!} \Omega^{\wedge N} \quad (76)$$

$$= \text{Pf}(\Omega_{\alpha\beta}) d^D x = |\det \Omega_{\alpha\beta}|^{1/2} d^D x. \quad (77)$$

We shall show that the basic equation (51) is satisfied if and only if ω_Ω is proportional to the volume form (76).

PROOF: We define the inverse Ω^{-1} of Ω , calculate the quantity $\frac{1}{2} \text{Tr}(\Omega^{-1} \mathcal{L}_X \Omega)$, then prove the basic formula (51).

- The symplectic form Ω defines an isomorphism from the tangent bundle TM to the cotangent bundle T^*M by

$$\Omega : X \mapsto i_X \Omega.$$

We can then define

$$X_\alpha := X^\beta \Omega_{\beta\alpha}. \quad (78)$$

The inverse $\Omega^{-1} : T^*M \longrightarrow TM$ is given by

$$X^\alpha = X_\beta \Omega^{\beta\alpha},$$

with

$$\Omega^{\alpha\beta} \Omega_{\beta\gamma} = \delta_{\gamma}^{\alpha}. \quad (79)$$

Note that in strict components, i.e. with $\Omega = \Omega_{AB} dx^A \wedge dx^B$ with $A < B$, X_A is not equal to $X^B \Omega_{BA}$.

- We compute

$$\begin{aligned} (\mathcal{L}_X \Omega)_{\alpha\beta} &= X^{\gamma} \Omega_{\alpha\beta,\gamma} + \Omega_{\gamma\beta} X^{\gamma}_{,\alpha} + \Omega_{\alpha\gamma} X^{\gamma}_{,\beta} \\ &= X_{\beta,\alpha} - X_{\alpha,\beta} \end{aligned} \quad (80)$$

using $d\Omega = 0$, that is $\Omega_{\beta\gamma,\alpha} + \Omega_{\gamma\alpha,\beta} + \Omega_{\alpha\beta,\gamma} = 0$. Hence

$$(\Omega^{-1} \mathcal{L}_X \Omega)^{\gamma}_{\beta} = \Omega^{\gamma\alpha} (X_{\beta,\alpha} - X_{\alpha,\beta})$$

and

$$\frac{1}{2} \text{Tr} (\Omega^{-1} \mathcal{L}_X \Omega) = \Omega^{\gamma\alpha} X_{\gamma,\alpha}. \quad (81)$$

- According to Darboux' theorem, there is a coordinate system (x^{α}) in which the volume form $\omega_{\Omega} = \Omega^{\wedge N} / N!$ is

$$\omega_{\Omega} = dx^1 \wedge \dots \wedge dx^{2N}, \quad (82)$$

and $\Omega = \Omega_{\alpha\beta} dx^{\alpha} \otimes dx^{\beta}$ with *constant* coefficients $\Omega_{\alpha\beta}$. The inverse matrix $\Omega^{\beta\alpha}$ is also made of constants, hence $\Omega^{\beta\alpha}_{,\gamma} = 0$.

In these coordinates

$$\begin{aligned} \mathcal{L}_X \omega_{\Omega} &= X^{\alpha}_{,\alpha} \omega_{\Omega} = (X_{\beta} \Omega^{\beta\alpha})_{,\alpha} \omega_{\Omega} \\ &= (X_{\beta,\alpha} \Omega^{\beta\alpha} + X_{\beta} \Omega^{\beta\alpha}_{,\alpha}) \omega_{\Omega} = X_{\beta,\alpha} \Omega^{\beta\alpha} \omega_{\Omega} \end{aligned} \quad (83)$$

and we conclude by using (81). ■

If X is a Hamiltonian vector field, then $\mathcal{L}_X \Omega = 0$ and $\mathcal{L}_X \omega_{\Omega} = 0$. The basic equation (51) is trivially satisfied.

4 Conclusion

Integration by parts is the key to the progress made in this paper for characterizing volume forms. It makes possible an infinitesimal characterization of the translation invariant symbol $\mathcal{D}$,

$$\int \mathcal{D}\varphi \frac{\delta U}{\delta \varphi(x)} = 0, \quad (84)$$

and is more powerful than its global translation (7),

$$\mathcal{D}(\varphi + \varphi_0) - \mathcal{D}\varphi = 0 . \quad (85)$$

The challenges we are now considering are the following:

- to extend to infinite-dimensional spaces the divergence formulae (50) and (51).
- to clarify the often observed relationship between the volume form and the Schrödinger equation satisfied by a functional integral.
- to develop issues mentioned briefly in this paper, in particular the Dobrushin-Lanford-Ruelle formula, and the annihilation/creation operators defined by the Malliavin formula.
- to derive the transformation laws of volume elements under the Cartan development mapping between two spaces of pointed paths on different manifolds.
- to extend the method from ordinary (bosonic) integration to Berezin (fermionic) integration.

Acknowledgments

Modern means of communication do not replace face-to-face brainstorming, and a group based in Paris and Austin functions due to means for travel expenses. We thank a steadfast friend John Tate who contributes to the annual visits of Pierre Cartier to Austin, and we thank the Jane and Roland Blumberg Centennial Professorship for partial support of Cécile DeWitt-Morette's visits to Paris. M. Berg is grateful to the Swedish Foundation for International Cooperation in Research and Higher Education.

References

- [1] N. Bourbaki, *Intégration* (Masson, Paris, 1982), Chapitre 9.
- [2] G. Gallavotti, *Statistical Mechanics: A Short Treatise* (Springer-Verlag, Berlin, 1999). Includes original DLR references.
- [3] P. Cartier, unpublished notes.
- [4] S. Janson, *Gaussian Hilbert Spaces* (Cambridge University Press, Cambridge, 1997).
- [5] P. Cartier and C. DeWitt-Morette, *J. Math. Phys.* **36**, 2237 (1995).
- [6] P. Cartier and C. DeWitt-Morette, *J. Math. Phys.* **41**, 4154 (2000).

- [7] M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, 1995).
- [8] C. DeWitt-Morette, *Commun. Math. Phys.* **28**, 47 (1972).
- [9] C. DeWitt-Morette, A. Maheshwari, and B. Nelson, *Phys. Rep.* **50**, 255 (1979).
- [10] D. McDuff, *Notices of AMS* **45**, 952 (1998).

VASSILIEV INVARIANTS AND FUNCTIONAL INTEGRATION

L.H. KAUFFMAN

*Department of Mathematics, Statistics and Computer Science,
University of Illinois at Chicago, 851 South Morgan Street,
Chicago, IL 60607-7045, USA
E-mail: kauffman@uic.edu*

This paper, dedicated to Hagen Kleinert, is an exposition of the relationship between the Witten functional integral and the theory of Vassiliev invariants of knots and links in three-dimensional space.

1 Introduction

In this article we want to show how Vassiliev invariants in knot theory arise naturally in context of the Witten functional integral. The relationship between Vassiliev invariants and the Witten integral has been known since Bar-Natan's thesis [1] where he discovered, through this connection, how to define Lie algebraic weight systems for these invariants.

The paper is a sequel to the Refs. [2–4] and an expanded version of a talk given at The Fifth Taiwan International Symposium on Statistical Physics (August 1999). In these papers we show more about the relationship of Vassiliev invariants and the Witten functional integral. In particular, we investigate how the Kontsevich integrals, used to give rigorous definitions of these invariants, arise as Feynman integrals in the perturbative expansion of the Witten functional integral; see also the work of Labastida and Pérez [5] on this same subject. Their work comes to an identical conclusion, interpreting the Kontsevich integrals in terms of the light-cone gauge and thereby extending the original work of Fröhlich and King [6]. The purpose of this paper is to give an exposition of the beginnings of these relationships and to introduce diagrammatic techniques that illuminate the connections.

This article is divided into two sections. First we discuss Vassiliev invariants and invariants of rigid vertex graphs and then we introduce the basic formalism and show how the functional integral is related directly to Vassiliev invariants.

It gives me great pleasure to dedicate this paper to Hagen Kleinert. His pioneering work [7] in the applications of functional integration to physical problems and his interest in topological work has sustained my own interest in this field since we met in 1996 in a conference on this very topic held in Cargèse, Corsica, and organized by Pierre Cartier and Cécile DeWitt-Morette [8].

2 Vassiliev Invariants and Invariants of Rigid Vertex Graphs

If $V(K)$ is a Laurent polynomial valued, or more generally, commutative ring valued invariant of knots, then it can be naturally extended to an invariant of rigid vertex graphs [9] by defining the invariant of graphs in terms of the knot invariant via an "unfolding" of the vertex. That is, we can regard the vertex as a "black box" and replace it by any tangle of our choice. Rigid vertex motions of the graph preserve the contents of the black box, and hence implicate ambient isotopies of the link obtained by replacing the black box by its contents. Invariants of knots and links that are evaluated on these replacements are then automatically rigid vertex invariants of the corresponding graphs. If we set up a collection of multiple replacements at the vertices with standard conventions for the insertions of the tangles, then a summation over all possible replacements can lead to a graph invariant with new coefficients corresponding to the different replacements. In this way each invariant of knots and links implicates a large collection of graph invariants [9,10]. The simplest tangle replacements for a 4-valent vertex are the two crossings, positive and negative, and the oriented smoothing. Let $V(K)$ be any invariant of knots and links. Extend V to the category of rigid vertex embeddings of 4-valent graphs by the formula

$$V(K_*) = aV(K_+) + bV(K_-) + cV(K_0), \quad (1)$$

where K_+ denotes a knot diagram K with a specific choice of positive crossing, K_- denotes a diagram identical to the first with the positive crossing replaced by a negative crossing, and K_* denotes a diagram identical to the first with the positive crossing replaced by a graphical node. This formula means that

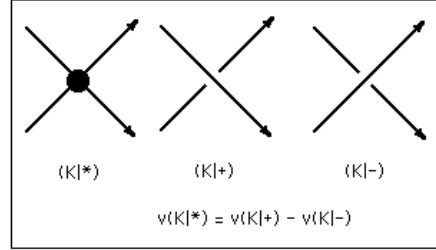


Figure 1. Exchange identity for Vassiliev invariants.

we define $V(G)$ for an embedded 4-valent graph G by taking the sum

$$V(G) = \sum_S a^{i_+(S)} b^{i_-(S)} c^{i_0(S)} V(S), \quad (2)$$

with the summation over all knots and links S obtained from G by replacing a node of G with either a crossing of positive or negative type, or with a smoothing of the crossing that replaces it by a planar embedding of non-touching segments (denoted 0). It is not hard to see that if $V(K)$ is an ambient isotopy invariant of knots, then, this extension is a rigid vertex isotopy invariant of graphs. In rigid vertex isotopy the cyclic order at the vertex is preserved, so that the vertex behaves like a rigid disk with flexible strings attached to it at specific points.

There is a rich class of graph invariants that can be studied in this manner. The Vassiliev invariants [11,12] constitute the important special case of these graph invariants where $a = +1$, $b = -1$ and $c = 0$:

$$V(K_*) = V(K_+) - V(K_-). \quad (3)$$

Call this formula the *exchange identity* for the Vassiliev invariant V (see Fig. 1). Thus $V(G)$ is a Vassiliev invariant if V is said to be of *finite type k* if $V(G) = 0$ whenever $|G| > k$ where $|G|$ denotes the number of (4-valent) nodes in the graph G . The notion of finite type is of extraordinary significance in studying these invariants. One reason for this is the following basic Lemma.

LEMMA. If a graph G has exactly k nodes, then the value of a Vassiliev invariant v_k of type k on G , i.e. $v_k(G)$, is independent of the embedding of G .

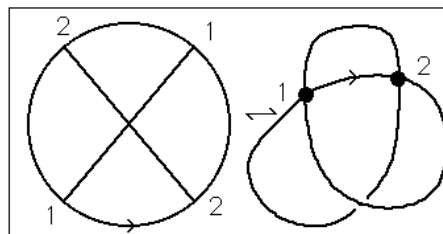


Figure 2. Chord diagrams.

PROOF. The different embeddings of G can be represented by link diagrams with some of the 4-valent vertices in the diagram corresponding to the nodes of G . It suffices to show that the value of $v_k(G)$ is unchanged under switching of a crossing. However, the exchange identity for v_k shows that this difference is equal to the evaluation of v_k on a graph with $k+1$ nodes and hence is equal to zero. This completes the proof.

The upshot of this Lemma is that Vassiliev invariants of type k are intimately involved with certain abstract evaluations of graphs with k nodes. In fact, there are restrictions (the four-term relations) on these evaluations demanded by the topology and it follows from results of Kontsevich [12] that such abstract evaluations actually determine the invariants. The knot invariants derived from classical Lie algebras are all built from Vassiliev invariants of finite type. All of this is directly related to the Witten functional integral [13].

In the next few figures we illustrate some of these main points. In Fig. 2 we show how one associates a so-called chord diagram to represent the abstract graph associated with an embedded graph. The chord diagram is a circle with arcs connecting those points on the circle that are welded to form the corresponding graph. In Fig. 3 we illustrate how the four-term relation is a consequence of topological invariance. In Fig. 4 we show how the four-term relation is a consequence of the abstract pattern of the commutator identity for a matrix Lie algebra. This shows that the four-term relation is directly related to a categorical generalisation of Lie algebras. Figure 5 illustrates how the weights are assigned to the chord diagrams in the Lie algebra case by inserting Lie algebra matrices into the circle and taking a trace of a sum of matrix products.

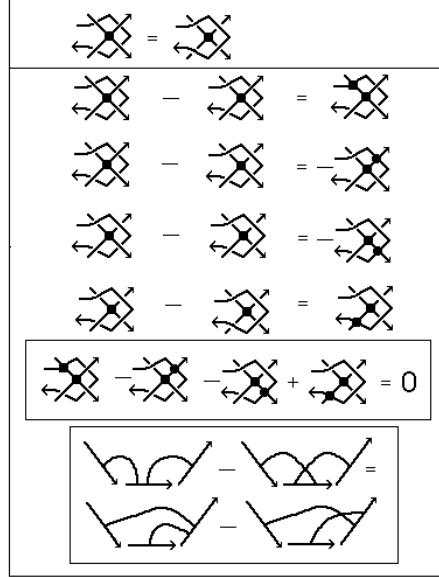


Figure 3. The four-term relation from topology.

3 Vassiliev Invariants and Witten Functional Integral

In Ref. [13] Edward Witten proposed a formulation of a class of 3-manifold invariants (see also Ref. [14]) which generalize Feynman integrals taking the form

$$Z(M) = \int DA e^{(ik/4\pi)S(M,A)}. \quad (4)$$

Here M denotes a 3-manifold without boundary and A is a gauge field, also called a gauge potential or gauge connection, defined on M . The gauge field is a one-form on a trivial G -bundle over M with values in a representation of the Lie algebra of G . The group G corresponding to this Lie algebra is said to be the gauge group. In this integral the action $S(M, A)$ is taken to be the integral over M of the trace of the Chern-Simons three-form $A \wedge dA + (2/3)A \wedge A \wedge A$, where the product is the wedge product of differential forms. $Z(M)$ integrates over all gauge fields modulo gauge equivalence. The formalism and internal logic of Witten's integral supports the existence of a large class of topological invariants of 3-manifolds and associated invariants of knots and links in these

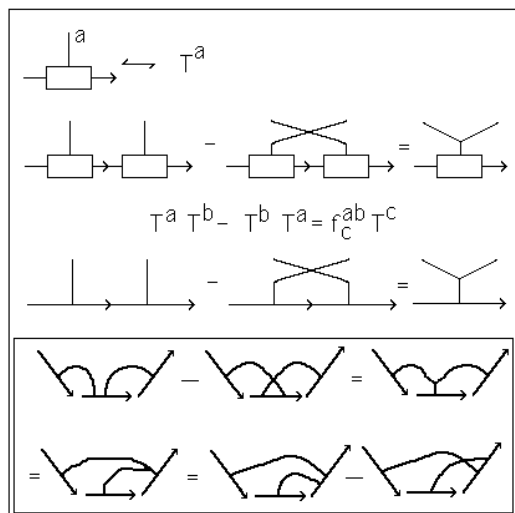


Figure 4. The four-term relation from categorical Lie algebra.

manifolds.

The invariants associated with this integral have been given rigorous combinatorial descriptions but questions and conjectures arising from the integral formulation are still outstanding. Specific conjectures about this integral take the form of just how it implicates invariants of links and 3-manifolds, and how these invariants behave in certain limits of the coupling constant k in the integral. Many conjectures of this sort can be verified through the combinatorial models. On the other hand, the really outstanding conjecture about the integral is that it exists! At the present time there is no measure theory or generalization of measure theory that supports it. Here is a formal structure of great beauty. It is also a structure whose consequences can be verified by a remarkable variety of alternative means.

We now look at the formalism of the Witten functional integral in more detail and see how it implicates invariants of knots and links corresponding to each classical Lie algebra. In order to accomplish this task, we need to introduce the Wilson loop. The Wilson loop is an exponentiated version of integrating the gauge field along a loop K in three-space that we take to be an embedding (knot) or a curve with transversal self-intersections. For this discussion, the Wilson loop will be denoted by the notation $W_K(A) = \langle K|A \rangle$

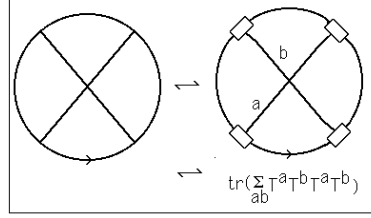


Figure 5. Calculating Lie algebra weights.

to stress the dependence on the loop K and the field A . It is usually indicated by the symbolism

$$W_K(A) = \langle K|A \rangle = \text{tr} \left(P e^{\oint_K A} \right). \quad (5)$$

Here P denotes path ordered integration. As we are integrating and exponentiating matrix valued functions, we must keep track of the order of the operations. The symbol tr denotes the trace of the resulting matrix. With the help of the Wilson loop functional on knots and links, Witten writes down a functional integral for link invariants in a 3-manifold M :

$$Z(M, K) = \int DA e^{(ik/4\pi)S(M,A)} \text{tr} \left(P e^{\oint_K A} \right) = \int DA e^{(ik/4\pi)S} \langle K|A \rangle. \quad (6)$$

Here $S(M, A)$ is the Chern-Simons Lagrangian as in the previous discussion. We abbreviate $S(M, A)$ as S and write $\langle K|A \rangle$ for the Wilson loop. Unless otherwise mentioned, the manifold M will be the three-dimensional sphere S^3 .

An analysis of the formalism of this functional integral reveals quite a bit about its role in knot theory. This analysis depends upon key facts relating the curvature of the gauge field to both the Wilson loop and the Chern-Simons Lagrangian. The idea for using the curvature in this way is due to Lee Smolin [15] (see also Ref. [16]). To this end, let us recall the local coordinate structure of the gauge field $A(x)$, where x is a point in three-space. We can write $A(x) = A_k^a(x) T_a dx^k$ where the index a ranges from 1 to m with the Lie algebra basis $\{T_1, T_2, T_3, \dots, T_m\}$ and the index k goes from 1 to 3. For each choice of a and k , $A_k^a(x)$ is a smooth function defined on three-space. In $A(x)$ we sum over the values of repeated indices. The Lie algebra generators T_a are matrices corresponding to a given representation of the Lie algebra of the gauge group G . We assume some properties of these matrices as follows:

- (1) $[T_a, T_b] = if^{abc}T_c$ where $[x, y] = xy - yx$, and the matrix of structure constants f^{abc} is totally antisymmetric. There is summation over repeated indices.
- (2) $\text{tr}(T_a T_b) = \delta_{ab}/2$ where δ_{ab} is the Kronecker delta, i.e. $\delta_{ab} = 1$ if $a = b$ and zero otherwise.

We also assume some facts about curvature (the reader may enjoy comparing with the exposition in Ref. [17]; but note the difference of conventions on the use of i in the Wilson loops and the curvature definitions). The first fact is the relation of Wilson loops and curvature for small loops:

FACT 1. The result of evaluating a Wilson loop about a very small planar circle around a point x is proportional to the area enclosed by this circle times the corresponding value of the curvature tensor of the gauge field evaluated at x . The curvature tensor is written $F_{rs}^a(x)T_a dx^r dy^s$. It is the local coordinate expression of $F = dA + A \wedge A$.

APPLICATION OF FACT 1. Consider a given Wilson line $\langle K|S \rangle$. Ask how its value will change if it is deformed infinitesimally in the neighborhood of a point x on the line. Approximate the change according to Fact 1, and regard the point x as the place of the curvature evaluation. Let $\delta\langle K|A \rangle$ denote the change in the value of the line. $\delta\langle K|A \rangle$ is given by the formula

$$\delta\langle K|A \rangle = dx^r dx^s F_{rs}^a(x) T_a \langle K|A \rangle. \quad (7)$$

This is the first-order approximation to the change in the Wilson line. In this formula it is understood that the Lie algebra matrices T_a are to be inserted into the Wilson line at the point x , and that we are summing over repeated indices. This means that each $T_a \langle K|A \rangle$ is a new Wilson line obtained from the original line $\langle K|A \rangle$ by leaving the form of the loop unchanged, but inserting the matrix T_a into that loop at the point x . In Fig. 6 we have illustrated this mode of insertion of Lie algebra into the Wilson loop. Here and in further illustrations in this section we use $W_K(A)$ to denote the Wilson loop. Note that in the diagrammatic version shown in Fig. 6 we have let small triangles with legs indicate dx^i . The legs correspond to indices just as in our work in the last section with Lie algebras and chord diagrams. The curvature tensor is indicated as a circle with three legs corresponding to the indices of F_a^{rs} .

NOTATION. In the diagrams in this section we have dropped mention-

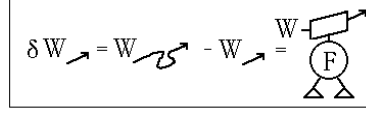


Figure 6. Lie algebra and curvature tensor insertion into the Wilson loop.

ing the factor of $1/4\pi$ that occurs in the integral. This convention saves space in the figures. In these figures L denotes the Chern-Simons Lagrangian.

REMARK. In thinking about the Wilson line (5), it is helpful to recall Euler's formula for the exponential:

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n. \quad (8)$$

The Wilson line is the limit, over partitions of the loop K , of products of the matrices $(1 + A(x))$ where x runs over the partition. Thus we can write symbolically

$$\langle K|A \rangle = \prod_{x \in K} (1 + A(x)) = \prod_{x \in K} (1 + A_k^a(x) T_a dx^k). \quad (9)$$

It is understood that a product of matrices around a closed loop connotes the trace of the product. The ordering is forced by the one-dimensional nature of the loop. Inserting a given matrix into this product at a point on the loop is then a well-defined concept. If T is a given matrix then it is understood that $T\langle K|A \rangle$ denotes the insertion of T into some point of the loop. In the case above, it is understood from the context in the formula that the insertion is to be performed at the point x indicated in the argument of the curvature.

REMARK. The previous remark implies the following formula for the variation of the Wilson loop with respect to the gauge field:

$$\frac{\delta \langle K|A \rangle}{\delta(A_k^a(x))} = dx^k T_a \langle K|A \rangle. \quad (10)$$

Varying the Wilson loop with respect to the gauge field results in inserting an infinitesimal Lie algebra element into the loop. Figure 7 gives a diagrammatic form for this formula. In that figure we use a capital D with up and down legs to denote the derivative $\delta/\delta(A_k^a(x))$. Insertions in the Wilson line are

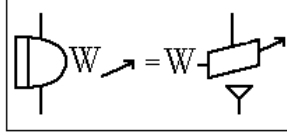


Figure 7. Differentiating the Wilson line.

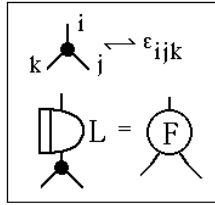


Figure 8. Variational formula for curvature.

indicated directly by matrix boxes placed in a representative bit of line.

PROOF.

$$\begin{aligned} \frac{\delta \langle K|A \rangle}{\delta(A_k^a(x))} &= \frac{\delta}{\delta(A_k^a(x))} \prod_{y \in K} (1 + A_k^a(y) T_a dy^k) = \prod_{y < x \in K} (1 + A_k^a(y) T_a dy^k) T_a dx^k \\ &\times \prod_{y > x \in K} (1 + A_k^a(y) T_a dy^k) = dx^k T_a \langle K|A \rangle. \end{aligned} \quad (11)$$

FACT 2. The variation of the Chern-Simons Lagrangian S with respect to the gauge potential at a given point in three-space is related to the values of the curvature tensor at that point by the following formula:

$$F_{rs}^a(x) = \epsilon_{rst} \frac{\delta S}{\delta(A_t^a(x))}. \quad (12)$$

Here ϵ_{abc} is the epsilon symbol for three indices, i.e. it is +1 for positive permutations of 123 and -1 for negative permutations of 123 and zero if any two indices are repeated. A diagrammatic representation for this formula is shown in Fig. 8. With these facts at hand we are prepared to determine how the Witten functional integral behaves under a small deformation of the loop K .

THEOREM 1. Let $Z(K) = Z(S^3, K)$ and let $\delta Z(K)$ denote the change of $Z(K)$ under an infinitesimal change in the loop K . Then

$$\delta Z(K) = (4\pi i/k) \int dA e^{(ik/4\pi)S} [\text{Vol}] T_a T_a \langle K|A \rangle, \quad (13)$$

where $\text{Vol} = \epsilon_{rst} dx^r dx^s dx^t$. The sum is taken over repeated indices, and the insertion is taken of the matrices $T_a T_a$ at the chosen point x on the loop K that is regarded as the center of the deformation. The volume element $\text{Vol} = \epsilon_{rst} dx^r dx^s dx^t$ is taken with regard to the infinitesimal directions of the loop deformation from this point on the original loop.

THEOREM 2. The same formula applies, with a different interpretation, to the case, where x is a double point of transversal self-intersection of a loop K , and the deformation consists in shifting one of the crossing segments perpendicularly to the plane of intersection so that the self-intersection point disappears. In this case, one T_a is inserted into each of the transversal crossing segments so that $T_a T_a \langle K|A \rangle$ denotes a Wilson loop with a self-intersection at x and insertions of T_a at $x + \epsilon_1$ and $x + \epsilon_2$ where ϵ_1 and ϵ_2 denote small displacements along the two arcs of K that intersect at x . In this case, the volume form is nonzero, with two directions coming from the plane of movement of one arc, and the perpendicular direction is the direction of the other arc.

PROOF.

$$\begin{aligned} \delta Z(K) &= \int DA e^{(ik/4\pi)S} \delta \langle K|A \rangle = \int DA e^{(ik/4\pi)S} dx^r dy^s F_{rs}^a(x) T_a \langle K|A \rangle \\ &= \int DA e^{(ik/4\pi)S} dx^r dy^s \epsilon_{rst} \frac{\delta S}{\delta(A_t^a(x))} T_a \langle K|A \rangle \\ &= (-4\pi i/k) \int DA \frac{\delta e^{(ik/4\pi)S}}{\delta(A_t^a(x))} \epsilon_{rst} dx^r dy^s T_a \langle K|A \rangle \\ &= (4\pi i/k) \int DA e^{(ik/4\pi)S} \epsilon_{rst} dx^r dy^s \frac{\delta T_a \langle K|A \rangle}{\delta(A_t^a(x))} \\ &= (4\pi i/k) \int DA e^{(ik/4\pi)S} [\text{Vol}] T_a T_a \langle K|A \rangle. \end{aligned} \quad (14)$$

This completes the formalism of the proof. In the case of part 2, a change of interpretation occurs at the point in the argument when the Wilson line is differentiated. Differentiating a self-intersecting Wilson line at a point of self-intersection is equivalent to differentiating the corresponding product of

$$\begin{aligned}
\delta z &= \int DA e^{ikL} \delta W \\
&= \int DA e^{ikL} \text{ (diagram with a loop labeled F and a crossing)} \\
&= \int DA e^{ikL} \text{ (diagram with a loop labeled L and a crossing)} \\
&= (-i/k) \int DA e^{ikL} \text{ (diagram with a loop labeled L and a crossing)} \\
&= (i/k) \int DA e^{ikL} \text{ (diagram with a loop labeled L and a crossing)} \\
&= (i/k) \int DA e^{ikL} \text{ (diagram with a loop labeled L and a crossing)}
\end{aligned}$$

Figure 9. Varying the functional integral by varying the line.

matrices with respect to a variable that occurs at two points in the product (corresponding to the two places where the loop passes through the point). One of these derivatives gives rise to a term with volume form equal to zero, the other term is the one that is described in part 2. This completes the proof of the theorem.

The formalism of this proof is illustrated in Fig. 9. In the case of switching a crossing the key point is to write the crossing-switch as a composition of first moving a segment to obtain a transversal intersection of the diagram with itself, and then to continue the motion to complete the switch. One then analyzes separately the case where x is a double point of transversal self-intersection of a loop K , and the deformation consists in shifting one of the crossing segments perpendicularly to the plane of intersection so that the self-intersection point disappears. In this case, one T_a is inserted into each of the transversal crossing segments so that $T^a T^a \langle K|A \rangle$ denotes a Wilson loop with a self-intersection at x and insertions of T^a at $x + \epsilon_1$ and $x + \epsilon_2$ as in part 2 of the theorem above. The first insertion is in the moving line, due to curvature. The second insertion is the consequence of differentiating the self-touching Wilson line. Since this line can be regarded as a product, the differentiation occurs twice at the point of intersection, and it is the second direction that produces the non-vanishing volume form.

Up to the choice of our conventions for constants, the switching formula,

$$\begin{aligned}
 Z \text{ (crossing with dot)} &= Z \text{ (crossing without dot)} - Z \text{ (crossing without dot)} \\
 &= (c/k) Z \text{ (crossing with self-touching loop)} + O(1/k^2)
 \end{aligned}$$

Figure 10. The difference formula.

as shown in Fig. 10, reads

$$\begin{aligned}
 Z(K_+) - Z(K_-) &= (4\pi i/k) \int DA e^{(ik/4\pi)S} T_a T_a \langle K_{**} | A \rangle \\
 &= (4\pi i/k) Z(T^a T^a K_{**}), \quad (15)
 \end{aligned}$$

where K_{**} denotes the result of replacing the crossing by a self-touching crossing. We distinguish this from adding a graphical node at this crossing by using the double star notation. A key point is to notice that the Lie algebra insertion for this difference is exactly what is done (in chord diagrams) to make the weight systems for Vassiliev invariants (without the framing compensation). Here we take formally the perturbative expansion of the Witten functional integral to obtain Vassiliev invariants as coefficients of the powers of $1/k^n$. Thus the formalism of the Witten functional integral takes one directly to these weight systems in the case of the classical Lie algebras. In this way the functional integral is central to the structure of the Vassiliev invariants.

Acknowledgments

The author would like to thank the National Science Foundation for support of this research under NSF Grant DMS-9205277.

References

- [1] D. Bar-Natan, *Perturbative Aspects of the Chern-Simons Topological Quantum Field Theory*, Ph.D. Thesis (Princeton University, 1991).
- [2] L.H. Kauffman, *Physica A* **281**, 173 (2000).

- [3] L.H. Kauffman, *Witten's Integral and the Kontsevich Integrals*, in *Particles, Fields, and Gravitation*, Proceedings of the Conference on Mathematical Physics at Lodz, Poland, Ed. J. Remblienski, AIP Conference Proceedings **453** (1998), p. 368.
- [4] L.H. Kauffman, *J. Math. Phys.* **36**, 2402 (1995).
- [5] J.M.F. Labastida and E. Pérez, *J. Math. Phys.* **39**, 5183 (1998).
- [6] J. Fröhlich and C. King, *Commun. Math. Phys.* **126**, 167 (1989).
- [7] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [8] C. DeWitt-Morette, P. Cartier, and A. Folacci, *Functional Integration - Basics and Applications*, NATO ASI Series, Series B: Physics Vol. **361** (1997).
- [9] L.H. Kauffman, *Amer. Math. Monthly* **95**, 195 (1988).
- [10] L.H. Kauffman and P. Vogel, *Journal of Knot Theory and Its Ramifications* **1**, 59 (1992).
- [11] J. Birman and X.S. Lin, *Invent. Math.* **111**, 225 (1993).
- [12] D. Bar-Natan, *Topology* **34**, 423 (1995).
- [13] E. Witten, *Commun. Math. Phys.* **121**, 351 (1989).
- [14] M.F. Atiyah, *The Geometry and Physics of Knots* (Cambridge University Press, Cambridge, 1990).
- [15] L. Smolin, *Mod. Phys. Lett. A* **4**, 1091 (1989).
- [16] P. Cotta-Ramusino, E. Guadagnini, M. Martellini, and M. Mintchev, *Nucl. Phys. B* **330**, 557 (1990).
- [17] L.H. Kauffman, *Knots and Physics*, 2nd ed. (World Scientific, Singapore, 1993).

Part II

Quantum Field Theory

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

HIGHER ALGEBRAIC GEOMETRIZATION EMERGING FROM NONCOMMUTATIVITY

Y. NE'EMAN

*School of Physics and Astronomy,
Raymond and Beverly Sackler Faculty of Exact Sciences,
Tel-Aviv University, Tel-Aviv 69978, Israel
E-mail: matildae@tauex.tau.ac.il*

We review the gradual geometrization which has occurred in fundamental physics from the discovery of special relativity in 1905 to the standard model in 1975. After discussing symmetry and ordinary supersymmetry, we introduce internal supersymmetry. Here the even and odd generators correspond to the form-calculus of a gauge theory with spontaneous symmetry breakdown, with the gauge field one-forms occupying the even submatrices and the Higgs fields zero-forms occupying the off-diagonal submatrices. The Grassmann superalgebra is not (super)-abelian and closes on some semi-simple subalgebra. We study two examples: the electroweak $SU(2/1)$, predicting the mass of the Higgs particle around 130 ± 10 GeV, and $P(4R)$ for Riemannian gravity. Internal supersymmetry does not operate on the physical Hilbert space and as a result of non-commutative geometry, the matter fields in its fibres relate to $Z(2)$ gradings other than that of quantum statistics (chirality in our examples).

1 Introduction - and the Physics of Time

Hagen Kleinert is sixty according to *classical* clocks, but this is clearly a misinterpretation of the data. Observations show that Hagen and Annemarie *have not aged at all*. Had they both been born with twins, he with a twin brother, and she with a twin sister, we might have been able to explain our paradox as a complicated extension of the *twin paradox*, conjecturing that we are now facing the travelling twins, who have just returned and replaced the sedentary couple. However, as the Kleinerts never had a twin couple, we must be facing some as yet uncharted and unidentified relativistic effect, perhaps related to some unknown aspect of quantum gravity - coupled à la Penrose

with the collapse of the state-vector and thus to the Everett-Wheeler Many-Worlds interpretation With the effect still shrouded in mystery, I have had to go along and behave as if I believed that Hagen is indeed getting older (wiser he certainly is) and I am happy to dedicate this study to his (classical) sixtieth birthday and wish Annemarie and him many happy returns. I shall do my best to be around when Hagen's classical age is 75, and partake in the next *Festschrift*, but, alas, I cannot make a firm commitment on this matter.

2 Steps in the Geometrization Process

Hagen's contributions to Physics in the last decade have mostly been of a geometrical nature, involving Riemannian manifolds with torsion - whether in 3 dimensions and the physics of materials (transitions between phases, e.g. melting [1]) or in 4 dimensions and issues relating to general relativity [2]. We physicists of the late XXth (and hopefully of the early XXIst) century have enjoyed the aesthetics and symmetries of the *geometrical representation*. Gradually, between 1905 and 1975, it has become the unique language of physics at the fundamental level [3]. I remind the reader that this takeover occurred in the following stages:

(a) Minkowski's 1907-08 geometrical reinterpretation of Einstein's (1905) special theory of relativity, namely of the symmetries of Maxwell's electromagnetism, as identified by Einstein,

(b) The Einstein-Grossmann application and extension of that model (1911) in a program aiming at reconciling Newtonian mechanics with the above symmetries of electromagnetism, leading to Einstein's construction of the general theory of relativity (GR, 1915) as the new (and fully geometrical) theory of gravity,

(c) The construction of a *gauge theory*, started in 1918 with H. Weyl's (failed) first attempt at a theory of electromagnetism (based on the assumption of local scale-invariance) unifiable with Einstein's gravity (i.e. geometrical). It was followed by his 1928 successful version, in which the geometry is that of a *fibre bundle*, with fibre group $U(1)$ realizing local invariance under transformations of the complex phase angle (introduced by quantum mechanics). This was then generalized (1953) to non-abelian groups by C.N. Yang and R.L. Mills and applied (1975) to the $S[U(2) \otimes U(3)]$ fibre group of the standard model. The latter emerged, on the one hand, as a result of our (1961) $SU(3)_{\text{flavor}}$ classification of the hadrons and the subsequent (1962-64) discovery of the structural mechanism to which it is due, namely the *quark*

model, followed first (1964-1972) by the introduction of (global) $SU(3)_{\text{color}}$ for the sake of preservation of Fermi-type quantum statistics and then by quantum chromodynamics (QCD), namely *local* $SU(3)_{\text{color}}$, after the discovery (1973) of *asymptotic freedom*. The geometrical nature of local gauge theories was emphasized (1974) by C.N. Yang and T.T. Wu.

(d) Two other developments (also launched in the early twenties) were suggested in the context of *further unification*:

I. Adding dimensions - the program suggested by T. Kaluza (1921) and by O. Klein (1924).

II. Following Einstein, adding an *antisymmetric piece to the metric or connection*.

Developments in the seventies in the unification program, namely in $N_s > 1$ supergravity and in superstring (or "M") theory have converged on a fusion of both these features. After the establishment of relativistic quantum field theory by its success (1948) in quantum electrodynamics (QED), the gravitational field of GR, representing the "fabric" of space-time, should by itself be treated as a quantum field of Bose-type. This thereby does not allow a role for an antisymmetric metric by the spin-statistics theorem. The necessary conditions, however, are induced through (1971-73) *supersymmetry, which adds fermionic degrees of freedom to any boson*. Gravity thereby becomes embedded in supergravity, with the antisymmetric characteristics of (II) represented by the presence of torsion. Finiteness considerations, namely the cancellation of chiral and dilational anomalies, then impose specific higher dimensionalities as in (I). The two programs – torsion and Kaluza-Klein dimensionalities – are thus presently actively pursued in the context of the 11-dimensional $N_s = 1$ *supergravity* constructed by E. Cremmer and B. Julia (1978), reducible in 4-dimensions to the $N_s = 8$ *maximal* or *saturated* supergravity, a version of supergravity which has been shown to represent the low-energy quantum field theory limit of "M-theory", the *state of the art* theory of post-Planck level and quantized gravity. [N_s is the "number of supersymmetries", i.e. the dimensionality of the internal degree of freedom, if any, carried by the Lorentz spinor multiplets of supersymmetry generators].

(e) An independent additional geometrical entry is due to Jean Thierry-Mieg. His 1979 thesis and related articles [4,5] identify the *ghost fields* of a Yang-Mills gauge theory – as conceived by R.P. Feynman (1962) in order to guarantee *off mass shell unitarity* and further developed by B.S. De Witt, L.D. Faddeev, and V.N. Popov – with *odd* elements of the *form calculus*, the

Grassmann "supercommutative" superalgebra of the gauge theory. Moreover, the "BRST" constraining superalgebra, linking together physical and ghost fields is seen to coincide with the *structural equations* of the fibre bundle.

(f) The superalgebraic system I describe in the rest of this article derives from the latter. It consists in a *nonsupercommutative extended superalgebra of forms*, defined over a fibre-bundle *whose base-space is split by a $Z(2)$ grading* but which does not necessarily coincide with the $Z(2)$ of quantum statistics (in my examples, it will be the $Z(2)$ of *chirality*). The first model of this type was discovered (1979) by the present author [6] and independently by David Fairlie [7] and involved *the simple Lie supergroup $SU(2/1)$ as an "internal supersymmetry", an irreducible algebraic extension of electroweak unification's (spontaneously-broken) local - $SU(2) \times U(1)$ symmetry*. The theory has been applied [8] to predict the mass of the Higgs meson, yielding $m(H) = 2m(W)$ in the exact (and unrenormalized) limit, while the inclusion of renormalization effects, as observed in couplings [9], yields as final result $m(H) = 130 \pm 15$ GeV. Note that 9 "events" have been observed at CERN in the fall of 2000 with a Higgs meson mass around 115 GeV.

3 Superalgebras, Supermatrices, and $Z(2)$ -Gradings

I first remind the reader of the main definitions and results relating to Lie [10] and to Grassmann (super) algebras [11]. The first involve the application of a $Z(2)$ -grading on the basis of the Lie superalgebra as a linear vector space; the $g(x)$ eigenvalue also determines the nature of the $Z(2)$ -graded super-Lie bracket $[x, y]$ and of the relevant super-Jacobi identity. Let the variable $E = \sqrt{1}$ represent the two elements of the finite group $Z(2)$. The superalgebra splits into two subspaces, labelled by that grading,

$$\begin{aligned}
 E &= \sqrt{1}, \quad g = \log_{-1}(E \in Z(2)_g), \\
 L &= L_0 + L_1, \\
 g(x \in L_0) &= 0, \quad g(y \in L_1) = 1, \\
 g([x, y]) &= g(x) \oplus g(y) \cong 2, \\
 [x, y] &= -(-1)^{g(x) \cdot g(y)} [y, x], \\
 [x, [y, z]] &= [x, y], \quad z] + (-1)^{g(x)g(y)} [y, [x, z]].
 \end{aligned} \tag{1}$$

In some cases, there also exists a Z -grading $z(L_a) \in Z$, where z is a "quantum number" which is additively preserved by the super-Lie-bracket, though the

nature of that bracket itself is still determined by the $Z(2)_z$ -binary grading within Z , $Z \supset Z(2)_z$,

$$L = \sum_i L^i, \text{ and for any } x \in L^a, y \in L^b, [x, y] \in L^{a+b}. \quad (2)$$

Lie superalgebras can always be (and generally are) represented in matrix form, organized in *quarters* Q_g according to

$$\begin{array}{cc|cc} A_0 & A_1 & & \\ \hline B_1 & B_0 & v^0 & v^1 \end{array} \quad (3)$$

the $g = 0$ and $g = 1$ generators thus spanning the squares along or off the main diagonal, respectively. With this supermatrix acting on a column-vector V split in two by some $Z(2)_v$, the $g = 0$ quarters are the endomorphisms of V , namely $A_0 = \text{End}(V^0)$, $B_0 = \text{End}(V^1)$, whereas the $g = 1$ quarters represent the homomorphisms between the two sectors in V , namely $A_1 = \text{Hom}(V_1, V_0)$, $B_1 = \text{Hom}(V_0, V_1)$. In supersymmetry, the $Z(2)_v$ is again the quantum statistics characteristic and correlates with the statistics of the Hilbert space particles in a supersymmetry study.

In the case of a Grassmann (super-commutative) algebra of *differential forms*, the Z grading and its odd-even partitioning $Z(2)_f \subset Z$, respectively, represent the total count in the applications of the exterior derivative d (or the number of differentials involved as factors), and the odd/even partitioning to which this degree belongs. If applied to differential forms arising in an anholonomic basis or in a supergroup manifold, the $Z(2)$ grading fixes the *exterior* (wedge) product according to the rules,

$$\begin{aligned} dx^a \wedge dx^b &= -(-1)^{g(a)g(b)} dx^b \wedge dx^a, \\ F_f &= \sum_a dx^{a_1} \wedge dx^{a_2} \wedge \cdots dx^{a_f} \mathcal{F}_{a_1 \cdot a_2 \cdots a_f}(x), \\ F_{f_1}^a \wedge F_{f_2}^b &= (-1)^{(f_1 \cdot f_2 + g(a)g(b))} F_{f_2}^b \wedge F_{f_1}^a. \end{aligned} \quad (4)$$

A third category of $Z(2)$ gradings describes the intrinsic Poincaré or Lorentz group $Z(2)_s$ -grading of the variables of space-time and its double-covering (spin) and the corresponding exterior-derivative operator, as in the case of the Salam-Strathdee "superspace" of supersymmetry. There would then be a need to characterize algebraic structures by $Z(2)_s$, i.e. yet another $Z(2)$, whose eigenvalues we denote by $s(x)$. As we do not deal here with "classical" [Golfand-Likhtman/Wess-Zumino] supersymmetry, we shall not use the $s(x)$.

We now discuss a coupling between superalgebras, in particular the case in which the Lie superalgebra matrices are valued over Grassmann superalgebras

of differential forms. In these "directly coupled" superalgebras (we use the direct product symbol), the multiplication is fixed by the definition of the $Z(2)$ grading as the base (-1) logarithm of the elements of $Z(2)$, namely the square-roots of the identity, so that for

$$\begin{aligned} A &= A_0 + A_1, \quad F = F_0 + F_1, \\ (A \otimes F)_0 &= A_0 \otimes F_0 \oplus A_1 \otimes F_1, \\ (A \otimes F)_1 &= A_1 \otimes F_0 \oplus A_0 \otimes F_1, \end{aligned} \quad (5)$$

and the direct product for matrix-elements in two matrices

$$(a \otimes p)(a' \otimes p') = (-1)^{f(p)g(a')}(aa' \otimes pp'), \quad (6)$$

with the sign fixed by the $Z(2)$ eigenvalues of p and a' , once a' has to move through p to get to a . In any case, the overall grading h is given by

$$h(a \otimes p) = g(a)f(p) + f^n(y \in r_n) \cong 2. \quad (7)$$

The simple and semi-simple Lie superalgebras have been classified by V. Kac [12].

4 The Quillen Superconnection

After I had conceived $SU(2/1)$, Jean Thierry-Mieg and I investigated the possibility that the system of forms (the Grassmann superalgebra) in a Yang-Mills theory, *when extended by Higgs fields (i.e. in cases of spontaneous breaking of local symmetry) might generate a nonsupercommutative (or "nonsuper-abelian") superalgebra*. Such an extended Grassmann superalgebra might sometimes happen to coincide with a simple Lie superalgebra [$SU(2/1)$ in the electroweak case, etc.]. The idea was partly triggered by the composition of the Lagrangian in such models, *with a term in the Higgs potential quartic in the Higgs field*: such a term could be reproduced by a Lagrangian quadratic in the curvatures, provided these curvatures be taken for a supergroup, in which the even directions $g = 0$ in the superalgebra's $Z(2)_g$ grading are spanned by the original gauge group, while the Higgs fields span the g -odd directions. At that stage, there was no such rederivation for the remaining part of the Higgs potential, namely the *spontaneous symmetry-breakdown triggering* term, quadratic in the Higgs fields and similar to a mass term - but with the inverted sign. In identifying the Higgs fields themselves with even elements in the Grassmann algebra's form-calculus $Z(2)_f$, we were limited at this stage [13], as we had not dared go beyond Thierry-Mieg's original

identification of the ghosts as vertical components of the gauge fields, packed into contracted one-forms (in the fibre's direction) and the view in which the Higgs fields are *ghosts of ghosts*, i.e. two-forms, twice vertical. For the group-elements to be fully bosonic and Lorentz-invariant, the parameters would coincide for the even subgroup with its ordinary scalar parameters, while the odd part would have the Lorentz-scalar anticommuting ghosts (one-forms in our geometric interpretation of ghosts and BRST).

More about our $SU(2/1)$ example. The group is homomorphic with $Osp(2/2)$, whose fundamental representation is 4-dimensional and fits the quarks [14]. Moreover, one is allowed to add one constant real number to the diagonal quantum numbers; but if by adding one gets integer values for Y_w and for the electric charges - the matrix reduces to a 3-dimensional one. The group thus "knows" that quarks have fractional charges while leptons carry integer ones. Note also that since $I_w = su(2)$ and $I_w\psi_R = 0$, we have for the supertrace $sTr(I_w^z) = 0$; also, as the electrically-charged leptons or quarks are all massive and thus appear both on the right-chiral and left-chiral eigenstates, we also have $str(Q) = 0$.

Some time later and with a more daring mathematical motivation, D. Quillen [15] *postulated* his theory of the *superconnection*, in which the matrix-elements in the odd ($g(a) = 1$) and even ($g(a) = 0$) submatrices of a superalgebra are valued over the Grassmann supercommutative zero-forms ($f = 0$) and one-forms ($f = 1$), respectively, the intertwined coupling thereby ensuring that *the total grading be odd everywhere*, $t = g + f = 1$. It was shown that the 1979 electroweak $SU(2/1)$ could naturally be recast in this mold [16].

5 Noncommutative Geometry: The Electro-Weak Higher Algebraic Geometrization

The third and last step has consisted in reproducing *the entire Yang-Mills Lagrangian with spontaneous symmetry breakdown* directly from one single invariant; in other words, developing a further generalization which has allowed doing it by squaring one single "curvature", the corresponding generalized two-form. It was provided by a variant of A. Connes' *Noncommutative Geometry* [17]. These further developments have drawn on a generalization of the concept of *parallel transport*, as realized by the application of a (covariant) derivative, namely a derivative plus a connection.

At the same time, it also resolved a seemingly paradoxical feature of the

original "internal supersymmetry" interpretation, namely under the action of the g -odd generators of the superalgebra, the absence in the particle Hilbert space of *boson to fermion* transitions and vice versa. Instead, Hilbert space has carried some other $Z(2)$ grading, unrelated to quantum statistics - *chirality* in the electroweak case, - so that the endomorphisms induced by the odd generators produce a change of chirality, while the even endomorphisms preserve it.

R. Coquereaux and F. Scheck [18,19] were the first to show that this interesting result - namely the interrelationship between physically different $Z(2)$ groups - one in the vector space upon which the transformations are enacted and one in the superalgebra - could be treated as a development of noncommutative geometry (NCG). It was shown that the 1979 electroweak $SU(2/1)$ could naturally be recast in this mold [18-20].

The new arena is a fibre bundle with a non-simply connected base space, namely a direct product of a two-point space $Z(2)$; the points $(1, -1)$ in this realization of $Z(2)$ are labelled L & R $B = Z(2) \otimes M(1/3) = M(1,3)_L \oplus M(1,3)_R$ [21]. The gauge-group is the same over the entire basis, but the two fibres carry different representations: for the $SU(2/1)$ example (we refer the reader to Ref. [14] for the detailed algebraic features of this symmetry), the upper (left-chiral) part of V is an $I_{\text{left}} = 1/2 = \text{isodoublet}$ ($\nu_{e/L}^0 | e^L$) with weak hypercharges $Y_w = -1$, whereas for the right-chiral $I_{\text{left}} = 0$, $Y_w = -2$ (the assignments were fixed at the time by application of the Gell-Mann/Nishijima rule, as adapted to weak interactions).

Among several new features in noncommutative geometry, the most relevant one is its generalization to discrete spaces of concepts originally related to differentiable manifolds. One such concept, relevant to our present issue, is *parallel transport* [20]. To move from one point $\nu^L(x_L)$ on the fibre, over the point $x_L \in M_L$ in the left-chiral space, to another point $e_L^-(x'_L)$ on that fibre but over a different position x'_L in the same left-chiral space, we just use the partial derivative ∂_μ for an infinitesimal move by a length $\epsilon^\mu(x)$ and later integrate. To preserve the self-parallelism of the fibre at different places, we add a connection $A_\mu(x_L)$ and use a covariant derivative $D_\mu = \partial_\mu + A_\mu(x_L)$. The same is true for moves on the right-chiral space. In principle, this motion is not different from the usual symmetric case - except for the interface with the chiral matter fields.

What do we do, however, to go from $\nu^L(x \in M_L)$ to $e^R(y \in M_R)$? Besides using the above means, we also have to "jump" between the two chiral spaces - or better, directly between the relevant fibres over them. This

is a discrete move and it will be achieved by a finite matrix, in $SU(2/1)$ by μ_6 (same as λ_6 in $SU(3)$), which relates $e_L^- \rightarrow e_R'^-$. It should be anti-hermitian, so that we define the *matrix-derivative* as $T := i\mu_6$. At the same time, however, we need to perform a discrete change in the fibre itself, i.e. transform $(I_w = 1/2, I_w^z = -1/2, Y_w = -1) \rightarrow (I_w = I_w^z = 0, Y_w = -2)$, a task for which an appropriate connection is required. It has to resemble Δ as to its Lorentz properties - i.e. it is a scalar. We also note its quantum numbers in $SU(2/1)$: $I_w = 1/2, Y_w = -1$. This is the Higgs field $\Phi(x)$! Altogether, we shall have yet another new piece in the covariant derivative. In the $SU(2/1)$ internal supersymmetry we have a fibre-bundle with structure group $SU(2/1)$ over a split basis $(M_L \oplus M_R)$ and get the expression for the overall curvature

$$\mathcal{R} = d\omega + \omega^2 = dA + A^2 + \Phi^2 + d\Phi + A\Phi + T\Phi = R_{YM} + D\Phi + "V^{1/2}", \quad (8)$$

where we regroup the terms in their traditional setup,

$$R_{YM} = dA + A^2, \quad D\Phi = d\Phi + A\Phi, \quad V = [(\Phi)^2]^2 + [T\Phi]^2. \quad (9)$$

Squaring that total curvature with its Clebsch-Gordan coefficients and applying $T^2 = -1$ yields the conventional Weinberg-Salam Hamiltonian

$$\begin{aligned} H_{YM} &= {}^*R_{YM} \wedge R_{YM}, \\ H(\Phi)_{\text{kinetic}} &= D\Phi^2, \\ H(\phi)_{\text{potential}} &= V_\Phi = -\mu\Phi^2 + \lambda(\Phi)^4. \end{aligned} \quad (10)$$

6 Higher Algebraic Geometrization and Riemannian Geometry

One of the macroscopic features of this Universe is its obeying the Riemannian constraint, namely,

$$D_\rho g_{\mu\nu} = Q_{\rho\mu\nu} = 0. \quad (11)$$

Following Smolin [22], we have conjectured that this describes the state of affairs at low-energy, arising through the degradation of the basic (high-energy) microscopic state, which is then unconstrained and endowed with more symmetry. Assuming the original and quantum-era Universe to have been *affine* [23-25] we may be able to throw some light on the symmetry-breaking mechanism. We have conjectured [26] that this symmetry breakdown occurred through a mechanism of the same type studied in this article.

I have found that the Higher Algebraic Geometrization is provided here by the simple superalgebra $p(4, R)$, a "hyper-exceptional" in Kac's list. The algebra of the homogeneous symmetry group $SL(4, R)$ on the tetrad frames will sit in the even quarters, i.e. A_0 , B_0 in Eq. (3), of the 8×8 matrices of the defining representation of $P(4, R)$, along the diagonal. $SL(4, R)$ will be in its covariant representation in A_0 , in the contravariant in B_0 . A_1 will contain the 10 symmetric matrices (out of 16) in $GL(4, R)$ and B_1 will contain the 6 antisymmetric ones.

The matrix-derivative will be given by a *unit matrix* in A_1 (or by a Minkowski metric, depending on the issue) and break $SL(4, R)$ down to $SO(4)$ or $SO(3, 1)$, i.e. to Riemannian geometry. To justify the introduction of the matrix-derivative we have to start with a chirality-split base space - but this is precisely what we have when we take a Dirac spinor $(1/2, 0) \oplus (0, 1/2)$ or a *world spinor* [27-29] with this lowest state. We may now write the full "extended curvature" of $P(4, R)$ - including the matrix-derivative piece. It includes the "SKY" [30-32] quadratic $SL(4, R)$ Lagrangian, the kinetic and gauge terms $D\Phi^+$ and $D\Phi^-$, respectively, for the two Higgs holonomic scalars (one a symmetric tensor in the frame indices, one an antisymmetric), a matrix-derivative generated $T\Phi^-$ which will trigger the spontaneous symmetry breakdown, and a term quadratic in the Higgs fields $\{\Phi^+\Phi^-\}$. I have described the physical effects in detail in Ref. [26], with results fitting the observed low-energy Riemannian system.

References

- [1] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [2] H. Kleinert, *Gen. Rel. Grav.* **32**, 769 (2000).
- [3] Y. Ne'eman, *Func. Diff. Eqs.* **5**, 19 (1998); talk delivered at the Intern. Conf. on Diff. Eqs., Ariel, Israel, 1998.
- [4] J. Thierry-Mieg, *J. Math. Phys.* **21**, 2834 (1980).
- [5] J. Thierry-Mieg, *Il Nuovo Cim. A* **56**, 396 (1980).
- [6] Y. Ne'eman, *Phys. Lett. B* **81**, 190 (1979).
- [7] D.B. Fairlie, *Phys. Lett. B* **82**, 97 (1979).
- [8] Y. Ne'eman, *Phys. Lett. B* **181**, 308 (1986).
- [9] D.S. Hwang, C.-Y. Lee, and Y. Ne'eman, *Int. J. Mod. Phys. A* **11**, 3509 (1996).

- [10] L. Corwin, Y. Ne'eman, and S. Sternberg, *Rev. Mod. Phys.* **47**, 573 (1975).
- [11] Y. Ne'eman and T. Regge, *Rivista Del Nuovo Cim.* **1**, 1 (1978); first issued as IAS Princeton and U. Texas ORO 3992 328 preprints.
- [12] V.G. Kac, *Func. Analysis and Appl.* **9**, 91 (1975); also *Comm. Math. Phys.* **53**, 31 (1977); see also V. Rittenberg, in *Group Theoretical Methods in Physics* (Proc. Tübingen, Germany, 1977), Eds. P. Kramer and A. Rieckers, Lecture Notes in Physics **79** (Springer, Berlin, 1977), p. 3.
- [13] J. Thierry-Mieg and Y. Ne'eman, *Proc. Nat. Acad. Sci. USA* **79**, 7068 (1982).
- [14] J. Thierry-Mieg and Y. Ne'eman, *Methods in Mathematical Physics*, (Proc. Aix en Provence and Salamanca, 1979), Eds. P.L. Garcia, A. Perez-Rendon, and J.M. Souriau, *Springer Lecture Notes in Mathematics* **836** (Springer, Berlin, 1980), p. 318.
- [15] D. Quillen, *Topology* **24**, 89 (1985).
- [16] A. Connes, in *The Interface of Mathematics and Particle Physics*, Eds. D. Quillen, G. Segal, and S. Tsou (Oxford University Press, Oxford, 1990).
- [17] Y. Ne'eman and S. Sternberg, *Proc. Nat. Acad. Sci. USA* **87**, 7875 (1990).
- [18] R. Coquereaux, R. Häussling, N.A. Papadopoulos, and F. Scheck, *Int. J. Mod. Phys. A* **7**, 2809 (1992).
- [19] R. Coquereaux, G. Esposito-Farese, and F. Scheck, *Int. J. Mod. Phys. A* **7**, 6555 (1992).
- [20] Y. Ne'eman, D.S. Hwang, and C.-Y. Lee, in *Group 21, Physical Applications and Mathematical Aspects of Geometry, Groups, and Algebras* **2** (Proc. XXI Inter. Coll. on Group Theoretical Methods in Physics Group 21), Eds. H.D. Doebner, W. Scherer, and C. Schulte (World Scientific, Singapore, 1997), p. 553.
- [21] A. Connes and J. Lott, *Nucl. Phys. B (Proc. Suppl.)* **18**, 29 (1990).
- [22] L. Smolin, *Nucl. Phys. B* **247**, 511 (1984).
- [23] Y. Ne'eman and D. Šijački, *Phys. Lett. B* **200**, 489 (1988).
- [24] C.-Y. Lee and Y. Ne'eman, *Phys. Lett. B* **242**, 59 (1990).
- [25] C.-Y. Lee, *Class. Quantum Grav.* **9**, 2001 (1992).
- [26] Y. Ne'eman, *Phys. Lett. B* **427**, 19 (1998).
- [27] Y. Ne'eman, *Proc. Nat. Acad. Sci. USA* **74**, 4157 (1977).
- [28] Y. Ne'eman, *Ann. Inst. H. Poincaré A* **28**, 369 (1978).

- [29] F.W. Hehl, J.D. McCrea, E.W. Mielke, and Y. Ne'eman, *Phys. Rep.* **258**, 1 (1995).
- [30] G. Stephenson, *Nuovo Cim.* **9**, 263 (1958).
- [31] C.W. Kilmister and D.J. Newman, *Proc. Cam. Phil. Soc.* **57**, 851 (1961).
- [32] C.N. Yang, *Phys. Rev. Lett.* **33**, 445 (1974).

DYNAMICAL FERMION MASSES UNDER THE INFLUENCE OF KALUZA-KLEIN FERMIONS IN RANDALL-SUNDRUM BACKGROUND

H. ABE

*Department of Physics, Hiroshima University, Higashi-Hiroshima,
Hiroshima 739-8526, Japan*

E-mail: abe@theo.phys.sci.hiroshima-u.ac.jp

T. INAGAKI

*Research Institute for Information Science and Education, Hiroshima University,
Higashi-Hiroshima,*

Hiroshima 739-8521, Japan

E-mail: inagaki@hiroshima-u.ac.jp

T. MUTA

*Department of Physics, Hiroshima University, Higashi-Hiroshima, Hiroshima
739-8526, Japan*

E-mail: muta@hiroshima-u.ac.jp

The dynamical fermion mass generation on the D3-brane in the Randall-Sundrum space-time is discussed in a model with bulk fermions in interaction with fermions on the branes. It is found that the dynamical fermion masses are generated at the natural (R.-S.) radius of the compactified extra-dimensional space and may be made small compared with masses of the Kaluza-Klein modes which are of order TeV.

1 Introduction

In the 1920's an interesting idea came up assuming the existence of an extra-dimensional space which eventually compactifies leaving our 4-dimensional space-time as a real world [1]. A few years ago, a proposal [2] for the mass

scale of the compactified space to be much smaller than the Planck scale gave a strong impact on the onset of studying phenomenological evidences of extra-dimensional effects [3]. There is a crucial problem, however, how such large extra dimensions are stabilized. Recently, Randall and Sundrum [4] (R.-S.) gave an alternative to the large extra dimension scenario. By introducing a specific curved bulk space they succeeded to get a mass scale much smaller than the Planck scale without relying on the fine-tuning.

In our present analysis we introduce bulk fermions in the R.-S. space-time [5] and see what effects could be observed. The bulk fermions interact with themselves as well as with fermions on the 4-dimensional branes through the exchange of the graviton and its Kaluza-Klein excited modes, or through the exchange of gauge bosons which may be assumed to exist in the bulk [5]. The interactions among fermions generated as a result of the exchange of all the Kaluza-Klein excited modes of the graviton or gauge bosons may be expressed as effective four-fermion interactions [3,6,7]. According to the four-fermion interactions we expect that the dynamical generation of fermion masses will take place. In the present communication we look for a possibility of the dynamical fermion mass generation under the influence of the bulk fermions through the effective four-fermion interactions in the R.-S. background. In the R.-S. background, fermion mass terms are forbidden by the $S^1/\mathbf{Z}_2$ symmetry. The possible source of fermion masses on the branes is two-fold, i.e. the dynamically generated fermion masses and masses of the Kaluza-Klein excited modes of the bulk fermions. The mass of the Kaluza-Klein excited modes is known to be of order of TeV [5].

We will review the dynamical fermion mass generation with bulk fermions in the *torus* compactified case [8] because most of the technical parts can be summarized in the torus (flat) extra dimension case. After that we proceed to the case of the R.-S. space-time in Section 3, the main part of this paper: It is found that the dynamical fermion masses are generated at the natural (R.-S.) radius of the compactified extra-dimensional space and may be made small compared with masses of the Kaluza-Klein modes which are of order of TeV because of the presence of the Randall-Sundrum warp factor.

2 Flat Extra Space

We assume an existence of 5-dimensional bulk fermions ψ in interaction with fermions L on the 4-dimensional brane. Effective interactions among these fermions can be given in the form of the four-fermion interaction. In fact it

is known that the exchange of the Kaluza-Klein excited modes of the bulk graviton results in effective four-fermion interactions [3]. After the Fierz transformation on the four-fermion interactions we generate the transition-type interactions. Accordingly we start with the following Lagrangian for our model [8]

$$\mathcal{L}^{(5)} = \bar{\psi} i \gamma^M \partial_M \psi + [\bar{L} i \gamma^\mu \partial_\mu L + g^2 \bar{\psi} \gamma^M L \bar{L} \gamma_M \psi] \delta(x^4), \quad (1)$$

where g is the coupling constant with mass dimension $-3/2$ and the index M runs from 0 to 4 while the index μ runs from 0 to 3. Fermions ψ and L are assumed to consist of N_f components.

After neglecting tensor interactions and introducing an auxiliary field $\sigma \sim \bar{\psi} L$, we compactify the bulk space on torus with radius R . Our 4-dimensional Lagrangian is rewritten as

$$\mathcal{L}^{(4)} = \bar{\Psi} (M + i \not{\partial}) \Psi - |\sigma|^2, \quad (2)$$

where $\Psi^t \equiv (L, \psi_0, \psi_1, \psi_{-1}, \dots)$ and

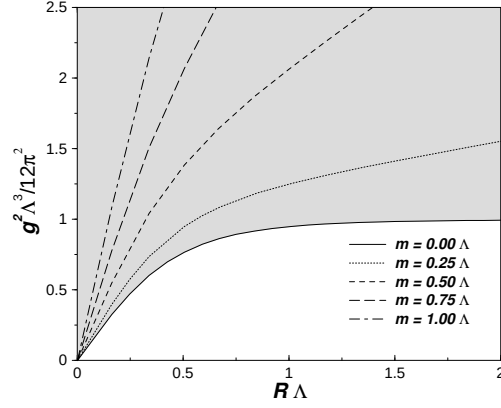
$$M \equiv \begin{pmatrix} 0 & m^* & m^* & m^* & \cdots \\ m & 0 & 0 & 0 & \cdots \\ m & 0 & \frac{1}{R} & 0 & \cdots \\ m & 0 & 0 & -\frac{1}{R} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad m = \frac{g\sigma}{\sqrt{2\pi R}}. \quad (3)$$

If σ acquires a non-vanishing vacuum expectation value, we replace σ in m by its vacuum expectation value $\langle \sigma \rangle$, i.e. $m = Ng \langle \sigma \rangle$. The eigenvalues of the matrix M determine the masses of 4-dimensional fermions. Obviously we find that the lightest eigenvalue is given by

$$\lambda_{\pm 0} = \pm |m| \quad \text{for} \quad |m| \ll 1/R. \quad (4)$$

Thus we conclude that there is a possibility of having the light fermion masses within our scheme being much smaller than the mass of the Kaluza-Klein modes of the bulk fermion. By performing the path integration for the fermion field Ψ , we find the effective potential for σ in the leading order of the $1/N_f$ expansion:

$$V(\sigma) = |\sigma|^2 - \frac{1}{2\pi^2} \int_0^\Lambda dx x^3 \ln [x^2 + |m|^2 (\pi x R) \coth(\pi x R)]$$

Figure 1. Critical radius as a function of g .

$$-\frac{1}{2\pi^2} \sum_{j=1}^{\infty} \int_0^{\Lambda} dx x^3 \ln \left[x^2 + \left(\frac{j}{R} \right)^2 \right]. \quad (5)$$

The gap equation to determine the vacuum expectation value $\langle |\sigma| \rangle$ of $|\sigma|$ reads

$$\frac{\partial V(\sigma)}{\partial |\sigma|} = 2|\sigma| \left\{ 1 - \frac{g^2}{2\pi^2} \int_0^{\Lambda} dx \frac{x^3}{2x \tanh(\pi x R) + g^2 |\sigma|^2} \right\} = 0. \quad (6)$$

By numerical integration of Eq. (6) we find that there exists a non-trivial solution for $|\sigma|$ for a suitable range of parameters g and R , and that the solution corresponds to the true minimum of the effective potential. Accordingly, the fermion mass is generated dynamically. Here the auxiliary field σ (or the composite field $\bar{L}\psi$) acquires a vacuum expectation value. Moreover it is easily confirmed that the phase transition associated with this symmetry breaking is of second order.

As shown in Eq. (4) the lowest fermion mass on the 4-dimensional brane is $m = Ng\langle |\sigma| \rangle$ where $\langle |\sigma| \rangle$ is determined by solving Eq. (6). The critical curve for $m = 0$ which represents the critical radius as a function of the coupling constant is shown in Fig. 1. If we assume $1/R \sim \text{TeV}$, light ($< \text{TeV}$) fermions are obtained in the region between the solid line and the dot-dashed line in Fig. 1 because it is natural that $R\Lambda \sim 1$. However, we have no idea how to justify $1/R \sim \text{TeV}$ and so we need to introduce a certain mechanism such as the Randall-Sundrum model.

3 Warped Extra Space

People wonder how, in theories with large ($\gg M_{\text{pl}}^{-1}$) extra dimensions, such large radii are stabilized. We have not yet come across with any satisfactory answer, while important results in many models depend essentially on the largeness of the radius (see the previous section). Randall and Sundrum [4] noticed that the existence of the brane leads to the curved bulk space and proposed the so-called Randall-Sundrum (R.-S.) model. Their model has *two* D3-branes on the $S_1/\mathbf{Z}_2$ orbifold fixed points and the ADS_5 between these branes:

$$ds^2 \equiv G_{MN} dx^M dx^N = e^{-2kb_0|y|} \eta_{\mu\nu} dx^\mu dx^\nu + b_0^2 dy^2, \quad (7)$$

where $k \sim M_{\text{pl}}$ is the gravity scale and $b_0^{-1} \sim 10^{-1}k - 10^{-2}k$ is the compactification scale. One of the most important results is the warp factor $e^{-kb_0/2}$ which is the suppression of the K.-K. masses of the bulk fields [5].

3.1 Bulk Fermions in R.-S. Background

Following Chang *et al.* [5] we derive the mode expansion of the bulk fermion in R.-S. space-time such that

$$\psi(x, y) = \frac{e^{\frac{3}{2}kb_0|y|}}{\sqrt{b_0}} \sum_n \left[\psi_L^{(n)}(x) \xi(y) + \psi_R^{(n)}(x) \eta(y) \right], \quad (8)$$

$$\begin{cases} \xi(y) = \sqrt{\frac{kb_0}{1-e^{-\frac{1}{2}kb_0}}} e^{-\frac{1}{2}kb_0|\frac{1}{2}-y|} \sin \frac{m_n}{kb_0} (1 - e^{kb_0|y|}) \\ \eta(y) = \sqrt{\frac{kb_0}{1-e^{-\frac{1}{2}kb_0}}} e^{-\frac{1}{2}kb_0|\frac{1}{2}-y|} \cos \frac{m_n}{kb_0} (1 - e^{kb_0|y|}) \end{cases}, \quad (9)$$

where $m_n \equiv n\pi kb_0/(e^{\frac{1}{2}kb_0} - 1)$ is b_0 times the mass of the n -th K.-K. mode. With this expansion the kinetic terms of the K.-K. modes become

$$\mathcal{L}_{\text{bulk}}^{(4)} = \int dy E \bar{\psi} i \gamma^M D_M \psi = \sum_n \left[\bar{\psi}^{(n)} i \not{\partial} \psi^{(n)} - \frac{m_n}{b_0} \bar{\psi}^{(n)} \psi^{(n)} \right]. \quad (10)$$

We can now apply previous results of our model with bulk fermions in the R.-S. background.

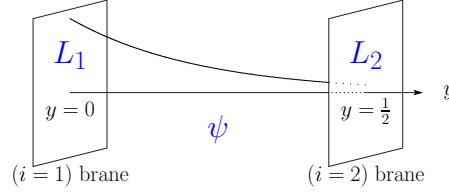


Figure 2. Four-fermion interaction model in R.-S. background.

3.2 Four-Fermion Interaction Model in R.-S. Background

When applying the previous torus-compactified model (1) to the case in the R.-S. background, we get

$$E_{\bar{M}}^M = \begin{pmatrix} e^{kb_0|y|}\eta_{\bar{\mu}}^{\mu} & 0 \\ 0 & b_0^{-1} \end{pmatrix}, \quad (11)$$

where $E_{\bar{M}}^M$ is the vielbein whose square becomes the metric G_{MN} . We introduce the bulk fermion ψ which propagates in the whole five-dimensional space as in the torus case, while *two* brane fermions L_i ($i = 1, 2$) propagate on an i -th “brane”.

We start with the Lagrangian in five dimensions,

$$\begin{aligned} \mathcal{L}^{(5)} = & E\bar{\psi}i\gamma^M D_M\psi + E_{(1)} \left[\bar{L}_1 i\gamma^\mu \partial_\mu L_1 + \frac{g_1^2}{N_f} \bar{\psi}\gamma^M L_1 \bar{L}_1 \gamma_M \psi \right] \delta(x^4) \\ & + E_{(2)} \left[\bar{L}_2 i\gamma^\mu \partial_\mu L_2 + \frac{g_2^2}{N_f} \bar{\psi}\gamma^M L_2 \bar{L}_2 \gamma_M \psi \right] \delta(x^4 - \tfrac{1}{2}), \end{aligned} \quad (12)$$

where $x^M = (x^\mu, x^4 = y)$, $\mu = (0, 1, 2, 3)$, g_i 's are the four-fermion couplings ($[g_i] = [-3/2]$), and N_f is the number of components of ψ and the L_i 's. Note that the bulk mass term is removed by the $S^1/\mathbf{Z}_2$ projection. Because Lorentz covariance should be preserved, we neglect the tensor (vector) interactions and obtain

$$\begin{aligned} \mathcal{L}^{(5)} = & E\bar{\psi}i\gamma^M D_M\psi + \left[\bar{L}_1 i\gamma^{\bar{\mu}} \partial_{\bar{\mu}} L_1 + \frac{g_1^2}{N_f} \bar{\psi}\gamma_5 L_1 \bar{L}_1 \gamma_5 \psi \right] \delta(x^4) \\ & + e^{-2kb_0} \left[e^{\frac{1}{2}kb_0} \bar{L}_2 i\gamma^{\bar{\mu}} \partial_{\bar{\mu}} L_2 + \frac{g_2^2}{N_f} \bar{\psi}\gamma_5 L_2 \bar{L}_2 \gamma_5 \psi \right] \delta(x^4 - \tfrac{1}{2}). \end{aligned} \quad (13)$$

After chiral rotation $\psi \rightarrow e^{i\pi\gamma^5/4}\psi$ and $L_i \rightarrow e^{i\pi\gamma^5/4}L_i$, we introduce auxiliary

fields σ_i to obtain

$$\begin{aligned} \mathcal{L}^{(5)} = & E\bar{\psi}i\gamma^M\partial_M\psi + \left[\bar{L}_1i\cancel{\partial}L_1 - N_f|\sigma_1|^2 + (g_1\sigma_1\bar{\psi}L_1 + h.c.)\right]\delta(x^4) \\ & + e^{-2kb_0}\left[e^{\frac{1}{2}kb_0}\bar{L}_2i\cancel{\partial}L_2 - N_f|\sigma_2|^2 + (g_2\sigma_2\bar{\psi}L_2 + h.c.)\right]\delta(x^4 - \frac{1}{2}). \end{aligned} \quad (14)$$

Now we substitute the mode expansion of ψ of Eq. (8) into the Lagrangian and integrate it over the extra-dimension. Taking $e^{-3kb_0/4}L_2 \rightarrow L_2$ and $e^{-kb_0}\sigma_2 \rightarrow \sigma_2$, the Lagrangian in four dimensions becomes

$$\begin{aligned} \mathcal{L}^{(4)} \equiv & \int dy \mathcal{L}^{(5)} \\ = & \sum_n \left[\bar{\psi}^{(n)}i\cancel{\partial}\psi^{(n)} - \frac{n}{R}\bar{\psi}^{(n)}\psi^{(n)} \right] + \bar{L}_1i\cancel{\partial}L_1 + \bar{L}_2i\cancel{\partial}L_2 - N_f(|\sigma_1|^2 + |\sigma_2|^2) \\ & + \sum_n \left(\mu_1\bar{\psi}_R^{(n)}L_1 + h.c. \right) + \sum_n \left((-1)^n\mu_2\bar{\psi}_R^{(n)}L_2 + h.c. \right), \end{aligned} \quad (15)$$

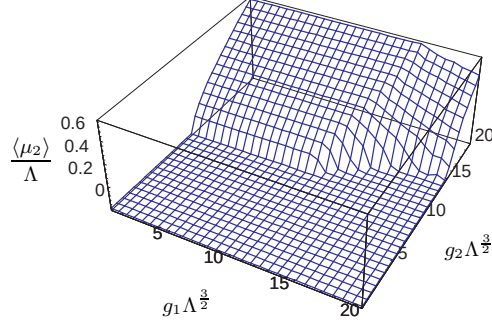
where

$$\mu_1 \equiv \frac{g_1\sigma_1}{\sqrt{b_0}}, \quad \mu_2 \equiv \frac{g_2\sigma_2}{\sqrt{b_0}}, \quad R \equiv \frac{b_0}{m_1} = \frac{e^{\frac{1}{2}kb_0} - 1}{\pi k} \sim (\text{TeV})^{-1}. \quad (16)$$

Here we note that the mass of the first K.-K. excited mode $1/R$ is of the order of TeV without fine-tuning because of the warp factor in the R.-S. metric. This is the only dependence on the factor.

After integrating out all fermionic degrees of freedom, we obtain the effective potential for $\sigma_i(\mu_i)$ as follows:

$$\begin{aligned} \bar{V}(\mu_1/\Lambda, \mu_2/\Lambda) &= [V(\mu_1/\Lambda, \mu_2/\Lambda) - V(0, 0)]/\Lambda^4 \\ &= 2\pi R\Lambda \left[\frac{(\mu_1/\Lambda)^2}{g_1^2\Lambda^3} + \frac{(\mu_2/\Lambda)^2}{g_2^2\Lambda^3} \right] - \frac{1}{2\pi^2} \int_0^1 dz z^3 \ln \left\{ 1 + \right. \\ &\quad \left. \frac{\pi R\Lambda}{z} \left[\left(\frac{\mu_1}{\Lambda} \right)^2 + \left(\frac{\mu_2}{\Lambda} \right)^2 \right] \cot(\pi z R\Lambda) + \frac{\pi^2}{2} (R\Lambda)^2 \left(\frac{\mu_1}{\Lambda} \right)^2 \left(\frac{\mu_2}{\Lambda} \right)^2 \right\}. \end{aligned} \quad (17)$$

Figure 3. $\langle \mu_2 \rangle$ as a function of g_1, g_2 .

3.3 Behavior of the Vacuum $\langle \mu_i \rangle$

The behavior of the vacuum is determined by solving the gap equations ($i, j = 1, 2; i \neq j$):

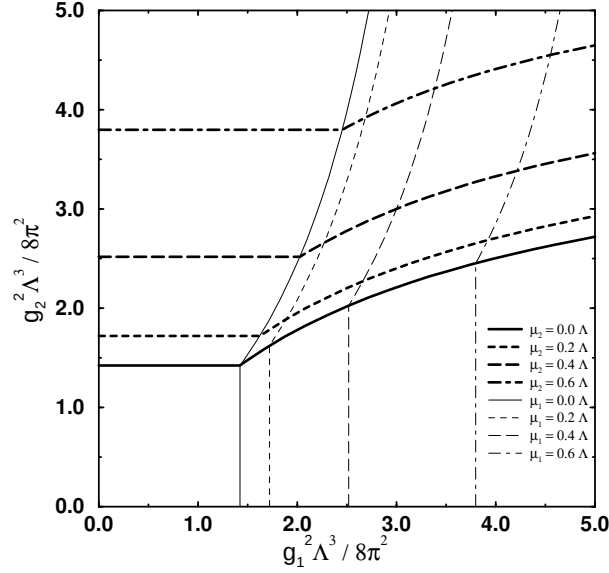
$$\begin{aligned} \frac{\partial \bar{V}}{\partial (\mu_i / \Lambda)} &= R\Lambda \frac{\mu_i}{\Lambda} \left\{ \frac{4\pi}{g_i^2 \Lambda^3} - \frac{1}{2\pi} \int_0^1 dz z^3 \right. \\ &\quad \times \frac{\pi R\Lambda \mu_j^2 \tanh(\pi z R\Lambda) + 2}{z \left[1 + \frac{\pi^2}{2} (R\Lambda)^2 \left(\frac{\mu_1}{\Lambda} \right) \left(\frac{\mu_2}{\Lambda} \right) \tanh(\pi z R\Lambda) + \pi R\Lambda \left[\left(\frac{\mu_1}{\Lambda} \right) + \left(\frac{\mu_2}{\Lambda} \right) \right]} \right\} \\ &= 0. \end{aligned} \quad (18)$$

The system has a second-order phase transition. As shown in Fig. 3, $\langle \mu_2 \rangle$ is a function of g_1 and g_2 for $R\Lambda = 1$. If $g_i \gtrsim g_{i,\text{critical}}$, we see that $\langle \mu_i \rangle \gtrsim 0$. The phase structure of this system is summarized in Fig. 4.

3.4 Effective Theory on $y = 1/2$ Brane

Now we realize that μ_1 and μ_2 can have a non-zero vacuum expectation value and that the Lagrangian (15) has the mixing term $\bar{\psi}_R L_i + \text{h.c.}$. The physics on the $i = 2$ brane is examined by integrating out the invisible field L_1 . Setting $\sqrt{2}\psi_L^{(n)} \equiv N^{(n)} - M^{(n)}$ and $\sqrt{2}\psi_R^{(n)} \equiv N^{(n)} + M^{(n)}$ for $n \neq 0$, the effective Lagrangian on the $y = 1/2$ brane is obtained as follows:

$$\mathcal{L}_{\text{eff}}^{(4)} = \bar{\Theta}_2 [i\bar{\psi} + M_2 + |\mu_1|^2 P] \Theta_2, \quad (19)$$

Figure 4. The phase structure of the vacuum $\langle \mu_2 \rangle$.

where $\Theta_2^T \equiv (L_2, \psi_R^{(0)}, N^{(1)}, M^{(1)}, \dots)$ and

$$M_2 \equiv \begin{pmatrix} 0 & \mu_2^* & -\mu_2^* & -\mu_2^* & \cdots \\ \mu_2 & 0 & 0 & 0 & \cdots \\ -\mu_2 & 0 & \frac{1}{R} & 0 & \cdots \\ -\mu_2 & 0 & 0 & -\frac{1}{R} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad P \equiv \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & (i\partial)^{-1} & (i\partial)^{-1} & (i\partial)^{-1} & \cdots \\ 0 & (i\partial)^{-1} & (i\partial)^{-1} & (i\partial)^{-1} & \cdots \\ 0 & (i\partial)^{-1} & (i\partial)^{-1} & (i\partial)^{-1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (20)$$

We have seen in the previous subsection that the dynamical fermion mass generation with ψ is a second-order phase transition, and we find the parameter region of (g_1, g_2) for $\langle \mu_2 \rangle \lesssim \Lambda$, $1/R \sim \text{TeV}$, $\langle \mu_1 \rangle = 0$. Thus in Eq. (19) we have a possibility to get a light Dirac fermion on the $i = 2$ brane.

4 Conclusion

In our model the dynamically generated fermion mass [9] is much smaller than the Planck scale because of the presence of the mass scale $1/R \sim \text{TeV}$ generated by the Randall-Sundrum warp factor. Furthermore, in spite of the presence of the mass scale $1/R \sim \text{TeV}$ in the theory, the fermion masses on the four-dimensional brane can be made smaller than this scale as a consequence of the interaction among the bulk and brane fermions: the mixing of the brane fermions with the bulk fermions does not lead to the lightest fermion masses of order $1/R$ and the dynamically generated fermion masses are not of order $1/R$. This result is obtained because the dynamical fermion masses generated in the second-order phase transition are small irrespective of $1/R$ near the critical radius. In our model the possibility of having low mass fermions resulted from the dynamical origin. This mechanism is quite different from the ones in other approaches in which low mass fermions are expected to show up as a result of the kinematical origins [10-14].

As an outlook we want to point out four items. The first item is whether the K.-K. gauge boson (or graviton) exchange induces suitable effective four-fermion interactions. The second is to understand the physics with $\langle \mu_1 \rangle \neq 0$ and the third is how to stabilize the radius $b_0 = 10 \sim 100k^{-1}$ with the pressure of the bulk fermion ψ [15-17]. The last is the extension of our model to electroweak symmetry breaking [6,7,18,19].

Acknowledgments

This paper is dedicated to the 60th birthday of Professor Hagen Kleinert who is a good old friend of one of the authors (T.M.). The authors would like to thank Hironori Miguchi, Koichi Yoshioka (Kyoto U.), Masahiro Yamaguchi (Tohoku U.), and Hiroaki Nakano (Niigata U.) for fruitful discussions and correspondence. They are also indebted to Tak Morozumi for useful comments. The present work was financially supported by the Monbusho Grant, Grant-in-Aid for Scientific Research (C) with contract number 11640280.

References

- [1] T. Kaluza, *Sitzungsber. d. Preuss. Akad. d. Wiss.*, 966 (1921); O. Klein, *Z. Phys.* **37**, 895 (1926).

- [2] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *Phys. Lett. B* **429**, 263 (1998); *Phys. Rev. D* **59**, 086004 (1999).
- [3] T. Han, J.D. Lykken, and R.-J. Zhang, *Phys. Rev. D* **59**, 105006 (1999).
- [4] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 3370 (1999).
- [5] S. Chang, J. Hisano, H. Nakano, N. Okada, and M. Yamaguchi, *Phys. Rev. D* **62**, 084025 (2000).
- [6] B.A. Dobrescu, *Phys. Lett. B* **461**, 99 (1999).
- [7] H.C. Cheng, B.A. Dobrescu, and C.T. Hill, eprint: hep-ph/9912343.
- [8] H. Abe, H. Miguchi, and T. Muta, *Mod. Phys. Lett. A* **15**, 445 (2000).
- [9] For a related critical argument, see e.g. H. Kleinert and B. Van den Bossche, *Phys. Lett. B* **474**, 336 (2000).
- [10] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, and J. March-Russel, eprint: hep-ph/9811448.
- [11] K.R. Dienes, E. Dudas, and T. Gherghetta, *Nucl. Phys. B* **557**, 25 (1999).
- [12] R.N. Mohapatra, S. Nandi, and A. Perez-Lorenzana, *Phys. Lett. B* **466**, 115 (1999).
- [13] A. Das and C.W. Kong, *Phys. Lett. B* **470**, 149 (1999).
- [14] K. Yoshioka, *Mod. Phys. Lett. A* **15**, 29 (2000).
- [15] S.A. Gundersen and F. Ravndal, *Ann. Phys. (N.Y.)* **182**, 90 (1988).
- [16] H. Abe, J. Hashida, T. Muta, and A. Purwanto, *Mod. Phys. Lett. A* **14**, 1033 (1999).
- [17] W.D. Goldberger and M.B. Wise, *Phys. Rev. Lett.* **83**, 4922 (1999).
- [18] N. Arkani-Hamed, H.-C. Cheng, B.A. Dobrescu, and L.J. Hall, *Phys. Rev. D* **62**, 096006 (2000).
- [19] M. Hashimoto, M. Tanabashi, and K. Yamawaki, eprint: hep-ph/0010260.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

THE ULTRAVIOLET FIXED-POINT IN QUANTUM ELECTRODYNAMICS – ADLER CONJECTURE: IS QED TRIVIAL?

R. ACHARYA

*Department of Physics & Astronomy,
Arizona State University Tempe, Arizona 85287, USA
E-mail: raghunath.acharya@asu.edu*

We review the question of triviality of QED due to Landau and recall the arguments of the Adler conjecture on the vanishing of the Callan-Symanzik function at the physical fine structure constant, $\beta(\alpha) = 0$.

As a result of the screening of charged particles by their interactions with virtual fermion-antifermion pairs in the vacuum state it is conceivable that there exists no interacting continuum limit of QED in 4 dimensions, a property called “triviality of the theory” by mathematical physicists. From the Renormalization Group (RG) point of view, triviality is a reflection of the *absence* of a nontrivial ultra-violet (UV) stable fixed-point in the Callan-Symanzik beta function. It has been emphasized by Wilson [1] that a nontrivial renormalizable theory can only be formulated, if the Callan-Symanzik function exhibits UV stable fixed-points at a nonzero coupling strength: Only such a fixed-point allows for a *nontrivial continuum limit* of the theory. This was demonstrated in the classic 1954 paper of Gell-Mann and Low [2]. Landau and Pomeranchuk [3] argued that QED is expected to be a *free* theory in this limit. Landau’s contention was recently substantiated by lattice calculations [4]. We are now pretty confident that, at infinite cutoff, perturbative QED suffers from complete screening and would have a vanishing fine structure constant. This is somewhat ironic since perturbative renormalization scores its greatest “triumph” in QED!

Nevertheless it is still conceivable that there does exist an UV stable fixed-point, after all, as first argued in the premise in the classic papers of

Adler [5], and Johnson and Baker [6], although subsequent investigation by Adler, Callan, Gross, and Jackiw and by Baker and Johnson [7] inclined again towards the *absence* of a UV fixed-point. This result was reemphasized by the analysis of Acharya and Narayana Swamy [7].

An UV fixed-point would only be possible, if non-perturbative effects changed the qualitative nature of the operator product expansion, or if there were a non-perturbative renormalization of the triangle anomaly [8]. In this note, we present once more an argument that the Callan-Symanzik beta function $\beta(\alpha)$ does have a nontrivial zero in QED [9], implying a nontrivial continuum limit.

Since the vacuum expectation value of the vector current $\mathcal{J}^\mu$ vanishes by Lorentz invariance, the charge Q must annihilate the vacuum state $|0\rangle$:

$$Q|0\rangle = 0. \quad (1)$$

The conservation of the vector current j^μ

$$\partial^\mu j_\mu(\mathbf{x}, t) = 0 \quad (2)$$

implies the *local* commutator

$$[Q, H(\mathbf{x}, t)] = 0. \quad (3)$$

The only assumption is that the surface terms at spatial infinity can be discarded. This is justified as long as there are *no* scalar Nambu-Goldstone bosons which could produce a long-range interaction necessary for a non-vanishing surface term at spatial infinity. In QED this is supposed to be the case.

Massless QED is scale-invariant at the classical level. After quantization, the divergence of the dilatation current D_μ is determined by the trace anomaly

$$\partial^\mu D_\mu = \frac{\beta(\alpha)}{\alpha} F_{\mu\nu} F^{\mu\nu}. \quad (4)$$

The dilatation charge

$$Q_D = \int d^3x D_0(\mathbf{x}, t) \quad (5)$$

satisfies the commutator relation

$$[Q_D, Q] = -id_Q Q \quad (6)$$

which defines the scale dimension of the charge Q , whose canonical value is $d_Q = 0$. We derive from Eqs. (3) and (6), invoking the Jacobi identity, the double commutator

$$[Q_D, [Q, H]] = 0, \quad (7)$$

and arrive at

$$[Q, \partial^\mu D_\mu] = 0. \quad (8)$$

The last step follows from

$$[H(\mathbf{x}, t), Q_D] = -i\partial_\mu D^\mu \neq 0 \quad (9)$$

by virtue of the trace anomaly (4).

Applying (8) to the vacuum state, we obtain

$$[Q, \partial^\mu D_\mu] |0\rangle = 0. \quad (10)$$

Using Eq. (1), this implies

$$Q\partial^\mu D_\mu |0\rangle = 0. \quad (11)$$

The divergence of the dilatation current is clearly a *local* operator. Moreover, the charge Q is *time-independent* (since the vector current is conserved) and has the following important properties:

- (a) It is a *constant* operator,
- (b) it is also a generator of pure phase rotations in the electron fields.

Therefore, the locality of $\partial^\mu D_\mu$ remains undisturbed by the multiplication with Q . We may thus invoke the Federbush-Johnson theorem [10] which applies to *any* local operator to conclude that

$$Q\partial^\mu D_\mu \equiv 0. \quad (12)$$

Since the charge Q must clearly be *non-vanishing* in QED, we obtain

$$\partial^\mu D_\mu \equiv 0. \quad (13)$$

Together with Eq. (4) this implies the result we wanted to prove:

$$\beta(\alpha) = 0. \quad (14)$$

This is the conclusion first drawn in the classic paper of Adler [5]. It remains an interesting problem to calculate $\beta(\alpha)$ non-perturbatively and verify

whether that (14) is true or not, and to find out the possible source of error in the previous simple line of arguments.

Acknowledgments

This brief essay is offered to celebrate the 60th birthday of Professor Hagen Kleinert (still youthful in appearance and “just smashing”). I have fond memories of his vitality, direct candor, and his innate brilliance. He was most gracious to give me an opportunity to come to Berlin in 1974 and 1975. On that occasion, I had many pleasant discussions with him on this unresolved problem and many other subjects.

References

- [1] K.G. Wilson, *Phys. Rev. B* **4**, 3174, 3184 (1971).
- [2] M. Gell-Mann and F. Low, *Phys. Rev.* **95**, 1300 (1954).
- [3] L.D. Landau and I.Ya. Pomeranchuk, *Dokl. Acad. Nauk. SSSR* **102**, 489 (1955).
- [4] S. Kim, J.B. Kogut, and M.-P. Lombardo, eprint: hep-lat/0009029 (2000).
- [5] S. Adler, *Phys. Rev. D* **5**, 3021 (1972).
- [6] K. Johnson and M. Baker, *Phys. Rev. D* **8**, 1110 (1973).
- [7] S. Adler, C. Callan, D. Gross, and R. Jackiw, *Phys. Rev. D* **6**, 2982 (1972); M. Baker and K. Johnson, *Physica A* **96**, 120 (1979); R. Acharya and P. Narayana Swamy, *Int. J. Mod. Phys. A* **12**, 3799 (1997).
- [8] S. Weinberg, *The Quantum Theory of Fields*, Vol. II (Cambridge Univ. Press, 1996).
- [9] C. Callan, *Phys. Rev. D* **2**, 1541 (1970); K. Symanzik, *Commun. Math. Phys.* **18**, 227 (1970).
- [10] P. Federbush and K. Johnson, *Phys. Rev.* **120**, 1926 (1960); B. Schroer (unpublished) (1960).

FROM Z OPERATOR TO $SO(10)$, NEUTRINO OSCILLATIONS, AND FERMI-DIRAC FUNCTIONS FOR QUARK PARTON DISTRIBUTIONS

F. BUCCELLA

*Dipartimento di Scienze Fisiche dell'Università di Napoli Federico II,
INFN Sezione di Napoli, Italy*

E-mail: franco.buccella@dsna1.na.infn.it

The oscillations advocated to explain the anomalies in solar and atmospheric neutrinos support $SO(10)$ gauge unification. In fact, within reasonable assumptions the highest matrix element of the Majorana mass matrix of right-handed neutrinos has a value in good agreement with the scale of B-L symmetry breaking in the $SO(10)$ theory with Pati-Salam intermediate symmetry. The present experimental knowledge on deep inelastic scattering supports an important role of the Pauli principle, which gives rise to for the correlation between the value of the second moment of each parton and the shape of its distribution.

1 $SO(10)$ and Neutrino Oscillations

Thirty years ago I met Hagen Kleinert at CERN and the very stimulating interaction with him and Carlos Alberto Savoy led our "Florentine" group to a breakthrough [1] in our search for the transformation from constituent to current quarks. Our results turned out to be in very good agreement with the previous phenomenological findings [2] on the chiral content of the baryon octet and on the mixing of the lower meson states.

Some years later Hagen invited me to give lectures in Berlin on the properties of the exceptional algebras coming from the underlying role of octonions [3]. This has been the starting point of my research on $SO(10)$, which can be easily found from $E(6)$ by considering 2×2 rather than 3×3 complex octonionic matrices [4].

Another German physicist, Christoph Wetterich, triggered my attention on the interesting phenomenological properties of $SO(10)$ [5]. Indeed Georgi

had discovered $SO(10)$ [6] before $SU(5)$ [7], but considered $SU(5)$ more appealing for its peculiar chiral properties [8].

The increasing precision for evaluating the gauge couplings and the lower limit on the proton lifetime excluded minimal $SU(5)$ [9] and asked at least for a modification, as its supersymmetric extension, which leads to a higher unification scale and to a value of α_s slightly larger than the experimental value.

An intriguing aspect of $SO(10)$ has been soon realized: the possibility to get massive neutrinos hierarchically lighter than the other fermions, through the see-saw mechanism [10].

By constructing positive invariants, which vanish for symmetry reasons in certain directions [11], we have been able to build a $SO(10)$ theory with $SU(4) \times SU(2) \times SU(2)$ intermediate symmetry [12] and to conclude [13] that the most promising signature of the unification might be solar neutrino oscillations.

At that time the only hint for that phenomenon came from the Homestake experiment [14], performed as proposed several years before by Bruno Pontecorvo [15], the inventor of neutrino oscillations [16].

We soon realized that the increase of the lower bound on the proton lifetime and the values of the gauge couplings would imply right-handed neutrino Majorana masses at the order of magnitude required by that phenomenon [17]. Later, a reduction of the neutrino flux has been found in the radiochemical GALLEX and SAGE experiments [18], as well as by measuring ν -e scattering at Kamiokande [19]. Finally strong evidence for ν_μ - ν_τ oscillation came from the study of atmospheric neutrinos with a square mass difference around $3.5 \times 10^{-3} \text{ (eV)}^2$ and maximal mixing [20].

In presence of several solutions with large mixing angles for solar neutrino oscillations [21], schemes with bimaximal mixing have been proposed either directly [22] for left-handed neutrinos or in the framework of see-saw models [23]: for the latter values of the matrix elements of the Majorana mass M^R of the right-handed neutrinos around $10^{10} - 10^{12} \text{ GeV}$ have been found.

We proposed a model [24] with a diagonal Dirac neutrino mass and vanishing diagonal matrix elements for M^R . As in Ref. [23] we found a negligible ν_e content in the heaviest ν^L , almost maximal mixing angle for solar neutrino oscillations and almost opposite eigenvalues for the mass of the two lightest neutrinos.

The vanishing of the 3×3 matrix element, which is a common feature

of Refs. [23,24], follows from the requirement of a not too large range for the non-vanishing matrix elements of M^R with a maximal mixing angle for $\nu_\mu - \nu_\tau$ and a diagonal Dirac mass matrix [25]. The highest matrix element of M^R , i.e. M_{23}^R , is around 7.5×10^{11} GeV. This is in good agreement with the scale of the symmetry breaking of B-L, 2.8×10^{11} , found in the theory with Pati-Salam intermediate symmetry [26].

We may conclude that the present evidence on neutrino oscillations is a good hint for $SO(10)$ gauge unification [27].

2 Pauli Principle and Parton Distributions

At the beginning of the 1970's, Murray Gell-Mann gave a seminar in Rome, discussing, apart from other topics, the importance of finding the transformation from constituent to current quarks. At the end of the talk I told him that we had found such a transformation some years ago at CERN [1]. He pointed out to us that we had found only a part of the transformation, which should also account for the presence of $q\bar{q}$ pairs, as seen in deep inelastic phenomena.

Our present knowledge is that the result found in Ref. [1] is the right description at $Q^2=0$, where baryons are a qqq state. At large Q^2 , protons and neutrons, the targets of deep inelastic phenomena, appear as continuous distributions of partons, including $\bar{q}$'s and gluons.

Richard Feynman, the father of path integrals, stated in a work with Field [28], when referring to the proton: "... the pairs $u\bar{u}$ expected to occur in the small x region (the "sea") are suppressed more than $d\bar{d}$ by the exclusion principle."

Despite a theoretical counterargument, this conjecture, as it happened often to Feynman, has been experimentally confirmed by the defect [29] in the Gottfried sum rule [30] and by the sign of the asymmetry found in the Drell-Yan production of muon pairs in pp and pn reactions [31]. Feynman and Field [28] assumed a different high- x behavior for u and d partons in the proton to comply with the dramatic fall at large x of the ratio $F_2^n(x)/F_2^p(x)$ [32]. This different behavior may be also a consequence of the Pauli principle, which may demand broader x -distributions for higher first moments.

This line of thought led Jacques Soffer and myself first to predict [33] the proportionality for $xg_1(x)$ and $F_2^p(x) - F_2^n(x)$, which holds for the contribution of the $u^\uparrow$ parton in the region dominated by valence quarks. This is the one with the highest first moment, expected to dominate at high x . It should hold for the other valence quarks, if $2u^\uparrow(x) = d(x)$, which is an approximately

good relation for their first moments. In fact the dominance of $u^\uparrow$ at high x explains both the behaviors of F_n^2/F_2^p and of $g_1^p(x)/F_1^p(x)$ [34], towards $1/4$ and 1 , respectively, as $x \rightarrow 1$.

After describing a large set of deep inelastic data with Fermi-Dirac functions for quark partons [35], a successful test for this assumption has been given in Ref. [36], by showing that the ratio between the third and the second moment of the valence partons is an increasing function of the second moment, as expected by the Pauli principle. We also found a higher second moment for s than for $\bar{s}$ [37], as previously advocated [38].

The consequences of the Pauli principle should disappear at higher Q^2 by the presence of the transverse degrees of freedom. This would imply a large dilution, the Boltzmann limit, where the shapes of the parton distribution would be independent of the first moments. However transverse degrees of freedom for partons of finite longitudinal momentum imply an additional energy, which may limit their role. The evolution equations [39] should be modified by the effect of Pauli blocking for the quarks and of the stimulated emission for gluons [40].

3 Acknowledgments

I am very happy to contribute to the book dedicated to Hagen, which gave me the occasion to recall the good old times of our collaboration and of time spent with him and Annemarie. It was not only very effective, but also very pleasant, since he is a personality rich of humour and enthusiasm, someone able to transmit these qualities to others. Bridging our national characters, his extrovert character influenced some shy aspects of mine, with the consequence of inducing a better confidence in myself which helped me in facing the challenging topics here described.

References

- [1] F. Buccella, E. Celeghini, H. Kleinert, C.A. Savoy, and E. Sorace, *Nuovo Cim. A* **69**, 133 (1970).
- [2] F. Buccella, M. De Maria, and M. Lusignoli, *Nucl. Phys. B* **6**, 430 (1968); C. Boldrighini, F. Buccella, E. Celeghini, E. Sorace, and L. Triolo, *Nucl. Phys. B* **22**, 651 (1970).
- [3] M. Gunaydin and F. Gursey, *J. Math. Phys.* **14**, 1651 (1973); F. Buccella, A. Della Selva, and A. Sciarrino, *J. Math. Phys.* **30**, 585 (1989).

- [4] F. Buccella, M. Falcioni, and A. Pugliese, *Lett. Nuovo Cim.* **18**, 441 (1977).
- [5] F. Buccella, L. Cocco, and C. Wetterich, *Nucl. Phys. B* **243**, 273 (1984).
- [6] H. Georgi, *Particle and Fields*, Ed. C.E. Carlson, (A.I.P., New York, 1975); H. Fritzsch and P. Minkowski, *Ann. Phys. (N.Y.)* **93**, 183 (1975).
- [7] H. Georgi and S.L. Glashow, *Phys. Rev. Lett.* **32**, 438 (1974).
- [8] H. Georgi, *Thirty Years of Elementary Particle Theory*, Ed. F. Buccella (Capri, 1992).
- [9] U. Amaldi, W. de Boer, and H. Furstenuau, *Phys. Lett. B* **260**, 447 (1991).
- [10] M. Gell-Mann, P. Ramond, and R. Slansky, *Supergravity* (North Holland, Amsterdam, 1980); T. Yanagida, *Unified Theory and the Baryon Number of the Universe*, Ed. O. Sawada *et al.*, KEK (1979).
- [11] F. Buccella and H. Ruegg, *Nuovo Cim. A* **67**, 61 (1982).
- [12] J.C. Pati and A. Salam, *Phys. Rev. D* **10**, 275 (1974); D. Chang, R.N. Mohapatra, and M.K. Parida, *Phys. Rev. Lett.* **52**, 1072 (1984); *Phys. Rev. D* **30**, 1052 (1984); R.N. Mohapatra and M.K. Parida, *Phys. Rev. D* **47**, 264 (1993); K.S. Babu and Q. Shafi, *Nucl. Phys. B* (Proc. Suppl.) **31**, 242 (1993).
- [13] F. Buccella, L. Cocco, A. Sciarrino, and T. Tuzi, *Nucl. Phys. B* **274**, 559 (1986).
- [14] R. Davis Jr. *et al.* (Homestake Coll.), *Phys. Rev. Lett.* **20**, 1205 (1968).
- [15] B. Pontecorvo, Report PD-205 National Research Council of Canada, Division of Atomic Energy, Chalk River (1946).
- [16] B. Pontecorvo, *Sov. Phys. JETP* **6**, 429 (1958); *Sov. Phys. JETP* **7**, 172 (1958); *Sov. Phys. JETP* **26**, 984 (1968).
- [17] F. Buccella, G. Miele, L. Rosa, P. Santorelli, and T. Tuzi, *Phys. Lett. B* **233**, 178 (1989).
- [18] P. Anselmann *et al.* (Gallex Coll.), *Phys. Lett. B* **342**, 440 (1995); J.N. Abdurashitov *et al.* (Sage Coll.), *Phys. Lett. B* **328**, 234 (1994).
- [19] Y. Fukuda *et al.* (Kamiokande Coll.), *Phys. Rev. Lett.* **77**, 1683 (1996).
- [20] Y. Fukuda *et al.* (SuperKamiokande Coll.), *Phys. Rev. Lett.* **81**, 397 (1998).
- [21] J.N. Bahcall, P.I. Krastev, and A.J. Smirnov, *Phys. Lett. B* **477**, 401 (2000).
- [22] H. Georgi and S.L. Glashow, *Phys. Rev. D* **61**, 097301 (2000).
- [23] M. Jezabek and Y. Sumino, *Phys. Lett. B* **440**, 327 (1998); B. Stech,

- Phys. Lett. B* **465**, 219 (1999).
- [24] M. Abud, F. Buccella, D. Falcone, G. Ricciardi, and F. Tramontano, *Mod. Phys. Lett. A* **15**, 15 (2000).
- [25] M. Abud and F. Buccella, *Int. J. Mod. Phys. A* **16**, 609 (2001).
- [26] F. Buccella and O. Pisanti, in *Neutrino Mixing*, Meeting in Honor of Samoil Bilenky's 70th Birthday, Turin, Italy, March 25-27, 1999, p. 62, eprint: hep-ph/9910447.
- [27] M. Abud and F. Buccella, *Results and Perspectives in Particle Physics* (La Thuile 2000); F. Buccella, *The Extreme Energy Limit* (Hanoi 2000).
- [28] R.P. Feynman and R.D. Field, *Phys. Rev. D* **15**, 2590 (1977); A. Niegawa and K. Sasaki, *Prog. Theor. Phys.* **54**, 192 (1975).
- [29] M. Arneodo *et al.* (NMC Coll.), *Phys. Rev. D* **50**, R1 (1994).
- [30] K. Gottfried, *Phys. Rev. Lett.* **18**, 1154 (1967).
- [31] A. Baldit *et al.* (NA51 Coll.), *Phys. Lett. B* **332**, 244 (1994).
- [32] T. Sloan, G. Smadja, and R. Voss, *Phys. Rep.* **102**, 405 (1988).
- [33] F. Buccella and J. Soffer, *Mod. Phys. Lett. A* **8**, 225 (1993).
- [34] J. Ashman *et al.* (EM Coll.), *Phys. Lett. B* **206**, 364 (1988); *Nucl. Phys. B* **328**, 1 (1989).
- [35] C. Bourrely, F. Buccella, G. Miele, G. Migliore, J. Soffer, and V. Tibullo, *Z. Phys. C* **62**, 431 (1994); C. Bourrely and J. Soffer, *Phys. Rev. D* **51**, 2108 (1995); *Nucl. Phys. B* **445**, 341 (1995); F. Buccella, G. Miele, G. Migliore, and V. Tibullo, *Z. Phys. C* **68**, 631 (1995); F. Buccella, G. Miele, and N. Tancredi, *Prog. Theor. Phys.* **96**, 749 (1996).
- [36] F. Buccella, I. Dorsner, O. Pisanti, L. Rosa, and P. Santorelli, *Mod. Phys. Lett. A* **13**, 441 (1998).
- [37] F. Buccella, O. Pisanti, and L. Rosa, eprint: hep-ph/0001159.
- [38] S.J. Brodsky and B.Q. Ma, *Phys. Lett. B* **381**, 317 (1996).
- [39] V.N. Gribov and L.N. Lipatov, *Sov. J. Nucl. Phys.* **15**, 438 (1972); G. Altarelli and G. Parisi, *Nucl. Phys. B* **126**, 298 (1977); Yu.L. Dokshitzer, D.I. Dyakonov, and S.I. Troyan, *Phys. Lett. B* **78**, 290 (1978); *Phys. Rep.* **58**, 269 (1980).
- [40] G. Mangano, G. Miele, and G. Migliore, *Nuovo Cim. A* **108**, 867 (1995).

A TILT AT CONSTITUENT QUARKS

C.A. SAVOY

*Service de Physique Théorique, C. E. de Saclay,
F-91191 Gif-sur-Yvette CEDEX, France
E-mail: savoy@spht.saclay.cea.fr*

The history of the transformation between current and constituent quarks, introduced 30 years ago to explain the pattern of the axial charge matrix elements and the saturation of current algebra, is briefly recollected. The work of Buccella, Kleinert *et al.* is succinctly recalled and its interpretation in terms of the so-called Melosh transformation, as well as more recent work to understand the surprisingly simple properties of the constituent quarks.

1 Once Upon a Time ...

In the winter of 1970, I was a PhD student at CERN having one of the chances of my life as I shared an office with Hagen Kleinert who was spending some months at CERN, before rejoining the FUB faculty. I was trying to understand how the Veneziano model and its infinite towers of “resonances” could match with chiral symmetries, while Hagen was interested in the symmetries of the newly discovered string theory. We then met Franco Buccella – another chance of my life – who had been contributing a lot to the understanding of hadron properties in terms of chiral symmetry. Franco told us about a suggestion by Gell-Mann to understand the relationship between the chiral group and quark model symmetries. We started a fruitful collaboration that resulted in the work [1] that I shall review here and that I consider as one of my best. Unfortunately, we were then not encouraged by the reaction of the CERN staff (ironically many of them began working two years later on the closely related Melosh transformation propagated from a collaboration of Gell-Mann with one of his PhD students). In particular, the censorship

asked us to change our title from “A Tilt Operator . . .”, suggested by Hagen, into “A Mixing Operator . . .”. But many other physicists acknowledged the interest of our paper. I wrote my thesis [2] on that a couple of years later.

Times were good as we worked a lot, I learned a lot, but we also had a lot of fun together. We would spend whole weekends with our algebraic calculations, but we would also leave on nice weekdays to ski in Le Crozet or elsewhere. And it was the time when Hagen met Annemarie. . .

Our research interests diverged since – but I have followed the impressive number of Hagen’s important contributions to many fields of theoretical physics. We seldom see each other now, but we keep being of those friends that only the “Sweet Bird of Youth” can make. After all, Hagen still is in his youth!

2 Constituent Quarks Versus Current Quarks

Chiral symmetry and its breakdown are the groundwork of hadron physics. These concepts perfectly harmonize with QCD and the existence of very light quarks. Current quarks are defined, and to a large extent observed, at short distances in the perturbative regime of QCD. They get their masses after the breakdown of the chirality dependent electroweak symmetry. But for reasons that remain obscure, two species (or flavours) of quarks are extremely light, while the strange quark is fairly light. The theory possesses an approximate $SU(3) \otimes SU(3)$ global chiral symmetry, with an almost exact $SU(2) \otimes SU(2)$ subgroup. At low energies, however, there occurs a phase transition with the spontaneous breakdown of the chiral symmetry to the chirality independent, flavour $SU(3)$, $SU(2)$ respectively, subgroup. Quarks, imprisoned inside hadron states, acquire a sort of effective mass under the action of strong confining forces. They have been called *constituent quarks*, as they define the quantum numbers and gross properties of the mesons and baryons. If I was to give a circumstantial dynamical description of chiral symmetry breaking in terms of current and constituent quarks, the applications of chiral symmetry in hadronic physics would become an impossible task. Fortunately, the fundamental approach of nonlinear realizations of spontaneously broken symmetries allows for a rigorous description of low energy interactions of the light degrees of freedom, the Goldstone bosons, among themselves as well as with the other hadrons. In the case of $SU(3) \otimes SU(3)$ the Goldstone bosons are the K ’s, π ’s and η . Let me concentrate on π ’s, henceforth restricting our attention to chiral $SU(2) \otimes SU(2)$.

Given the behavior of a hadron under the (diagonal, chiral independent) isospin $SU(2)$ subgroup of the chiral $SU(2) \otimes SU(2)$, the effective Lagrangian describing its pionic interactions, can be established in a systematic way. As a matter of fact, this approach is so general that it overlooks much of the peculiar dynamics of the system under consideration. Several other important properties of hadrons are kind of orthogonal to this approach and have to be included by different considerations. At this point I would like to evade for a while any reference to recent work on this question to go thirty years back in time and describe our work [1].

At that time the concept of quarks was rather unsatisfactory. The surprising success of the (non-relativistic) quark model in the description of the static properties of hadrons was withering away because of both the puzzle of quark confinement and the inability to deal with some problems. Indeed, the study of vector and axial currents from weak interactions, the current algebra approach, the PCAC idea and its link to the spontaneous breakdown of chiral symmetry, with pions as Goldstone bosons, were attracting attention and driving the naive quark model into a corner.

In the old-fashioned quark model, the hadronic states were approached as a series of resonances. This becomes vague if resonances are broad and ineffectual at high energies where many states are superposed, but it remains a fruitful framework at low energies. A qualitative understanding of the hadron spectrum is obtained by Wigner-classifying states according to irreducible representations of the $SU(6) \otimes O(3)$ group, or its $SU(4) \otimes O(3)$ subgroup if only non-strange quarks are considered, as I choose to do here. The building blocks are the four states of spin and isospin of u and d constituent quarks, and $O(3)$ is the group generated by the internal angular momentum $\vec{L}$, the spin of hadrons being given by $\vec{J} = \vec{L} + \vec{S}$, where $\vec{S}$ is a generator of $SU(4)$ corresponding to the "sum of the spins of the quarks". The expected mass degeneracies are badly approximated, but the agreement with the observed masses can be improved by allowing for $SU(4) \otimes O(3)$ dependent mass-shifts. In other words, $SU(4) \otimes O(3)$ is a rudimentary symmetry but a fair *classification group* for hadronic states. One can consider the linear action on hadrons to define the $SU(4)$ charges. In terms of the constituent quark "fields", these charges can be simply realized as bilinears if free-like anticommutation relations are assumed for these "fields". As generators of an approximated symmetry, these charges are not observable, *a priori*.

Quite apart stood current algebra and PCAC. Vector and axial currents

are measured in weak transitions. The vector and axial charges (here restricted for simplicity to the non-strange sector) define a $SU(2) \otimes SU(2)$ algebra. They can be realized in terms of free quark fields and their canonical commutation relations – though this fact was proved afterwards, by the observation of quark-partons in deep inelastic scattering – but this fundamental realization in terms of (perturbative) current quarks hardly allows for computations of transitions between hadron resonances.

Nevertheless, the fact that these charges close the chiral $SU(2) \otimes SU(2)$ algebra must give useful constraints. If we assume that this algebra is linearly realized in the narrow resonance approximation, it can be transformed in a series of sum rules for the axial and vector charge matrix elements. Because of the spontaneous breaking of the chiral symmetry, these matrix elements are to be taken between states of very different mass, thus corresponding to some amount of momentum transfer. In order to approach the appropriate zero momentum transfer limit, the sum rules are written in the $p \rightarrow \infty$ frame. This can be done even better by taking light like charges and collinear states.

The matrix elements of the axial charges are not only measured in weak interactions. Through the idea of PCAC, which follows from the existence of light pseudoscalar mesons, the Goldstone bosons for the broken axial symmetries, the matrix elements of the axial charges or the corresponding light like object, $\hat{Q}_5^a$ are related to the S -matrix for the pionic transition $\alpha \rightarrow \beta + \pi$ (in a collinear frame) by:

$$\langle \alpha, \pi^a | S | \beta \rangle \propto (m_\beta^2 - m_\alpha^2) \frac{1}{f_\pi} \langle \alpha | \hat{Q}_5^a | \beta \rangle, \quad (1)$$

where the proportionality factor includes normalizations and the momentum conservation δ -function. Therefore the chiral algebra can be tested from the experimental data on pionic decays of hadrons.

3 The Mixing Operator Z

Thus, the problem is how to saturate (to realize) the chiral algebra in terms of hadron states. The crucial observation is that the $SU(4)$ algebra of classification of hadrons contains a $SU(2) \otimes SU(2)$ subalgebra. Hadrons are assembled into simple irreducible representations of the classification algebra. If the chiral algebra is to be linearly realized by these hadronic states, it should be related to the classification $SU(2) \otimes SU(2)$ subalgebra by a unitary transformation that mixes different irreducible representations of the classification

algebra – as stressed by Gell-Mann in the late 60's. Let us call e^{iZ} the unitary operator that transforms the charges of the chiral and the classification $SU(2) \otimes SU(2)$ algebras into each other.

The idea of representation mixing in the realization (or saturation) of the chiral algebra was not quite new and some important work already existed when Buccella, Kleinert, and myself undertook a thorough analysis of the problem. In 1970, at CERN, we proposed an explicit algebraic form for the *mixing operator* Z , so defining its matrix elements between $SU(4) \otimes O(3)$ representations [1]. This permitted, through Eq. (1), to calculate a very large number of pionic transitions in terms of a few unknown quantities, in fair agreement with the existing data. Let me first review our result in its original formulation.

First consider meson states in the $(\mathbf{15}, L)$ representations of $SU(4) \otimes O(3)$ (strange mesons were directly included in the original transformation with mesons in the $\mathbf{35}$ of $SU(6)$). These meson states $|M\rangle$ are to be defined in the $p_3 \rightarrow \infty$ momentum frame. Consider the $SU(2) \otimes SU(2)$ subgroup of $SU(4)$ generated by the isospin charges I^a , $a = 1, 2, 3$, and the axial charges X^a . We wrote the axial charge of the $SU(2) \otimes SU(2)$ current algebra as:

$$Q_5^a = e^{i\theta Z} X^a e^{-i\theta Z}, \quad (2)$$

where the operator Z was postulated, as inferred from the experimental data, to have the form:

$$Z = \frac{1}{k} [\vec{W} \times \vec{k}]_3. \quad (3)$$

Here the W -spin operator is a "flipped" version of the spin in $SU(4)$. If $\vec{S} = (\vec{\sigma} + \vec{\bar{\sigma}})/2$, then we have

$$\vec{W} = (\vec{\sigma} - \vec{\bar{\sigma}})/2 \quad (4)$$

with a $(-)$ sign for the antiquark spin. In this representation we obtain

$$X^a = \frac{1}{2}(\tau^a \sigma_3 - \bar{\tau}^a \bar{\sigma}_3). \quad (5)$$

The operator $\vec{k}$ can be interpreted as the relative transverse momentum of the quarks inside the mesons, while the scalar function $\theta(k)$ has been chosen later in my PhD thesis as $\arctan(k/m)$, leading to a good fit to the data. The important properties are that the action of $\vec{W}$ on $SU(4)$ (or $SU(6)$) states the property $(L = 1, L_3 = \pm 1)$ of $\vec{k}$. From these properties, the matrix

elements $\langle M' | Q_5^a | M \rangle$ can be calculated in terms of a few parameters from the expression:

$$\begin{aligned} Q_5^a &= X^a \cos \theta - Y^a \sin \theta, \\ Y^a &= i[Z, X^a] = \frac{1}{2k} (\tau^a \vec{\sigma} - \bar{\tau}^a \vec{\sigma}) \cdot \vec{k}. \end{aligned} \quad (6)$$

It is worth noticing that $\{I^a, X^a, Y^a, Z\}$ close a $Sp(4)$ algebra.

The matrix elements so-calculated can now be compared with the meson decays to pions of Eq. (1). This turns out to give a very fair description of the PDG tables, both for the overall widths and for each helicity coupling.

A similar operator can be introduced for baryons, which are classified in the **56**, **70** and **20** representations of $SU(6)$. The basic property of the proposed operator Z is that it transforms in the **35** representation of $SU(6)$. Indeed, this is sufficient in order to preserve some nice $SU(6)$ predictions for the axial charge matrix elements between **56** states, the lowest octet and decuplet of baryons. The meson-baryon couplings are then calculated. By approximating the meson-baryon cross-section by resonances, one can derive the celebrated Johnson-Treiman relation from the selection rules defined by Eq. (2), and many other properties that are consistent with the data. Interestingly enough, if the unknown parameters are estimated in the non-relativistic harmonic oscillator quark model, one obtains a nice agreement with data by fitting $\langle k^2/m^2 \rangle \simeq 1$ for the fundamental states, which looks reasonable. I refer to a good textbook [3] for a review of these comparisons.

4 Wigner Rotations: A Tilt at Quarks

With few exceptions, the real impact of this issue on the particle community only occurred in 1973 after the interesting PhD thesis by Melosh [4] – who earnestly acknowledged that his first work was wrong and that he corrected it afterwards. By that time the development of QCD as the gauge theory of strong interactions and the paramount experiments at SLAC had transformed the quark concept into reality. It could cause some surprise that Melosh mostly discussed *free* quarks. Without going into the detailed reasoning let me recall the work of Melosh as reformulated in a subsequent approach [5].

Consider quark states $|p, h, \tau\rangle_q$, with momentum p , helicity h and third component τ of isospin. It is defined in terms of the rest state $|\vec{p}, h, \tau\rangle_q$ by

an appropriate boost,

$$|p, h, \tau\rangle_q = L(p) |\dot{p}, h, \tau\rangle_q. \quad (7)$$

Then, consider light-like charges defined as follows

$$Q_\alpha^a = \int d^4x \delta(x^0 + x^3) q^\dagger(x) \frac{\tau^a \hat{\sigma}_\alpha}{2} (1 + \gamma_0 \gamma_3) q(x),$$

$$\hat{\sigma}_0 = 1, \quad \hat{\sigma}_3 = \gamma_5, \quad \hat{\sigma}_{1,2} = \gamma_{1,2} \quad (\tau^0 = 1). \quad (8)$$

They close an $SU(4)$ algebra (the 16th charge is the baryon number) if (null-plane) canonical commutation relations are postulated for the relevant components of the quark field. It is possible to choose the boost $L(p)$ in Eq. (7), which defines the spin operator, so that the charges (8) act on the states in Eq. (7) in the following way:

$$Q_\alpha^a |p, h, \tau\rangle_q = \frac{1}{2} (\sigma_\alpha)_h^{h'} (\tau^a)_\tau^{\tau'} |p, h', \tau'\rangle_q, \quad (9)$$

namely, they transform as the $\mathbf{4}$ representation of $SU(4)$. Analogously, the corresponding antiquark states transform as the $\bar{\mathbf{4}}$ representation:

$$Q_\alpha^a |p, h, \tau\rangle_{\bar{q}} = \frac{\varepsilon_a}{2} (\sigma_\alpha^*)_h^{h'} (\tau^{a*})_\tau^{\tau'} |p, h', \tau'\rangle_{\bar{q}}, \quad (10)$$

where $\varepsilon_0 = \varepsilon_3 = -\varepsilon_1 = -\varepsilon_2 = 1$.

The clue to the choice of $L(p)$ is the invariance of the null-plane: $L(p)$ has to be an element of the subgroup $E_{(+)}$ of the Poincaré group that leaves the null-plane invariant. It is important to notice that $L(p)$ depends on the “effective mass” m of the quark.

Next, we construct “hadron” states out of the quark states defined by Eqs. (7)-(10), provisionally called “current” quarks. We must define a state $|P, J, J_3, T, T_3\rangle$. The hadron spin $\vec{J}$ is defined by using a boost $L(P)$ analogous to that introduced in Eq. (7) for the quark states. The multiquark states are then projected into irreducible representations of the Poincaré group by the use of the appropriate Clebsch-Gordan coefficients. It is possible to introduce an orbital angular momentum $\vec{L}$ and covariant c.m. relative momentum $\vec{k}$ for the quarks, with $\vec{J} = \vec{L} + \vec{S}$. But, in order to define $\vec{S}$ as the sum of the spins of the quarks one has to redefine the quark states and consider the Wigner rotated ones:

$$|p(P), h, \tau\rangle_q = R(p, P)_h^{h'} |p, h', \tau\rangle_q. \quad (11)$$

At this point one can introduce *classification charges* $\tilde{Q}_\alpha^a$ acting on the constituent quark states $|p(P), h, \tau\rangle_q$ with matrices analogous to Eqs. (9, 10). They close an $SU(4)$ group that classify, together with the $\vec{L}$ generated $O(3)$, the hadronic states in a simple way. From the transformation (11) between current and constituent quarks, one deduces for the one-quark states

$$Q_\alpha^a = R^{-1}(p, P) \tilde{Q}_\alpha^a R(p, P). \quad (12)$$

For hadronic states consisting of a quark-antiquark pair (mesons) or three quarks (baryons), it follows:

$$\begin{aligned} Q_\alpha^a &= e^{-iZ} \tilde{Q}_\alpha^a e^{iZ}, \\ e^{iZ} &= \otimes R(p_i, P). \end{aligned} \quad (13)$$

Therefore, the unitary transformation between the light like currents and the classification charges takes the simple form of a direct product of Wigner rotations. Even if its form is well defined one cannot calculate its matrix elements between hadron states since we do not know their wave functions. The only model-independent results are obtained by determining the properties of the operator Z in Eq. (13) under the classification group generated by $\{\tilde{Q}_\alpha^a\}$, and then applying the Wigner-Eckart theorem. It turns out that these algebraic properties of Z are precisely those proposed by Buccella, Kleinert *et al.* in 1970.

In practice, we concentrate on the observable charges Q_α^a , with $a = 1, 2, 3$ and $\alpha = 0, 3$. They close the chiral $SU(2) \times SU(2)$, and Q_3^a matrix elements are related to pionic transitions as given by Eq. (1). Can we conclude that Eqs. (12) and (13) define the transformation between current and constituent quarks? Certainly not, as should be clear from the above derivation. Both quark states are supposed to have some effective mass m , which in the fits turn out to be of order $\mathcal{O}(300 \text{ MeV})$. This is a typical constituent quark mass. The transformations (12) and (13) have a very simple interpretation indeed. Wigner rotations are applied to the constituents to harmonize the definition of their spins so that they can be added up in the definition of hadron states. It has an impact on the calculation of pionic transitions and solves (to a large extent) some puzzles of the naive quark model that would roughly correspond to neglect these Wigner rotations and therefore the e^{iZ} transformation. The fact that hadron resonances can be described so simply in terms of constituent quarks is to be taken as an empirically rewarding approximation.

The fundamental assumption is however encoded in Eqs. (9) and (10). Constituent quarks, at least after a careful definition, behave under the action of $SU(2) \times SU(2)$ charges just like the light free quarks. This is rather surprising since their “effective mass” m provides a measure of the chiral symmetry breaking. And the success of the approach of Buccella *et al.* (and many others following the work by Melosh) basically relies on this simple behavior of the constituent quarks.

5 “After Many a Summer ...”

In the 80’s, Georgi and Manohar have suggested [6] to write an effective Lagrangian for constituent quarks, which they call the “chiral quark model”, that also incorporates the $SU(2) \otimes SU(2)$ spontaneous breaking by including the pions in the standard formalism of nonlinear realizations. More recent papers by Weinberg [7] then provide an interesting interpretation of the successes of the constituent quark approach to pionic transitions. It is based on previous work by himself, and he reformulated the problem of the saturation of the chiral algebra, without direct use of the axial charges. Indeed, the axial symmetries being badly broken, it is questionable whether the charges and their algebra can be even defined. Equation (1) is introduced as a definition of *pionic reduced matrix elements* $\langle \beta | Q_5^a | \alpha \rangle$. The so-defined matrices Q_5^a transform like the pion, i.e. as a vector under the conserved isospin generators. Then Weinberg considers the amplitudes for $\pi^i + \alpha \rightarrow \pi^j + \beta$. From the Adler-Weisberger sum rules, in which the amplitudes near the threshold are fixed by low-energy theorems and related through resonance saturation of dispersion relations to the matrix elements in Eq. (1), Weinberg derives the commutator of the matrices Q_5^a and shows that they close the $SU(2) \otimes SU(2)$ algebra of the broken chiral symmetry together with the isospin charges.

As a matter of fact, the approach is general in the sense that it applies to reduced collinear Goldstone boson amplitudes that can be shown to compose the Lie algebra of the whole broken symmetry group along with the unbroken symmetry generators. The key ingredient is the assumption on the asymptotic behavior of the amplitudes. Weinberg refers to these restored algebras as “mended symmetries”. Along the same lines Weinberg derived commutation relations involving the $(\text{mass})^2$ operator from superconvergence relations that follow from unsubtracted dispersion relations for amplitudes with “exotic” quantum numbers.

Motivated by the work of Georgi and Manohar, Weinberg then considered

pion-(constituent) quark scattering and derived the algebraic relations for reduced pionic couplings and masses. It turns out that "constituent quarks" saturate by themselves the sum rules, so that they define an irreducible representation of $SU(2) \otimes SU(2)$. This is precisely the content of Eqs. (9) and (10). Of course, an understanding of the concept of constituent quarks on more fundamental bases is still lacking.

In a sense, the analysis started by Melosh and topped by the above mentioned recent work, provides a physical interpretation for the mixing operator defined by Buccella *et al.* in 1970. One could wonder why such an operator, with simple algebraic properties postulated *ab initio*, was essentially correct. The answer is that, as often, simplicity is a guide to find theories which reproduce a certain amount of *experimental* data!

Before I put a period to this recollection, let me stress the completely new insight on this problem we have at present. Most of the enlightenment comes from the developments in QCD. Current quarks are appropriately defined in the short distance, chiral-symmetric regime of strong interactions. In order to retrace the simple phenomenological properties of the confined constituent quarks discussed here, back to those current quarks, one needs non-perturbative approaches. Of course, we think we know why quarks are confined, still we are not able to give an accurate formulation of confinement. The transition to the broken chiral symmetry phase at low energies as well as some hadronic matrix elements can be studied and estimated by powerful non-perturbative methods. An obvious one is lattice computation. An instructive approximation is provided by the $1/N$ development as shown, e.g. in some recent work by Weinberg. Recently, the confinement property was given a new formulation in terms of dualities, in supersymmetric gauge theories. But it will certainly take a while before one could theoretically understand the idea of constituent quarks.

In the issue, the intense investigation, on phenomenological grounds, of the realization of chiral symmetry has eventually led to a simple formulation of the problem in terms of constituent quark properties. This provides an adequate framework to study the subject from the fundamental viewpoint: current quarks and QCD.

References

- [1] F. Buccella, H. Kleinert, C.A. Savoy, E. Celeghini, and E. Sorace, *Nuovo Cim. A* **69**, 133 (1970).

-
- [2] C.A. Savoy, *On the Relation Between the Algebra of Currents and the Classification Group for Hadrons*, PhD Dissertation, University of Geneva (1973).
 - [3] F.E. Close, *An Introduction to Quarks and Partons* (Academic Press, New York, 1979).
 - [4] H.J. Melosh, *Phys. Rev. D* **9**, 1095 (1974).
 - [5] F. Buccella, C.A. Savoy, and P. Sorba, *Lett. Nuovo Cim.* **10**, 455 (1974).
 - [6] H. Georgi and A. Manohar, *Nucl. Phys. B* **234**, 189 (1984).
 - [7] S. Weinberg, *Phys. Rev.* **177**, 1083 (1969); *Phys. Rev. Lett.* **65**, 1177 (1990); *Phys. Rev. Lett.* **65**, 1181 (1990); *Phys. Rev. Lett.* **67**, 3473 (1991).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

THE BREAKING OF ISOSPIN AND THE ρ - ω SYSTEM

H. FRITZSCH

*Sektion Physik, Ludwig-Maximilians-Universität München,
Theresienstrasse 37, D-80333 München, Germany
E-mail: bm@theorie.physik.uni-muenchen.de*

Simple quark models for the low-lying vector mesons suggest a mixing between the u - and d -flavors and a violation of the isospin symmetry for the ρ - ω system much stronger than observed. It is shown that the chiral dynamics, especially the QCD anomaly, is responsible for a restoration of the isospin symmetry in the ρ - ω system.

Although there are no doubts that all observed strong-interaction phenomena can be described within the theory of QCD, a quantitative description of the strong interaction phenomena in the low-energy sector is still lacking. Only a few features of the low-energy phenomena have been partially understood by the lattice gauge theory approach.

The low-energy sector of the physics of the strong interactions is dominated by the low-lying pseudoscalar mesons (π, K, η, η') and the low-lying vector mesons (ρ, ω, K^*, ϕ). It is well known that the structures of the quark wave functions of the pseudoscalar mesons (0^{-+}) and of the vector mesons (1^{--}) differ substantially.

In the vector meson channel there is a strong mixing between the eight components of the $SU(3)$ octet (wave function: $(\bar{u}u + \bar{d}d - 2\bar{s}s)/\sqrt{6}$) and of the $SU(3)$ singlet (wave function: $(\bar{u}u + \bar{d}d + \bar{s}s)/\sqrt{3}$). The mixing strength is such that the mass eigenstates ω and ϕ have approximately the quark contents $(\bar{u}u + \bar{d}d)/\sqrt{2}$ and $\bar{s}s$, respectively. While this feature looks peculiar when viewed upon from the platform of the underlying $SU(3)$ symmetry, it finds a simple interpretation if one takes into account the Zweig rule [1], which states that the mixing must take place in such a way that quark lines are neither destroyed nor created.

On the other hand the pseudoscalar mesons follow the pattern prescribed by the $SU(3)$ symmetry in the absence of singlet-octet mixing. The neutral mass eigenstates η and η' are nearly an $SU(3)$ octet or $SU(3)$ singlet:

$$\eta \approx \frac{1}{\sqrt{6}} (\bar{u}u + \bar{d}d - 2\bar{s}s), \quad (1)$$

$$\eta' \approx \frac{1}{\sqrt{3}} (\bar{u}u + \bar{d}d + \bar{s}s). \quad (2)$$

This indicates a large violation of the Zweig rule in the 0^{-+} channel [2,3]. Large transitions between the various $(\bar{q}q)$ -configurations must take place. In QCD the strong mixing effects are related to the spontaneous breaking of the chiral $U(1)$ symmetry normally attributed to instantons. Effectively the mass term for the pseudoscalar mesons can be written as follows, neglecting the effects of symmetry breaking in the gluonic mixing term [4-6],

$$M_{\bar{q}q}^2 = \begin{pmatrix} M_u^2 & 0 & 0 \\ 0 & M_d^2 & 0 \\ 0 & 0 & M_s^2 \end{pmatrix} + \lambda \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad (3)$$

where M_u^2, M_d^2 and M_s^2 are the M^2 -values of the masses of quark composition $\bar{u}u, \bar{d}d$ and $\bar{s}s$, respectively.

It is well known that the mass and mixing pattern of the 0^{-+} -mesons is described by such an ansatz [2]. The parameter λ which describes the mixing strength due to the gluonic forces, is essentially given by the η' -mass: $\lambda \cong 0.24 \text{ GeV}^2$. Since λ is large compared to the strength of $SU(3)$ violation given by the s -quark mass, large mixing phenomena are present in the 0^{-+} channel, as seen in the corresponding wave functions.

The situation is different in the vector meson 1^{--} channel. Here, the gluonic mixing term is substantially smaller than the strength of $SU(3)$ violation such that the Zweig rule is valid to a good approximation. If one describes the mass matrix for the vector mesons in a similar way as for the pseudoscalar, we have

$$M_{\bar{q}q} = \begin{pmatrix} M(\bar{u}u) & 0 & 0 \\ 0 & M(\bar{d}d) & 0 \\ 0 & 0 & M(\bar{s}s) \end{pmatrix} + \tilde{\lambda} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad (4)$$

where $M(\bar{q}q)$ denotes the mass of a vector meson with quark composition $\bar{q}q$ in the absence of the mixing term. The magnitude of the mixing term $\tilde{\lambda}$ can be obtained in a number of different ways, e.g by considering the ρ_0 - ω

mass difference. Neglecting the isospin violation caused by the m_d - m_u mass splitting, the gluonic mixing term is responsible for the ρ_0 - ω mass shift:

$$M_\omega - M_\rho = 2\tilde{\lambda}, \quad (5)$$

$$\tilde{\lambda} \cong 6.0 \pm 0.5 \text{ MeV}. \quad (6)$$

In QCD the isospin symmetry is violated by the mass splitting between the u - and d -quark. Typical estimates give:

$$\frac{m_d - m_u}{(m_d + m_u)/2} \cong 0.58. \quad (7)$$

The observed smallness of isospin symmetry breaking effects is usually attributed to the fact that the mass difference $m_d - m_u$ is small compared to the QCD scale Λ_{QCD} . However, in the case of the vector mesons the QCD interaction enters in two different ways:

a) In the chiral limit of vanishing quark masses the masses of the vector mesons are solely due to the QCD interaction, i.e. $M = \text{const} \cdot \Lambda_{\text{QCD}}$.

b) The QCD mixing term will lead to a mixing among the various flavour components such that the $SU(3)$ singlet (quark composition $(\bar{u}u + \bar{d}d + \bar{s}s)/\sqrt{3}$) is lifted upwards compared to the two other neutral components given by the wave functions $(\bar{u}u - \bar{d}d)/\sqrt{2}$ and $(\bar{u}u + \bar{d}d - 2\bar{s}s)/\sqrt{6}$. The corresponding mass shift is given by $3\tilde{\lambda}$.

We approach the real world by first introducing the mass of the strange quark. As soon as m_s becomes larger than $3\tilde{\lambda}$, substantial singlet-octet mixing sets in, and the mass of one vector meson increases until it reaches the observed value of the ϕ -mass. At the same time the Zweig rule, which is strongly violated in the chiral $SU(3)_L \times SU(3)_R$ limit becomes more and more valid.

The validity of the Zweig rule is determined by the ratio $m_s/\tilde{\lambda}$. If this ratio vanishes, the Zweig rule is violated strongly. In reality, taking m_s (1GeV) $\approx$ 150 MeV, the ratio $m_s/\tilde{\lambda}$ is about 25, implying that the Zweig rule is nearly exact.

In a second step we introduce the light quark masses m_u and m_d . We concentrate on the non-strange vector mesons. If the gluonic mixing interaction were turned off, the mass eigenstates would be $v_u = |\bar{u}u\rangle$ and $v_d = |\bar{d}d\rangle$. The masses of these mesons are given by:

$$M(v_u) = \langle v_u | H^0 + m_u \bar{u}u | v_u \rangle, \quad (8)$$

$$M(v_d) = \langle v_d | H^0 + m_d \bar{d}d | v_d \rangle. \quad (9)$$

Here H^0 is the QCD-Hamiltonian in the chiral limit $m_u = m_d = 0$. Thus the masses can be written as

$$M(v_u) = M_0 + 2m_u c, \quad (10)$$

$$M(v_d) = M_0 + 2m_d c. \quad (11)$$

The constant c is given by the expectation value of $\bar{q}q$. The introduction of the light quark masses induces positive mass shifts for both v_u and v_d . These mass shifts can be estimated by considering the corresponding mass shifts of the charged K^* -mesons. One finds [7,8]:

$$M(v_d) - M(v_u) \cong 2(m_d - m_u)c \cong 1.7 \text{ MeV}. \quad (12)$$

It is remarkable that this mass shift is of similar order of magnitude as the mass shift between the isosinglet and isotriplet state in the chiral limit, where isospin symmetry is valid. This implies that the strength of the gluonic mixing term is comparable to the $\Delta I = 1$ mass term. It follows that the eigenstates of the mass operator taking both the violation of isospin and the gluonic mixing into account will not quite be eigenstates of the isospin symmetry.

For the ρ_0 - ω system the mass operator takes the form:

$$M = \begin{pmatrix} M(v_u) & 0 \\ 0 & M(v_d) \end{pmatrix} + \tilde{\lambda} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad (13)$$

Using $M(v_u) = M(\bar{u}u)$, $M(v_d) = M(\bar{d}d)$ and $\tilde{\lambda} = 5.9 \text{ MeV}$, we find

$$\begin{aligned} |\rho_0\rangle &= 0.997 \left| \frac{1}{\sqrt{2}}(\bar{u}u - \bar{d}d) \right\rangle + 0.071 \left| \frac{1}{\sqrt{2}}(\bar{u}u + \bar{d}d) \right\rangle, \\ |\omega\rangle &= -0.071 \left| \frac{1}{\sqrt{2}}(\bar{u}u - \bar{d}d) \right\rangle + 0.997 \left| \frac{1}{\sqrt{2}}(\bar{u}u + \bar{d}d) \right\rangle. \end{aligned}$$

The mixing angle α describing the strength of the triplet-singlet mixing is about -4.1° , i.e. a sizeable violation of isospin symmetry is obtained. Neither is the ρ_0 -meson an isospin triplet, nor is the ω -meson an isospin singlet.

The conclusions follow directly from the observed smallness of the gluonic mixing in the vector meson channel and the $m_u - m_d$ mass splitting, as observed e.g. in the mass spectrum of the K^* -mesons. Nevertheless they are in direct conflict with observed facts. According to Eq. (13), the probability of the ρ_0 -meson to be an $I = |\bar{u}u + \bar{d}d\rangle/\sqrt{2}$ -state is $\sin^2 \alpha \cong 0.51\%$. Taking into account the observed branching ratio for the decay $\omega \rightarrow \pi^+\pi^-$, $\text{BR} \cong (2.21 \pm 0.30)\%$, this probability is bound to be less than 0.12% , in disagreement

with the value derived above. Obviously our theoretical estimate cannot be correct.

We consider the discrepancy described above as a serious challenge for our understanding of the low-energy sector of QCD.

Conclusions

The mass difference $\Delta M = M(v_d) - M(v_u)$ must be smaller than estimated above. In order to reproduce the observed branching ratio for the decay $\omega \rightarrow \pi^+ \pi^-$, ΔM cannot exceed 0.82 MeV, implying that our simple estimates based on quark model considerations cannot be correct. This can be seen as follows. We consider the following two-point functions

$$\begin{aligned} u_{\mu\nu} &= \langle 0 | \bar{u}(x) \gamma_\mu u(x) \bar{u}(y) \gamma_\nu u(y) | 0 \rangle, \\ d_{\mu\nu} &= \langle 0 | \bar{d}(x) \gamma_\mu d(x) \bar{d}(y) \gamma_\nu d(y) | 0 \rangle, \\ m_{\mu\nu} &= \langle 0 | \bar{d}(x) \gamma_\mu d(x) \bar{u}(y) \gamma_\nu u(y) | 0 \rangle. \end{aligned} \quad (14)$$

The mixed spectral function $m_{\mu\nu}$ is expected to be essentially zero in the low-energy region, since the two different currents can communicate only via intermediate gluonic mesons. In perturbative QCD these states would be represented by three gluons. The vanishing of $m_{\mu\nu}$ implies the validity of the Zweig rule.

The spectral functions $u_{\mu\nu}$ and $d_{\mu\nu}$ are strongly dominated at low energies by the ρ_0 - and ω -resonances. The actual intermediate states contributing to the two-point functions are 2π - and 3π -states. However, a violation of the isospin symmetry due to the $u - d$ -quark mass splitting does not show up in the π -meson spectrum. The $\pi^+ - \pi^0$ mass splitting is solely due to the electromagnetic interaction. It follows that resonant (2π) or (3π) states, i.e. the ρ - ω -resonances, cannot display the effects of the isospin violation either, and the mass difference $\Delta M = M(v_d) - M(v_u)$ must be very small.

Although the isospin symmetry is broken explicitly by the $u - d$ mass terms, this symmetry violation does not show up in the ρ - ω sector. The isospin symmetry breaking is shielded by the pion dynamics. Effectively, the symmetry is restored by dynamical effects. Here the gluon anomaly plays an important role. The effect of a dynamical symmetry restoration by nonperturbation effects discussed here might be reproduced by lattice simulations. It might be that similar symmetry restoration effects are present in other situations, for example in the electroweak sector, which is sensitive to the

dynamics in the TeV region.

Acknowledgments

I am happy that this paper could be included in the volume dedicated to my friend Hagen Kleinert on the occasion of his 60th birthday. We never collaborated, but our worldlines crossed regularly e.g. at CERN and in Berlin. We met for a longer time at Caltech in Pasadena in 1973/74, where Hagen and Annemarie spent a sabbatical year during the time when I was developing QCD with Murray Gell-Mann. This is where Hagen encountered Dick Feynman with whom he published a paper a few years later. I have problems seeing Hagen as a sixty-years old colleague, since he appears like a senior post-doc. In Aspen Hagen would never be granted the price-reduction on the lift-ticket offered to anyone above 60, unless he takes his birth certificate along - but also in ten years from now Hagen will have the same problem.

This work is supported by the Deutsche Forschungsgemeinschaft, DFG-No. FR 412/25-2.

References

- [1] S. Okubo, *Phys. Lett.* **5**, 163 (1963); G. Zweig, CERN Report No. 8419/TH 414 (1964); J. Iizuka, *Prog. Theor. Phys. Suppl.* **37-8**, 21 (1996).
- [2] H. Fritzsch and P. Minkowski, *Nuovo Cim.* **30**, 393 (1975).
- [3] G. Veneziano, *Mod. Phys. Lett. A* **4**, 1605 (1989); G.M. Shore and G. Veneziano, *Nucl. Phys. B* **381**, 23 (1992).
- [4] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Phys. Lett. B* **47**, 365 (1973).
- [5] G. 't Hooft, *Phys. Rev. D* **14** 3432 (1976).
- [6] M.A. Shifman, *Phys. Rep.* **209**, 341 (1986).
- [7] M.D. Scadron, *Phys. Rev. D* **29**, 2076 (1984).
- [8] H. Fritzsch and A. Müller, eprint: hep-ph/0011278.

ANALYTIC CONFINEMENT AND REGGE SPECTRUM

G.V. EFIMOV

*Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research,
Dubna, Russia*

E-mail:efimovg@thsun1.jinr.ru

Using a simple relativistic QFT model of scalar fields, we demonstrate that the analytic confinement, where propagators of initial constituent particles are entire analytic functions in the complex p^2 -plane, and the weak coupling constant lead to the linear Regge behavior of two- and three-particle bound states.

1 Introduction

The problem of Regge trajectories in particle physics is an active area of investigation beginning in the 1960's [1]. Experimental data show that meson and baryon Regge trajectories are almost linear, though the latter can only be approximatively linear [2]. Standard theoretical calculations which give the linear spectrum for meson and baryon mass squares are based on a relativized Schrödinger equation with an appropriate linearly increasing potential [3,4]. Perturbative QCD approaches have shown that Regge trajectories are non-linear [5,6].

We show in this article, using a simple relativistic model of scalar fields, that in the framework of QFT analytic confinement of constituent particles (propagators are entire analytic functions in the momentum p^2 -space) and weak coupling constant (the Bethe-Salpeter equation can be used) lead to the linear Regge spectrum of bound states.

Thus, if the QCD vacuum results in analytic confinement of quarks and gluons [7–9] and the QCD coupling constant α_s is small, hadron bound states are expected to have, at least asymptotically, the linear Regge spectrum, and masses of these states can be calculated by the Bethe-Salpeter equation.

2 The Model

We consider a simple quantum field model with confinement and demonstrate the properties of possible bound states that can be interpreted as standard physical particles. Let $\Phi(x)$ and $\varphi(x)$ be two scalar fields with the Lagrangian in the Euclidean metrics

$$\begin{aligned} \mathcal{L}(x) = & -\Phi^+(x)S^{-1}(\square)\Phi(x) - \frac{1}{2}\varphi(x)D^{-1}(\square)\varphi(x) \\ & - g\Phi^+(x)\Phi(x)\varphi(x) - g\varphi^3(x) \end{aligned} \quad (1)$$

with

$$S^{-1}(\square) = \frac{\Lambda^2}{\epsilon} \cdot e^{\frac{\square}{\Lambda^2}}, \quad D^{-1}(\square) = \Lambda^2 e^{\frac{\square}{\Lambda^2}}.$$

The equation for a free field $A(x) = (\Phi(x), \varphi(x))$ looks like

$$e^{\frac{\square}{\Lambda^2}} A(x) = 0, \quad \text{or} \quad e^{-\frac{p^2}{\Lambda^2}} \tilde{A}(p) = 0. \quad (2)$$

The solutions are $A(x) \equiv 0$ ($\varphi(x) \equiv 0$ and $\Phi(x) \equiv 0$), because the function $e^{-p^2/\Lambda^2} \neq 0$, i.e. it has no zeroes for any real or complex p^2 . Exactly this property means analytic confinement. The scale of the confinement region is characterized by the parameter Λ . The fields $\Phi(x)$ and $\varphi(x)$ can exist in a virtual state only, so that they can be called *virton fields* [11]. In addition, one can say that the field $\Phi(x)$ describes constituent particles (scalar “quarks”), and the field $\varphi(x)$ particles-carriers (scalar “gluons”). The parameter $\epsilon \ll 1$ means that “the mass” of “the quark” Φ is much larger then “the mass” of “the gluon” ϕ . In other words, the Compton wavelength of “the massive quark” Φ is smaller than the diameter of the confinement region.

The propagators in momentum and x -spaces look like

$$\begin{aligned} \tilde{S}(p^2) &= \frac{\epsilon}{\Lambda^2} e^{-\frac{p^2}{\Lambda^2}}, \quad \tilde{D}(p^2) = \frac{1}{\Lambda^2} e^{-\frac{p^2}{\Lambda^2}}, \\ S(x) &= \frac{\Lambda^2 \epsilon}{(4\pi)^2} e^{-\frac{\Lambda^2}{4} x^2}, \quad D(x) = \frac{\Lambda^2}{(4\pi)^2} e^{-\frac{\Lambda^2}{4} x^2}. \end{aligned} \quad (3)$$

They are entire analytic functions in the complex p^2 -plane. This guarantees the confinement of “particles” $\Phi(x)$ and $\varphi(x)$ in each perturbation order in the dimensionless coupling constant

$$\lambda = \left(\frac{g}{4\pi\Lambda} \right)^2 \leq 1. \quad (4)$$

The mechanism of bound states can be described in the following way [10].
Let us consider the partition function

$$Z = \iiint \delta\Phi \delta\Phi^+ \delta\varphi e^{-(\Phi^+ S^{-1} \Phi) - \frac{1}{2}(\varphi D^{-1} \varphi) - g(\Phi^+ \Phi \varphi) - g\varphi^3}.$$

After integration over ϕ in the one-particle exchange approximation, valid for small coupling constants λ , Z reads

$$Z = \iint \delta\Phi \delta\Phi^+ e^{-(\Phi^+ S^{-1} \Phi) + L_2[\Phi] + L_3[\Phi] + O[\Phi^8]}, \quad (5)$$

where

$$\begin{aligned} L_2[\Phi] &= g^2(\Phi_1^+ \Phi_1 D_{12} \Phi_2^+ \Phi_2), \\ L_3[\Phi] &= g^4(\Phi_1^+ \Phi_1 \Phi_2^+ \Phi_2 \Phi_3^+ \Phi_3 \Gamma_{123}). \end{aligned}$$

Here,

$$\Phi_j^+ = \Phi(x_j), \quad D_{ij} = D(x_i - x_j),$$

and

$$\Gamma_{123} = \Gamma(x_1, x_2, x_3) = \int dy D(x_1 - y) D(x_2 - y) D(x_3 - y),$$

where the integration over dx_j ($j = 1, 2, 3$) is implied.

The term

$$L_4[\Phi] = g^4(\Phi_1^+ \Phi_1 D_{12} \Phi_2^+ S_{23} \Phi_3 D_{34} \Phi_4^+ \Phi_4)$$

is important in the Faddeev equations for a three-body problem. In our case, it is small in comparison with L_3 according to our assumption $\epsilon \ll 1$.

3 Two-Particle Bound States

Two-particle bound states are defined by the term L_2 which can be transformed as

$$L_2 = g^2(\Phi_1^+ \Phi_2 \sqrt{D_{12}} \cdot \sqrt{D_{12}} \Phi_2^+ \Phi_1).$$

Let us use the Gaussian representation

$$e^{L_2} = \iint \delta B \delta B^+ e^{-(B_{12}^+ B_{12}) + g[(B_{12}^+ \Phi_1^+ \Phi_2 \sqrt{D_{12}}) + (\sqrt{D_{12}} \Phi_2^+ \Phi_1 B_{12})]},$$

where the bilocal field $B_{12} = B(x_1, x_2)$ is introduced.

We substitute this representation into the partition function (5) and integrate over Φ . The result is

$$Z = \iint \delta B \delta B^+ e^{-(B_{12}^+ B_{12}) + 2g^2 (B_{12} \sqrt{D_{12}} S_{11'} S_{22'} \sqrt{D_{1'2'}} B_{1'2'}) + O[g^3]} \quad (6)$$

with $S_{11'} = S(x_1 - x'_1)$, and so on. We introduce the variables

$$x_1 = x + \xi, \quad x_2 = x - \xi$$

and denote $B(x_1, x_2) = \frac{1}{4} \mathcal{B}(x, \xi)$. After some calculations with (3) we get

$$\begin{aligned} & 2g^2 (B_{12} \sqrt{D_{12}} S_{11'} S_{22'} \sqrt{D_{1'2'}} B_{1'2'}) \\ &= \frac{32g^2 \Lambda^4 \epsilon^2}{(4\pi)^4} \iint dx dx' \iint d\xi d\xi' \mathcal{B}(x, \xi) e^{-\frac{\Lambda^2}{2} (x-x')^2} K(\xi, \xi') \mathcal{B}(x', \xi'). \end{aligned}$$

Here the kernel $K(\xi, \xi')$ can be diagonalized

$$K(\xi, \xi') = e^{-\Lambda^2 (\xi^2 - (\xi\xi') + \xi'^2)} = \sum_Q U_Q(\xi) \kappa_Q U_Q(\xi'), \quad (7)$$

$$\kappa_Q = \kappa_{nl} = \left(\frac{2\pi}{\Lambda^2 (2 + \sqrt{3})} \right)^2 \cdot \left(\frac{1}{2 + \sqrt{3}} \right)^{2n+l}$$

(for the eigenfunctions $U_Q(\xi)$ see the Appendix).

The numbers $Q = (nl\{\mu\})$ can be considered as radial n , orbital l , and magnetic $\{\mu\} = (\mu_1, \dots, \mu_l)$ quantum numbers.

It should be stressed that the diagonalization of the kernel $K(\xi, \xi')$ is nothing else but the solution of the Bethe-Salpeter equation in the one-boson exchange when the propagators are defined by (3). To go to the standard form of the Bethe-Salpeter equation, we have to introduce a new function $U_Q(y) = \sqrt{D(y)} \Psi_Q(y)$ and to pass to the momentum representation [10].

Let us introduce the function

$$\mathcal{B}_Q(x) = 4 \int d\xi U_Q(\xi) \mathcal{B}(x, \xi), \quad \tilde{\mathcal{B}}_Q(p) = \int dx e^{ipx} \mathcal{B}_Q(x). \quad (8)$$

Then the Gaussian quadratic measure defining the partition function (6) gets the form

$$Z = \prod_Q \delta \tilde{\mathcal{B}}_Q \tilde{\mathcal{B}}_Q^+ \exp \left\{ - \int \frac{dp}{(2\pi)^4} \sum_Q \tilde{\mathcal{B}}_Q^+(p) [1 - \Pi_Q(p^2)] \tilde{\mathcal{B}}_Q(p) \right\} \quad (9)$$

with

$$\Pi_Q(p^2) = \frac{\lambda}{\lambda_{2c}} \left(\frac{1}{2 + \sqrt{3}} \right)^{2n+l} e^{-\frac{p^2}{2\Lambda^2}}, \quad \lambda_{2c} = \frac{(2 + \sqrt{3})^2}{2\epsilon^2}. \quad (10)$$

The equation

$$1 = \Pi_Q(-M_Q^2) \quad (11)$$

defines the mass $M_Q = M_{nl}$ of two-particle bound states with quantum numbers Q :

$$M_Q^2 = M_{nl}^2 = \Lambda^2 2 \ln \frac{\lambda_{2c}}{\lambda} + (2n + l) \Lambda^2 2 \ln(2 + \sqrt{3}). \quad (12)$$

One can see that this spectrum has a purely linear Regge behavior.

The function $1 - \Pi_Q(p^2)$ defines the Gaussian measure or kinetic term in the functional integral (9). To represent this function in the standard form, let us develop it in the vicinity of the point $p^2 = -M_Q^2$:

$$1 - \Pi_Q(p^2) = Z_Q(p^2 + M_Q^2) + O((p^2 + M_Q^2)^2)$$

with

$$Z_Q = -\Pi'_Q(-M_Q^2) = \frac{\lambda}{\lambda_c} \cdot \frac{e^{\frac{M_Q^2}{2\Lambda^2}}}{(2 + \sqrt{3})^{l+2n}} \cdot \frac{1}{2\Lambda^2}.$$

Thus, the kinetic term in Z [Eq. (9)] reads

$$\begin{aligned} & \left(\tilde{\mathcal{B}}_Q^+(p) [1 - \Pi_Q(p^2)] \tilde{\mathcal{B}}_Q(p) \right) \\ &= Z_Q \left(\tilde{\mathcal{B}}_Q^+(p) [(p^2 + M_Q^2) + O((p^2 + M_Q^2)^2)] \tilde{\mathcal{B}}_Q(p) \right), \end{aligned}$$

and after the renormalization

$$\tilde{\mathcal{B}}_Q(p) \rightarrow Z_Q^{-1/2} \tilde{\mathcal{B}}_Q(p)$$

it gets the standard form.

4 Three-Particle Bound States

Three-particle bound states are defined by the term $L_3[\Phi]$ in (5) which can be transformed as

$$L_3 = g^4 (\Phi_1^+ \Phi_2^+ \Phi_3^+ \sqrt{\Gamma_{123}} \cdot \sqrt{\Gamma_{123}} \Phi_1 \Phi_2 \Phi_3)$$

with

$$\Gamma_{123} = \Gamma(x_1, x_2, x_3) = \frac{\Lambda^2}{9(4\pi)^4} e^{-\frac{\lambda^2}{2}(3\xi_1^2 + \xi_2^2)},$$

where

$$x_1 = x + \xi_1 + \xi_2, \quad x_2 = x + \xi_1 - \xi_2, \quad x_3 = x - 2\xi_1. \quad (13)$$

We use the Gaussian representation

$$e^{L_3} = \iint \delta H \delta H^+ e^{-(H_{123}^+ H_{123}) + g^4 [(H_{123}^+ \sqrt{\Gamma_{123}} \Phi_1 \Phi_2 \Phi_3) + (\Phi_1^+ \Phi_2^+ \Phi_3^+ \sqrt{\Gamma_{123}} H_{123})]},$$

where $\Phi_j = \Phi(x_j)$, $j = 1, 2, 3$, and $H = H_{123} = H(x_1, x_2, x_3)$ are tri-local fields. We substitute this representation into (5). After integration over Φ , the partition function reads

$$Z = \iint \delta \Phi \delta \Phi^+ e^{-(\Phi^+ S^{-1} \Phi) + L_3} = \iint \delta H \delta H^+ e^{-(H_{123}^+ H_{123}) + g^4 W[H] + O(g^6)},$$

where

$$g^4 W[H] = 6g^4 (H_{123}^+ \sqrt{\Gamma_{123}} S_{11'} S_{22'} S_{33'} \sqrt{\Gamma_{1'2'3'}} H_{1'2'3'}).$$

Using the notation

$$\tilde{H}(p; \xi) = \tilde{H}(p; \xi_1, \xi_2) = \int dx e^{ipx} H(x + \xi_1 + \xi_2, x + \xi_1 - \xi_2, x - 2\xi_1),$$

one can get after some calculations

$$g^4 W[H] = \lambda^2 C \int \frac{dp}{(2\pi)^4} \iint d\xi \iint d\xi' \tilde{H}^+(p; \xi) \Sigma(p; \xi, \xi') \tilde{H}(p; \xi'),$$

$$\Sigma(p; \xi, \xi') = e^{-\frac{p^2}{3\Lambda^2}} \cdot K_1(\xi_1, \xi'_1) \cdot K_2(\xi_2, \xi'_2),$$

where C is a constant and

$$K_j(\xi_j, \xi'_j) = e^{-\frac{c_j \lambda^2}{4} [3\xi_j^2 - 4\xi_j \xi'_j + \xi_j'^2]} = \sum_{Q_j} U_{Q_j}^{(j)}(\xi_j) \kappa_{Q_j}^{(j)} U_{Q_j}^{(j)}(\xi'_j), \quad (j = 1, 2),$$

with $c_1 = 3$ and $c_2 = 1$, and

$$\kappa_{Q_j}^{(j)} = \left(\frac{4\pi}{c_j \Lambda^2 (3 + \sqrt{5})} \right)^2 \left(\frac{2}{3 + \sqrt{5}} \right)^{2n_j + l_j}.$$

Thus, we get

$$(H_{123}^+ H_{123}) = \int \frac{dp}{(2\pi)^4} \sum_{Q_1 Q_2} \tilde{H}_{Q_1 Q_2}^+(p) \tilde{H}_{Q_1 Q_2}(p),$$

$$g^4 W[H] = \int \frac{dp}{(2\pi)^4} \sum_{Q_1 Q_2} \tilde{H}_{Q_1 Q_2}^+(p) \Sigma_{Q_1+Q_2}(p^2) \tilde{H}_{Q_1 Q_2}(p),$$

where

$$\tilde{H}_{Q_1 Q_2}(p) = \iint d\xi_1 d\xi_2 U_{Q_1}^{(1)}(\xi_1) U_{Q_2}^{(2)}(\xi_2) \tilde{H}(p; \xi_1, \xi_2),$$

$$\Sigma_{Q_1+Q_2}(p^2) = \Sigma_{nl} = e^{-\frac{p^2}{3\Lambda^2}} \cdot \left(\frac{\lambda}{\lambda_{3c}} \right)^2 \cdot \left(\frac{2}{3 + \sqrt{5}} \right)^{2(n_1+n_2)+(l_1+l_2)},$$

$$\lambda_{3c}^2 = \frac{3(3 + \sqrt{5})^4}{2^5 \epsilon^3}, \quad n = n_1 + n_2, \quad l = l_1 + l_2.$$

The Gaussian quadratic measure over the fields $\tilde{\mathcal{H}}_{Q_1 Q_2}(p) = \tilde{\mathcal{H}}_Q(p)$ looks like

$$\prod_Q \delta \tilde{\mathcal{H}}_Q \tilde{\mathcal{H}}_Q^+ \exp \left\{ - \int \frac{dp}{(2\pi)^4} \sum_Q \tilde{\mathcal{H}}_Q^+(p) [1 - \Sigma_{nl}(p^2)] \tilde{\mathcal{H}}_Q(p) \right\}. \quad (14)$$

Here $Q = \{Q_1, Q_2\}$ is the condensed index and $n = n_1 + n_2$ and $l = l_1 + l_2$.

The equation

$$1 = \Sigma_{nl}(-\mathcal{M}_{nl}^2) \quad (15)$$

defines the mass $\mathcal{M}_Q = \mathcal{M}_{nl}$ of three-particle bound states with quantum numbers $Q = nl\{\mu\}$. The spectrum has a purely linear Regge behavior:

$$\mathcal{M}_{nl}^2 = \Lambda^2 3 \ln \frac{\lambda_{3c}}{\lambda} + (2n + l) \Lambda^2 3 \ln \left(\frac{3 + \sqrt{5}}{2} \right). \quad (16)$$

We observe that the states $\tilde{\mathcal{H}}_Q(p) = \tilde{\mathcal{H}}_{Q_1 Q_2}(p)$ are degenerated, so that the states with fixed sums $n = n_1 + n_2$ and $l = l_1 + l_2$ have the same mass.

Now we can represent

$$1 - \Sigma_{nl}(p^2) = Z_{nl}(p^2 + \mathcal{M}_Q^2) + O((p^2 + \mathcal{M}_Q^2)^2),$$

and after the renormalization,

$$\tilde{\mathcal{H}}_Q(p) \rightarrow Z_Q^{-1/2} \tilde{\mathcal{H}}_Q(p)$$

has the standard form.

5 Conclusion

Finally, we can conclude that the analytic confinement leads to a reasonable picture of bound states. Bound states exist for small coupling constants $\lambda < \lambda_c < 1$, and their masses grow when $\lambda \rightarrow 0$ as

$$M_Q \sim \Lambda \sqrt{\ln \frac{\lambda_c}{\lambda}}.$$

This means that the size of the confinement region $r_{\text{cnf}} \sim 1/\Lambda$ exceeds remarkably the Compton length of all bound states for small λ . In other words, the physical particles described by the fields $B_Q(x)$ and $H_Q(x)$ and all physical transformations with them take place in the confinement region.

Analytic confinement leads to the purely linear Regge spectrum for all two- and three-particle bound states with quantum numbers $Q = (nl)$, and the slope of the Regge trajectories is defined by the scale of the confinement region Λ and does not depend on the coupling constant λ . Besides, the slopes of two- and three-particle Regge trajectories are close to each other

$$M_{nl}^2 \sim (2n + l) \cdot \Lambda^2 \cdot 2.633\dots, \quad \mathcal{M}_{nl}^2 \sim (2n + l) \cdot \Lambda^2 \cdot 2.887\dots$$

This idea can be used in QCD. Let us assume that gluon vacuum QCD is realized by the selfdual homogeneous classical field. This assumption leads to analytic confinement of quarks and gluons, and we get the linear Regge spectrum for quark and glueball bound states [7,8,12].

Appendix

Let us consider the kernel

$$K = K(y_1, y_2) = e^{-ay_1^2 + 2by_1y_2 - ay_2^2}, \quad a > b, \quad (17)$$

with

$$\text{Tr } K = \int dy K(y, y) = \int dy e^{-2(a-b)y^2} = \frac{\pi^2}{4(a-b)^2} < \infty.$$

The eigenfunctions of the problem

$$KU_Q = \kappa_Q U_Q, \quad Q = \{nl\{\mu\}\} = \{nl\{\mu_1 \dots \mu_l\}\}$$

or

$$\int dy_2 K(y_1, y_2) U_Q(y_2) = \kappa_Q U_Q(y_1)$$

are the following

$$U_Q = U_{nl\{\mu\}}(y) = N_{nl} T_{l\{\mu\}}(n_y) (y^2)^{\frac{l}{2}} L_n^{(l+1)}(2\beta y^2) e^{-\beta y^2}. \quad (18)$$

Here, $n_y = y/\sqrt{y^2}$ and

$$\beta = \sqrt{a^2 - b^2}, \quad N_{nl} = \frac{\sqrt{2^l(l+1)}}{\pi} (2\beta)^{1+\frac{l}{2}} \sqrt{\frac{\Gamma(n+1)}{\Gamma(n+l+2)}}.$$

The functions $T_{l\{\mu\}}(n)$ satisfy the conditions

$$T_{l\{\mu_1 \mu_2 \dots \mu_l\}}(n) = T_{l\{\mu_2 \mu_1 \dots \mu_l\}}(n), \quad T_{l\{\mu \mu \mu_3 \dots \mu_l\}}(n) = 0,$$

$$\sum_{\{\mu\}} T_{l\{\mu\}}(n_1) T_{l\{\mu\}}(n_2) = \frac{1}{2^l} C_l^1((n_1 n_2)), \quad C_l^1(1) = l+1,$$

where $C_l^1(t)$ are the Gegenbauer polynomials, and

$$\int dn T_{l\{\mu\}}(n) T_{l'\{\mu'\}}(n) = \delta_{ll'} \delta_{\{\mu\}\{\mu'\}} \frac{2\pi^2}{2^l(l+1)}.$$

The eigenvalues are

$$\kappa_Q = \kappa_{nl} = \kappa_0 \cdot \left(\frac{b}{a + \sqrt{a^2 - b^2}} \right)^{2n+l}, \quad \kappa_0 = \frac{\pi^2}{(a + \sqrt{a^2 - b^2})^2}. \quad (19)$$

References

- [1] P.D.B. Collins and E.J. Squires, *Regge Poles in Particle Physics* (Springer-Verlag, Berlin, 1968).
- [2] A. Tang and J.W. Norbury, *Phys. Rev. D* **62**, 016006 (2000).
- [3] S. Godfrey and N. Isgur, *Phys. Rev. D* **32**, 189 (1985).
- [4] D. Kahana, K. Maung Maung, and J.W. Norbury, *Phys. Rev. D* **48**, 3408 (1993).
- [5] E. Di Salvo, L. Kondratyuk, and P. Saracco, *Z. Phys. C* **69**, 149 (1995).

-
- [6] W.K. Tang, *Phys. Rev. D* **48**, 2019 (1993).
 - [7] G.V. Efimov and S.N. Nedelko, *Phys. Rev. D* **51**, 174 (1995); *Eur. Phys. J. C* **1**, 343 (1998).
 - [8] Ja.V. Burdanov, G.V. Efimov, S.N. Nedelko, and S.A. Solunin, *Phys. Rev. D* **54**, 4483 (1996).
 - [9] G.V. Efimov, A.C. Kalloniatis, and S.N. Nedelko, *Phys. Rev. D* **59**, 014026 (1999).
 - [10] G.V. Efimov, eprint: hep-ph/9907483.
 - [11] G.V. Efimov and M.A. Ivanov, *The Quark Confinement Model of Hadrons* (IOP Publishing, London, 1993).
 - [12] Ja.V. Burdanov and G.V. Efimov, eprint: hep-ph/0009027.

RECURSIVE GRAPHICAL CONSTRUCTION OF TADPOLE-FREE FEYNMAN DIAGRAMS AND THEIR WEIGHTS IN ϕ^4 -THEORY

A. PELSTER AND K. GLAUM

*Institut für Theoretische Physik, Freie Universität Berlin,
Arnimallee 14, D-14195 Berlin, Germany*

E-mails: pelster@physik.fu-berlin.de, glaum@physik.fu-berlin.de

We review different approaches to the graphical generation of the tadpole-free Feynman diagrams of the self-energy and the one-particle irreducible four-point function. These are needed for calculating the critical exponents of the Euclidean multicomponent scalar ϕ^4 -theory with renormalization techniques in $d = 4 - \epsilon$ dimensions.

1 Introduction

In 1982 Hagen Kleinert proposed a program for systematically constructing all Feynman diagrams of a field theory together with their proper weights by graphically solving a set of functional differential equations [1]. It relies on considering a Feynman diagram as a functional of its graphical elements, i.e. its lines and vertices. Functional derivatives with respect to these graphical elements are represented by removing lines or vertices of a Feynman diagram in all possible ways. With these graphical operations, the program proceeds in four steps. First, a nonlinear functional differential equation for the free energy is derived as a consequence of the field equations. Subsequently, this functional differential equation is converted to a recursion relation for the loop expansion coefficients of the free energy. From its graphical solution, the connected vacuum diagrams are constructed. Finally, all diagrams of n -point functions are obtained by removing lines or vertices from the connected vacuum diagrams. This program was recently used to systematically generate all

Feynman diagrams of QED [2] and of ϕ^4 -theory in the disordered, symmetric phase [3] and the ordered, broken-symmetry phase [4,5].

The present article reviews a modification of this program [6]. The aim is to construct directly the Feynman diagrams of n -point functions which are relevant for the renormalization of a field theory. To this end we consider the self-energy and the one-particle irreducible four-point function of the Euclidean multicomponent scalar ϕ^4 -theory as functionals of the free correlation function. As such they obey a closed set of functional differential equations which can be turned into graphical recursion relations. These are solved order by order in the number of loops, producing all one-particle irreducible diagrams with their proper weights. A subsequent absorption of all tadpole corrections in the lines leads to modified graphical recursion relations for the tadpole-free one-particle irreducible diagrams which are needed for calculating the critical exponents with renormalization techniques in $d = 4 - \epsilon$ dimensions. Finally, we elucidate how our procedure is related to the method of higher functional Legendre transformations which was also investigated in Ref. [1].

2 Scalar ϕ^4 -Theory

Consider a self-interacting scalar field ϕ with N components in d Euclidean dimensions whose thermal fluctuations are controlled by the energy functional

$$E[\phi] = \frac{1}{2} \int_{12} G_{12}^{-1} \phi_1 \phi_2 + \frac{1}{4!} \int_{1234} V_{1234} \phi_1 \phi_2 \phi_3 \phi_4. \quad (1)$$

In this short-hand notation, the spatial arguments and tensor indices of the field ϕ , the bilocal kernel G^{-1} , and the quartic interaction V are indicated by simple number indices, i.e.

$$1 \equiv \{x_1, \alpha_1\}, \quad \int_1 \equiv \sum_{\alpha_1} \int d^d x_1, \quad \phi_1 \equiv \phi_{\alpha_1}(x_1), \\ G_{12}^{-1} \equiv G_{\alpha_1, \alpha_2}^{-1}(x_1, x_2), \quad V_{1234} \equiv V_{\alpha_1, \alpha_2, \alpha_3, \alpha_4}(x_1, x_2, x_3, x_4). \quad (2)$$

The kernel is a functional matrix G^{-1} , while the interaction V is a functional tensor, both being symmetric in their respective indices. The energy functional (1) describes d -dimensional Euclidean ϕ^4 -theories generically. These are models for a family of universality classes of continuous phase transitions, such as the $O(N)$ -symmetric ϕ^4 -theory which serves to derive the critical

phenomena in dilute polymer solutions ($N = 0$), Ising- and Heisenberg-like magnets ($N = 1, 3$), and superfluids ($N = 2$). In all these cases, the energy functional (1) is specified by the bilocal kernel

$$G_{\alpha_1, \alpha_2}^{-1}(x_1, x_2) = \delta_{\alpha_1, \alpha_2} (-\partial_{x_1}^2 + m^2) \delta(x_1 - x_2), \quad (3)$$

and by the quartic interaction

$$V_{\alpha_1, \alpha_2, \alpha_3, \alpha_4}(x_1, x_2, x_3, x_4) = \frac{g}{3} \{ \delta_{\alpha_1, \alpha_2} \delta_{\alpha_3, \alpha_4} + \delta_{\alpha_1, \alpha_3} \delta_{\alpha_2, \alpha_4} + \delta_{\alpha_1, \alpha_4} \delta_{\alpha_2, \alpha_3} \} \\ \times \delta(x_1 - x_2) \delta(x_1 - x_3) \delta(x_1 - x_4), \quad (4)$$

where the bare mass m^2 is proportional to the temperature distance from the critical point, and g denotes the bare coupling constant. In this article we leave the kernel G^{-1} in the energy functional (1) completely general, except for the symmetry with respect to its indices, and insert the physical value (3) only at the end. By doing so, we consider all statistical quantities derived from (1) as functionals of the free correlation function G which is the functional inverse of the kernel G^{-1} :

$$\int_2 G_{12} G_{23}^{-1} = \delta_{13}. \quad (5)$$

This allows us to introduce functional derivatives with respect to G whose basic rule reflects the symmetry of its indices:

$$\frac{\delta G_{12}}{\delta G_{34}} = \frac{1}{2} \{ \delta_{13} \delta_{42} + \delta_{14} \delta_{32} \}. \quad (6)$$

Such functional derivatives are represented graphically by removing one line from a Feynman diagram in all possible ways [1-5]. Thereby each line in a Feynman diagram represents a free correlation function

$$1 \text{ --- } 2 \equiv G_{12}, \quad (7)$$

and each vertex represents an integral over the interaction

$$\begin{array}{c} \diagup \quad \diagdown \\ \times \end{array} \equiv - \int_{1234} V_{1234}. \quad (8)$$

Thus the differentiation (6) is illustrated graphically as

$$\frac{\delta}{\delta G_{34}} 1 \text{ --- } 2 = \frac{1}{2} \left\{ 1 \text{ --- } 3 \text{ --- } 4 \text{ --- } 2 + 1 \text{ --- } 4 \text{ --- } 3 \text{ --- } 2 \right\}, \quad (9)$$

where the elements of Feynman diagrams are extended by an open dot with two labeled line ends representing the delta function:

$$1 \dashv\dashv 2 = \delta_{12} . \quad (10)$$

3 Connected Vacuum Diagrams

By using natural units in which the Boltzmann constant k_B times the temperature T equals unity, the partition function is determined as a functional integral over the Boltzmann weight $e^{-E[\phi]}$, i.e.

$$Z = \int \mathcal{D}\phi e^{-E[\phi]} , \quad (11)$$

and may be evaluated perturbatively as a power series in the interaction V . From this we obtain the negative free energy $W = \ln Z$ as an expansion

$$W = \sum_{l=1}^{\infty} W^{(l)} , \quad (12)$$

where the coefficients $W^{(l)}$ for each loop order $l \geq 2$ may be displayed as connected vacuum diagrams constructed from the lines (7) and the vertices (8). As has been elaborated in detail in Ref. [3], the connected vacuum diagrams contributing to $W^{(l)}$ follow together with their weights from a graphical recursion relation which can be written diagrammatically for $l \geq 2$ as

$$\begin{aligned} W^{(l+1)} = & \frac{1}{6l} \frac{\delta^2 W^{(l)}}{\delta_1 \text{---} 2 \delta_3 \text{---} 4} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \diagup \\ \diagdown \\ \diagup \\ \diagdown \end{array} + \frac{1}{2l} \frac{\delta W^{(l)}}{\delta_1 \text{---} 2} \begin{array}{c} 1 \\ 2 \end{array} \begin{array}{c} \diagup \\ \diagdown \end{array} \bigcirc \\ & + \frac{1}{6l} \sum_{k=1}^{l-2} \frac{\delta W^{(l-k)}}{\delta_1 \text{---} 2} \begin{array}{c} 1 \\ 2 \end{array} \begin{array}{c} \diagup \\ \diagdown \end{array} \begin{array}{c} 3 \\ 4 \end{array} \frac{\delta W^{(k+1)}}{\delta_3 \text{---} 4} . \end{aligned} \quad (13)$$

This is iterated starting from the two-loop contribution

$$W^{(2)} = \frac{1}{8} \bigcirc \bigcirc . \quad (14)$$

The right-hand side of (13) contains three different graphical operations. The first two are linear and involve one- or two-line amputations of the previous

perturbative order. The third operation is nonlinear and mixes two different one-line amputations of lower orders.^a

The connected vacuum diagrams resulting from the graphical recursion relation (13) together with their weights are shown up to $l = 5$ loops in Ref. [3]. There we observed that the nonlinear operation in (13) does not lead to topologically new diagrams. It only corrects the weights generated from the first two linear operations. Continuing the solution of the graphical recursion relation (13) to higher loops is an arduous task. We have therefore automatized the procedure in Ref. [3] by computer algebra with the help of a unique matrix notation for Feynman diagrams. The corresponding MATHEMATICA program and higher-order results up to $l = 7$ are available on the internet [7].

4 One-Particle Irreducible Diagrams

Once the connected vacuum diagrams are known, the diagrams of the fully-interacting two-point function

$$\mathbf{G}_{12} = \frac{1}{Z} \int \mathcal{D}\phi \phi_1 \phi_2 e^{-E[\phi]} \quad (15)$$

and the connected four-point function

$$\mathbf{G}_{1234}^c = \frac{1}{Z} \int \mathcal{D}\phi \phi_1 \phi_2 \phi_3 \phi_4 e^{-E[\phi]} - \mathbf{G}_{12} \mathbf{G}_{34} - \mathbf{G}_{13} \mathbf{G}_{24} - \mathbf{G}_{14} \mathbf{G}_{23} \quad (16)$$

follow from removing one or two lines in all possible ways, respectively [3,7]:

$$\mathbf{G}_{12} = 2 \int_{34} G_{13} G_{24} \frac{\delta W}{\delta G_{34}}, \quad (17)$$

$$\mathbf{G}_{1234}^c = 2 \int_{56} G_{35} G_{46} \frac{\delta \mathbf{G}_{12}}{\delta G_{56}} - \mathbf{G}_{13} \mathbf{G}_{24} - \mathbf{G}_{14} \mathbf{G}_{23}. \quad (18)$$

Note that removing a line from the diagrams of $\mathbf{G}_{12}$ according to the first term of (18) leads to disconnected diagrams which are cancelled by the second and the third term to yield the connected diagrams contributing to $\mathbf{G}_{1234}^c$. Dropping all diagrams of $\mathbf{G}_{12}$ and $\mathbf{G}_{1234}^c$ which would fall into two pieces

^aNote that the first two operations in (13) were already used in Ref. [8] as heuristic algorithms to generate all topological different connected vacuum diagrams of the ϕ^4 -theory up to $l = 8$ loops with a computer program. Their corresponding weights were then determined by combinatorial means with a second computer program.

by removing a line, one obtains the one-particle irreducible diagrams of the self-energy

$$\Sigma_{12} = G_{12}^{-1} - \mathbf{G}_{12}^{-1}, \quad (19)$$

with $\mathbf{G}^{-1}$ being the functional inverse of $\mathbf{G}$

$$\int_2 \mathbf{G}_{12} \mathbf{G}_{23}^{-1} = \delta_{13}, \quad (20)$$

and the one-particle irreducible four-point function

$$\Gamma_{1234} = - \int_{5678} \mathbf{G}_{15}^{-1} \mathbf{G}_{26}^{-1} \mathbf{G}_{37}^{-1} \mathbf{G}_{48}^{-1} \mathbf{G}_{5678}^c. \quad (21)$$

Along these lines all diagrams and their weights were constructed which are relevant for the five-loop renormalization of the ϕ^4 -theory in $d = 4 - \epsilon$ dimensions [8-10].

At this stage the question arises whether the one-particle irreducible diagrams can also be generated in a direct graphical way without any reference to the connected vacuum diagrams. This is indeed possible and will be elaborated in detail in Ref. [6]. Here we restrict ourselves to present some of the results. To this end we extend the elements of Feynman diagrams by a double straight line representing the fully-interacting two-point function

$$1 \text{ --- } 2 = \mathbf{G}_{12}, \quad (22)$$

and by a 2- and 4-vertex with a big open dot representing the self-energy and the one-particle irreducible four-point function

$$1 \text{ --- } \bigcirc \text{ --- } 2 = \Sigma_{12}, \quad (23)$$

$$\begin{array}{c} 2 \quad 3 \\ \diagdown \quad \diagup \\ \bigcirc \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} = \Gamma_{1234}, \quad (24)$$

respectively. It turns out that the self-energy Σ follows from an integral equation which reads graphically

$$1 \text{ --- } \bigcirc \text{ --- } 2 = \frac{1}{2} \text{ } 1 \text{ --- } \bigcirc \text{ --- } 2 + \frac{1}{6} \text{ } 1 \text{ --- } \bigcirc \text{ --- } 2. \quad (25)$$

Thereby the connected two-point function $\mathbf{G}$ is obtained from (19) according to the Dyson equation

$$1 \text{ --- } 2 = 1 \text{ --- } 2 + 1 \text{ --- } \bigcirc \text{ --- } 2. \quad (26)$$

Solving the Eqs. (25) and (26) iteratively necessitates the knowledge of the one-particle irreducible four-point function Γ . Its diagrams could be determined from

$$\begin{aligned} \text{Diagram 1} &= 2 \frac{\delta \text{Diagram 2}}{\delta 3 \text{ --- } 4} + 2 \text{Diagram 3} \frac{\delta \text{Diagram 4}}{\delta 5 \text{ --- } 6} \text{Diagram 5} \\ &- 2 \text{Diagram 3} \frac{\delta \text{Diagram 4}}{\delta 5 \text{ --- } 4} - 2 \frac{\delta \text{Diagram 4}}{\delta 3 \text{ --- } 6} \text{Diagram 5}. \end{aligned} \quad (27)$$

However, such a procedure would have one disadvantage: removing a line in the diagrams of the self-energy Σ also leads to one-particle reducible diagrams which are later on cancelled by the third and the fourth term in (27). As the number of undesired one-particle reducible diagrams occurring at an intermediate step of the calculation increases with the perturbative order, this procedure is quite inefficient in determining the diagrams of the one-particle irreducible four-point function Γ . By inserting (25) in (27) it turns out that we can derive another equation for Γ whose iterative solution only involves one-particle irreducible diagrams:

$$\begin{aligned} \text{Diagram 1} &= \text{Diagram 2} + \frac{1}{3} \text{Diagram 3} \frac{\delta \text{Diagram 4}}{\delta 6 \text{ --- } 7} \\ &+ \frac{1}{2} \text{Diagram 5} + \frac{1}{2} \text{Diagram 6} + \frac{1}{2} \text{Diagram 7} \\ &+ \frac{1}{6} \text{Diagram 8} + \frac{1}{6} \text{Diagram 9} + \frac{1}{6} \text{Diagram 10}. \end{aligned} \quad (28)$$

Note that Eqs. (25), (26), and (28) only generate one-particle irreducible diagrams for Σ and Γ , if all lower loop orders contain one-particle irreducible diagrams. By induction this establishes that Σ and Γ only consist of one-particle irreducible diagrams.

5 Tadpole-Free One-Particle Irreducible Diagrams

In order to reduce the number of diagrams, we aim at substituting the free correlation function G by a modified one $\tilde{G}$. If we would fix $\tilde{G}$ according to

$$\tilde{G}_{12}^{-1} = G_{12}^{-1} + \frac{1}{2} \int_{34} V_{1234} \tilde{G}_{34}, \quad (29)$$

it would contain all repetitive one-loop corrections. This method was established in Ref. [11] to get rid of all one-particle irreducible diagrams carrying tadpole corrections when calculating the β -function of the vacuum energy density up to five loops. Here, however, we go one step further by demanding instead of (29)

$$\tilde{G}_{12}^{-1} = G_{12}^{-1} + \frac{1}{2} \int_{34} V_{1234} G_{34}, \quad (30)$$

which amounts to absorbing all momentum-independent line corrections into the mass [12]. The tadpole corrections of the modified correlation function $\tilde{G}$ arising from (30) were already treated perturbatively in Ref. [4]. It was shown that they lead to additional diagrams which cancel order by order all diagrams of one-particle irreducible n -point functions carrying any kind of tadpole correction. In Ref. [6] we elaborate that these inefficient cancellations can be circumvented as the remaining tadpole-free one-particle irreducible diagrams directly follow from a closed set of functional differential equations. To this end we consider the modified self-energy

$$\tilde{\Sigma}_{12} = \tilde{G}_{12}^{-1} - G_{12}^{-1} \quad (31)$$

and the one-particle irreducible four-point function Γ as functionals of the modified correlation function $\tilde{G}$. Extending the elements of Feynman diagrams by a wiggly line representing the modified correlation function

$$1 \rightsquigarrow 2 = \tilde{G}_{12} \quad (32)$$

and by a 2-vertex with a wiggly dot representing the tadpole-free self-energy

$$1 \sim \text{dot} \sim 2 = \tilde{\Sigma}_{12} \quad (33)$$

our results read as follows. The tadpole-free self-energy $\tilde{\Sigma}$ obeys the integral equation

$$1 \sim \text{dot} \sim 2 = \frac{1}{6} 1 \sim \text{bubble} \sim 2, \quad (34)$$

where the fully-interacting two-point function $\mathbf{G}^c$ is obtained from (31) according to the modified Dyson equation

$$1 \text{ --- } 2 = 1 \text{ ~~~~~ } 2 + 1 \text{ --- } \text{tadpole} \text{ --- } 2. \quad (35)$$

Then the one-particle irreducible four-point function Γ obeys

$$\begin{aligned} \text{diagram} &= \text{diagram} + \frac{1}{3} \text{diagram} \frac{\delta \text{diagram}}{\delta \text{diagram}} \\ &+ \frac{1}{6} \text{diagram} \frac{\delta \text{diagram}}{\delta \text{diagram}} + \frac{1}{12} \text{diagram} \frac{\delta \text{diagram}}{\delta \text{diagram}} \\ &+ \frac{1}{2} \text{diagram} + \frac{1}{2} \text{diagram} + \frac{1}{2} \text{diagram} \\ &+ \frac{1}{6} \text{diagram} + \frac{1}{6} \text{diagram} + \frac{1}{6} \text{diagram}. \end{aligned} \quad (36)$$

Iteratively solving Eqs. (34)–(36) in a graphical way leads to all tadpole-free diagrams of the self-energy and of the one-particle irreducible four-point function together with their weights. Up to five loops they are listed in Appendix A of Ref. [10]. Their respective quadratic and logarithmic divergences in $d = 4 - \epsilon$ dimensions contribute to the $1/\epsilon$ -poles of the renormalization constants of the field ϕ , the coupling constant g and the mass m^2 within the minimal subtraction scheme, so that they determine the critical exponents of scalar ϕ^4 -theory.

6 One-Particle Irreducible Diagrams without Line Corrections

The number of diagrams can be even further reduced by substituting the free correlation function G by the fully-interacting two-point function $\mathbf{G}$ itself. Such a substitution was investigated in Ref. [1] within the framework of higher functional Legendre transformations. By considering the self-energy Σ and the one-particle irreducible four-point function Γ as functionals of the fully-

interacting two-point function $\mathbf{G}$, it can be shown that they obey [1,6]

$$1 \text{---} \bigcirc \text{---} 2 = \frac{1}{2} 1 \text{---} \bigcirc \text{---} 2 + \frac{1}{6} 1 \text{---} \bigcirc \text{---} 2, \quad (37)$$

$$\begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} \begin{array}{c} 3 \\ \diagdown \quad \diagup \\ 4 \end{array} = 2 \frac{\delta 1 \text{---} \bigcirc \text{---} 2}{\delta 3 \text{---} 4} + \begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 1 \quad 6 \end{array} \begin{array}{c} 5 \\ \diagdown \quad \diagup \\ 6 \end{array} \frac{\delta 5 \text{---} \bigcirc \text{---} 6}{\delta 3 \text{---} 4}. \quad (38)$$

Inserting (37) in (38) it turns out that the one-particle irreducible four-point function also follows from [6]

$$\begin{aligned} \begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} \begin{array}{c} 3 \\ \diagdown \quad \diagup \\ 4 \end{array} &= \begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} \begin{array}{c} 3 \\ \diagdown \quad \diagup \\ 4 \end{array} + \frac{1}{3} 1 \text{---} \bigcirc \text{---} 2 \frac{\delta 5 \text{---} \begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 5 \quad 4 \end{array} \begin{array}{c} 3 \\ \diagdown \quad \diagup \\ 4 \end{array}}{\delta 6 \text{---} 7} + \frac{1}{6} 1 \text{---} \bigcirc \text{---} 2 \frac{\delta 5 \text{---} \begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 5 \quad 4 \end{array} \begin{array}{c} 3 \\ \diagdown \quad \diagup \\ 4 \end{array}}{\delta 6 \text{---} 7} \\ &+ \frac{1}{2} \begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} \begin{array}{c} 3 \\ \diagdown \quad \diagup \\ 4 \end{array} + \frac{1}{2} \begin{array}{c} 3 \\ \diagup \quad \diagdown \\ 1 \quad 2 \end{array} \begin{array}{c} 4 \\ \diagdown \quad \diagup \\ 2 \end{array} + \frac{1}{2} \begin{array}{c} 4 \\ \diagup \quad \diagdown \\ 1 \quad 3 \end{array} \begin{array}{c} 2 \\ \diagdown \quad \diagup \\ 3 \end{array} \\ &+ \frac{1}{6} \begin{array}{c} 2 \\ \diagup \quad \diagdown \\ 1 \quad 4 \end{array} \begin{array}{c} 3 \\ \diagdown \quad \diagup \\ 4 \end{array} + \frac{1}{6} \begin{array}{c} 3 \\ \diagup \quad \diagdown \\ 1 \quad 2 \end{array} \begin{array}{c} 4 \\ \diagdown \quad \diagup \\ 2 \end{array} + \frac{1}{6} \begin{array}{c} 4 \\ \diagup \quad \diagdown \\ 1 \quad 3 \end{array} \begin{array}{c} 2 \\ \diagdown \quad \diagup \\ 3 \end{array}. \quad (39) \end{aligned}$$

The graphical solution of the functional differential equations (37) and (38) or (39) leads to all one-particle irreducible diagrams which do not contain any line corrections. Once they are generated, we can recover the tadpole-free one-particle irreducible diagrams according to the following algorithm [6]. At first we subtract the one-loop correction from the self-energy in order to obtain the tadpole-free self-energy

$$1 \text{---} \text{---} 2 = 1 \text{---} \bigcirc \text{---} 2 - \frac{1}{2} 1 \text{---} \bigcirc \text{---} 2. \quad (40)$$

Iterating the modified Dyson equation (35), we then determine the tadpole-free diagrams of the fully-interacting two-point function $\mathbf{G}$ where the lines represent the modified correlation function (32). Finally we insert this expansion into the one-particle irreducible diagrams of Σ and Γ determined from (37) and (38) or (39).

7 Summary

We have reviewed different approaches to graphically generate tadpole-free one-particle irreducible diagrams together with their weights which are needed for calculating the critical exponents of ϕ^4 -theory. One approach is based on the graphical solution of Eqs. (34)–(36), the latter being more complex. The alternative approach consists of the simpler Eqs. (37) and (38) at the expense of subsequent iterations of the modified Dyson equation (35) with (40). In order to decide which of those approaches is more efficient, it is necessary to perform further analytic studies or to automatize them by computer algebra.

Acknowledgments

A.P. is deeply indebted to Professor Dr. Hagen Kleinert for the opportunity to work with him as a scientific assistant. His universal physical knowledge, his many brilliant ideas and his truly encouraging personality cause a thrilling scientific environment in his research group. K.G. thanks Professor Kleinert for supervising the research for his diploma thesis, which led to the derivation of the results of this article were derived. Both of us wish him a happy 60th birthday. Finally we thank Michael Bachmann, Dr. Boris Kastening, and Dr. Bruno van den Bossche for their interest in the recursive graphical construction of Feynman diagrams.

References

- [1] H. Kleinert, *Fortschr. Phys.* **30**, 187 and 351 (1982).
- [2] M. Bachmann, H. Kleinert, and A. Pelster, *Phys. Rev. D* **61**, 085017 (2000).
- [3] H. Kleinert, A. Pelster, B. Kastening, and M. Bachmann, *Phys. Rev. E* **62**, 1537 (2000).
- [4] B. Kastening, *Phys. Rev. E* **61**, 3501 (2000).
- [5] A. Pelster and H. Kleinert, eprint: hep-th/0006153.
- [6] A. Pelster, H. Kleinert, and K. Glaum, in preparation.
- [7] <http://www.physik.fu-berlin.de/~kleinert/294/programs>.
- [8] J. Neu, Diploma Thesis (in German), FU-Berlin (1990).
- [9] H. Kleinert, J. Neu, V. Schulte-Frohlinde, K.G. Chetyrkin, and S.A. Larin, *Phys. Lett. B* **272**, 39 (1991); *Phys. Lett. B* **319**, 545(E) (1993).

- [10] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [11] B. Kastening, *Phys. Rev. D* **54**, 3965 (1996); *Phys. Rev. D* **57**, 3567 (1998).
- [12] J. Küster and G. Münster, *Z. Phys. C* **73**, 551 (1997).

RECURSIVE CONSTRUCTION OF FEYNMAN GRAPHS IN SPONTANEOUSLY BROKEN $O(N)$ -SYMMETRIC ϕ^4 -THEORY

B. KASTENING

*Institut für Theoretische Physik, Freie Universität Berlin, Arnimallee 14,
D-14195 Berlin, Germany*

E-mail: ka@physik.fu-berlin.de

We consider $O(N)$ -symmetric ϕ^4 -theory in its spontaneously broken phase and investigate how the corresponding one-particle irreducible Feynman graphs with arbitrary numbers of external legs can recursively be constructed. In particular, we sketch the derivation of the necessary identities for the effective energy Γ (or, with trivial modifications, effective action), which subsequently can be converted into recursion relation for graphs. Although the resulting relations will not be as concise as their counterpart in the $N = 1$ case, they nevertheless should be a useful means for graph generation once implemented on a computer.

First let me note that it is a pleasure for me to contribute to the anniversary edition in honor of Prof. Hagen Kleinert. Let this be a small but original contribution to one of the many areas of theoretical physics he is and he has been working on through the course of the years. The article can be viewed as a direct formalized consequence of his work in Refs. [1,2].

The $O(N)$ -symmetric ϕ^4 -theory is a field theoretic model which serves not only as a toy model in particle physics, but also covers an important subset of universality classes in the context of critical phenomena. While several universal quantities can be derived upon working in the symmetric (or disordered) phase, there are a number of amplitude ratios whose calculation necessitates to consider the spontaneously broken (or ordered) phase. This results in the context of perturbation theory in a vast increase of diagrams to be considered even in low loop orders. Part of this increase is due to three-point vertices arising in the ordered phase, which also happens already in the

$N = 1$ case. The other part of the increase is due to the presence of two different propagators instead of only one in the $N = 1$ case. For a classic reference, see Ref. [3]. A more recent reference is Ref. [4], which also contains a list of related references.

In the context of critical phenomena (and not only there), we are mainly interested in one-particle irreducible (1PI) diagrams. There are several ways to obtain these diagrams for $N > 1$ (for $N = 1$, see Ref. [5] or, for an alternative approach, see Ref. [6]). Among them are:

- (1) Use the appropriate Feynman rules.
- (2) First derive the graphs for single-component ϕ^4 -theory, then replace each propagator by a sum of Higgs and Goldstone propagators and throw away those graphs that are not permitted.
- (3) Write down recursion relations for the connected graphs and throw away all non-1PI graphs.
- (4) Introduce a mixed propagator which is Higgs at one end and Goldstone at the other and translate the recursion relations for the single-component theory into recursion relations for the case at hand. Set the mixed propagator to zero in the resulting diagrams, so only permitted diagrams survive.
- (5) Develop recursion relations for the broken-symmetry 1PI graphs.

Strategy 1 is fraught with the danger of making mistakes, both by hand or when programming. Strategies 2-4 produce a huge number of diagrams at intermediate stages that are going to be dropped eventually. Strategy 5 is the one suggested here. It certainly results in rather elaborate equations, as we will see below. However, once everything is automated, it appears to be superior to the other methods. This is especially true if even the generation of the necessary equations is automated, a task left for future work, since in this brief report, we present equations derived "by hand".

Straightforwardly generalizing earlier work in Refs. [5,7] (we use conventions introduced there), let the energy be given by

$$\begin{aligned}
 E[\phi, \chi, J, I, H, G] = & C + \int_1 (J_1 - \hat{J}_1) \phi_1 + \int_1 (I_1 - \hat{I}_1) \chi_1 \\
 & + \frac{1}{2} \int_{12} H_{12}^{-1} \phi_1 \phi_2 + \frac{1}{2} \int_{12} G_{12}^{-1} \chi_1 \chi_2 + \frac{1}{6} \int_{123} S_{123} \phi_1 \phi_2 \phi_3 + \frac{1}{2} \int_{123} T_{123} \phi_1 \chi_2 \chi_3 \\
 & + \frac{1}{24} \int_{1234} L_{1234} \phi_1 \phi_2 \phi_3 \phi_4 + \frac{1}{4} \int_{1234} M_{1234} \phi_1 \phi_2 \chi_3 \chi_4 + \frac{1}{24} \int_{1234} N_{1234} \chi_1 \chi_2 \chi_3 \chi_4
 \end{aligned} \tag{1}$$

with appropriately symmetrized bare propagators H and G and interactions S , T , L , M , and N . The indices represent real-space arguments as well as group indices, while the integration signs stand not only for space integrations, but also for summations over group indices. The currents $\hat{J}$ and $\hat{I}$ will be used to define Legendre transforms while keeping the option of having non-zero currents J and I in the effective energy.

Define the functionals Z and W by

$$Z[\hat{J}, \hat{I}, H, G] = \exp W[\hat{J}, \hat{I}, H, G] = \int D\phi D\chi \exp(-E[\phi, \chi, \hat{J}, \hat{I}, H, G]). \quad (2)$$

There is an infinite set of identities we can derive for W . A few of the simplest ones appear particularly useful for our purposes. They are

$$\begin{aligned} 0 &= e^{-W[\hat{J}, \hat{I}, H, G]} \int D\phi D\chi \frac{\delta}{\delta\phi_1} e^{-E[\phi, \chi, \hat{J}, \hat{I}, H, G]} \\ &= \hat{J}_1 - J_1 - \int_2 H_{12}^{-1} \frac{\delta W}{\delta\hat{J}_2} + \int_{23} S_{123} \frac{\delta W}{\delta H_{23}^{-1}} + \int_{23} T_{123} \frac{\delta W}{\delta G_{23}^{-1}} \\ &\quad + \frac{1}{3} \int_{234} L_{1234} \left(\frac{\delta^2 W}{\delta\hat{J}_2 \delta H_{34}^{-1}} + \frac{\delta W}{\delta\hat{J}_2} \frac{\delta W}{\delta H_{34}^{-1}} \right) \\ &\quad + \int_{234} M_{1234} \left(\frac{\delta^2 W}{\delta\hat{J}_2 \delta G_{34}^{-1}} + \frac{\delta W}{\delta\hat{J}_2} \frac{\delta W}{\delta G_{34}^{-1}} \right), \end{aligned} \quad (3)$$

$$\begin{aligned} 0 &= e^{-W[\hat{J}, \hat{I}, H, G]} \int D\phi D\chi \frac{\delta}{\delta\phi_1} \left(\phi_2 e^{-E[\phi, \chi, \hat{J}, \hat{I}, H, G]} \right) \\ &= \delta_{12} + (\hat{J}_1 - J_1) \frac{\delta W}{\delta\hat{J}_2} + 2 \int_3 H_{13}^{-1} \frac{\delta W}{\delta H_{23}^{-1}} \\ &\quad + \int_{34} S_{134} \left(\frac{\delta^2 W}{\delta\hat{J}_2 \delta H_{34}^{-1}} + \frac{\delta W}{\delta\hat{J}_2} \frac{\delta W}{\delta H_{34}^{-1}} \right) \\ &\quad + \int_{34} T_{134} \left(\frac{\delta^2 W}{\delta\hat{J}_2 \delta G_{34}^{-1}} + \frac{\delta W}{\delta\hat{J}_2} \frac{\delta W}{\delta G_{34}^{-1}} \right) \\ &\quad - \frac{2}{3} \int_{345} L_{1345} \left(\frac{\delta^2 W}{\delta H_{23}^{-1} \delta H_{45}^{-1}} + \frac{\delta W}{\delta H_{23}^{-1}} \frac{\delta W}{\delta H_{45}^{-1}} \right) \\ &\quad - 2 \int_{345} M_{1345} \left(\frac{\delta^2 W}{\delta H_{23}^{-1} \delta G_{45}^{-1}} + \frac{\delta W}{\delta H_{23}^{-1}} \frac{\delta W}{\delta G_{45}^{-1}} \right), \end{aligned} \quad (4)$$

$$\begin{aligned}
0 &= e^{-W[jh, \hat{I}, H, G]} \int D\phi D\chi \frac{\delta}{\delta\chi_1} \left(\chi_2 e^{-E[\phi, \chi, \hat{J}, \hat{I}, H, G]} \right) \\
&= \delta_{12} + (\hat{I}_1 - I_1) \frac{\delta W}{\delta \hat{I}_2} + 2 \int_3 G_{13}^{-1} \frac{\delta W}{\delta G_{23}^{-1}} \\
&\quad + 2 \int_{34} T_{134} \left(\frac{\delta^2 W}{\delta \hat{J}_3 \delta G_{24}^{-1}} + \frac{\delta W}{\delta \hat{J}_3} \frac{\delta W}{\delta G_{24}^{-1}} \right) \\
&\quad - 2 \int_{345} M_{4513} \left(\frac{\delta^2 W}{\delta G_{23}^{-1} \delta H_{45}^{-1}} + \frac{\delta W}{\delta G_{23}^{-1}} \frac{\delta W}{\delta H_{45}^{-1}} \right) \\
&\quad - \frac{2}{3} \int_{345} N_{1345} \left(\frac{\delta^2 W}{\delta G_{23}^{-1} \delta G_{45}^{-1}} + \frac{\delta W}{\delta G_{23}^{-1}} \frac{\delta W}{\delta G_{45}^{-1}} \right). \quad (5)
\end{aligned}$$

Next, define W_0 and W_I by

$$W_0 = W_{S,T,L,M,N=0}, \quad W = W_0 + W_I. \quad (6)$$

For W_0 , the identities (3)-(5) reduce to

$$\frac{\delta W_0}{\delta J_1} = \int_2 H_{12} J_2, \quad \frac{\delta W_0}{\delta I_1} = \int_2 G_{12} I_2, \quad (7)$$

$$\frac{\delta W_0}{\delta H_{12}^{-1}} = -\frac{1}{2} \left(H_{12} + \int_3 H_{13} J_3 \int_4 H_{24} J_4 \right), \quad (8)$$

$$\frac{\delta W_0}{\delta G_{12}^{-1}} = -\frac{1}{2} \left(G_{12} + \int_3 G_{13} I_3 \int_4 G_{24} I_4 \right). \quad (9)$$

With appropriate normalization of the path integration measure, this is solved by

$$\begin{aligned}
W_0 &= -C - \frac{1}{2} \int_1 (\ln H^{-1})_{11} - \frac{1}{2} \int_1 (\ln G^{-1})_{11} \\
&\quad + \frac{1}{2} \int_{12} H_{12} (J_1 - \hat{J}_1) (J_2 - \hat{J}_2) + \frac{1}{2} \int_{12} G_{12} (I_1 - \hat{I}_1) (I_2 - \hat{I}_2). \quad (10)
\end{aligned}$$

Now the identities (3)-(5) may be translated into identities for W_I and then into recursion relations for the connected graphs of the theory. One possibility to obtain the 1PI graphs would be to just throw away the non-1PI graphs. However, this would generate a vast number of connected, but non-1PI graphs to be thrown away later anyway and therefore will very soon be an exploding task. Instead, let us right away turn to the effective energy, the generating functional of 1PI graphs.

Define the effective energy Γ by a twofold Legendre transformation,

$$\Gamma[\sigma, \pi, H, G] = -W[\hat{J}, \hat{I}, H, G] + \int_1 \hat{J}_1 \sigma_1 + \int_1 \hat{I}_1 \pi_1, \quad (11)$$

with new independent variables

$$\sigma_1 = \left(\frac{\delta W}{\delta \hat{J}_1} \right)_{\hat{I}HG}, \quad \pi_1 = \left(\frac{\delta W}{\delta \hat{I}_1} \right)_{\hat{J}HG}. \quad (12)$$

As usual,

$$\left(\frac{\delta \Gamma}{\delta \sigma_1} \right)_{\pi HG} = \hat{J}_1, \quad \left(\frac{\delta \Gamma}{\delta \pi_1} \right)_{\sigma HG} = \hat{I}_1. \quad (13)$$

Define further Γ_0 and Γ_I by

$$\Gamma_0 = \Gamma_{S,T,L,M,N=0}, \quad \Gamma = \Gamma_0 + \Gamma_I. \quad (14)$$

With

$$\sigma_1 = \frac{\delta W_0}{\delta \hat{J}_1} = \int_2 H_{12}(\hat{J}_2 - J_2), \quad \pi_1 = \frac{\delta W_0}{\delta \hat{I}_1} = \int_2 G_{12}(\hat{I}_2 - I_2) \quad (15)$$

we have

$$\hat{J}_1 = J_1 + \int_2 H_{12}^{-1} \sigma_2, \quad \hat{I}_1 = I_1 + \int_2 G_{12}^{-1} \pi_2, \quad (16)$$

and therefore with (10)

$$\begin{aligned} \Gamma_0 = C &+ \frac{1}{2} \int_1 (\ln H^{-1})_{11} + \frac{1}{2} \int_1 (\ln G^{-1})_{11} + \int_1 J_1 \sigma_1 + \int_1 I_1 \pi_1 \\ &+ \frac{1}{2} \int_{12} H_{12}^{-1} \sigma_1 \sigma_2 + \frac{1}{2} \int_{12} G_{12}^{-1} \pi_1 \pi_2. \end{aligned} \quad (17)$$

This is the only place in Γ , where J and I still appear, in other words: Γ_I is independent of J and I .

Now we can translate (3)-(5) into identities for Γ_I , which in turn can be converted into recursion relations for Feynman graphs. The main complication arises from the fact that we now deal with a twofold Legendre transform, which destroys some of the ease we are used to from dealing with the $N = 1$ case [5], since the relations between the second derivatives of W and Γ become more complicated. Instead of translating (3)-(5) in full generality, we use for each situation the simplest identity available.

Let us for the remainder invoke a generalized Einstein summation convention, where we sum/integrate over the variables represented by repeated indices. For the generation of diagrams with at least one external field σ , we translate the simplest identity (3) and obtain

$$\begin{aligned}
0 = & \frac{\delta\Gamma_I}{\delta\sigma_1}\sigma_1 - \frac{1}{2}S_{123}\sigma_1H_{23} - \frac{1}{2}S_{123}\sigma_1\sigma_2\sigma_3 - \frac{1}{2}T_{123}\sigma_1G_{23} - \frac{1}{2}T_{123}\sigma_1\pi_2\pi_3 \\
& - \frac{1}{2}L_{1234}\sigma_1\sigma_2H_{34} - \frac{1}{6}L_{1234}\sigma_1\sigma_2\sigma_3\sigma_4 - \frac{1}{2}M_{1234}\sigma_1\sigma_2G_{34} - \frac{1}{2}M_{1234}\sigma_1\sigma_2\pi_3\pi_4 \\
& + S_{123}\sigma_1H_{24}H_{35}\frac{\delta\Gamma_I}{\delta H_{45}} + T_{123}\sigma_1G_{24}G_{35}\frac{\delta\Gamma_I}{\delta G_{45}} \\
& + L_{1234}\sigma_1\sigma_2H_{35}H_{46}\frac{\delta\Gamma_I}{\delta H_{56}} + L_{1234}\sigma_1H_{25}H_{36}H_{47}\Gamma_{657}^{\sigma H} \\
& + 2M_{1234}\sigma_1\pi_3H_{25}G_{46}\frac{\delta^2\Gamma_I}{\delta\sigma_5\delta\pi_6} + 2M_{1234}\sigma_1\pi_3H_{25}G_{46}\Gamma_{67}^G G_{78}\frac{\delta^2\Gamma_I}{\delta\sigma_5\delta\pi_8} \\
& + 2M_{1234}\sigma_1\pi_3H_{25}\frac{\delta\Gamma_I}{\delta H_{56}}H_{67}G_{48}\frac{\delta^2\Gamma_I}{\delta\sigma_7\delta\pi_8} \\
& + 2M_{1234}\sigma_1\pi_3H_{25}\frac{\delta\Gamma_I}{\delta H_{56}}H_{67}G_{48}\Gamma_{89}G_{90}\frac{\delta^2\Gamma_I}{\delta\sigma_7\delta\pi_0} \\
& + M_{1234}\sigma_1\sigma_2G_{35}G_{46}\frac{\delta\Gamma_I}{\delta G_{56}} + M_{1234}\sigma_1H_{25}G_{36}G_{47}\Gamma_{657}^{\sigma G}. \tag{18}
\end{aligned}$$

For the generation of diagrams with no Higgs fields σ and no Higgs propagator H , we may translate (5) with $T, M = 0$ and obtain

$$\begin{aligned}
0 = & \frac{\delta\Gamma_I}{\delta\pi_1}\pi_1 + 2\frac{\delta\Gamma_I}{\delta G_{12}}G_{12} - \frac{1}{6}N_{1234}\pi_1\pi_2\pi_3\pi_4 - N_{1234}G_{12}\pi_3\pi_4 - \frac{1}{2}N_{1234}G_{12}G_{34} \\
& + 2N_{123456}\pi_1\pi_2G_{35}G_{46}\frac{\delta\Gamma_I}{\delta G_{56}} + 2N_{1234}G_{12}G_{35}G_{46}\frac{\delta\Gamma_I}{\delta G_{56}} \\
& + \frac{4}{3}N_{1234}\pi_1G_{25}G_{36}G_{47}\frac{\delta^2\Gamma_I}{\delta\pi_5\delta G_{67}} + \frac{2}{3}N_{1234}G_{15}G_{26}G_{37}G_{48}\frac{\delta^2\Gamma_I}{\delta G_{56}\delta G_{78}} \\
& - \frac{2}{3}N_{1234}G_{15}G_{26}\frac{\delta\Gamma_I}{\delta G_{56}}G_{37}G_{48}\frac{\delta\Gamma_I}{\delta G_{78}} - \frac{8}{3}N_{1234}\pi_1G_{25}\frac{\delta\Gamma_I}{\delta G_{56}}G_{67}\frac{\delta^2\Gamma_I}{\delta\pi_7\delta G_{89}}G_{83}G_{94} \\
& - \frac{2}{3}N_{1234}G_{15}G_{26}\frac{\delta^2\Gamma_I}{\delta G_{56}\delta\pi_7}G_{78}\frac{\delta^2\Gamma_I}{\delta\pi_8\delta G_{90}}G_{93}G_{04} \\
& + \frac{4}{3}N_{1234}G_{15}G_{26}\frac{\delta^2\Gamma_I}{\delta G_{56}\delta\pi_7}G_{78}\frac{\delta\Gamma_I}{\delta G_{89}}G_{90}\frac{\delta^2\Gamma_I}{\delta\pi_0\delta G_{12}}G_{13}G_{24}, \tag{19}
\end{aligned}$$

where it is understood that in all terms we have set $T, M = 0$.

For the generation of all other 1PI diagrams, i.e. diagrams with no Higgs field σ , but at least one Higgs propagator H , we translate (4) with $\sigma = 0$ and obtain

$$\begin{aligned}
0 = & 2 \frac{\delta \Gamma_I}{\delta H_{12}} H_{12} + 2 T_{123} H_{14} \frac{\delta^2 \Gamma_I}{\delta \sigma_4 \delta \pi_5} G_{52} \pi_3 + 2 T_{123} H_{14} \frac{\delta^2 \Gamma_I}{\delta \sigma_4 \delta \pi_5} G_{56} \Gamma_{67}^G G_{72} \pi_3 \\
& - 4 T_{123} H_{14} \frac{\delta \Gamma_I}{\delta H_{45}} H_{56} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \delta \pi_7} G_{72} \pi_3 - 4 T_{123} H_{14} \frac{\delta \Gamma_I}{\delta H_{45}} H_{56} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \delta \pi_7} G_{78} \Gamma_{89}^G G_{92} \pi_3 \\
& + S_{123} H_{14} H_{25} H_{36} \Gamma_{456}^{\sigma H} + T_{123} H_{14} G_{25} G_{36} \Gamma_{456}^{\sigma G} - \frac{1}{2} L_{1234} H_{12} H_{34} \\
& + 2 L_{1234} H_{12} H_{35} H_{46} \frac{\delta \Gamma_I}{\delta H_{56}} - \frac{2}{3} L_{1234} H_{15} H_{26} \frac{\delta \Gamma_I}{\delta H_{56}} H_{37} H_{48} \frac{\delta \Gamma_I}{\delta H_{78}} \\
& - \frac{2}{3} L_{1234} H_{15} H_{26} H_{37} H_{48} \Gamma_{5678}^{HH} + 2 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{73} \pi_4 \\
& + 2 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{78} \Gamma_{89}^G G_{93} \pi_4 \\
& - 2 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \sigma_7} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \delta \pi_9} G_{93} \pi_4 \\
& + 4 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \sigma_7} H_{78} \frac{\delta \Gamma_I}{\delta H_{89}} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta \pi_{\bar{1}}} G_{\bar{1}3} \pi_4 \\
& + 2 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta \pi_{\bar{1}}} G_{\bar{1}3} \pi_4 \\
& - 4 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta \Gamma_I}{\delta H_{0\bar{1}}} H_{\bar{1}2} \frac{\delta^2 \Gamma_I}{\delta \sigma_2 \delta \pi_{\bar{3}}} G_{\bar{3}3} \pi_4 \\
& + 2 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{78} \Gamma_{89}^G G_{90} \frac{\delta^2 \Gamma_I}{\delta \pi_0 \delta \sigma_{\bar{1}}} H_{\bar{1}2} \frac{\delta^2 \Gamma_I}{\delta \sigma_2 \delta \pi_{\bar{3}}} G_{\bar{3}3} \pi_4 \\
& - 4 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{78} \Gamma_{89}^G G_{90} \frac{\delta^2 \Gamma_I}{\delta \pi_0 \delta \sigma_{\bar{1}}} H_{\bar{1}2} \frac{\delta \Gamma_I}{\delta H_{2\bar{3}}} H_{34} \frac{\delta^2 \Gamma_I}{\delta \sigma_4 \delta \pi_{\bar{5}}} G_{\bar{5}3} \pi_4 \\
& - 2 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \sigma_7} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \delta \pi_9} G_{90} \Gamma_{0\bar{1}}^G G_{\bar{1}3} \pi_4 \\
& + 4 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \sigma_7} H_{78} \frac{\delta \Gamma_I}{\delta H_{89}} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta \pi_{\bar{1}}} G_{\bar{1}2} \Gamma_{2\bar{3}}^G G_{\bar{3}3} \pi_4 \\
& + 2 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta \pi_{\bar{1}}} G_{\bar{1}2} \Gamma_{2\bar{3}}^G G_{\bar{3}3} \pi_4 \\
& - 4 M_{1234} H_{15} H_{26} \frac{\delta^2 \Gamma_I}{\delta H_{56} \delta \pi_7} G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta \Gamma_I}{\delta H_{0\bar{1}}} H_{\bar{1}2} \frac{\delta^2 \Gamma_I}{\delta \sigma_2 \delta \pi_{\bar{3}}} G_{\bar{3}4} \Gamma_{4\bar{5}}^G G_{\bar{5}3} \pi_4
\end{aligned}$$

$$\begin{aligned}
& +2M_{1234}H_{15}H_{26}\frac{\delta^2\Gamma_I}{\delta H_{56}\delta\pi_7}G_{78}\Gamma_{89}^G G_{90}\frac{\delta^2\Gamma_I}{\delta\pi_0\delta\sigma_1}H_{12}\frac{\delta^2\Gamma_I}{\delta\sigma_2\delta\pi_3}G_{34}\Gamma_{45}^G G_{53}\pi_4 \\
& -4M_{1234}H_{15}H_{26}\frac{\delta^2\Gamma_I}{\delta H_{56}\delta\pi_7}G_{78}\Gamma_{89}^G G_{90}\frac{\delta^2\Gamma_I}{\delta\pi_0\delta\sigma_1}H_{12}\frac{\delta\Gamma_I}{\delta H_{23}}H_{34}\frac{\delta^2\Gamma_I}{\delta\sigma_4\delta\pi_5} \\
& \quad \times G_{56}\Gamma_{67}^G G_{73}\pi_4 \\
& +M_{1234}H_{15}H_{26}G_{37}G_{48}\Gamma_{5678}^{HG},
\end{aligned} \tag{20}$$

where it is understood that in all terms we have set $\sigma = 0$.

For diagrams without Goldstone field π or propagator G , we may also use the identities in Ref. [5], of course.

We still have to define the quantities $\Gamma^{\sigma H}$, $\Gamma^{\sigma G}$, Γ^{HH} , Γ^{HG} , and Γ^G appearing in the identities (18)-(20). The first four are given by

$$\begin{aligned}
\Gamma_{123}^{\sigma H} = & \frac{\delta^2\Gamma_I}{\delta\sigma_1\delta H_{23}} - \frac{\delta^2\Gamma_I}{\delta\sigma_1\delta\pi_4}G_{45}\frac{\delta^2\Gamma_I}{\delta\pi_5\delta H_{23}} - \frac{\delta^2\Gamma_I}{\delta\sigma_1\delta\pi_4}G_{45}\Gamma_{56}^G G_{67}\frac{\delta^2\Gamma_I}{\delta\pi_7\delta H_{23}} \\
& -2\frac{\delta\Gamma_I}{\delta H_{14}}H_{45}\frac{\delta^2\Gamma_I}{\delta\sigma_5\delta H_{23}} + 2\frac{\delta\Gamma_I}{\delta H_{14}}H_{45}\frac{\delta^2\Gamma_I}{\delta\sigma_5\delta\pi_6}G_{67}\frac{\delta^2\Gamma_I}{\delta\pi_7\delta H_{23}} \\
& +2\frac{\delta\Gamma_I}{\delta H_{14}}H_{45}\frac{\delta^2\Gamma_I}{\delta\sigma_5\delta\pi_6}G_{67}\Gamma_{78}^G G_{89}\frac{\delta^2\Gamma_I}{\delta\pi_9\delta H_{23}},
\end{aligned} \tag{21}$$

$$\begin{aligned}
\Gamma_{123}^{\sigma G} = & \frac{\delta\Gamma_I}{\delta\sigma_1\delta G_{23}} - \frac{\delta^2\Gamma_I}{\delta\sigma_1\delta\pi_4}G_{45}\frac{\delta^2\Gamma_I}{\delta\pi_5\delta G_{23}} - \frac{\delta^2\Gamma_I}{\delta\sigma_1\delta\pi_4}G_{45}\Gamma_{56}^G G_{67}\frac{\delta^2\Gamma_I}{\delta\pi_7\delta G_{23}} \\
& -2\frac{\delta\Gamma_I}{\delta H_{14}}H_{45}\frac{\delta\Gamma_I}{\delta\sigma_5\delta G_{23}} + 2\frac{\delta\Gamma_I}{\delta H_{14}}H_{45}\frac{\delta^2\Gamma_I}{\delta\sigma_5\delta\pi_6}G_{67}\frac{\delta^2\Gamma_I}{\delta\pi_7\delta G_{23}} \\
& +2\frac{\delta\Gamma_I}{\delta H_{14}}H_{45}\frac{\delta^2\Gamma_I}{\delta\sigma_5\delta\pi_6}G_{67}\Gamma_{78}^G G_{89}\frac{\delta^2\Gamma_I}{\delta\pi_9\delta G_{23}},
\end{aligned} \tag{22}$$

$$\begin{aligned}
\Gamma_{1234}^{HH} = & -\frac{\delta^2\Gamma_I}{\delta H_{12}\delta H_{34}} + \frac{\delta^2\Gamma_I}{\delta H_{12}\delta\pi_5}G_{56}\frac{\delta^2\Gamma_I}{\delta\pi_6\delta H_{34}} \\
& +\frac{\delta^2\Gamma_I}{\delta H_{12}\delta\pi_5}G_{56}\Gamma_{67}^G G_{78}\frac{\delta^2\Gamma_I}{\delta\pi_8\delta H_{34}} + \frac{\delta^2\Gamma_I}{\delta H_{12}\delta\sigma_5}H_{56}\frac{\delta^2\Gamma_I}{\delta\sigma_6\delta H_{34}} \\
& -2\frac{\delta^2\Gamma_I}{\delta H_{12}\delta\sigma_5}H_{56}\frac{\delta\Gamma_I}{\delta H_{67}}H_{78}\frac{\delta^2\Gamma_I}{\delta\sigma_8\delta H_{34}} \\
& +\frac{\delta^2\Gamma_I}{\delta H_{12}\delta\pi_5}G_{56}\frac{\delta^2\Gamma_I}{\delta\pi_6\delta\sigma_7}H_{78}\frac{\delta^2\Gamma_I}{\delta\sigma_8\delta\pi_9}G_{90}\frac{\delta^2\Gamma_I}{\delta\pi_0\delta H_{34}} \\
& -2\frac{\delta^2\Gamma_I}{\delta H_{12}\delta\pi_5}G_{56}\frac{\delta^2\Gamma_I}{\delta\pi_6\delta\sigma_7}H_{78}\frac{\delta\Gamma_I}{\delta H_{89}}H_{90}\frac{\delta^2\Gamma_I}{\delta\sigma_0\delta\pi_1}G_{12}\frac{\delta^2\Gamma_I}{\delta\pi_2\delta H_{34}}
\end{aligned}$$

$$\begin{aligned}
& -2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta \Gamma_I}{\delta H_{0\bar{1}}} H_{1\bar{2}} \frac{\delta^2 \Gamma_I}{\delta \sigma_2 \delta \pi_3} G_{34} \Gamma_{45}^G G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta H_{34}} \\
& -2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \delta \pi_7} G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta H_{34}} \\
& +4 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta \Gamma_I}{\delta H_{67}} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \delta \pi_7} G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta H_{34}} \\
& -2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \delta \pi_7} G_{78} \Gamma_{89}^G G_{90} \frac{\delta^2 \Gamma_I}{\delta \pi_0 \delta H_{34}} \\
& +4 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta \Gamma_I}{\delta H_{67}} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \delta \pi_9} G_{90} \Gamma_{0\bar{1}}^G G_{1\bar{2}} \frac{\delta^2 \Gamma_I}{\delta \pi_2 \delta H_{34}} \\
& -2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta \sigma_7} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \delta \sigma_9} G_{90} \Gamma_{0\bar{1}}^G G_{1\bar{2}} \frac{\delta^2 \Gamma_I}{\delta \pi_2 \delta H_{34}} \\
& +4 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta \sigma_7} H_{78} \frac{\delta \Gamma_I}{\delta H_{89}} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta \sigma_{\bar{1}}} G_{1\bar{2}} \Gamma_{23}^G G_{34} \frac{\delta^2 \Gamma_I}{\delta \pi_4 \delta H_{12}}, \quad (23)
\end{aligned}$$

$$\begin{aligned}
\Gamma_{1234}^{HG} = & -\frac{\delta^2 \Gamma_I}{\delta H_{12} \delta G_{34}} + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta G_{34}} \\
& + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta G_{34}} - \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \delta G_{34}} \\
& + 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta \Gamma_I}{\delta H_{67}} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \delta G_{34}} + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \sigma_7} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \delta G_{34}} \\
& - 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \sigma_7} H_{78} \frac{\delta \Gamma_I}{\delta H_{89}} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta G_{34}} \\
& + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \sigma_9} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta G_{34}} \\
& - 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \sigma_9} H_{90} \frac{\delta \Gamma_I}{\delta H_{0\bar{1}}} H_{1\bar{2}} \frac{\delta^2 \Gamma_I}{\delta \sigma_2 \delta G_{34}} \\
& - \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \pi_7} G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta G_{34}} \\
& + 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta \Gamma_I}{\delta H_{67}} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \pi_9} G_{90} \frac{\delta^2 \Gamma_I}{\delta \pi_0 \delta G_{34}} \\
& + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta \sigma_7} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \pi_9} G_{90} \frac{\delta^2 \Gamma_I}{\delta \pi_0 \delta G_{34}} \\
& - 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta \sigma_7} H_{78} \frac{\delta \Gamma_I}{\delta H_{89}} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \pi_{\bar{1}}} G_{1\bar{2}} \frac{\delta^2 \Gamma_I}{\delta \pi_2 \delta G_{34}}
\end{aligned}$$

$$\begin{aligned}
& + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta \pi_1} G_{12} \frac{\delta^2 \Gamma_I}{\delta \pi_2 \delta G_{34}} \\
& - 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta \Gamma_I}{\delta H_{01}} H_{12} \frac{\delta^2 \Gamma_I}{\delta \sigma_2 \delta \pi_3} G_{34} \frac{\delta^2 \Gamma_I}{\delta \pi_4 \delta G_{34}} \\
& - \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta^2 \Gamma_I}{\delta \sigma_6 \delta \pi_7} G_{78} \Gamma_{89}^G G_{90} \frac{\delta^2 \Gamma_I}{\delta \pi_0 \delta G_{34}} \\
& + 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \sigma_5} H_{56} \frac{\delta \Gamma_I}{\delta H_{67}} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \delta \pi_9} G_{90} \Gamma_{01}^G G_{12} \frac{\delta^2 \Gamma_I}{\delta \pi_2 \delta G_{34}} \\
& + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta \sigma_7} H_{78} \frac{\delta^2 \Gamma_I}{\delta \sigma_8 \pi_9} G_{90} \Gamma_{01}^G G_{12} \frac{\delta^2 \Gamma_I}{\delta \pi_2 \delta G_{34}} \\
& - 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \frac{\delta^2 \Gamma_I}{\delta \pi_6 \delta \sigma_7} H_{78} \frac{\delta \Gamma_I}{\delta H_{89}} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \pi_1} G_{12} \Gamma_{23}^G G_{34} \frac{\delta^2 \Gamma_I}{\delta \pi_4 \delta G_{34}} \\
& + \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta^2 \Gamma_I}{\delta \sigma_0 \delta \pi_1} G_{12} \frac{\delta^2 \Gamma_I}{\delta \pi_2 \delta G_{34}} \\
& - 2 \frac{\delta^2 \Gamma_I}{\delta H_{12} \delta \pi_5} G_{56} \Gamma_{67}^G G_{78} \frac{\delta^2 \Gamma_I}{\delta \pi_8 \delta \sigma_9} H_{90} \frac{\delta \Gamma_I}{\delta H_{01}} H_{12} \frac{\delta^2 \Gamma_I}{\delta \sigma_2 \delta \pi_3} G_{34} \frac{\delta^2 \Gamma_I}{\delta \pi_4 \delta G_{34}}, \quad (24)
\end{aligned}$$

while for Γ^G holds

$$\Gamma_{12}^G = - \frac{\delta^2 \Gamma_I}{\delta \pi_1 \delta \pi_2} - \Gamma_{13}^G G_{34} \frac{\delta^2 \Gamma_I}{\delta \pi_4 \delta \pi_2}. \quad (25)$$

To obtain recursion relations for the 1PI diagrams, it is most useful to separate Eqs. (18)-(25) by numbers of loops and by powers of σ and π . The resulting equations recursively generate all 1PI Feynman graphs, including the vacuum graphs (i.e. graphs with zero powers of σ and π).

We can hardly praise our results for conciseness. However, once the derivation of our equations as well as their implementation in terms of recursion relations are automated, they provide a straightforward and safe means of obtaining all 1PI Feynman graphs for the spontaneously broken $O(N)$ ϕ^4 -theory. Nevertheless it would be very useful if one could find a simpler set of equations, a task left for future work.

References

- [1] H. Kleinert, *Fortschr. Phys.* **30**, 187 (1982).
- [2] H. Kleinert, *Fortschr. Phys.* **30**, 351 (1982).

- [3] C. Bagnuls, C. Bervillier, D.I. Meiron, and B.G. Nickel, *Phys. Rev. B* **35**, 3585 (1987).
- [4] H. Kleinert and B. Van den Bossche, eprint: cond-mat/0011329.
- [5] B. Kastening, *Phys. Rev. E* **61**, 3501 (2000), eprint: hep-th/9908172.
- [6] A. Pelster and H. Kleinert, eprint: hep-th/0006153.
- [7] H. Kleinert, A. Pelster, B. Kastening, and M. Bachmann, *Phys. Rev. E* **62**, 1537 (2000), eprint: hep-th/9907168.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

CRITICAL BEHAVIOR OF CORRELATION FUNCTIONS AND ASYMPTOTIC EXPANSIONS OF FEYNMAN AMPLITUDES

A.P.C. MALBOUISSON

*Centro Brasileiro de Pesquisas Físicas-CBPF,
Rua Dr. Xavier Sigaud 150, Rio de Janeiro, RJ 22290-180, Brazil
E-mail: adolfo@lafex.cbpf.br*

We present a connection between the critical behavior of correlation functions and the general theory of asymptotic behaviors of Feynman amplitudes. Using the Mellin representation of Feynman integrals, an asymptotic expansion for a generic Feynman amplitude can be obtained for any set of invariants going to zero or to ∞ . If we take all masses going to zero in Euclidean metric, the truncated expansion has a rest compatible with the convergence of the series. In analogy to the application of field theory to critical phenomena, we consider from our general asymptotic expansions the critical behavior of correlation functions, in particular the critical behavior of the two-point function.

1 Introduction

In field theory, a special situation arises for Euclidean Green's functions in the momentum representation, when vanishingly small values for the external momenta are considered besides the zero-mass limit for the fields. In this case, we speak of the infrared (divergent) behavior of correlation functions. These divergences, which are seen as a "pathological" behavior in the context of applications of field theories to particle physics, are associated with the large-distance correlations in statistical systems and play a crucial role in the study of critical phenomena and phase transitions in such systems.

In this note we adopt a mathematical physicist's point of view, in the framework of perturbative field theory, starting from the observation that infrared (critical) behaviors of correlation functions in Euclidean field theories

may be seen as a special case of a general class of asymptotic behavior of Feynman amplitudes, as some of the involved masses tend to zero. The use of the perturbative method can be justified in applications of field theory to critical phenomena, for the examples of models of field theory that have been found to give relevant informations. These informations are controlled by the free field fixed-point, or by fixed-points that approach the free field fixed-point in some limit. This means that the Feynman diagram approach to field theory plays an important role in understanding physical situations in critical phenomena. As we have stressed above, the large-distance correlations in statistical systems are particularly important, as they play a crucial role in the study of phase transitions. In field theory language, these large-distance correlations manifest themselves as infrared behaviors of correlation functions, which are in perturbative language a particular case of asymptotic behaviors of Feynman amplitudes. This is one of the reasons why the analysis presented in this note could be interesting for the perturbative field theoretical approach to critical phenomena. For a complete account on the application of field theoretical methods to critical phenomena the reader is referred to the books by Kleinert [1].

Divergent large-distance behaviors of renormalized field theories containing massless fields and infrared divergences received a large amount of attention over the last decades. Historically, in applications to particle physics, they have been considered as an undesirable feature, a kind of "illness" of the theory which should be "cured" at any price. Actually these divergences appear at different levels. For Green's functions in Minkowskian metric it has been shown a long time ago that for some theories (e.g. QED₄) Green's functions exist at the zero-mass limit for some particles, as distributions on the 4-momenta. This means that Green's functions are well defined quantities in the infrared limit [2]. For particles on mass shell, Green's functions generally do not have a limit for those theories, even if they are well defined off mass shell. The oldest and best known examples are infrared divergences in scattering amplitudes in QED. Since the work of Bloch and Nordsieck [3], this problem has been investigated exhaustively [4,5].

It is worthwhile to emphasize that, in contrast to what happens in applications to particle physics, in applications of field theory to critical phenomena both ultraviolet divergences and infrared behaviors need not to be "cured". The ultraviolet cutoff is related to the inverse of some fundamental length of the system such as the atomic scale and the infrared behaviors of correlation functions describe directly the approach to critical points.

We make use of Mellin transform techniques to represent Feynman integrals, along similar lines as it has been done to study renormalization and asymptotic behaviors of scattering amplitudes [6–8], and to study the heat kernel expansion [9]. To fix our framework we consider a theory involving scalar fields $\varphi_i(x)$ having masses m_i , defined on a Euclidean space. For simplicity we may think of a single scalar field $\varphi(x)$ having a mass m . A generic Feynman graph G is a set of I internal lines, L loops, q connected components (a graph is disconnected if $q > 1$) and n vertices linked by some (polynomial) potential. To each vertex are attributed external momenta $\{p_i\}$ and internal ones $\{k_a\}$. A subgraph $S \subset G$ is defined as a graph, where all lines, vertices, and loops belong to G and a quotient graph G/S is a graph obtained from G reducing S to a point. A q -tree of the diagram G is a subgraph of G having q connected components, without loops and linking all vertices of G . The cases $q = 1$ (1-trees) and $q = 2$ (2-trees) are of particular interest for us.

The Feynman amplitude $G(\{a_k\})$ corresponding to the diagram G is a function of the set of invariants $\{a_k\}$ built from external momenta, $\sum_i p_i^2$, and squared masses m_i^2 ; it is defined in the Schwinger-Bogoliubov representation by [2,10]

$$G(a_k) = \int_0^\infty \prod_{i=1}^I d\alpha_i U^{-\frac{D}{2}}(\alpha) e^{-\frac{V(\alpha)}{U(\alpha)}}, \quad (1)$$

where D is the Euclidean space dimension with a positive metric.

In the above formula, the Symanzik polynomials $U(\alpha)$ and $V(\alpha)$ are constructed from the graph G by the prescription

$$U(\alpha) = \sum_{1.T} \prod_{i \notin 1.T} \alpha_i \quad (2)$$

and

$$V(\alpha) = \sum_{2.T} \left(\sum p_j \right)^2 \left(\prod_{i \notin 2.T} \alpha_i \right) + \left(\sum_{j \in G} m_j^2 \alpha_j \right) U(\alpha), \quad (3)$$

where the symbols $\sum_{1.T}$ and $\sum_{2.T}$ mean summation over the 1-trees and 2-trees of G , respectively. The sum $\sum p_j$ in Eq. (3) is the total external momentum entering one of the 2-tree connected components (any one of them equivalently, by momentum conservation). Notice that $U(\alpha)$ and $V(\alpha)$

are homogeneous polynomials in the α -variables, of degrees L and $L + 1$, respectively.

2 Mellin Representation and Asymptotic Expansions of Feynman Amplitudes

In the following we have in mind the physical situation of the infrared behavior, but we would like to emphasize that our method is quite general, in the sense that it applies to any asymptotic limit in Euclidean metric (any choice of the subset a_l below), for arbitrarily given external momenta, generic or exceptional, and for arbitrary vanishing or finite masses. If we perform a scale transformation on the subset $\{a_l\}$ of invariants, $a_l \rightarrow \lambda a_l$, the polynomial V splits into two parts,

$$V(\lambda a_m) = \lambda W(a_l, \alpha) + R(a_q, \alpha), \quad (4)$$

where the polynomials $W(a_l, \alpha)$ and $R(a_q, \alpha)$ are also homogeneous of degree $L + 1$ in the α -variables.

To be concrete we consider here a special situation with the external momenta $\{p_i\}$ fixed and we investigate the limit $\lambda \rightarrow 0$ corresponding to vanishing masses. In this case W is just the second term in Eq. (3). As we have noted above, the method applies along the same lines to any other class of asymptotic behavior. Indeed we note that from a dimensional argument,

$$G\left(\frac{a_l}{\lambda}, a_q\right) = \lambda^\omega G(a_l, \lambda a_q), \quad (5)$$

where $\omega = I - DL/2$. This means that the study of a given subset going to zero is equivalent to study the $\lambda \rightarrow \infty$ limit on the complementary subset of invariants.

Under the λ -scaling performed in Eq. (4) G becomes a function of λ , $G(\lambda)$, and its Mellin transform, $M(z) = \int_0^\infty d\lambda \lambda^{-z-1} G(\lambda)$ may be written in the form

$$M(z) = \Gamma(-z) \int_0^\infty \prod_{i=1}^I d\alpha_i U^{-\frac{D}{2}} e^{-\frac{R}{U}} \left(\frac{W}{U}\right)^z. \quad (6)$$

The scaled amplitude $G(\lambda)$ associated to the Feynman graph G may be obtained by the inverse Mellin transform,

$$G(\lambda) = \frac{1}{2i\pi} \int_{\sigma-i\infty}^{\sigma+i\infty} dz \lambda^z M(z), \quad (7)$$

where $\sigma = \text{Re}(z) < 0$ belongs to the analyticity domain of $M(z)$.

Since the integrand of Eq. (7) vanishes exponentially at $\sigma \pm i\infty$, due to the behavior of $\Gamma(z)$ at large values of $\text{Im } z$, the integration contour may be displaced to the right by Cauchy's theorem, picking up successively the poles of the integrand, provided we can desingularize the integral in Eq. (6). Such a problem has been studied by an appropriate choice of local coordinates [11] and also using Hepp sectors and a multiple Mellin representation [6].

From these works it has been possible to show that $M(z)$ has a meromorphic structure of the form

$$M(z) = \sum_{n,q} \frac{A_{nq} q!}{(z-n)^{q+1}}. \quad (8)$$

It results from the displacement of the integration contour in the inverse Mellin transform, an expansion for small values of λ , of the form

$$G(\lambda) = \sum_{n=n_0}^N \lambda^n \sum_{q=0}^{q_{\max}(n)} A_{nq} \ln^q(\lambda) + R_N(\lambda), \quad (9)$$

where the coefficients $A_n(\{p\})$ and the powers of logarithms come from the residues at the poles $z = n$.

The rest of the expansion $R_N(\lambda)$ is given by

$$R_N(\lambda) = \int_{-\infty}^{+\infty} \frac{d(\text{Im } z)}{2i\pi} \lambda^z \Gamma(-z) F(z), \quad (10)$$

with

$$N < \text{Re}(z) < N+1, \quad \text{Re}(z) = N + \eta, \quad 0 < \eta < 1 \quad (11)$$

and where

$$F(z) = \int_0^\infty \prod_{i=1}^I d\alpha_i U^{-\frac{D}{2}} e^{-\frac{B}{U}} \left(\frac{W}{U} \right)^z. \quad (12)$$

It is a rather difficult task to perform explicitly the α -integrations in Eq. (6) above for a general Feynman amplitude. As this calculation will not be necessary for our purposes, we give the appropriate references for the interested reader [6,7,12,13]. It is shown in these papers that renormalized Feynman amplitudes can be expressed as finite sums of convergent integrals which are exactly of the same type as those of convergent diagrams, provided the various integration variables associated to the remainders of renormalization Taylor

operators are renamed as supplementary Hepp-sector variables. In the following we keep the notations corresponding to convergent graphs, which means that the results are valid for convergent as well as for renormalized divergent diagrams.

We have shown the convergence of asymptotic expansions of general Feynman amplitudes in another article [14], obtaining a bound for the remainder of the expansion. In the particular case of all masses going to zero (infrared behavior), we have shown that for $I - DL/2 > 0$ (which is just the condition for UV convergence for the dimensionally regularized amplitude), the rest of the asymptotic expansion may be written in the form

$$|R_N(\lambda)| < K_1 K_N(\{p\})(\mu^2)^N \lambda^N, \quad (13)$$

where K_1 and K_N are finite constants and μ is a finite mass scale. The scaling parameter λ is arbitrarily small in the limit of vanishing masses. Therefore the factor $(\lambda\mu^2)^N$ in the bound above makes the sequence of the remainders $R_N(\lambda)$ converge to zero as $N \rightarrow \infty$, which is a condition for the convergence of the asymptotic expansion.

3 The 2-Point Function Critical Behavior

In another article we have shown [14] that one can obtain a convergent series from Eq. (9) as $N \rightarrow \infty$. Let us specify to the limit of all masses going to zero, and consider for simplicity the case of a single field having mass m . The analysis below can be generalized without difficulty to the case of several fields having different masses. We also consider dimensionally regularized amplitudes, that is, we take the Euclidean space dimension D to be such that the amplitudes are formally defined as convergent integrals, divergences appearing later as singularities for some diagrams.

For the 2-point function $G^{(2)}(p^2, m^2)$, the only nonzero invariant of the type $(\sum_i p_i)^2$ contributing to the construction of the Symanzik polynomial $V(\alpha)$ in Eq. (3) is p^2 . This may be seen if we note that for any diagram G contributing to the two-point function, the whole set of two-trees in the definition of $V(\alpha)$ in Eq. (3) divides into two classes, in which the total external momentum entering one of its connected components is either p^2 or zero (named respectively *relevant* and *irrelevant* two-trees). In this case, after introducing a fixed mass scale μ , it is easy to see that the inverse Mellin transform, Eq. (7), may be rewritten in terms of the variable p^2/m^2 . Then the small mass behavior of a graph G contributing to the two-point function,

$G^{(2)}(p^2, m^2)$, has the form

$$G(p^2, m^2) = \sum_{n=-\omega}^{\infty} \left(\frac{m^2}{p^2} \right)^n \sum_{q=0}^{q_{\max}(n)} A_{nq}(\mu^2) \left[-\ln \left(\frac{m^2}{p^2} \right) \right]^q. \quad (14)$$

We remember that according to Eqs. (6) and (7) the expansion above comes from the inverse Mellin transform,

$$G\left(\frac{m^2}{p^2}\right) = \frac{1}{2i\pi} \int_{\sigma-i\infty}^{\sigma+i\infty} dz \left(\frac{p^2}{m^2}\right)^z \Gamma(-z) \\ \times \int_0^\infty \prod_{i=1}^I d\alpha_i U^{-\frac{D}{2}} e^{-\mu^2 \sum_{i \in G} \alpha_i} \left(\frac{W'}{U}\right)^z, \quad (15)$$

where $W' = \mu^2(\sum_{j \in G} \alpha_j)U(\alpha)$, $R'(\alpha) = \sum'_{2.T} \prod'_{i \notin 2.T} \alpha_i$, and the notations $\sum'$ and $\prod'$ indicate respectively summation and product over *relevant* two-trees.

The coefficients $A_{nq}(\mu^2)$ in Eq. (14) come from the meromorphic structure of the Mellin transform displayed in Eq. (8). It is an extremely hard task to determine explicitly all these coefficients, which is equivalent to completely desingularize the integral over the α -variables in Eq. (15) respective to z . The coefficients $A_{\omega q}$ in Ref. [10] and Ref. [6] corresponding to the leading poles of the Mellin transform have been studied in the case of the behavior of Feynman amplitudes at large momenta, which is mathematically equivalent to the case of vanishing masses studied here. We adapt the method used in the above mentioned works to get an expression for the leading coefficients in the case of the small mass behavior. In the following we give only the general lines of the method, the results we have obtained and the definitions of the basic objects. The calculations are very involved and the full mathematical details in the case of large momenta behavior are in the above quoted references. These calculations can be adapted to our case without major difficulties. The main tool used to perform the analytic continuation of the Mellin transform is the generalized Taylor operator, a generalization of the operators used in field theory for the purpose of renormalization. It is defined as follows: given a function $f(x)$, such that $x^{-\nu}f(x)$ is infinitely differentiable at $x = 0$, we define the generalized Taylor operator τ^n as

$$\tau_x^n f(x) = x^{-\lambda-\epsilon} T^{n+\lambda} [x^{\lambda+\epsilon} f(x)], \quad (16)$$

where T is the usual Taylor operator, $\lambda \geq -E'(\nu)$ is an integer, $E'(\nu)$ is the smallest integer $\geq \text{Re}(\nu)$ and $\epsilon = E'(\nu) - \nu$. For any subdiagram $S \subseteq G$ this corresponds to a generalized Taylor operator defined by

$$\tau_S^n f(\{\alpha\}) = [\tau_\rho^n f(\{\alpha\})]_{\alpha_i = \rho^{2\alpha_i}, i \in S} \Big|_{\rho=1}. \quad (17)$$

A basic quantity associated to the diagram G playing a role in the desingularization procedure has the form

$$\prod_{S \subseteq G} \left[1 - \tau_S^{I(S)} \right] \left\{ U^{\frac{D}{2}}(\alpha) \left[\frac{W'(\alpha)}{U(\alpha)} \right]^z \right\}, \quad (18)$$

where the product runs over subdiagrams S of the graph G , including G itself. Although the τ operators do not commute, it can be shown that the complete product $\prod_{S \subseteq G} (1 - \tau_S^{I(S)}[.])$ is independent of the order of application of the factors upon the function between the brackets, $[.]$.

The procedure follows along lines parallel as was done in Ref. [10] and in Ref. [6] for the large momenta behavior. We obtain for the leading coefficients $A_{\omega q}$ the expression

$$\begin{aligned} A_{\omega q} &= \frac{1}{q!} \sum_{\{S_1, \dots, S_k\}} \frac{2^k}{(k-q-1)!} \xi_{G/S_1} \tilde{\xi}_{S_1/S_2} \dots \tilde{\xi}_{S_{k-1}/S_k} \\ &\times \left[\frac{d^{k-q-1}}{dz^{k-q-1}} [\Gamma(-z) \tilde{\gamma}_{S_k}(z)] \right]_{z=-n_0}. \end{aligned} \quad (19)$$

In the above equation, ω is the Weinberg leading power

$$\omega = \text{Sup}_G [\omega(S)], \quad (20)$$

where Sup_G runs over the superficial degree of divergence of all essential subdiagrams of G , $\omega(S) = L(S)D - 2I(S)$ and $q_{\max}(\omega) = Q - 1$, Q being the number of elements in the largest set of nested leading subdiagrams. The sum runs over all forests $\{S_1, \dots, S_k\}$ of $k(> q)$ nested *leading* subdiagrams $S_1 \supset S_2 \supset \dots \supset S_k$ (we remember that *leading* subdiagrams are those whose superficial degree of divergence equals ω). The quantities ξ , $\tilde{\xi}$, $\tilde{\gamma}$ are obtained from the subdiagrams $S \subseteq G$ by the formulas

$$\xi_S = \int_\kappa \prod_{i \in S} d\alpha_i e^{-\mu^2 \sum_{i \in S} \alpha_i} \prod_{S' \subseteq S} \left[1 - \tau_{S'}^{-2I(S')} \right] U_S^{\frac{D}{2}}(\alpha), \quad (21)$$

$$\tilde{\xi}_S = -\frac{d\xi_S}{d\mu^2} = \mu^2 \sum_{i \in S} \xi_{S_i}, \quad (22)$$

$$\begin{aligned} \gamma_{\tilde{S}_k} &= \frac{\Gamma(-\frac{z}{2})}{\Gamma(-z)} \mu^2 \int_0^\infty \prod_{i \in S} d\alpha_i e^{-\mu^2 \sum_{j \in S} \alpha_j} \\ &\times \prod_{S' \subseteq S} \left[1 - \tau_{S'}^{-2I(S')} \right] \left[\frac{W'_S}{U_S}(\alpha) \right]^{\frac{z}{2}} \left[\sum_{i \in S} \alpha_i \right] U_S^{-\frac{D}{2}}(\alpha), \end{aligned} \quad (23)$$

where S_i is the diagram obtained from S by inserting a two leg vertex (a mass insertion μ^2) on the line i . Particular cases for the quantities in the above equations are $\varrho_{G/G} = 1$ and $\varrho_{G/S} = 0$ if G itself is leading. The Feynman amplitude corresponding to S_i is simply given by

$$G_{S_i} = \int_0^\infty \prod_{i \in S} d\alpha_i U^{-\frac{D}{2}}(\alpha) (\alpha_i) e^{-\frac{V(\alpha)}{U(\alpha)}}. \quad (24)$$

The various factors in Eqs. (19), (22), and (23) can be reorganized to write the leading coefficients in the more convenient form

$$\begin{aligned} A_{\omega q} &= \frac{1}{q!} \sum_{\{S_1, \dots, S_k\}} \frac{2^k}{(k-q-1)!} \\ &\times \frac{d^{k-q-1}}{dz^{k-q-1}} \left[\int_0^\infty \prod_{i \in G} d\alpha_i e^{-\mu^2 \sum_{i \in G} \alpha_i} A(\alpha; z) \right]_{z=\omega}, \end{aligned} \quad (25)$$

where the function $A(\alpha; z)$ is defined by

$$\begin{aligned} A(\alpha; z) &= \Gamma\left(-\frac{z}{2}\right) R'_G \left[U_{G/S_1}^{-\frac{D}{2}} U_{S_1/S_2}^{-\frac{D}{2}} \dots U_{S_{k-1}/S_k}^{-\frac{D}{2}} U_{S_k}^{-\frac{D}{2}} \right. \\ &\times \left. \left[\sum_{S_1/S_2} \alpha_i \right] \dots \left[\sum_{S_{k-1}/S_k} \alpha_i \right] \left[\sum_{S_k} \alpha_i \right] (W'_{S_k}/U_{S_k})^{\frac{z}{2}} \right], \end{aligned} \quad (26)$$

and where, taking the convention $S_0 \equiv G$, R'_G is the operator (the order $-2I(T)$ is understood for each τ operator corresponding to a subdiagram T)

$$R'_G = \prod_{l=1}^k \left[\prod_{T_l \subseteq S_{l-1}/S_l} (1 - \tau_{T_l}) \right] \prod_{T \subseteq S_k} (1 - \tau). \quad (27)$$

The operator R'_G does not change the homogeneity properties of the functions upon which it acts. So, remembering the homogeneity properties of the polynomials U and W' , and noting that $L(G/S_1) + L(S_1/S_2) + \dots + L(S_{k-1}/S_k) + L(S_k) = L(G)$, we see from Eq. (26) that $A(\alpha; z)$ is a homogeneous function in the α -variables of degree $L(G)D/2 + k + z/2$. Then taking spherical coordinates in α -space we may write from the preceding equations

$$A_{\omega q} = \frac{1}{q!} \sum_{\{S_1, \dots, S_k\}} \frac{2^{k-1}}{(k-q-1)!} \frac{d^{k-q-1}}{dz^{k-q-1}} \left[\Gamma\left(-\frac{z}{2}\right) \int d\Omega \int_0^\infty d\varrho e^{\varrho \mu^2 f(\Omega)} \right. \\ \left. \times \varrho^{I - \frac{L(G)D}{2} + k + \frac{z}{2} - 1} g(\Omega; z) \right]_{z=\omega}. \quad (28)$$

The integral over ϱ in the equation above may be expressed in terms of the Γ -function, and $f(\Omega)$, $g(\Omega; z)$ are functions of the angular variables Ω depending on the specific topological characteristics of the graph G considered. We get for the leading coefficients $A_{\omega q}$

$$A_{\omega q} = \frac{1}{q!} \sum_{\{S_1, \dots, S_k\}} \frac{2^{k-1}}{(k-q-1)!} \frac{d^{k-q-1}}{dz^{k-q-1}} \left[\Gamma\left(-\frac{z}{2}\right) \int d\Omega g(\Omega; z) \right. \\ \left. \times (\mu^2 f(\Omega))^{-(I - \frac{L(G)D}{2} + k + \frac{z}{2})} \Gamma\left(I - \frac{L(G)D}{2} + k + \frac{z}{2}\right) \right]_{z=\omega}. \quad (29)$$

In the very neighbourhood of criticality the contribution of the amplitude G is given by the leading term in the expansion (14) which corresponds to the highest powers of m^2/p^2 and of $-\ln(m^2/p^2)$. This means that for very small values of m^2 we have

$$G(p^2, m^2) \approx A_{\omega Q} \left(\frac{m^2}{p^2}\right)^{-\omega} \left[-\ln\left(\frac{m^2}{p^2}\right)\right]^{Q-1}, \quad (30)$$

where we remember that ω is the Weinberg leading power, given by Eq. (20), and Q is the number of elements in the largest set of leading subdiagrams of G . To get the coefficient $A_{\omega Q}$ in Eq. (30) from Eq. (29), we note that the sum has only one term, corresponding to the nest $\{S_1, \dots, S_Q\}$ and a zeroth-order derivative. We obtain

$$A_{\omega Q} = \frac{2^{Q-1}}{(Q-1)!} \Gamma\left(-\frac{\omega}{2}\right) g(\Omega; z) \Gamma\left(I - \frac{L(G)D}{2} + k + \frac{z}{2}\right) \\ \times \int d\Omega g(\Omega; \omega) [\mu^2 f(\Omega)]^{-[I - \frac{L(G)D}{2} + k + \frac{z}{2}]}. \quad (31)$$

Now, if G itself is leading it does not contribute to the expansion, since in this case $\varrho_{G/S} = 0$ for every $S \subset G$. If G is not leading, $\omega(G) = L(G) - 2I(G) < \omega$, and there exists a $\delta > 0$ such that $\omega = L(G)D - 2I(G) + \delta$. The arguments of the Γ -functions in Eq. (31) above are respectively $-\omega/2 = (2I(G) - L(G)D - \delta)/2$ and $\delta + Q$. Since $\delta > 0$ and $Q \geq 1$, singularities in $A_{\omega Q}$ come from the factor $\Gamma(-\omega/2)$. Thus at fixed space dimension Δ (for instance $\Delta = 3$), if the diagram G has a topological structure such that $L(G)\Delta - 2I(G) + \delta = 2n$, n an integer ≥ 0 , the corresponding singularity of the Γ -function above, $\Gamma(-\omega/2)$, should be removed by a renormalization procedure. This may be done, as usual, taking $D = \Delta - \epsilon$, and subtracting the pole at $\epsilon = 0$, leaving some regular function $\Gamma_{\text{Ren}}(\Delta)$. The result in space dimension Δ for the coefficient $A_{\omega Q}$ reads

$$A_{\omega Q}(G) = \frac{2^{Q(G)-1}}{(Q(G)-1)!} \Gamma_{\text{Ren};G}(\Delta) (\mu^2)^{-[\delta+Q(G)]} \Gamma(\delta + Q(G)) \\ \times \int d\Omega g_G(\Omega; \omega) [f_G(\Omega)]^{-[\delta+Q(G)]}. \quad (32)$$

In the above equation we have displayed explicitly the dependence of the various quantities on the Feynman amplitude G we have considered.

Thus the behavior of the two-point function near criticality is described by an expression having the form

$$G^{(2)}(p^2, m^2) \approx \sum_G A_{\omega Q}(G) \left(\frac{m^2}{p^2} \right)^{-\omega(G)} \left[-\ln \left(\frac{m^2}{p^2} \right) \right]^{Q(G)-1}, \quad (33)$$

where the symbol $\sum_G$ means summation over the whole set of Feynman diagrams contributing to the two-point function. The quantities under the summation symbol can be obtained by explicitly calculating each Feynman diagram G . This result holds for any scalar field theory without derivative couplings.

References

- [1] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines*, Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [2] J.H. Lowenstein and W. Zimmermann, *Nucl. Phys. B* **86**, 77 (1975).
- [3] F. Bloch and A. Nordsieck, *Phys. Rev.* **52**, 54 (1937).

- [4] D.R. Yennie, S.C. Frautshi, and H. Suura, *Ann. Phys.* **13**, 379 (1961).
- [5] T.W.B. Kibble, *Phys. Rev.* **133**, 1527 (1968); *Phys. Rev.* **174**, 1882 (1968); *Phys. Rev.* **175**, 1624 (1968).
- [6] M.C. Bergère, C. de Calan, and A.P.C. Malbouisson, *Comm. Math. Phys.* **62**, 137 (1978).
- [7] C. de Calan and A.P.C. Malbouisson, *Ann. Inst. Henri Poincaré* **32**, 91 (1980).
- [8] C. de Calan and A.P.C. Malbouisson, *Comm. Math. Phys.* **90**, 413 (1983).
- [9] A.P.C. Malbouisson, *J. Math. Phys.* **35**, 479 (1994).
- [10] M.C. Bergère and Y.M.P. Lam, *Comm. Math. Phys.* **39**, 1 (1974).
- [11] A.N. Vartchenko, *Functional Analysis and Its Applications* **10**, 13 (1976).
- [12] M.C. Bergère and J.B. Zuber, *Comm. Math. Phys.* **35**, 113 (1974).
- [13] M.C. Bergère and Y.M.P. Lam, *J. Math. Phys.* **17**, 1546 (1976).
- [14] A.P.C. Malbouisson, *J. Phys. A* **33**, 3587 (2000).

GAUGE SYMMETRY AND NEURAL NETWORKS

T. MATSUI

Department of Physics, Kinki University, Higashi-Osaka, 8502 Japan

E-mail: matsui@phys.kindai.ac.jp

We propose a new model of neural network. It consists of spin variables to describe the state of neurons as in the Hopfield model and new gauge variables to describe the state of synapses. The model possesses local gauge symmetry and resembles lattice gauge theory of high-energy physics. Time dependence of synapses describes the process of learning. The mean-field theory predicts a new phase corresponding to the confinement phase, in which the brain loses the ability of learning and memory.

1 Introduction

The Hopfield model of neural network [1] succeeds to explain some basic functions of human brain such as the associated memory. However, to be a more realistic model, at least the following points shall be taken into account:

- effects of external stimulations through eyes and ears on neurons,
- effects of time variations of synapses on neurons.

The second point is essential to describe the function of learning, since the possible patterns to memorize are completely determined according to the strengths of synapse connections among neurons as long as they are time-independent. Their time-dependence induces the process of learning itself.

In Section 2, we review the Hopfield model briefly. Then, in Section 3, we propose a new model of neural network, in which the strengths of synapse connections are regarded as gauge connections varying in time according to the gauge principle. By using the mean-field theory, we see that the model predicts a new state of the brain in which both learning and memory are impossible. An outlook is presented in Section 4.

2 Hopfield Model

Let us briefly review the framework of the Hopfield model, where the energy $E_H(\{S_i\})$ is given by

$$E_H(\{S_i\}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N J_{ij} S_i S_j. \quad (1)$$

Here, $S_i = \pm 1$ is the Ising spin variable to describe the state of the i -th neural cell ($i = 1, 2, \dots, N$). When excited, it is $S_i = 1$, when unexcited we have $S_i = -1$. J_{ij} is a given real constant that expresses the strength of synapse connection for the signal propagating from the j -th cell to the i -th cell.

The time evolution of $S_i(t)$ for every discrete time interval ϵ (often set unity) is governed by the following equation:

$$S_i(t + \epsilon) = \text{sgn} \left[-\frac{\partial E_H}{\partial S_i}(t) \right] = \text{sgn} \left[\sum_j J_{ij} S_j(t) \right]. \quad (2)$$

Thus, $J_{ij} > 0$ tries to proliferate (un)excited cells, while $J_{ij} < 0$ prefers mixtures of excited and unexcited ones. If the system converges into a certain configuration of $\{S_i\}$ after a sufficiently long time, it corresponds to recalling certain pattern. Such a configuration should be a stationary point of E_H , i.e. $\partial E_H / \partial S_i = 0$ for every i . All these configurations are determined once the coefficients J_{ij} are given.

Practically speaking, the rule (2) may not necessarily hold over the whole time interval due to unavoidable errors in signal propagations. Such a situation may be simulated by adding random noises $\eta_i(t)$ into the square bracket in the right-hand side of (2), whose strength can be identified as a fictitious "temperature" T . If T is large, the error in signal propagations occurs frequently. Thus it is interesting to study statistical mechanics of the system E_H by using the Boltzmann distribution. The partition function Z_H is given by

$$Z_H = \prod_i \sum_{S_i = \pm 1} \exp(-\beta E_H), \quad \beta \equiv 1/T. \quad (3)$$

In the case that all J_{ij} are positive, the system has two phases:

- a ferromagnetic phase below a certain critical temperature T_c , $T < T_c$, in which there is a long-range order, and the average $\langle S_i \rangle \neq 0$,

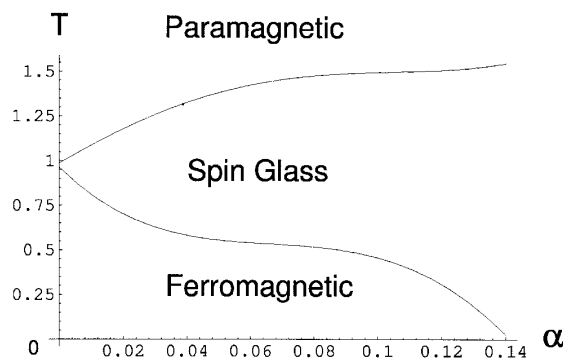


Figure 1. Phase structure of the Hopfield model in the $\alpha(\equiv M/N) - T$ plane.

Table 1. Phases of the Hopfield model.

Phase	$\sum_i \langle S_i \rangle$	$\sum_i \langle S_i \rangle^2$	Property
Ferromagnetic	$\neq 0$	$\neq 0$	memory
Spin glass	0	$\neq 0$	false memory
Paramagnetic	0	0	no memory

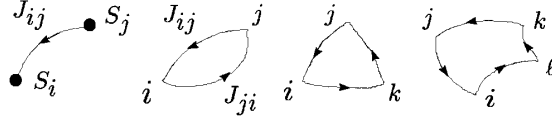
- a paramagnetic phase above T_c , $T > T_c$, in which S_i are random and $\langle S_i \rangle = 0$.

The ferromagnetic phase corresponds to the state of clear memory, while in the paramagnetic phase no definite patterns can persist. If J_{ij} is complicated, there arises a spin-glass phase as we shall see.

Explicitly, let us fix J_{ij} according to Hebb's rule as

$$J_{ij} = \frac{1}{N} \sum_{\alpha=1}^M \xi_i^\alpha \xi_j^\alpha, \quad (4)$$

where we prepare M patterns $S_i = \xi_i^\alpha$ ($\alpha = 1, \dots, M$) to recall. The replica method gives rise to the phase diagram shown in Fig. 1 [2]. Each phase is explained in Table 1.

Figure 2. Graphical representation of each term in E of Eq. (6).

3 New Model with Local Gauge Symmetry

As pointed out in Section 1, one needs the time variation of J_{ij} to incorporate the function of learning. There are various approaches for this point. We regard both S_i and J_{ij} as dynamical variables and treat them on an equal footing. Let us assume that their time dependence is controlled such that they reach a local minimum of the new energy $E(\{S_i, J_{ij}\})$. In order to determine E , we impose the condition that E is locally gauge invariant under the following gauge transformation:

$$S_i \rightarrow S'_i \equiv V_i S_i, \quad J_{ij} \rightarrow J'_{ij} \equiv V_i J_{ij} V_j, \quad E(\{S'_i, J'_{ij}\}) = E(\{S_i, J_{ij}\}), \quad (5)$$

where $V_i = \pm 1$ is the $Z(2)$ variable associated with the i -th cell. Since J_{ij} describes the state of the synapse connecting the i -th and the j -th cells, it is natural to regard it as the connection of gauge theory. The neural network may possess certain conserved quantities in association with the long-term memory. The local gauge symmetry we address may respect such a conservation law. This point will be reported in detail in a separate publication [3]. It is often stressed that the connections J_{ij} and J_{ji} are independent (asymmetric). Then a general form of $E(\{S_i, J_{ij}\})$ and the partition function Z may be given by

$$\begin{aligned} E = & -\frac{1}{2} \sum_{i,j} S_i J_{ij} S_j + \frac{g_2}{2} \sum_{i,j} J_{ij} J_{ij} \\ & + \frac{g_3}{3!} \sum_{i,j,k} J_{ij} J_{jk} J_{ki} + \frac{g_4}{4!} \sum_{i,j,k,\ell} J_{ij} J_{jk} J_{k\ell} J_{\ell i} + \dots, \\ Z = & \prod_i \sum_{S_i = \pm 1} \prod_{i \neq j} \int dJ_{ij} \exp(-\beta E), \quad \beta \equiv 1/T. \end{aligned} \quad (6)$$

Since $V_i^2 = 1$, each term of E is gauge invariant (see Fig. 2). E takes a form very similar to the lattice gauge theory in particle physics, where S_i

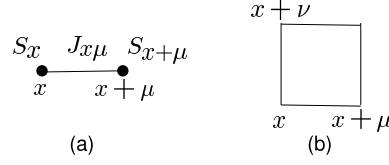


Figure 3. Graphical representation of E of Eq. (7): (a) λ term; (b) g^{-2} term.

corresponds to a matter field and J_{ij} to an exponentiated gauge field. If the parameters $g_2, g_3, g_4, \dots$ are set to zero, E reduces to E_H of (1).

3.1 Model I

To be explicit, we need to specify the model further. Let us first consider the $Z(2)$ Higgs gauge model on a 3D cubic lattice,

$$E_I = -\lambda \sum_x \sum_{\mu=1}^3 S_{x+\mu} J_{x\mu} S_x - \frac{1}{g^2} \sum_x \sum_{\mu < \nu} J_{x\mu} J_{x+\mu, \nu} J_{x+\nu, \mu} J_{x\nu},$$

$$Z_I = \prod_x \sum_{S_x=\pm 1} \prod_{x, \mu} \sum_{J_{x\mu}=\pm 1} \exp(-\beta E_I) \equiv \exp(-\beta F_I), \quad (7)$$

where x denotes the lattice site on which S_x lives, and μ ($= 1, 2, 3$) both the direction and the unit vector. We consider only the connections between the nearest-neighbor sites $(x, x + \mu)$ and treat them as a symmetric $Z(2)$ variable on a link $(x, x + \mu)$; $J_{x, x+\mu} = J_{x+\mu, x} \equiv J_{x\mu} = \pm 1$. The λ term and the $1/g^2$ -term are depicted in Fig. 3.

The time evolution of $J_{x\mu}$ may be given by the similar rule as (2),

$$S_x(t + \epsilon) = \text{sgn} \left[-\frac{\partial E_I}{\partial S_x}(t) + \eta_x(t) \right],$$

$$J_{x\mu}(t + \alpha\epsilon) = \text{sgn} \left[-\frac{\partial E_I}{\partial J_{x\mu}}(t) + \zeta_{x\mu}(t) \right], \quad (8)$$

where α sets the ratio of the two time scales for S_x and $J_{x\mu}$. We report our study of (8) elsewhere [4].

Now let us study the phase diagram of E_I by using the mean-field theory, which is formulated as a variational principle as follows. Let us introduce a variational energy E_0 . Then the Jensen-Peierls inequality gives rise to

$$F_I \leq F_0 + \langle E_I - E_0 \rangle_0,$$

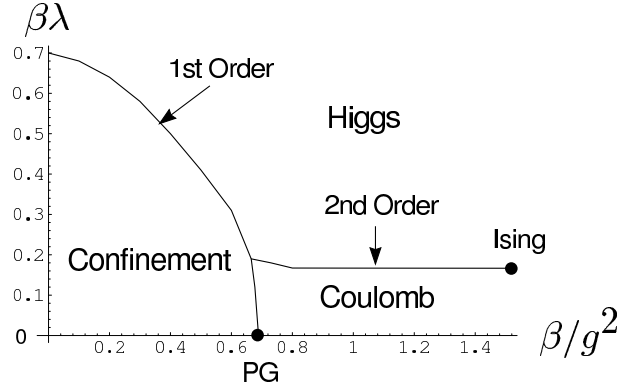


Figure 4. Phase diagram of Model I of Eq. (7). The point marked as Ising locates the second-order transition point of the Ising model. The point PG locates the first-order transition point of the $Z(2)$ pure gauge theory.

$$\begin{aligned} Z_0 &= \text{Tr} \exp(-\beta E_0) \equiv \exp(-\beta F_0), \\ \langle O \rangle_0 &= Z_0^{-1} \text{Tr} O \exp(-\beta E_0), \end{aligned} \quad (9)$$

where Tr implies $\prod_x \sum_{S_x=\pm 1} \prod_{x,\mu} \sum_{J_{x\mu}=\pm 1}$. We choose the variational parameters in E_0 so that the right-hand-side of inequality reaches the minimum. For E_0 we assume the translational invariance of mean fields and employ the single-site and single-link energy,

$$E_0 = - \sum_x \sum_{\mu} W_{x\mu} J_{x\mu} - \sum_x h_x S_x, \quad (10)$$

with the two variational parameters $W_{x\mu} = W$ and $h_x = h$. The result is given in Table 2 and Fig. 4. As shown in Table 2, one may take $\langle S_x \rangle$ as an order parameter to judge whether the system succeeds to recall definite patterns, and $\langle J_{x\mu} \rangle$ is taken to judge whether the system is able to learn some new patterns. In the confinement phase, neither memory nor learning is possible. This phase is missing in the Hopfield model. We note that Monte Carlo simulations of the 3D $Z(2)$ Higgs gauge model exhibit these three phases, but the phase boundary of Higgs and confinement phases does not continue to $\beta/g^2 = 0$ but terminates at some finite value. These two phases can reach each other smoothly by contouring the end point. This “complementarity” reflects that $|J_{x\mu}| = 1$. This is proved by a rigorous treatment [5], but not predicted correctly in our variational treatment [6].

Table 2. Phases of Model I of Eq. (7).

Phase	$\langle S_x \rangle$	$\langle J_{x\mu} \rangle$	Memory	Learning	Hopfield Model
Higgs	$\neq 0$	$\neq 0$	yes	yes	Ferromagnetic
Coulomb	0	$\neq 0$	no	yes	Paramagnetic
Confinement	0	0	no	no	not available

In order to answer to what extent these results are trustworthy, we introduce and study two other models, Model II and Model III.

3.2 Model II

The framework and the energy E_{II} of Model II is the same as in Model I, but here we allow the additional state $J_{x\mu} = 0$ which describes the possibility that the connection between x and $x + \mu$ is missing:

$$E_{\text{II}} = E_{\text{I}}, \quad Z_{\text{II}} = \prod_x \sum_{S_x = \pm 1} \prod_{x, \mu} \sum_{J_{x\mu} = 0, \pm 1} \exp(-\beta E_{\text{II}}). \quad (11)$$

The phase diagram calculated by the similar mean-field theory is shown in Fig. 5. The global structure remains the same as in Fig. 4, although the region of the confinement phase is enlarged as expected since the added states clearly favor this phase.

3.3 Model III

In Model III, we introduce two independent $Z(2)$ variables $J_{x\mu}$ and $\bar{J}_{x\mu}$ for the synapse between x and $x + \mu$ to take into account their independence as

$$J_{x\mu} \equiv J_{x, x+\mu}, \quad \bar{J}_{x\mu} \equiv J_{x+\mu, x}. \quad (12)$$

We also define $J_{x, -\mu} \equiv J_{x-\mu, x}$. The energy E_{III} is then given by

$$E_{\text{III}} = -\lambda \sum_x \left(\sum_{\pm\mu} \bar{J}_{x, \pm\mu} S_{x \pm \mu} \right) \left(\sum_{\pm\nu} J_{x, \pm\nu} S_{x \pm \nu} \right) - \frac{1}{g^2} \sum_x \sum_{\mu < \nu} [\bar{J}_{x\mu} \bar{J}_{x+\mu, \nu} J_{x+\nu, \mu} J_{x\nu} + (\mu \leftrightarrow \nu)]. \quad (13)$$

We note that the expression $J_{ij} S_i S_j$ in $E_{\text{H, I, II}}$ washes out the asymmetry $J_{ij} \neq J_{ji}$, while the first term of (13) reflects it. Each term in E is depicted in

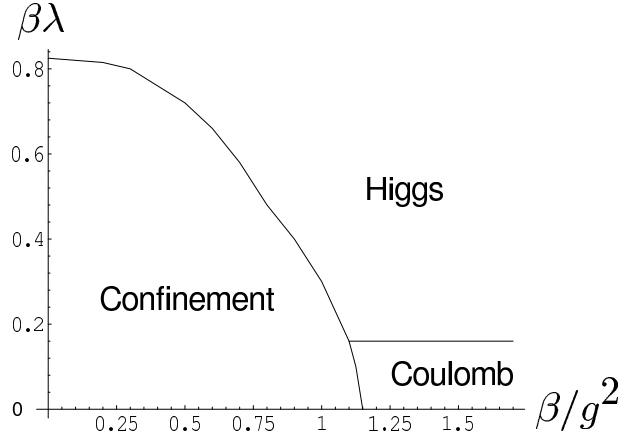


Figure 5. Phase diagrams of Model II of Eq. (11).

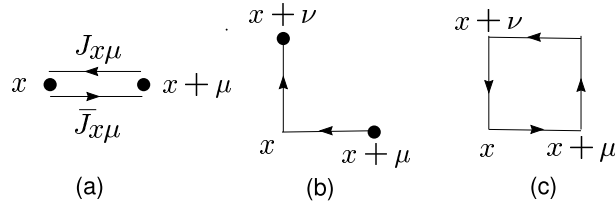
Figure 6. Graphical representation of Model III of Eq. (13): (a) $J_{x\mu}$ and $\bar{J}_{x\mu}$; (b) λ term; (c) g^{-2} term.

Fig. 6. The phase diagram in mean-field theory is shown in Fig. 7. The global structure still remains unchanged, although the region of the confinement phase is diminished considerably. This may be understood since the first term in E_{III} is bilinear in J_{ij} , in contrast to $E_{\text{H, I, II}}$, and it favors nonvanishing J_{ij} .

4 Summary and Outlook

Our results may be summarized with some outlook as follows:

- Due to the dynamical variables J_{ij} , a new confinement phase appears at high temperatures, which describes the new state of no ability of learning

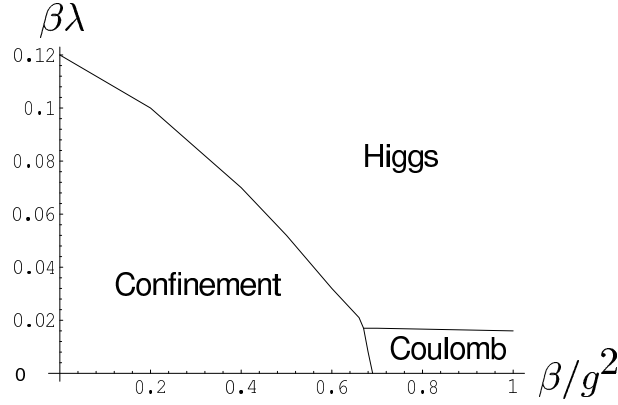


Figure 7. Phase diagrams of Model III of Eq. (13).

and memory.

- To describe the spin-glass phase, further study of long-range correlations and/or frustrations is necessary.
- Relaxing of $J_{ij} = \pm 1(,0)$ to $-\infty < J_{ij} < \infty$ may be interesting, but requires a detailed form of the energy.
- Study of the time evolution of J_{ij} and S_i may describe the mechanism of learning such as the process to forget the patterns.
- Study of the effect of local gauge symmetry on brain function on a “quantum” level is interesting. Introduction of gauged versions of quantum brain dynamics [7] and cellular automata with Penrose’s idea [8] may be the first step.

Acknowledgments

I thank Prof. Hagen Kleinert for fruitful discussions on various fields of physics during my pleasant stay at the Freie Universität Berlin in 1983-1990. I also appreciate discussions with Dr. Kazuhiko Sakakibara and Mr. Motohiro Kemuriyama.

References

- [1] J.J. Hopfield, *Proc. Nat. Acad. Sci. USA* **79**, 2554 (1982).
- [2] D.J. Amit, H. Gutfreund, and H. Sompolinsky, *Ann. Phys. (N.Y.)* **173**,

30 (1987).

- [3] T. Matsui and K. Sakakibara, in preparation.
- [4] M. Kemuriyama and T. Matsui, in preparation.
- [5] E. Fradkin and S. Shenker, *Phys. Rev. D* **19**, 3682 (1979).
- [6] E. Brézin and J.M. Drouffe, *Nucl. Phys. B* **200**, 93 (1982).
- [7] C. Stuart, Y. Takahashi, and H. Umezawa, *J. Theor. Biol.* **71**, 605 (1978).
- [8] S. Hameroff and R. Penrose, *J. Consciousness Study* **3**, 36 (1996).

Part III

Variational Perturbation Theory

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

NOTE ON THE PATH-INTEGRAL VARIATIONAL APPROACH IN MANY-BODY THEORY

J.T. DEVREESE

*Theoretische Fysica van de Vaste Stof, Universiteit Antwerpen (UIA),
B-2610 Antwerpen, Belgium;
Universiteit Antwerpen (RUCA), B-2020 Antwerpen, Belgium;
TU Eindhoven, 5600 MB Eindhoven, The Netherlands
E-mail: devreese@uia.ua.ac.be*

I discuss how a variational approach can be extended to systems of identical particles (in particular fermions) within the path-integral treatment. The applicability of the many-body variational principle for path integrals is illustrated for different model systems and is shown to depend crucially on whether or not a model system possesses the proper symmetry with respect to permutations of identical particles.

1 Introduction

In the path-integral formulation of quantum mechanics, the so-called *Jensen-Feynman inequality* provides an upper bound to the free energy of a quantum system, if properly applied. It was introduced [1] in Feynman's path-integral approach to the Fröhlich polaron (see formula (8.40) in Ref. [2]):

$$F \leq F_M + \frac{1}{\beta} \langle S - S_M \rangle_{S_M} \text{ if } S, S_M \text{ are real.} \quad (1)$$

In the variational functional, F and S are the free energy and the action functional^a of the system under consideration, whereas F_M and S_M are the free energy and the action functional of a model system; the temperature is described by the parameter $\beta = 1/k_B T$. Angular brackets mean a weighted

^aIt is implicitly assumed that the action functional and the path integral are expressed in the imaginary-time variable. This convention is followed throughout the present paper.

average over the paths [2]:

$$\langle \bullet \rangle_{S_M} \equiv \frac{\int \bullet \exp(-S_M) D(\text{path})}{\int \exp(-S_M) D(\text{path})}. \quad (2)$$

A rigorous argument to prove the inequality (1) is based on the convex nature of the exponential $\exp(x)$ of a *real* stochastic variable x (see e.g. Fig. 11-1 in Ref. [3]), which leads to $\langle e^x \rangle \geq e^{\langle x \rangle}$ with $\langle x \rangle$ being the weighted average of x .

Apart from Feynman's variational treatment of the ground-state energy of a polaron, the path-integral approach based on the Jensen-Feynman inequality was successfully applied to a series of problems [5], e.g. to the calculation of the effective classical partition function [4,6] and of quantum corrections to the free energy of nonlinear systems [7,8], to the description of all critical exponents observable in second-order phase transitions [9], and to the problem of bipolaron stability [10,11].

The derivation of the Jensen-Feynman inequality crucially depends on the assumption that both the action and the trial action are *real* functionals. As already recognized by Feynman (see Ref. [3], p. 308), its application to a *polaron in a magnetic field* therefore becomes problematic, because the action functional S for a polaron in a magnetic field (and any reasonable trial action S_M) is no longer real-valued. A discussion of this problem lies beyond the scope of the present paper. For more details on the status of this problem, see the literature [12–18].

The Jensen-Feynman inequality is reminiscent of the Bogoliubov inequality [19,20], which provides the following upper bound to the free energy F of a system described by the Hamiltonian H

$$F \leq F_M - \frac{1}{\beta} \frac{\text{Tr} [(H - H_M) e^{-\beta H_M}]}{\text{Tr} (e^{-\beta H_M})} \text{ if } H, H_M \text{ are Hermitian}, \quad (3)$$

where H_M is the Hamiltonian of some trial system with free energy F_M . The Rayleigh-Ritz variational principle (see e.g. Ref. [21], p. 172) for the ground-state energy $E \leq \langle \Psi_M | H | \Psi_M \rangle / \langle \Psi_M | \Psi_M \rangle$ with a trial state $|\Psi_M\rangle$ is the zero-temperature limit of the Bogoliubov inequality.

The condition that S and S_M are real in (1) is not necessarily equivalent to the requirement that H and H_M are Hermitian in (3). If the Hamiltonians H and H_M in the Bogoliubov inequality are Hermitian operators corresponding to Lagrangians L and L_M in the Jensen-Feynman inequality, then the one-to-one correspondence between (1) and (3) guarantees the validity of the

Jensen-Feynman inequality, even if the action functionals are not real (e.g. for a particle in a magnetic field).

However, both inequalities do not necessarily have the same physical content: for a system with action S it is not always possible to derive a corresponding Hamiltonian. For example, for the Fröhlich polaron (in the absence of a magnetic field) the strength of the Jensen-Feynman inequality lies in the fact that it remains valid after the elimination of the phonons, with a *retarded* effective action functional, for which no corresponding Hamiltonian representation is known. In the operator formulation, the phonon elimination can formally be realized with ordered-operator calculus, but this approach involves *non-Hermitian* effective operators in the electron variables.

Fermion systems (with parallel spins) form an important class of systems for which the Jensen-Feynman inequality is not directly applicable (whereas the Bogoliubov inequality remains valid in the Hilbert space of antisymmetric states under permutations of the particle coordinates). The reason is that the path integral for *fermions* with parallel spin, *if expressed in the full coordinate space*, is a superposition of path integrals with all possible permutations of the particle coordinates, with negative signs for all odd permutations. For *bosons*, no negative signs result from the permutations, and the application of the Jensen-Feynman inequality presents no problems. Therefore, only the many-fermion problem will be explicitly addressed below.

2 Path Integral Approach for Many-Body Systems

Recent studies on the path-integral approach to the many-body problem for a fixed number of *identical particles* by Brosens, Lemmens and Devreese [22] have allowed to calculate the Feynman-Kac functional on a state space for N indistinguishable particles, which was found by imposing an ordering on the configuration space, and the introduction of a set of boundary conditions in this state space. The path integral (in the imaginary-time variable) for identical particles was shown to be positive within this state space. This implies (see subsection 3.1 for more details) that a many-body extension of the Jensen-Feynman inequality was found, which can be used to evaluate the partition function for interacting identical particles (Ref. [22], p. 4476, reference [48]). This many-body variational principle for path integrals was applied to the study of thermodynamical properties of a spin-polarized gas of bosons (Ref. [24], abstract, Eq. (3); Ref. [25], Eq. (13)). The applicability of the variational principle as formulated in Ref. [22] for many-body problems

was discussed in relation to the analysis of correlations (Ref. [26], p. 1641) and thermodynamical properties (Ref. [27], p. 3911) of a confined gas of harmonically interacting spin-polarized fermions.

The many-body variational principle for path integrals was also used recently in order to calculate the ground-state energy and the optical absorption spectrum of a many-polaron system, confined to a quantum dot (Ref. [28], p. 306).

The remainder of this paper addresses the question *which choice of model actions is allowed* in order to treat specific systems of interacting bosons and fermions. I will give some examples, illustrating that the applicability of the many-body variational principle for path integrals crucially depends on whether or not a model system possesses the proper symmetry properties with respect to permutations of identical particles. The requirements analyzed in this article are qualitatively new as compared to the Feynman variational principle of Refs. [1–3].

3 Many-Body Variational Principle for Path Integrals

Let a many-fermion system be described by the action functional $S[\bar{\mathbf{x}}(t)]$, where $\bar{\mathbf{x}} \equiv \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ are the coordinate vectors of fermions. The partition function Z_F of a many-fermion system can be expressed as a path integral:

$$Z_F = \sum_P \frac{(-1)^P}{N!} \int d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{P\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \exp \{-S[\bar{\mathbf{x}}(t)]\}, \quad (4)$$

where the summation is over all elements P of the permutation group. The weight $(-1)^P$ is the character of the representation, i.e. +1 for even permutations and -1 for odd permutations (for the case of fermions).

3.1 Model Systems with Local Potentials

In Ref. [22], a many-body problem was analyzed for a local potential $V(\bar{\mathbf{x}})$ (including interparticle interactions) with the action functional in the imaginary-time representation

$$S[\bar{\mathbf{x}}(t)] = \frac{1}{\hbar} \int_0^{\hbar\beta} dt \left[\frac{m}{2} \sum_{j=1}^N \dot{\mathbf{x}}_j^2(t) + V(\bar{\mathbf{x}}(t)) \right]. \quad (5)$$

Note that this action is invariant under the permutations of any two fermions at any (imaginary) time, since the potential cannot make a distinction between identical particles.

If the potential $V(\bar{\mathbf{x}}(t))$ is invariant with respect to the permutations of the *Cartesian components of the particle coordinates* [22,23], the many-body propagator is obtained by four *independent* processes per pair of particles, defined on a state space (D_n^3 in the notations of Ref. [22]) with well-defined boundary conditions (see Eqs. (4.16) and (4.17) in Ref. [22] for details). These processes were shown to give *positive* contributions to the propagator. Hence, the propagator itself is *positive* on D_n^3 , implying that the Jensen-Feynman inequality can be used to estimate the partition function for interacting identical particles.^b The partition function can then be represented (apart from a normalizing factor) in the form

$$Z_F = \int_{D_n^3} d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \exp \{-S[\bar{\mathbf{x}}(t)]\}, \quad \bar{\mathbf{x}}(t) \in D_n^3. \quad (6)$$

Consider now a model system with the action functional in the imaginary-time representation

$$S_M[\bar{\mathbf{x}}(t)] = \frac{1}{\hbar} \int_0^{\hbar\beta} dt \left[\frac{m}{2} \sum_{j=1}^N \dot{\mathbf{x}}_j^2(t) + V_M(\bar{\mathbf{x}}(t)) \right], \quad (7)$$

where the *model* potential $V_M(\bar{\mathbf{x}})$ contains some variational parameters, and allows for an analytical calculation of the path integral. Suppose furthermore that it is invariant with respect to the permutations of the components of the particle positions. The path-integral expression for the partition function of the model system is thus:

$$Z_M = \int_{D_n^3} d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \exp \{-S_M[\bar{\mathbf{x}}(t)]\}, \quad \bar{\mathbf{x}}(t) \in D_n^3. \quad (8)$$

One can represent (6) as follows:

^bIf the potential is only invariant under permutations of the particle coordinates, the subprocesses are not linearly independent, and transitions between the subprocesses have to be taken into account. The actual analysis in terms of the state space D_n^3 is then only feasible in practice for a very limited number of fermions. In this case, an overcomplete space covered by all even permutations of the particle coordinates is more appropriate.

$$\begin{aligned}
Z_F &\equiv \int_{D_n^3} d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \exp \{ -S_M[\bar{\mathbf{x}}(t)] - (S[\bar{\mathbf{x}}(t)] - S_M[\bar{\mathbf{x}}(t)]) \} \\
&= Z_M \langle \exp \{ -(S[\bar{\mathbf{x}}(t)] - S_M[\bar{\mathbf{x}}(t)]) \} \rangle_{S_M}.
\end{aligned} \tag{9}$$

Here, the angular brackets denote the quantum statistical expectation value:

$$\langle \bullet \rangle_{S_M} \equiv [Z_M]^{-1} \int_{D_n^3} d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \bullet \exp \{ -S_M[\bar{\mathbf{x}}(t)] \}, \tag{10}$$

analogously to Eq. (2). The key element of this definition is that the path integrals in Eqs. (8-10) are defined on the *same* state space D_n^3 , which stems from the symmetry properties of the true action $S[\bar{\mathbf{x}}(t)]$.

Taking into account that the propagators are positive on the domain D_n^3 , one obtains the inequality

$$Z_F \geq Z_M \exp \{ -\langle S[\bar{\mathbf{x}}(t)] - S_M[\bar{\mathbf{x}}(t)] \rangle_{S_M} \}, \tag{11}$$

which is readily converted into an upper bound for the free energy

$$F_F \leq F_M + \frac{1}{\beta} \langle S[\bar{\mathbf{x}}(t)] - S_M[\bar{\mathbf{x}}(t)] \rangle_{S_M}. \tag{12}$$

This *many-body variational principle for path integrals* is formally very similar to the Jensen-Feynman inequality (1). The difference between the Eqs. (12) and (1) lies in the definition of the expectation values. In (11) and (12) the expectation value (10) is defined over a subdomain D_n^3 of the configuration space, whereas the expectation value (2) in (1) is defined over the full configuration space. However, because the symmetry properties allow to unfold the state space into the full configuration space, the restriction to the state space can be omitted in the calculation. The state space (D_n^3 in this example) only serves the goal to check whether the action and the trial action have the correct symmetry properties.

3.2 Model Systems with Retarded Effective Interactions

The action functional $S[\bar{\mathbf{x}}(t)]$ of the system under study can contain a retarded effective interaction. This is the case, e.g. for a system of N polarons after the phonon variables have been integrated out. Such many-fermion systems substantially differ from those considered above in subsection 3.1, and a different class of model systems seems to be appropriate. For this purpose,

we consider a model system consisting of fermions interacting with *auxiliary fictitious particles*.

Such a model system has the action functional in the imaginary-time representation

$$S_M = \frac{1}{\hbar} \int_0^{\hbar\beta} L_M(t) dt. \quad (13)$$

The model "Lagrangian" is chosen in the form:

$$L_M = L_F(\bar{\mathbf{x}}) + L_f(\bar{\mathbf{y}}) + L_{F-f}(\bar{\mathbf{x}}, \bar{\mathbf{y}}), \quad (14)$$

where $\{\mathbf{x}_j\} \equiv \bar{\mathbf{x}}$ are the coordinate vectors of the fermions, and $\{\mathbf{y}_j\} \equiv \bar{\mathbf{y}}$ are the coordinate vectors of the fictitious particles. The "Lagrangians" $L_F(\bar{\mathbf{x}})$, $L_f(\bar{\mathbf{y}})$ and $L_{F-f}(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ describe fermions, fictitious particles, and the interaction between the fermions and the fictitious particles, respectively. Here, the discussion is limited to the case of *distinguishable* fictitious particles for the sake of simplicity.

The partition function Z_M of the model system can be written as the following path integral:

$$Z_M = \sum_P \frac{\xi^P}{N!} \int d\bar{\mathbf{x}} \int d\bar{\mathbf{y}} \int_{\bar{\mathbf{x}}}^{P\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \int_{\bar{\mathbf{y}}}^{P\bar{\mathbf{y}}} D\bar{\mathbf{y}}(t) \exp\{-S_M\}. \quad (15)$$

Integrating out the coordinates of the fictitious particles, the partition function (15) takes the form

$$Z_M = Z_0 Z_f, \quad (16)$$

$$Z_0 = \sum_P \frac{\xi^P}{N!} \int d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{P\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \exp\{-S_0[\bar{\mathbf{x}}(t)]\}, \quad (17)$$

where Z_f is the partition function of the system of fictitious particles:

$$Z_f = \int d\bar{\mathbf{y}} \int_{\bar{\mathbf{y}}}^{P\bar{\mathbf{y}}} D\bar{\mathbf{y}}(t) \exp\left\{-\frac{1}{\hbar} \int_0^{\hbar\beta} L_f(\bar{\mathbf{y}}) dt\right\}, \quad (18)$$

and $S_0[\bar{\mathbf{x}}(t)]$ is an effective action which only depends on the fermion variables,

$$S_0[\bar{\mathbf{x}}(t)] \equiv \frac{1}{\hbar} \int_0^{\hbar\beta} (L_F(t) dt + \Phi_0[\bar{\mathbf{x}}(t)]). \quad (19)$$

The last term in (19) is referred to as the *influence phase* of the fictitious particles. It is defined as

$$\begin{aligned} \exp \{-\Phi_0 [\bar{\mathbf{x}}(t)]\} &\equiv \\ &\equiv [Z_f]^{-1} \int d\bar{\mathbf{y}} \int_{\bar{\mathbf{y}}}^{\bar{\mathbf{y}}} D\bar{\mathbf{y}}(t) \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} [L_f(\bar{\mathbf{y}}) + L_{e-f}(\bar{\mathbf{x}}, \bar{\mathbf{y}})] dt \right\}. \end{aligned} \quad (20)$$

For interaction Lagrangians $L_{e-f}(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ which are quadratic in $\bar{\mathbf{y}}$, the influence phase can be shown to take the form of a *retarded effective interaction*:

$$\Phi_0 [\bar{\mathbf{x}}(t)] = \int_0^{\hbar\beta} dt \int_0^{\hbar\beta} ds K(t, s) \mathbf{X}(t) \cdot \mathbf{X}(s), \quad (21)$$

where $K(t, s)$ depends on two time variables (t, s) , while $\mathbf{X}(t)$ is a linear function of the fermion coordinates $\bar{\mathbf{x}}(t)$ (see the next section for specific examples).

If the action functional $S[\bar{\mathbf{x}}(t)]$ of the system under study satisfies the permutation symmetry conditions discussed in the previous subsection, its partition function can be represented as a path integral over the space state D_n^3 in the form (6):

$$\begin{aligned} Z_F &\equiv \int_{D_n^3} d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \exp \{-S_0[\bar{\mathbf{x}}(t)] - (S[\bar{\mathbf{x}}(t)] - S_0[\bar{\mathbf{x}}(t)])\}, \\ &= Z_0 \langle \exp \{-(S[\bar{\mathbf{x}}(t)] - S_0[\bar{\mathbf{x}}(t)])\} \rangle_{S_0}. \end{aligned} \quad (22)$$

If the model action $S_0[\bar{\mathbf{x}}(t)]$ also possesses the above symmetry properties with respect to permutations, its partition function (17) can also be written in the form of a path integral over the domain D_n^3 ,

$$Z_0 = \int_{D_n^3} d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \exp \{-S_0[\bar{\mathbf{x}}(t)]\}, \quad (23)$$

and the quantum statistical expectation value in (22) can be defined as

$$\langle \bullet \rangle_{S_0} \equiv [Z_0]^{-1} \int_{D_n^3} d\bar{\mathbf{x}} \int_{\bar{\mathbf{x}}}^{\bar{\mathbf{x}}} D\bar{\mathbf{x}}(t) \bullet \exp \{-S_0[\bar{\mathbf{x}}(t)]\}, \quad (24)$$

analogously to the definition (2). The fact that the propagators are positive in the integration domain D_n^3 guarantees that the inequality

$$Z_F \geq Z_0 \exp \{ \langle -(S[\bar{\mathbf{x}}(t)] - S_0[\bar{\mathbf{x}}(t)]) \rangle_{S_0} \} \quad (25)$$

holds true. Consequently we obtain an upper bound for the free energy similar to Eq. (12):

$$F_F \leq F_v \equiv F_0 + \frac{1}{\beta} \langle S[\bar{\mathbf{x}}(t)] - S_0[\bar{\mathbf{x}}(t)] \rangle_{S_0}, \quad (26)$$

where $F_0 = -(\ln Z_0)/\beta$ and $F_F = -(\ln Z_F)/\beta$.

Because of the presence of the retarded action (21) resulting from the elimination of the fictitious particles, one should guarantee that the functional $S_0[\bar{\mathbf{x}}(t)]$ has the required symmetry with respect to permutations which allows that the many-body processes, related to the quantum statistical expectation value (24), are restricted to the state space D_n^3 at any time. This condition is an essential ingredient for the justification of the *many-body variational principle for path integrals* (12). Like in the case (12) of local potentials, the upper bound (26) to the free energy is formally very similar to the Jensen-Feynman inequality (1): the difference lies again in the definition of the expectation values. In (25) and (26) the expectation value (24) is defined over the subdomain D_n^3 of the configuration space, whereas the expectation value (2) in (1) is defined over the full configuration space. However, like in the previous subsection, the state space (D_n^3 in this case) is only needed to check whether the symmetry of the action and the trial action allows to apply the inequality. The calculation can be performed over the total configuration space by unfolding the state space.

4 Examples: Non-Interacting Fermions as a Test Case for a Many-Body Variational Principle with Path Integrals

4.1 Model System with Each Fermion Harmonically Interacting with One Fictitious Particle

In order to illustrate the applicability of the many-body variational principle for path integrals (26), and in particular the need of the correct symmetry requirements for the model action, we first consider a very simple system of $N = \sum_{\sigma=\pm 1/2} N_\sigma$ non-interacting fermions, described by the Lagrangian

$$L_F = \frac{m}{2} \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} \dot{\mathbf{x}}_{j,\sigma}^2, \quad (27)$$

where N_σ is the number of electrons with spin component $\sigma = \pm 1/2$. The ground-state energy for the system with the classical "Lagrangian" (27) is

elementary:

$$E = \frac{3}{5} E_F, \quad E_F \equiv \frac{\hbar^2 k_F^2}{2m}, \quad (28)$$

with the Fermi wave number k_F and the Fermi energy E_F .

We now examine whether the many-body variational principle for path integrals (26) indeed provides an upper bound to the correct ground-state energy (28).

For the model system (14) we choose a Lagrangian L_M , in which each fermion harmonically interacts with *one fictitious particle*. We do this uncritically, deliberately overlooking the problem of the required symmetry of the state space of the model action and choose

$$L_M = \frac{m}{2} \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} \dot{\mathbf{x}}_{j,\sigma}^2 + \frac{M}{2} \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} \dot{\mathbf{y}}_{j,\sigma}^2 + \frac{k}{2} \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} (\mathbf{x}_{j,\sigma} - \mathbf{y}_{j,\sigma})^2, \quad (29)$$

where M is the mass of the fictitious particles, and k is the force constant of the elastic bond between a fermion and its accompanying fictitious particle. We introduce the following notations:

$$w = \sqrt{\frac{k}{M}}, \quad v = \sqrt{\frac{k}{\mu}}, \quad \mu = \frac{mM}{m+M}.$$

For this particular case, the elimination of the fictitious particles leads to the influence phase (20):

$$\begin{aligned} \Phi_0 [\bar{\mathbf{x}}(t)] = & -\frac{Mw^3}{8\hbar} \int_0^{\hbar\beta} dt \int_0^{\hbar\beta} ds \frac{\cosh w \left(|t-s| - \frac{\hbar\beta}{2} \right)}{\sinh \frac{1}{2} \hbar\beta w} \\ & \times \sum_{\sigma=\pm 1} \sum_{j=1}^{N_\sigma} [\mathbf{x}_{j,\sigma}(t) - \mathbf{x}_{j,\sigma}(s)]^2, \end{aligned} \quad (30)$$

with a quadratic effective retarded self-interaction for each fermion.

Then we apply the many-body variational principle for path integrals (26) naively to the chosen model system. The free energy F_v from this inequality is calculated analytically. The parameters M and k of the model "Lagrangian" (14) are then found by minimizing the value of the supposed upper bound

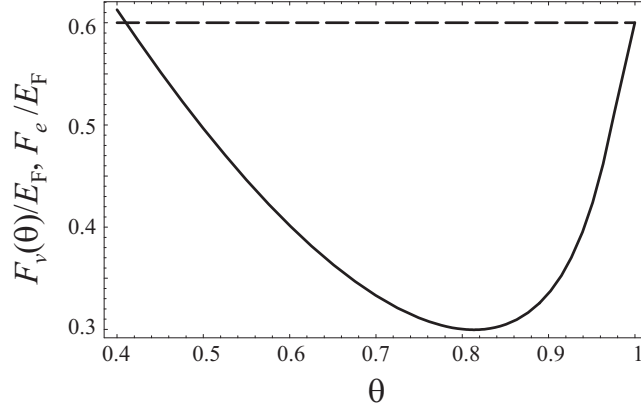


Figure 1. Free energy $F_v(\theta)$ ($\theta = w/v$) for non-interacting fermions [according to the inequality (26)] (solid line) compared with the exact free energy F_e (dashed line) at the optimal value of the variational frequency parameter $v = 1.83$ (in units of $E_F/\hbar$, where $E_F = 49$ meV is the Fermi energy for the density $n_e = 5 \times 10^{19} \text{ cm}^{-3}$, rather arbitrarily chosen); $\beta = 16.4$ (in units of E_F^{-1}). The mass m of a fermion is taken to be m_0 , the bare electron mass.

F_v . This calculation has shown that the many-body variational principle for path integrals (26) is *violated* for this model system, as clearly illustrated in Fig. 1.

What was wrong in the above approach? The answer is immediate: “Of course, we forgot to check whether the model action has the required symmetry!” The model “Lagrangian” (29) is *not symmetrical* with respect to the permutations of the fermion coordinates $\mathbf{x}_{j,\sigma}$, because each of them is linked with a particular fictitious particle.

4.2 Model System with Each Fermion Having Equal Elastic Bonds with All Fictitious Particles

As a second example, we study the many-body variational principle for path integrals (26) for $N = \sum_{\sigma=\pm 1/2} N_\sigma$ non-interacting electrons in a parabolic confinement described by the Lagrangian (in the imaginary-time variable)

$$L = \frac{m}{2} \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} (\dot{\mathbf{x}}_{j,\sigma}^2 + \Omega_0^2 \mathbf{x}_{j,\sigma}^2). \quad (31)$$

In particular, the case $\Omega_0 \rightarrow 0$ will be analyzed (a translationally invariant system, equivalent to the free-fermion model in subsection 4.1). The exact ground-state energy of this system can immediately be written down:

$$E_0(\Omega_0, N) = \hbar\Omega_0 \sum_{\sigma=\pm 1/2} \left\{ N_\sigma \left[n_0(N_\sigma) + \frac{5}{2} \right] - \frac{1}{4} N_0(n_0(N_\sigma)) [n_0(N_\sigma) + 4] \right\}. \quad (32)$$

Here, $n_0(N_\sigma)$ is the number of the upper fully occupied energy level for N_σ fermions with spin component σ . The number of fermions in all closed shells $[N_0(n_0(N_\sigma)) \leq N_\sigma]$ is

$$N_0(n_0) \equiv \sum_{n=0}^{n_0} \frac{(n+1)(n+2)}{2} = \frac{1}{6} (n_0+1)(n_0+2)(n_0+3). \quad (33)$$

A model system is now considered, which consists of particles in a harmonic confinement potential with elastic interparticle interactions as studied in Ref. [29]. The "Lagrangian" of this model system is chosen in the form

$$\begin{aligned} L_M = & \frac{m}{2} \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} (\dot{\mathbf{x}}_{j,\sigma}^2 + \Omega^2 \mathbf{x}_{j,\sigma}^2) + \frac{m\omega^2}{4} \sum_{\sigma,\tau=\pm 1/2} \sum_{j,l=1}^{N_\sigma} (\mathbf{x}_{j,\sigma} - \mathbf{x}_{l,\tau})^2 \\ & + \frac{M}{2} \sum_{l=1}^{N_B} (\dot{\mathbf{y}}_l^2 + \Omega_B^2 \mathbf{y}_l^2) + \frac{M\omega_B^2}{4} \sum_{j=1}^{N_\sigma} \sum_{l=1}^{N_B} (\mathbf{y}_j - \mathbf{y}_l)^2 \\ & + \frac{k}{2} \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} \sum_{l=1}^{N_B} (\mathbf{x}_{j,\sigma} - \mathbf{y}_l)^2. \end{aligned} \quad (34)$$

The frequencies Ω , ω , Ω_B , ω_B , the mass M of fictitious particles, and the force constant k are treated as *variational parameters*. It is important to stress, that in this model system each fermion has identical elastic bonds with *all* fictitious particles, and therefore permutations of the fermion coordinates leave the "Lagrangian" (34) invariant.

After integration over the paths of the fictitious particles, the partition function (16) becomes

$$Z_f = \left[\sinh \left(\frac{\hbar\beta\tilde{\Omega}_B}{2} \right) \right]^3 \left[\sinh \left(\frac{\hbar\beta\omega_B}{2} \right) \right]^{3N_B-3}, \quad (35)$$

with

$$\tilde{\Omega}_B = \sqrt{\Omega^2 + \frac{kN_B}{m}}, \quad w_B = \sqrt{\Omega^2 - N_B\omega^2 + \frac{kN_B}{m}}. \quad (36)$$

The action (19) in the imaginary time representation takes the form

$$S_0 [\bar{\mathbf{x}}(t)] \equiv S_{F0} [\bar{\mathbf{x}}(t)] + \Phi_0 [\bar{\mathbf{x}}(t)], \quad (37)$$

$$S_{F0} [\bar{\mathbf{x}}(t)] = \frac{1}{\hbar} \int_0^{\hbar\beta} dt \left[\sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} \frac{m}{2} \dot{\mathbf{x}}_{j,\sigma}^2(t) + \frac{m\omega^2 N^2}{2} \mathbf{X}^2 \right], \quad (38)$$

with the coordinate

$$\mathbf{X} \equiv \sum_{\sigma=\pm 1/2} \sum_{j=1}^{N_\sigma} \mathbf{x}_{j,\sigma}, \quad (39)$$

and the "influence phase" (20) for the present model becomes

$$\Phi_0 [\bar{\mathbf{x}}(t)] = \frac{k^2 N^2 N_B^2}{4m_B \hbar \tilde{\Omega}_B} \int_0^{\hbar\beta} dt \int_0^{\hbar\beta} ds \frac{\cosh \left[\tilde{\Omega}_B \left(|t-s| - \frac{\hbar\beta}{2} \right) \right]}{\sinh \left(\frac{\hbar\beta \tilde{\Omega}_B}{2} \right)} \mathbf{X}(t) \cdot \mathbf{X}(s). \quad (40)$$

As distinct from (29), this model Lagrangian (34) is *invariant with respect to permutations of the components of all fermion coordinates*. Hence, for the chosen model system the symmetry conditions on the action are fulfilled (as formulated in subsection 4.1), which ensures the validity of the many-body variational principle for path integrals (26).

The functional [resulting from (26) at zero temperature] for the ground-state energy of N fermions in a parabolic confinement with the confinement frequency Ω_0 takes the form

$$E_v(\Omega_1, \Omega_2, w, w_B) = \hbar \left\{ \frac{\Omega_0^2 + w^2}{2w^2} \left[\tilde{E}(w, N) - \frac{3}{2}w \right] + \frac{3}{2}(\Omega_1 + \Omega_2 - w_B) \right. \\ \left. + \frac{3}{4}(\Omega_0^2 - \Omega_1^2 - \Omega_2^2 + w_B^2) \sum_{i=1}^2 \frac{a_i^2}{\Omega_i} + \frac{3\gamma^2}{4w_B} \sum_{i=1}^2 \frac{a_i^2}{\Omega_i(\Omega_i + w_B)} \right\}, \quad (41)$$

where the following notations are used:

$$a_1 = \left(\frac{\Omega_1^2 - w_B^2}{\Omega_1^2 - \Omega_2^2} \right)^{1/2}, \quad a_2 = \left(\frac{w_B^2 - \Omega_2^2}{\Omega_1^2 - \Omega_2^2} \right)^{1/2}, \quad (42)$$

$$\gamma = [(\Omega_1^2 - w_B^2)(w_B^2 - \Omega_2^2)]^{1/2}, \quad (43)$$

and

$$\begin{aligned} \tilde{E}(w, N) = \hbar w \sum_{\sigma=\pm 1/2} \left\{ N_{\sigma} \left[n_0(N_{\sigma}) + \frac{5}{2} \right] \right. \\ \left. - \frac{1}{4} N_0(n_0(N_{\sigma})) [n_0(N_{\sigma}) + 4] \right\}. \end{aligned} \quad (44)$$

The difference between the upper bound to the ground-state energy (41) and the exact ground-state energy (32) is clearly positive:

$$E_v(\Omega_1, \Omega_2, w, w_B) - E_0(\Omega_0, N) = \frac{3\hbar}{4} \frac{(\Omega_0 - z)^2}{z}, \quad (45)$$

with

$$z = \frac{\Omega_1 \Omega_2 + \Omega_0^2}{\Omega_1 + \Omega_2}, \quad (46)$$

and it follows from (45) that the inequality

$$E_v(\Omega_1, \Omega_2, w, w_B) \geq E_0(\Omega_0, N) \quad (47)$$

holds true for any values of the variational parameters, *in accordance with the many-body variational principle for path integrals* (26).

The minimal value of the functional (41), which is achieved at $z = \Omega_0$, coincides with the exact ground-state energy (32) of N non-interacting fermions in a parabolic confinement with the frequency Ω_0 . This result confirms the applicability of the many-body variational principle for path integrals (26) to the many-fermion system under consideration, with a model system, whose "Lagrangian" (34) is symmetric under the permutations of the components of the fermion positions.

Thus, the many-body variational principle for path integrals (26) is satisfied for a system of non-interacting fermions in a parabolic confinement potential (including the translationally invariant case $\Omega_0 = 0$), when the model system with the "Lagrangian" (34) is considered.

Clearly, the Lagrangian (31) was not considered for its own sake, since it is trivial to treat. It was merely presented as a test case to illustrate our many-body variational principle for path integrals with two non-trivial trial actions which elucidate the crucial role of the correct symmetry requirements.

5 Conclusions

Summarizing, the applicability of the many-body variational principle for path integrals (26) crucially depends on the symmetry of the model system with respect to permutations of identical particles.

The invariance of the action functional $S_0[\bar{\mathbf{x}}(t)]$ and of the model action with respect to permutations of components of the positions of any two fermions at any (imaginary) time ensures that the fermion propagator is positive on the state space D_n^3 .

It should be noted that these rather stringent symmetry conditions are used as an example. A more general variational principle for identical particles, not limited to D_n^3 but to a much larger subspace of the configuration space, will be presented in future publications.

The main result of the present analysis is that a many-body variational principle for path integrals (26) can be found in the framework of the many-body path-integral approach, even for retarded effective interactions, provided that the model action and the true action have the appropriate symmetry properties under permutations of the particle coordinates.

Acknowledgments

I like to thank V.M. Fomin for interesting discussions during the preparation of the manuscript, and S.N. Klimin for assistance with the numerical aspects. Stimulating interaction with F. Brosens and L. Lemmens in the frame of our collaborations on path integrals for many-body systems is gratefully acknowledged. This work has been supported by the BOF NOI (UA-UIA), GOA BOF UA 2000, IUAP, FWO-V projects G.0287.95, G.0071.98, G.0274.01N and the W.O.G. WO.025.99N (Belgium).

References

- [1] R.P. Feynman, *Phys. Rev.* **97**, 660 (1955).
- [2] R.P. Feynman, *Statistical Mechanics. A Set of Lectures* (W.A. Benjamin, Reading, Massachusetts, 1972).
- [3] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
- [4] R.P. Feynman and H. Kleinert, *Phys. Rev. A* **34**, 5080 (1986).

- [5] L.S. Schulman, *Techniques and Applications of Path Integration* (J. Wiley & Sons, New York, 1981).
- [6] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [7] R. Giachetti and V. Tognetti, *Phys. Rev. Lett.* **55**, 912 (1985).
- [8] R. Giachetti and V. Tognetti, *Phys. Rev. B* **33**, 7647 (1986).
- [9] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [10] G. Verbist, F.M. Peeters, and J.T. Devreese, *Solid State Communications* **76**, 1005 (1990).
- [11] G. Verbist, F.M. Peeters, and J.T. Devreese, *Phys. Rev. B* **43**, 2712 (1991).
- [12] R.W. Hellwarth and P.M. Platzman, *Phys. Rev.* **128**, 1599 (1962).
- [13] J.T. Marshall and M.S. Chawla, *Phys. Rev. B* **2**, 4283 (1970).
- [14] V.I. Sheka, L.S. Khazan, and E.V. Mozdor, *Fiz. Tverd. Tela (Leningrad)* **18**, 3240 (1976) [*Sov. Phys. Solid State* **18**, 1890 (1976)].
- [15] F.M. Peeters and J.T. Devreese, *Phys. Rev. B* **25** 7281, 7302 (1982).
- [16] D.M. Larsen, *Phys. Rev. B* **32**, 2657 (1985).
- [17] J.T. Devreese and F. Brosens, *Phys. Rev. B* **45**, 6459 (1992).
- [18] M. Bachmann, H. Kleinert, and A. Pelster, *Phys. Rev. A* **62**, 52509 (2000).
- [19] I.A. Kvasnikov, *Akad. Nauk SSSR* **110**, 755 (1956).
- [20] D. Ruelle, *Statistical Mechanics: Rigorous Results* (Benjamin, New York, 1969).
- [21] L.I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955).
- [22] L.F. Lemmens, F. Brosens, and J.T. Devreese, *Phys. Rev. E* **53**, 4467 (1996).
- [23] L.F. Lemmens, F. Brosens, and J.T. Devreese, *Phys. Rev. E* **55**, 7813 (1997) (Erratum).
- [24] J. Tempere, F. Brosens, L.F. Lemmens, and J.T. Devreese, *Solid State Communications* **107**, 51 (1998).
- [25] J. Tempere, F. Brosens, L.F. Lemmens, and J.T. Devreese, *Phys. Rev. A* **61**, 043605 (2000).
- [26] F. Brosens, J.T. Devreese, and L.F. Lemmens, *Phys. Rev. E* **58**, 1634 (1998).
- [27] S. Foulon, F. Brosens, J.T. Devreese, and L.F. Lemmens, *Phys. Rev. E* **59**, 3911 (1999).

- [28] J.T. Devreese, S.N. Klimin, V.M. Fomin, and F. Brosens, *Solid State Communications* **114**, 305 (2000).
- [29] F. Brosens, J.T. Devreese, and L.F. Lemmens, *Phys. Rev. E* **55**, 227 (1997); *ibid.* **55**, 6795 (1997); *ibid.* **58**, 1634 (1998).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

VARIATIONAL PERTURBATION THEORY: A POWERFUL METHOD FOR DERIVING STRONG-COUPLING EXPANSIONS

W. JANKE

*Institut für Theoretische Physik, Universität Leipzig,
Augustusplatz 10/11, D-04109 Leipzig, Germany
E-mail: wolfgang.janke@itp.uni-leipzig.de*

In this contribution, an overview of Kleinert's variational perturbation theory will be given. Contrary to standard perturbative approaches, this method yields converging approximations uniformly in the coupling strength of anharmonic terms. In standard quantum mechanics, the simplest example is the one-dimensional anharmonic oscillator. It is shown how variational perturbation theory can be exploited to calculate the convergent *strong*-coupling expansion from the divergent *weak*-coupling perturbation series. By recalling the Duru-Kleinert path-integral solution of the three-dimensional hydrogen problem, I make use of the mapping between Coulomb systems in three dimensions and oscillator systems in four dimensions to also derive the quantum mechanical strong-coupling expansion of the ground-state energy of the hydrogen problem allowing for an isotropic *gr*-perturbation.

1 Introduction

Most perturbation expansions in physics are divergent asymptotic series whose large-order coefficients grow factorially. Typical examples include the perturbative calculation of the anomalous magnetic moment of the electron in quantum electrodynamics, field theoretical ϵ -expansions of critical exponents in statistical physics, and the low-field expansions of the Stark and Zeeman effects in atomic physics.

The study of the large-order behavior of such asymptotic series and the problem of developing appropriate resummation techniques has a long history in Hagen Kleinert's research group. In fact, shortly after having finished my

diploma thesis under his guidance and having jointly solved the “particle-in-a-box” problem with path-integral techniques [1], this was one of the first research topics we started working on together. After an intense initial period we returned to this problem many times [2–4] and collected over the years quite an extensive set of lecture notes [5]. While working on systematic improvements of the variational approximation [6,7] underlying an effective classical description of quantum statistics, Kleinert [8] realized in 1993 that this could lead to a powerful resummation scheme for divergent perturbation series. Remembering our joint work on the variational scheme [9,10], here our interests met again and we began to work [11–13] on what will be briefly described in the following.

The paradigm for a divergent asymptotic perturbation series expansion is the ground-state energy $E^{(0)}(g)$ of an anharmonic oscillator with potential $V(x) = \omega^2 x^2/2 + gx^4/4$ ($\omega^2, g > 0$). The weak-coupling Rayleigh-Schrödinger perturbation series takes the form

$$E^{(0)}(g) = \omega \sum_{k=0}^{\infty} E_k^{(0)} \left(\frac{g/4}{\omega^3} \right)^k, \quad (1)$$

where the coefficients $E_k^{(0)} = 1/2, 3/4, -21/8, 333/16, -30885/128, \dots$ can be shown to grow asymptotically as [14]

$$E_k^{(0)} = -(1/\pi)(6/\pi)^{1/2}(-3)^k k^{-1/2} k! [1 + \mathcal{O}(1/k)]. \quad (2)$$

Only for very weak couplings g , a direct evaluation of the power-series expansion truncated at a finite order $k \propto 1/g$ can yield a reasonably good approximation [5] with an error of the order of $\exp(-\text{const}/g)$. At larger couplings the partial sums become very erratic and hence completely useless, unless some resummation procedure is applied. The accuracy of standard techniques such as Padé or Borel resummation [5] also deteriorates, however, quite rapidly in the strong-coupling limit, where $E^{(0)}(g)$ has an expansion of the form

$$E^{(0)}(g) = (g/4)^{1/3} \left[\alpha_0 + \alpha_1 (4\omega^3/g)^{2/3} + \alpha_2 (4\omega^3/g)^{4/3} + \dots \right]. \quad (3)$$

2 Variational Perturbation Theory

Variational perturbation theory [7,8,15], on the other hand, yields a sequence of *exponentially* fast converging approximations *uniformly* in g (Refs. [11–

13,16]), and thus provides a powerful method for calculating the strong-coupling expansion coefficients α_i in (3). As mentioned in the historical remarks above, the starting point of this approach was the variational principle for evaluating quantum partition functions in the path-integral formulation [6,7]. While in many applications the accuracy was found to be excellent over a wide range of temperatures [10], slight deviations from exact [6,7,9,10] or computer simulation [17] results at very low temperatures motivated a systematic study of higher-order corrections [7,8,15].

In the zero-temperature limit the higher-order calculations simplify and lead to a resummation scheme for the energy eigenvalues which can be summarized as follows. First, the harmonic term of the potential is decomposed into $\omega^2 x^2 = \Omega^2 x^2 + (\omega^2 - \Omega^2) x^2$, where Ω is a trial frequency to be optimized later, and the potential is rewritten as $V(x) = \Omega^2 x^2/2 + g(-2\sigma x^2/\Omega + x^4)/4$, with $\sigma = \Omega(\Omega^2 - \omega^2)/g$. Keeping σ fixed, one then performs a perturbation expansion of $\hat{E}_N^{(0)} \equiv E_N^{(0)}/\Omega$ in powers of $\hat{g} \equiv g/\Omega^3$,

$$\hat{E}_N^{(0)}(\hat{g}, \sigma) = \sum_{k=0}^N \varepsilon_k^{(0)}(\sigma) (\hat{g}/4)^k, \quad (4)$$

$$\varepsilon_k^{(0)}(\sigma) = \sum_{j=0}^k E_j^{(0)} \binom{(1-3j)/2}{k-j} (-4\sigma)^{k-j}, \quad (5)$$

which can be readily derived by inserting $\omega = (\Omega^2 - g\sigma/\Omega)^{1/2} = \Omega(1 - \hat{g}\sigma)^{1/2}$ into the original perturbation series (1) and reexpanding in powers of $\hat{g}$. By construction the truncated power series $W_N(g, \Omega) \equiv \Omega \hat{E}_N^{(0)}(\hat{g}, \sigma)$ becomes independent of Ω in the limit $N \rightarrow \infty$. At any finite order, however, it *does* depend on Ω , the approximation having its fastest speed of convergence where it depends least on Ω , i.e. at points where $\partial W_N/\partial \Omega = 0$. If we denote the order-dependent optimal value of Ω by Ω_N , the quantity $W_N(g, \Omega_N)$ is the new approximation to $E^{(0)}(g)$.

While W_N is a polynomial in Ω of degree $3N$ with g -dependent coefficients, it can be proven [11] that its derivative with respect to Ω admits a compact, factorized representation, namely $\partial W_N/\partial \Omega = (\hat{g}/4)^N P_N(\sigma)$, where $P_N(\sigma) = -2d\varepsilon_{N+1}^{(0)}(\sigma)/d\sigma$ is a polynomial of degree N in σ . This observation simplifies the calculations considerably and shows that the optimal solutions Ω_N depend only trivially on g through $\sigma_N = \Omega_N(\Omega_N^2 - \omega^2)/g$. The order-dependent

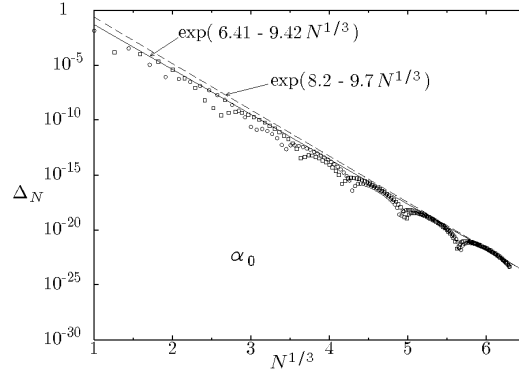


Figure 1. Exponential accuracy as shown by the rapidly decreasing error $\Delta_N = |\alpha_0^{(N)} - \alpha_0|$ of the N th approximant (7) for the strong-coupling coefficient α_0 , where $N = 1, \dots, 251$ is the order of the weak-coupling expansion.

optimal values of σ are found to be well fitted by

$$\sigma_N = cN \left(1 + 6.85/N^{2/3} \right), \quad (6)$$

with $c = 0.186\,047\,272\dots$ determined analytically (cf. Section 3). This observation suggested [12] that the variational resummation scheme can be taken directly to the strong-coupling limit by introducing the reduced frequency $\hat{\omega} = \omega/\Omega$, rewriting the approximation as $W_N = (g/\hat{g})^{1/3} w_N(\hat{g}, \hat{\omega}^2)$, and expanding the function $w_N(\hat{g}, \hat{\omega}^2)$ in powers of $\hat{\omega}^2 = (\omega^3/g)^{2/3} \hat{g}^{2/3}$. As a result, for $E^{(0)}(g) \approx W_N(g)$ we find an expansion of the form (3) with coefficients [12]

$$\alpha_n^{(N)} = (\hat{g}/4)^{(2n-1)/3} \sum_{k=0}^N (-1)^{k+n} \sum_{j=0}^{k-n} E_j^{(0)} \binom{(1-3j)/2}{k-j} \binom{k-j}{n} (-\hat{g}/4)^j. \quad (7)$$

If this is evaluated at $\hat{g} = 1/\sigma_N$ with σ_N given in (6), we obtain the exponentially fast approach to the exact limit as shown in Fig. 1 for α_0 . Notice the oscillatory modulations. The computation of the higher-order coefficients α_n proceeds similarly and the results up to $n = 22$ can be found in Table 1 of Ref. [12]. Our results for α_0 and α_1 based on the first 251 weak-coupling expansion coefficients are compared in Table 1 with other recent estimates. So far the more mathematically motivated resummation scheme of Ref. [18]

Table 1. Estimates of the leading strong-coupling expansion coefficients α_n .

α_0	Ref.
0.667 986 259 155 777 108 270 96	[12]
0.667 986 259 155 777 108 270 962 016 919 86	[18]
0.667 986 259 155 777 108 270 962 016 919 860 199 4	[19]
0.667 986 259 155 777 108 270 962 016 919 860 199 430 404 936 984 ...	[21]
α_1	Ref.
0.143 668 783 380 864 910 020 3	[12]
0.143 668 783 380 864 910 020 319 127 583 17	[18]
0.143 668 783 380 864 910 020 319 127 583 168 634 2	[19]

gives more accurate numbers than those in Ref. [12]. However, by applying extrapolation techniques to the sequences for α_n [19], the accuracy of our estimates can be further considerably improved. Specifically we employed Wynn's ϵ -algorithm [20] where the extrapolants $\epsilon_k^{(n)}$ are defined recursively by $\epsilon_k^{(-1)} = 0$, $\epsilon_k^{(0)} = W_k$, and $\epsilon_k^{(n+1)} = \epsilon_{k+1}^{(n-1)} + 1/(\epsilon_{k+1}^{(n)} - \epsilon_k^{(n)})$.

3 Convergence Behavior

It is well known that the ground-state energy $E^{(0)}(g)$ of the quartic anharmonic oscillator satisfies a subtracted dispersion relation which implies an integral representation for the perturbation coefficients [14],

$$E_k^{(0)} = \frac{4^k}{2\pi i} \int_0^{-\infty} \frac{dg}{g^{k+1}} \text{disc} E^{(0)}(g), \quad (8)$$

where $\text{disc} E^{(0)}(g) = 2i\text{Im} E^{(0)}(g - i\eta)$ denotes the discontinuity across the left-hand cut in the complex g -plane. For large k , only its $g \rightarrow 0^-$ behavior is relevant and a semiclassical calculation yields $\text{disc} E^{(0)}(g) \approx -2i\omega(6/\pi)^{1/2}(-4\omega^3/3g)^{1/2} \exp(4\omega^3/3g)$, which in turn implies the large-order behavior (2) of $E_k^{(0)}$.

The reexpanded series (5) is obtained from (1) by replacing $\omega \rightarrow \Omega(1 - \sigma\hat{g})^{1/2}$, which, in terms of the coupling constant, amounts to $\bar{g} \equiv g/\omega^3 \rightarrow \hat{g}/(1 - \sigma\hat{g})^{3/2}$. This implies [13] a dispersion relation for $\hat{E}^{(0)} \equiv E^{(0)}/\Omega$ and

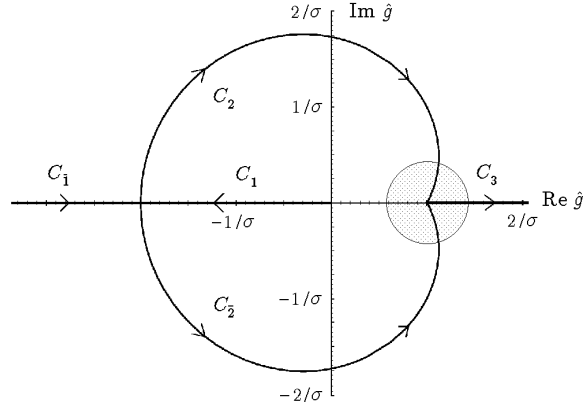


Figure 2. Cuts in the complex $\hat{g}$ -plane. The shaded area shows the circle of convergence of the strong-coupling expansion.

hence the coefficients $\varepsilon_k^{(0)}$,

$$\varepsilon_k^{(0)} = \frac{4^k}{2\pi i} \int_C \frac{d\hat{g}}{\hat{g}^{k+1}} \text{disc}_C \hat{E}^{(0)}(\hat{g}), \quad (9)$$

where $\text{disc}_C \hat{E}^{(0)}(\hat{g})$ is the discontinuity across the cuts C in the complex $\hat{g}$ -plane shown in Fig. 2. The cuts C run along the contours $C_1, C_{\bar{1}}, C_2, C_{\bar{2}}$, the images of the left-hand cut in the complex g -plane, and C_3 , originating from the square root of $1 - \sigma\hat{g}$ in the mapping from $\bar{g}$ to $\hat{g}$.

Let us now discuss the contributions of the various cuts to the N th term S_N . For the cut C_1 and the empirically observed optimal solutions $\sigma_N = cN(1 + b/N^{2/3})$, a saddle-point approximation shows [13] that this term gives a convergent contribution, $S_N(C_1) \propto e^{-[-b \log(-\gamma) + (cg)^{-2/3}]N^{1/3}}$, if one chooses $c = 0.186\,047\,272\dots$ and $\gamma = -0.242\,964\,029\dots$. Inserting the fitted value of $b = 6.85$, this yields an exponent of $-b \log(-\gamma) = 9.7$, in rough agreement with the convergence seen in Fig. 1. If this would be the only contribution, the convergence behavior could be changed at will by varying the parameter b . For $b < 6.85$, a slower convergence is indeed observed. The convergence cannot be improved, however, by choosing $b > 6.85$, since the optimal convergence is limited by the contributions of the other cuts.

The cut $C_{\bar{1}}$ is still harmless; it contributes a last term $S_N(C_{\bar{1}})$ of the negligible order $e^{-N \log N}$. The cuts $C_{2,\bar{2},3}$, however, deserve a careful con-

sideration. If they would really start at $\sigma\hat{g} = 1$, the leading behavior would be $\varepsilon_k^{(0)}(C_{2,\bar{2},3}) \propto \sigma^k$, and therefore $S_N(C_{2,\bar{2},3}) \propto (\sigma\hat{g})^N$, which would be in contradiction to the empirically observed convergence in the strong-coupling limit. The important point is that the cuts in Fig. 2 do not really reach the point $\sigma\hat{g} = 1$. There exists a small circle of radius $\Delta\hat{g} > 0$ in which $\hat{E}^{(0)}(\hat{g})$ has no singularities at all, a consequence of the fact that the strong-coupling expansion (3) converges for $g > g_s$. The complex conjugate pair of singularities gives a contribution $S_N(C_{2,\bar{2},3}) \approx e^{-N^{1/3}a \cos \theta} \cos(N^{1/3}a \sin \theta)$, with $a = 1/(|\bar{g}_s|c)^{2/3}$. By analyzing the convergence behavior of the strong-coupling series we find $|\bar{g}_s| \approx 0.160$ and $\theta \approx -0.467$, which implies an asymptotic falloff of $e^{-9.23N^{1/3}}$ for the envelope, and furthermore also explains the oscillations in the data [13].

4 Three-Dimensional Coulomb Systems

It is well known that Coulomb systems in *three* dimensions (3D) and oscillator systems in *four* dimensions (4D) are closely related to each other [22,23]. This property has been exploited in many ways; in particular it was one of the clues for the solution of the path-integral for the hydrogen atom by Duru and Kleinert [23]. One recent example is a simplified analysis of the large-order behavior of weak-coupling expansions for perturbed Coulomb systems [2,3]. Here we shall concentrate on strong-coupling expansions which are usually more difficult to derive. Specifically we consider the 3D Coulomb system with the Hamiltonian

$$H_C = \frac{1}{2}\vec{p}^2 - \frac{1}{r} + gr, \quad (10)$$

which is related to the 4D anharmonic oscillator with the Hamiltonian

$$H = \frac{1}{2}\mathbf{p}^2 + \frac{\omega^2}{2}\mathbf{x}^2 + \lambda(\mathbf{x}^2)^2. \quad (11)$$

In particular the ground-state energies of the two systems can be mapped onto each other. If E denotes the ground-state energy of H_C and ϵ is the ground-state energy of H , then the relation reads [2]

$$\epsilon = \epsilon(\omega, \lambda) = 1, \quad (12)$$

with

$$\omega^2 = -E/2, \quad \lambda = g/16. \quad (13)$$

By simple scaling arguments the energy E of the Coulomb systems is easily seen to possess a weak-coupling series expansion in powers of g ,

$$E = E_0 + E_1 g + E_2 g^2 + \dots, \quad (14)$$

and a strong-coupling expansion of the form

$$E = g^{2/3} \left[a_0 + a_1 g^{-1/3} + a_2 g^{-2/3} + \dots \right]. \quad (15)$$

Similarly, for the 4D anharmonic oscillator the weak-coupling expansion reads

$$\epsilon = \omega \left[\epsilon_0 + \epsilon_1 (\lambda/\omega^3) + \epsilon_2 (\lambda/\omega^3)^2 + \dots \right], \quad (16)$$

and the strong-coupling expansion takes the form

$$\epsilon = \lambda^{1/3} \left[\alpha_0 + \alpha_1 (\lambda/\omega^3)^{-2/3} + \alpha_2 (\lambda/\omega^3)^{-4/3} + \dots \right]. \quad (17)$$

4.1 4D Anharmonic Oscillator Strong-Coupling Expansion

To calculate the strong-coupling coefficients a_n in (15) by using the relations (10)-(13) we first have to derive the strong-coupling expansion of the 4D anharmonic oscillator. Here this task is accomplished by means of variational perturbation theory as described in the previous section for the one-dimensional case. In four dimensions the necessary input information, which is the weak-coupling perturbation coefficients ϵ_i , can also be easily generated to very high orders by applying recursion relations first discussed by Bender and Wu [14]. The present calculation is based on an expansion up to order 290. The resummation scheme of variational perturbation theory then yields again exponentially fast convergent sequences $\alpha_n^{(N)}$ for the strong-coupling coefficients α_n and, as discussed above, their accuracy may be further improved by applying standard extrapolation techniques such as Wynn's ϵ -algorithm [20]. This procedure was applied [24] to obtain the expansion coefficients shown in Table 2.

4.2 3D Coulomb Strong-Coupling Expansion

By making use of the relations (12) and (13) it is straightforward to express the expansion coefficients for the Coulomb system in terms of the expansion coefficients for the anharmonic oscillator. For the mapping between the weak-coupling series see, e.g., Ref. [2]. Similarly, by inserting the strong-coupling

Table 2. Coefficients α_n of the strong-coupling expansion (17) for the ground-state energy of the 4D anharmonic oscillator (11). The symbol # denotes the number of digits obtained.

n	α_n	#
0	3.398 150 176 027 696 746 352 787 969 422 624 006 593 241 57	45
1	0.447 038 467 415 823 402 400 410 319 616 607 612 580 077 2	43
2	-0.015 633 102 347 011 889 354 006 985 272 609 625 470 865	42
3	0.000 806 409 491 306 496 503 927 969 548 053 909 647 16	41
4	-0.000 039 561 514 296 026 965 992 526 411 214 682 179 66	41
5	0.000 001 484 265 174 534 240 244 510 299 896 097 807 89	41
6	-0.000 000 013 262 160 340 168 018 805 647 156 427 568 43	41
7	-0.000 000 004 230 227 654 282 595 813 731 333 132 587 68	41
8	0.000 000 000 479 462 248 736 079 997 207 517 503 895 6	40
9	-0.000 000 000 029 933 252 179 913 943 227 242 901 209 2	40
10	0.000 000 000 000 777 862 757 469 125 859 005 974 522 4	40
...	...	...
50	$\approx 6 \times 10^{-56}$	57

expansion (17) into (12) and using (13) to replace ω and λ by E and g , we obtain

$$\hat{\alpha}_0 + \hat{\alpha}_1 \frac{E}{g^{2/3}} + \hat{\alpha}_2 \left(\frac{E}{g^{2/3}} \right)^2 + \dots = \frac{16^{1/3}}{g^{1/3}}, \quad (18)$$

where

$$\hat{\alpha}_i \equiv \left(-\frac{16^{2/3}}{2} \right)^i \alpha_i \quad (19)$$

are rescaled strong-coupling coefficients of the anharmonic oscillator. Next we insert the strong-coupling expansion (15) of E and equate equal powers of $g^{-1/3}$. By defining the auxiliary sums

$$S_k \equiv \sum_{n=0}^k \binom{n}{k} \hat{\alpha}_n a_0^{n-k}, \quad (20)$$

Table 3. Coefficients a_n of the strong-coupling expansion (15) for the ground-state energy of the 3D Coulomb system (10) with gr -perturbation. The symbol # denotes the number of digits obtained.

n	a_n	#
0	1.855 757 081 489 238 479	19
1	-1.051 866 501 087 132 24	18
2	-0.186 184 039 301 421 2	16
3	-0.036 881 399 705 264	15
4	-0.003 821 866 269 00	14
5	0.000 631 796 785	12
6	0.000 303 537 651	12
7	0.000 008 501 572	12
8	-0.000 022 637 496	12
9	-0.000 004 975 49	11
10	0.000 001 184 37	11

where $\binom{n}{k}$ denotes the standard binomial coefficient, this leads to the equation

$$S_0 = 0, \quad (21)$$

which determines a_0 as one of the roots of S_0 , and in addition to a set of recursive relations for the higher coefficients a_k , with $k > 0$,

$$\begin{aligned}
a_1 &= 16^{1/3}/S_1, \\
a_2 &= -a_1^2 S_2/S_1, \\
a_3 &= -(2a_1 a_2 S_2 + a_1^3 S_3)/S_1, \\
a_4 &= -[(a_2^2 + 2a_1 a_3)S_2 + 3a_1^2 a_2 S_3 + a_1^4 S_4]/S_1, \\
a_5 &= -[2(a_1 a_4 + a_2 a_3)S_2 + 3(a_1^2 a_3 + a_2^2 a_1)S_3 + 4a_1^3 a_2 S_4 + a_1^5 S_5]/S_1,
\end{aligned} \quad (22)$$

and so on.

By inserting the strong-coupling expansion coefficients for the anharmonic oscillator into Eq. (20), compiled in Table 2, it is now straightforward to solve (21) for a_0 and then to evaluate the explicit expressions (22). The results up to the 10th order in $g^{-1/3}$ are collected in Table 3. By using a different method, Fernández [25] obtained: $a_0 = 1.855\,757\,081\,49$, $a_1 = -1.051\,866\,501\,1$, $a_2 =$

$-0.186\,184\,039\,3$, $a_3 = -0.036\,881\,399\,7$, and $a_4 = -0.003\,821\,866\,3$.

4.3 Expansion around g_0 with $E(g_0) = 0$

Another solution of (18) is obviously given by

$$E = 0, \quad (23)$$

$$g_0 = 16/\alpha_0^3 = 0.407\,748\dots \quad (24)$$

Here $\alpha_0 = 3.398\,150\dots$ was used from Table 2. The numerical value (24) is thus the coupling constant for which the ground-state energy of the perturbed Coulomb problem (10) assumes the special value zero. This may be used as the starting point for a systematic expansion of the form

$$E = b_1(g/g_0 - 1) + b_2(g/g_0 - 1)^2 + \dots, \quad (25)$$

with

$$b_1 = -\frac{1}{3} \frac{\hat{\alpha}_0}{\hat{\alpha}_1} g_0^{2/3}, \quad (26)$$

$$b_2 = -\frac{\hat{\alpha}_2}{\hat{\alpha}_1} \frac{b_1^2}{g_0^{2/3}}, \quad (27)$$

and so on. Using the numbers for α_n given in Table 2 this leads to $b_1 = 0.438\,855\dots$ and $b_2 = -0.038\,886\dots$. A comparison of this “ $(g - g_0)$ ”-expansion with the strong-coupling series is shown in Fig. 3. Notice the very good agreement even for relatively small coupling constants g .

5 Conclusions

Variational perturbation theory is a perfect tool for converting the divergent weak-coupling perturbation series of anharmonic systems into a sequence of exponentially fast converging approximations for the strong-coupling expansion. For the anharmonic oscillator, the empirically observed convergence behavior with superimposed oscillations can be explained by identifying the relevant singularities in a dispersion integral representation. A combination of variational perturbation theory with the mapping of 3D Coulomb onto 4D oscillator systems allows to compute with high accuracy the strong-coupling expansion for the ground-state energy E of a perturbed Coulomb system. As a by-product of this approach, a novel expansion around the coupling constant g_0 defined by $E(g_0) = 0$ can also be derived.

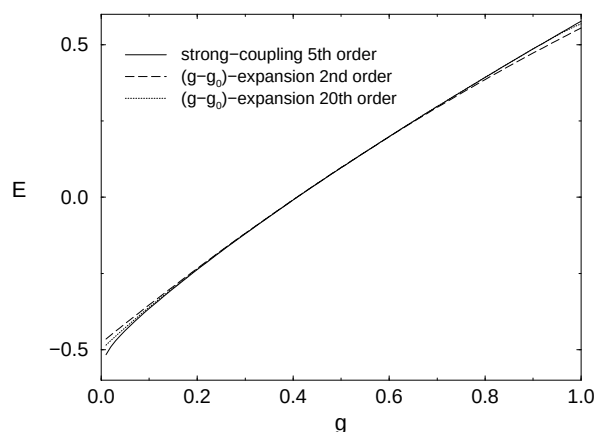


Figure 3. Comparison of strong-coupling and $(g - g_0)$ -expansions for the ground-state energy E of the Coulomb system (10).

In order to focus on the main points, I have here illustrated the basic ideas of variational perturbation theory with rather simple examples from quantum mechanics. The approach is, however, by no means limited to quantum mechanics only. In fact, in the past few years many fascinating applications to field theoretic models and critical phenomena, in particular to the precise calculation of critical exponents, have been worked out by Kleinert and his collaborators. Since these applications are by now far too numerous to be reviewed in this contribution, I refer the reader to several other articles in this chapter of the *Festschrift* dealing with the most successful directions of this ongoing line of research.

Acknowledgments

Before closing this article I wish to thank Hagen Kleinert for his guidance and friendship over many years. Being my “Diplom-” and “Doktorvater” he directed my scientific way at an early stage, and I am very grateful for his advice and suggestions. I remember with a smile often having sit next to him in a seminar or colloquium when, after he had a glance on some notes I gave him earlier, he suddenly passed a small sheet of paper saying just a few words: “Das ist ein Knüller” – “This is sensational”. I knew of course that this hardly could be true with all the problems we were trying to solve, but

perhaps this gives a glimpse on how enthusiastic Hagen Kleinert is about his research work and on how stimulating the atmosphere around him can be. For the years to come I wish him the power to maintain this optimistic attitude and to keep up with – or even further increase – the enormous productivity of the past years reflected at least partially in the many references to his work in this *Festschrift* – Happy 60th Birthday and all the best!

References

- [1] W. Janke and H. Kleinert, *Lett. Nuovo Cim.* **25**, 297 (1979).
- [2] W. Janke and H. Kleinert, *Phys. Rev. A* **42**, 2792 (1990).
- [3] W. Janke, *Phys. Lett. A* **143**, 107 (1990); *Phys. Lett. A* **144**, 116 (1990); *Phys. Rev. A* **41**, 6071 (1990).
- [4] H. Kleinert, S. Thoms, and W. Janke, *Phys. Rev. A* **55**, 915 (1997).
- [5] W. Janke and H. Kleinert, *Resummation of Divergent Perturbation Series - An Introduction to Theory and a Guide to Practical Applications*, lecture notes, 300 pages (unpublished).
- [6] R.P. Feynman and H. Kleinert, *Phys. Rev. A* **34**, 5080 (1986); R. Giachetti and V. Tognetti, *Phys. Rev. Lett.* **55**, 912 (1985); *Phys. Rev. B* **33**, 7647 (1986).
- [7] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [8] H. Kleinert, *Phys. Lett. A* **173**, 332 (1993).
- [9] W. Janke and H. Kleinert, *Phys. Lett. A* **118**, 371 (1986); *Chem. Phys. Lett.* **137**, 162 (1987); W. Janke and B.K. Cheng, *Phys. Lett. A* **129**, 140 (1988).
- [10] W. Janke, in *Path Integrals from meV to MeV*, Eds. V. Sa-yakanit *et al.* (World Scientific, Singapore, 1989), p. 355.
- [11] W. Janke and H. Kleinert, *Phys. Lett. A* **199**, 287 (1995).
- [12] W. Janke and H. Kleinert, *Phys. Rev. Lett.* **75**, 2787 (1995).
- [13] H. Kleinert and W. Janke, *Phys. Lett. A* **206**, 283 (1995).
- [14] C.M. Bender and T.T. Wu, *Phys. Rev.* **184**, 1231 (1969); *Phys. Rev. D* **7**, 1620 (1973).
- [15] H. Kleinert and H. Meyer, *Phys. Lett. A* **184**, 319 (1994).
- [16] For related work, see I.R.C. Buckley, A. Duncan, and H.F. Jones, *Phys. Rev. D* **47**, 2554 (1993); A. Duncan and H.F. Jones, *Phys. Rev. D* **47**, 2560 (1993); C.M. Bender, A. Duncan, and H.F. Jones, *Phys. Rev. D* **49**, 4219 (1994); C. Arvanitis, H.F. Jones, and C.S. Parker, *Phys. Rev.*

- D* **52**, 3704 (1995); R. Guida, K. Konishi, and H. Suzuki, *Ann. Phys.* **241**, 152 (1995); *ibid.* **249**, 109 (1996).
- [17] W. Janke and T. Sauer, *Phys. Lett. A* **197**, 335 (1995).
- [18] E.J. Weniger, *Phys. Rev. Lett.* **77**, 2859 (1996).
- [19] W. Janke and H. Kleinert, to be published.
- [20] A.J. Guttmann, in *Phase Transitions and Critical Phenomena*, Vol. 13, Eds. C. Domb and J.L. Lebowitz (Academic Press, New York, 1989), p. 1.
- [21] F. Vinette and J. Čížek, *J. Math. Phys.* **32**, 3392 (1991).
- [22] E. Schrödinger, *Proc. Roy. Irish Academy A* **46**, 183 (1941).
- [23] H. Duru and H. Kleinert, *Phys. Lett. B* **84**, 185 (1979); *Fortschr. Phys.* **30**, 401 (1982).
- [24] W. Janke, Leipzig preprint, to be published.
- [25] F.M. Fernández, *Phys. Lett. A* **166**, 173 (1992).

VARIATIONAL PERTURBATION THEORY FOR THE GROUND-STATE WAVE FUNCTION

A. PELSTER AND F. WEISSBACH

*Institut für Theoretische Physik, Freie Universität Berlin,
Arnimallee 14, D-14195 Berlin, Germany*

E-mails: pelster@physik.fu-berlin.de, florian.weissbach@physik.fu-berlin.de

We evaluate perturbatively the density matrix in the low-temperature limit and thus the ground-state wave function of the anharmonic oscillator up to second order in the coupling constant. We then employ Kleinert's variational perturbation theory to determine the ground-state wave function for all coupling strengths.

1 Introduction

Variational perturbation theory as developed by Kleinert [1] provides a systematic algorithm to evaluate perturbation series at all coupling strengths including the strong-coupling limit $g \rightarrow \infty$. It was thoroughly investigated for the ground-state energy of the anharmonic oscillator up to the 250th order [2,3], and its convergence was found to be exponentially fast and uniform. A similar systematic study has not yet been performed for the ground-state wave function. A first-order variational approach was set up by Kunihiro [4]. However, his method did not satisfactorily deal with certain problems in the variational procedure which will be discussed in detail below.

In this work we improve Kunihiro's first-order calculation and extend the treatment to the second order in the coupling strength. The ground-state wave function of the anharmonic oscillator is calculated from the low-temperature limit of the diagonal elements of the density matrix. A variational evaluation of the density matrix for the double-well potential has already been performed for finite temperatures in Ref. [5].

2 Perturbation Theory

Consider a quantum mechanical point particle of mass M moving in the one-dimensional anharmonic oscillator potential

$$V(x) = \frac{M}{2}\omega^2 x^2 + gx^4, \quad (1)$$

where ω denotes the frequency and g the coupling constant. We determine its ground-state wave function $\Psi(x)$ by evaluating the low-temperature limit of the diagonal elements of the density matrix

$$\Psi(x) = \lim_{\beta \rightarrow \infty} \sqrt{\rho(x, x)}, \quad (2)$$

which is defined by

$$\rho(x_b, x_a) = \frac{(x_b | \hbar\beta | x_a | 0)}{Z}. \quad (3)$$

Here $(x_b | \hbar\beta | x_a | 0)$ denotes the imaginary-time evolution amplitude with the path-integral representation [1]

$$\begin{aligned} (x_b | \hbar\beta | x_a | 0) &= \int_{x(0)=x_a}^{x(\hbar\beta)=x_b} \mathcal{D}x \\ &\times \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau \left[\frac{M}{2} \dot{x}^2(\tau) + \frac{M}{2} \omega^2 x^2(\tau) + gx^4(\tau) \right] \right\}, \quad (4) \end{aligned}$$

and Z denotes the partition function

$$Z = \int_{-\infty}^{+\infty} dx (x | \hbar\beta | x | 0). \quad (5)$$

By expanding Eq. (4) in powers of the coupling constant g we obtain the perturbation series

$$\begin{aligned} (x_b | \hbar\beta | x_a | 0) &= (x_b | \hbar\beta | x_a | 0)_\omega \left[1 - \frac{g}{\hbar} \int_0^{\hbar\beta} d\tau_1 \langle x^4(\tau_1) \rangle_\omega \right. \\ &\quad \left. + \frac{g^2}{2\hbar^2} \int_0^{\hbar\beta} d\tau_1 \int_0^{\hbar\beta} d\tau_2 \langle x^4(\tau_1) x^4(\tau_2) \rangle_\omega + \dots \right], \quad (6) \end{aligned}$$

where we have introduced the harmonic imaginary-time evolution amplitude

$$(x_b \hbar\beta | x_a 0)_\omega \equiv \int_{x(0)=x_a}^{x(\hbar\beta)=x_b} \mathcal{D}x \times \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau \left[\frac{M}{2} \dot{x}^2(\tau) + \frac{M}{2} \omega^2 x^2(\tau) \right] \right\}, \quad (7)$$

and the harmonic expectation value for an arbitrary functional $F[x]$ of the path $x(\tau)$:

$$\langle F[x] \rangle_\omega \equiv \frac{1}{(x_b \hbar\beta | x_a 0)_\omega} \int_{x(0)=x_a}^{x(\hbar\beta)=x_b} \mathcal{D}x F[x] \times \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau \left[\frac{M}{2} \dot{x}^2(\tau) + \frac{M}{2} \omega^2 x^2(\tau) \right] \right\}. \quad (8)$$

The latter is evaluated with the help of the generating functional for the harmonic oscillator, whose path-integral representation reads

$$(x_b \hbar\beta | x_a 0)_\omega[j] = \int_{x(0)=x_a}^{x(\hbar\beta)=x_b} \mathcal{D}x \exp \left\{ -\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau \left[\frac{M}{2} \dot{x}^2(\tau) + \frac{M}{2} \omega^2 x^2(\tau) - j(\tau)x(\tau) \right] \right\}, \quad (9)$$

leading to [1]

$$(x_b \hbar\beta | x_a 0)_\omega[j] = (x_b \hbar\beta | x_a 0)_\omega \exp \left[\frac{1}{\hbar} \int_0^{\hbar\beta} d\tau_1 x_{cl}(\tau_1) j(\tau_1) + \frac{1}{2\hbar^2} \int_0^{\hbar\beta} d\tau_1 \int_0^{\hbar\beta} d\tau_2 G(\tau_1, \tau_2) j(\tau_1) j(\tau_2) \right], \quad (10)$$

with

$$(x_b \hbar\beta | x_a 0)_\omega = \sqrt{\frac{M\omega}{2\pi\hbar \sinh \hbar\beta\omega}} \times \exp \left\{ -\frac{M\omega}{2\hbar \sinh \hbar\beta\omega} [(x_a^2 + x_b^2) \cosh \hbar\beta\omega - 2x_a x_b] \right\}. \quad (11)$$

In Eq. (10) we have introduced the classical path

$$x_{\text{cl}}(\tau) \equiv \frac{x_a \sinh(\hbar\beta - \tau)\omega + x_b \sinh \omega\tau}{\sinh \hbar\beta\omega}, \quad (12)$$

and the Green function

$$G(\tau_1, \tau_2) \equiv \frac{\hbar}{2M\omega \sinh \hbar\beta\omega} [\theta(\tau_1 - \tau_2) \sinh(\hbar\beta - \tau_1)\omega \sinh \omega\tau_2 + \theta(\tau_2 - \tau_1) \sinh(\hbar\beta - \tau_2)\omega \sinh \omega\tau_1] . \quad (13)$$

We follow Ref. [6] and evaluate harmonic expectation values of polynomials in x arising from the generating functional (10) according to Wick's theorem. Let us illustrate the procedure to reduce the power of the polynomial by the example of the harmonic expectation value

$$\langle x^n(\tau_1) x^m(\tau_2) \rangle_\omega . \quad (14)$$

- (i) Contracting $x(\tau_1)$ with $x^{n-1}(\tau_1)$ and $x^m(\tau_2)$ leads to a Green function $G(\tau_1, \tau_1)$ and $G(\tau_1, \tau_2)$ with multiplicity $n-1$ and m , respectively. The rest of the polynomial remains within the harmonic expectation value, leading to $\langle x^{n-2}(\tau_1) x^m(\tau_2) \rangle_\omega$ and $\langle x^{n-1}(\tau_1) x^{m-1}(\tau_2) \rangle_\omega$.
- (ii) If $n > 1$, extract one $x(\tau_1)$ from the expectation value giving $x_{\text{cl}}(\tau_1)$ multiplied by $\langle x^{n-1}(\tau_1) x^m(\tau_2) \rangle_\omega$.
- (iii) Add the terms (i) and (ii).
- (iv) Repeat the previous steps until only products of expectation values $\langle x(\tau_1) \rangle_\omega = x_{\text{cl}}(\tau_1)$ remain.

With the help of this procedure, the first-order harmonic expectation value $\langle x^4(\tau_1) \rangle_\omega$ is reduced to

$$\langle x^4(\tau_1) \rangle_\omega = x_{\text{cl}}(\tau_1) \langle x^3(\tau_1) \rangle_\omega + 3 G(\tau_1, \tau_1) \langle x^2(\tau_1) \rangle_\omega . \quad (15)$$

Furthermore, we find

$$\langle x^3(\tau_1) \rangle_\omega = x_{\text{cl}}(\tau_1) \langle x^2(\tau_1) \rangle_\omega + 2 G(\tau_1, \tau_1) x_{\text{cl}}(\tau_1) \quad (16)$$

and

$$\langle x^2(\tau_1) \rangle_\omega = x_{\text{cl}}^2(\tau_1) + G(\tau_1, \tau_1) . \quad (17)$$

Combining Eqs. (15)–(17) we obtain in first order

$$\langle x^4(\tau_1) \rangle_\omega = x_{\text{cl}}^4(\tau_1) + 6 x_{\text{cl}}^2(\tau_1) G(\tau_1, \tau_1) + 3 G^2(\tau_1, \tau_1) . \quad (18)$$

The second-order harmonic expectation value requires considerably more effort and finally leads to

$$\begin{aligned}
 \langle x^4(\tau_1) x^4(\tau_2) \rangle_\omega &= x_{\text{cl}}^4(\tau_1) x_{\text{cl}}^4(\tau_2) + 16 x_{\text{cl}}^3(\tau_1) G(\tau_1, \tau_2) x_{\text{cl}}^3(\tau_2) \\
 &+ 12 x_{\text{cl}}^2(\tau_1) G(\tau_1, \tau_1) x_{\text{cl}}^4(\tau_2) + 72 x_{\text{cl}}^2(\tau_1) G^2(\tau_1, \tau_2) x_{\text{cl}}^2(\tau_2) \\
 &+ 36 x_{\text{cl}}^2(\tau_1) G(\tau_1, \tau_1) G(\tau_2, \tau_2) x_{\text{cl}}^2(\tau_2) + 96 x_{\text{cl}}^3(\tau_1) G(\tau_1, \tau_2) G(\tau_2, \tau_2) x_{\text{cl}}(\tau_2) \\
 &+ 6 G^2(\tau_1, \tau_1) x_{\text{cl}}^4(\tau_2) + 96 x_{\text{cl}}(\tau_1) G^3(\tau_1, \tau_2) x_{\text{cl}}(\tau_2) \\
 &+ 144 x_{\text{cl}}(\tau_1) G(\tau_1, \tau_1) G(\tau_1, \tau_2) G(\tau_2, \tau_2) x_{\text{cl}}(\tau_2) + 9 G^2(\tau_1, \tau_1) G^2(\tau_2, \tau_2) \\
 &+ 36 G^2(\tau_1, \tau_1) x_{\text{cl}}^2(\tau_2) G(\tau_2, \tau_2) + 144 x_{\text{cl}}^2(\tau_1) G^2(\tau_1, \tau_2) G(\tau_2, \tau_2) \\
 &+ 72 G(\tau_1, \tau_1) G^2(\tau_1, \tau_2) G(\tau_2, \tau_2) + 24 G^4(\tau_1, \tau_2). \quad (19)
 \end{aligned}$$

The contractions can be illustrated by Feynman diagrams with the following rules. A vertex represents the integration over τ

$$\times = \int_0^{\hbar\beta} d\tau, \quad (20)$$

a line denotes the Green function

$$1 \text{ --- } 2 = G(\tau_1, \tau_2), \quad (21)$$

and a cross pictures a classical path

$$\times \text{ --- } 1 = x_{\text{cl}}(\tau_1). \quad (22)$$

Inserting the harmonic expectation values (18) and (19) into the perturbation expansion (6) leads in first order to the diagrams

$$\int_0^{\hbar\beta} d\tau_1 \langle x^4(\tau_1) \rangle_\omega \equiv \begin{array}{c} \times \\ | \\ \times \text{ --- } \times \\ | \\ \times \end{array} + 6 \begin{array}{c} \times \\ | \\ \times \text{ --- } \times \end{array} + 3 \begin{array}{c} \times \text{ --- } \times \\ | \\ \times \end{array}, \quad (23)$$

whereas the second-order terms are

$$\begin{aligned}
 \int_0^{\hbar\beta} d\tau_1 \int_0^{\hbar\beta} d\tau_2 \langle x^4(\tau_1) x^4(\tau_2) \rangle_\omega &\equiv \begin{array}{c} \times \\ | \\ \times \text{ --- } \times \end{array} \begin{array}{c} \times \\ | \\ \times \text{ --- } \times \end{array} + 16 \begin{array}{c} \times \text{ --- } \times \\ | \quad | \\ \times \quad \times \end{array} \\
 &+ 12 \begin{array}{c} \times \\ | \\ \times \text{ --- } \times \end{array} \begin{array}{c} \times \\ | \\ \times \end{array} + 72 \begin{array}{c} \times \text{ --- } \times \\ | \quad | \\ \times \quad \times \end{array} + 36 \begin{array}{c} \times \text{ --- } \times \\ | \quad | \\ \times \quad \times \end{array}
 \end{aligned}$$

$$\begin{aligned}
& +96 \times \text{diagram}_1 + 6 \times \text{diagram}_2 + 96 \times \text{diagram}_3 \\
& +144 \times \text{diagram}_4 + 36 \times \text{diagram}_5 + 144 \times \text{diagram}_6 \\
& +72 \times \text{diagram}_7 + 24 \times \text{diagram}_8 + 9 \times \text{diagram}_9 . \quad (24)
\end{aligned}$$

We observe that contributions from both connected and disconnected Feynman diagrams appear. The disconnected diagrams vanish once we rewrite the imaginary-time evolution amplitude in the form

$$(x_b \hbar\beta | x_a 0) = (x_b \hbar\beta | x_a 0)_\omega \exp [W(x_b, \hbar\beta; x_a, 0)] , \quad (25)$$

where the exponent $W(x_b, \hbar\beta; x_a, 0)$ contains only the connected Feynman diagrams. We obtain from (6) and (23)–(25) the expansion

$$\begin{aligned}
W(x_b, \hbar\beta; x_a, 0) = & -\frac{g}{\hbar} \left(\times \text{diagram}_{10} \times + 6 \times \text{diagram}_{11} \times + 3 \times \text{diagram}_{12} \right) \\
& + \frac{g^2}{2\hbar^2} \left(8 \times \text{diagram}_{13} \times + 36 \times \text{diagram}_{14} \times + 48 \times \text{diagram}_{15} \times \right. \\
& + 48 \times \text{diagram}_{16} \times + 72 \times \text{diagram}_{17} \times + 72 \times \text{diagram}_{18} \times \\
& \left. + 36 \times \text{diagram}_{19} + 12 \times \text{diagram}_{20} \right) + \dots , \quad (26)
\end{aligned}$$

where disconnected diagrams are indeed no longer present. As mentioned above, we restrict ourselves to the low-temperature limit of the diagonal elements of the density matrix which determine the ground-state wave function. In order to evaluate the various contributions in (26), we need the classical path (12) and the Green function (13) in the low-temperature limit:

$$\begin{aligned}
\lim_{\beta \rightarrow \infty} x_{cl}(\tau) &= x \left(e^{-\omega\tau} + e^{-\omega(\hbar\beta-\tau)} \right) , \quad (27) \\
\lim_{\beta \rightarrow \infty} G(\tau_1, \tau_2) &= \frac{\hbar}{2M\omega} \left[\theta(\tau_1 - \tau_2) e^{-\omega(\tau_1-\tau_2)} + \theta(\tau_2 - \tau_1) e^{-\omega(\tau_2-\tau_1)} \right]
\end{aligned}$$

$$\left. -e^{-\omega(\tau_1+\tau_2)} - e^{-2\hbar\beta\omega+\omega(\tau_1+\tau_2)} \right] . \quad (28)$$

Computing with these expressions the Feynman diagrams in (26), the low-temperature limit of the imaginary-time evolution amplitude (25) reads together with (11)

$$\begin{aligned} \lim_{\beta \rightarrow \infty} \langle x | \hbar\beta | x \rangle = \lim_{\beta \rightarrow \infty} \sqrt{\frac{M\omega}{\hbar\pi}} \exp \left[-\frac{\hbar\beta\omega}{2} - \frac{M\omega}{\hbar} x^2 + \frac{g}{\hbar} \left(\frac{9\hbar^2}{8M^2\omega^3} \right. \right. \\ \left. \left. - \frac{3\hbar^3\beta}{4M^2\omega^2} - \frac{3\hbar}{2M\omega^2} x^2 - \frac{1}{2\omega} x^4 \right) + \frac{g^2}{2\hbar^2} \left(\frac{21\hbar^5\beta}{4M^4\omega^5} - \frac{205\hbar^4}{16M^4\omega^6} \right. \right. \\ \left. \left. + \frac{21\hbar^3}{2M^3\omega^5} x^2 + \frac{11\hbar^2}{4M^2\omega^4} x^4 + \frac{\hbar}{3M\omega^3} x^6 \right) + \dots \right] . \quad (29) \end{aligned}$$

According to (5), the partition function Z follows from (29) by performing an integration with respect to x . This results in

$$\lim_{\beta \rightarrow \infty} Z = \lim_{\beta \rightarrow \infty} \exp \left(-\frac{\hbar\beta\omega}{2} - \frac{3g\hbar^2\beta}{4M^2\omega^2} + \frac{21g^2\hbar^3\beta}{8M^4\omega^5} + \dots \right) . \quad (30)$$

Inserting (29) and (30) into (3) we observe a cancellation of all terms which would diverge in the low-temperature limit $\beta \rightarrow \infty$. Thus the diagonal elements of the density matrix read in this limit

$$\begin{aligned} \lim_{\beta \rightarrow \infty} \rho(x, x) = \sqrt{\frac{M\omega}{\hbar\pi}} \exp \left[-\frac{M\omega}{\hbar} x^2 + \frac{g}{\hbar} \left(\frac{9\hbar^2}{8M^2\omega^3} - \frac{3\hbar}{2M\omega^2} x^2 - \frac{1}{2\omega} x^4 \right) \right. \\ \left. + \frac{g^2}{2\hbar^2} \left(-\frac{205\hbar^4}{16M^4\omega^6} + \frac{21\hbar^4}{2M^3\omega^5} x^2 + \frac{11\hbar^2}{4M^2\omega^4} x^4 + \frac{\hbar}{3M\omega^3} x^6 \right) + \dots \right] . \quad (31) \end{aligned}$$

By taking the square root and expanding the exponential term up to second order in the coupling strength g , we derive the second-order ground-state wave function (2):

$$\begin{aligned} \Psi(x) = \left(\frac{M\omega}{\hbar\pi} \right)^{1/4} \exp \left(-\frac{M\omega}{2\hbar} x^2 \right) \left[1 - \frac{g}{\hbar} \left(-\frac{9\hbar^2}{16M^2\omega^3} \right. \right. \\ \left. \left. + \frac{3\hbar}{4M\omega^2} x^2 + \frac{1}{4\omega} x^4 \right) + \frac{g^2}{2\hbar^2} \left(-\frac{1559\hbar^4}{256M^4\omega^6} + \frac{141\hbar^3}{32M^3\omega^5} x^2 \right. \right. \\ \left. \left. + \frac{53\hbar^2}{32M^2\omega^4} x^4 + \frac{13\hbar}{24M\omega^3} x^6 + \frac{1}{16\omega^2} x^8 \right) + \dots \right] . \quad (32) \end{aligned}$$

This result corresponds to the solution of the Bender-Wu recursion relation [7] for the ground-state wave function which is normalized such that

$$\int_{-\infty}^{+\infty} dx \Psi^2(x) = 1 \quad (33)$$

holds up to second order in the coupling strength g .

3 Variational Perturbation Theory

Variational perturbation theory enables us to evaluate the ground-state wave function for all values of the coupling constant g and even in the strong-coupling limit $g \rightarrow \infty$. To this end we simply add and subtract a harmonic oscillator of trial frequency Ω to the anharmonic oscillator potential (1):

$$V(x) = \frac{M}{2} \Omega^2 x^2 + g \frac{M}{2} \frac{\omega^2 - \Omega^2}{g} x^2 + g x^4. \quad (34)$$

We now treat the second term as if it was of the order of the coupling constant g . The result is obtained most simply by substituting the frequency ω in the original anharmonic oscillator potential (1) according to Kleinert's square root trick [1]

$$\omega \rightarrow \Omega \sqrt{1 + gr}, \quad (35)$$

where we have

$$r \equiv \frac{\omega^2 - \Omega^2}{g\Omega^2}. \quad (36)$$

Writing the ground-state wave function (32) in the form $\Psi(x) = \exp[W(x)]$ with the cumulant expansion

$$W(x) = \exp \left[\frac{1}{4} \log \left(\frac{M\omega}{\hbar\pi} \right) - \frac{M\omega}{2\hbar} x^2 + \frac{g}{\hbar} \left(\frac{9\hbar^2}{16M^2\omega^3} - \frac{3\hbar}{4M\omega^2} x^2 - \frac{1}{16\omega} x^4 \right) + \frac{g^2}{2\hbar^2} \left(-\frac{205\hbar^4}{32M^4\omega^6} + \frac{21\hbar^4}{4M^3\omega^5} x^2 + \frac{11\hbar^2}{8M^2\omega^4} x^4 + \frac{\hbar}{6M\omega^3} x^6 \right) + \dots \right], \quad (37)$$

we apply the trick (35) to $W(x)$, and reexpand in powers of g at fixed r . Afterwards r is substituted according to (36). Thus we obtain in the first order

$$\Psi^{(1)}(x, \Omega) = \exp \left[\frac{1}{4} \log \left(\frac{M\Omega}{\hbar\pi} \right) - \frac{1}{8} + \frac{\omega^2}{8\Omega^2} - \frac{M\Omega}{4\hbar} \left(1 + \frac{\omega^2}{\Omega^2} \right) x^2 \right]$$

$$+\frac{g}{\hbar}\left(\frac{9\hbar^2}{16M^2\Omega^3}-\frac{3\hbar}{4M\Omega^2}x^2-\frac{1}{4\Omega}x^4\right)\Bigg], \quad (38)$$

whereas the second-order expansion reads

$$\begin{aligned} \Psi^{(2)}(x, \Omega) = \exp \Bigg\{ & \frac{1}{4} \log \left(\frac{M\Omega}{\hbar\pi} \right) - \frac{1}{16} + \frac{\omega^2}{4\Omega^2} - \frac{\omega^4}{16\Omega^4} \\ & - \frac{M\Omega}{2\hbar} \left(\frac{3}{8} + \frac{3\omega^2}{4\Omega^2} - \frac{\omega^4}{8\Omega^4} \right) x^2 + \frac{g}{\hbar} \left[\frac{9\hbar^2}{16M^2\Omega^3} \left(\frac{5}{2} - \frac{3\omega^2}{2\Omega^2} \right) \right. \\ & \left. - \frac{3\hbar}{4M\Omega^2} \left(2 - \frac{\omega^2}{\Omega^2} \right) x^2 - \frac{1}{4\Omega} \left(\frac{3}{2} - \frac{\omega^2}{2\Omega^2} \right) x^4 \right] \\ & \left. + \frac{g^2}{2\hbar^2} \left[-\frac{205\hbar^4}{32M^4\Omega^6} + \frac{21\hbar^3}{2M^3\Omega^5} x^2 + \frac{11\hbar^2}{8M^2\Omega^4} x^4 + \frac{\hbar}{6M\Omega^3} x^6 \right] \right\}. \quad (40) \end{aligned}$$

Both in first and in second order, the ground-state wave function depends on the artificially introduced frequency parameter Ω . According to the principle of minimal sensitivity [8] we minimize its influence on $\Psi^{(n)}(x, \Omega)$ by searching for local extrema of $\Psi^{(n)}(x, \Omega)$ with respect to Ω . As we have written the wave function in the form $\Psi^{(n)}(x, \Omega) = \exp[W^{(n)}(x, \Omega)]$, it is sufficient to take into account just the inner derivative of $\Psi^{(n)}(x, \Omega)$, i.e. we obtain the condition $\partial W^{(n)}(x, \Omega)/\partial\Omega = 0$.

It turns out in the first order $n = 1$ that this equation has two solutions for $x < 0.684$ and for $x > 0.780$, however in the interval $0.684 < x < 0.780$, $\Psi^{(1)}(x, \Omega)$ does not have any extremum [4]. In accordance with the principle of minimal sensitivity we look for turning points on that interval instead, i.e. we solve $\partial^2 W^{(1)}(x, \Omega)/\partial\Omega^2 = 0$. Fig. 1 shows how the curve for the turning points links the extremal branches. Now we have to choose which one of the branches of $\Omega^1(x)$ we take into account. Inserting the lower branch for $x > 0.780$ into the wave function (38) leads to unphysical results as the ground-state wave function explodes dramatically. Thus we choose the upper branch for $x > 0.780$. For $x < 0.684$ the wave function becomes rather independent of the choice of Ω . As we are looking for a function $\Omega^{(1)}(x)$ which is as smooth as possible, we choose the lower branch for $x < 0.684$. Fig. 1 shows all branches of Ω and highlights our final choice by a solid line.

For the second order $n = 2$ there are no real positive solutions of the equation $\partial W^{(2)}(x, \Omega)/\partial\Omega = 0$ on the interval $x = [0, 4]$. Once more we have to look for turning points instead and solve $\partial^2 W^{(2)}(x, \Omega)/\partial\Omega^2 = 0$. This equation has two positive solutions on this interval, so we get two branches for the solution $\Omega^{(2)}(x)$ (see Fig. 2). Again we have to choose one of these two

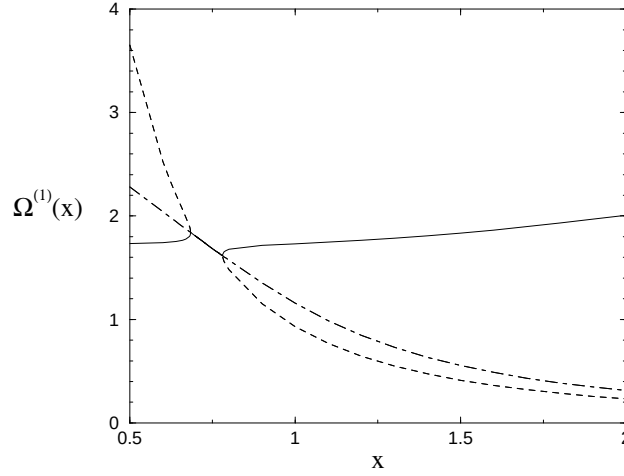


Figure 1. First-order results for the variational parameter Ω at the intermediate coupling $g = 1/2$. The extremal branches for $x < 0.684$ and for $x > 0.780$ (solid lines and dashed lines) are obtained from the equation $\partial W^{(1)}(x, \Omega)/\partial \Omega = 0$. For $0.684 < x < 0.780$ there are no real positive solutions of this equation. Thus we look in this interval also for turning points, i.e. we determine real positive solutions of the equation $\partial^2 W^{(1)}(x, \Omega)/\partial \Omega^2 = 0$. The curve for the turning points on the entire interval lies between the two other branches (dot-dashed line) and fills the gap. Thus we can take those branches into account which provide us with the most continuous function $\Omega^{(1)}(x)$, i.e. the solid line.

branches. Relying on a similar argument as for the first-order approximation we choose the upper branch for $x > 0.8$ and the lower one for $x < 0.8$. The perturbation series converges so quickly that the curves for the first and second order as well as the exact ground-state wave function are not distinguishable on the plots. To see the difference we determine the mean square deviation from the exact numerical solution

$$D^{(n)} = 2 \int_0^\infty dx \left[\Psi^{(n)}(x) - \Psi^{\text{ex}}(x) \right]^2, \quad (41)$$

where the index n denotes the order. The integration is performed numerically. The factor 2 is introduced for symmetry reasons, since we restrict our calculations to the positive x -axis. It turns out that the mean square deviation $D^{(2)} = 6.8 \times 10^{-7}$ is smaller than $D^{(1)} = 1.1 \times 10^{-5}$ by a factor of 0.063, which indicates that variational perturbation theory converges very quickly also for the ground-state wave function. The same applies to both

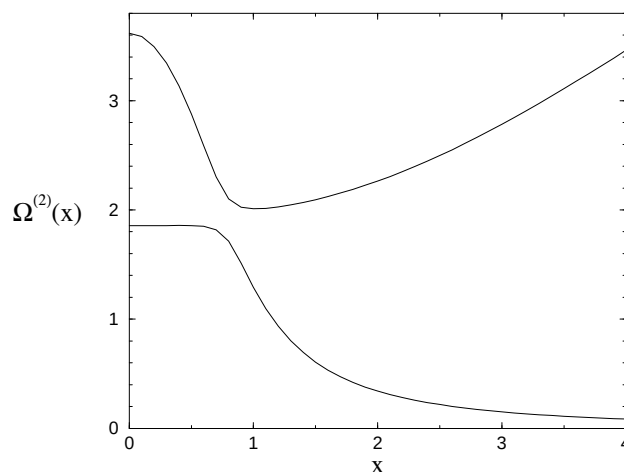


Figure 2. The two positive branches of $\Omega^{(2)}(x)$ on the interval $[0, 4]$ for intermediate coupling $g = 1/2$ obtained by solving the turning point equation $\partial^2 W^{(2)}(x, \Omega)/\partial \Omega^2 = 0$. In order to achieve the smoothest function we choose the lower branch for $x < 0.8$ and the upper branch for $x > 0.8$. This choice is justified by the results of our first-order calculation (see Fig. 1).

weak and strong coupling as is illustrated in Fig. 3 which shows the second-order ground-state wave function $\Psi^{(2)}(x)$ for $g = 0.1$, $g = 1/2$, and $g = 50$, respectively.

Note that variational perturbation theory does not preserve the normalization of the wave function. Although the perturbative ground-state wave function (32) is still normalized in the usual sense, this normalization is spoilt by extremizing $\Psi(x, \Omega)$ with respect to the frequency parameter Ω . Thus we have to normalize the variational ground-state wave function at the end.

Acknowledgments

Both of us congratulate Professor Kleinert to his 60th birthday. F.W. thanks Professor Kleinert for supervising the research for his diploma thesis where the results of this article were derived. Moreover F.W. is very grateful to his family for their patience and their support. Finally both authors thank Michael Bachmann who always found time for fruitful discussions on variational perturbation theory.

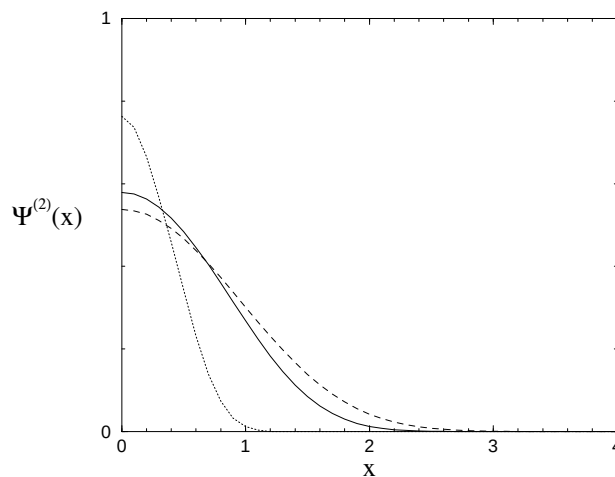


Figure 3. The normalized second-order ground-state wave function for weak coupling (dashed, $g = 0.1$), for intermediate coupling (solid, $g = 1/2$), and for strong coupling (dotted, $g = 50$).

References

- [1] H. Kleinert, *Path Integrals in Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [2] W. Janke and H. Kleinert, *Phys. Rev. Lett.* **75**, 2787 (1995).
- [3] H. Kleinert and W. Janke, *Phys. Lett. A* **206**, 283 (1995).
- [4] T. Kunihiro, *Phys. Rev. Lett.* **78**, 3229 (1997).
- [5] M. Bachmann, H. Kleinert, and A. Pelster, *Phys. Rev. A* **60**, 3429 (1999).
- [6] H. Kleinert, A. Pelster, and M. Bachmann, *Phys. Rev. E* **60**, 2510 (1999).
- [7] C.M. Bender and T.T. Wu, *Phys. Rev.* **184**, 1231 (1969); *Phys. Rev. D* **7**, 1620 (1973).
- [8] P.M. Stevenson, *Phys. Rev. D* **23**, 2916 (1981).

UNIFORMLY SUITABLE ESTIMATION FOR THERMODYNAMIC VALUES

I.D. FERANCHUK AND A. IVANOV

*Department of Theoretical Physics, Belarussian State University,
4 Fr. Skaryny av., 220080 Minsk, Republic of Belarus*

E-mail: ilya@fer.minsk.by

The operator method and the cumulant expansion are used for the approximate calculation of partition functions and free energy in quantum statistics. It is shown for some model systems that the zeroth approximation of the method interpolates these values with rather good accuracy in the entire range of both the Hamiltonian parameters and temperature. The proposed method allows one to calculate also the corrections to the zeroth approximation.

1 Introduction

An important contribution of Professor Hagen Kleinert and his research group concerns the development of non-perturbative methods for quantum field theory and quantum statistics [1]. His analysis of this problem is based mainly on the functional formulation of quantum mechanics. The Green function and the density matrix are the main objects in such an approach and that allows one to calculate the observable physical values directly, without an intermediate calculation of the energy spectrum for all stationary states of the system. Unfortunately, this great advantage of the method is rather restricted for the non-perturbative theories because of the difficulties when calculating the path integrals with non-quadratic action.

On the other hand, the operator method (OM) can be used for the non-perturbative solution of the Schrödinger equation [2]. It permits one to find the uniformly fitted interpolation and convergent series for eigenvalues of quite complicated Hamiltonians. But this approach is rather difficult for

calculations in quantum statistics, where the physical values are defined by the energy spectrum as a whole.

The purpose of the present work is to generalize the OM for the calculation of thermodynamic values in quantum statistics. Namely, we will show that the application of OM for the eigenvalues together with the cumulant expansion [3] for the summation on the quantum numbers of the system permits us to calculate the uniformly suitable estimation for the partition function and free energy and to find the second-order corrections for them in the entire range of temperature and Hamiltonian parameters.

2 Formulation of the Problem

First we want to define more precisely what we mean under "uniformly suitable estimation" (USE). Let us consider some physical variable with eigenvalues $F_n(\lambda)$ depending on the quantum number n and the physical parameter λ . Let us also take the function $F_n^{(0)}(\lambda)$ as the USE for $F_n(\lambda)$ with the following inequality being fulfilled in the entire range of variation of n and λ :

$$\left| \frac{F_n^{(0)}(\lambda) - F_n(\lambda)}{F_n(\lambda)} \right| \leq \xi^{(0)}. \quad (1)$$

Here the parameter $\xi^{(0)} < 1$ is supposed to be independent on the values n, λ and defines the accuracy of USE. In principle, we also suppose that there is a method for calculating a sequence of the functions $F_n^{(s)}(\lambda)$ corresponding to the decreasing sequence of the parameters $\xi^{(s)}$; $s = 0, 1, 2, \dots$, so that

$$\lim_{s \rightarrow \infty} F_n^{(s)}(\lambda) = F_n(\lambda). \quad (2)$$

It seems that the definition in Eq. (1) is not constructive because the exact values $F_n(\lambda)$ are unknown. However, there are several possibilities to estimate the value $\xi^{(0)}$. In particular, one can compare the asymptotic series for $F_n(\lambda)$ in different limiting cases of the parameter or the quantum number with the corresponding expansions of the function $F_n^{(0)}(\lambda)$. Besides, the difference between $F_n^{(1)}(\lambda)$ and $F_n^{(0)}(\lambda)$ can also be considered as an estimation for $\xi^{(0)}$. As an example, we mention the USE for the eigenvalues of the Hamiltonian for various physical systems calculated on the basis of the OM [1,4,5].

Let us consider from this point of view the thermodynamic perturbation theory in the Schrödinger representation of quantum statistics. Usually it is formulated for the free energy of the system [6] and the leading terms are the

following

$$F(\lambda, \beta) = F_0 + \lambda \sum_n V_{nn} w_n + \lambda^2 \sum_n \sum_{m \neq n} \frac{|V_{mn}|^2 w_n}{E_n^{(0)} - E_m^{(0)}} + \frac{1}{2} \beta \lambda^2 \left[\left(\sum_n V_{nn} w_n \right)^2 - \sum_n V_{nn}^2 w_n \right] + \dots \quad (3)$$

Here we have introduced $\beta = 1/k_B T$, where T is the temperature and k the Boltzmann constant; F, F_0 are the exact and approximate free energies respectively; $E_n^{(0)}$ are the eigenvalues of the unperturbed Hamiltonian and V_{mn} are the matrix elements of the perturbation operator with the following form of the total Hamiltonian

$$H = H_0 + \lambda V, \quad (4)$$

and $w_n = \exp [\beta(F_0 - E_n^{(0)})]$ is the unperturbed density matrix.

With any fixed number of terms, these series do not yield the free energy in the whole range of the temperature and the perturbation parameter even for the simplest cases. Let us illustrate this by means of the model Hamiltonian used in Ref. [7] for the analysis of convergence of usual perturbation series

$$\begin{aligned} H &= H_0 + \lambda V = \frac{1}{2}(\hat{p}^2 + x^2) + \lambda x^2, \\ E_n^{(0)} &= n + \frac{1}{2}, \quad F_0 = \frac{1}{\beta} \ln [2 \sinh (\beta/2)], \\ V_{mn} &= \left(n + \frac{1}{2} \right) \delta_{m,n} + \frac{1}{2} \left[\sqrt{(n+1)(n+2)} \delta_{m,n+2} \right. \\ &\quad \left. + \sqrt{n(n-1)} \delta_{m,n-2} \right], \end{aligned} \quad (5)$$

with $\delta_{m,n}$ being the Kronecker symbol.

Certainly, the exact free energy is well known for this model ($F = \ln [2 \sinh (\beta/2 \sqrt{1+2\lambda})]/\beta$) but if we use these matrix elements in formula (3), rather simple calculations lead to the following result

$$\begin{aligned} F(\lambda, \beta) &= \frac{1}{\beta} \ln [2 \sinh \beta/2] + \frac{\lambda}{2} \coth \beta/2 \\ &\quad - \frac{\lambda^2}{8} \left[1 + \coth \beta/2 + \frac{1}{2 \sinh^2 \beta/2} (1 + 2\beta) \right] + \dots \end{aligned} \quad (6)$$

It is evident that this series does not satisfy the USE criteria in the plane of both parameters. In the low-temperature limit ($\beta \rightarrow \infty$), formula (6) leads to the power series of λ for the ground-state energy and this series diverges in the range of $\lambda > 1/2$ because of a singular point of the exact eigenvalue [7] in the complex plane of λ . When the temperature increases ($\beta \rightarrow 0$), the second-order correction becomes singular ($\sim -\lambda^2/4\beta^2$), although the exact free energy has no similar singularity.

So, instead of Eq. (3), our objective is to formulate another regular method which permits to find the USE for the free energy of the quantum system with an arbitrary energy spectrum $E(n)$.

3 Cumulant Expansion for the Quantum Partition Function

The partition function of some quantum system,

$$Z(\beta) = \sum_{n=0}^{\infty} \exp[-\beta E(n)], \quad (7)$$

can be represented in the operator form. Let us introduce the basic set of the state vectors as the eigenfunctions of the excitation number operator

$$\hat{n}|n\rangle = \hat{a}^+ \hat{a}|n\rangle = n|n\rangle, \quad (8)$$

with the creation $\hat{a}^+$ and annihilation $\hat{a}$ operators, corresponding to the harmonic oscillator with arbitrary frequency ω .

Then, we can write

$$Z(\beta) = \langle v | \exp[-\beta E(\hat{n}) + v\hat{n} - \ln N(v)] | v \rangle. \quad (9)$$

Here $|v\rangle$ is the normalized state vector depending on the arbitrary parameter v having the physical meaning of the inverse temperature:

$$|v\rangle = \sqrt{N(v)} \sum_{n=0}^{\infty} e^{-vn} |n\rangle, \quad N(v) = 1 - e^{-v}. \quad (10)$$

Let us remind that the cumulant expansion is valid for an arbitrary exponential operator when averaging any normalized state vector [3],

$$\langle \exp(\hat{A}) \rangle = \exp \left[\sum_{n=1}^{\infty} \frac{K_n}{n!} \right], \quad (11)$$

where the cumulants K_n are expressed in terms of the moments of the operator $\hat{A}$. This expansion is the strict one and every cumulant corresponds to the partial summation of usual power series. The first few terms in Eq. (11) are [3]

$$\begin{aligned} K_1 &= \langle \hat{A} \rangle, \\ K_2 &= \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2, \\ K_3 &= \langle \hat{A}^3 \rangle - 3\langle \hat{A} \rangle \langle \hat{A}^2 \rangle + 2\langle \hat{A} \rangle^3. \end{aligned} \quad (12)$$

If we apply now two terms of the cumulant expansion to the partition function, the representation (9) can be transformed as

$$\begin{aligned} Z(\beta) &\simeq \tilde{Z}(\beta, v) = \exp \left[\langle v | \hat{R} | v \rangle - \ln N(v) + \frac{1}{2} \left(\langle v | \hat{R}^2 | v \rangle - \langle v | \hat{R} | v \rangle^2 \right) \right], \\ \hat{R} &= -\beta E(\hat{n}) + v \hat{n}. \end{aligned} \quad (13)$$

A further analytical consideration is possible if we suppose that the fluctuation of the number of excitations with respect to its averaged value is limited in the entire range of the temperature; the self-consistency of this assumption will be tested by the final result. Then all values in formula (13) can be expanded in the series of this fluctuation and the following result can be obtained up to second order:

$$\begin{aligned} \tilde{Z}(\beta, v) &\simeq \exp \left\{ -\beta E(\bar{n}) + v \bar{n} - \ln N(v) \right. \\ &\quad \left. - \frac{\bar{n}(\bar{n}+1)}{2} [\beta E''(\bar{n}) + \beta^2 E(\bar{n}) E''(\bar{n}) + (\beta E'(\bar{n}) - v)^2] \right\}. \end{aligned} \quad (14)$$

Here the value $\bar{n}$ has the physical meaning of the averaged number of excitations and it is defined by the formula

$$\bar{n} = \frac{e^{-v}}{1 - e^{-v}}, \quad e^{-v} = \frac{\bar{n}}{\bar{n} + 1}. \quad (15)$$

A more convenient way is to consider the value $\bar{n}$ as the variational parameter instead of v and to find $\bar{n} = \bar{n}(\beta)$ from the minimum condition for the free energy calculated in the zeroth order of the cumulant expansion. In that way the following expansion can be obtained with the desired accuracy:

$$F(\beta) = -\frac{1}{\beta} \ln Z(\beta) = F^{(0)} + \Delta F + \dots,$$

$$F^{(0)} = E(\bar{n}) - \frac{1}{\beta}[(\bar{n} + 1) \ln(\bar{n} + 1) - \bar{n} \ln \bar{n}], \quad (16)$$

$$\Delta F = \frac{\bar{n}(\bar{n} + 1)}{2} E''(\bar{n})[1 + \beta E(\bar{n})].$$

The value $\bar{n}(\beta)$ in the formula (16) should be calculated as the solution of the equation

$$\beta E'(\bar{n}) = \ln \frac{\bar{n} + 1}{\bar{n}}. \quad (17)$$

The formulas (16) and (17) define the USE for the free energy of the quantum system with known energy spectrum. They can be generalized for systems with several degrees of freedom.

4 Operator Method for the Estimation of the Eigenvalues

In order to make the above result in practice more effective, it should be supplemented by some uniformly suitable estimation for the eigenvalues considered as a function of quantum numbers and parameters of the Hamiltonian $\hat{H}$ of the system. For this purpose we use the zeroth approximation of the OM [2] when the same operators $\hat{a}, \hat{a}^+$, used in formula (8), are introduced to $\hat{H}$ by means of the canonical transformation for the operators of coordinate and momentum

$$\hat{x} = \frac{1}{\sqrt{2\omega}}(\hat{a} + \hat{a}^+), \quad \hat{p} = -i\sqrt{\frac{\omega}{2}}(\hat{a} - \hat{a}^+), \quad (18)$$

with the frequency ω considered as variational parameter.

Then one has to select, in the total Hamiltonian, that part $\hat{H}_0$ which is diagonal with respect to the excitation number operator $\hat{n} = \hat{a}^+ \hat{a}$:

$$\hat{H} = \hat{H}_0(\hat{n}) + \hat{H}_1, \quad [\hat{H}_0, \hat{n}] = 0. \quad (19)$$

Thus the OM zeroth approximation for the energy levels of the system is defined as the evident eigenvalues of the operator

$$\hat{H}_0|n\rangle = E^{(0)}(n, \omega_n)|n\rangle. \quad (20)$$

The frequency parameters ω_n are the solutions of the variational equations for every quantum number

$$\frac{\partial E^{(0)}(n, \omega_n)}{\partial \omega_n} = 0. \quad (21)$$

As shown for a number of physical systems [2,4,5], the function $E^{(0)}(n, \omega_n)$ defines the USE for the exact energy spectrum with rather good accuracy. Thus, Eqs. (20) and (21), together with (16) and (17), permit us to calculate the USE for the quantum free energy as the solution of the system of algebraic equations.

In order to illustrate how this method works, let us consider the Hamiltonian corresponding to the problem of the anharmonic oscillator,

$$\hat{H} = \frac{1}{2}(\hat{p}^2 + \hat{x}^2) + \mu\hat{x}^2 + \lambda\hat{x}^4. \quad (22)$$

In this case the transformation (18) leads to the function

$$E^{(0)}(n, \omega_n) = \frac{2n+1}{4} \left[\omega_n + \frac{1}{\omega_n}(1+2\mu) \right] + \frac{3\lambda}{4\omega_n^2}(2n^2+2n+1). \quad (23)$$

The parameter ω_n is the solution of the cubic equation

$$\omega_n^3 - \omega_n(1+2\mu) - \frac{6\lambda(2n^2+2n+1)}{2n+1} = 0. \quad (24)$$

Then Eq. (17) for the averaged number of the excitations, i.e. the saddle point for the partition function, reads

$$\frac{\beta}{2} \left[\frac{1}{\omega_{\bar{n}}}(\omega_{\bar{n}}^2 + 1 + 2\mu) + \frac{3\lambda}{\omega_{\bar{n}}^2}(2\bar{n}+1) \right] = \ln \frac{\bar{n}+1}{\bar{n}}, \quad (25)$$

and the free energy takes the following form:

$$F^{(0)}(\beta, \lambda, \mu) = \frac{1}{4}\bar{n} \left(\omega_{\bar{n}} - \frac{1+2\mu}{\omega_{\bar{n}}} \right) + \frac{1}{8} \left(\omega_{\bar{n}} + \frac{1+2\mu}{\omega_{\bar{n}}} \right) - \frac{3\lambda}{2\omega_{\bar{n}}^2}\bar{n}(2\bar{n}+1) - \frac{1}{\beta} \ln \bar{n}. \quad (26)$$

In this work, we restrict ourselves to analytical calculations only. Let us consider some limiting cases for the solutions of equations (22) to (24). If, for example, the anharmonic parameter λ is zero, equations (22) and (23) lead to the exact values

$$\omega_n = \sqrt{1+2\mu}, \quad E^{(0)}(n, \omega_n) = E_n = \sqrt{1+2\mu} \left(n + \frac{1}{2} \right). \quad (27)$$

Eq. (24) has also the analytical solution

$$\bar{n} = \frac{e^{-\nu}}{1 - e^{-\nu}}, \quad \nu = \beta\sqrt{1+2\mu}. \quad (28)$$

When substituting this value, Eq. (16) transforms to the expression

$$F = \frac{1}{\beta} \ln [2 \sinh (\beta/2 \sqrt{1+2\mu})], \quad (29)$$

which coincides with the exact one as distinct from the usual thermodynamic perturbation theory.

Let us consider also the limiting cases of low ($\beta \rightarrow \infty$) and high ($\beta \rightarrow 0$) temperature. Using these conditions one can simplify the formulae (17) and (22) to (26) to the following form. In the case of $\beta \rightarrow \infty$ and thus

$$\bar{n} \simeq e^{-\beta\nu_0} \ll 1,$$

we have

$$\begin{aligned} \omega_0^3 - \omega_0(1+2\mu) - 6\lambda &= 0, \\ E(n, \omega_n) = E(0, \omega_0) &= \frac{1}{8} \left(3\omega_0 + \frac{1+2\mu}{\omega_0} \right) \simeq F^{(0)}, \\ \nu_0 = E'(0, \omega_0) &= \frac{1}{4} \left(3\omega_0 + \frac{1+2\mu}{\omega_0} \right), \\ \Delta F &\simeq \frac{1}{2} \bar{n} \frac{3\lambda}{\omega_0^2} \beta E(0, \omega_0) \simeq \frac{1}{32} \left(\omega_0 - \frac{1+2\mu}{\omega_0} \right) \left(3\omega_0 + \frac{1+2\mu}{\omega_0} \right) \beta e^{-\beta\nu_0}, \\ \left| \frac{\Delta F}{F^{(0)}} \right| &= \frac{1}{4} \left| \left(\omega_0 - \frac{1+2\mu}{\omega_0} \right) \right| \beta e^{\beta\nu_0} \ll 1. \end{aligned}$$

In the opposite case of high temperature $\beta \rightarrow 0$ and

$$\bar{n} = \left(\frac{36}{\lambda} \right)^{1/4} \left(\frac{3}{4\beta} \right)^{3/4} \gg 1,$$

we obtain, leaving only the major $\bar{n}$ part in Eq. (24) and neglecting the term proportional to $\omega_{\bar{n}}$,

$$\begin{aligned} \omega_{\bar{n}} &\simeq (6\lambda\bar{n})^{1/3}, \\ \bar{E} = E(\bar{n}, (6\lambda\bar{n})^{1/3}) &\simeq \frac{9}{2} \left(\frac{\lambda}{36} \right)^{1/3} \bar{n}^{4/3}, \\ F^{(0)} &\simeq \bar{E} - \frac{1}{\beta} \left[\bar{n} \ln \frac{\bar{n}+1}{\bar{n}} + \ln \bar{n} \right] \simeq \frac{3}{4\beta} \ln \beta, \\ \Delta F &\simeq \frac{\bar{n}^2}{2} E''(1 + \beta E) \simeq \frac{7}{24\beta}, \end{aligned}$$

$$\left| \frac{\Delta F}{F^{(0)}} \right| \simeq \frac{7}{18|\ln \beta|} \ll 1.$$

So, the second-order corrections are small in both limiting cases. That means that the system of algebraic equations (22) to (26) defines the USE for the anharmonic oscillator.

Finally, the proposed method can be generalized to quantum systems with many internal degrees of freedom. This permits to consider analytically the thermodynamic characteristics of real molecular gases.

References

- [1] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [2] I.D. Feranchuk, L.I. Komarov, I.V. Nichipor, and A.P. Ulyanenko, *Ann. Phys. (N.Y.)* **238**, 370 (1995).
- [3] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1951).
- [4] I.D. Feranchuk and L.X. Hai, *Phys. Lett. A* **137**, 385 (1989).
- [5] I.D. Feranchuk and A.L. Tolstik, *J. Phys. A: Math. Gen.* **32**, 2115 (1999).
- [6] L.D. Landau and E.M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976).
- [7] F.M. Fernández and E.A. Castro, *Phys. Lett. A* **91**, 339 (1982).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

**PATH-INTEGRAL AND PERTURBATION METHODS FOR
DEBYE-WALLER FACTORS OBSERVED BY EXTENDED
X-RAY-ABSORPTION FINE STRUCTURE SPECTROSCOPY**

T. YOKOYAMA

*Department of Chemistry, Graduate School of Science, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

E-mail: toshi@chem.s.u-tokyo.ac.jp

Extended x-ray-absorption fine-structure spectroscopy provides information on the radial distribution function (RDF) around the x-ray absorbing atom in molecules and solids. The cumulants of RDF in gaseous Br₂ are evaluated using the second-order perturbation and path integral effective classical potential (ECP) methods, and the results are compared to the experimental data. As an application of the ECP method to solid surfaces, anisotropic and anharmonic vibration of thin Cu films is examined. The calculated results suggest that the out-of-planar vibration is more enhanced and more anharmonic than the lateral one, this being consistent with the experiments.

1 Introduction

Recently great attention has been paid to the temperature dependence of extended x-ray-absorption fine-structure (EXAFS), which provides information on thermal vibrations including anharmonicity. Experimentally, the EXAFS Debye-Waller factors or the cumulants of the radial distribution function (RDF) around the x-ray absorbing atom can be obtained. The theory concerning the relationship between the EXAFS cumulants and vibrational potentials has been developed based on the theories of quantum statistical first-order perturbation (P1) [1–5] and the path integral effective classical potential (ECP) [6–8]. The perturbation theory includes several intrinsic difficulties for practical use: complicated formulation and computation of many-atom systems and less reliability in large anharmonic systems. For instance,

the fourth-order cumulant from P1 is negative in the case of a positive fourth-order force constant and needs the second-order perturbation (P2) to yield meaningful results. On the other hand, the path integral ECP method [9] can appropriately overcome such difficulties in the P1 theory, although on the other hand the ECP method cannot describe quantum statistical anharmonicity at very low temperature.

Here, I will present the formulas of the EXAFS cumulants for diatomic systems using the P2 theory. This work was kindly motivated by Prof. Kleinert. When I visited Berlin in 1998, I showed him the ECP and P1 results. He was not satisfied with the results because he knew that the fourth-order cumulant requires the P2 theory and that the disagreement of the P1 result with the ECP theory and the experiment is trivial. He immediately suggested that I should derive the P2 formula. The P2 results given here are compared to those by the ECP theory and the experiment of the Br_2 molecule. I will also show my recent results concerning anisotropic and anharmonic vibration of surface atoms for thin Cu films calculated with the ECP method. The results are compared to our experimental data.

2 Diatomic Br_2

2.1 Second-Order Perturbation Formulas of the EXAFS Cumulants

Let us first derive the P2 formulas of the EXAFS cumulants for the diatomic system up to the fourth order. The thermal average of some physical quantity $\langle \mathcal{A} \rangle$ is expressed using a trace as

$$\langle \mathcal{A} \rangle = \frac{1}{Z} \text{Tr} \mathcal{A} e^{-\beta(\mathcal{H}_0 + \mathcal{H}')}, \quad (1)$$

where $Z = \text{Tr} e^{-\beta(\mathcal{H}_0 + \mathcal{H}')}$ is the partition function, $\mathcal{H}_0$ and $\mathcal{H}'$ are the non-perturbed and perturbed Hamiltonian operators, respectively, and $\beta = (k_B T)^{-1}$ (k_B the Boltzmann constant and T the temperature). In the P2 perturbation, the trace is given as

$$\begin{aligned} \text{Tr} \mathcal{A} e^{-\beta(\mathcal{H}_0 + \mathcal{H}')} &\cong \text{Tr} \mathcal{A} e^{-\beta \mathcal{H}_0} \left(1 - \int_0^\beta \tilde{H}'(t_1) dt_1 \right. \\ &\quad \left. + \int_0^\beta \int_0^{t_1} \tilde{H}'(t_1) \tilde{H}'(t_2) dt_2 dt_1 \right), \end{aligned} \quad (2)$$

where $\tilde{H}'(\beta) = e^{\beta\mathcal{H}_0} \mathcal{H}' e^{-\beta\mathcal{H}_0}$. The partition function is obtained in a similar equation [$\mathcal{A}=1$ in Eq. (2)]. When the non-perturbed Hamiltonian is exactly solved, the second-order thermal average can be analytically evaluated.

The interatomic potential $V(r)$ in diatomic Br_2 is appropriately described by the fourth-order polynomial as

$$V(r) = \frac{1}{2}\kappa_0(r-r_0)^2 - \kappa_3(r-r_0)^3 + \kappa_4(r-r_0)^4 + \dots \quad (3)$$

Here, κ_0 , κ_3 , and κ_4 are the harmonic, third-order and fourth-order force constants, respectively. The second-order perturbation is important for even-order cumulants since the P1 and P2 terms provide the same order of magnitude. On the other hand, odd-order cumulants may be described within P1. The third- and fourth-order terms in Eq. (3) are regarded as a perturbed Hamiltonian, while the harmonic term can be used as a non-perturbed one. Using the eigenvalues and eigenfunctions of the non-perturbed harmonic oscillator, the integral of Eq. (2) is analytically evaluated. The second-order partition function is consequently obtained as

$$Z = Z^{(0)} + Z^{(1)} + Z^{(2)} + \dots, \quad (4)$$

$$\begin{aligned} Z^{(0)} &= \frac{1}{1-z}, \\ Z^{(1)} &= \frac{\kappa_4\sigma_0^4}{k_B T} \frac{3(1+z)^2}{(1-z)^3}, \\ Z^{(2)} &= \frac{\kappa_3^2\sigma_0^6}{(\hbar\omega)(k_B T)} \frac{11z^2 + 38z + 11}{(1-z)^3}, \end{aligned}$$

where $\omega = \sqrt{\kappa_0/\mu}$ (μ is the reduced mass), $\sigma_0^2 = \hbar/(2\mu\omega)$ ($\hbar$ is the Planck constant divided by 2π), and $z = e^{-\beta\hbar\omega}$. The consequent formulas up to the fourth-order cumulant (P1 for C_1 and C_3 , and P2 for C_2 and C_4) are given as follows:

$$C_1 = \langle r - r_0 \rangle \cong C_1^{(1)} = \frac{6\kappa_3\sigma_0^4}{\hbar\omega} \frac{1+z}{1-z}, \quad (5)$$

$$C_2 = \langle (r - r_0 - C_1)^2 \rangle \cong C_2^{(0)} + C_2^{(1)} + C_2^{(2)}, \quad (6)$$

$$\begin{aligned} C_2^{(0)} &= \sigma_0^2 \frac{1+z}{1-z}, \\ C_2^{(1)} &= -\frac{\kappa_4\sigma_0^6}{\hbar\omega} \frac{12(1+z)^2}{(1-z)^2} - \frac{\kappa_4\sigma_0^6}{k_B T} \frac{24z(1+z)}{(1-z)^3}, \end{aligned}$$

$$C_2^{(2)} = \frac{\kappa_3^2 \sigma_0^8}{(\hbar\omega)^2} \frac{4(13z^2 + 58z + 13)}{(1-z)^2} + \frac{\kappa_3^2 \sigma_0^8}{(\hbar\omega)(k_B T)} \frac{24z(1+z)}{(1-z)^3},$$

$$C_3 = \langle (r - r_0 - C_1)^3 \rangle \cong C_3^{(1)} = \frac{\kappa_3 \sigma_0^6}{\hbar\omega} \frac{4(z^2 + 10z + 1)}{(1-z)^2}, \quad (7)$$

$$C_4 = \langle (r - r_0 - C_1)^4 \rangle - 3C_2^2 \cong C_4^{(1)} + C_4^{(2)}, \quad (8)$$

$$C_4^{(1)} = -\frac{\kappa_4 \sigma_0^8}{\hbar\omega} \frac{12(z^3 + 9z^2 + 9z + 1)}{(1-z)^3} - \frac{\kappa_4 \sigma_0^8}{k_B T} \frac{144z^2}{(1-z)^4},$$

$$C_4^{(2)} = \frac{\kappa_3^2 \sigma_0^{10}}{(\hbar\omega)^2} \frac{12(5z^3 + 109z^2 + 109z + 5)}{(1-z)^3} + \frac{\kappa_3^2 \sigma_0^{10}}{(\hbar\omega)(k_B T)} \frac{720z^2}{(1-z)^4}.$$

Note that the first-order perturbation terms of the even-order cumulants, $C_2^{(1)}$ and $C_4^{(2)}$ are negative, while $C_2^{(2)}$ and $C_4^{(1)}$ are positive.

2.2 Results

The EXAFS cumulants C_2 , C_3 , and C_4 for the intramolecular Br–Br shell were calculated numerically. The ECP formulation is omitted since it is fully established and found in the famous textbook of Prof. Kleinert [9]. The vibrational data [10] of $\kappa_0=2.459$ (mdyn/Å), $\kappa_3=1.756$ (mdyn/Å²) and $\kappa_4=1.058$ (mdyn/Å³) for Br₂ were employed. Figure 1 shows the ECP results, together with the P1 and P2 ones (the harmonic results are also given as HA for C_2). The classical and experimental EXAFS data are also plotted.

In the C_2 plot, the P1 result is slightly underestimated, especially at high temperature compared to the ECP one, because of a negative P1 value, while P2 agrees very well with ECP. Although C_4 is negative in P1, the P2 result approaches ECP. However, there seems to be still some underestimation for C_4 in P2 compared to ECP and this indicates that a higher-order perturbation calculation would be required.

In the case of C_3 , similar trends are observed. The P1 method estimates a little smaller value than ECP, again indicating a requirement of higher-order perturbation theory. The classical method shows essentially the same results as the ECP one at $T > 200$ K. As the temperature goes down from ~ 200 K, the classical value converges to zero monotonically as a function of T^2 and it gradually deviates from the two quantum statistical methods. This is caused by the zero-point anharmonicity. At $T < 100$ K, ECP begins to deviate from

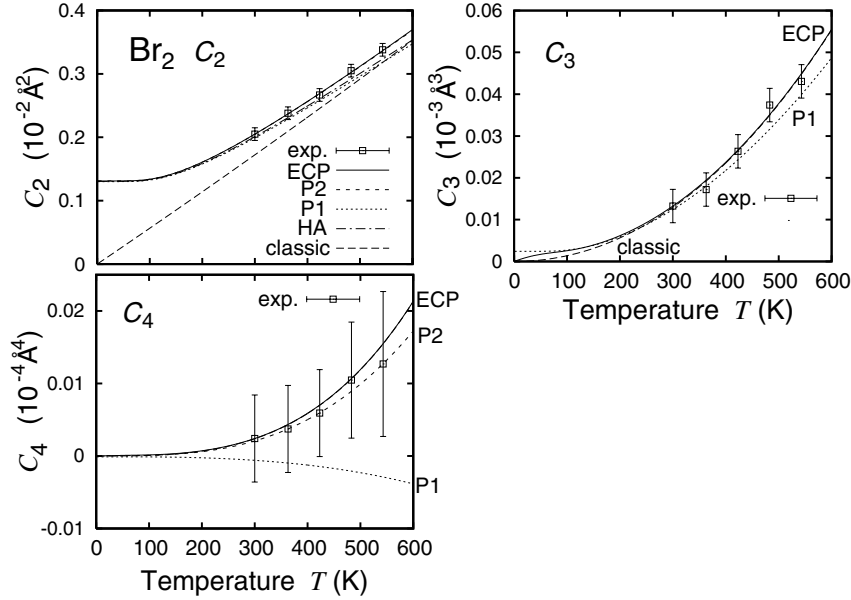


Figure 1. Temperature dependence of C_2 , C_3 , and C_4 of Br_2 evaluated by the ECP (solid lines), the classical (classic; long-dashed lines), and the P1 (dotted lines for C_2 , C_3 and C_4) and P2 (short-dashed lines for C_2 and C_4) methods, together with the experimental data (squares with error bars). For C_2 , the results of the harmonic approximation are also given (HA; dot-dashed line).

P1 and converges to zero at $T = 0$ K. In the present pair potential, there should be some finite (positive) C_3 at $T = 0$ K, implying that the P1 method predicts more appropriate C_3 at very low temperatures.

3 Anisotropic Vibration of a Thin Cu(111) Film

3.1 Computational Details

In the ECP calculations of bulk Cu, the formalism [12] is available within the low coupling approximation. The embedded-atom method (EAM) potential used here was taken from Ref. [11]. The normal vibrational analysis was initially performed to obtain the pure quantum mechanical fluctuations [12] using a cubic Brillouine zone $[-2\pi/a_0, 2\pi/a_0]$ (with the *fcc* lattice constant a_0) by sampling about 10^6 phonons. The classical-like *NPT* (constant pressure and temperature in a closed system) Monte Carlo (MC) calculations

were subsequently performed using the effective classical potential instead of the bare potential for 256 Cu atoms (4^3 *fcc* unit cells). Three-dimensional periodicity was imposed. 20 000 MC steps were calculated to reach the equilibrium and further 10 000 MC steps were performed to obtain the EXAFS cumulants, where each MC step contained 256-time movements of atoms and one-time variation of the lattice constant.

In the calculations of a thin Cu film, there exists no three-dimensional periodicity which leads to difficulties in applying the usual formula [12]. However, in order to obtain an estimation for the EXAFS cumulants to be compared with the experimental data, the quantum mechanical correction in ECP can be replaced by the bulk values, although this is rather crude from a pure theoretical point of view. This would give some overestimation of the vibrational amplitude of surface phonons at low temperature since the eigenfrequencies should be smaller than the bulk ones. Within this approximation, the ECP calculations can be similarly performed. A six-layer Cu(111) film was assumed, where each layer contained 48 atoms in a rectangular lattice and the lowest layer was assumed to be vibrationally fixed. Similar *NPT* MC simulations were performed to obtain the EXAFS cumulants of intra- and interlayer atom pairs.

3.2 Results

The calculated EXAFS cumulants for the first-nearest neighbor (NN) Cu-Cu shell of bulk Cu and the thin Cu(111) film are shown in Fig. 2. Therein, the obtained quantities for each layer are separately given for the thin film. Note that the first-and-second layer (12) distance as well as all the intra-layer distances, i.e. 11, 22, 33, and 44, are contracted compared to the bulk one, while the second-and-third (23) and third-and-fourth (34) layer distances are elongated. The thermal expansion, C_2 and C_3 are found to be significantly enhanced for the first-and-second layer.

When comparing the present theoretical results with the experimental data [14], polarization dependent Cu (and also Ni) *K*-edge EXAFS spectra were taken for ultrathin Cu(111) [or Ni(111)] films grown epitaxially on HOPG (highly oriented pyrolytic graphite) by varying the thickness. By using the results of thickness dependence of anisotropic C_2 and C_3 , the surface out-of-planar and in-planar components can be extracted. Moreover, by assuming (roughly) that all other contributions than the in-planar first-and-first (11) and the out-of-planar first-and-second (12) ones are equal to the bulk

Table 1. Effective Debye temperature Θ_D and difference of the third-order EXAFS cumulants ΔC_3 (difference between 100 and 300 K) for the surface out-of-planar, surface in-planar and bulk vibrations. Both experimental and theoretical results are given.

	out-of-plane		in-plane		bulk	
	Θ_D	ΔC_3	Θ_D	ΔC_3	Θ_D	ΔC_3
Calc.	272	3.65	290	2.47	313	1.48
Exptl.	262(25)	3.8(8)	322(30)	3.1(6)	338	1.62

one, three components of 11, 12 and bulk were successfully determined. For detailed procedures of the data analysis see Ref. [14]. The results are summarized in Table 1. In the experimental work, the obtained quantities are the difference between 100 and 300 K. For C_2 , the ΔC_2 values are replaced by the effective Debye temperature Θ_D , which is similarly evaluated by using the correlated Debye model [13]. Although the difference between the out-of-planar and in-planar vibrations is underestimated in the calculations, both results agree well with each other semi-quantitatively. It is concluded both experimentally and theoretically that the surface out-of-planar vibration is significantly softer and more anharmonic than the inner, lateral or bulk ones. As a reason for the underestimation in the calculations, one can suppose that the present theoretical calculations treat perfect films, while the actual films should contain many defects and some roughening might occur already at room temperature. This is indicated by the experimental findings [14] of smaller coordination number and larger C_2 . Since the surface area becomes wider in the presence of defects, the surface Debye temperature would effectively be lowered. One should also note that the surface Debye temperature of Cu(100) determined by LEED (low energy electron diffraction) [15] is 235 K, which is still lower than the EXAFS result of 262 K.

It is essentially important to recall intrinsic differences between EXAFS and diffraction. The diffraction techniques such as LEED provide information on absolute displacements with respect to the lattice, while EXAFS gives relative ones between x-ray absorbing and neighboring atoms. Although within the correlated Debye model the Debye temperature should be equal to the one given by the diffraction, C_2 and C_3 are the parameters obtained uniquely by EXAFS. It is natural that the out-of-planar vibration is enhanced at the

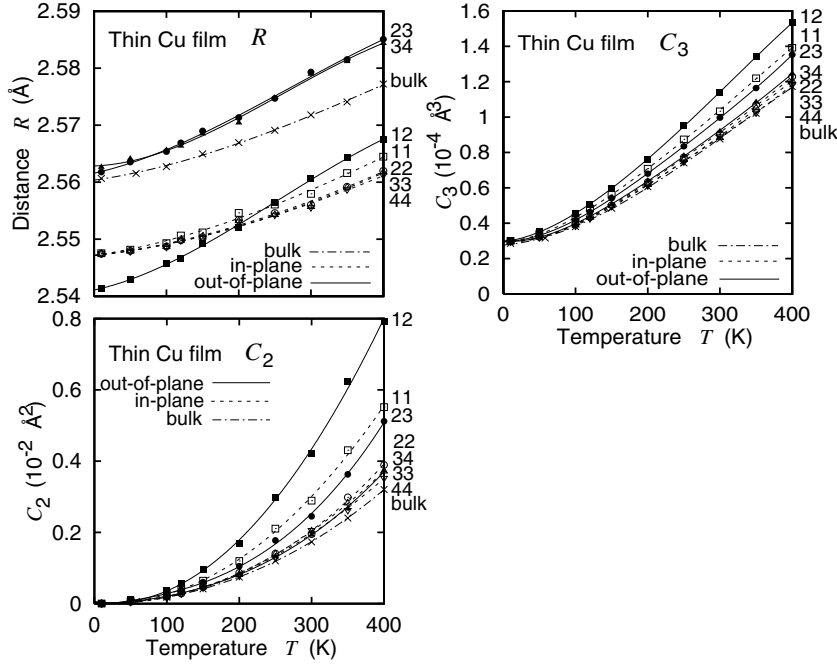


Figure 2. Temperature dependence of the interatomic distance R and the second- and third-order cumulants, C_2 and C_3 , for several first-NN Cu-Cu shells in a 6 ML Cu(111) film, together with the bulk values. For instance, 11 implies the first-and-first intra-layer distance, while 12 indicates the first-and-second interlayer distance (11: open square; 22: open circle; 33: upward triangle; 44: downward triangle; 12: filled square; 23: filled circle; 34: filled upward triangle; intra-layer: dashed line; interlayer: solid line; bulk: cross and dot-dashed line).

surface as long as the absolute displacement is discussed. This is simply ascribed to a lack of atoms upwards at the surface. On the contrary, from a local point of view, the out-of-planar motion is not always enhanced as can be seen in Refs. [16,17]. In the $p4g(2 \times 2)\text{N}/\text{Ni}(001)$ surface, the out-of-planar N-Ni bond was found to be much stiffer and less anharmonic than the lateral ones [16], and similarly the out-of-planar S-Ni bond in $c(2 \times 2)\text{S}/\text{Ni}(110)$ is stiffer and less anharmonic than the lateral ones [17]. In the present case, the surface out-of-planar bond is found to be weaker and more anharmonic than the in-planar one. The aim of this work is to obtain a hint of surface melting, which is an initial stage of bulk melting. The enhancement of the vibrational amplitude and anharmonicity along the surface normal is an important con-

clusion since a roughening transition should occur through the hopping of surface atoms for which not only the amplitude but also the anharmonicity is essential.

4 Discussion and Conclusions

In the present calculations for a one-dimensional system of Br_2 , some disadvantage of the ECP method was seen in the estimation of C_3 at low temperature. In Fig. 1(b), a strange decrease in $C_3(\text{ECP})$ was found at a temperature less than ~ 100 K. This is because in the ECP method the vibrational properties tend to be harmonic at the 0 K limit. It should also be true for C_4 , although C_4 is essentially zero and less important at low temperature. The perturbation theory can predict C_3 more accurate at low temperature. The ECP method is not reliable for strong quantum systems. In order to obtain such information, one has to perform quantum statistical perturbation calculations or more sophisticated path-integral Monte Carlo simulations. At higher temperatures, however, the P1 and P2 theory is insufficient to describe appropriate anharmonicity. Higher-order expansion would be required, and in the EXAFS analysis of usual systems common low-temperature anharmonicity is not very important. It can be concluded that the ECP method is more practical and reliable for most systems with one degree of vibrational freedom as well as many degrees of freedoms.

In summary, the ECP theory can treat the quantum effect, anharmonicity, many degrees of vibrational freedom, three-dimensional periodicity, higher-NN interaction and many-body interaction. It is essentially difficult or impossible to include all the contributions to other theories, at least in a practical sense of numerical calculations. It is also noted again that the EAM matches the ECP theory excellently, this allowing to investigate vibrational properties of metals quantum statistically.

In the simulations of thin $\text{Cu}(111)$ film, thermal vibration and local thermal expansion are larger than those for their corresponding bulk metals. The relative motions focused on the surface local bonds are enhanced in the surface normal direction. These findings are consistent with the EXAFS experiments. The enhanced vibrational amplitude and anharmonicity in the surface normal direction could be a trigger of a roughening transition, surface melting, and consequent bulk melting.

Acknowledgments

I gratefully acknowledge Prof. H. Kleinert of the Freie Universität Berlin for invaluable advice and discussion during and after my visit of Berlin in 1998. I would heartfully wish him a happy 60th birthday.

References

- [1] H. Rabus, Ph.D. thesis (Department of Physics, Freie Universität Berlin, 1991).
- [2] A.I. Frenkel and J.J. Rehr, *Phys. Rev. B* **48**, 585 (1993).
- [3] T. Fujikawa and T. Miyanaga, *J. Phys. Soc. Jpn.* **62**, 4108 (1993).
- [4] T. Yokoyama, K. Kobayashi, T. Ohta, and A. Ugawa, *Phys. Rev. B* **53**, 6111 (1996).
- [5] T. Yokoyama, Y. Yonamoto, T. Ohta, and A. Ugawa, *Phys. Rev. B* **54**, 6921 (1996).
- [6] T. Fujikawa, T. Miyanaga, and T. Suzuki, *J. Phys. Soc. Jpn.* **66**, 2897 (1997).
- [7] T. Yokoyama, *Phys. Rev. B* **57**, 3423 (1998).
- [8] T. Yokoyama, *J. Synchrotron Radiat.* **6**, 323 (1999).
- [9] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [10] K.P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV: Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).
- [11] S.M. Foiles, M.I. Basks, and M.S. Daw, *Phys. Rev. B* **33**, 7983 (1986).
- [12] A. Cuccoli, R. Giachetti, V. Tognetti, R. Vaia, and P. Verrucchi, *J. Phys. Condens. Matter* **7**, 7891 (1995).
- [13] G. Beni and P.M. Platzman, *Phys. Rev. B* **14**, 1514 (1976).
- [14] M. Kiguchi, T. Yokoyama, D. Matsumura, O. Endo, H. Kondoh, and T. Ohta, *Phys. Rev. B* **61**, 14020 (2000).
- [15] S. Müller, A. Kinne, M. Kottcke, R. Metzler, P. Bayer, L. Hammer, and K. Heinz, *Phys. Rev. Lett.* **75**, 2859 (1995).
- [16] L. Wenzel, D. Arvanitis, H. Rabus, T. Lederer, K. Baberschke, and G. Comelli, *Phys. Rev. Lett.* **64**, 1765 (1990).
- [17] T. Yokoyama, H. Hamamatsu, Y. Kitajima, Y. Takata, S. Yagi, and T. Ohta, *Surf. Sci.* **313**, 197 (1994).

THE HIGHEST-DERIVATIVE VERSION OF VARIATIONAL PERTURBATION THEORY

B. HAMPRECHT AND A. PELSTER

*Institut für Theoretische Physik, Freie Universität Berlin,
Arnimallee 14, D-14195 Berlin, Germany*

E-mails: bodo.hamprecht@physik.fu-berlin.de, axel.pelster@physik.fu-berlin.de

We systematically investigate different versions of variational perturbation theory by forcing not only the first or second but also higher derivatives of the approximant with respect to the variational parameter to vanish. The choice of the highest-derivative version turns out to be the most successful one for approximating the ground-state energy of the anharmonic oscillator. It is therefore used to determine the critical exponent α of the specific heat in superfluid ^4He in agreement with the value measured in recent space shuttle experiments.

1 Introduction

The perturbative treatment of quantum statistical or field-theoretical problems renders in general results in the form of divergent infinite power series in some coupling constant g . Typically, the coefficients of these series grow factorially with alternating signs leading to a zero convergence radius. Various resummation schemes may be applied to obtain finite results for all values of the coupling constant g , even in the strong-coupling limit $g \rightarrow \infty$ (for an overview see Chap. 16 of the book of Kleinert and Schulte-Frohlinde [1] and the references therein). Most successful is a recent systematic development by Kleinert [2], extending the variational method of Feynman and Kleinert [3] which was set up for calculating the effective classical potential in quantum statistics. It has been thoroughly tested for the ground-state energy of the anharmonic oscillator and shown to converge exponentially fast and uniform to the correct result [4]. This was encouraging enough to apply the method

also to divergent series which arise from renormalizing the ϕ^4 -theory of critical phenomena [1,5-9], where the perturbation coefficients are available up to six and partly to seven loops in $d = 3$ [10-12] and up to five loops in $d = 4 - \epsilon$ dimensions [13]. The method yielded finite results with a smooth dependence on the order. Furthermore, the theoretical results are in excellent agreement with the only experimental value available so far with an appropriate accuracy, the critical exponent α governing the behavior of the specific heat near the superfluid phase transition of ^4He [14,15].

Let us briefly recall the method. Consider some function $f(g)$ which is perturbatively obtained for a small coupling constant g as the divergent weak-coupling series

$$f(g) = \sum_{n=0}^{\infty} c_n g^n, \quad (1)$$

where the c_n denote the respective expansion coefficients. Kleinert's variational perturbation theory [2] replaces the series (1) by^a

$$f^N(g, p, q, w) = \left(\frac{g}{w}\right)^{pq} \sum_{n=0}^N c_n w^n \sum_{k=0}^{N-n} (-1)^k \binom{p - \frac{n}{q}}{k} \left(1 - \left(\frac{w}{g}\right)^q\right)^k. \quad (2)$$

By doing so, the newly introduced parameters N , p , q , w are specified as follows. N represents the order of the expansion and will be increased step by step as far as the knowledge of the weak-coupling coefficients c_n permits. The parameters p and q determine the asymptotic behavior of the function $f(g)$ in the strong-coupling limit $g \rightarrow \infty$ according to

$$f(g) = g^{pq} \sum_{j=0}^{\infty} b_j g^{-jq}, \quad (3)$$

where the b_j denote the strong-coupling coefficients. In quantum statistics the parameters p and q are usually known, e.g. from dimensional arguments, whereas in statistical field theory they are related to unknown critical exponents, so they have to be determined self-consistently from the weak-coupling coefficients c_n [5,6]. Finally, as the variational parameter w is introduced such that the approximant (2) will not depend on it in the limit $N \rightarrow \infty$, it should

^aIn contrast to the standard notation [2] the parameter q in Eq. (2) has been chosen in such a way that it coincides with the critical exponent Ω describing the approach to scaling in the field theory of critical phenomena [1].

be determined according to the principle of minimal sensitivity [16]. This principle is, however, no clear-cut mathematical statement and may therefore be implemented differently, giving rise to different versions of Kleinert's variational perturbation theory. Here we investigate and compare the versions, which define minimal sensitivity by the vanishing of a derivative of (2) with respect to w to some order k . Such a version is considered to be good, if its results converge well to the true value in the quantum statistical case, especially when p is known, but q is taken to be unknown, since this will anticipate the situation for the field-theoretical application. Another desirable feature can be seen in a very smooth dependence on the order N , since this will greatly enhance the possibility to extrapolate field-theoretical results to $N = \infty$. Finally, the simpler the version's defining prescription, the more easily it may be generalized to field-theoretical applications.

To begin with we focus our attention to the versions of variational perturbation theory in the strong-coupling limit $g \rightarrow \infty$. Comparing the Eqs. (2) and (3) gives, for the leading strong-coupling coefficient b_0 , the expression

$$b_0^N = (-1)^N w^{-pq} \sum_{n=0}^N c_n (-w)^n \binom{p-1-\frac{n}{q}}{N-n}, \quad (4)$$

where the variational parameter w is still to be optimized. For the ground-state energy of the anharmonic oscillator potential

$$V(x) = \frac{1}{2}x^2 + gx^4 \quad (5)$$

the leading strong-coupling coefficient has the numerical value [17]

$$b_0 = 0.667\,986\,259\,155\,777\,108\,270\,96\dots, \quad (6)$$

and dimensional arguments require $q = 2/3$, $p = 1/2$. The weak-coupling coefficients c_n are derived in the Appendix (see Table 5).

2 The Highest-Derivative Version

Traditionally, the expression (4) for b_0 is made stationary by forcing its first or second derivative with respect to the variational parameter w to vanish, depending on whether the order N is odd or even. Here we investigate whether some higher derivatives may be used instead. To this end we consider, for every order N , all derivatives $\partial^k b_0^N / \partial w^k$ with $k = 1, \dots, N+4$ and determine

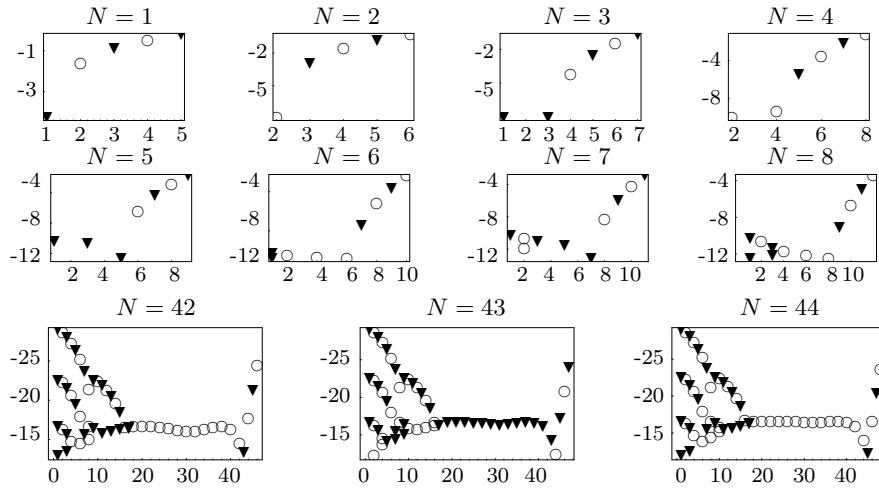


Figure 1. The approach of b_0^N to its true value b_0 is measured by the quantity $\log |b_0^N - b_0|$ which is plotted over k , where k appears in the condition $\partial^k b_0^N / \partial w^k = 0$ determining the variational parameter w . Some typical values for the order N have been depicted with different symbols ▼ and ○ used to plot odd and even k , respectively. Optimal results are obtained for $k = 1$ or $k = 2$, but with no simple rule as to which of the two has to be chosen. Taking $k = N$, however, the same quality is obtained with no ambiguity of choice. Also for $k = N - 2$, the outcome is very reasonable. Note that for some of the lower values of k there is more than one solution and that for some k , like e.g. for $k = N - 1$, there is no solution at all.

all real positive zeros w . For each of these zeros, the quantity $\log |b_0^N - b_0|$ measures the quality of approximating b_0 in Eq. (6) by b_0^N in Eq. (4) (see Fig. 1).

Optimal results are obtained for $k = N$. Moreover, the prescription is unique, since for $k = N$ there exists only one real positive zero w of $\partial^k b_0^N / \partial w^k$. It should be noted that for $k = N - 1$ there never exists a real positive zero. Another unique and almost optimal prescription would be $k = N - 2$ which works well for all $N > 2$. The results for $k = 1$ and $k = 2$ are of comparable quality with the values for $k = N$. Although only one of them has a real positive zero up to $N = 5$, the equations for $k = 1$ and $k = 2$ tend to have more than one solution for larger N , the one with the largest value for w always producing the best result. But there is no simple rule to decide whether $k = 1$ or $k = 2$ gives the better value. The initial indication of using $k = 1$ for odd and $k = 2$ for even orders does not carry through beyond $N = 6$.

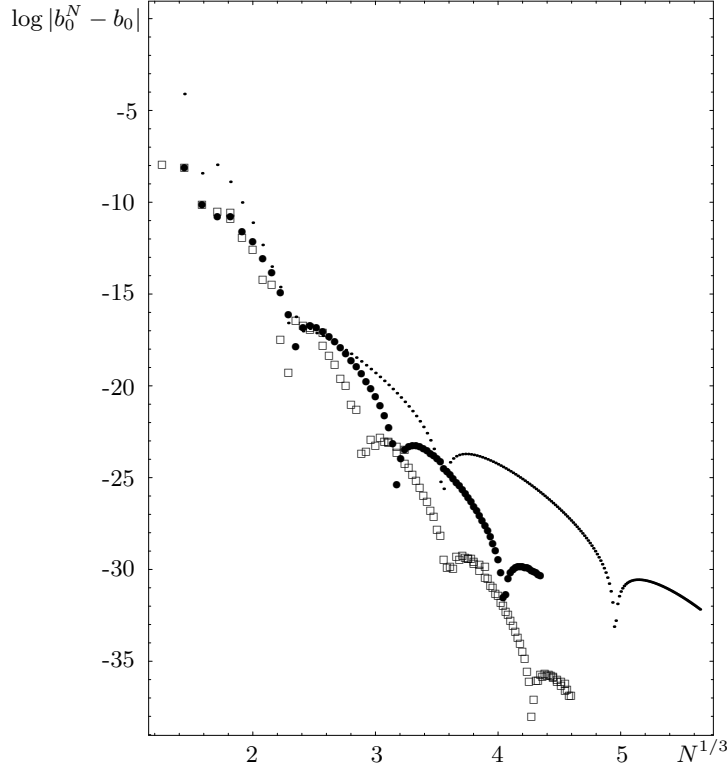


Figure 2. The error $\log |b_0^N - b_0|$ plotted over $N^{1/3}$ as obtained by various methods: • for $k = N$, ◦ for a quadratic approximation to $k = N - 2$, and ◻ for optimal values from $k = 1$ or $k = 2$, respectively.

In Fig. 2 we compare the results for $k = N$ with the optimal branch chosen judiciously from the equations for $k = 1$ and $k = 2$. We also show a quadratic approximation to the $k = N - 2$ equation where only the last three terms of the condition $\partial^{N-2} b_0^N / \partial w^{N-2} = 0$ are kept, leaving us with a quadratic equation for the variational parameter w . It can be seen that this approximation is of high quality, as was to be expected. This may be understood as follows: the coefficients of the w -expansion of b_0^N in Eq. (4) alternate in sign and grow faster than $n!$, which is the same for the weak-coupling coefficients c_n from which the former are derived. Therefore the series (4), or any derivative thereof will preferably become zero, if neighbouring terms in the sum nearly

cancel each other. The failure of such cancellation among some high-order neighbours cannot easily be compensated by lower-order terms because of the large-order behavior of the coefficients. Hence the vanishing of the derivative for $k = N - 2$ can well be approximated with little loss of accuracy by choosing w in such a way, that only the three highest-order terms cancel each other. Thus we conclude that $k = N - 2$ represents the highest non-trivial derivative version for applications in statistical field theory, where we usually have $p = 0$, so the strong-coupling coefficient b_0^N in Eq. (4) is a polynomial in w . It becomes the simpler the higher derivatives are being used.

3 Self-Consistent Determination of the Parameter q

In this section we evaluate again the leading coefficient b_0 in the strong-coupling expansion for the ground-state energy of the anharmonic oscillator. But this time the parameter q is not fixed to its proper value $q = 2/3$. Instead we draw this information from the same weak-coupling coefficients c_n , whereas p is set to its known value $p = 1/2$. In order to determine q we need an additional equation. This usually results from a self-consistent reasoning by evaluating the logarithmic derivative of the strong-coupling series (3) of $f(g)$ in the strong-coupling limit $g \rightarrow \infty$ [5,6]:

$$\lim_{g \rightarrow \infty} \frac{gf'(g)}{f(g)} = \lim_{g \rightarrow \infty} \frac{g^{pq} \sum_{j=0}^{\infty} b_j q(p-j) g^{-jq}}{g^{pq} \sum_{j=0}^{\infty} b_j g^{-jq}} = pq = \frac{q}{2}. \quad (7)$$

Defining $F(g) := gf'(g)/f(g)$, the weak-coupling coefficients γ_n of $F(g)$ are determined from the corresponding c_n of $f(g)$ in Eq. (1). The strong-coupling limit of $F(g)$ can then be constructed with the help of variational perturbation theory by using γ_n instead of c_n in quite an analogous way as before. Since $F(g)$ tends to the constant value $q/2$ in the strong-coupling limit $g \rightarrow \infty$, its parameter p has to be zero, whereas q must have the same value as before. Thus, from the Eqs. (4) and (7), we obtain for the strong-coupling limit of $F(g)$

$$(-1)^N \sum_{n=0}^N \gamma_n (-w)^n \binom{-1 - \frac{n}{q}}{N - n} = \frac{q}{2}, \quad (8)$$

where we determine the variational parameter w on the left-hand side by demanding that the k th derivative with respect to w must vanish, k being chosen appropriately. This set of equations can be rearranged into two polynomials of q and w which have to vanish simultaneously. The complete set of complex zeros becomes legion even for moderate orders N , their number growing with N^2 . Constraining ourselves to real and positive solutions for q in the expected range $0.5 < q < 0.9$, we find, in contrast to the previous case, a very regular behavior. Exactly one unique solution exists for all N up to $N = 36$ and for all k if $N + k$ is an odd integer, whereas there is no solution within this range if $N + k$ is even. The solutions for different k approach the value $q = 2/3$ closer for smaller k . As k increases, the loss of accuracy, however, is tolerable, as can be seen from Fig. 3, where again the error $\log |b_0^N - b_0|$ is plotted over $N^{1/3}$, the approximation b_0^N being obtained as before with the highest-derivative version, but using the self-consistently determined q instead of the exact $q = 2/3$. Some loss of accuracy is the price to pay, if we want to make one of both equations linear by choosing $k = N - 1$, which is now the highest non-trivial derivative. Notice that the penalty is relatively low for small orders N , which in field-theoretical models are the only ones available at present.

4 Critical Exponent α for Liquid Helium

After having analyzed the highest-derivative versions of variational perturbation theory by the example of the ground-state energy of the anharmonic oscillator, we apply this method to liquid ^4He . In particular, we consider its superfluid state near the transition point T_c , in order to calculate the critical exponent α governing the power behavior $|T - T_c|^{-\alpha}$ of the specific heat. Within the framework of the ϕ^4 -theory of critical phenomena, the superfluid phase transition of ^4He is described by a complex order-parameter field ϕ whose bare energy functional is of the Ginzburg-Landau type and reads in $d = 3$ dimensions [1]:

$$E[\phi_B] = \int d^3x \left\{ \frac{1}{2} [\partial\phi_B(x)]^2 + \frac{1}{2} m_B^2 \phi_B^2(x) + \frac{\pi}{5} g_B [\phi_B^2(x)]^2 \right\}. \quad (9)$$

By calculating the Feynman diagrams, one encounters divergencies which are removed by analytic regularization [18]. A subsequent renormalization leads to renormalized values of mass, coupling constant and field which are related to the bare input quantities by the respective renormalization constants Z_m ,

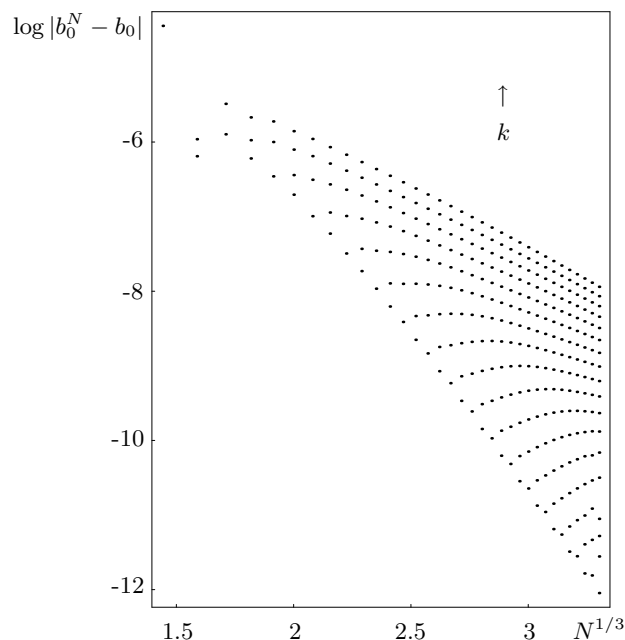


Figure 3. For every N several q -values are obtained by setting the k th derivative of the strong-coupling limit of $F(g) = g f'(g)/f(g)$ with respect to the variational parameter w equal to zero. For each of these self-consistent values of q the error $\log |b_0^N - b_0|$ is plotted over $N^{1/3}$. Smaller k give better results, but the highest k -values up to $k = N - 1$ are also tolerable.

Z_g and Z_ϕ :

$$m_B^2 = m^2 Z_m Z_\phi^{-1}, \quad g_B = g Z_g Z_\phi^{-2}, \quad \phi_B = \phi Z_\phi^{1/2}. \quad (10)$$

In the literature, one finds expansions for these renormalization constants and for certain logarithmic derivatives, the so-called renormalization group functions, up to six and partly up to seven loops [10-12]. All these expansions can be written in terms of the dimensionless bare coupling constant $\bar{g}_B = g_B/m$. For instance, one obtains, for the dimensionless renormalized coupling constant $\bar{g} = g/m$, the series

$$\bar{g}(\bar{g}_B) = \sum_{n=0}^{\infty} \kappa_n \bar{g}_B^n. \quad (11)$$

The logarithmic derivative of the square mass ratio

$$\eta_m(\bar{g}_B) = -\frac{d \log \frac{m^2}{m_B^2}}{d \log \bar{g}_B} \quad (12)$$

reads correspondingly

$$\eta_m(\bar{g}_B) = \sum_{n=0}^{\infty} \lambda_n \bar{g}_B^n. \quad (13)$$

Its weak-coupling coefficients κ_n and λ_n have been listed in Table 1, respectively. In the field-theoretic description of second-order phase transitions, the renormalized square mass m^2 vanishes near the transition point T_c . Thus the approach to scaling corresponds to the strong-coupling limit of the dimensionless bare coupling constant $\bar{g}_B$ [1,5-9]. Expecting to obtain finite results for both (11) and (13), when the dimensionless bare coupling constant $\bar{g}_B$ tends to infinity, the strong-coupling expansions for $\bar{g}(\bar{g}_B)$ and $\eta_m(\bar{g}_B)$ should read

$$\bar{g}(\bar{g}_B) = \bar{g} + \sum_{j=1}^{\infty} \beta_j \bar{g}_B^{-j\Omega}, \quad (14)$$

$$\eta_m(\bar{g}_B) = \eta_m + \sum_{j=1}^{\infty} \gamma_j \bar{g}_B^{-j\Omega}. \quad (15)$$

Note that strong-coupling expansions like (14) and (15) are governed by one and the same critical exponent Ω [1], a common feature of second-order phase transitions. Thus from either of these series Ω may be extracted in various ways, taking logarithmic derivatives with respect to the dimensionless bare coupling constant $\bar{g}_B$. The strong-coupling behavior (14) of the dimensionless renormalized coupling constant $\bar{g}(\bar{g}_B)$ implies, for instance,

$$\lim_{\bar{g}_B \rightarrow \infty} \frac{d \log \bar{g}(\bar{g}_B)}{d \log \bar{g}_B} = 0, \quad (16)$$

$$\lim_{\bar{g}_B \rightarrow \infty} \frac{d \log \bar{g}'(\bar{g}_B)}{d \log \bar{g}_B} = -\Omega - 1, \quad (17)$$

with corresponding results for $\eta_m(\bar{g}_B)$, as can easily be seen from Eq. (15). Forcing these relations upon the weak-coupling expansions as well, we obtain equations to which the strong-coupling limit of variational perturbation theory can be applied. We use its highest-derivative version as explained above

Table 1. Weak-coupling coefficients for $\bar{g}(\bar{g}_B)$ and $\eta_m(\bar{g}_B)$ from 6- and 7-loop perturbation theory [10–12] (above the tripple line) and from extrapolations based on instanton calculations [7,18] (below the tripple line).

Weak-Coupling Coefficients		
	$\eta_m = \sum \lambda_n \bar{g}_B^n$	$\bar{g} = \sum \kappa_n \bar{g}_B^n$
n	λ_n	κ_n
0	0	0
1	4	1
2	−46.814 814 814 814 82	−10
3	667.389 693 519 318 6	120.148 148 148 148 1
4	−10 792.618 387 448 09	−1 642.256 264 617 284
5	191 274.332 379 005 1	24 816.045 615 887 86
6	−3 644 347.117 315 811	−407 363.539 559 348
7	7 378 080.984 900 002	
7		$7.180\,326\,000\,784\,143 \times E^6$
8	$-1.575\,312\,899\,817\,985 \times E^9$	$-1.347\,981\,388\,551\,665 \times E^8$
9	$3.529\,433\,822\,947\,775 \times E^{10}$	$2.679\,259\,494\,762\,891 \times E^9$
10	$-8.269\,004\,838\,427\,051 \times E^{11}$	$-5.612\,792\,935\,327\,37 \times E^{10}$
11	$2.020\,940\,372\,700\,199 \times E^{13}$	$1.234\,985\,211\,831\,956 \times E^{12}$
12	$-5.143\,391\,961\,273\,287 \times E^{14}$	$-2.846\,297\,378\,479\,211 \times E^{13}$
13	$1.360\,628\,154\,286\,311 \times E^{16}$	$6.837\,009\,909\,070\,767 \times E^{14}$
14	$-3.734\,409\,500\,947\,708 \times E^{17}$	$-1.699\,763\,366\,416\,082 \times E^{16}$
15	$1.062\,320\,774\,475\,501 \times E^{19}$	$4.354\,392\,675\,760\,15 \times E^{17}$
16	$-3.132\,670\,751\,194\,59 \times E^{20}$	$-1.151\,470\,001\,163\,625 \times E^{19}$
17	$9.584\,565\,077\,380\,565 \times E^{21}$	$3.166\,449\,929\,095\,059 \times E^{20}$
18	$-3.044\,620\,939\,704\,96 \times E^{23}$	$-9.135\,728\,488\,698\,395 \times E^{21}$
19	$1.004\,119\,260\,969\,699 \times E^{25}$	$2.780\,684\,275\,579\,437 \times E^{23}$
20	$-3.435\,379\,417\,189\,77 \times E^{26}$	$-8.925\,916\,735\,892\,876 \times E^{24}$
21	$1.217\,685\,465\,986\,233 \times E^{28}$	$3.006\,538\,451\,329\,012 \times E^{26}$
22	$-4.465\,282\,353\,357\,697 \times E^{29}$	$-1.055\,6247\,984\,629\,23 \times E^{28}$
23	$1.691\,830\,517\,012\,908 \times E^{31}$	$3.841\,496\,075\,948\,586 \times E^{29}$
24	$-6.616\,032\,677\,503\,413 \times E^{32}$	$-1.443\,458\,944\,165\,208 \times E^{31}$

and find good agreement between the Ω_N -values from each of the four equations (see Table 2). In fact they are supposed to approach the same limit as $N \rightarrow \infty$. We notice that the members of each pair obtained from either

Table 2. The critical exponent Ω as obtained from the strong-coupling limit of the first or the second logarithmic derivatives of the weak-coupling series for $\bar{g}(\bar{g}_B)$ and $\eta_m(\bar{g}_B)$, respectively.

Ω -Values				
N	$\Omega(\bar{g}, 1)$	$\Omega(\bar{g}, 2)$	$\Omega(\eta_m, 1)$	$\Omega(\eta_m, 2)$
2	0.730 495	0.735 397	0.715 930	0.721 303
3	0.751 627	0.751 166	0.714 270	0.712 744
4	0.757 596	0.757 762	0.703 544	0.700 880
5	0.763 865	0.763 975	0.705 086	0.704 576
6			0.714 586	0.714 540

the η_m - or the $\bar{g}$ -expansion are almost the same, whereas the different pairs have not yet converged so well for $N \leq 6$. Even though the data from the $\bar{g}$ -expansion look more promising because of their smoother behavior, we still use the results Ω_N from the η_m -expansion alone to calculate the critical exponent α , and do not use any information from the $\bar{g}$ -expansion at this stage. We evaluate the critical exponent η_m from

$$\eta_m^N = (-1)^N \sum_{n=0}^N \lambda_n (-w_N)^n \binom{-1 - \frac{n}{\Omega_N}}{N - n}, \quad (18)$$

with the variational parameter w_N fixed by

$$w_N = -\frac{\lambda_{N-1}(N-1+\Omega_N)}{\lambda_N N \Omega_N}. \quad (19)$$

Here the Ω_N are supplied from the $\Omega(\eta_m, 1)$ column of Table 2. The corresponding results are shown in Table 3 along with the corresponding values for the critical exponent α of the specific heat which follows from

$$\alpha = \frac{1 - 2\eta_m}{\eta_m - 2}. \quad (20)$$

These are values quite close to the experimental data $\alpha_{\text{exp}} = -0.01056 \pm 0.00038$ [15], thus giving support for the method used.

5 Improvement by Extrapolation

Higher-order perturbation coefficients for $\bar{g}(\bar{g}_B)$ and $\eta_m(\bar{g}_B)$ are known approximately (see Table 1). They have been derived from instanton calcu-

Table 3. The critical exponent α for the specific heat of liquid helium, derived from the perturbation expansion for η_m using the highest-derivative version of variational perturbation theory.

α -Values		
N	η_m	α
2	0.490 834	+0.0121
3	0.513 786	−0.0185
4	0.522 480	−0.0304
5	0.522 651	−0.0307
6	0.519 592	−0.0265

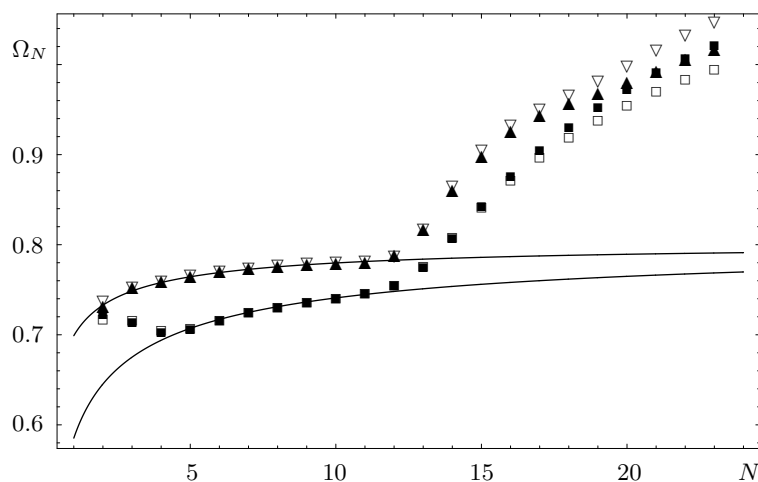


Figure 4. The critical exponent Ω derived from the strong-coupling limit of the first and the second logarithmic derivative of the weak-coupling expansion of $\bar{g}(\bar{g}_B)$ ($\blacktriangle$, ∇) and of $\eta_m(\bar{g}_B)$ ($\blacksquare$, $\square$), respectively. The solid lines are extrapolations with $\Omega_\infty = 0.8$ fitted to the $\bar{g}$ - and η_m -points in the reliable interval $3 < N < 12$, respectively.

lations, which were fitted to the weak-coupling data [7,18]. Extending the previous calculation on this basis up to the order of $N = 24$, good convergence of all four different sets of the Ω_N is found. We notice, however, a kink at about $N = 12$ in Fig. 4, which strongly suggests, that the extrapolation is no longer reliable beyond this point. For large N the results Ω_N have the

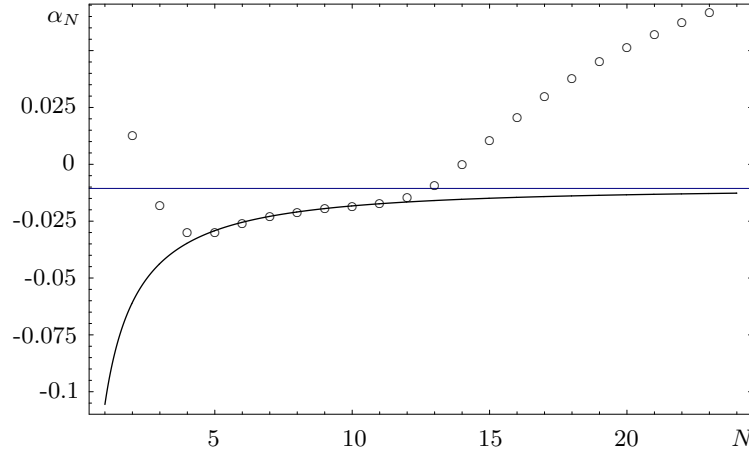


Figure 5. The critical exponent α of the specific heat of liquid ^4He near the superfluid phase transition plotted against the order N . The straight horizontal line represents the experimental value of $\alpha_{\text{exp}} = -0.0106$. The curved line is an extrapolation of the results from the range of $3 < N < 12$.

asymptotic form [1,5–9]

$$\Omega_N = \Omega_\infty - a \exp(-b N^{1-\Omega_\infty}), \quad (21)$$

where the constants a and b are different for the η_m - and the $\bar{g}$ -expansion. Unfortunately, the fit is extremely insensitive to the value of Ω_∞ : all values in a broad interval around $\Omega_\infty = 0.8$ would fit the data up to $N = 12$ from both series with very much the same quality (see Fig. 4). In Fig. 5 the resulting α -values are plotted against the asymptotic form

$$\alpha_N = \alpha_\infty - c \exp(-d N^{1-\Omega_\infty}), \quad (22)$$

for $\alpha_\infty = -0.0106$ and $\Omega_\infty = 0.8$. Up to the kink at $N = 12$, the data seem to approach the experimental number excellently.

6 Quality Check for the Highest-Derivative Version

The four different methods, discussed before to extract the critical exponent Ω from the data, should all give results which converge to the same limit as $N \rightarrow \infty$. We show now that the highest-derivative version reveals a better convergence than the traditional method which uses the first or second derivatives instead. The details are reported in Table 4, where the variance

σ_Ω between the values $\Omega_N^{(1)}$, $\Omega_N^{(2)}$ from different methods is listed for specific ranges of N :

$$\sigma_\Omega = \sqrt{\frac{1}{N_1 - N_0 + 1} \sum_{N=N_0}^{N_1} (\Omega_N^{(1)} - \Omega_N^{(2)})^2}. \quad (23)$$

In the first row of Table 4 we list the Ω -values extracted from the weak-coupling expansion (11) of $\bar{g}(\bar{g}_B)$ according to the methods (16), (17) using the highest-derivative version. The same quantity evaluated in the traditional way, i.e. using the first derivative for even N and the second derivative for odd N , is shown, for comparison, in the second row. It can be seen that both methods are of the same quality in the range of $2 \leq N \leq 6$, for which the calculation is based on the weak-coupling coefficients from perturbation theory, and also in the range of $7 \leq N \leq 12$, based on extrapolated values. Only for $N > 12$, where the extrapolation is no more reliable, anyhow, the highest-derivative version seems to be worse than the one obtained by the traditional method. However, in order to obtain these apparently better results using low order derivatives, the rules of the game had to be changed somewhat, taking the second derivative throughout for even and odd order N , as soon as N becomes larger than 12. The next two rows show the corresponding comparison based on the weak-coupling expansion (13) of $\eta_m(\bar{g}_B)$. Here it turns out that the highest-derivative version gives results of significantly better quality than the traditional method, up to $N = 12$. Finally, in the last two rows the deviations of the Ω -values from the weak-coupling expansions (11) of $\bar{g}(\bar{g}_B)$ and (13) of $\eta_m(\bar{g}_B)$, extracted according to (16), are shown in comparison for the highest-derivative version (above) and the traditional method (below). It can be seen that both methods are more or less comparable in quality, the highest-derivative version giving slightly better results.

Appendix

The perturbative solution of the time-independent Schrödinger equation

$$H\Psi_m := (H_0 + gV)\Psi_m = E_m\Psi_m \quad (24)$$

usually leads to infinite sums for the expansion coefficients of the energy eigenvalues $E_m(g)$ and the state vectors $\Psi_m(g)$. But if the perturbing potential V happens to be band-diagonal in the basis of the eigenvectors of the unper-

Table 4. Comparison of quality for the highest-derivative version and for the traditional method, based on convergence of the critical exponent Ω obtained in two different ways from two different sets of data.

		σ_Ω		
data	method	$2 \leq N \leq 6$	$7 \leq N \leq 12$	$13 \leq N \leq 24$
$\bar{g}(\bar{g}_B)$	high	0.002 159	0.000 570	0.017 682
$\bar{g}(\bar{g}_B)$	low	0.002 245	0.000 522	0.000 274
$\eta_m(\bar{g}_B)$	high	0.002 749	0.000 258	0.016 251
$\eta_m(\bar{g}_B)$	low	0.008 214	0.008 121	0.005 915
both	high	0.046 663	0.040 836	0.037 912
both	low	0.046 549	0.057 049	0.055 477

turbed Hamiltonian H_0 , then all these sums become finite and the eigenvalues $E_m(g)$ can be determined recursively to all orders.

Here we assume that the spectrum $\epsilon_0^{(0)}, \epsilon_0^{(1)}, \epsilon_0^{(2)}, \dots, \epsilon_0^{(m)}, \dots$ of the unperturbed Hamiltonian H_0 is non-degenerate and we denote the corresponding eigenvectors by $|0\rangle, |1\rangle, |2\rangle, \dots, |m\rangle, \dots$. Following the usual procedure, the energy eigenvalues $E_m(g)$ and the state vectors $\Psi_m(g)$ are expanded in powers of g :

$$\Psi_m(g) = \sum_{n,i} \gamma_{n,i}^{(m)} g^i \alpha_n |n\rangle, \quad (25)$$

$$E_m(g) = \sum_j \epsilon_j^{(m)} g^j, \quad (26)$$

where the α_n are inserted into the definition, for later computational convenience. They will only show up in the coefficients of the state vector Ψ_m , but not in the expansion of the energy eigenvalues E_m which we are looking for. Without loss of generality the normalization of the state vectors Ψ_m is chosen such that $\langle \Psi_m | m \rangle = \alpha_m$ to all orders, so we have

$$\gamma_{m,i}^{(m)} = \delta_{i,0}, \quad \gamma_{k,0}^{(m)} = \delta_{k,m}. \quad (27)$$

Inserting the ansatz (25), (26) into the Schrödinger equation (24), projecting the result onto the base vector $\langle k | \alpha_k$ and extracting the coefficient of g^j , we

obtain the relation:

$$\gamma_{k,i}^{(m)} \epsilon_0^{(k)} + \sum_n \frac{\alpha_n}{\alpha_k} V_{k,n} \gamma_{n,i-1}^{(m)} = \sum_j \epsilon_j^{(m)} \gamma_{k,i-j}^{(m)}, \quad (28)$$

with the matrix elements

$$V_{k,n} := \langle k | V | n \rangle. \quad (29)$$

For $i = 0$ this equation is identically satisfied. For $i \neq 0$ we get the following two relations, one for $k = m$ and the other one for $k \neq m$:

$$\epsilon_i^{(m)} = \sum_n \gamma_{m+n,i-1}^{(m)} W_{m,n}, \quad (30)$$

$$\gamma_{k,i}^{(m)} = \frac{1}{\epsilon_0^{(k)} - \epsilon_0^{(m)}} \left[\sum_{j=1}^{i-1} \epsilon_j^{(m)} \gamma_{k,i-j}^{(m)} - \sum_n \gamma_{k+n,i-1}^{(m)} W_{k,n} \right]. \quad (31)$$

In these expressions all sums are finite if the potential V is band-diagonal such that the augmented matrix elements

$$W_{m,n} := V_{m,m+n} \frac{\alpha_{m+n}}{\alpha_m} \quad (32)$$

are different from zero only for $-d < n < d$, where d is some finite number. For any quantum number m these relations may now be applied recursively with respect to the order, leading to $\epsilon_1^{(m)}, \gamma_{k,1}^{(m)}, \epsilon_2^{(m)}, \gamma_{k,2}^{(m)}, \dots, \epsilon_r^{(m)}, \gamma_{k,r}^{(m)}, \dots$ in turn. Notice that for given m and r only a finite number of the $\gamma_{k,r}^{(m)}$ is non-zero.

Considering in particular the anharmonic oscillator with the unperturbed Hamiltonian $H_0 = (p^2 + x^2)/2$ and the perturbing potential $V = x^4$, we may choose $\alpha_k = \sqrt{2^k/k!}$ to get the non-vanishing augmented matrix elements from (29) and (32). An easy way to calculate the augmented matrix elements starts from the ones for $V = x$, where we have

$$W_{k,-1}^x = k/2, \quad (33)$$

$$W_{k,1}^x = 1. \quad (34)$$

From here the augmented matrix elements $W_{k,j}^{x^l}$ for any polynomial $V = x^l$ can be determined recursively by applying the rules of matrix multiplication:

$$W_{k,j}^{x^l} = \sum_i W_{k,i}^x W_{k+i,j-i}^{x^{l-1}}. \quad (35)$$

Table 5. The first ten coefficients $\epsilon_k^{(0)}$ of the weak-coupling expansion for the ground-state energy $E_0(g)$ of the anharmonic oscillator.

The Anharmonic Oscillator	
k	$\epsilon_k^{(0)}$
0	$1/2$
1	$3/4$
2	$-21/8$
3	$333/16$
4	$-30\,885/128$
5	$916\,731/256$
6	$-65\,518\,401/1\,024$
7	$2\,723\,294\,673/2\,048$
8	$-1\,030\,495\,099\,053/32\,768$
9	$54\,626\,982\,511\,455/65\,536$
10	$-6\,417\,007\,431\,590\,595/262\,144$

Thus we obtain for $V = x^4$:

$$W_{k,-4}^{x^4} = k(k-1)(k-2)(k-3)/16, \quad (36)$$

$$W_{k,-2}^{x^4} = k(2k-1)(k-1)/4, \quad (37)$$

$$W_{k,0}^{x^4} = 3(2k^2 + 2k + 1)/4, \quad (38)$$

$$W_{k,2}^{x^4} = 2k + 3, \quad (39)$$

$$W_{k,4}^{x^4} = 1. \quad (40)$$

Being rational numbers they are suitable for recursive calculations even up to very high orders. In Table 5 we have listed the first ten coefficients $\epsilon_k^{(0)}$ for the expansion of the ground-state energy $E_0(g)$.

Acknowledgments

It is a particular pleasure for B.H. to contribute to this volume in honor of Hagen Kleinert, looking back gratefully on four decades of friendship originating in the physics undergraduate courses of Technische Universität Hannover, incorporating a common year at the University of Colorado at Boulder as well

as the last 32 years as colleagues together at the Freie Universität Berlin. This is written with great admiration for Hagen's deep dedication to the miracles of nature and their physical interpretation.

References

- [1] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [2] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [3] R.P. Feynman and H. Kleinert, *Phys. Rev. A* **34**, 5080 (1986).
- [4] W. Janke and H. Kleinert, *Phys. Rev. Lett.* **75**, 2787 (1995); H. Kleinert and W. Janke, *Phys. Lett. A* **206**, 283 (1995).
- [5] H. Kleinert, *Phys. Rev. D* **57**, 2264 (1998); Addendum, *Phys. Rev. D*, **58**, 107702 (1998).
- [6] H. Kleinert, *Phys. Lett. B* **434**, 74 (1998).
- [7] H. Kleinert, *Phys. Rev. D* **60**, 085001 (1999).
- [8] H. Kleinert *Phys. Lett. B* **463**, 69 (1999).
- [9] H. Kleinert and V. Schulte-Frohlinde, *J. Phys. A* **34**, 1037 (2001).
- [10] B.G. Nickel, D.I. Meiron, and G.A. Baker Jr., University of Guelph, preprint (1977), http://www.physik.fu-berlin.de/~kleinert/kleiner_reb8/programs/programs.html.
- [11] S.A. Antonenko and A.I. Sokolov, *Phys. Rev. E* **51**, 1894 (1995).
- [12] D.B. Murray and B.G. Nickel, University of Guelph, preprint (1998).
- [13] H. Kleinert, J. Neu, V. Schulte-Frohlinde, K.G. Chetyrkin, and S.A. Larin, *Phys. Lett. B* **272**, 39 (1991); *Phys. Lett. B* **319**, 545(E) (1993).
- [14] H. Kleinert, *Phys. Lett. A* **277**, 205 (2000).
- [15] J.A. Lipa, D.R. Swanson, J.A. Nissen, T.C.P. Chui, and U.E. Israelsson, *Phys. Rev. Lett.* **76**, 944 (1996); J.A. Lipa, D.R. Swanson, J.A. Nissen, Z.K. Geng, P.R. Williamson, D.A. Stricker, T.C.P. Chui, U.E. Israelsson, and M. Larson, *Phys. Rev. Lett.* **84**, 4894 (2000). In the second paper, the value for the critical exponent of the specific heat of superfluid ^4He given in the first paper was corrected to $\alpha_{\text{exp}} = -0.01056 \pm 0.0004$.
- [16] P.M. Stevenson, *Phys. Rev. D* **23**, 2916 (1981).
- [17] F. Vinette and J. Čížek, *J. Math. Phys.* **32**, 3392 (1991).
- [18] G. Parisi, *J. Stat. Phys.* **23**, 49 (1980).

FAST-CONVERGENT RESUMMATION ALGORITHM AND CRITICAL EXPONENTS OF ϕ^4 -THEORY IN THREE DIMENSIONS

F. JASCH

*Theoretische Polymerphysik, Albert-Ludwig-Universität Freiburg,
D-79104 Freiburg i.Br., Germany
E-mail: florian@physik.uni-freiburg.de*

During 1998 I had the opportunity to work towards my Diploma's degree under the guidance of Professor Kleinert. The result was a new algorithm for calculating highly accurate critical exponents from divergent perturbation expansions of field theories which I would like to summarize on this occasion. Typically, we possess L expansion coefficients of such a divergent series in powers of the bare coupling constant g_B , plus two more informations: The knowledge of the large-order behavior proportional to $(-\alpha)^k k! k^\beta g_B^k$, with a known growth parameter α , and the knowledge of the approach to scaling being of the type $c + c'/g_B^\omega$, with constants c, c' and a critical exponent of approach ω . The latter information leads to an increase in the speed of convergence and a high accuracy of the results. The algorithm is applied to the six- and seven-loop expansions for the critical exponents of $O(N)$ -symmetric ϕ^4 -theories in three dimensions, and the result for the critical exponent α is compared with a recent satellite experiment.

1 Introduction

The field-theoretic approach to critical phenomena provides us with power series expansions for the critical exponents of a wide variety of universality classes.

When inserted into the renormalization group equations, these expansions are supposed to determine the critical exponents via their values at an infrared-stable fixed point $g = g^*$. The latter step is nontrivial since the expansions are divergent and require a resummation, for which sophisticated

methods have been developed [1]. The resummation methods use the information from the known large-order behavior $(-\alpha)^k k! k^\beta g_B^k$ of the expansions and analytic mapping techniques to obtain quite accurate results.

A completely different resummation procedure was developed on the basis of variational perturbation theory [2] to the expansions in powers of the *bare coupling constant*, which goes to infinity at the critical point.

This method converges as fast as the previous ones, even though it makes only use of the fact that the power series for the critical exponents approach their constant critical value in the form $c + c'/g_B^\omega$, where c, c' are constants, and ω is the critical exponent of the approach to scaling.

We may therefore expect that a resummation method which incorporates both informations should lead to results with an even higher accuracy, and it is our purpose to present such a method in the form of a simple algorithm [3].

2 The Problem

Mathematically, the problem we want to solve is the following: Let

$$f_L(g_B) = \sum_{k=0}^L f_k g_B^k \quad (1)$$

be the first L terms of a divergent asymptotic expansion

$$f(g_B) = \sum_{k=0}^{\infty} f_k g_B^k \quad (2)$$

with the large-order behavior of the expansion coefficients

$$f_k \stackrel{k \rightarrow \infty}{\sim} \gamma k! (-\alpha)^k k^\beta [1 + \mathcal{O}(1/k)]. \quad (3)$$

Suppose furthermore that $f(g_B)$, possesses a strong-coupling expansion of the type

$$f(g_B) = g_B^s \sum_{k=0}^{\infty} b_k g_B^{-k\omega}, \quad (4)$$

which is assumed to have some finite convergence radius $|g_B| \geq g_B^{\text{conv}}$.

We are interested in an efficient method to determine the strong-coupling coefficients b_k from the known coefficients f_k of the asymptotic expansion.

3 Hyper-Borel Transformation

It turns out that ordinary Borel resummation is not well suited to solve this problem. Therefore we constructed an algorithm which is based on a generalization to be called *hyper-Borel transformation* [4] defined by

$$\tilde{B}(y) = \sum_{k=0}^{\infty} \tilde{B}_k y^k, \quad (5)$$

with coefficients

$$\tilde{B}_k \equiv \omega \frac{\Gamma(k(1/\omega - 1) + \beta_0)}{\Gamma(k/\omega - s/\omega) \Gamma(\beta_0)} f_k. \quad (6)$$

3.1 General Properties

The inverse of this transformation is given by the double integral

$$f(g_B) = \frac{\Gamma(\beta_0)}{2\pi i} \oint_C dt e^t t^{-\beta_0} \int_0^\infty \frac{dy}{y} \left[\frac{g_B}{yt^{(1-\omega)/\omega}} \right]^s \exp \left[\frac{yt^{(1-\omega)/\omega}}{g_B} \right]^\omega \tilde{B}(y), \quad (7)$$

as can easily be shown with the help of the integral representation of the inverse Gamma function

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i} \int_C dt e^t t^{-z}. \quad (8)$$

The transformation possesses a free parameter β_0 which can be used to optimize the approximation $f_L(g_B)$ at each order L . The power ω of the strong-coupling expansion is assumed to lie in the interval $0 < \omega < 1$, as it does in the upcoming physical applications.

The hyper-Borel transformation has the desired property of allowing for a resummation of $f_L(g_B)$ with the full sequence of powers of g_B in the strong-coupling expansion (4).

Our transform $\tilde{B}(y)$ shares with the ordinary Borel transform the property of being analytic at the origin and its radius of convergence is determined by the singularity on the negative real axis at

$$y_s = -\frac{1}{\sigma} \equiv -\frac{1}{\alpha} \frac{1}{\omega(1-\omega)^{1/\omega-1}}. \quad (9)$$

3.2 Resummation Procedure

A resummation procedure can now be set up on the basis of the transform $\tilde{B}(y)$. The inverse transformation (7) contains an integral over the entire positive axis, requiring an analytic continuation of the Taylor expansion of $\tilde{B}(y)$ beyond the convergence radius.

The reason for introducing the transform $\tilde{B}(y)$ was to allow us to reproduce the complete power sequence in the strong-coupling expansion (4), with a leading power g_B^s and a subleading sequence of powers $g_B^{s-k\omega}$, $k = 1, 2, 3, \dots$. This is achieved by removing a factor $e^{-\rho\sigma y}$ with $\rho, \sigma > 0$ from the truncated series (5) of our transform $\tilde{B}(y)$. Furthermore, by removing a second simple prefactor of the form $(1 + \sigma y)^{-\delta}$, we weaken the leading singularity in the hyper-Borel complex y -plane, which determines the large order behavior (3). The remaining series has still a finite radius of convergence. To achieve convergence on the entire positive y -axis for which we must do the integral (7), we reexpand the remaining series of y in powers of the conformal mapping $\kappa(y)$ given by

$$\kappa(y) = \frac{\sigma y}{1 + \sigma y}, \quad (10)$$

which maps a shifted right half of the complex y -plane with $\Re[y] \geq -1/2\sigma$ onto the unit circle in the complex κ -plane. Thus we reexpand $\tilde{B}(y)$ in the following way:

$$\tilde{B}(y) \equiv \sum_{k=0}^{\infty} \tilde{B}_k y^k = e^{-\rho\sigma y} [1 + \sigma y]^{-\delta} \sum_{k=0}^{\infty} h_k \kappa^k(y) = e^{-\rho\sigma y} \sum_{k=0}^{\infty} h_k \frac{(\sigma y)^k}{(1 + \sigma y)^{k+\delta}}. \quad (11)$$

The inverse hyper-Borel transform of $\tilde{B}(y)$ is now found by forming the integrals of the expansion functions in (11)

$$I_n(g_B) = \frac{\Gamma(\beta_0)}{2\pi i} \oint_C dt e^t t^{-\beta_0} \times \int_0^\infty \frac{dy}{y} \left[\frac{g_B}{y t^{1/\omega-1}} \right]^s \exp \left[-\frac{y t^{1/\omega-1}}{g_B} \right]^\omega e^{-\rho\sigma y} \frac{(\sigma y)^n}{(1 + \sigma y)^{n+\delta}}, \quad (12)$$

so that the approximants $f_L^a(g_B)$ may be written as

$$f_L^a(g_B) = \sum_{n=0}^L h_n I_n(g_B). \quad (13)$$

The convergent strong-coupling expansion of $I_n(g_B)$ is obtained by performing a Taylor series expansion of the exponential function in (12), which is an expansion in powers of $1/g_B^\omega$. After integrating over t and y we obtain an expansion

$$I_n(g_B) = g_B^s \sum_{k=0}^{\infty} b_k^{(n)} g_B^{-k\omega}, \quad (14)$$

which has indeed the same power sequence as the strong-coupling expansion (4) of the function $f(g_B)$ to be resummed and the expansion coefficients are given by

$$b_k^{(n)} = \frac{(-1)^k}{k!} \frac{\sigma^{s-k\omega} \Gamma(\beta_0)}{\Gamma[(\omega-1)k + \beta_0 + (1/\omega-1)s]} i_k^{(n)}, \quad (15)$$

where $i_k^{(n)}$ denotes the integral

$$i_k^{(n)} = \int_0^\infty dy e^{-\rho y} (1+y)^{-\delta-n} y^{k\omega+n-s-1}. \quad (16)$$

For large k , the integral on the right-hand side of (16) can be estimated with the help of the saddle-point approximation, which shows that the strong-coupling expansion (4) has a finite convergence radius

$$|g_B| \geq \frac{(\rho\sigma)^\omega}{(1-\omega)^{1-\omega}}, \quad (17)$$

implying that the basis functions $I_n(g_B)$, and certainly also $f(g_B)$ itself, possess additional singularities beside $g_B = 0$. The parameter ρ will be optimally adjusted to match the positions of these singularities.

3.3 Convergence Properties of Resummed Series

We shall now discuss the speed of convergence of the resummation procedure. For this it will be sufficient to estimate the convergence of the strong-coupling coefficients b_k^L of the approximations $f_L(g_B)$ against the true b_k in (4). The convergence for arbitrary values of g_B will always be better than that. Such an estimate is possible by looking at the large- n behavior of the expansion coefficients $b_k^{(n)}$ in the strong-coupling expansion of $I_n(g_B)$ in (14). This is determined by the saddle-point approximation to the integral $i_k^{(n)}$ in Eq. (16),

which we rewrite as

$$i_k^{(n)} = \int_0^\infty dy e^{-\rho y - n \ln(1+1/y)} (1+y)^{-\delta} y^{k\omega-s-1}. \quad (18)$$

The saddle point lies at

$$y_s = \sqrt{\frac{n}{\rho}} [1 + \mathcal{O}(1/\sqrt{n})]. \quad (19)$$

At this point, the total exponent in the integrand is

$$-\rho y_s - n \ln \left(1 + \frac{1}{y_s}\right) = -2\sqrt{\rho n} [1 + \mathcal{O}(1/\sqrt{n})], \quad (20)$$

implying the large- n behavior

$$b_k^{(n)} \stackrel{n \rightarrow \infty}{\sim} \text{const.} \times n^{k\omega-s-1-\delta} e^{-2\sqrt{\rho n}} [1 + \mathcal{O}(1/\sqrt{n})]. \quad (21)$$

The coefficients b_k^L of the approximations $f_L^a(g_B)$ are linear combinations of the coefficients $b_k^{(n)}$ of the basis functions $I_n(g_B)$:

$$b_k^L = \sum_{n=0}^L b_k^{(n)} h_n. \quad (22)$$

The speed of convergence with which the b_k^L 's approach b_k as the number L goes to infinity is governed by the growth with n of the reexpansion coefficients h_n and of the coefficients $b_k^{(n)}$ in Eq. (21). We shall see that for the series to be resummed, the reexpansion coefficients h_n will grow at most like some power n^r , implying that the approximations b_k^L approach their $L \rightarrow \infty$ -limit b_k with an error proportional to

$$b_k^L - b_k \sim L^{r+k\omega-s-\delta-1/2} \times e^{-2\sqrt{\rho L}}. \quad (23)$$

This is the important advantage of the present resummation method with respect to variational perturbation theory [2,5], where the error decreases merely like $e^{-\text{const} \times L^{1-\omega}}$ with $1-\omega$ close to $1/4$.

4 Resummation of Ground-State Energy of the Anharmonic Oscillator

Before beginning with the resummation of the perturbation expansions for the critical exponents of ϕ^4 -field theories, it will be useful to obtain a feeling for the quality of the above-developed resummation procedure, in particular

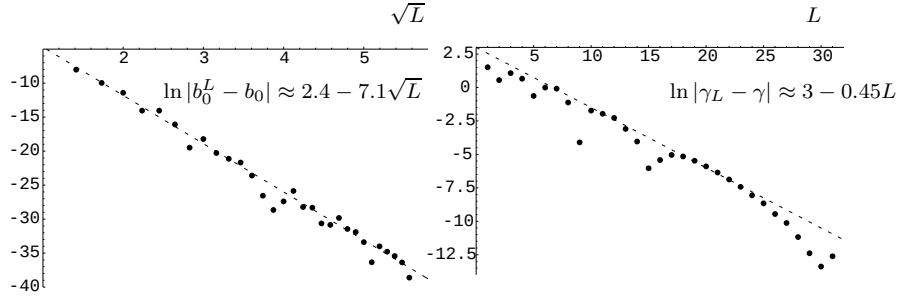


Figure 1. Logarithmic plot of the convergence behavior of the successive approximations to the prefactor γ in the large-order behavior (3), and of the leading strong-coupling coefficient b_0^L .

for the significance of the parameters upon the speed of convergence. We do this by resumming the often-used example of an asymptotic series, the perturbation expansion of the ground-state energy of the anharmonic oscillator with Hamiltonian

$$H = \frac{p^2}{2} + m^2 \frac{x^2}{2} + gx^4. \quad (24)$$

The ground state of this quantum mechanical system has an asymptotic expansion with a large-order behavior of the form (3), where the growth parameters are given by $\alpha = 3, \beta = -1/2, \gamma = \sqrt{6/\pi^3}$, and it also possesses a strong-coupling expansion (4) with the parameters $s = 1/3, \omega = 2/3$. We fix the parameters in our resummation method by the condition that the approximants $E_0^L(g)$ of the ground-state energy $E_0(g)$ obey Eqs. (3) and (4) with the same parameters as $E_0(g)$. The best choice of β_0 will be made differently depending on the regions of g .

Let us test the convergence of our algorithm at small negative coupling constants g , i.e. near the tip of the left-hand cut in the complex g -plane. We do this by calculating the prefactor γ in the large-order behavior (3). In this case the convergence turns out to be fastest by giving the parameter β_0 a small value, i.e. $\beta_0 = 2$.

The values of the approximants γ_L are shown in Fig. 1. They converge exponentially fast against the exact limiting value. The convergence of the strong-coupling coefficients b_k^L is given by the stretched exponential $\approx e^{-\text{const} \times \sqrt{L}}$ [see Eq. (23)], rather than $\approx e^{-\text{const} \times L^{1/3}}$ for variational perturbation theory. The latter is seen on the right-hand side of Fig. 1.

Table 1. Strong-coupling coefficients b_n of the 70-th order approximants $E_{70}^0(g) = \sum_{n=0}^{70} h_n I_n(g)$ to the ground-state energy $E^0(g)$ of the anharmonic oscillator. They have the same accuracy as the variational perturbative calculations up to order 251 in Refs. [2,6].

n	b_n
0	0.667 986 259 155 777 108 270 962 02
1	0.143 668 783 380 864 910 020 319
2	-0.008 627 565 680 802 279 127 963
3	0.000 818 208 905 756 349 542 41
4	-0.000 082 429 217 130 077 219 91
5	0.000 008 069 494 235 040 964 75
6	-0.000 000 727 977 005 945 772 63
7	0.000 000 056 145 997 222 351 17
8	-0.000 000 002 949 562 732 709 36
9	-0.000 000 000 064 215 331 956 97
10	0.000 000 000 048 214 263 789 07

We have applied our resummation method to the first 10 strong-coupling coefficients using the expansion coefficients f_k up to order 70. The results are shown in Table 1. Comparison with a similar table in Refs. [2,6] shows that the new resummation method yields in 70th order the same accuracy as variational perturbation theory did in 251st order.

5 Resummation of Critical Exponents

Having convinced ourselves of the fast convergence of our new resummation method, let us now turn to the perturbation expansions of the $O(N)$ -symmetric ϕ^4 -theories in powers of the bare coupling constant λ_B in $D = 3$ dimensions. If we introduce the dimensionless bare coupling constant $g_B \equiv \lambda_B/m$, where m is the renormalized mass, the critical exponents are defined by

$$\begin{aligned} \eta &= g_B \frac{d}{dg_B} \log Z_\Phi \Big|_{g_B=\infty}, \\ 2 - \nu^{-1} &= g_B \frac{d}{dg_B} \log \frac{m_0^2}{m^2} \Big|_{g_B=\infty}. \end{aligned} \quad (25)$$

Table 2. Critical exponents of the $O(N)$ -symmetric ϕ^4 -theory from our new resummation method.

n	γ	η	ν	ω
0	1.1604[8] (4){0.075}	0.0285[6](4){0.037}	0.5881[8](4){0.075}	0.803[3]{1}
1	1.2403[8] (4){0.110}	0.0335[6](3){0.043}	0.6303[8](4){0.065}	0.792[3]{1}
2	1.3164[8] (5){0.033}	0.0349[8](5){0.042}	0.6704[7](4){0.098}	0.784[3]{1}
3	1.3882[10](7){0.210}	0.0350[8](5){0.043}	0.7062[7](4){0.110}	0.783[3]{1}

The expansions of the field renormalization constant Z_ϕ and the bare mass m_0 have been calculated up to seventh order in g_B in the literature [7]. When approaching the critical point, the renormalized mass m tends to zero, so that the problem is to find the strong-coupling limit $g_B \rightarrow \infty$ of these expansions. In order to have the critical exponents approach a constant value, the power s in Eq. (4) must be set equal to zero.

In contrast to the quantum mechanical discussion in the last section, the exponent ω governing the approach to the scaling limit is now unknown, and must also be determined from the available perturbation expansions. As in Ref. [5,8], we solve this problem by using the fact that the existence of a critical point implies the renormalized coupling constant g in powers of g_B to converge against a constant renormalized coupling g^* for $m \rightarrow 0$.

The convergence against a fixed coupling g^* occurs only for the correct value of ω in the resummation functions $I_n(g_B, \omega)$. At different values, $g(g_B)$ has some strong-coupling power behavior g_B^s with $s \neq 0$. We may therefore determine ω by forming a series for the power s ,

$$s = \frac{d \log g(g_B)}{d \log g_B} = \frac{g_B}{g} g'(g_B), \quad (26)$$

resumming this for various values of ω in the basis functions, and finding the critical exponent ω from the zero of s . Alternatively, since $g(g_B) \rightarrow g^*$, we can just as well resum the series for $-gs$, which coincides with the β -function of renormalization group theory [not to be confused with the growth parameter β in (3)]

$$\beta(g_B) \equiv -g_B \frac{dg(g_B)}{dg_B}. \quad (27)$$

The results for the critical exponents of all $O(N)$ -symmetries are shown in Table 2. The total error is indicated in the square brackets. It is deduced

from the error of resummation of the critical exponent at a fixed value of ω indicated in the parentheses, and from the error $\Delta\omega$ of ω , using the derivative of the exponent with respect to ω given in curly brackets. Symbolically, the relation between these errors is

$$[\dots] = (\dots) + \Delta\omega\{\dots\}. \quad (28)$$

The accuracy of our results can be judged by comparison with the most accurately measured critical exponent α parameterizing the divergence of the specific heat of superfluid helium at the λ -transition by $|T_c - T|^{-\alpha}$. By going into a vicinity of the critical temperature with $\Delta T \approx 10^{-8}$ K, a recent satellite experiment has provided us with the value [9] $\alpha = -0.01056 \pm 0.00038$. Our value for α is deduced from ν in Table 2 via the hyper-scaling relation $\alpha = 2 - 3\nu$ to $\alpha = -0.0112 \pm 0.0021$, in good agreement with the experimental number.

References

- [1] R. Guida and J. Zinn-Justin, *J. Phys. A* **31**, 8130 (1998). See also the textbook: H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [2] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [3] F. Jasch and H. Kleinert, *J. Math. Phys.* **42**, 52 (2001), eprint: cond-mat/9906246.
- [4] This transformation has never been investigated in the literature, although it is contained in a class of quite general mathematical transformations introduced in the textbook of G.H. Hardy, *Divergent Series* (Oxford University Press, Oxford, 1949) in the context of *moment constant methods*. These comprise transformations $B(y) = \sum f_k y^k / \mu_k$, where the μ_k are given by a Stieltjes integral $\mu_k = \int_0^\infty x^k d\chi(x)$ and χ is a bounded and increasing function of x guaranteeing the convergence of the Stieltjes integral. This definition includes our transformation for the somewhat complicated choice $d\chi(x) = (\Gamma(\beta)/2\pi i)x^{-s-1} \oint_C dt e^{t+x\omega t^{1-\omega}} t^{s(1-1/\omega)-\beta_0} dx$.
- [5] H. Kleinert, *Phys. Rev. D* **57**, 2264 (1998); Addendum: *ibid.* **58**, 107702 (1998).
- [6] W. Janke and H. Kleinert, *Phys. Rev. Lett.* **75**, 2787 (1995).
- [7] D.B. Murray and B.G. Nickel, University of Guelph, preprint (1998).

- [8] H. Kleinert, *Phys. Rev. D* **60**, 85001 (1999).
- [9] J.A. Lipa, D.R. Swanson, J.A. Nissen, T.C.P. Chui, and U.E. Israelsson, *Phys. Rev. Lett.* **76**, 944 (1996); J.A. Lipa, D.R. Swanson, J.A. Nissen, Z.K. Geng, P.R. Williamson, D.A. Stricker, T.C.P. Chui, U.E. Israelsson, and M. Larson, *Phys. Rev. Lett.* **84**, 4894 (2000).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

CRITICAL EXPONENT α OF SUPERFLUID HELIUM FROM VARIATIONAL STRONG-COUPPLING THEORY

B. VAN DEN BOSSCHE

*Physique Nucléaire Théorique, B5, Université de Liège, 4000 Liège, Belgium,
and
Institut für Theoretische Physik, Freie Universität Berlin, Arnimallee 14, D-14195
Berlin, Germany
E-mail: bossche@physik.fu-berlin.de*

The variational strong-coupling theory of Hagen Kleinert is used to determine the critical exponent α of superfluid helium. It is shown that applying the theory to $\exp(\alpha)$, a highly accurate value of α is obtained.

1 Introduction

The variational approach of Feynman and Kleinert [1] has been systematically improved by Kleinert in Ref. [2]. It has been extended to field theory for the determination of critical exponents in $D = 3$ dimensions in Ref. [3] and in $4 - \epsilon$ dimensions in Refs. [4,5]. For a review, see Ref. [6]. Recently, it has been shown that the theory is applicable for the determination of amplitude functions and ratios [7].

Having learned variational perturbation theory directly from its inventor, it is a pleasure to dedicate a contribution on this subject to him. More precisely, I shall focus on the theoretical determination of the critical exponent α of superfluid helium, and show that a negative α is obtained, already at the two-loop level. Our choice of studying this comes from the fact it is probably the best well-known measured quantity. It was obtained in a zero-gravity experiment by Lipa *et al.* [8], who parameterized the specific heat as follows

(we use the second of the references quoted in Ref. [8]):

$$C^\pm = \frac{A^\pm}{\alpha} |t|^{-\alpha} (1 + D|t|^\Delta + E|t|^{2\Delta}) + B, \quad t = T/T_c - 1, \quad (1)$$

$$\alpha = -0.01056 \pm 0.0004, \quad (2)$$

with $\Delta = 0.5$, $A^+/A^- = 1.0442 \pm 0.001$, $A^-/\alpha = -525.03$, $D = -0.00687$, $E = 0.2152$ and $B = 538.55$ (J/mol K). Apart from B , this parameterization is an approximation to the Wegner expansion form

$$F = F_\pm |t|^\chi \left(1 + a_{0,1}|t|^{\Delta_0} + a_{0,2}|t|^{2\Delta_0} + a_{0,3}|t|^{3\Delta_0} + \dots \right. \\ \left. + a_{1,1}|t|^{\Delta_1} + a_{1,2}|t|^{2\Delta_1} + a_{1,3}|t|^{3\Delta_1} + \dots \right) \quad (3)$$

with χ a combination of critical exponents and $F_\pm$ denoting the leading amplitude above and below T_c , respectively. Compared to this general Wegner expansion, higher powers in $\Delta \equiv \Delta_0$ have been neglected in (1), as well as daughter powers $\Delta_i, i \geq 1$. This will be also the case in the present work, where we shall take into account only one exponent Δ , related to the more well-known critical exponent ω by the relation $\Delta = \omega\nu$.

2 Model and Algorithm

The critical behavior of many different physical systems can be described by an $O(N)$ -symmetric ϕ^4 -theory. In particular, the case $N = 0$ describes polymers, $N = 1$ the Ising transition (a universality class which comprises binary fluids, liquid-vapor transitions and antiferromagnets), $N = 2$ the superfluid transition in helium, $N = 3$ isotropic magnets (transition of the Heisenberg type), and $N = 4$ phase transition of Higgs fields at finite temperature. The field energy is given by the Ginzburg-Landau functional

$$\mathcal{H} = \int d^D x \left[\frac{1}{2} (\nabla \phi_B)^2 + \frac{1}{2} m_B^2 \phi_B^2 + \frac{g_B}{4!} (\phi_B^2)^2 \right]. \quad (4)$$

As mentioned above, we shall restrict our analysis to the case of superfluid helium, for which the field ϕ_B has $N = 2$ components. The subscript B stands for bare. We shall work in the minimal subtraction (MS) scheme in $4 - \epsilon$ dimensions with ϵ -expansion. For this reason, the square of the bare mass m_B^2 goes to zero at the transition (the critical value of the bare mass is identically zero), linearly with the temperature: $m_B^2 \equiv t = (T/T_c - 1)$, hence the name "reduced temperature" for t .

To save space, we shall not reproduce the whole algorithm of variational perturbation theory reviewed in Refs. [6,7]. We sketch below only the main points. Let us start with a function f whose expansion in terms of the bare coupling constant is known up to a given order L :

$$f \approx f_L(\bar{g}_B) = \sum_{i=0}^L f_i \bar{g}_B^i, \quad (5)$$

with $\bar{g}_B$ a reduced coupling constant to be defined later. Then, provided we know that the function f goes to a constant as the bare coupling constant goes to infinity, variational perturbation theory indicates that the value of the function f at the critical point is given by

$$f_L(\bar{g}_B \rightarrow \infty) = \text{opt}_{\hat{g}_B} \left[\sum_{i=0}^L f_i \hat{g}_B^i \sum_{j=0}^{L-i} \binom{-i\epsilon/\omega}{j} (-1)^j \right]. \quad (6)$$

This simple formula replaces the well-known but rather involved resummation procedure for the renormalized power series. In particular, for critical exponents, the ϵ -expansion of Eq. (6) reproduces the ϵ -expansion obtained using the renormalized theory (before resummation).

The operator $\text{opt}_{\hat{g}_B}$ is referred to as "optimization", and has a particular meaning: The even-loop orders are optimized by extrema (they may be minima or maxima depending on the critical exponents to be investigated), the odd-loop orders by turning-points, obtained by equating the second derivative of (6) to zero.

We are now ready to determine the exponent α for superfluid helium. We proceed in two steps. The first step (Section 3) is dedicated to two- and three-loop orders, where we can calculate the strong-coupling limit $\lim_{g_B \rightarrow \infty} \alpha(g_B)$ analytically. The second step is given in Section 4, where we shall present numerical results up to five loops.

The starting point of our analysis are the five-loop calculations of Refs. [9,10]. Working within the MS scheme means that only the poles in ϵ are removed. This allows us to identify the renormalization constants Z_{m^2} , Z_g , and Z_ϕ which relate the bare and renormalized mass, coupling constant, and field, respectively:

$$m_B^2 = m^2 \frac{Z_{m^2}}{Z_\phi}, \quad (7)$$

$$g_B = \mu^\epsilon \frac{Z_g}{Z_\phi^2} g, \quad (8)$$

$$\phi_B = Z_\phi^{1/2} \phi. \quad (9)$$

In the MS scheme, the renormalization constants depend only on g , or, upon inverting (8), on $\bar{g}_B \equiv g_B/|m|$. This statement comes from the identification of the scale μ with the renormalized mass $|m|$ and setting $D = 3$, or $\epsilon = 1$. At the critical point, the renormalized mass goes to zero: The critical theory corresponds to the strong-coupling limit of the bare theory.

From these considerations and the relevant expansions for the $O(2)$ -symmetric theory taken from Ref. [6], we obtain

$$\begin{aligned} g = \bar{g}_B - \frac{10}{3} \bar{g}_B^2 + \frac{130}{9} \bar{g}_B^3 - \frac{[6017 + 384\zeta(3)]}{81} \bar{g}_B^4 \\ + \frac{[420505 + 78432\zeta(3) - 5760\zeta(4) + 36480\zeta(5)]}{972} \bar{g}_B^5 \\ - [26929681 + 9514768\zeta(3) + 92928\zeta(3)^2 - 1260960\zeta(4) \\ + 8001280\zeta(5) - 912000\zeta(6) + 3386880\zeta(7)] \frac{\bar{g}_B^6}{9722}, \end{aligned} \quad (10)$$

$$\begin{aligned} \alpha = \frac{1}{2} - \bar{g}_B + \frac{7}{2} \bar{g}_B^2 - \frac{677}{36} \bar{g}_B^3 + \frac{[81913 + 4272\zeta(3) + 2304\zeta(4)]}{648} \bar{g}_B^4 \\ - [311381 + 46896\zeta(3) - 3520\zeta(3)^2 + 18492\zeta(4) \\ + 12480\zeta(5) + 15200\zeta(6)] \frac{\bar{g}_B^5}{324}, \end{aligned} \quad (11)$$

where we have redefined g and $\bar{g}_B$ so as to absorb a factor $1/(4\pi)^2$. To save space, we have also written these expressions only for $\epsilon = 1$. We have however checked that, by keeping ϵ everywhere, the usual ϵ -expansion of the critical exponents would have been recovered. Taking the exponential of (11), and reexpanding up to the fifth order, we have

$$\begin{aligned} \frac{\exp(\alpha)}{\sqrt{e}} = 1 - \bar{g}_B + 4\bar{g}_B^2 - \frac{809}{36} \bar{g}_B^3 + \frac{[99229 + 4272\zeta(3) + 2304\zeta(4)]}{648} \bar{g}_B^4 \\ - [3788857 + 490320\zeta(3) - 35200\zeta(3)^2 + 196440\zeta(4) \\ + 124800\zeta(5) + 152000\zeta(6)] \frac{\bar{g}_B^5}{3240}. \end{aligned} \quad (12)$$

The aim of this article is to show that the strong-coupling limit of the critical exponent α from the series (12) yields a much better result than the original series (11).

3 Analytical Evaluation of α up to Three Loops

Following the general procedure described in Refs. [3–7], the critical exponent ω of the approach to scaling can be extracted from (10) by considering its logarithmic derivative. The result has been obtained analytically at the two-loop level in Ref. [4] and at the three-loop level in Ref. [7]. Specializing to the case of superfluid helium, the solutions read, with $\omega = 1/(\rho - 1)$, and where the subscript indicates the loop order,

$$\rho_2 = 4\sqrt{\frac{2}{5}}, \quad (13)$$

$$\rho_3 = \frac{50311 + 1152\zeta(3)}{2[12907 - 3456\zeta(3)]} - 3 \frac{\sqrt{3[1039 + 128\zeta(3)][62779 + 2688\zeta(3)]}}{[12907 - 3456\zeta(3)]} \times \cos \left[-\frac{2}{3}\pi + \frac{1}{3}\text{arcsec}(\theta) \right], \quad (14)$$

with

$$\theta = \frac{400[62779 + 2688\zeta(3)]^{3/2}}{\sqrt{3117 + 384\zeta(3)}\{55818649 + 768\zeta(3)[-118163 + 15552\zeta(3)]\}}. \quad (15)$$

Numerically, we have $\rho_2 \approx 2.52982$ and $\rho_3 \approx 2.38683$, to which correspond $\omega_2 \approx 0.65367$ and $\omega_3 \approx 0.721069$.

Using (6) to two loops, together with (13), the strong-coupling limit of (11) and (12) is

$$\alpha_2 = \frac{3}{70} \approx 0.0428571, \quad (16)$$

$$\alpha_2 = \frac{1}{2} - \ln\left(\frac{5}{3}\right) \approx -0.0108256, \quad (17)$$

respectively, from which it is clear that the strong-coupling limit of the direct series (11) fails to give a negative value of α at the two-loop level, while the strong-coupling limit of $\exp(\alpha)$ leads to such a negative value. Moreover, this value is extremely close to the experimental result (2).

The three-loop order is also determined analytically and confirms the two-loop conclusion: Using (6) to three loops, together with (14) and (15), the strong-coupling limit of (11) and (12) is

$$\alpha_3 = \frac{450097 + 77826\rho_3 - 127218\rho_3^2 + 8988\rho_3^3}{916658} \approx 0.0363442, \quad (18)$$

Table 1. Critical exponent ω of the approach to scaling for different loop orders L .

L	2	3	4	5
ρ_L	2.52982	2.38683	2.36773	2.32803
ω_L	0.653671	0.721069	0.731141	0.752997

Table 2. Critical exponent α for different loop orders L .

L	2	3	4	5
α_L from Eq. (11)	0.0428571	0.0363441	0.0214116	0.0189176
α_L from Eq. (12)	-0.0108256	-0.0116665	-0.0110531	-0.0116324

$$\alpha_3 = \frac{1}{2} + \ln \left(\frac{648337 + 56280\rho_3 - 93144\rho_3^2 + 10320\rho_3^3}{654481} \right) \approx -0.0116664, \quad (19)$$

respectively. The second result lies again very close to the experimental value (2), whereas the direct evaluation would have led to a positive value. The excellent agreement of the strong-coupling limit of (12) is confirmed by the numerical five-loop calculation of the next section.

4 Numerical Evaluation of α to Five Loops

It is not possible to evaluate analytically the strong-coupling limit of a given function above the three-loop order since the defining equation for ρ , obtained taking the logarithmic derivative of (10) and applying the algorithm (6), is higher than cubic. Nothing can, however, prevent us from calculating the algorithm numerically. We have done it for the four- and five-loop orders. To facilitate the comparison with the previous section, we recall also below the numerical value of the second- and third-loop order.

With $N = 2$ and $\epsilon = 1$, we obtain for ρ_L and ω_L the values shown in Table 1. Applying (6) to (11) and (12), together with the numerical values of Table 1, we obtain for the critical exponent α results which are summarized in Table 2. We see that the strong-coupling limit of α , as given by the direct series (11), decreases, as expected, but is still positive at the five-loop order. A

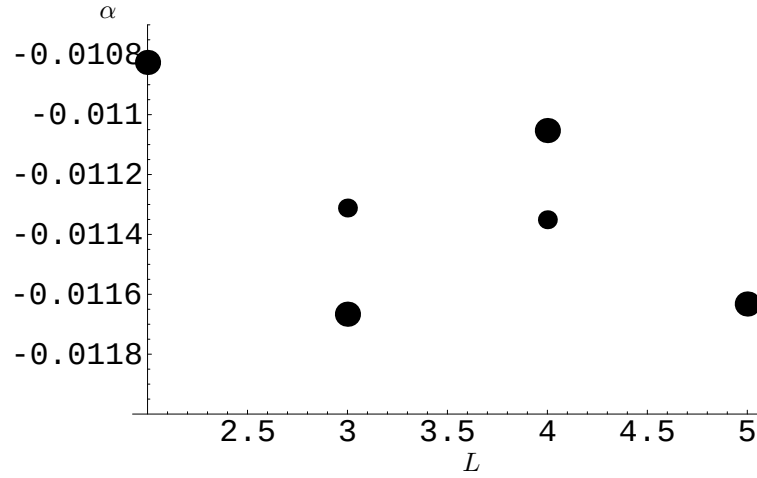


Figure 1. Critical exponent α of superfluid helium as a function of the loop order L . Thick dots denote the exponent based on the strong-coupling limit of Eq. (12) while thin dots are Shanks-improved results [12].

nontrivial extrapolation procedure, such as the one employed in Refs. [3,5,6], is needed.

The calculation of α , based on (12), gives negative results, not too far from the experimental result. Moreover, the results seem to alternate around a given value, see Fig. 1. An ordinary procedure for accelerating the convergence seems to be applicable. In our case, it is tempting to perform a Shanks transformation [12]. Denoting by α^* the improved value, we can construct

$$\alpha_3^* = \frac{\alpha_2\alpha_4 - \alpha_3^2}{\alpha_2 + \alpha_4 - 2\alpha_3} = -0.011312, \quad (20)$$

$$\alpha_4^* = \frac{\alpha_3\alpha_5 - \alpha_4^2}{\alpha_3 + \alpha_5 - 2\alpha_4} = -0.011351. \quad (21)$$

In Figure 1, we have plotted the values given in the last row of Table 2 as well as the improved values of Eqs. (20) and (21).

It would be extremely interesting to obtain the six-loop order of α . Then a third Shanks-improved point α_5^* could be obtained, and an iterated Shanks transformation could be performed, allowing to obtain $\alpha^{**} = [\alpha_3^*\alpha_5^* - (\alpha_4^*)^2]/(\alpha_3^* + \alpha_5^* - 2\alpha_4^*)$. This would probably lead to an extremely precise value of α . Due to the large number of Feynman diagrams to be evaluated, this task is, however, not manageable at present using available

techniques, and calls for new ideas.

Another interesting work would be to check if the Borel resummation of the ϵ -expansion of $\exp(\alpha)$ would lead to a similar improvement of the convergence as variational perturbation theory.

5 Conclusion

In this contribution, we have applied variational perturbation theory to the determination of the critical exponent α of superfluid helium from five-loop ϵ -expansions. Previous studies were based on Borel resummation [11], or on variational perturbation theory together with a suitable extrapolation procedure [3,5,6]. The approach followed here was based on a full self-consistent calculation, using the same loop order for the critical exponent of the approach to scaling ω as for the exponent α itself. For the direct resummation of α , this has proven to be of less accuracy than the extrapolation approach [5,6]. However, we have shown here that an appropriate choice of the function of α , in this work we have chosen $\exp[\alpha(\bar{g}_B)]$, to be evaluated in the strong-coupling limit $\bar{g}_B$ is able to drastically improve the convergence. Our work does not solve the question of which function of a critical exponent gives the fastest convergence of strong-coupling theory, nor have we checked whether the exponential function can be used to obtain the other exponents more accurately than before. The present note may, however, be seen as a first step of a program towards optimizing the functions of the critical exponents to be evaluated.

Acknowledgments

I thank the Alexander von Humboldt Foundation and the F.N.R.S. (Belgium) for support. I am indebted to Professor Kleinert for numerous clarifications and discussions.

References

- [1] R.P. Feynman and H. Kleinert, *Phys. Rev. A* **34**, 5080 (1986).
- [2] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995); variational perturbation theory is developed in Chapters 5 and 17.

- [3] H. Kleinert, *Phys. Rev. D* **57**, 2264 (1998); Addendum *ibid.* **58**, 107702 (1998); *ibid.* **60**, 085001 (1999).
- [4] H. Kleinert, *Phys. Lett. B* **434**, 74 (1998); **463**, 69 (1999).
- [5] H. Kleinert and V. Schulte-Frohlinde, *J. Phys. A* **34**, 1037 (2001), eprint: cond-mat/9907214.
- [6] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001) [<http://www.physik.fu-berlin.de/~kleinert/b8>].
- [7] H. Kleinert and B. Van den Bossche, eprint: cond-mat/0011329.
- [8] J.A. Lipa, D.R. Swanson, J.A. Nissen, T.C.P. Chui, and U.E. Israelsson, *Phys. Rev. Lett.* **76**, 944 (1996); J.A. Lipa, D.R. Swanson, J.A. Nissen, Z.K. Geng, P.R. Williamson, D.A. Stricker, T.C.P. Chui, U.E. Israelsson, and M. Larson, *Phys. Rev. Lett.* **84**, 4894 (2000). In the second paper, the value of α given in the first paper was corrected to $\alpha = -0.01056 \pm 0.0004$.
- [9] K.G. Chetyrkin, S.G. Gorishny, S.A. Larin, and F.V. Tkachov, *Phys. Lett. B* **132**, 351 (1983).
- [10] H. Kleinert, J. Neu, V. Schulte-Frohlinde, K.G. Chetyrkin, and S.A. Larin, *Phys. Lett. B* **272**, 39 (1991); Erratum: *Phys. Lett. B* **319**, 545(E) (1993).
- [11] R. Guida and J. Zinn-Justin, *J. Phys. A* **31**, 8103 (1998), eprint: cond-mat/9803240 (1998). See also J. Zinn-Justin, eprint: hep-th/0002136.
- [12] C. Brezinski and M. Redivo-Zaglia, *Extrapolation Methods, Theory, and Practice*, in *Studies in Computational Mathematics*, Eds. C. Brezinski and L. Wuytack (North-Holland, Amsterdam, 1991).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

FIVE-LOOP EXPANSION OF THE ϕ^4 -THEORY AND CRITICAL EXPONENTS FROM STRONG-COUPLING THEORY

V. SCHULTE-FROHLINDE

*Lyman Laboratory, Department of Physics, Harvard University,
12 Oxford Street Cambridge, MA 02138, USA*

E-mail: frohlinde@cmts.harvard.edu

The full analytic reevaluation of all diagrams up to five loops of the $O(N)$ -symmetric ϕ^4 -theory led to the correction of the ε -expansions of the β -function and the anomalous dimensions. These expansions were ideal testing grounds for various resummation techniques. Especially Hagen Kleinert's strong-coupling approach led to a reformulation of the renormalization group theory in terms of the bare parameters.

1 Introduction

The scalar quantum field theory with ϕ^4 -interaction correctly describes many experimentally observable features of critical phenomena. Field theoretic renormalization group techniques [1] in $D = 4 - \varepsilon$ dimensions [2–4] combined with Borel resummation methods of the resulting ε -expansions [5] lead to extremely accurate determinations of the critical exponents of all $O(N)$ universality classes. The renormalization group (RG) functions of the ϕ^4 -theory were first calculated analytically close to four dimensions using dimensional regularization [6] and the minimal subtraction (MS) scheme [7] in three- and four-loop approximations [8,9]. This calculation was extended to the five-loop level [10–12] after the ingenious invention of special reduction algorithms for the integrals [13,14]. The critical exponents were obtained as ε -expansions [3] up to the order ε^5 . These expansions have to be evaluated for $\varepsilon = 1$ in order to obtain results in three dimensions.

When the analytic five-loop calculation of the β -function and the anomalous dimensions was completed in 1983/1984, Hagen Kleinert had the idea to use the new algorithms to automatize the calculation of Feynman diagrams and their ε -expansions. In 1989, this idea then became a thesis project for my colleague Joachim Neu and me. Our first step in this rather lengthy project was an independent recalculation of the five-loop perturbation series using the same techniques [10,13]. Unfortunately, we could not reproduce the results for some of the diagrams. Hagen Kleinert sent us to Moscow to discuss our results, a trip which led to the discovery of errors in six of the 135 diagrams and to our first publication [15]. In the subsequent years, the perturbation expansions for the critical exponents were used to study old and new resummation methods leading among other results [16] to Kleinert's strong-coupling approach to the renormalization group [17,18].

Here, we will summarize the five-loop calculations [15] and then present the strong-coupling approach to resum the ε -expansion of the critical exponents [19]. Details can be found in our textbook [20].

2 Five-Loop Expansion of the ϕ^4 -Theory

We consider the $O(N)$ -symmetric theory of N -dimensional real scalar fields ϕ_B with the Lagrangian

$$L_B(x) = \frac{1}{2} [\partial\phi_B(x)]^2 + \frac{1}{2} m_B^2 \phi_B^2(x) + (4\pi)^2 \frac{\lambda_B}{4!} [\phi_B^2(x)]^2, \quad (1)$$

in Euclidean space with $D = 4 - \varepsilon$ dimensions. The bare (unrenormalized) coupling constant λ_B and mass m_B are expressed via renormalized ones as

$$\lambda_B = \mu^\varepsilon Z_g g = \mu^\varepsilon \frac{Z_4}{(Z_2)^2} g, \quad m_B^2 = Z_{m^2} m^2 = \frac{Z_{\phi^2}}{Z_2} m^2. \quad (2)$$

Here μ is the unit of mass in dimensional regularization and Z_4 , Z_2 , Z_{m^2} , Z_g are the renormalization constants of the vertex function, propagator, mass, and coupling constant, respectively, with Z_{ϕ^2} being the renormalization constant of the two-point function obtained from the propagator by the insertion of the vertex $(-\phi^2)$ in all possible ways [9]. In the MS-scheme the renormalization constants do not depend on dimensional parameters and are expressible as series in $1/\varepsilon$ with purely g -dependent coefficients:

$$Z_i = 1 + \sum_{k=1}^{\infty} \frac{Z_{i,k}(g)}{\varepsilon^k}, \quad (3)$$

where $i = g, m^2, 2, 4, \phi^2$. The β -function and the anomalous dimensions entering the RG equations are expressed in the standard way as follows:

$$\beta(g) = \left. \frac{d g}{d \ln \mu} \right|_{\lambda_B} = -\varepsilon g + g \frac{\partial Z_{g,1}}{\partial g}, \quad (4)$$

$$\gamma_m = \left. \frac{d \ln m}{d \ln \mu} \right|_{\lambda_B} = -\frac{d \ln Z_{m^2}}{d \ln \mu^2} = \frac{1}{2} g \frac{\partial Z_{m^2,1}}{\partial g}, \quad (5)$$

$$\gamma_i(g) = \left. \frac{d \ln Z_i}{d \ln \mu^2} \right|_{\lambda_B} = -\frac{1}{2} g \frac{\partial Z_{i,1}}{\partial g}, \quad i = 2, 4, \phi^2. \quad (6)$$

To determine all RG functions up to five loops we calculated the five-loop approximation to the three constants Z_2 , Z_4 , and Z_{ϕ^2} . The constant Z_2 contains the counterterms of the 11 five-loop propagator diagrams. The constant Z_4 receives contributions from 124 vertex diagrams. Of these diagrams, 90 contribute to Z_{ϕ^2} after appropriate changes of combinatorial factors.

We have used the same methods as in the previous works [10,13] to calculate the counterterms from the dimensionally regularized Feynman integrals, namely, the method of infrared rearrangement [21], the Gegenbauer polynomial x -space technique (GPXT) [14], the integration-by-parts algorithm [22], and the R - and R^* -operations [23]. These methods allow to proceed with the calculation of massless integrals with only one external momentum. The renormalization is carried out recursively and for each Feynman diagram separately. The higher-order diagrams are then algebraically reduced to one-loop integrations by the integration-by-parts algorithms. Restrictions of the applicability of these algorithms have so far prevented the complete automatization on a computer.

Some of the diagrams do not follow the general scheme. Three diagrams were calculated analytically first [11] by using the so-called method of uniqueness, later the same results were obtained by using the Gegenbauer polynomials in x -space together with several non-trivial tricks [24]. A detailed description of the calculations including the diagramwise results is presented elsewhere [20].

The analytic results of the five-loop approximations to the RG functions $\beta(g)$, $\gamma_2(g)$ and $\gamma_m(g)$ are expansions in g with N -dependent coefficients. The number ε appears only once in the β -function. These RG functions can now be used to calculate the ε -expansions of the critical exponents which describe

the behavior of a statistical system near the critical point of the second-order phase transition [4]. Close to the critical temperature $T = T_C$, the asymptotic behavior of the correlation function for $|\mathbf{x}| \rightarrow \infty$ has the form

$$\Gamma(\mathbf{x}) \sim \frac{e^{-|\mathbf{x}|/\xi}}{|\mathbf{x}|^{D-2+\eta}} . \quad (7)$$

Close to T_C , the correlation length ξ behaves for $\tau = T - T_C \rightarrow 0$ as

$$\xi \sim \tau^{-\nu} (1 + \text{const} \cdot \tau^{\omega\nu} + \dots) . \quad (8)$$

The three critical exponents η , ν , and ω defined in this way completely specify the critical behavior of the system.

The behavior (7) and (8) is found for the ϕ^4 -theory if $\mu \rightarrow 0$ as $T \rightarrow T_C$. In this limit the coupling constant g approaches the so-called infrared-stable fixed point which is determined by the condition

$$\beta(g^*) = 0, \quad \beta'(g^*) = [\partial\beta(g)/\partial g]_{g=g^*} > 0 . \quad (9)$$

The fixed point g^* is determined as an expansion in ε :

$$g^* = \sum_{k=1}^{\infty} g^{(k)} \varepsilon^k . \quad (10)$$

Approaching the fixed point, the renormalized mass goes to zero such that $\xi = 1/m$ behaves like (8). The resulting formulas for the critical exponents are

$$\eta = 2\gamma_2(g^*) , \quad 1/\nu = 2[1 - \gamma_m(g^*)] , \quad \omega = \beta'(g^*) , \quad (11)$$

each emerging as an ε -expansion up to order ε^5 [15,20].

It is known that the ε -expansions are asymptotic series, and special resummation techniques [5,25] should be applied to obtain reliable estimates of the critical exponents. One such technique will be described now.

3 Strong-Coupling Theory

In 1998, Hagen Kleinert has developed a new approach [26,27] to critical exponents of field theories based on the strong-coupling limit of variational perturbation expansions [28,29]. This limit is relevant for critical phenomena if the renormalization constants are expressed in terms of the unrenormalized

coupling constant since the infrared-stable fixed point is approached for infinite g_B : $g(g_B) \rightarrow g^*$ for $g_B \rightarrow \infty$. This idea has been applied successfully to $O(N)$ -symmetric ϕ^4 -theories in three and $4 - \varepsilon$ dimensions [17–19], yielding the three fundamental critical exponents ν, η, ω with high accuracy.

From model studies of perturbation expansions of the anharmonic oscillator it is known that variational perturbation expansions possess good strong-coupling limits [30,31], with a speed of convergence governed by the convergence radius of the strong-coupling expansion [32,33]. This has enabled Hagen Kleinert to set up an algorithm [29] for deriving uniformly convergent approximations to functions of which one knows a few initial Taylor coefficients and an important scaling property: the functions approach a constant value with a given inverse power of the variable. The renormalized coupling constant g and the critical exponents of a ϕ^4 -theory have precisely this property as a function of the bare coupling constant g_B . In $D = 4 - \varepsilon$ dimensions the approach is parameterized as follows [26]

$$g(g_B) = g^* - \frac{\text{const}}{g_B^{\omega/\varepsilon}} + \dots, \quad (12)$$

where g^* is the infrared-stable fixed point, and ω is called the critical exponent of the approach to scaling [compare Eqs. (8) and (11)]. This exponent is universal, governing the approach to scaling of every function $F(g)$,

$$f(g_B) = F(g(g_B)) = F(g^*) + F'(g^*) \times \frac{\text{const}}{g_B} \equiv f^* + \frac{\text{const}'}{g_B^{\omega/\varepsilon}}. \quad (13)$$

Strong-coupling theory is designed to calculate f^* and ω . Let $f(g_B)$ be a function with this behavior and suppose that we know its first $L+1$ expansion terms,

$$f_L(g_B) = \sum_{l=0}^L a_l g_B^l. \quad (14)$$

More specifically than in Eq. (12), we assume that $f(g_B)$ approaches its constant strong-coupling limit f^* in the form of an inverse power series

$$f_M(g_B) = \sum_{m=0}^M b_m (g_B^{-2/q})^m, \quad (15)$$

with a finite radius of convergence [34]. Then the L th approximation to the

value f^* is obtained from the strong-coupling formula [17,26,27]

$$f_L^* = \text{opt}_{\hat{g}_B} \left[\sum_{l=0}^L a_l v_l \hat{g}_B^l \right]. \quad (16)$$

The quantities

$$v_l \equiv \sum_{k=0}^{L-l} \binom{-ql/2}{k} (-1)^k \quad (17)$$

are simply binomial expansions of $(1-1)^{-ql/2}$ up to the order $L-l$. The expression in brackets in Eq. (16) has to be optimized in the variational parameter $\hat{g}_B$. The optimum is the smoothest among all real extrema. If there are no real extrema, the turning points serve the same purpose.

3.1 Application to Renormalization Constants and Critical Exponents

Going back to Eqs. (1) and (2) we now set the scale parameter μ equal to the physical mass m and consider all quantities as functions of $g_B = \lambda_B/m^\varepsilon$. Now, instead of μ , we let m_B^2 go to zero like $\tau = \text{const} \times (T - T_c)$ as the temperature T approaches the critical temperature T_c , and assume that also m^2 goes to zero, and thus g_B to infinity. The latter assumption turns out to be self-consistent. Assuming the theory to scale as suggested by experiments, we now determine the value of the renormalized coupling constant g in the strong-coupling limit $g_B \rightarrow \infty$, and also of the exponent ω , assuming the behavior (12). First we apply formula (16) to the logarithmic derivative $s(g_B)$ of the function $g(g_B)$:

$$s(g_B) \equiv g_B g'(g_B)/g(g_B). \quad (18)$$

Setting $s_L^* = 0$ determines the approximation ω_L to ω .

The other critical exponents are found as follows. If we assume that the ratios m^2/m_B^2 and ϕ^2/ϕ_B^2 have a limiting power-law behavior for small m

$$\frac{m^2}{m_B^2} \propto g_B^{-\eta_m/\varepsilon} \propto m^{\eta_m}, \quad \frac{\phi^2}{\phi_B^2} \propto g_B^{\eta/\varepsilon} \propto m^{-\eta}, \quad (19)$$

the powers η_m and η can be calculated from the strong-coupling limits of the logarithmic derivatives

$$\eta_m(g_B) = -\varepsilon \frac{d}{d \log g_B} \log \frac{m^2}{m_B^2}, \quad \eta(g_B) = \varepsilon \frac{d}{d \log g_B} \log \frac{\phi^2}{\phi_B^2}. \quad (20)$$

When approaching the second-order phase transition, where the bare mass m_B^2 vanishes like $\tau \equiv (T - T_c)$, the physical mass m^2 vanishes with a different power of τ . This power is obtained from the first equation in (19), which shows that $m \propto \tau^{1/(2-\eta_m)}$. In experiments one observes that the correlation length of fluctuations $\xi = 1/m$ increases near T_c like $\tau^{-\nu}$. A comparison with the previous equation shows that the critical exponent ν is equal to $1/(2-\eta_m)$. Similarly we see from the second equation in (19) that the scaling dimension $D/2 - 1$ of the free field ϕ_B for $T \rightarrow T_c$ is changed in the strong-coupling limit to $D/2 - 1 + \eta/2$, the number η being the anomalous dimension of the field. This implies a change in the large-distance behavior of the correlation functions $\langle \phi(x)\phi(0) \rangle$ at T_c from the free-field behavior r^{-D+2} to $r^{-D+2-\eta}$. The results from the renormalization group are recovered from assumption (19). Comparison with Eq. (11) shows that $\eta_m = 2\gamma_m$, whereas η is the same as above.

Let us mention that this procedure leads to resummed expressions which have the same ε -expansions as those found by renormalization group techniques.

3.2 Five-Loop Results

In a first step, we determine the parameter ω such that the logarithmic derivative of $g(g_B)$ approaches zero for $g_B \rightarrow \infty$. We therefore insert the coefficients of the power series of $s(g_B)$ from Eq. (18) into Eq. (16) and determine $q = 2/\omega$ for $L = 2, 3, 4, 5$, such that $s_L^* = 0$. The resulting ε -expansion for the approach-to-scaling parameter ω reproduces the well-known ε -expansion [15] up to the corresponding order. In Fig. 1a), the approximations ω_L are plotted against the number of loops L for $\varepsilon = 1$ and $N = 3$. Apparently, the five-loop results are still some distance away from a constant $L \rightarrow \infty$ -limit. The slow approach to the limit calls for a suitable extrapolation method. The convergence behavior in the limit $L \rightarrow \infty$ was determined [26] to be of the general form

$$f^*(L) \approx f^* + \text{const} \times e^{-cL^{1-\omega}}. \quad (21)$$

We plot the approximations s_L for a given ω near the expected critical exponent against L , and fit the points by the theoretical curve (21) to determine the limit s^* . Then ω is varied, and the plots are repeated until s^* is zero. The resulting ω is the desired critical exponent, and the associated plot is shown in Fig. 1b). Since the optimal variational parameter $\hat{g}_B$ comes from minima and

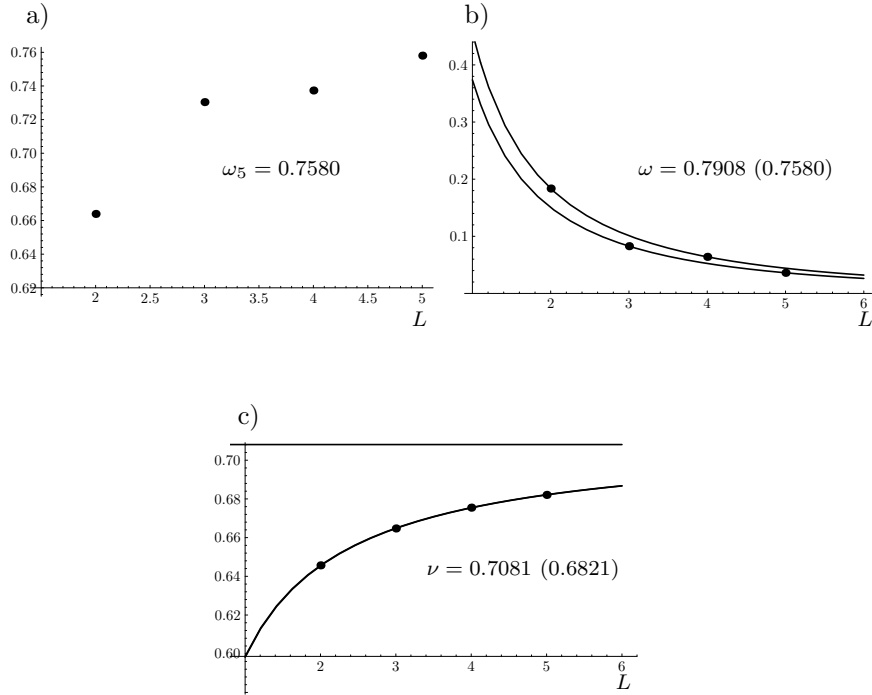


Figure 1. a) Critical exponent ω of approach to scaling calculated from $s_L^* = 0$, plotted against the order of approximation L for $N = 3$. b) Extrapolation of the solutions of the equation $s_L^* = 0$ to $L \rightarrow \infty$ with the help of Eq. (21). The value of ω for which $s_L^* \rightarrow 0$ for $L \rightarrow \infty$ determines $\omega = 2/q$. The best extrapolating function is $s_L = -6.8 \times 10^{-7} + 156.916e^{-5.84 L^{0.2092}}$. c) Determination of the critical exponent ν plotted as a function of L . The extrapolating function is $\nu_L = 0.7081 - 4.0104e^{-3.6012 L^{0.2092}}$, the horizontal line indicates the value of ν_∞ .

turning points for even and odd approximants in alternate order, the points are best fitted by two different curves. The resulting ω -values are listed in Table 1. They are used to derive the strong-coupling limits for the exponents ν , γ and η . For the calculation of the critical exponent ν , we find the five-loop expansion for $\nu(g_B)$ using the relation $\nu(g_B) = 1/[2 - \eta_m(g_B)]$. From this we calculate the strong-coupling approximations ν_L for $L = 2, 3, 4, 5$. After extrapolating these to infinite L , we obtain the numbers listed for different universality classes $O(N)$ in Table 1. The corresponding extrapolation fits are

Table 1. Critical exponents of five-loop strong-coupling theory and comparison with the results from Borel-type resummation (GZ) [33], and from variational perturbation theory in $D = 3$ dimensions [27]. The parentheses behind each number show the five-loop approximation to see the extrapolation distance.

	VPT, $D = 4 - \varepsilon$	Borel-Res. (GZ)	VPT 3D
	$\omega_\infty(\omega_5)$		
$N = 0$	0.80345(0.7448)	0.828 ± 0.023	0.810
$N = 1$	0.7998(0.7485)	0.814 ± 0.018	0.805
$N = 2$	0.7948(0.7530)	0.802 ± 0.018	0.800
$N = 3$	0.7908(0.7580)	0.794 ± 0.018	0.797
	$\nu_\infty(\nu_5)$ (I)		
$N = 0$	0.5874(0.5809)	0.5875 ± 0.0018	0.5883
$N = 1$	0.6292(0.6171)	0.6293 ± 0.0026	0.6305
$N = 2$	0.6697(0.6509)	0.6685 ± 0.0040	0.6710
$N = 3$	0.7081(0.6821)	0.7050 ± 0.0055	0.7075
	$\eta_\infty(\eta_5)$ (I)		
$N = 0$	0.0316(0.0234)	0.0300 ± 0.0060	0.03215
$N = 1$	0.0373(0.0308)	0.0360 ± 0.0060	0.03370
$N = 2$	0.0396(0.0365)	0.0385 ± 0.0065	0.03480
$N = 3$	0.0367(0.0409)	0.0380 ± 0.0060	0.03447
	$\gamma_\infty(\gamma_5)$		
$N = 0$	1.1576(1.1503)	1.1575 ± 0.0050	1.616
$N = 1$	1.2349(1.2194)	1.2360 ± 0.0040	1.241
$N = 2$	1.31045(1.2846)	1.3120 ± 0.0085	1.318
$N = 3$	1.3830(1.3452)	1.3830 ± 0.0135	1.390

plotted in Fig. 1c). Similarly, estimations for the exponents η and γ can be obtained [19]. In Table 1 the resulting values are compared to those found by Borel resummation in $D = 4 - \varepsilon$ dimensions and by the same strong-coupling approach in $D = 3$ dimensions.

4 Conclusion

Instead of expressing the renormalization group functions in the renormalized coupling constant g_B , we can reexpand in the bare coupling constant g . This allows applying strong-coupling theory to the five-loop perturbation expan-

sions of $O(N)$ -symmetric ϕ^4 -theories in $4 - \varepsilon$ dimensions. Satisfactory values for all critical exponents are obtained.

Acknowledgments

We thank the German Academic Exchange Service (DAAD) for support.

References

- [1] N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience, New York, 1980).
- [2] K.G. Wilson, *Phys. Rev. B* **4**, 3184 (1971); K.G. Wilson and J.B. Kogut, *Phys. Rep. C* **12**, 75 (1974).
- [3] K.G. Wilson and M.E. Fisher, *Phys. Rev. Lett.* **28**, 240 (1972).
- [4] E. Brézin, J.C. Le Guillou, and J. Zinn-Justin, in *Phase Transitions and Critical Phenomena*, Vol. 6, Eds. C. Domb and M.S. Green (Academic Press, New York, 1976).
- [5] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 1st ed. (Clarendon Press, Oxford, 1989).
- [6] G. 't Hooft and M. Veltman, *Nucl. Phys. B* **44**, 189 (1972).
- [7] G. 't Hooft, *Nucl. Phys. B* **61**, 455 (1973).
- [8] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, and B.G. Nickel, *Phys. Lett. A* **44**, 227 (1973).
- [9] A.A. Vladimirov, D.I. Kazakov, and O.V. Tarasov, *Sov. Phys. JETP* **50**, 521 (1979); Preprint JINR E2-12249, Dubna, 1979.
- [10] K.G. Chetyrkin, A.L. Kataev, and F.V. Tkachov, *Phys. Lett. B* **99**, 147 (1981); *ibid.* **101**, 457 (1981) (Erratum).
- [11] D.I. Kazakov, *Phys. Lett. B* **133**, 406 (1983); *Teor. Mat. Fiz.* **58**, 343 (1984); *Dubna lecture notes*, E2-84-410 (1984).
- [12] S.G. Gorishny, S.A. Larin, and F.V. Tkachov, *Phys. Lett. A* **101**, 120 (1984).
- [13] K.G. Chetyrkin, S.G. Gorishny, S.A. Larin, and F.V. Tkachov, *Phys. Lett. B* **132**, 351 (1983); Preprint INR P-0453, Moscow, 1986.
- [14] K.G. Chetyrkin and F.V. Tkachov, Preprint INR P-118, Moscow, 1979; K.G. Chetyrkin, A.L. Kataev, and F.V. Tkachov, *Nucl. Phys. B* **174**, 345 (1980).
- [15] H. Kleinert, J. Neu, V. Schulte-Frohlinde, K.G. Chetyrkin, and S.A. Larin, *Phys. Lett. B* **272**, 39 (1991); H. Kleinert and V. Schulte-

- Frohlinde, *Phys. Lett. B* **342**, 284 (1995).
- [16] H. Kleinert, S. Thoms, and V. Schulte-Frohlinde, *Phys. Rev. B* **56**, 14428 (1997).
- [17] H. Kleinert, *Phys. Lett. B* **434**, 74 (1998).
- [18] H. Kleinert, *Phys. Lett. B* **463**, 69 (1999).
- [19] H. Kleinert and V. Schulte-Frohlinde, *J. Phys. A* **34**, 1037 (2001), eprint: cond-mat/9907214.
- [20] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [21] A.A. Vladimirov, *Teor. Mat. Fiz.* **43**, 210 (1980).
- [22] F.V. Tkachov, *Phys. Lett. B* **100**, 65 (1981); K.G. Chetyrkin and F.V. Tkachov, *Nucl. Phys. B* **192**, 159 (1981).
- [23] K.G. Chetyrkin and F.V. Tkachov, *Phys. Lett. B* **114**, 340 (1982); K.G. Chetyrkin and V.A. Smirnov, *Phys. Lett. B* **144**, 410 (1984).
- [24] D.J. Broadhurst, *Massless Scalar Feynman Diagrams: Five Loops and Beyond*, Open University Preprint OUT - 4102 - 18 1985, Milton Keynes, U.K.
- [25] D.I. Kazakov and D.V. Shirkov, *Fortschr. Phys.* **28**, 465 (1980); J. Zinn-Justin, *Phys. Rep.* **70**, 3 (1981).
- [26] H. Kleinert, *Phys. Rev. D* **57**, 2264 (1998); Add.: *ibid.* **58**, 1077 (1998).
- [27] H. Kleinert, *Phys. Rev. D* **60**, 85001 (1999).
- [28] H. Kleinert, *Phys. Lett. A* **173**, 332 (1993).
- [29] Details of strong-coupling theory are found in Chapter 5 of the textbook: H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [30] W. Janke and H. Kleinert, *Phys. Lett. A* **199**, 287 (1995).
- [31] W. Janke and H. Kleinert, *Phys. Rev. Lett.* **75**, 2787 (1995).
- [32] H. Kleinert and W. Janke, *Phys. Lett. A* **206**, 283 (1995).
- [33] R. Guida and J. Zinn-Justin, *J. Phys. A* **31**, 8103 (1998).
- [34] H. Kleinert, *Phys. Lett. A* **207**, 133 (1995).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

Part IV

Phase Transitions and Critical Phenomena

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

NONASYMPTOTIC CRITICAL BEHAVIOR FROM FIELD THEORY

C. BAGNULS

*Service de Physique de l'Etat Condensé,
C. E. Saclay, F91191 Gif-sur-Yvette Cedex, France
E-mail: bagnuls@spec.saclay.cea.fr*

C. BERVILLIER

*Service de Physique Théorique,
C. E. Saclay, F91191 Gif-sur-Yvette Cedex, France
E-mail: bervil@spht.saclay.cea.fr*

We present and discuss the results of calculations up to five- or six-loop orders of nonasymptotic critical behavior, above and below T_c , within the field-theoretical framework.

1 Introduction

When talking about critical behavior, one usually thinks of critical exponents (power laws), and eventually of corrections to scaling, all notions being strictly related to the unprecise definition of an asymptotic critical domain. In fact, criticality may be observed beyond that theoretical domain and this makes it sometimes difficult to compare theory and experiments [1]. For example, it is possible that some systems undergo a retarded crossover [2] from classical to Ising-like critical behaviors. In such cases, the critical domain would be much larger than for, say, pure fluids. Consequently, many correction-to-scaling terms should be introduced. It is then very likely that the series would not converge. For that reason, nonasymptotic theoretical expressions of critical behavior are required to describe such systems.

Apparently it is not widely known that, beyond the estimations of the critical exponents, the renormalization group (RG) theory [3] is also adapted to provide us with nonasymptotic forms of the critical behavior, especially when a crossover phenomenon occurs (the crossover is then characterized by the competition of two fixed points).

We briefly present here the principles of calculations done within the massive field theoretical framework in three dimensions ($d = 3$) [4] and which have yielded accurate nonasymptotic forms of the susceptibility $\chi(\tau)$ and the specific heat $C(\tau)$ for $\tau = (T - T_c)/T_c > 0$ and $\tau < 0$, of the correlation length $\xi(\tau)$ for $\tau > 0$, and of the coexistence curve $M(\tau)$ for $\tau < 0$ [5,6]. The calculations presented here have induced, directly or indirectly, several subsequent publications [7,8]. We hope that this text will encourage further works on nonasymptotic critical behavior. In particular, we think that variational perturbation theory used recently to estimate universal exponents [9] and amplitude ratios [10], could be an advantageous tool.

Let us first specify that "Nonasymptotic critical behavior" means performing a resummation of an infinite series of correction-to-scaling terms which are expected [11] in the asymptotic expression of any singular quantity such as $\xi(\tau)$. Particularly, for $\tau \rightarrow 0^{+,-}$, we have:

$$\xi(\tau) = \xi_0^{+,-} |\tau|^{-\nu} \left[1 + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{\xi^{+,-}}^{(n,m)} |\tau|^{n\Delta_m} \right], \quad (1)$$

in which ν is a critical exponent, $\xi_0^{+,-}$ stands for the leading critical amplitudes in the two phases and the coefficients $a_{\xi^{+,-}}^{(n,m)}$ correspond to the amplitudes of the confluent corrections to scaling controlled by the exponents Δ_m ($m = 1, 2, \dots, \infty$). Those exponents (ν and Δ_m) arise in a linear study [11] of the RG transformation in the vicinity of a fixed point: the solutions of the eigenvalue problem provide us with some positive (say one, λ_0 , for simplicity^a) and infinitely many negative eigenvalues (λ_m for $m = 1, 2, \dots, \infty$). Then we have $\nu = 1/\lambda_0$, $\Delta_m = -\nu\lambda_m$ ($m = 1, 2, \dots, \infty$). The case $m = 1$ corresponds to the first correction-to-scaling term associated with the largest negative eigenvalue and the usual notations are ω (for $-\lambda_1$) and $\Delta = \omega\nu$ (for Δ_1). In a linear study of the RG, discarding the next-to-leading correction

^aThe number of positive eigenvalues depends on the fixed point considered. Since we are interested in a critical point (but not a multi-critical point), there is only one positive eigenvalue and the fixed point referred to in that case is the famous Wilson-Fisher [12] fixed point.

terms in Eq. (1), it usually holds that

$$\xi(\tau) \simeq \xi_0^{+,-} |\tau|^{-\nu} \left[1 + a_{\xi}^{+,-} |\tau|^{\Delta} \right], \quad (2)$$

with the universality of the ratios ξ_0^+/ξ_0^- and a_{ξ}^+/a_{ξ}^- . Eq. (2) is only valid asymptotically, close to the critical point.

Notice that the infinite sum in Eq. (1) does not converge for large values of τ . Thus, to get a useful nonasymptotic expression of the critical part of ξ , we must consider a resummation procedure. It is provided by the RG theory. However the framework we use implies an approximation: "from field theory" means that only one family of correction-to-scaling terms (associated to $m = 1$) is accounted for. Thus, instead of Eq. (1), our effective expression for $\xi(\tau)$ is:

$$\xi_{FT}(\tau) = \xi_0^{+,-} |\tau|^{-\nu} \left[1 + \sum_{n=1}^{\infty} a_{\xi^{+,-}}^{(n)} |\tau|^{n\Delta} \right]. \quad (3)$$

As one could see in comparing our calculations with experimental [13] or Monte Carlo [14] data, the approximation of field theory does not prevent the study from yielding physically useful nonasymptotic critical behaviors. In fact, in the case of Eq. (3), the range $0 < \tau < \infty$ corresponds to an *interpolation between two fixed points* and, consequently, the crossover is described by universal functions [15].

2 Principles of the Calculations

One starts from the "bare" or unrenormalized ϕ^4 -Hamiltonian in d Euclidean dimensions corresponding to the scalar field theory to be renormalized:

$$H = \int d^d x \left[\frac{1}{2} \left\{ (\nabla \phi_0)^2 + r_0 \phi_0^2 \right\} + \frac{g_0}{4!} \phi_0^4 \right], \quad (4)$$

in which ϕ_0 depends on x and may eventually represent a vector with n components. Thus Eq. (4) is supposed to satisfy the $O(n)$ -symmetry. The bare coupling g_0 is dimensionful and "measured" in units of Λ (the ultra-violet cutoff):

$$g_0 = u_0 \Lambda^{\epsilon},$$

in which $\epsilon = 4 - d$ and u_0 is dimensionless.

There exist two kinds of renormalization schemes for the scalar field theory: the massive and the Weinberg scheme [16]. In the massive scheme, the unit of reference is provided by the mass parameter $m = \xi^{-1}$. In this framework the critical theory, corresponding to $m = 0$, is not defined. On the contrary, in the Weinberg scheme, one first defines the critical theory (massless theory) and the unit of length scale is provided by the inverse of some arbitrary momentum-subtraction-point parameter μ . In that renormalization scheme, the "soft-mass" parameter t is introduced via the renormalization of insertions of the ϕ_0^2 -operator within the vertex-functions; when different from zero, t is (linearly) proportional to the reduced temperature scale τ defined above. Though we have done our calculations within the massive framework — because the longest available^b perturbative series [17] have been obtained within the massive scheme directly in $d = 3$ — the presentation of the principles of the calculation is simpler within the Weinberg scheme. In that scheme, the renormalization conditions correspond to the following (re)-definitions:

$$\phi_0 = [Z_3(u)]^{1/2} \phi, \quad (5)$$

$$(\phi_0)^2 = \frac{Z_3(u)}{Z_2(u)} (\phi)^2, \quad (6)$$

$$u_0 \Lambda^\epsilon = \mu^\epsilon u \frac{Z_1(u)}{[Z_3(u)]^2}, \quad (7)$$

$$r_0 = r_{0c} + \frac{Z_2(u)}{Z_3(u)} t, \quad (8)$$

in which u is the renormalized coupling and r_{0c} is defined by:

$$\Gamma_0^{(0,2)}(p; r_0, g_0) \Big|_{p=0, r_0=r_{0c}} = 0, \quad (9)$$

where the subscript 0 refers to the bare theory.

Up to analytical terms which are usually neglected when studying critical phenomena, the quantity $r_0 - r_{0c}$ is proportional to the physical parameter τ :

$$\frac{r_0 - r_{0c}}{\Lambda^2} = \theta \tau + O(\tau^2), \quad (10)$$

where θ is a nonuniversal factor.

^bThe unpublished Guelph report [17] may be obtained via the web site http://www.physik.fu-berlin.de/~kleinert/kleiner_reb8/programs/programs.html.

The renormalized N -point vertex-functions with L -insertions of the $(\phi)^2$ operator are related to their unrenormalized counterparts as follows:

$$\Gamma^{(L,N)}(\{q,p\};u,\mu) = [Z_3(u)]^{N/2} \left[\frac{Z_2(u)}{Z_3(u)} \right]^L \Gamma_0^{(L,N)}(\{q,p\};r_0,g_0).$$

It does not matter for the following that the renormalization functions $Z_i(u)$ are defined by renormalization conditions on the 2-point and 4-point vertex-functions, considered at some subtraction momentum point expressed in terms of the only dimensionful (momentum-like) parameter μ or by a "minimal" subtraction procedure (such as the subtraction of poles located at $\epsilon = 0$).

In field theory, the RG originates from the arbitrariness of the subtraction procedure for a given bare theory. Hence, the renormalized quantities u and t become functions of the renormalization parameter $l = -\ln(\mu/\Lambda)$. Consequently, Eq. (8) must be understood as follows:

$$r_0 = r_{0c} + \frac{Z_2[u(l)]}{Z_3[u(l)]} t(l). \quad (11)$$

Now, by imposing that $t(l)$ remains a fixed quantity (say $t(l) = 1$), one relates the evolution of $u(l)$ to the approach to the critical point of the bare (physical) theory (defined by $r_0 \rightarrow r_{0c}$). Then, with $t(l) = 1$, Eq. (11) shows that, for $r_0 = r_{0c}$, $u(l)$ must take a particular value u^* , so that $Z_2(u^*)/Z_3(u^*)$ vanishes. Of course, u^* is the nontrivial zero of the famous β -function:

$$\beta(u) = - \left. \frac{du(l)}{dl} \right|_{u_0} \quad (12)$$

with $\omega = d\beta(u)/du|_{u=u^*}$ being positive, so that $u(l) \xrightarrow{l \rightarrow \infty} u^*$.

The pure scaling (power law) regime of vertex-functions corresponds to $u(l) = u^*$ and the first correction-to-scaling term [as in Eq. (2)] to a linear correction proportional to $u(l) - u^*$. As $u(l)$ moves further away from u^* , more and more correction terms must be included but then a nonlinear study is required. It is a matter of fact that the domain $0 < u(l) < u^*$ corresponds to the entire domain $\infty > r_0 - r_{0c} > 0$. Therefore, if one re-sums perturbative series in powers of $u(l)$ in the range $]0, u^*[$, one implicitly obtains nonasymptotic critical answers which interpolate^c between a classical critical behavior

^cWe have also performed calculations [18] for $u > u^*$. In this range, the sign of the first correction-to-scaling term is changed and corresponds to the Ising model [19].

(when $u(l)$ is small) and, say, an Ising-like critical behavior (for the $O(1)$ -symmetry) when $u(l)$ approaches u^* . It remains to invert Eq. (11) to express these answers under the forms of functions of $r_0 - r_{0c}$ [or of τ , via Eq. (10)] which is the genuine physical “measure” of the distance to the critical point.

In order to get the best possible accuracy, we have looked at the available calculations up to relatively high orders of perturbation series. There are two kinds of such calculations:

- (1) analytically up to fifth order in the Weinberg scheme with dimensional regularization and minimal subtractions [20].
- (2) numerically up to sixth order for $d = 3$ in the massive scheme [17].

In both cases, only the renormalization functions Z_i are considered, because the theoretical interest is usually focused on the critical exponents, which are obtained via the Z_i 's. For example the series expansion for the critical exponent η is given by:

$$\eta(u) = \beta(u) \frac{d}{du} \ln Z_3(u), \quad (13)$$

once considered at $u = u^*$.

However, we are not simply interested in the critical exponents but in complete functions such as ξ and χ . Now, only in the second case, the renormalization conditions are such that ξ and χ are known in terms of the Z_i 's. This is not true in the first case^d.

We denote the renormalized parameters of the massive scheme by g and m (instead of u and t). Their relations to the bare parameters are similar to those given by Eqs. (5)–(9) except that, in addition to the change $u \rightarrow g$, Eqs. (7) and (8) now read:

$$g_0 = m^\epsilon g \frac{Z_1(g)}{[Z_3(g)]^2}, \quad (14)$$

$$r_0 = \delta m^2 + \frac{m^2}{Z_3(g)}. \quad (15)$$

The mass shift δm^2 is defined by a subtraction condition^e which avoids the

^dThis is why the amplitude functions are known only up to three loop order in this scheme [7].

^eThis eliminates the quadratic ultra-violet divergences occurring at $d = 4$.

explicit consideration of r_{0c} via Eq. (9), namely:

$$\Gamma^{(0,2)}(0; m, g) = m^2. \quad (16)$$

The other subtraction conditions^f which define the Z_i 's read:

$$\begin{aligned} \frac{d}{dp^2} \Gamma^{(0,2)}(p; m, g) \Big|_{p=0} &= 1, \\ \Gamma^{(0,4)}(\{0\}; m, g) &= m^\epsilon g, \\ \Gamma^{(1,2)}(\{0, 0\}; m, g) &= 1, \end{aligned}$$

so that the physical (bare) quantities ξ and χ are given by:

$$\begin{aligned} \xi^{-1}(g) &= m = g_0 \frac{[Z_3(g)]^2}{gZ_1(g)}, \\ \chi^{-1}(g) &= Z_3^{-1} m^2 = g_0^2 \frac{[Z_3(g)]^3}{[gZ_1(g)]^2}. \end{aligned}$$

The re-summations of the perturbative series for those quantities have been done using the technique initiated by Le Guillou and Zinn-Justin [21] after having taken into account the singularities of the Z_i 's at the fixed point g^* . They may be easily treated by writing, e.g. for $Z_3(g)$ which has a singularity at g^* of the form $(g^* - g)^{\eta/\omega}$:

$$Z_3(g) = Z_3(y) \exp \left\{ \int_y^g \frac{\eta(x)}{\beta(x)} dx \right\},$$

in which y is some small value of g , the definitions of $\beta(x)$ and $\eta(x)$ being unchanged in their forms compared to Eqs. (12) and (13). Let us mention that some difficulties could be encountered in the resummation procedure due to nonanalytic confluent singularities [22] in the β -function at g^* , but they have not been numerically observed yet.

Thus, in the homogeneous phase, the physical quantities ξ and χ can be easily estimated as functions of g in the range $]0, g^*[$ from the calculated series [17]. However our aim was to obtain those quantities as functions of τ (i.e. of $r_0 - r_{0c}$). Now, the massive framework uses $m \propto \tau^\nu$ instead of $t \propto \tau$, therefore the linear relation to τ is lost. In order to reintroduce it, we use the

^fThese eliminate the logarithmic ultra-violet divergences occuring at $d = 4$.

fact that for zero external momenta

$$[Z_2(g)]^{-1} = \Gamma_0^{(1,2)}(r_0, g_0) = \left. \frac{\partial}{\partial r_0} \Gamma_0^{(0,2)}(r_0, g_0) \right|_{g_0}.$$

Using Eqs. (14)–(16), we reexpress this in the following form:

$$\frac{d(r_0/g_0^2)}{dg} = Z_2(g) \frac{d}{dg} \left\{ \frac{[Z_3(g)]^3}{[gZ_1(g)]^2} \right\},$$

which, after integration, allows us to (implicitly) define an effective (and nonperturbative) critical value r'_{0c} by referring to the fixed-point value g^* :

$$\tilde{t}(g) \equiv \frac{r_0 - r'_{0c}}{g_0^2} = - \int_g^{g^*} dx Z_2(x) \frac{d}{dx} \left\{ \frac{[Z_3(x)]^3}{[xZ_1(x)]^2} \right\}. \quad (17)$$

The integrand of Eq. (17) may be estimated using the same procedure as before and the integration has been done numerically yielding the numerical evolution of $\tilde{t}(g)$ in the interval $]0, g^*[$. The final results (the functions $\xi(\tilde{t})$ and $\chi(\tilde{t})$) were obtained after a fitting procedure of the implicit form $\xi(g)$, $\chi(g)$ and $\tilde{t}(g)$. This summarizes the calculations done in the homogeneous phase [5] which included also the specific heat $C(\tilde{t})$, whose perturbative series was previously [23] extracted from the Guelph report [17]; the calculations were performed for the symmetries $n = 1, 2$, and 3.

3 Calculations in the Inhomogeneous Phase and the Critical Bare Mass

“... it is more difficult to calculate physical quantities in the ordered phase because the theory is parameterized in terms of the disordered phase correlation length m^{-1} which is singular at T_c . Also the normalization of correlation functions is singular at T_c ” [24].

We have calculated [6] the perturbative series for the free energy directly at $d = 3$ using the numerous already-estimated Feynman integrals of the massive scheme [17] and new kinds of integrals which have been estimated for the occasion.

Because the free energy is generally written in terms of $T - T_c$, we have been led to explicitly consider the delicate question of the critical bare mass. Indeed, it is known that the perturbative series of super-renormalizable massless field theories (such as $\phi_{d<4}^4$) develop infrared singularities which are usu-

ally simply ignored within the ϵ -expansion framework. In 1973, using a dimensional regularization, Symanzik [25] has shown that the critical bare mass — which has the form $r_{0c} = g_0^{2/\epsilon} f(\epsilon)$ in which $f(\epsilon)$ has poles at $\epsilon = 2/k$ ($k = 1, 2, \dots, \infty$) — is in fact an infrared regulator for the theory. However the final result (free of infrared divergences) is no longer perturbative (e.g. logarithms of g_0 appear at $d = 3$). Though the nonperturbative nature of r_{0c} is an important aspect of the RG theory [26], this question may be circumvented when looking at the critical behavior, since T_c is a nonuniversal quantity. Thus its explicit determination is not required, only the difference $T - T_c$ is needed. Hence, provided that Eq. (9) is again satisfied, one may redefine r_{0c} [as done in Eq. (17)]. Consequently, it is allowed [6] to perform a particular mass-shift $r_0 = r'_0 + \delta r_0(\epsilon)$ in such a way that $\delta r_0(\epsilon)$ subtracts the poles occurring at $\epsilon = 1$, and to fix afterwards the critical temperature in terms of r'_0 .

The series for the free energy have then been obtained graph by graph up to five loops according to the following rules:

- (1) Graphs involving only ϕ^4 -vertices which were already estimated [17] with the mass-shift parameter δm^2 [defined by Eqs. (15), (16)] have been re-expressed to account for the mass-shift parameter $\delta r_0(\epsilon)$.
- (2) New Feynman integrals at $d = 3$ have been found with their weights, involving:
 - exclusively ϕ^3 -vertices have been calculated and compared to existing estimates [27],
 - ϕ^3 -vertices mixed with a single ϕ^4 -vertex have been estimated for the first time for the occasion.

Those series for the free energy have been used by Guida and Zinn-Justin [8] to give an accurate estimation of the scaled equation of state. But this kind of consideration does not account for any correction-to-scaling term and the comparison with experiments is not easy. Instead, we are again interested in actually measurable quantities like the susceptibility χ , the specific heat C , and the spontaneous magnetization M in the inhomogeneous phase. We have not considered the correlation length ξ in this phase because the required Feynman integrals were not calculated by Nickel *et al.* [17]. This quantity has been considered afterwards at $d = 3$ but up to 3-loop order only [28].

Because the renormalization procedure is unchanged in going into the broken-symmetry phase, the critical singularities at the fixed point g^* may be

taken into account with the same renormalization functions $Z_i(g)$ as defined previously. There, the relation between the linear measure of the distance to T_c and the unchanged renormalized coupling g is different. Instead of Eq. (17) we obtain:

$$\tilde{t}'(g) = - \int_g^{g^*} dx \left\{ Z_2(x) \frac{d}{dx} \left\{ \frac{[Z_3(x)]^3}{[xZ_1(x)]^2} \right\} [1 - U(x)] \right\},$$

in which $U(g)$ is given [6] as a power series in g .

Obviously, our nonasymptotic study of the critical behavior accounts for all universal properties expected when $\tau \rightarrow 0$. Consequently, as a by-product, estimates of universal combinations of leading critical amplitudes were given for the first time from the five loop order at $d = 3$. The recent careful re-estimations [8,29] of those universal combinations from the same series have mainly reduced the error-bars. We also gave for the first time accurate estimates of some [30] of the universal ratios of the first confluent correction-to-scaling [6].

Acknowledgments

A seminar given by one of us (C. Be.) in the group of Prof. Hagen Kleinert encouraged us to present again our works on nonasymptotic critical behavior. We congratulate Prof. Hagen Kleinert on the occasion of his 60th birthday.

References

- [1] H. Güttinger and D.S. Cannell, *Phys. Rev. A* **24**, 3188 (1981).
- [2] Y. Levin and M.E. Fisher, *Physica A* **225**, 164 (1996); N.V. Brilliantov, C. Bagnuls, and C. Bervillier, *Phys. Lett. A* **245**, 274 (1998); C. Bagnuls and C. Bervillier, *Cond. Matt. Phys.* **3**, 559 (2000).
- [3] K.G. Wilson and J.B. Kogut, *Phys. Rep. C* **12**, 77 (1974).
- [4] G. Parisi, *J. Stat. Phys.* **23**, 49 (1980).
- [5] C. Bagnuls and C. Bervillier, *Phys. Rev. B* **32**, 7209 (1985); *J. Physique Lett.* **45**, L95 (1984).
- [6] C. Bagnuls, C. Bervillier, D.I. Meiron, and B.G. Nickel, *Phys. Rev. B* **35**, 3585 (1987); see also: Addendum-erratum, eprint: hep-th/0006187.
- [7] R. Schloms and V. Dohm, *Nucl. Phys. B* **328**, 639 (1989); *Phys. Rev. B* **42**, 6142 (1990); H.J. Krause, R. Schloms, and V. Dohm, *Z. Phys. B* **79**, 287 (1990); F.J. Halfkann and V. Dohm, *Z. Phys. B* **89**, 79 (1992);

- S.S.C. Burnett, M. Strösser, and V. Dohm, *Nucl. Phys. B* **504**, 665 (1997); *Nucl. Phys. B* **509**, 729(E) (1998); M. Strösser, S.A. Larin, and V. Dohm, *Nucl. Phys. B* **540**, 654 (1999); M. Strösser, M. Mönnigmann, and V. Dohm, *Physica B* **284**, 41 (2000).
- [8] R. Guida and J. Zinn-Justin, *Nucl. Phys. B* **489**, 626 (1997).
- [9] H. Kleinert, *Phys. Lett. B* **434**, 74 (1998); *Phys. Lett. B* **463**, 69 (1999); *Phys. Rev. D* **57**, 2264 (1998); Addendum: *Phys. Rev. D* **58**, 107702 (1998).
- [10] H. Kleinert and B. Van den Bossche, eprint: cond-mat/0011329.
- [11] F.J. Wegner, *Phys. Rev. B* **5**, 4529 (1972); and in *Phase Transitions and Critical Phenomena*, Vol. VI, Eds. C. Domb and M.S. Green (Academic Press, New York, 1976), p. 7.
- [12] K.G. Wilson and M.E. Fisher, *Phys. Rev. Lett.* **28**, 240 (1972).
- [13] C. Bagnuls and C. Bervillier, *Phys. Rev. Lett.* **58**, 435 (1987); C. Bagnuls, C. Bervillier, and Y. Garrabos, *J. Physique Lett.* **45**, L127 (1984); C.W. Garland, G. Nounesis, M.J. Young, and R.J. Birgeneau, *Phys. Rev. E* **47**, 1918 (1993).
- [14] K. Binder and E. Luijten, *Comp. Phys. Comm.* **127**, 126 (2000).
- [15] See also: A. Pelissetto, P. Rossi, and E. Vicari, *Phys. Rev. E* **58**, 7146 (1998); S. Caracciolo, M.S. Causo, A. Pelissetto, P. Rossi, and E. Vicari, *Nucl. Phys. B (Proc. Suppl.)* **73**, 757 (1999); and the recent review: A. Pelissetto and E. Vicari, eprint: cond-mat/0012164; on the crossover phenomenon in fluids and fluid mixtures, see for example: M.A. Anisimov and J.V. Sengers, in *Equations of State for Fluids and Fluid Mixtures*, Eds. J.V. Sengers, R.F. Kayser, C.J. Peters, and H.J. White Jr (Elsevier, Amsterdam, 2000), p. 381; see also the discussion: C. Bagnuls and C. Bervillier, *Phys. Rev. Lett.* **76**, 4094 (1996); M.A. Anisimov, A.A. Povodyrev, V.D. Kulikov, and J.V. Sengers, *ibid.* 4095 (1996).
- [16] S. Weinberg, *Phys. Rev. D* **8**, 3497 (1973).
- [17] B.G. Nickel, D.I. Meiron, and G.A. Baker Jr., Guelph University preprint, unpublished (1977).
- [18] C. Bagnuls and C. Bervillier, *Phys. Lett. A* **195**, 163 (1994).
- [19] C. Bagnuls and C. Bervillier, *Phys. Rev. B* **41**, 402 (1990); A. Liu and M.E. Fisher, *J. Stat. Phys.* **58**, 431 (1990); B.G. Nickel and B. Sharpe, *J. Phys. A* **12**, 1819 (1979).
- [20] K.G. Chetyrkin, A.L. Kataev, and F.V. Tkachov, *Phys. Lett. B* **99**, 147 (1981); *Phys. Lett. B* **101**, 457 (1981); K.G. Chetyrkin, S.G. Gorishny, S.A. Larin, and F.V. Tkachov, *Phys. Lett. B* **132**, 351 (1983); S.G.

- Gorishny, S.A. Larin, and F.V. Tkachov, *Phys. Lett. A* **101**, 120 (1984); H. Kleinert, J. Neu, V. Schulte-Frohlinde, K.G. Chetyrkin, and S.A. Larin, *Phys. Lett. B* **272**, 39 (1991); Erratum: *Phys. Lett. B* **319**, 545 (1993).
- [21] J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev. Lett.* **39**, 95 (1977).
- [22] B.G. Nickel, in *Phase Transitions*, Eds. M. Lévy, J.C. Le Guillou and J. Zinn-Justin (Plenum Press, New York and London, 1982), p. 291; C. Bagnuls and C. Bervillier, *J. Phys. Stud.* **1**, 366 (1997); A. Pelissetto, and E. Vicari, *Nucl. Phys. B (Proc. Suppl.)* **73**, 775 (1999); M. Caselle, A. Pelissetto, and E. Vicari, eprint: hep-th/0010228.
- [23] C. Bervillier and C. Godrèche, *Phys. Rev. B* **21**, 5427 (1980).
- [24] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 1st ed. (Oxford University Press, 1989), p. 614.
- [25] K. Symanzik, *Lett. Nuovo Cim.* **8**, 771 (1973); see also: G. Parisi, *Nucl. Phys. B* **150**, 163 (1979); R. Jackiw and S. Templeton, *Phys. Rev. D* **23**, 2291 (1981); M.C. Bergère and F. David, *Ann. Phys. (N.Y.)* **142**, 416 (1982).
- [26] C. Bagnuls and C. Bervillier, *Phys. Rev. Lett.* **60**, 1464 (1988).
- [27] J. Reeve, A.J. Guttmann, and B. Keck, *Phys. Rev. B* **26**, 3923 (1982); J. Reeve, *J. Phys. A* **15**, L521 (1982).
- [28] C. Gutfeld, J. Küster, and G. Münster, *Nucl. Phys. B* **479**, 654 (1996).
- [29] R. Guida and J. Zinn-Justin, *J. Phys. A* **31**, 8103 (1998).
- [30] See also: C. Bagnuls and C. Bervillier, *Phys. Rev. B* **24**, 1226 (1981).

NONANALYTICITY OF THE BETA-FUNCTION AND SYSTEMATIC ERRORS IN FIELD-THEORETIC CALCULATIONS OF CRITICAL QUANTITIES

M. CASELLE

*Dip. di Fisica dell'Università di Torino and INFN, Sez. Torino,
I-10125 Torino, Italy
E-mail: caselle@to.infn.it*

A. PELISSETTO

*Dip. di Fisica dell'Università di Roma "La Sapienza" and INFN, Sez. Roma 1,
I-00185 Roma, Italy
E-mail: andrea.pelissetto@roma1.infn.it*

E. VICARI

*Dip. di Fisica dell'Università di Pisa and INFN, Sez. Pisa,
I-56100 Pisa, Italy
E-mail: vicari@df.unipi.it*

By considering the fixed-dimension perturbative expansion, we discuss the nonanalyticity of the renormalization group functions at the fixed point and its consequences for the numerical determination of critical quantities.

1 Introduction

In the last thirty years there has been a significant progress in the understanding of critical phenomena. It has been realized that the behavior in the neighborhood of a critical phase transition, i.e. a transition characterized by long-range correlations, is determined by very few properties: the space dimensionality, the range of the interactions, the number of components of the order parameter, and the symmetry of the Hamiltonian. This means

that physically different systems may have the same critical behavior. For instance, a simple fluid at the liquid-vapor transition and a uniaxial magnet at the Curie point behave identically: critical exponents, dimensionless amplitude ratios, and scaling functions are numerically equal. This phenomenon, which is referred to as universality, has been understood within the Wilson's renormalization group (RG) approach. The conceptual setting is thus quite well established, and the theory of critical phenomena has reached the maturity of well-verified theories like, for instance, QED or the standard model of weak interactions. Nonetheless, it is important to improve experiments and theoretical calculations in order to understand the limits of validity of these theories. In order to test QED and the standard model several experiments have provided accurate estimates that can be directly compared with the theoretical predictions. The most classical ones are the experiments on the g -factor of electrons and muons, and on the Lamb shift in hydrogen. In the theory of critical phenomena, the superfluid transition in ^4He plays a very special role, since it is essentially the only case in which one can determine a critical exponent with an accuracy of 10^{-4} . This is due to a combination of extremely favorable conditions: the singularity in the compressibility of the fluid is particularly weak; it is possible to prepare very pure samples; experiments may be performed in a microgravity environment (for instance on the Space Shuttle), thereby reducing the gravity-induced broadening of the transition. A recent experiment [1,2] obtained an extremely accurate estimate^a of the critical exponent of the specific heat, $\alpha = -0.01056(38)$. This result should be compared with the most precise theoretical estimates: the analysis of high-temperature (HT) expansions gives [3,4] $-0.0146(8)$, $-0.0150(17)$; Monte Carlo (MC) simulations give [3,5] $-0.0148(15)$, $-0.0169(33)$; the analysis of the $d = 3$ perturbative expansion gives $-0.0112(21)$ (variational perturbation theory [6–8]) and $-0.011(4)$ (Borel resummation [9]). There is a clear discrepancy between the most accurate theoretical estimates and the experimental result. However, in order to understand whether the difference is truly significant, we must ask the question: Are the quoted errors reliable? Our experience, looking backward in time, is that there is a natural tendency to be overconfident in one's own results, and thus to systematically underestimate the errors: As Hagen Kleinert [8] put it, each one has a tendency to

^aThe original result reported in Ref. [1] was incorrect. The new estimate is reported in Ref. [2]. The error which is reported in the text is a private communication of J. Lipa to the authors of Ref. [3].

apply the “principle of maximal optimism”. Clearly, further theoretical and experimental investigation is needed to settle the problem.

In order to set correct error bars, it is necessary, although clearly not sufficient, to have a good understanding of the possible sources of systematic error. In MC and HT works, most of the systematic error is due to the nonanalytic corrections to scaling. Indeed, in N -vector systems, there are corrections t^Δ (t is the reduced temperature) to the leading scaling behavior, with $\Delta \approx 0.5$ – 0.6 in the physically relevant cases $0 \leq N \leq 4$. In the analysis of the MC data and of the HT series, these slowly decaying corrections require careful extrapolations, in the absence of which precise but incorrect results are obtained. To give an example, we report here some recent results for the four-point renormalized coupling g^* in the three-dimensional Ising model (see the discussion in Sec. 5 of Ref. [10] and Fig. 1 reported there):

MC, no nonanalytic corrections [11]:	$g^* = 25.0(5);$
MC, with nonanalytic corrections [10]:	$g^* = 23.7(2);$
HT, no nonanalytic corrections [12]:	$g^* = 24.5(2);$
HT, with nonanalytic corrections [13,10]:	$g^* = 23.69(10), 23.55(15).$

For comparison, perturbative field theory gives $g^* = 23.64(7)$ [9], while a recent analysis of improved HT expansions gives $g^* = 23.49(4)$ [14]. Clearly, neglecting the corrections to scaling introduces a large systematic error. And, even worse, there is no way to evaluate it, unless one assumes that nonanalytic corrections are really there.

A solution to these problems is represented by the improved models [3–5,14–18] which are such that the leading scaling correction (approximately) vanishes. The systematic errors are now sensibly reduced and one obtains more reliable estimates.

MC and HT analyses, although different in practice, are very similar in spirit, and indeed they are affected by the same type of systematic errors. In order to assess the reliability of the results, it is thus important to have a different approach to compare with. Field theory provides it and indeed independent estimates can be obtained by using a variety of different methods: the ϵ -expansion pioneered by Wilson and Fisher, the fixed-dimension expansion proposed by Parisi, the perturbative expansion in the minimal-subtraction scheme without ϵ -expansion proposed by Dohm, and the so-called exact RG (there are many different versions, see, e.g. Ref. [19]), which essentially consists in *approximately* solving nonperturbatively the RG equations.

Here we will focus on the fixed-dimension expansion method, which is the one providing the most precise estimates, and, together with the ϵ -expansion, has been most widely used. We consider the standard ϕ^4 -theory with N -vector fields and discuss the role of the singularities of the RG functions at the critical point in the numerical determination of critical quantities.

2 Singularities of the RG Functions

An important controversial issue [20–24] in the field-theoretic (FT) approach in fixed dimension is the presence of nonanalyticities at the fixed point g^* , which is defined as the zero of the β -function. The question was clarified long ago by Nickel [21] who gave a simple argument to show that nonanalytic terms should in principle be present in the β -function. The same argument applies also to other series, like those defining the critical exponents: any RG function is expected to be singular at the fixed point.

To understand the problem, let us consider the four-point renormalized coupling g as a function of the temperature T . For $T \rightarrow T_c$ we can write down an expansion of the form

$$g = g^* \left[1 + a_1 t + a_2 t^2 + \dots + b_1 t^\Delta + b_2 t^{2\Delta} + \dots + c_1 t^{\Delta+1} + \dots + d_1 t^{\Delta_2} + \dots + e_1 t^\gamma + \dots \right], \quad (1)$$

where $\Delta, \Delta_2, \dots$ are subleading exponents and t is the reduced temperature $t \equiv (T - T_c)/T_c$. The corrections proportional to t^γ are due to the presence of an analytic background in the free energy. On general grounds, we expect that $a_1 = a_2 = a_3 = \dots = 0$. Indeed, these analytic corrections arise from the nonlinearity of the scaling fields, and their effect can be eliminated in the Green's functions by an appropriate change of variables [25]. For dimensionless RG-invariant quantities such as g , the leading term is universal and therefore independent of the scaling fields, so that no analytic term can be generated. Analytic correction factors to the singular correction terms are generally present, and therefore the constants c_i in Eq. (1) are expected to be nonzero.

Starting from Eq. (1) it is easy to compute the β -function. Since the mass gap m scales analogously, we obtain for $\Delta < \gamma$ (this condition is usually, but

not always, satisfied^b) the following expansion:

$$\beta(g) = m \frac{\partial g}{\partial m} = \alpha_1 \Delta g + \alpha_2 (\Delta g)^2 + \dots + \beta_1 (\Delta g)^{\frac{1}{\Delta}} + \beta_2 (\Delta g)^{\frac{2}{\Delta}} + \dots + \gamma_1 (\Delta g)^{1+\frac{1}{\Delta}} + \dots + \delta_1 (\Delta g)^{\frac{\Delta-2}{\Delta}} + \dots + \zeta_1 (\Delta g)^{\frac{\Delta}{\Delta}} + \dots, \quad (2)$$

where $\Delta g = g^* - g$. It is easy to verify the well-known fact that $\alpha_1 = -\Delta/\nu \equiv -\omega$ and that, if $a_1 = a_2 = \dots = 0$ in Eq. (1), then $\beta_1 = \beta_2 = \dots = 0$. Equation (2) clearly shows the presence of several nonanalytic terms with exponents depending on $1/\Delta$, Δ_i/Δ , and γ/Δ .

As pointed out by Alan Sokal [22,23], the nonanalyticity of the RG functions can also be understood within Wilson's RG approach. We repeat here his argument. Consider the Gaussian fixed point which, for $3 \leq d < 4$, has a two-dimensional unstable manifold $\mathcal{M}_u$: the two unstable directions correspond to the interactions ϕ^2 and ϕ^4 . Notice that continuum field theories are in a one-to-one correspondence with Hamiltonians on $\mathcal{M}_u$ and that the FT RG is nothing but Wilson's RG restricted to $\mathcal{M}_u$. But now $\mathcal{M}_u$ has no special status at the nontrivial fixed point. In particular, there is no reason why it should approach it exactly along the leading irrelevant direction. Barring miracles, the approach should be along a generic direction, which is tangent to the leading irrelevant one, but which has nonzero components along any of the irrelevant directions. If this happens, nonanalytic terms are present in any RG function.

In order to clarify the issue, Ref. [10] determined the asymptotic behavior of $\beta(g)$ for $g \rightarrow g^*$ in the continuum field theory for $N \rightarrow \infty$ and $2 < d < 4$, and showed that Eq. (2) holds and that the expected nonanalytic terms are indeed present. In Ref. [26] the computation was extended to two dimensions, finding again nonanalytic terms.

The presence of nonanalyticities gives rise to systematic deviations in FT estimates, as it did in MC and HT studies. In the next section we will present a test-case and we will discuss the type of deviations one should expect.

3 A Simple Example

In order to understand the role of the nonanalytic terms in the analyses of the FT perturbative expansions we consider a simple zero-dimensional example.

^bIn some models, for instance in the 2D Ising model, $\Delta > \gamma$. In this case, Eq. (2) is still correct [26] if γ and Δ are interchanged and $\alpha_1 = -\gamma/\nu$.

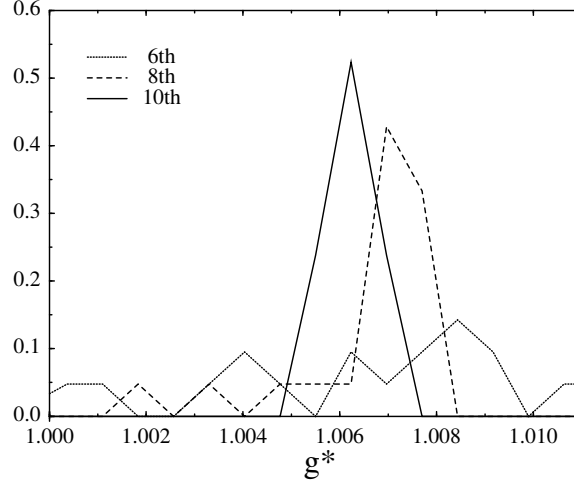


Figure 1. Distribution of the results for the resummations of g^* for $p = 1/10$, $c = -Z_1/5$.

Define

$$f(g; c, p) = c(1 - g)^{1+p} + \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} x^4 \exp \left[-\frac{x^2}{2} - \frac{gx^4}{24} \right]. \quad (3)$$

For $c \neq 0$ and p not integer, this function has a branch point for $g = 1$ and thus it should mimic the behavior we expect for FT expansions. For $g \rightarrow 1$, we have

$$f(g; c, p) \approx Z_0 - Z_1(1 - g) + c(1 - g)^{1+p} + O((1 - g)^2), \quad (4)$$

where $Z_0 = 1.37556014$, $Z_1 = -0.679325$. We wish to repeat here the same steps performed in the calculation of g^* and ω . Therefore, we determine g^* and Z_1 , by solving the equations:

$$f(g^*; c, p) = Z_0, \quad Z_1 = f'(g^*; c, p). \quad (5)$$

Of course, $f(g; c, p)$ is replaced with an appropriate resummation of its perturbative expansion. We use here the resummation scheme proposed in Ref. [27] that makes explicit use of the location of the Borel-transform singularity, but similar results are obtained extending the Borel transform by means of Padé approximants (note that one could also use the perturbative series in the *bare*

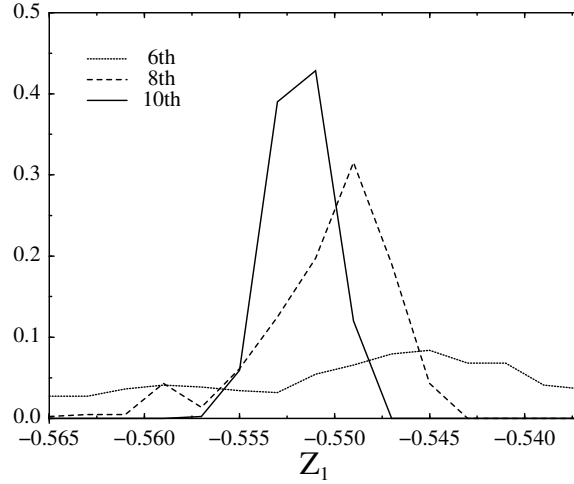


Figure 2. Distribution of the results for the resummations of Z_1 for $p = 1/10$, $c = -Z_1/5$.

coupling [28]). The mean values and errors are determined by using the procedure of Ref. [29]. In the absence of the nonanalytic term, i.e. for $c = 0$, using the n th-order expansion, we obtain

$$\begin{aligned} n = 6: & \quad g^* = 1.00025(131) \quad Z_1 = -0.6791(178), \\ n = 8: & \quad g^* = 0.99997(10) \quad Z_1 = -0.6800(18), \\ n = 10: & \quad g^* = 1.00000(1) \quad Z_1 = -0.6791(2). \end{aligned}$$

There is good agreement, the precision increases by a factor of 10 every two orders, and the error bars are correct.

The next step is to consider the role of the nonanalytic corrections, by adding a term that is small compared to the analytic one. We choose $c = -Z_1/5$.

Now, for $p = 1/10$ we obtain

$$\begin{aligned} n = 6: & \quad g^* = 1.0043(62) \quad Z_1 = -0.550(20), \\ n = 8: & \quad g^* = 1.0066(15) \quad Z_1 = -0.550(3), \\ n = 10: & \quad g^* = 1.0062(5) \quad Z_1 = -0.552(2). \end{aligned}$$

In this case the agreement is poor, especially for Z_1 , and, even worse, the errors are completely incorrect. This can be understood from Figs. 1 and 2

where we show the distribution of the approximants that are used. These distributions are nicely peaked, but unfortunately at an incorrect value of g^* and Z_1 . Thus, in the presence of these (strong) nonanalyticities, the fact that the approximants have a narrow distribution is not a good indication that the results are reliable. Also, the stability of the results with the number of terms of the series is completely misleading. As we shall discuss below, this is what we believe is happening in two dimensions.

If we consider instead a weak nonanalyticity, i.e. $p \approx 1$, the discrepancies we have found for $p = 1/10$ are much smaller, although still present. For instance, for $p = 9/10$ we have

$$\begin{aligned} n = 6: & \quad g^* = 1.013(45) \quad Z_1 = -0.755(270), \\ n = 8: & \quad g^* = 1.014(7) \quad Z_1 = -0.655(23), \\ n = 10: & \quad g^* = 1.006(4) \quad Z_1 = -0.668(2). \end{aligned}$$

In this case the results are consistent with the exact values, although the errors are still slightly underestimated. As expected, the largest discrepancies are observed for Z_1 .

4 Conclusions

We have shown that nonanalytic terms may give rise to systematic deviations and a systematic underestimate of the error bars. Now, what should we expect in the interesting two- and three-dimensional cases?

In three dimensions $\Delta \approx 0.5$ and Δ_2/Δ is approximately 2 [30]. Thus, the leading nonanalytic term has the exponent Δ_2/Δ and is not very different from an analytic one. As a consequence, we expect small corrections and indeed the FT results are in substantial agreement with the estimates obtained in MC and HT studies. However, small differences are observed for γ and ω for $N = 0$ and $N = 1$:

$$\begin{array}{llll} N = 0 & \gamma_{\text{FT}} = 1.1596(20) & \text{Ref. [9],} & \gamma_{\text{MC}} = 1.1575(6) \quad \text{Ref. [31],} \\ N = 0 & \Delta_{\text{FT}} = 0.478(10) & \text{Ref. [9],} & \Delta_{\text{MC}} = 0.517(7)_{-0}^{+10} \quad \text{Ref. [32],} \\ N = 1 & \gamma_{\text{FT}} = 1.2396(13) & \text{Ref. [9],} & \gamma_{\text{HT}} = 1.2371(4) \quad \text{Ref. [14],} \\ N = 1 & \Delta_{\text{FT}} = 0.504(8) & \text{Ref. [9],} & \Delta_{\text{MC}} = 0.533(6) \quad \text{Ref. [33].} \end{array}$$

There are slight differences, especially for ω , but still at the level of a few error bars. Note that, as discussed in Ref. [10], part of the error may be due to a slightly incorrect estimate of g^* . Using the estimate of g^* obtained from the analysis of the HT expansions, the FT estimates change towards the HT and MC values.

Larger discrepancies are observed in two dimensions. For $N \geq 3$ it is easy to predict the behavior of the RG functions at the critical point. Indeed, the theory is massive for all temperatures. The critical behavior is controlled by the zero-temperature Gaussian point and can be studied in perturbation theory in the corresponding N -vector model. One finds only logarithmic corrections to the purely Gaussian behavior. It follows that the operators have dimensions that coincide with their naive (engineering) dimensions, apart from logarithmic multiplicative corrections related to the so-called anomalous dimensions. The leading irrelevant operator has dimension two [34,35] and thus, for $m \rightarrow 0$, we expect

$$g(m) = g^* \left\{ 1 + c m^2 (-\ln m^2)^\zeta \left[1 + O\left(\frac{\ln(-\ln m^2)}{\ln m^2}\right) \right] \right\}, \quad (6)$$

where ζ is an exponent related to the anomalous dimension of the leading irrelevant operator, and c is a constant. Therefore,

$$\beta(g) = m \frac{\partial g}{\partial m} = -2\Delta g \left(1 + \frac{\zeta}{\ln \Delta g} + \dots \right), \quad (7)$$

with $\Delta g \equiv g^* - g$. Clearly there are logarithmic corrections in this case and therefore, we expect large deviations in the determinations of ω , which should be 2. These deviations are indeed observed: for $N = 3$ the analysis of the five-loop series [36] yields the estimate $\beta'(g^*) = 1.33(2)$, which is very different from the expected result $\beta'(g^*) = 2$.

For $N = 1$, we can repeat Nickel's analysis, using the fact that, from conformal field theory, we can compute the RG dimensions of all relevant and irrelevant operators. Using these results we predict that^c

$$\beta(g) = -\frac{7}{4}\Delta g \left(1 + b_1 |\Delta g|^{1/7} + b_2 |\Delta g|^{2/7} + b_3 |\Delta g|^{3/7} + \dots \right), \quad (8)$$

where $\Delta g \equiv g^* - g$. Such an expansion is confirmed by an analysis of the lattice Ising model. Again we find strong nonanalyticities, and correspondingly we expect large deviations. And, indeed such large deviations are observed: the analysis of the five-loop series gives $g^* = 15.39(25)$ [36] and $\beta'(g^*) = 1.31(3)$, to be compared with the exact prediction $\beta'(g^*) = 7/4$ and the estimates $g^* = 14.69735(3)$ in Ref. [37] and $g^* = 14.6975(1)$ in Ref. [38].

^cSometimes it has been claimed that the leading irrelevant operator in the Ising universality class has $\omega = 4/3$, so that $\beta'(g^*) = 4/3$. This claim is incorrect. Indeed, such an operator can only appear in nonunitary extensions of the Ising model, but not in the standard (unitary) ϕ^4 -field theory. For a detailed discussion, see Appendix A of Ref. [26].

References

- [1] J.A. Lipa, D.R. Swanson, J.A. Nissen, T.C.P. Chui, and U.E. Israelsson, *Phys. Rev. Lett.* **76**, 944 (1996).
- [2] J.A. Lipa, D.R. Swanson, J.A. Nissen, Z.K. Geng, P.R. Williamson, D.A. Stricker, T.C.P. Chui, U.E. Israelsson, and M. Larson, *Phys. Rev. Lett.* **84**, 4894 (2000).
- [3] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi, and E. Vicari, eprint: cond-mat/0010360.
- [4] M. Campostrini, A. Pelissetto, P. Rossi, and E. Vicari, *Phys. Rev. B* **61**, 5905 (2000).
- [5] M. Hasenbusch and T. Török, *J. Phys. A* **32**, 6361 (1999).
- [6] H. Kleinert, eprint: cond-mat/9906107.
- [7] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [8] F. Jasch and H. Kleinert, *J. Math. Phys.* **42**, 52 (2001), eprint: cond-mat/9906246v2.
- [9] R. Guida and J. Zinn-Justin, *J. Phys. A* **31**, 8103 (1998).
- [10] A. Pelissetto and E. Vicari, *Nucl. Phys. B* **519**, 626 (1998).
- [11] G.A. Baker Jr. and N. Kawashima, *J. Phys. A* **29**, 7183 (1996).
- [12] S. Zinn, S.-N. Lai, and M. E. Fisher, *Phys. Rev. E* **54**, 1176 (1996).
- [13] P. Butera and M. Comi, *Phys. Rev. E* **55**, 6391 (1997).
- [14] M. Campostrini, A. Pelissetto, P. Rossi, and E. Vicari, *Phys. Rev. E* **60**, 3526 (1999).
- [15] J.H. Chen, M.E. Fisher, and B.G. Nickel, *Phys. Rev. Lett.* **48**, 630 (1982).
- [16] M. Hasenbusch, K. Pinn, and S. Vinti, *Phys. Rev. B* **59**, 11471 (1999).
- [17] H.G. Ballesteros, L.A. Fernández, V. Martín-Mayor, A. Muñoz Sudupe, G. Parisi, and J.J. Ruiz-Lorenzo, *J. Phys. A* **32**, 1 (1999).
- [18] M. Hasenbusch, *J. Phys. A* **32**, 4851 (1999).
- [19] C. Bagnuls and C. Bervillier, eprint: hep-th/0002034.
- [20] G. Parisi, *J. Stat. Phys.* **23**, 49 (1980).
- [21] B.G. Nickel, in *Phase Transitions*, Eds. M. Lévy, J.C. Le Guillou, and J. Zinn-Justin (Plenum Press, New York and London, 1982).
- [22] A.D. Sokal, *Europhys. Lett.* **27**, 661 (1994).
- [23] B. Li, N. Madras, and A.D. Sokal, *J. Stat. Phys.* **80**, 661 (1995).
- [24] C. Bagnuls and C. Bervillier, *J. Phys. Stud.* **1**, 366 (1997).
- [25] A. Aharony and M.E. Fisher, *Phys. Rev. B* **27**, 4394 (1983).

- [26] P. Calabrese, M. Caselle, A. Celi, A. Pelissetto, and E. Vicari, *J. Phys. A* **33**, 8155 (2000), eprint: hep-th/0005254 .
- [27] J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev. Lett.* **39**, 95 (1977); *Phys. Rev. B* **21**, 3976 (1980).
- [28] H. Kleinert, *Phys. Lett. B* **434**, 74 (1998); *ibid.* **463**, 69 (1999); *Phys. Rev. D* **57**, 2264 (1998), **58**, 107702 (1998).
- [29] J.M. Carmona, A. Pelissetto, and E. Vicari, *Phys. Rev. B* **61**, 15136 (2000).
- [30] G.R. Golner and E.K. Riedel, *Phys. Lett. A* **58**, 11 (1976); K.E. Newman and E.K. Riedel, *Phys. Rev. B* **30**, 6615 (1984).
- [31] S. Caracciolo, M.S. Causo, and A. Pelissetto, *Phys. Rev. E* **57**, R1215 (1998).
- [32] P. Belohorec and B.G. Nickel, unpublished Guelph University report (1997).
- [33] M. Hasenbusch, *Monte Carlo Studies of the Three-Dimensional Ising Model*, Habilitationsschrift, Humboldt-Universität zu Berlin (1999).
- [34] E. Brézin, J. Zinn-Justin, and J.C. Le Guillou, *Phys. Rev. B* **14**, 4976 (1976).
- [35] K. Symanzik, *Nucl. Phys. B* **226**, 187 and 205 (1983).
- [36] E.V. Orlov and A.I. Sokolov, *Fiz. Tverd. Tela* **42**, 2087 (2000). A shorter English version appeared as eprint: hep-th/0003140.
- [37] M. Caselle, M. Hasenbusch, A. Pelissetto, and E. Vicari, *J. Phys. A* **33**, 8171 (2000), eprint: hep-th/0003049.
- [38] J. Balog, M. Niedermaier, F. Niedermayer, A. Patrascioiu, E. Seiler, and P. Weisz, *Nucl. Phys. B* **583**, 614 (2000).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

**A REMARK ON THE NUMERICAL VALIDATION OF
TRIVIALITY FOR SCALAR FIELD THEORIES
BY HIGH-TEMPERATURE EXPANSIONS**

P. BUTERA AND M. COMI

*Istituto Nazionale di Fisica Nucleare,
Dipartimento di Fisica, Università di Milano-Bicocca,
3 Piazza della Scienza, 20126 Milano, Italy
E-mails: butera@mi.infn.it; comi@mi.infn.it*

We suggest a simple modification of the usual procedures of analyzing the high-temperature (strong-coupling or hopping-parameter) expansions of the renormalized four-point coupling constant in the $\vec{\phi}_4^4$ lattice scalar field theory. As a result we can more convincingly validate numerically the triviality of the continuum limit taken from the high-temperature phase.

There has been a steady accumulation of suggestive numerical and analytical evidence, but not yet a complete rigorous proof that the continuum limit of the lattice regularized N -component $\vec{\phi}_4^4$ -theory describes a free (or “trivial”) field theory [1–4]. The basic clues of this paradoxical no-interaction property were indicated almost half a century ago [5], but more stringent studies of this conjecture had to wait for the developments of the Renormalization Group (RG) theory [6]. The modern rigorous analyses of Refs. [1,2,7–10] have finally proved that the $\vec{\phi}_d^4$ -theory is non-trivial [11] in $d \leq 3$ and trivial in $d \geq 5$ dimensions, at least for not too large values of N . Triviality is expected to occur also in $d = 4$ dimensions. Since, however, the rigorous results in this direction are still partial, various routes to recover an interesting continuum theory have also been explored [12,13].

The lattice Euclidean $\vec{\phi}_4^4$ -theory with $O(N)$ symmetry is defined by the action [14]

$$S = \sum_x \left\{ -\beta \sum_{\mu} \vec{\phi}_x \cdot \vec{\phi}_{x+\mu} + \vec{\phi}_x^2 + \lambda(\vec{\phi}_x^2 - 1)^2 \right\}, \quad (1)$$

where $\vec{\phi}_x$ is a real N -component field at the lattice site x and μ is the unit vector in the μ direction and $\lambda \geq 0$.

For $\beta \uparrow \beta_c(N, \lambda)$, at fixed positive λ the model has a critical point where a second-order transition occurs from a high-temperature (HT) paramagnetic phase to a low-temperature ferromagnetic phase. In the $\lambda \rightarrow \infty$ limit, the model leads to the lattice nonlinear $O(N)$ -symmetric σ -model or, equivalently, to the N -vector spin model.

The construction of a continuum limit of the lattice theory is reduced to the determination of its critical properties. Here we shall consider only the continuum limit taken from the HT phase. In the context of the RG theory a detailed description is obtained for the asymptotic cutoff dependence of the correlation functions in terms of the weak-coupling expansion of the theory's beta function.

If we set $\tau(N, \lambda) = 1 - \beta/\beta_c(N, \lambda)$, the perturbative RG theory yields [15] the following critical behavior as $\tau(N, \lambda) \downarrow 0$ for the correlation length

$$\xi^2(\beta, N, \lambda) = A_\xi^2(N) \frac{|\ln(\tau(N, \lambda))|^{G(N)}}{\tau(N, \lambda)} \left[1 + O\left(\ln(|\ln(\tau)|)/\ln(\tau)\right) \right], \quad (2)$$

where $G(N) = N + 2/N + 8$.

The asymptotic behavior of the susceptibility is completely similar:

$$\chi(\beta, N, \lambda) = A_\chi(N) \frac{|\ln(\tau(N, \lambda))|^{G(N)}}{\tau(N, \lambda)} \left[1 + O\left(\ln(|\ln(\tau)|)/\ln(\tau)\right) \right]. \quad (3)$$

The fourth derivative of the free energy at zero field $\chi_4(\beta, N, \lambda)$ has the behavior

$$\chi_4(\beta, N, \lambda) = A_4(N) \frac{|\ln(\tau(N, \lambda))|^{4G(N)-1}}{\tau(N, \lambda)^4} \left[1 + O\left(\ln(|\ln(\tau)|)/\ln(\tau)\right) \right]. \quad (4)$$

In terms of χ , ξ^2 and χ_4 , the dimensionless renormalized 4-point coupling constant $g_r(N, \lambda)$ is defined by the critical value of the ratio

$$g_r(\beta, N, \lambda) = -\frac{\chi_4(\beta, N, \lambda)}{\xi^4(\beta, N, \lambda)\chi^2(\beta, N, \lambda)} \quad (5)$$

as $\tau(N, \lambda) \downarrow 0$.

It can be shown that $g_r(\beta, N, \lambda)$ is non-negative [16] for all β . If $g_r(\beta, N, \lambda)$ vanishes as $\tau(N, \lambda) \downarrow 0$, the continuum limit of the lattice model taken from the high-temperature phase describes a (generalized)-free-field theory [17],

namely a theory where the connected parts of the four-point and higher-point functions vanish.

As $\tau(N, \lambda) \downarrow 0$, the perturbative RG yields the leading asymptotic behavior, with a well-specified universal amplitude [14]

$$g_r(\beta, N, \lambda) \approx \frac{c(N)}{|\ln(\tau(N, \lambda))|} \left[1 + O\left(\ln(|\ln(\tau)|)/\ln(\tau)\right) \right], \quad (6)$$

where $c(N) = 2/b_1(N)$ and $b_1(N) = N + 8/48\pi^2$ is the first non-vanishing coefficient of the beta-function.

Therefore the perturbative RG theory implies that $g_r(\beta, N, \lambda) \rightarrow 0$ as $\tau(N, \lambda) \downarrow 0$, namely that the continuum $\bar{\phi}_4^4$ -model is trivial.

Since the validity of the results in Eqs. (2), (3), (4), and (6) is based upon the (unwarranted) perturbative determination of the beta function, it is interesting, at least, to try to confirm them within different approximation schemes. To this purpose, various HT or strong-coupling expansion analyses [6,14,18–22] have been performed. Many extensive Monte Carlo (MC) lattice simulations [23–31] have also been carried out and progressively refined over the years, due to the rapid evolution of computers and the improvement of algorithms and data analysis. Until now, both the stochastic simulation and the HT series studies have been generally carried out in a completely parallel way. For instance, in the case of the $4d$ self-avoiding walk model (namely, the $\bar{\phi}_4^4$ -model for $N = 0$ and $\lambda \rightarrow \infty$) a HT expansion of χ up to order β^{13} on the hypercubic lattice has been analyzed [18] in order to detect the logarithmic factor predicted by the RG theory in Eq. (3) and to estimate its exponent. For the same purpose, χ and ξ^2 have been measured [23,30] in high precision MC simulations. Analogous studies [19,21] have been devoted to the $4d$ Ising model (namely, the $N = 1$ case for $\lambda \rightarrow \infty$) using series $\mathcal{O}(\beta^{17})$ for χ and χ_4 on the hypercubic lattice. The MC simulations of Refs. [24,25,28] have tried to show directly the consistency of the estimates of $g_r(\beta, N, \lambda)$ with the elusive asymptotic behavior shown in Eq. (6). A somewhat different approach, based on the scaling properties of the partition function zeroes in the complex β plane, has been adopted in the simulations of Ref. [31]. Moreover, various analytical or semianalytical approaches [32,33] have also been pursued.

All of these non-perturbative calculations have given results consistent, or at least not in contrast, with the predicted critical behaviors of χ , ξ , and g_r . However, the cited computations are somewhat limited in their extent, since only the $N = 0$ and $N = 1$ cases for $\lambda \rightarrow \infty$ and the $N = 1$ case for finite λ

to order β^{10} have been considered until now [20]. Moreover, it is a common experience that it is difficult to uncover numerically a logarithmic behavior or, more generally, a logarithmic correction to a power behavior. Indeed, as the computations proceed deeply into the asymptotic regime, their reliability decreases and the uncertainties of their results often reach almost the same order of magnitude as the effects that have to be revealed. In the case of the HT expansions, we have also observed that the methods of Refs. [18,21], which were very effective in the $N = 0$ and $N = 1$ cases, are not as successful when $N > 1$.

In this note, we do not present new data, but reconsider the HT expansions calculated through order β^{14} , more than a decade ago, by Lüscher and Weisz [14] for χ , ξ^2 , and χ_4 on the hypercubic lattice. They have produced and analyzed these series to obtain a bound on the Higgs particle mass as a consequence of the triviality of the scalar sector in the standard model. From the outset, they used also the assumption of validity of the perturbative RG and therefore avoided to place too much confidence in the HT series within the critical regime. We make no such assumption, but rather suggest a slightly different and hopefully more convincing way of analyzing the series, which takes advantage of the specific smoothness features of the HT expansion approach and, in the end, also turns out to be completely consistent with the RG results. We study how accurately an obvious consequence of Eq. (6), rather than the equation itself, is confirmed by computations.

One can see that Eq. (6) implies

$$F(\beta, N, \lambda) = \tau(N, \lambda) \frac{d}{d\beta} \left[\frac{1}{g_r(\beta, N, \lambda)} \right] = \frac{b_1(N)}{2} + O\left(\ln(|\ln(\tau)|)/\ln^3(\tau)\right) \quad (7)$$

as $\tau(N, \lambda) \downarrow 0$.

In order to confirm triviality, we then show at least that $F(\beta, N, \lambda)$ has a finite limit $\tilde{F}(N, \lambda)$, as $\tau(N, \lambda) \downarrow 0$. The analysis is even more compelling if

- i) $\tilde{F}(N, \lambda) \approx \tilde{F}(N)$, namely if the quantity $\tilde{F}(N, \lambda)$ does not depend on λ , as required by universality, and if, moreover,
- ii) $2\tilde{F}(N)/b_1(N) \approx 1$, namely if, unlike in previous approaches, it is possible to show complete quantitative consistency between the strong-coupling estimate of g_r , including the universal amplitude $2/b_1(N)$, and the weak-coupling RG prediction of Eqs. (6) and (7).

Since the HT series coefficients can be written as simple rational functions of N [34], we can easily repeat the analysis on a wide range of values of N

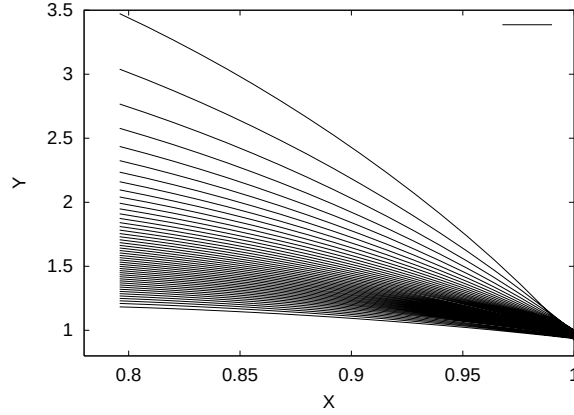


Figure 1. The quantity $y = 2\tilde{F}(\beta, N, \lambda)/b_1(N)$ versus the scaled variable $x = \beta/\beta_c(N, \lambda)$, with $N = 4$. Going from the top to the bottom of the figure, the various curves correspond to increasing values of λ between 0 and ∞ .

and thus further corroborate this result.

Let us stress that the whole analysis cannot be easily performed in the context of a MC simulation, whereas it is completely straightforward in a HT series approach.

The main results of our procedure can be summarized into a couple of figures. In Fig. 1 we have plotted for $N = 1$ the quantity $y = 2\tilde{F}(\beta, N, \lambda)/b_1(N)$ versus the scaled variable $x = \beta/\beta_c(N, \lambda)$, in order to be able to compare the curves obtained for various fixed values of λ . The values of $\beta_c(N, \lambda)$ used here are estimated by an analysis of the susceptibility expansions. We have calculated $F(\beta, N, \lambda)$ by simply forming Padé approximants (PA) of its HT expansion. For each value of λ , we have plotted only the highest non-defective PA, namely the [6/6] or the [6/7] approximants, as appropriate. The other approximants of sufficiently high order have the same behavior and are not reported in the figure. As $x \rightarrow 1$, the various curves obtained in this way appear to tend to unity, independently of λ , within a good approximation, thus confirming i). We expect that the residual small spread of the limiting values would be significantly reduced if we could further reduce the uncertainties in the determination of $\beta_c(N, \lambda)$ and if we could devise approximants more accurately allowing for the slowly vanishing corrections to scaling indicated in Eq. (7). Of course, these improvements are strictly related.

We have performed the above calculation for various values of N . Figure 2

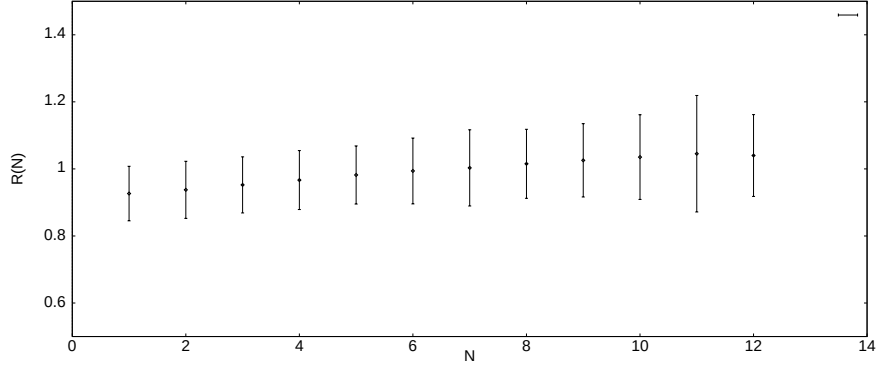


Figure 2. The ratio $R(N) = 2\tilde{F}(N)/b_1(N)$ versus N .

shows the ratio $R(N) = 2\tilde{F}(N)/b_1(N)$ versus N . For each value of N the reported error reflects the spread of this quantity. Within a fair approximation, $R(N)$ appears to be unity over a wide range of values of N , thus confirming ii).

Therefore both figures indicate a good quantitative agreement with the asymptotic formula Eq. (6), obtained within the perturbative RG approach. These very general results are unlikely to be accidental and completely confirm the conventional expectations concerning triviality.

The HT series we have used in this first test are unfortunately too short to make a more refined analysis possible. The favourable results, however, encourage to resume this study as soon as our systematic work of HT series extension by the linked-cluster method [35] will make new longer expansions for this model available. We do not expect results to be qualitatively different from this preliminary study, but significant quantitative improvements.

At this point, it is interesting to recall that a similar result on the consistency between the strong coupling and the weak coupling approaches has already been reported for the $O(N)$ symmetric lattice nonlinear σ -model in two dimensions. Indeed, also for this model, the first perturbative coefficient of the beta- and of the gamma-functions has been computed [36] with a good accuracy, starting from a strong-coupling expansion.

In conclusion, we have shown that a small modification of the current procedures of numerical analysis is sufficient to shift the emphasis from a difficult qualitative question, namely how accurately an elusive logarithmic behavior is reproduced by an approximation scheme of inevitably limited

precision, to a more quantitative issue. Our technique of analysis for HT series is not more involved than the usual ones, while, for all values of N , it seems to produce a more convincing numerical validation of the perturbative-RG triviality predictions within the strong-coupling approach.

References

- [1] A.D. Sokal, *Ann. Inst. H. Poincaré A* **37**, 317 (1982).
- [2] A. Fernandez, J. Fröhlich, and A. Sokal, *Random Walks, Critical Phenomena, and Triviality in Quantum Field Theory* (Springer Verlag, Berlin, 1992).
- [3] K. Symanzik, *J. Phys. (France)* **43**, Suppl. C3, 254 (1982).
- [4] D. Callaway, *Phys. Rep.* **167**, 241 (1988).
- [5] L.D. Landau and I. Pomeranchuk, *Dokl. Akad. Nauk. SSSR* **102**, 489 (1955); I. Pomeranchuk, V.V. Sudakov, and K.A. Ter Martirosian, *Phys. Rev.* **103**, 784 (1956).
- [6] K.G. Wilson and J.B. Kogut, *Phys. Rep. C* **12**, 75 (1974).
- [7] M. Aizenmann, *Phys. Rev. Lett.* **47**, 1 (1981).
- [8] J. Fröhlich, *Nucl. Phys. B* **200**, 281 (1982).
- [9] T. Hara, *J. Stat. Phys.* **47**, 57 (1987); T. Hara and H. Tasaki, *J. Stat. Phys.* **47**, 99 (1987).
- [10] S.B. Shlosman, *Sov. Phys. Dokl.* **33**, 905 (1988).
- [11] J.P. Eckmann and R. Epstein, *Commun. Math. Phys.* **64**, 95 (1979).
- [12] G. Gallavotti and V. Rivasseau, *Ann. Inst. H. Poincaré* **40**, 185 (1984).
- [13] P. Cea, M. Consoli, L. Cosmai, and P.M. Stevenson, *Mod. Phys. Lett. A* **14**, 1673 (1999), and references therein.
- [14] M. Lüscher and P. Weisz, *Nucl. Phys. B* **300**, 325 (1988); *Nucl. Phys. B* **290**, 25 (1987).
- [15] E. Brézin, J.C. Le Guillou, and J. Zinn-Justin, in *Phase Transitions and Critical Phenomena*, Vol. VI, Eds. C. Domb and M.S. Green (Academic Press, New York, 1976); J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 3rd ed. (Clarendon, Oxford, 1996).
- [16] J.L. Lebowitz, *Commun. Math. Phys.* **35**, 87 (1974).
- [17] C.M. Newman, *Commun. Math. Phys.* **41**, 1 (1975).
- [18] A.J. Guttmann, *J. Phys. A* **11**, L103 (1978).
- [19] D.S. Gaunt, M.F. Sykes, and S. McKenzie, *J. Phys. A* **12**, 871 (1979).
- [20] G.A. Baker Jr. and J.M. Kincaid, *Phys. Rev. Lett.* **42**, 1431 (1979); *ibid. (E)* **44**, 434 (1980); *J. Stat. Phys.* **24**, 469 (1981).

- [21] A. Vladikas and C.C. Wong, *Phys. Lett. B* **189**, 154 (1987).
- [22] C.M. Bender, F. Cooper, G.S. Guralnik, R. Roskies, and D.H. Sharp, *Phys. Rev. D* **23**, 2976 (1981); *ibid.* **23**, 2999 (1981); G.A. Baker Jr., L.P. Benofy, F. Cooper, and D. Preston, *Nucl. Phys. B* **210**, 273 (1982); C.M. Bender and H.F. Jones, *Phys. Rev. D* **38**, 2526 (1988).
- [23] C.A. de Carvalho, J. Fröhlich, and S. Caracciolo, *Nucl. Phys. B* **215**, 209 (1983).
- [24] B. Freedman, P. Smolenski, and D. Weingarten, *Phys. Lett. B* **113**, 481 (1982).
- [25] I.A. Fox and I.G. Halliday, *Phys. Lett. B* **159**, 148 (1985).
- [26] I.T. Drummond, S. Duane, and R.R. Horgan, *Nucl. Phys. B* **280**, 25 (1987).
- [27] W. Bernreuther, M. Göckeler, and M. Kremer, *Nucl. Phys. B* **295**, 211 (1988).
- [28] C. Frick, K. Jansen, J. Jersak, I. Montvay, G. Münster, and P. Seufferling, *Nucl. Phys. B* **331**, 515 (1990).
- [29] J.K. Kim and A. Patrascioiu, *Phys. Rev. D* **47**, 2588 (1992).
- [30] P. Grassberger, R. Hegger, and L. Schäfer, *J. Phys. A* **27**, 7265 (1994).
- [31] R. Kenna and C.B. Lang, *Phys. Rev. E* **49**, 5012 (1994).
- [32] W. Bardeen and M. Moshe, *Phys. Rev. D* **28**, 1372 (1983).
- [33] D. Gromes, *Z. Phys. C* **71**, 347 (1996).
- [34] P. Butera, M. Comi, and G. Marchesini, *Phys. Rev. B* **41**, 11494 (1990).
- [35] P. Butera and M. Comi, *Phys. Rev. B* **47**, 11969 (1993); *ibid.* **B 50**, 3052 (1994); *ibid.* **B 52**, 6185 (1995); *ibid.* **E 55**, 6391 (1997); *ibid.* **B 56**, 8212 (1997); *ibid.* **B 58**, 11552 (1998); *ibid.* **B 60**, 6749 (1999); *Nucl. Phys. B (Proc. Suppl.)* **63** A-C, 643 (1998); eprint: hep-lat/0006009.
- [36] P. Butera and M. Comi, *Phys. Rev. B* **54**, 15828 (1996).

MULTILOOP ϕ^4 -THEORY AT CRITICALITY

J.A. GRACEY

*Theoretical Physics Division, Department of Mathematical Sciences,
University of Liverpool, Peach Street, Liverpool, L69 7ZF, UK*

E-mail: jag@amtp.liv.ac.uk

We review the contribution of Hagen Kleinert's group to the computation of the $\overline{\text{MS}}$ renormalization group functions of four-dimensional $O(N)$ ϕ^4 -theory at *five* loops. The structure of the β -function beyond this order is also discussed from the point of view of recent developments in connecting knot theory with the transcendental numbers which appear at three and higher loops as well as the large N expansion.

1 Introduction

It is a pleasure and an honour to participate in this publication celebrating Professor Dr. Hagen Kleinert's sixtieth birthday and to review one of his many contributions related to multiloop calculations underlying critical phenomena [1]. In particular we will recall the role a particular four-dimensional scalar field theory plays in understanding phase transitions occurring in nature. For example, ϕ^4 -field theory endowed with an $O(N)$ internal symmetry in three space-time dimensions underpins the statistical properties of long polymer chains, ($N = 0$), it relates to phase transitions in Ising like systems and the physics of classical fluid liquid vapour transition, ($N = 1$), it deals with Helium superfluid transition, ($N = 2$), and ferromagnetic systems, ($N = 3$) [2].

A key element to understand the physics of these various phase transitions experimentally and theoretically are the fundamental critical exponents. These are universal quantities from the point of view of the renormalization group equation and govern the scaling behavior of, say, the specific heat or susceptibility. Indeed they are presently being measured more accurately and

hence current theoretical input must be ingenious enough to compete with the progress being made. Since the critical exponents of the underlying quantum field theory are simply related to the renormalization group functions evaluated at the critical coupling or temperature, the issue is one of computing quantities such as the β -function or field anomalous dimensions to as high a loop order as is humanly or computerly possible. There are several approaches to this problem in relation to scalar field theories such as ϕ^4 -theory which is renormalizable in four dimensions but superrenormalizable in three dimensions. For instance, one can compute directly in the three-dimensional theory (see, for example, Refs. [3–5] and references therein). Alternatively, the four-dimensional theory can be renormalized to determine the renormalization group functions in the $\overline{\text{MS}}$ scheme. These are used to deduce the critical exponents as series in ϵ , where $d = 4 - 2\epsilon$ is the space-time dimension which then need to be resummed since the expressions are formally divergent or asymptotic. Indeed there are various ways of dealing with this resummation problem to improve numerical accuracy for the exponents though the standard method is Padé-Borel summation. Alternative approaches have been developed more recently by Kleinert, for example, which are based on a strong-coupling method of a variational technique which does not make use of renormalization methods and has been applied to the strictly three-dimensional model [6] and the $(4 - 2\epsilon)$ -dimensional model [7–9]. Numerical results obtained for the final three-dimensional critical exponents from this new approach are impressive and also competitive with other series improvement techniques. However, the main aim of this article is to recall the computation of Kleinert *et al.* [1] and some issues concerning the renormalization group equations of scalar ϕ^4 -field theories in four and other dimensions, since resummation techniques rely heavily upon having the explicit information at hand, as well as important recent insights into the structures which lie beyond the five-loop results of Ref. [1].

The paper is organised as follows. Background to ϕ^4 -theory is discussed in Section 2, where the five-loop results of Ref. [1] are reviewed. In Section 3 the structure of the four-dimensional $\overline{\text{MS}}$ renormalization group functions at six and more loops is examined in relation to connections with knot theory and the large N expansion. Concluding remarks are contained in Section 4.

2 The ϕ^4 -Theory

The underlying quantum field theory which governs the above critical phenomena is ϕ^4 -theory whose (massless) Lagrangian is

$$L = \frac{1}{2} \partial_\mu \phi_0^a \partial^\mu \phi_0^a + \frac{16\pi^2}{4!} g_0 (\phi_0^a \phi_0^a)^2, \quad (1)$$

where the field ϕ_0^a and the coupling constant g_0 are bare and $1 \leq a \leq N$. The factor of $16\pi^2$ associated with the coupling constant is included so that the expansion of the renormalization group functions is in terms of g rather than the usual $g/(16\pi^2)$. It is instructive to compare Eq. (1), reformulated in terms of renormalized parameters, with the criticality version. Introducing the renormalized parameters $\phi = \phi_0 \sqrt{Z_\phi}$ and $g = g_0 Z_g$, we get

$$L = \frac{Z_\phi}{2} \partial_\mu \phi^a \partial^\mu \phi^a + \frac{16\pi^2}{4!} \tilde{\mu}^{4-d} g Z_g Z_\phi^2 (\phi^a \phi^a)^2, \quad (2)$$

where Z_ϕ and Z_g are computed in a regularized version of the theory. Since the four-dimensional models are related to lower-dimensional models, one uses here dimensional regularization where the arising singularities will appear as poles in ϵ . Moreover, they are subtracted in a (modified) minimal way which allows to compute the renormalization group functions to as high a loop order as possible. The scale $\tilde{\mu}$ is introduced in Eq. (2) to ensure the coupling constant to remain dimensionless in d dimensions.

By contrast in the critical region in d dimensions the action S governing the phase transition has a different form. If ξ is some length scale at the fixed point, then S is formally

$$S = \int d^d x \left[\frac{\xi^{-\eta}}{2} \partial_\mu \phi^a \partial^\mu \phi^a + \frac{\xi^\chi}{2} \sigma \phi^a \phi^a - \frac{3\xi^{\chi_\sigma^2} \sigma^2}{32\pi^2 g} - \frac{\sigma}{2\lambda} + \dots \right], \quad (3)$$

where the auxiliary field σ has been introduced in order to have a trivalent interaction at criticality. If one ignores the higher terms, the elimination of the trivalent interaction would restore the four-point interaction of Eq. (1). The omitted terms involve composite operators of the fields ϕ^a , σ and their derivatives. Each will have its own coupling constant which will either be relevant, irrelevant or marginal at criticality. For instance, the coupling constant λ corresponds to the coupling of the linear term in σ . If one ignores for the moment the presence of the scaling dimension with each term, then the interaction which dominates the transition is $\sigma \phi^2$. The remaining terms

correspond to quadratic or linear terms. They are present to *illustrate* an important feature of the relation of ϕ^4 -theory to lower-dimensional models. The additional coupling λ in fact corresponds to the coupling constant of the two-dimensional $O(N)$ nonlinear σ model whose Lagrangian can be written in a form analogous to Eq. (1)

$$L = \frac{1}{2} \partial_\mu \phi_0^a \partial^\mu \phi_0^a + \frac{\sigma_0}{2} \left(\phi_0^a \phi_0^a - \frac{1}{\lambda_0} \right), \quad (4)$$

which is renormalizable in two dimensions, but whose renormalization group functions can also be used to determine the critical exponents of the three-dimensional transitions. The point is that, at criticality, Eq. (3) is the full underlying theory. It is related to the boundary dimension models, where one reduces the space-time dimensionality from four, σ^2 becomes irrelevant in two dimensions whereas σ becomes relevant. On the contrary, when approaching four dimensions, σ^2 is relevant but σ becomes an irrelevant operator. In other words both models are equivalent at the appropriate critical point of their β -function. Thus either model can be used to determine critical exponents.

The power of the scaling dimension ξ in each term of Eq. (3) represents the anomalous dimension of that operator which is generated by radiative corrections in the quantum theory^a. For instance, η is the ϕ -field anomalous dimension and is measured experimentally. In field theory, its numerical value is a reflection of the size of all radiative corrections. Therefore, by simple dimensional analysis, where α is the full dimension of ϕ , the results of Refs. [10,11] are

$$\alpha = \mu - 1 + \frac{1}{2}\eta, \quad (5)$$

where $d = 2\mu$. Likewise the anomalous dimension of the trivalent interaction is defined to be χ , giving [10,11]

$$\beta = 2 - \eta - \chi, \quad (6)$$

where β is the full σ field dimension. For the remaining two terms one finds the respective scaling laws

$$\chi_{\sigma^2} = 2\mu - 2\beta - 2\omega, \quad \beta = 2\mu - \frac{1}{\nu}. \quad (7)$$

^aThe term linear in σ does not have an associated anomalous operator dimension as this is already incorporated in its dimension β . The scaling law which arises from this term is recorded later.

Having discussed the relation of the underlying theories and simply comparing Eqs. (2) and (3), there appears to be a connection with the coefficients of the kinetic operators. This is indeed the case which is readily established through the critical renormalization group equation and is documented, for example, in Ref. [2]. If the anomalous dimension is $\gamma(g)$, it is defined from the wave function renormalization constant through

$$\gamma(g) = \beta(g) \frac{\partial \ln Z_\phi}{\partial g}, \quad (8)$$

where $\beta(g) = \tilde{\mu} \partial g / \partial \tilde{\mu}$ is the usual β -function. Then the critical renormalization group equation gives in our conventions

$$\eta = \gamma(g_c), \quad (9)$$

where g_c is the d -dimensional non-trivial fixed point of the ϕ^4 β -function which represents the underlying phase transition of the model. The other exponents have analogous relations. For instance, in our conventions

$$\omega = -\frac{1}{2}\beta'(g_c), \quad \frac{1}{\nu} = -\beta'(\lambda_c), \quad (10)$$

where λ_c is the d -dimensional non-trivial fixed point of the $O(N)$ nonlinear σ model.

Therefore, having argued this relation to be between the usual renormalization constants and the critical exponents of the phase transition, one can provide the renormalization group functions to very high precision. The best current state is the five-loop work of Prof. Dr. Kleinert and collaborators [1]. Earlier calculations at lower orders were carried out in Refs. [12–16]. However, some initial attempts [15,16] at the five-loop calculation contained errors in several simple integrals which were observed and corrected in Ref. [1]. As a testimony to the huge calculation of Ref. [1], it is worth quoting the full five-loop $\overline{\text{MS}}$ result for the d -dimensional β -function which is

$$\begin{aligned} \beta(g) = & (d-4)\frac{g}{2} + [N+8]\frac{g^2}{6} - [3N+14]\frac{g^3}{6} \\ & + [33N^2 + 922N + 2960 + 96(5N+22)\zeta(3)]\frac{g^4}{432} \\ & + [5N^3 - 6320N^2 - 80456N - 196648 \\ & - 96(63N^2 + 764N + 2332)\zeta(3) + 288(5N+22)(N+8)\zeta(4) \\ & - 1920(2N^2 + 55N + 186)\zeta(5)]\frac{g^5}{7776} \end{aligned}$$

$$\begin{aligned}
& + [13N^4 + 12578N^3 + 808496N^2 + 6646336N + 13177344 \\
& \quad - 16(9N^4 - 1248N^3 - 67640N^2 - 552280N - 1314336)\zeta(3) \\
& \quad - 768(6N^3 + 59N^2 - 446N - 3264)\zeta^2(3) \\
& \quad - 288(63N^3 + 1388N^2 + 9532N + 21120)\zeta(4) \\
& \quad + 256(305N^3 + 7466N^2 + 66986N + 165084)\zeta(5) \\
& \quad - 9600(N + 8)(2N^2 + 55N + 186)\zeta(6) \\
& \quad + 112896(14N^2 + 189N + 526)\zeta(7)] \frac{g^6}{124416} + O(g^7), \quad (11)
\end{aligned}$$

where $\zeta(z)$ is the Riemann zeta function. The term $(d-4)g$, which corresponds to the dimension of the coupling, has been included to demonstrate the existence in d dimensions of a non-trivial value for g_c . This was used, together with the other renormalization group functions, to determine series for η , ν and ω , whose resummed three-dimensional values are in agreement with experiment and other methods [1].

3 Six Loops and Beyond

Given the need for the more accurate evaluation of critical exponents because of better experimental precision it is worth reviewing insights into the problem of tackling the extension of Eq. (11) to six loops and beyond. Two major approaches have recently been developed which attack different parts of the quintic polynomial in N which will appear as the six-loop coefficient of Eq. (11). The first of these is based on the observation that there appears to be a connection between abstract knot theory and number theory with the value of the Feynman diagrams when calculated in dimensional regularization. The initial breakthrough was by Kreimer in Ref. [17], where it was shown that the momentum routing in a Feynman graph could be associated with a knot link diagram. Skeining such link diagrams appropriately allowed one to decompose these into either a set of unknots or a set of unknots plus a prime knot. The remarkable and elegant feature which emerged was that the simple pole in ϵ of the corresponding Feynman diagram had only rational numbers in the former case but in the latter situation when a prime knot was involved, the Feynman diagram contained in addition to rationals a transcendental number such as $\zeta(3)$. In essence [17], if a Feynman graph skeined to a (prime) $(2, n)$ -torus knot then its associated pole part contained $\zeta(n)$. This beautiful connection has since been studied extensively and the higher torus

knots contain new zeta irreducible double and triple sums [18]. For instance, the next prime torus knot beyond the $(2, n)$ -set is the $(3, 4)$ -torus knot which is associated in Feynman diagrams to the double sum

$$U_{6,2} = \sum_{n>m>0}^{\infty} \frac{(-1)^{n-m}}{n^6 m^2} . \quad (12)$$

This number had previously been investigated in Refs. [19,20] where it remained a puzzle since it could not be reduced to a series of products of ordinary $\zeta(n)$'s of the same level of transcendentality. In the knot context it turns out that the braid word structure of the associated prime knot has a simple correspondence with the nested sum structure which is not reducible to lower $\zeta(n)$'s [17,18]. Whilst these revolutionary ideas were developed without reference to a particular field theory or its symmetry properties, it was not clear whether such new zeta irreducible numbers would in fact arise in the renormalization group functions of a four-dimensional theory. However, it was shown in ϕ^4 -theory for $N = 1$ in Ref. [18] that certain diagrams at six (and seven) loops with no subgraph divergences had a non-trivial knot number structure beyond the $(2, n)$ -torus knot level. Indeed $U_{6,2}$ arose in that part of the six-loop β -function polynomial which was N -independent and its coefficient was calculated explicitly. Therefore, computing higher-order corrections to Eq. (11) would have to account for this new feature. For instance, knowing that such structures will exist could allow one to exploit it as a basis for performing such calculations.

The second method of gaining insight into the form of the six-loop and higher $\overline{\text{MS}}$ β -function is to determine the coefficients of the leading and next to leading terms of the polynomial in N at each loop order. This is provided by the large N method developed originally for the $O(N)$ σ model in Refs. [10,11,21]. There the d -dimensional critical exponents themselves were computed in successive powers of $1/N$ to three terms in the series for η and ν . Since they are expressed as functions of $d = 2\mu$, one can extract through the critical renormalization group equation information on the coefficients of the corresponding renormalization group functions in $4 - 2\epsilon$ dimensions and compare it with the explicit $\overline{\text{MS}}$ perturbative results as a consequence of the critical point equivalence. To the orders each of these is computed to, there is exact agreement between both. More importantly, this connection can be used to gain information on the coefficients of the renormalization group functions going beyond those currently calculated at the orders in $1/N$ which are

available. To achieve this for four-dimensional ϕ^4 -theory, one requires knowledge of the location of the fixed point g_c in d dimensions at the appropriate order in $1/N$. This therefore requires the critical exponent ω at $O(1/N^2)$ which was calculated in Ref. [22]. The method is based on the Lagrangian of the form

$$L = \frac{1}{2} \partial_\mu \phi^a \partial^\mu \phi^a + \frac{1}{2} \sigma \phi^a \phi^a - \frac{3\sigma^2}{32\pi^2 g}. \quad (13)$$

It is used to apply the uniqueness method of integration [10,11] when computing the large set of Feynman diagrams which occur at $O(1/N^2)$. Thus, one can determine the critical exponent ω as [22]

$$\omega_1 = (2\mu - 1)^2 \eta_1 \quad (14)$$

and

$$\begin{aligned} \omega_2 = - & \left[\frac{4(\mu^2 - 5\mu + 5)(2\mu - 3)^2(\mu - 1)\mu^2[\bar{\Psi}(\mu) + \bar{\Psi}^2(\mu)]}{(\mu - 2)^3(\mu - 3)} \right. \\ & + \frac{16\mu(2\mu - 3)^2}{(\mu - 2)^3(\mu - 3)^2\eta_1} \\ & + \frac{3(4\mu^5 - 48\mu^4 + 241\mu^3 - 549\mu^2 + 566\mu - 216)(\mu - 1)\mu^2\hat{\Theta}(\mu)}{2(\mu - 2)^3(\mu - 3)} \\ & + [16\mu^{10} - 240\mu^9 + 1608\mu^8 - 6316\mu^7 + 15861\mu^6 \\ & - 25804\mu^5 + 26111\mu^4 - 14508\mu^3 + 2756\mu^2 \\ & + 672\mu - 144]/[(\mu - 2)^4(\mu - 3)^2]\bar{\Psi}(\mu) \\ & - [144\mu^{14} - 2816\mu^{13} + 24792\mu^{12} - 130032\mu^{11} + 452961\mu^{10} \\ & - 1105060\mu^9 + 1936168\mu^8 - 2447910\mu^7 + 2194071\mu^6 \\ & - 1320318\mu^5 + 460364\mu^4 - 43444\mu^3 - 26280\mu^2 \\ & \left. + 8208\mu - 864]/[2(2\mu - 3)(\mu - 1)(\mu - 2)^5(\mu - 3)^2\mu] \right] \eta_1^2, \quad (15) \end{aligned}$$

where $\omega = \sum_{i=0}^{\infty} \omega_i/N^i$ and $\omega_0 = \mu - 2$. The various variables and functions are defined by

$$\begin{aligned} \eta_1 &= \frac{2(\mu - 2)\Gamma(2\mu - 1)}{\mu\Gamma(1 - \mu)\Gamma^3(\mu)}, \\ \bar{\Psi}(\mu) &= \psi(2\mu - 3) + \psi(3 - \mu) - \psi(\mu - 1) - \psi(1), \\ \hat{\Theta}(\mu) &= \psi'(\mu - 1) - \psi'(1), \end{aligned}$$

$$\bar{\Phi}(\mu) = \psi'(2\mu - 3) - \psi'(3 - \mu) - \psi'(\mu - 1) + \psi'(1), \quad (16)$$

where $\psi(x) = \frac{d}{dx} \ln \Gamma(x)$ and $\Gamma(x)$ is the Euler Γ -function. So, for example, if we represent all orders $\overline{\text{MS}}$ ϕ^4 β -function at $O(1/N^2)$ as

$$\beta(g) = \frac{1}{2}(d-4)g + (a_1N + b_1)g^2 + \sum_{r=2}^{\infty} (a_rN + b_r)N^{r-2}g^{r+1}, \quad (17)$$

then we find the *new* $\overline{\text{MS}}$ coefficients

$$\begin{aligned} a_6 &= [29 + 528\zeta(3) - 432\zeta(4)]/1866240, \\ a_7 &= [61 + 80\zeta(3) + 1584\zeta(4) - 1728\zeta(5)]/26873856, \\ a_8 &= -[5760\zeta(6) - 6336\zeta(5) - 240\zeta(4) \\ &\quad + 1152\zeta^2(3) - 208\zeta(3) - 125]/376233984, \\ b_6 &= -[28160\zeta(7) - 95200\zeta(6) + 150336\zeta(5) + 6912\zeta(4)\zeta(3) \\ &\quad - 14112\zeta(4) - 24064\zeta^2(3) - 11880\zeta(3) - 5661]/466560, \\ b_7 &= -[8520960\zeta(8) - 32724480\zeta(7) + 43286400\zeta(6) + 3993600\zeta(5)\zeta(3) \\ &\quad - 31998720\zeta(5) - 8663040\zeta(4)\zeta(3) - 2538432\zeta(4) \\ &\quad + 11381760\zeta^2(3) + 7461168\zeta(3) + 1125439]/403107840, \\ b_8 &= -[9210880\zeta(9) - 41166720\zeta(8) + 61054080\zeta(7) + 4300800\zeta(6)\zeta(3) \\ &\quad - 38500800\zeta(6) + 2995200\zeta(5)\zeta(4) - 19553280\zeta(5)\zeta(3) \\ &\quad + 11519040\zeta(5) + 17072640\zeta(4)\zeta(3) + 5863104\zeta(4) \\ &\quad + 542720\zeta^3(3) - 10141440\zeta^2(3) \\ &\quad - 4518336\zeta(3) - 717083]/1410877440. \end{aligned} \quad (18)$$

Similarly, if one represents the $\overline{\text{MS}}$ field anomalous dimension at $O(1/N^3)$ as

$$\gamma(g) = \sum_{r=1}^{\infty} (c_rN^2 + d_rN + e_r)N^{r-2}g^{r+1}, \quad (19)$$

where $e_1 \equiv 0$, the large N results give

$$\begin{aligned} e_9 &= [1560674304\zeta(10) - 12534896640\zeta(9) + 11070010560\zeta(8) \\ &\quad + 1732018176\zeta(7)\zeta(3) + 581961984\zeta(7) - 3411394560\zeta(6)\zeta(3) \\ &\quad - 2684240640\zeta(6) + 209534976\zeta^2(5) - 1567752192\zeta(5)\zeta(4) \\ &\quad + 1754664960\zeta(5)\zeta(3) - 975533568\zeta(5) - 9289728\zeta(4)\zeta^2(3) \\ &\quad + 1310201856\zeta(4)\zeta(3) + 1636615872\zeta(4) - 137158656\zeta^3(3) \end{aligned}$$

$$\begin{aligned}
& - 1708996608\zeta^2(3) + 294403968\zeta(3) \\
& - 89800704U_{62} - 341350433]/1950396973056 .
\end{aligned} \tag{20}$$

The previous term in the series at this order, e_8 , was given in Ref. [23] and like the expression for e_9 , it contains the zeta irreducible $U_{6,2}$. However, it is important to recognise that the first appearance of this number in the *full* anomalous dimension will be at a lower-loop order than the ninth order of e_8 .

4 Discussion

Whilst the knot theory insight into the higher-order structure of the $\overline{\text{MS}}$ ϕ^4 four-dimensional renormalization group functions is quite impressive, true progress in this area will only be represented by the provision of the *full* result at six loops. This would require a huge amount of tedious computation since, for instance, one needs to determine the *finite* part of the large number of five-loop diagrams as they will contribute when multiplied by the one-loop vertex counterterms. Therefore, we believe such a result will not appear in the foreseeable future and hence Eq. (11) remains the current state of the art. Nevertheless, to emphasise Prof. Dr. Kleinert's continued interest and impressive contribution to this field, it is worth mentioning an extension of the calculation of Eq. (11) to a model which involves, in addition to an $O(N)$ ϕ^4 -interaction, a cubic interaction. The critical exponents for this double coupling model were computed again to five loops in $\overline{\text{MS}}$ in Ref. [24] to explore in detail the stability of a variety of fixed points which occur in this model since they correspond to phase transitions in three-dimensional cubic crystals.

References

- [1] H. Kleinert, J. Neu, V. Schulte-Frohlinde, K.G. Chetyrkin, and S.A. Larin, *Phys. Lett. B* **272**, 39 (1991); *ibid.* **319**, 545(E) (1993).
- [2] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 1st ed. (Clarendon Press, Oxford, 1989).
- [3] J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev. B* **21**, 3976 (1980).
- [4] G.A. Baker Jr., B.G. Nickel, and D.I. Meiron, *Phys. Rev. B* **17**, 1365 (1978).
- [5] S.A. Antonenko and A.I. Sokolov, *Phys. Rev. E* **51**, 1894 (1995).
- [6] H. Kleinert, *Phys. Rev. D* **60**, 085001 (1999).
- [7] H. Kleinert, *Phys. Lett. B* **434**, 74 (1998).

- [8] H. Kleinert, *Phys. Lett. B* **463**, 69 (1999).
- [9] H. Kleinert and V. Schulte-Frohlinde, eprint: cond-mat/9907214.
- [10] A.N. Vasil'ev, Yu.M. Pis'mak, and J.R. Honkonen, *Theor. Math. Phys.* **46**, 157 (1981).
- [11] A.N. Vasil'ev, Yu.M. Pis'mak, and J.R. Honkonen, *Theor. Math. Phys.* **47**, 465 (1981).
- [12] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, and B.G. Nickel, *Phys. Lett. A* **44**, 227 (1973).
- [13] A.A. Vladimirov, D.I. Kazakov, and O.V. Tarasov, *Sov. Phys. JETP* **50**, 521 (1979).
- [14] F.M. Dittes, Yu.A. Kubyshin, and O.V. Tarasov, *Theor. Math. Phys.* **37**, 879 (1978).
- [15] K.G. Chetyrkin, A.L. Kataev, and F.V. Tkachov, *Phys. Lett. B* **99**, 147 (1981); *ibid.* **101**, 457(E) (1981).
- [16] K.G. Chetyrkin, S.G. Gorishny, S.A. Larin, and F.V. Tkachov, *Phys. Lett. B* **132**, 351 (1983).
- [17] D. Kreimer, *Phys. Lett. B* **354**, 117 (1995).
- [18] D.J. Broadhurst and D. Kreimer, *Int. J. Mod. Phys. C* **6**, 519 (1995).
- [19] D.J. Broadhurst, *Z. Phys. C* **32**, 249 (1986).
- [20] D.J. Broadhurst, *Z. Phys. C* **41**, 81 (1988).
- [21] A.N. Vasil'ev, Yu.M. Pis'mak, and J.R. Honkonen, *Theor. Math. Phys.* **50**, 127 (1982).
- [22] D.J. Broadhurst, J.A. Gracey, and D. Kreimer, *Z. Phys. C* **75**, 559 (1997).
- [23] J.A. Gracey, *Nucl. Instrum. Methods A* **389**, 361 (1997).
- [24] H. Kleinert and V. Schulte-Frohlinde, *Phys. Lett. B* **342**, 284 (1995).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

PHASE ORDERING DYNAMICS OF ϕ^4 -THEORY WITH HAMILTONIAN EQUATIONS OF MOTION

B. ZHENG¹, V. LINKE², AND S. TRIMPER¹

¹ *Fachbereich Physik, Universität Halle, 06099 Halle, Germany*

² *Fachbereich Physik, Freie Universität Berlin, 14195 Berlin, Germany*

*E-mails: zheng@hera.physik.uni-halle.de, linke@physik.fu-berlin.de,
trimper@hera.physik.uni-halle.de*

Phase ordering dynamics of the $(2+1)$ - and $(3+1)$ -dimensional ϕ^4 -theory with Hamiltonian equations of motion is investigated numerically. Dynamic scaling is confirmed. The dynamic exponent z is different from that of the Ising model with dynamics of model A, while the exponent λ is the same.

1 Introduction

It is believed that macroscopic properties of many particle systems could, in principle, be described by microscopic deterministic equations of motion (e.g. Newton, Hamiltonian, and Heisenberg equations), if all interactions, boundary conditions, and initial states could be taken into account. In practice, however, it is very difficult to solve these equations, except for some simple cases. Therefore, statistical mechanics was developed to deal effectively with such systems. Usually, ensemble theories are appropriate for equilibrium states, but they are inadequate for non-equilibrium states, where a general theory does not exist. In many cases, stochastic dynamics, e.g. following from Langevin-type equations of motion or Monte Carlo dynamics, are approximate theories. Anyway, it is an open question whether microscopic equations of motion could really produce the results of statistical mechanics, or vice versa (see Refs. [1–5]).

With the development of computers, it becomes gradually more and more possible to solve microscopic deterministic equations *numerically*. This at-

tracts scientists of different fields. The study of microscopic fundamental dynamics aims on the one hand to test statistical mechanics, and on the other hand to explore new physics. For example, if a system is isolated, there is only internal interaction, and periodic boundary conditions can be adopted. So, computations are greatly simplified. To achieve ergodicity, the system should start from random initial states. Recently, such effort has been made for the $O(N)$ vector model and XY model [5–7]. The results support that deterministic Hamiltonian equations correctly describe second-order phase transitions. The estimated static critical exponents are consistent with those calculated from canonical ensembles (e.g. see the recent text book [8]). More interestingly, the macroscopic short-time (non-equilibrium) dynamic behavior of the $(2+1)$ -dimensional ϕ^4 -theory at *criticality* has also been investigated and dynamic scaling is found [9,10]. The results indicate that Hamiltonian dynamics in two dimensions with random initial states is in the same universality class of Monte Carlo dynamics of model A.

In a similar spirit, phase ordering dynamics of the $(2+1)$ -dimensional ϕ^4 -theory with Hamiltonian equations of motion has been investigated in Ref. [11]. Assuming random initial states, there is a minimum energy density which is above the real minimum energy density of the system. Starting from this minimum energy density (note that energy is conserved), which is well below the critical energy density, phase ordering occurs. Dynamic scaling behavior is found. The dynamic exponent z is different from that of model A dynamics, but the exponent λ governing the power-law decay of the autocorrelation looks the same. It is somewhat interesting that the scaling function of the equal-time spatial correlation function is the same as that of the Ising model with model A dynamics. All results are independent of the parameters in the system.

The purpose of this article is twofold: Firstly, we generalize the computations to $(3+1)$ dimensions, which is important because our realistic world is in $(3+1)$ dimensions. Furthermore, in phase ordering of model A dynamics, the dynamic exponent z is dimension-independent but the exponent λ is dimension-dependent. It is interesting to see whether this property is kept in Hamiltonian dynamics. Attention will also be put on whether the scaling function of the equal-time spatial correlation function in three dimensions is the same as the one of the Ising model with model A dynamics. Secondly, to achieve more confidence on our conclusions, we will reexamine the results for $(2+1)$ dimensions obtained in Ref. [11], using somewhat different, more careful approaches. Since the computations in $(3+1)$ dimensions are very

time consuming, more accurate data are obtained in $(2 + 1)$ dimensions.

2 Phase Ordering Dynamics

In the following, we outline phase ordering dynamics with Hamiltonian equations of motion. For a recent review of general ordering dynamics, readers are referred to Ref. [12].

2.1 The Model

For an isolated system, the Hamiltonian of the $(d + 1)$ -dimensional ϕ^4 -theory on a square or cubic lattice is

$$H = \sum_i \left[\frac{1}{2} \pi_i^2 + \frac{1}{2} \sum_{\mu} (\phi_{i+\mu} - \phi_i)^2 - \frac{1}{2} m^2 \phi_i^2 + \frac{1}{4!} g \phi_i^4 \right] \quad (1)$$

with $\pi_i = \dot{\phi}_i$. It leads to the equations of motion

$$\ddot{\phi}_i = \sum_{\mu} (\phi_{i+\mu} + \phi_{i-\mu} - 2\phi_i) + m^2 \phi_i - \frac{1}{3!} g \phi_i^3. \quad (2)$$

Here μ represents spatial directions and energy is conserved in these equations. The solutions are supposed to generate a microcanonical ensemble. The temperature could be defined as the averaged kinetic energy. For the *non-equilibrium* dynamic system, however, total energy is a more convenient controlling parameter, since it is conserved and can be taken as an input from initial states. For given parameters m^2 and g , there exists a critical energy density ϵ_c , separating the ordered phase (below ϵ_c) and the disordered phase (above ϵ_c). The phase transition is of second order.

We should emphasize that a Langevin equation at zero temperature is also "deterministic" in the sense that there is no noise. But it is essentially different from the Hamiltonian equations (2). The former describes relaxation towards equilibrium at zero temperature for a non-isolated system, but the latter contains full physics at all temperatures for an isolated system.

The order parameter of the ϕ^4 -theory is the magnetization. The time-dependent magnetization $M \equiv M^{(1)}(t)$ and its second moment $M^{(2)}$ are defined as

$$M^{(k)}(t) = \frac{1}{L^{dk}} \left\langle \left[\sum_i \phi_i(t) \right]^k \right\rangle, \quad k = 1, 2. \quad (3)$$

L is the lattice size and d is the spatial dimension. Here, the average is *over initial configurations*, which means that it is a real sample average and different from the time average in equilibrium. Following ordering dynamics with stochastic equations [12], we consider the dynamic process that the system, initially in a *disordered* state but with an energy density well below ϵ_c , is suddenly released to evolve according to Eq. (2). For simplicity, we set the initial kinetic energy to zero, i.e. $\dot{\phi}_i(0) = 0$. To generate a random initial configuration $\{\phi_i(0)\}$, we first fix the magnitude $|\phi_i(0)| \equiv c$, then randomly give the sign to $\phi_i(0)$ with the restriction of a fixed magnetization in units of c , and finally the constant c is determined by the given energy. In case of stochastic dynamics, scaling behavior of phase ordering is dominated by the fixed point $(T_I, T_F) = (\infty, 0)$ with T_I being the initial temperature and T_F being the temperature after quenching [12]. In Hamiltonian dynamics, the energy density cannot be taken to the real minimum $e_{\min} = -3m^4/2g$ since the system does not move. Actually, for the initial states described above, the energy is given by

$$V = \sum_i \left[(d - \frac{1}{2}m^2)\phi_i^2 + \frac{1}{4!}g\phi_i^4 \right]. \quad (4)$$

For the case of $d < m^2/2$ phase ordering occurs when the initial magnetization is set to zero for an energy density well below the critical point ϵ_c , due to the competition of two ordered states [11]. The scaling behavior is dominated by the minimum energy density $v_{\min} = V_{\min}/L^d$, which is a kind of fixed point. Above $v_{\min}$, there are extra corrections to scaling. From now, we redefine the energy density $e_{\min}$ as zero. Then the fixed point is $\epsilon_0 = v_{\min} - e_{\min}$. In this article, we consider only the energy density at exactly the fixed point ϵ_0 .

2.2 Dynamic Scaling Behavior

Let us first consider the case of the initial magnetization $m_0 = 0$. An important observable is the equal-time correlation function

$$C(r, t) = \frac{1}{L^d} \left\langle \sum_i \phi_i(t) \phi_{i+r}(t) \right\rangle. \quad (5)$$

The scaling hypothesis is that, at the *late* stage of the time evolution, $C(r, t)$ obeys the scaling form

$$C(r, t) = f(r/t^{1/z}), \quad (6)$$

where z is the so-called dynamic exponent. Here, "late" is meant in a microscopic sense. In other words, when the domain size ($\sim t^{1/z}$) is big enough in units of the lattice spacing, scaling behavior emerges. At finite t , of course, there may be corrections to scaling which are generally not universal. They may induce difficulties for observing scaling behavior and uncertainties in the determination of the critical exponents.

Simple understanding of the scaling behavior of $C(r, t)$ can be achieved from the second moment of the magnetization. Integrating over r in Eq. (6), we obtain the power law behavior

$$M^{(2)}(t) \sim t^{d/z} . \quad (7)$$

Another interesting observable is the auto-correlation function

$$A(t) = \frac{1}{L^d} \left\langle \sum_i \phi_i(0) \phi_i(t) \right\rangle . \quad (8)$$

The scaling hypothesis leads to the power law

$$A(t) \sim t^{-\lambda/z} , \quad (9)$$

which implies that ordering dynamics is in some sense "critical". Here λ is another independent exponent.

For the discussions above, the initial magnetization m_0 is set to zero. If m_0 is non-zero, the system reaches a unique ordered state within a finite time. If m_0 is infinitesimally small, however, the time for reaching the ordered state is also infinite and scaling behavior can still be expected, at least at relatively early times (in a macroscopic sense). In this case, an interesting observable is the magnetization itself. It increases by a power law

$$M(t) \sim t^\theta, \quad \theta = (d - \lambda)/z. \quad (10)$$

The exponent θ can be written as x_0/z , with x_0 being the scaling dimension of m_0 . This power-law behavior has been investigated in critical dynamics [13,14]. The interesting point here is that θ is related to the exponent λ which governs the power-law decay of the auto-correlation. By combining measurements of θ and λ , one can also estimate the dynamic exponent z .

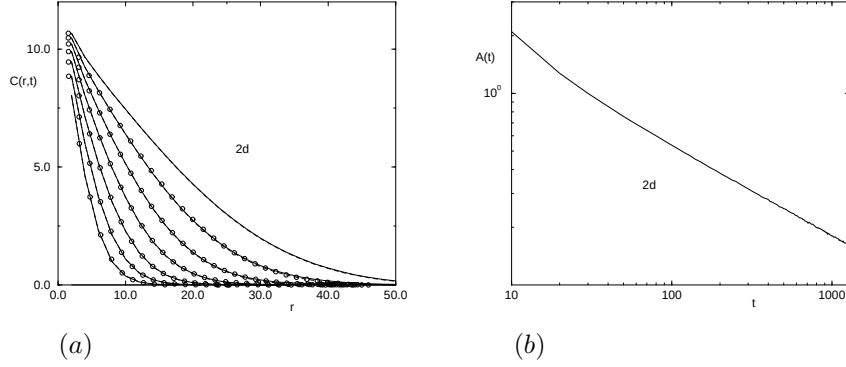


Figure 1. (a) $C(r, t)$ in two dimensions obtained with $L = 256$ and $\Delta t = 0.01$ is plotted with solid lines for $t = 20, 40, 80, 160, 320, 640$, and 1280 (from left). Circles fitted to the curve at the time t indicate the data at the time $2t$ with r being rescaled by a factor $2^{-1/z}$. $C(r, t)$ obtained with $\Delta t = 0.02$ is also plotted with dashed lines. These overlap almost completely with the solid lines. (b) $A(t)$ obtained with $\Delta t = 0.01$ (the solid line) on a log-log scale. The curve for $\Delta t = 0.02$ overlaps completely with that for $\Delta t = 0.01$.

3 Numerical Results

To solve the equations of motion (2) numerically, we discretize $\ddot{\phi}_i$ by $[\phi_i(t + \Delta t) + \phi_i(t - \Delta t) - 2\phi_i(t)]/(\Delta t)^2$. Starting from an initial configuration, we update the equations of motion up to a certain maximum time $t_{\max}$. Then we repeat the procedure with other initial configurations. Reasonable results in two dimensions are obtained mainly with $\Delta t = 0.05$ up to $t_{\max} = 640$ and a lattice size $L = 521$ [11]. 200 samples of initial configurations are used for averaging. For three dimensions, we also perform the computations with $\Delta t = 0.05$ up to $t_{\max} = 640$ but with a lattice size $L = 125$. Here, 50 samples are taken for averaging. We have also carried out some computations with other Δt 's and lattice sizes to confirm the results. At the time $t_{\max}$, the equal-time correlation function $C(r, t)$ decays to nearly zero at $r \sim 45$ and this indicates that the finite-size effect with $L = 128$ is already small. Furthermore, in order to gain more confidence in our conclusions, especially whether our $t_{\max}$ has really reached the scaling regime, we perform more accurate computations in two dimensions (compared with those in Ref. [11]) with a lattice size $L = 256$ and $\Delta t = 0.02, 0.01$ up to $t_{\max} = 1280$. The number of samples for averaging is 600. In this article, somewhat different and more careful approaches will be adopted.

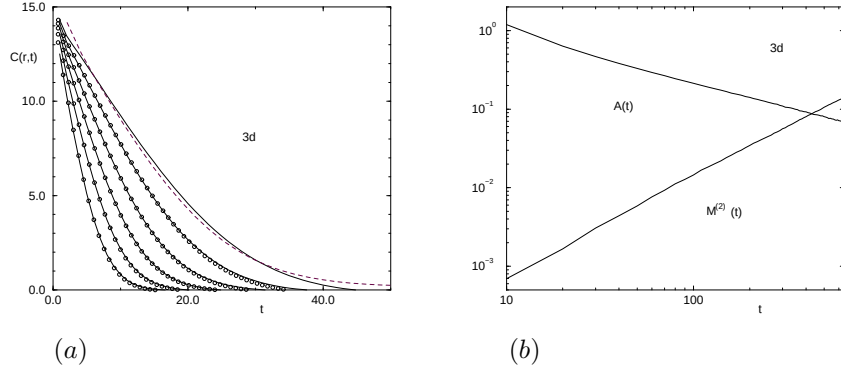


Figure 2. (a) $C(r, t)$ in three dimensions obtained with $L = 128$ and $\Delta t = 0.05$ is plotted with solid lines for $t = 20, 40, 80, 160, 320$, and 640 (from left). Circles fitted to the curve at the time t display the data at the time $2t$ with r being rescaled by a factor $2^{-1/z}$. The dashed line represents the scaling function in two dimensions. (b) $A(t)$ and $M^{(2)}(t)$ on a log-log scale.

In Fig. 1(a), the equal-time correlation function $C(r, t)$ in two dimensions is displayed. Solid lines are obtained with $\Delta t = 0.01$ and, from left to right, the time t is 20, 40, 80, 160, 320, 640, and 1280. Data for $\Delta t = 0.02$ are also plotted with dashed lines, but they almost completely overlap with the solid lines. For the curve of $t = 1280$, $C(r, t)$ decays to nearly zero at $r \sim 50$. Therefore, we conclude that the finite-size effect for the lattice size $L = 256$ should already be negligibly small. To confirm this, we have also compared the data with those in Ref. [11]. Our data also show that the finite Δt effect for $\Delta t = 0.05$ is negligible, too. According to the scaling form (6), from data collapse of $C(r, t)$ at different t 's, one can estimate the dynamic exponent z . As observed in Ref. [11], the effective dynamic exponent $z(t)$ shows a small dependence on the time t . To explore this behavior and extract confidently the value of z , we perform scaling collapse of $C(r, t)$ with the time t and $2t$. In Fig. 1(a), circles fitted to a solid line of the time t are the data of the time $2t$ with r being rescaled by a factor $2^{-1/z}$, i.e. $C(r, t) = C(r2^{1/z}, 2t)$. The dynamic exponent $z(t)$ is determined by the best fitting of the circles to the corresponding solid line. We see clearly that the data collapse nicely.

Figure 2(a) shows $C(r, t)$ in three dimensions. Scaling collapse is also observed, even though for larger r it is not as good as in two dimensions. This can be neither a finite-size effect nor a finite Δt effect, since it exists also for small t 's. To see the trend of $z(t)$ as the time t evolves, we plot in

Fig. 3(a) the effective exponent $z(t)$ against $1/t$. For two dimensions, $z(t)$ starting from a value around 3 gradually *decreases* and reaches 2.63(2) at $t = 640$ (i.e. obtained with data of $C(r, t)$ at the time $t = 640$ and $2t = 1280$). Assuming that the behavior of $z(t)$ will not change essentially after $t = 1280$, the extrapolated value of z to the infinite time t is estimated to be 2.6(1). Interestingly, for three dimensions the exponent $z(t)$ starting from a value around 2.5 *increases* slowly, but stabilizes at 2.7 after $t = 80$. A good estimate of z is $z = 2.7(1)$. Within statistical errors, the values of the dynamic exponent z in two and three dimensions coincide with each other, thus indicating that the dynamic exponent z is dimension-independent. This can also be seen from the joining of two different curves at relatively large times in Fig. 3(a).

In the case of the Ising model with Monte Carlo dynamics, the effective exponent $z(t)$ in two dimensions converges to $z = 2$ rather fast (see Ref. [15]), but it is relatively slow in three dimensions due to corrections to scaling. It might be somewhat general that phase ordering dynamics in three dimensions is somewhat more complicated than in two dimensions.

An interesting fact is that even though the dynamic exponent z of the ϕ^4 -theory in two dimensions with Hamiltonian dynamics is different from that of the Ising model with Monte Carlo dynamics, the scaling function $f(x)$ in Eq. (6) is the same [11]. However, this is probably only by chance since it is *not* the case in three dimensions. The scaling function $f(x)$ of the three-dimensional ϕ^4 -theory with Hamiltonian dynamics is different not only from that of the two-dimensional but also from that of the three-dimensional Ising model with Monte Carlo dynamics. The dashed line in Fig. 2(a) shows the $f(x)$ of the two-dimensional ϕ^4 -theory. In general, Hamiltonian dynamics for isolated systems differs indeed from stochastic dynamics for non-isolated systems.

For a simple understanding of the correlation function $C(r, t)$, one can measure the time-dependent second moment $M^{(2)}(t)$. The scaling form results in a power-law behavior for $M^{(2)}(t)$ and from the slope in log-log scale one can estimate the corresponding exponent. Such an approach is rather typical and useful in critical dynamics [14]. It can also be applied in ordering dynamics, but this is less efficient. In critical dynamics, in the scaling collapse of $C(r, t)$, one has to determine two exponents, the dynamic exponent z and the static exponent $2\beta/\nu$. Therefore it is efficient to read out directly the exponent $(d - 2\beta/\nu)/z$ from the slope of $M^{(2)}(t)$ in log-log scale [14].

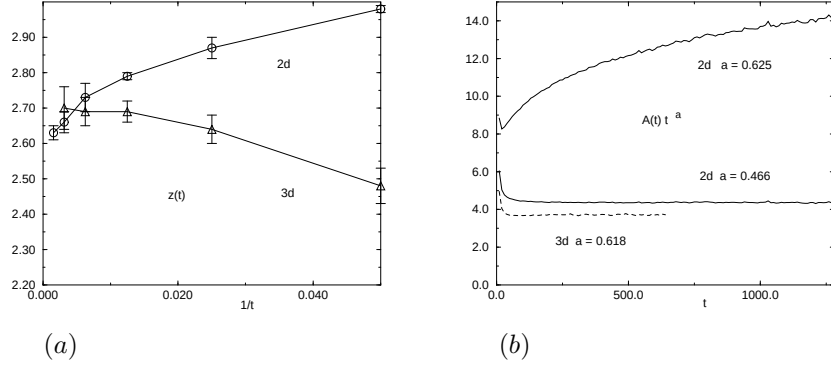


Figure 3. (a) The effective dynamic exponent $z(t)$ measured from scaling collapse of $C(r, t)$ with the times t and $2t$. (b) Taking $a = \lambda/z$, $A(t)t^a$ tends to a constant.

However, in ordering dynamics the “static” exponent $2\beta/\nu = 0$ and the scaling collapse of $C(r, t)$ is only a one parameter fit. Measurements of $M^{(2)}(t)$ do not show any advantage since it is not self-averaged and there is a larger fluctuation for bigger lattices (see the data in Ref. [11]). Anyway, in Fig. 2(b) we have plotted the second moment in log-log scale for the three-dimensional ϕ^4 -theory. The power-law behavior is seen after $t \sim 80$ and this is consistent with Fig. 3(a). According to Eq. (7), the resulting dynamic exponent is $z = 2.5(2)$, consistent within errors with $z = 2.7(1)$ measured from $C(r, t)$.

Another interesting exponent in ordering dynamics is λ governing the power-law decay of the auto-correlation $A(t)$ in Eq. (9). The measurements of the auto-correlation in ordering dynamics is easier than in critical dynamics since the fluctuation is much smaller. The results for the ϕ^4 -theory in two and three dimensions are shown in Fig. 1(b) and 2(b). In order to see how the effective exponent λ/z depends on the time t , we have measured the slope of the curves in the time interval $[t, 2t]$. The results are given in Table 1. For both, two and three dimensions, the exponent λ/z becomes stable after $t = 160$. The final values are $\lambda/z = 0.466(3)$ and $0.618(4)$ for two and three dimensions, respectively. To show that our estimates of λ/z are indeed reasonable, we plot $A(t)t^a$ in Fig. 3(b) as a function of the time t . A correct value $a = \lambda/z$ should result in a constant for $A(t)t^a$, at least for larger times. Such a behavior is nicely seen from the lower solid line and the dashed line for two and three dimensions. To confirm that the value $\lambda/z = 0.466(3)$ for two dimensions is really different from $\lambda/z = 0.625$ for stochastic dynamics, the corresponding curve with $a = 0.625$ is also displayed there (the upper

Table 1. The exponent λ/z measured in the time interval $[t, 2t]$ from the auto-correlations in two and three dimensions.

t	40	80	160	320	640
$2D$	0.508(1)	0.492(1)	0.469(7)	0.461(6)	0.463(6)
$3D$	0.633(4)	0.609(1)	0.617(3)	0.619(7)	

solid line). Obviously, it does not tend to a constant.

From measurements of z (from $C(r, t)$) and λ/z , we estimate the exponent $\lambda = 1.21(5)$ and $1.67(6)$ for two and three dimensions, respectively. For stochastic dynamics, the theoretical prediction for two dimensions is $\lambda = 1.25$ [12,16], but in Monte Carlo simulations it is usually slightly smaller [15]. Extrapolation is needed to obtain a value very close to 1.25. There is always some uncertainty in extrapolation. Therefore, we tend to claim that λ of the ϕ^4 -theory in two dimensions with Hamiltonian dynamics is the same as that of stochastic dynamics. In three dimensions, our $\lambda = 1.67(6)$ agrees very well with the "best" theoretical prediction 1.67 for stochastic dynamics [12,17]. Numerical measurements of λ for stochastic dynamics in three dimensions look somewhat problematic and the results fluctuate around the theoretical values.

To complete our investigation, we have also simulated the initial increase of the magnetization in Eq. (10). Since the exponent θ is relatively big, compared with that in critical dynamics [11,14], we need to prepare a very small initial magnetization m_0 .

In Fig. 4, the magnetization in three dimensions is plotted on a log-log scale for $m_0 = 0.00123, 0.00245$, and 0.00491 (from below), respectively. The power-law behavior is observed after $t \sim 50$. From the slope, we measure the exponent θ . Within statistical errors, we cannot find any m_0 dependence of θ . The value of θ is estimated to be $0.55(2)$. With θ in hand, combining $\lambda/z = 0.618(4)$, we obtain another value for the dynamic exponent, $z = 2.6(1)$.

In Table 2, all exponents measured for the ϕ^4 -theory with Hamiltonian dynamics are summarized. Results for two dimensions are taken from Ref. [11], but λ/z , λ and z from $C(r, t)$ are slightly modified. Different measurements in two and three dimensions suggest that $z = 2.6(1)$ is a good estimate for the dynamic exponent. Since the critical exponent θ in phase ordering dynamics is different from the case of critical dynamics [14], it has not yet got enough attention, even though it has been mentioned in Ref. [12]. One reason might

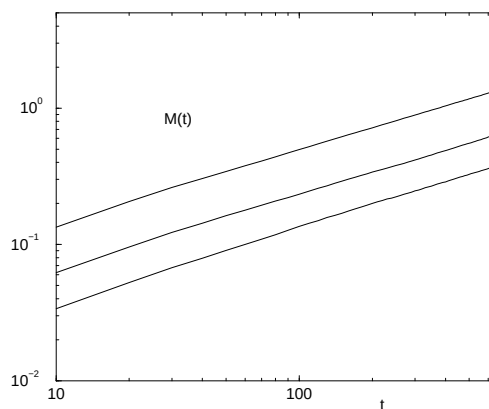


Figure 4. The magnetization in three dimensions on a log-log scale. The lattice size is $L = 128$. From below, $m_0 = 0.00123$, 0.00245 , and 0.00491 .

be that, in ordering dynamics!, increasing of the magnetization is expected if a non-zero initial value m_0 is set, but in critical dynamics, this is anomalous. Anyway, we think θ is interesting since it gives another independent estimate for the dynamic exponent z or λ .

4 Conclusions

We have numerically solved the Hamiltonian equations of motion for the two- and three-dimensional ϕ^4 -theory with random initial states. Phase ordering dynamics is carefully investigated. Scaling behavior is confirmed. The dynamic exponent z is dimension-independent. Different measurements yield a value $z = 2.6(1)$ which differs from $z = 2$ for stochastic dynamics of model A. The scaling function for the equal-time spatial correlation function is dimension-dependent, and in general it is also different from that of stochastic dynamics of model A (this is the same probably only by chance in two dimensions). However, the exponent λ of Hamiltonian dynamics is the same as that of stochastic dynamics of model A.

Acknowledgments

The present work is supported in part by DFG, TR 300/3-1, and by DFG, GRK 271.

Table 2. Exponents of the ϕ^4 -theory with Hamiltonian dynamics. To calculate λ , z measured from $C(r, t)$ is taken as input.

			z			λ
	θ	λ/z	$d/(\lambda/z + \theta)$	$C(r, t)$	$M^{(2)}$	
2D	0.31(1)	0.466(3)	2.6(1)	2.6(1)	2.6(1)	1.21(5)
3D	0.55(2)	0.618(4)	2.6(1)	2.7(1)	2.5(2)	1.67(6)

References

- [1] E. Fermi, J. Pasta, and S. Ulam, in *Collected Papers of Enrico Fermi*, Ed. E. Segré (Univ. Chicago, Chicago, 1965).
- [2] J. Ford, *Phys. Rep.* **213**, 271 (1992).
- [3] D. Escande, H. Kantz, R. Livi, and S. Ruffo, *J. Statist. Phys.* **76**, 605 (1994).
- [4] M. Antoni and S. Ruffo, *Phys. Rev. E* **52**, 2361 (1995).
- [5] L. Caiani, L. Casetti, and M. Pettini, *J. Phys. A* **31**, 3357 (1998).
- [6] L. Caiani, L. Casetti, C. Clementi, G. Pettini, M. Pettini, and R. Gatto, *Phys. Rev. E* **57**, 3886 (1998).
- [7] X. Leoncini and A.D. Verga, *Phys. Rev. E* **57**, 6377 (1998).
- [8] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [9] B. Zheng, M. Schulz, and S. Trimper, *Phys. Rev. Lett.* **82**, 1891 (1999).
- [10] B. Zheng, *Mod. Phys. Lett. B* **13**, 631 (1999).
- [11] B. Zheng, *Phys. Rev. E* **61**, 153 (2000).
- [12] A.J. Bray, *Adv. Phys.* **43**, 357 (1994), and references therein.
- [13] H.K. Janssen, B. Schaub, and B. Schmittmann, *Z. Phys. B* **73**, 539 (1989).
- [14] B. Zheng, *Int. J. Mod. Phys. B* **12**, 1419 (1998).
- [15] K. Humayun and A.J. Bray, *J. Phys. A* **24**, 1915 (1991).
- [16] D.S. Fisher and D.A. Huse, *Phys. Rev. B* **38**, 373 (1988).
- [17] F. Liu and G.F. Mazenko, *Phys. Rev. B* **44**, 9185 (1991).

PHASE TRANSITION IN THE RANDOM ANISOTROPY MODEL

M. DUDKA

*Institute for Condensed Matter Physics, Ukrainian Acad. Sci.,
UA-79011 Lviv, Ukraine
E-mail: maxdudka@icmp.lviv.ua*

R. FOLK

*Institut für Theoretische Physik, Johannes-Kepler-Universität Linz,
A-4040 Linz, Austria
E-mail: folk@tphys.uni-linz.ac.at*

YU. HOLOVATCH

*Institute for Condensed Matter Physics, Ukrainian Acad. Sci., UA-79011 Lviv and
Ivan Franko National University of Lviv, UA-79005 Lviv, Ukraine
E-mail: hol@icmp.lviv.ua*

The influence of a local anisotropy of random orientation on a ferromagnetic phase transition is studied for two cases of anisotropy axis distribution. To this end a model of a random anisotropy magnet is analyzed by means of the field theoretical renormalization group approach in two-loop approximation refined by a resummation of the asymptotic series. The one-loop result of Aharony indicating the absence of a second-order phase transition for an isotropic distribution of random anisotropy axis at space dimension $d < 4$ is corroborated. For a cubic distribution the accessible stable fixed point leads to disordered Ising-like critical exponents.

1 Introduction

Modern understanding of universal properties of matter in the vicinity of critical points is mainly due to the application of renormalization group (RG) ideas [1]. Applied to the problems of condensed matter physics in the early

1970s, the RG technique proved to be a powerful tool to study critical phenomena. For example, expressions for critical exponents governing the magnetic phase transition in regular systems are known by now with record accuracy both for isotropic [2] [$O(m)$ symmetrical] and cubic [3] magnets. The RG approach also sheds light on the influence of structural disorder on ferromagnetism. In the present article we will apply the field theoretical RG approach to study peculiarities of magnetic behavior influenced by disorder in a form of random anisotropy axis [4]. It is a special pleasure for us to dedicate this paper to Prof. Hagen Kleinert on the occasion of his 60th anniversary. His contribution to the field is hard to be overestimated.

Although an influence of a weak quenched structural disorder on universal properties of a ferromagnetic phase transition has already been a problem of intensive study for several decades, there remains a number of unsettled questions. Here, one should distinguish between random site, random field and random anisotropy magnet. A weak quenched disorder preserves the second-order phase transition in three-dimensional ($d = 3$) random site magnets [5] but can destroy this transition in random field systems [6] for $d < 4$. The situation for the random anisotropy magnets is not so clear.

Typical examples of random-anisotropy magnets are amorphous rare-earth - transition metal alloys. Some of these systems order magnetically and for the description of the ordered structure it has been proposed [4] to consider a regular lattice of magnetic ions, each of them being subjected to a local anisotropy of random orientation. The Hamiltonian of this random anisotropy model (RAM) reads [4]

$$\mathcal{H} = - \sum_{\mathbf{R}, \mathbf{R}'} J_{\mathbf{R}, \mathbf{R}'} \vec{S}_{\mathbf{R}} \vec{S}_{\mathbf{R}'} - D_0 \sum_{\mathbf{R}} (\hat{x}_{\mathbf{R}} \vec{S}_{\mathbf{R}})^2, \quad (1)$$

where $\vec{S}_{\mathbf{R}}$ is an m -component vector on a lattice site $\mathbf{R}$, $J_{\mathbf{R}, \mathbf{R}'}$ is an exchange interaction, D_0 is an anisotropy strength, and $\hat{x}_{\mathbf{R}}$ is a unit vector pointing in the local (quenched) random direction of an uniaxial anisotropy.

The model has been investigated by a variety of techniques including mean-field theory [7], computer simulations [8], $1/m$ -expansion [9], renormalization group ε -expansion [10–12]. The limit case of an infinite anisotropy has been subject to a detailed study as well [13,14]. However the nature of the low-temperature phase in RAM is not completely clear up to now, although several low-temperature phases were discussed like ferromagnetic ordering [7,8], spin-glass phase [8,9], and quasi long-range ordering [15].

The nature of ordering is connected with the distribution of the random variables $\hat{x}_{\mathbf{R}}$ in Eq. (1). For an isotropic distribution arguments similar to those applied by Imry and Ma [16] for a random-field Ising model bring about the absence of ferromagnetic order for space dimensions $d < 4$ [12,17], whereas anisotropic distributions may lead to a ferromagnetic order [18].

Application of the Wilson RG technique to RAM with the isotropic distribution of a local anisotropy axis suggests [10] the possibility of "runaway" solutions of the recursion equations. Such a behavior has been interpreted as a smeared transition. However this result was obtained in first order of the ε -expansion and remains to be confirmed also in higher orders.

Here, we will report results obtained by means of the field theoretical RG technique in two-loop approximation refined by a resummation of the resulting asymptotic series. We will consider two cases of distribution of the random anisotropy axis and show that a ferromagnetic second-order phase transition takes place only when the distribution is non-isotropic. Moreover we will show that the RAM provides another example of a disordered model, where the only possible new critical behavior is of "random Ising" type, similar to the site-diluted magnets [5]. More detailed results can be found in Refs. [19,20].

2 Isotropic Case

In order to deal with quenched disorder, one way to obtain the effective Hamiltonian of a RAM is to make use of the replica trick. For a given configuration of quenched random variables $\hat{x}_{\mathbf{R}}$ in Eq. (1) the partition function may then be written in the form of a functional integral of a Gibbs distribution depending on $\hat{x}_{\mathbf{R}}$. To average over configurations, one should complete the model by choosing a certain distribution of $\hat{x}_{\mathbf{R}}$. We will analyze two cases: first the isotropic case, where the random vector $\hat{x}$ points with equal probability in any direction of the m -dimensional hyperspace, and second the cubic case, where $\hat{x}$ lies along the edges of the m -dimensional hypercube. Other distributions may be considered as well. In the first case the distribution function reads:

$$p(\hat{x}) \equiv \left(\int d^m \hat{x} \right)^{-1} = \frac{\Gamma(m/2)}{2\pi^{m/2}}. \quad (2)$$

Following the above described program, one ends up with the replica $n \rightarrow 0$ limit of the effective Hamiltonian [10]:

$$\mathcal{H}_{\text{eff}} = - \int d^d R \left\{ \frac{1}{2} [\mu_0^2 |\vec{\varphi}|^2 + |\vec{\nabla} \vec{\varphi}|^2] + u_0 |\vec{\varphi}|^4 + v_0 \sum_{\alpha=1}^n |\vec{\phi}^\alpha|^4 + w_0 \sum_{\alpha, \beta=1}^n \sum_{i, j=1}^m \phi_i^\alpha \phi_j^\alpha \phi_i^\beta \phi_j^\beta \right\}, \quad (3)$$

where μ_0^2 and u_0, v_0, w_0 are defined by D_0 and familiar bare couplings of an m -vector model, and $\vec{\phi}^\alpha \equiv \vec{\phi}_{\mathbf{R}}^\alpha$ is an m -dimensional vector, $|\vec{\varphi}|^2 = \sum_{\alpha=1}^n |\vec{\phi}^\alpha|^2$. The bare couplings are restricted by $u_0 > 0, v_0 > 0, w_0 < 0$. Furthermore, values of u_0 and w_0 are related to appropriate cumulants of the distribution function (2) and their ratio equals $w_0/u_0 = -m$. Note that the symmetry of the u_0 and v_0 terms corresponds to the random site m -vector model [21]. However the u_0 -term has an opposite sign.

In order to study long-distance properties of the Hamiltonian (3), we use the field theoretical RG approach [1]. Here the critical point of a system corresponds to a stable fixed point (FP) of the RG transformation. We apply the massive field theory renormalization scheme [22] performing renormalization at fixed space dimension d and zero external momenta. In two-loop approximation we get [19] expressions for the RG functions in form of asymptotic series in renormalized couplings u, v, w .

As it was mentioned in the introduction, the only known RG results for RAM with isotropic distribution of the local anisotropy axis so far are those obtained in first order in ε [10]. In total one obtains eight fixed points. All FPs with $u > 0, v > 0, w < 0$ appear to be unstable for $\varepsilon > 0$ except of the "polymer" $O(n=0)$ FP III which is stable for all m (see Fig. 1). However the presence of a stable FP is not a sufficient condition for a second-order phase transition. In order to be physically relevant, the FP should be accessible from the initial values of couplings. This is not the case for the location of FPs shown in Fig. 1. Indeed starting from the region of physical initial conditions (denoted by the cross in Fig. 1) in the plane of $v = 0$ one would have to cross the separatrix joining the unstable FPs I and VI. This is not possible and so one never reaches the stable FP III. As far as both FPs I and VI are strongly unstable with respect to v , FP III is not accessible for arbitrary positive v either. Finally, the runaway solutions of the RG equations show that the second-order phase transition is absent in the model. The main question of

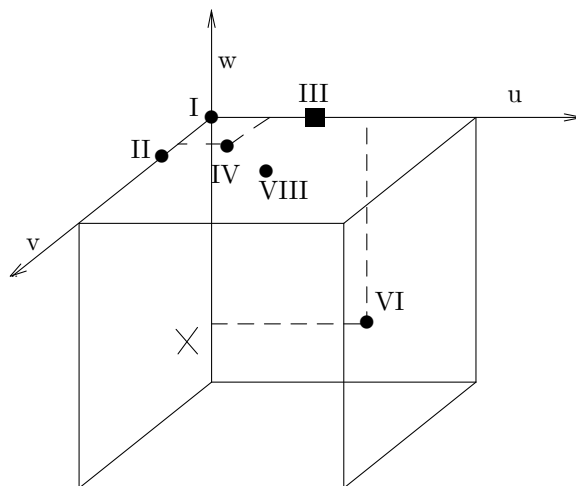


Figure 1. Fixed points of the RAM with isotropic distribution of a local anisotropy axis. The fixed points located in the octant $u > 0, v > 0, w < 0$ are shown. The filled box shows the stable fixed point, the cross denotes typical initial values of couplings.

Table 1. Resummed values of the fixed points and critical exponents for the isotropic case in two-loop approximation for $d = 3$. We absorb the value of a one-loop integral into the normalization of the couplings.

FP	m	u^*	v^*	w^*	ν	η
I	$\forall m$	0	0	0		
II	2	0	0.9107	0	0.663	0.027
	3	0	0.8102	0	0.693	0.027
	4	0	0.7275	0	0.724	0.027
III	$\forall m$	1.1857	0	0	0.590	0.023
IV	2	-0.0322	0.9454	0	0.668	0.027
	3	0.1733	0.6460	0	0.659	0.027
	4	0.2867	0.4851	0	0.653	0.028
VI	2	1.4650	0	-1.6278	0.449	-0.028
VIII	2	0.7517	0.7072	-0.3984	0.626	0.031
	3	0.8031	0.5463	-0.3305	0.620	0.029
	4	0.8349	0.4545	-0.2888	0.617	0.029

interest is whether the above described picture of runaway solutions is not an artifact of the ε -expansion. To check this, we use a more refined analysis of the FPs and their stability, considering the series for RG functions directly at $d = 3$ [22]. It is known that series of this type are at best asymptotic and a resummation procedure has to be applied to obtain reliable data on their basis. We make use of Padé-Borel resummation techniques [23], first writing the RG functions as resolvent series [24] in one auxiliary variable and then performing the resummation. Numerical values of the FPs are given in Table 1. Resummed two-loop results qualitatively confirm the picture obtained in first-order ε -expansion: the stability of the FPs does not change after the resummation. This supports the conjecture of Aharony [10] that an accessible stable FP for the RAM with isotropic distribution of the local anisotropy axis is absent. In the table, we list values of correlation length and pair correlation function critical exponents ν and η which are resummed in a similar way. As they are calculated in unstable FPs, they have rather to be considered as effective ones.

3 Cubic Case

Let us now consider the second example of anisotropy axis distribution, when the vector $\hat{x}_{\mathbf{R}}$ in Eq. (1) points only along one of the $2m$ directions of axes $\hat{k}_i$ of a cubic lattice:

$$p(\hat{x}) = \frac{1}{2m} \sum_{i=1}^m [\delta^{(m)}(\hat{x} - \hat{k}_i) + \delta^{(m)}(\hat{x} + \hat{k}_i)]. \quad (4)$$

The *rationale* for such a choice is to mimic the situation when an amorphous magnet still “remembers” the initial (cubic) lattice structure. Repeating the procedure described in the previous section, one ends up with the following effective Hamiltonian which is of interest in the limit $n \rightarrow 0$ [10]:

$$\begin{aligned} \mathcal{H}_{\text{eff}} = - \int d^d R \left\{ \frac{1}{2} \left[\mu_0^2 |\vec{\varphi}|^2 + |\vec{\nabla} \vec{\varphi}|^2 \right] + u_0 |\vec{\varphi}|^4 + v_0 \sum_{\alpha=1}^n |\vec{\phi}^\alpha|^4 \right. \\ \left. + w_0 \sum_{i=1}^m \sum_{\alpha, \beta=1}^n \phi_i^{\alpha 2} \phi_i^{\beta 2} + y_0 \sum_{i=1}^m \sum_{\alpha=1}^n \phi_i^{\alpha 4} \right\}. \end{aligned} \quad (5)$$

Here, the bare couplings are $u_0 > 0$, $v_0 > 0$, $w_0 < 0$. The y_0 term is generated when the RG transformation is applied and may be of either sign.

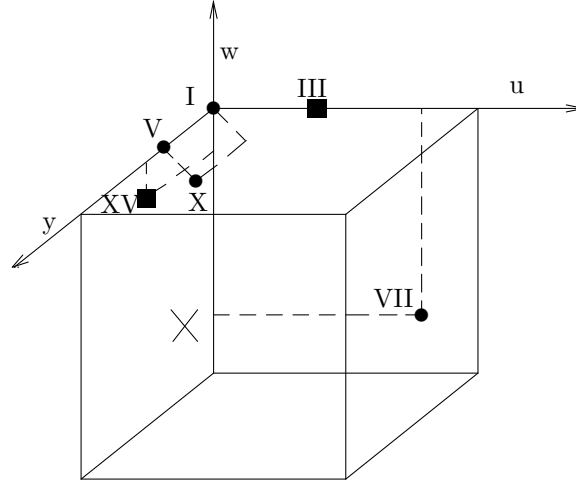


Figure 2. Fixed points of the RAM with distribution of a local anisotropy axis along hypercube axes for $v = 0$. The only fixed points located in the region $u > 0, w < 0$ are shown. Filled boxes show the stable fixed points, the cross denotes typical initial values of the couplings.

The symmetry of the w_0 terms differs in Eqs. (3) and (5). Furthermore, values of w_0 and u_0 differ for Hamiltonians (3) and (5), but their ratio equals $-m$ again.

We apply the massive field theory renormalization scheme [22] and get the RG functions in two-loop approximation [20]. As in the previous case we reproduce the first-order ε -results [10]. Now one gets 14 FPs. However, in first order of the ε -expansion all FPs with $u > 0, v > 0, w < 0$ appear to be unstable for $\varepsilon > 0$, except of the “polymer” $O(n = 0)$ FP III which is stable for all m but not accessible (see Fig. 2). Now the account of the ε^2 -terms qualitatively changes the picture. Indeed, the system of equations for the FPs appears to be degenerated at the one-loop level. As known from other cases in two-loop order, this leads to the appearance of a new FP which is stable and is expressed by a $\sqrt{\varepsilon}$ series [21]. The possibility of such a scenario was predicted already in Refs. [18]. However it remained unclear whether there exist any other accessible stable FPs.

Applying Padé-Borel resummation, we get 16 FPs. Values of the FPs with $u^* > 0, v^* > 0, w^* < 0$ are listed in Table 2. The last FP XV in Table 2 corresponds to the stable FP of the $\sqrt{\varepsilon}$ -expansion. It has coordinates with $u^* = v^* = 0, w^* < 0$ and $y^* > 0$ and is accessible from the typical initial

Table 2. Resummed values of the FPs and critical exponents for cubic distribution in two-loop approximation for $d = 3$. We absorb the value of a one-loop integral into the normalization of the couplings.

FP	m	u^*	v^*	w^*	y^*	ν	η
I	$\forall m$	0	0	0	0	1/2	0
II	2	0	0.9107	0	0	0.663	0.027
	3	0	0.8102	0	0	0.693	0.027
	4	0	0.7275	0	0	0.720	0.026
III	$\forall m$	1.1857	0	0	0	0.590	0.023
V	$\forall m$	0	0	0	1.0339	0.628	0.026
VI	3	0.1733	0.6460	0	0	0.659	0.027
	4	0.2867	0.4851	0	0	0.653	0.027
VII	$\forall m$	2.1112	0	-2.1112	0	1/2	0
VIII	2	0	1.5508	0	-1.0339	0.628	0.026
	3	0	0.8393	0	-0.0485	0.693	0.027
	4	0	0.5259	0	0.3624	0.709	0.026
IX	3	0.1695	0.7096	0	-0.1022	0.659	0.027
	4	0.2751	0.4190	0	0.1432	0.653	0.027
X	$\forall m$	0.6678	0	-0.6678	1.0339	0.628	0.026
XV	$\forall m$	0	0	-0.4401	1.5933	0.676	0.031

values of couplings (shown by a cross in Fig. 2). Applying the resummation procedure, we have not found any other stable FPs in the region of interest. The effective Hamiltonian (5) at $u = v = 0$ in the replica limit $n \rightarrow 0$ reduces to a product of m effective Hamiltonians of a weakly diluted quenched random site Ising model. This means that for any value of $m > 1$ the system is characterized by the same set of critical exponents as those of a weakly diluted random site quenched Ising model. In Table 2, we give values of critical exponents in the other FPs as well: if the flows from initial values of couplings pass near these FPs, one may observe an effective critical behavior governed by these critical exponents.

4 Conclusions

We applied the field theoretical RG approach to analyze the critical behavior of a model of random anisotropy magnets with isotropic and cubic distribu-

tions of a local anisotropy axis. The origin of a low-temperature phase in this model is not completely clear. General arguments based on an estimate of the energy for formation of magnetic domains [16] lead to the conclusion that for $d < 4$ a ferromagnetic order is absent [12,17]. However, these arguments do not take into account the entropy which may be important for disordered systems [14]. Furthermore, they do not apply for anisotropic distributions of the random axis [18].

In the RG analysis the absence of a ferromagnetic second-order phase transition corresponds to the lack of a stable FP of the RG transformation. However in the case of RAM with isotropic distribution of a local anisotropy axis the scenario differs. Our two-loop calculation leads to a $O(n = 0)$ symmetric FP which is stable for any value of m for both isotropic and cubic distributions of a random anisotropy axis. Note that this FP is not accessible from the initial values of the couplings. We checked the location of the FPs up to second order in the ε -expansion and by means of a fixed $d = 3$ technique refined by Padé-Borel resummation.

In the case of isotropic distribution of a random anisotropy axis our analysis supports the conjecture of Aharony [10] based on results linear in ε about runaway solutions of the RG equations. For the cubic distribution we get two stable FPs. One of them (FP III in Fig. 2) is not accessible as in the isotropic case. But the disordered Ising-like FP (FP XV in Fig. 2) may be reached from the initial values of couplings. Applying the resummation procedure we have not found any other stable FPs in the region of interest. This means that RAM with cubic distributions of a random anisotropy axis is governed by a set of critical exponents of a weakly site diluted quenched Ising model [21,25].

To conclude we want to attract attention to a certain similarity in the critical behavior of both random-site [21] and random-anisotropy [4] quenched magnets: if there appears a *new critical behavior* at all, it *is always governed by critical exponents of site-diluted Ising type*. The above calculations of a "phase diagram" of RAM are based on two-loop expansions improved by a resummation technique. Once the qualitative picture becomes clear, there is no need to go into higher orders of a perturbation theory as far as the critical exponents of the site-diluted Ising model are known by now with high accuracy [25].

Acknowledgments

Yu. H. acknowledges helpful discussions with Mykola Shpot. This work has been supported in part by "Österreichische Nationalbank Jubiläumsfonds" through the grant No. 7694.

References

- [1] D.J. Amit, *Field Theory, the Renormalization Group, and Critical Phenomena* (World Scientific, Singapore, 1984); J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 1st ed. (Oxford University Press, Oxford, 1989); H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [2] H. Kleinert *et al.*, *Phys. Lett. B* **272**, 39 (1991); *ibid.* **319**, 545 (1993); R. Guida and J. Zinn-Justin, *J. Phys. A* **31**, 8103 (1998).
- [3] H. Kleinert and V. Schulte-Frohlinde, *Phys. Lett. B* **342**, 284 (1995); J.M. Carmona, A. Pelissetto, and E. Vicari, *Phys. Rev. B* **61**, 15136 (2000).
- [4] R. Harris, M. Plischke, and M.J. Zuckermann, *Phys. Rev. Lett.* **31**, 160 (1973).
- [5] A weak quenched random site dilution is irrelevant, if the heat capacity critical exponent α_{pure} of the undiluted system is negative (see A.B. Harris, *J. Phys. C* **7**, 1671 (1974)). For positive α_{pure} dilution causes change in the universality class (for a recent review see e.g. R. Folk, Yu. Holovatch, and T. Yavors'kii, *Phys. Rev. B* **61**, 15114 (2000)).
- [6] See e.g. reviews by D.P. Belanger and A.P. Young, *JMMM* **100**, 272 (1991); D.P. Belanger, eprint: cond-mat/0009029.
- [7] B. Derrida and J. Vannimenus, *J. Phys. C* **13**, 3261 (1980); R. Harris and D. Zobin, *J. Phys. F* **7**, 337 (1977); E. Callen, Y.J. Liu, and J.R. Cullen, *Phys. Rev. B* **16**, 263 (1977); J.D. Patterson, G.R. Gruzalski, and D.J. Sellmyer, *Phys. Rev. B* **18**, 1377 (1978); R. Alben, J.J. Becker, and M.C. Chi, *J. Appl. Phys.* **49**, 1653 (1978).
- [8] M.C. Chi and R. Alben, *J. Appl. Phys.* **48**, 2987 (1977); M.C. Chi and T. Egami, *J. Appl. Phys.* **50**, 1651 (1979); R. Harris and S.H. Sung, *J. Phys. F* **8**, L299 (1978); C. Jayaprakash and S. Kirkpatrick, *Phys. Rev. B* **21**, 4072 (1980).
- [9] Y.Y. Goldschmidt, *Nucl. Phys. B* **225**, 123 (1983); Y.Y. Goldschmidt, *Phys. Rev. B* **30**, 1632 (1984); A. Khurana, A. Jagannathan, and J.M.

- Kosterlitz, *Nucl. Phys. B* **240**, 1 (1984); A. Jagannathan, B. Shaub, and J.M. Kosterlitz, *Nucl. Phys. B* **265**, 324 (1986); Y.Y. Goldschmidt and A. Aharony, *Phys. Rev. B* **32**, 264 (1985).
- [10] A. Aharony, *Phys. Rev. B* **12**, 1038 (1975).
- [11] J.H. Chen and T.C. Lubensky, *Phys. Rev. B* **16**, 2106 (1977).
- [12] R.A. Pelcovits, E. Pytte, and J. Rudnick, *Phys. Rev. Lett.* **40**, 476 (1978); R.A. Pelcovits, *Phys. Rev. B* **19**, 465 (1979); E. Pytte, *Phys. Rev. B* **18**, 5046 (1978).
- [13] A.J. Bray and M.A. Moore, *J. Phys. C* **18**, L139 (1985); A.B. Harris, R.G. Caflisch, and J.R. Banavar, *Phys. Rev. B* **35**, 4929 (1987).
- [14] K.M. Fischer and A. Zippelius, *J. Phys. C* **18**, L1139 (1985).
- [15] A. Aharony and E. Pytte, *Phys. Rev. Lett.* **19**, 1583 (1980); R. Fisch, *Phys. Rev. B* **57**, 269 (1998); D.E. Feldman, *JETP Lett.* **70**, 135 (1999); J. Chakrabati, *Phys. Rev. Lett.* **81**, 385 (1998).
- [16] Y. Imry and S.-K. Ma, *Phys. Rev. Lett.* **35**, 1399 (1975).
- [17] S.-K. Ma and J. Rudnick, *Phys. Rev. Lett.* **40**, 589 (1978).
- [18] D. Mukamel and G. Grinstein, *Phys. Rev. B* **25**, 381 (1982); A.L. Korzhenevskii and A.A. Luzhkov, *Sov. Phys. JETP* **94**, 250 (1988).
- [19] M. Dudka, R. Folk, and Yu. Holovatch, to appear in *Cond. Matt. Phys.* **4**, (2001).
- [20] M. Dudka, R. Folk, and Yu. Holovatch, in preparation.
- [21] G. Grinstein and A. Luther, *Phys. Rev. B* **13**, 1329 (1976).
- [22] G. Parisi (1973) (unpublished); *J. Stat. Phys.* **23**, 49 (1980).
- [23] G.A. Baker Jr., B.G. Nickel, and D.I. Meiron, *Phys. Rev. B* **17**, 1365 (1978).
- [24] P.J.S. Watson, *J. Phys. A* **7**, L167 (1974).
- [25] Up to now the critical exponents of a weakly diluted quenched Ising model are calculated in RG technique with record six-loop accuracy in massive theory scheme: A. Pelissetto and E. Vicari, *Phys. Rev. B* **62**, 6393 (2000). In minimal subtraction scheme five-loop results are available: B.N. Shalaev, S.A. Antonenko, and A.I. Sokolov, *Phys. Lett. A* **230**, 105 (1997); R. Folk, Yu. Holovatch, and T. Yavors'kii, *J. Phys. Stud.* **2**, 213 (1998); *JETP Lett.* **69**, 747 (1999).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

GAPS AND MAGNETIZATION PLATEAUS IN LOW-DIMENSIONAL QUANTUM SPIN SYSTEMS

K.-H. MÜTTER

Fachbereich Physik, Bergische Universität Wuppertal, Germany

E-mail: muetter@theorie.physik.uni-wuppertal.de

The Lieb-Schulz-Mattis theorem predicts the existence of “soft modes” (zero energy excitations) in quasi one-dimensional quantum spin systems exposed to an homogeneous external field. The soft modes appear as zeroes in the dispersion curve and as singularities in the static structure factors. The critical behavior at the (field dependent) soft mode momenta is described by conformal field theory. Breaking of translation invariance, e.g. by the modulation of the external field of the nearest-neighbor coupling, is shown to yield an efficient mechanism to generate gaps and plateaus in the magnetization curve. The plateaus appear at those magnetizations where the period of the modulation coincides with the soft mode momentum. Spin ladder systems develop a characteristic sequence of magnetization plateaus. Experimental results on plateaus in magnetization curves (so far in two- and three-dimensional compounds) are discussed.

1 Introduction

Gaps in the energy spectrum play a crucial role in condensed matter physics. Fundamental properties in superconductivity or in the fractional quantum Hall effect originate from the existence of a gap between the ground state and the excited states. The corresponding ground-state wave function describes cooperative phenomena, where infinitely many degrees of freedom act together for example to build the Bose condensate of Cooper pairs in the BCS theory or the condensate of composite fermions or bosons in the fractional quantum Hall effect.

In this report, I present recent results on the mechanisms, which lead to the formation of gaps in the energy spectrum of low-dimensional quantum

spin systems. The corresponding Hamiltonians

$$H = \sum_l J_l H_l, \quad (1)$$

$$H_l = \sum_i \vec{S}_i \vec{S}_{i+l} \quad (2)$$

are built up from isotropic couplings of spin S operators $\vec{S}_i$ and $\vec{S}_{i+l}$ at sites i and $i+l$. Many exact and numerical results on the one-dimensional system with nearest-neighbor couplings have been accumulated during the last 70 years:

- (1) The spin 1/2 system has been solved by Bethe [1] in 1931 with his famous ansatz which allows the computation of the energy eigenvalues. Quite recently an exact computation of the specific heat has been presented by Klümper [2]. An analytic derivation of the low lying excitations – the so called two-spinon spectrum – and their impact on the dynamical structure factor in the ground state has been achieved by Bougourzi, Karbach, and Müller [3]. The spin 1/2 system with nearest-neighbor couplings is gapless. The critical properties – coded in the η -exponents of the static structure factor – are well described by the predictions of conformal field theory [4], which relates η to the finite-size behavior of the energy eigenvalues [5,6].
- (2) The spin 1 system is not solvable with the Bethe ansatz or similar techniques like the Yang-Baxter equations. Numerical results were obtained by exact diagonalization with a Lanczos or conjugate gradient algorithm and recently by means of the density matrix renormalization group (DMRG) [7]. Haldane [8] formulated in 1983 his famous conjecture that quantum spin chains with integer spin $S = 1, 2, \dots$ have a gap, whereas chains with half integer spin $S = 1/2, 3/2, \dots$ are gapless. Recent numerical results [9] for spin 1, spin 3/2 and spin 2 chains support the Haldane conjecture.
- (3) During the last 10 years, also ladder systems [10] – quantum spin systems between one and two dimensions – have been studied with numerical methods like DMRG. Concerning the ground-state properties, the following peculiar property has been found: Ladder systems with an even number of legs ($l = 2, 4, 6, \dots$) have a gap; those with an odd number ($l = 1, 3, 5, \dots$) have no gap.
- (4) The two-dimensional spin 1/2 system has been studied intensively after

the discovery of high T_c superconductivity motivated by the following fact: The undoped material La_2CuO_4 shows antiferromagnetic order at low temperatures in the CuO planes. The corresponding order parameter - the "staggered" magnetization, given by the static structure factor at momentum $\vec{p} = (\pi, \pi)$ - is nonvanishing in two dimensions [11] in contrast to the one-dimensional case. We have studied the transition from two to one dimension in a model with different nearest-neighbor couplings $J_{\parallel}$ and $J_{\perp}$ along the horizontal and vertical directions. We find that the staggered magnetization is non-negative for all couplings $\alpha = J_{\perp}/J_{\parallel} > 0$ and vanishes with an infinite slope for $\alpha = 0$ (see Ref. [12]).

In this work, I will briefly review in Section 2 the Lieb-Schultz-Mattis theorem [13] and its implications on the existence of soft modes in one-dimensional quantum spin systems. Then I will demonstrate in Section 3 that breaking of translational invariance is a very efficient mechanism for the formation of gaps and magnetization plateaus. Section 4 is devoted to spin ladder systems. It turns out that each spin ladder system with l -legs possesses a characteristic sequence of magnetization plateaus. In section 5 I will report on some compounds where experimentalists have found plateaus in the magnetization curve. Unfortunately, these compounds seem to have a two- or higher-dimensional coupling structure.

2 The Lieb-Schultz-Mattis Theorem and the Appearance of Soft Modes in One-Dimensional Quantum Spin Systems

Let me specify first my notion of a ground-state gap: There is no unitary operator U , which creates, from the ground state $|0\rangle$, *new* states

$$|n\rangle = U^n|0\rangle \quad (3)$$

with energy expectation values

$$\langle n|H|n\rangle - \langle 0|H|0\rangle = O(N^{-1}), \quad (4)$$

coinciding with the ground-state energy $E_0 = \langle 0|H|0\rangle$ in the thermodynamical limit. Of course, the new states $|n\rangle$ should differ from the ground state by their quantum numbers. For translation invariant one-dimensional systems

over a finite range L , the Hamiltonian is

$$H = \sum_{l=1}^{l=L} J_l H_l + B \sum_i S_i^{(3)}. \quad (5)$$

Lieb, Schultz, and Mattis [13] have proposed the following operator U :

$$U = \exp \left\{ \frac{2\pi i}{N} \sum_{i=1}^N i S_i^{(3)} \right\}. \quad (6)$$

The Hamiltonian commutes with the total spin

$$S_3 = \sum_{i=1}^N S_i^{(3)} \quad (7)$$

and, in the presence of a magnetic field B , the ground state is magnetized: $M = S_3/N$. The magnetization $M = M(B)$ follows from the magnetization curve. Lieb, Schultz, and Mattis proved that the new states $|n\rangle = U^n|0\rangle$ obey Eq. (4). Following the argument of Affleck and Lieb [14], the quantum numbers of the new states can be seen by applying the translation operator

$$T|n\rangle = e^{i[nq(M)+p_0]}|n\rangle, \quad (8)$$

i.e. the momentum of the new states,

$$p_n(M) = nq(M) + p_0(M), \quad (9)$$

differs from the ground-state momentum $p_0(M)$ by

$$q(M) = \pi(1 - 2M), \quad (10)$$

unless $nq(M)$ is a multiple of 2π .

In the following, I will call the state $|n\rangle$ the n 'th soft mode. Soft modes appear as zeros in energy differences defining the dispersion curve

$$\omega(q, M, N) = E(q + p_0, M, N) - E(p_0, M, N). \quad (11)$$

In Fig. 1 I show the dispersion curve for $M = 1/4$ in a model with nearest- and next-to-nearest-neighbor coupling with $\alpha = 0$ and $\alpha = 1/2$, where $\alpha = J_2/J_1$ (see Ref. [15]). The first soft mode ($n = 1$) is clearly visible at $q = \pi/2$. The second soft mode ($n = 2$) at $q = \pi$ still suffers from large finite-size effects.

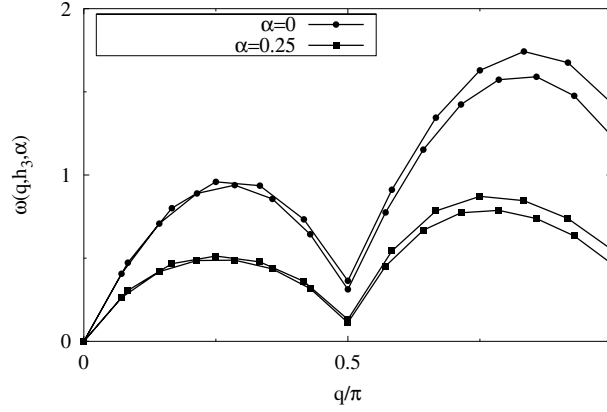


Figure 1. Dispersion curve (11) for $M = 1/4$, $\alpha = 0, 1/4$ on finite systems with $N = 24, 28$. The dips at $q = \pi/2, \pi$ occur at momenta of the first ($n = 1$) and second ($n = 2$) soft mode.

Conformal field theory makes a prediction on the finite-size dependence at the soft mode

$$\omega(q = q(M), M, N) \stackrel{N \rightarrow \infty}{\sim} \frac{\Omega(M)}{N}, \quad (12)$$

which can be verified numerically. Moreover the coefficient $\Omega(M)$ is related to the critical η -exponent

$$\eta(M) = \frac{\Omega(M)}{\pi v(M)}, \quad (13)$$

where $v(M)$ is the spin wave velocity

$$v(M) = \lim_{N \rightarrow \infty} \left[E \left(p_0 + \frac{2\pi}{N}, M, N \right) - E(p_0, M, N) \right]. \quad (14)$$

For the nearest-neighbor model ($\alpha = 0$), the energy differences, which enter on the right hand side of Eq. (13) can be computed by means of the Bethe ansatz on very large systems ($N = 10^4$). Therefore, in this case, the field dependence of the critical exponent $\eta = \eta(M)$ is very well known (see Fig. 2) [16]. The η -exponent describes the divergence in the static structure factor at the soft mode $q = q(M)$. The latter are defined as ground-state expectation values

$$S_0(q, M) = \langle 0 | O(q) O(-q) | 0 \rangle \quad (15)$$

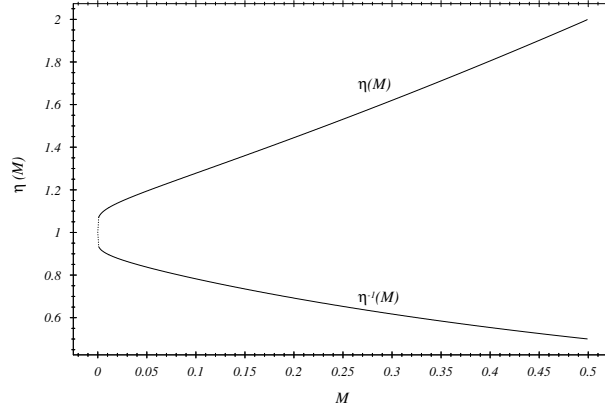


Figure 2. The critical η -exponent (13) versus magnetization M for $\alpha = 0$. The energy differences (12) and (14) were computed with the Bethe ansatz on large systems ($N = 10^4$).

of appropriate Fourier transformed operators

$$O(q) = \sum_j e^{iqj} O_j, \quad (16)$$

$$O_j =: S_j^{(3)}, (\vec{S}_j \vec{S}_{j+1}), \dots, \quad (17)$$

$$S_0(q(M), M, N) \xrightarrow{N \rightarrow \infty} N^{1-\eta(M)}. \quad (18)$$

Figure 3 shows the q -dependence of the static structure factor for the dimer operator $\vec{S}_j \vec{S}_{j+1}$ [15]. A clear singularity is seen at the first soft mode ($q = \pi/2$). No singularity is visible at the second soft mode ($q = \pi$). According to the Lieb-Schultz-Mattis theorem, the position of the soft modes does not depend on the couplings (e.g. on $\alpha = J_2/J_1$). The critical exponents $\eta(M, \alpha)$, however, do depend on α . For α large enough, the singularity at the first soft mode is weakened and the second soft mode becomes visible.

3 Breaking of Translation Invariance: A Mechanism for the Formation of Gaps and Magnetization Plateaus

Translation invariance is essential for the validity of the Lieb-Schultz-Mattis theorem. Breaking of translation invariance is a vital attack against the soft modes. It has severe consequences.

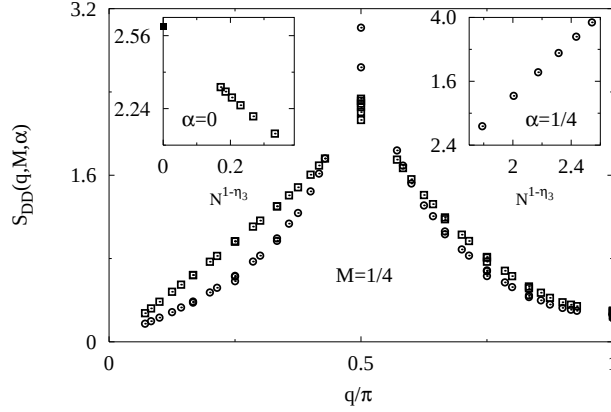


Figure 3. The static dimer structure factor (15) for $M = 1/4$, $\alpha = 0, 1/4$ on finite systems with $N = 24, 28$. A peak occurs at the first soft mode $q = \pi/2$. No peak is visible at the second soft mode $q = \pi$. The insets demonstrate the validity of (18) derived from conformal field theory.

Let us consider the nearest-neighbor model

$$H = H_1 + BS^{(3)}(0) + \delta \left(S^{(3)}(q) + S^{(3)}(-q) \right) \quad (19)$$

in a homogenous field B and a modulated field of strength δ :

$$S^{(3)}(q) = \sum_{j=1}^N S_j^{(3)} e^{iqj}. \quad (20)$$

In Fig. 4(a) we see what happens if one switches on the perturbation δ with wave number $q = \pi/2$. For $\delta = 0$ we see a smooth magnetization curve $M = M(B)$. Indeed this curve has been computed at the beginning of the sixties by C.N. Yang and C.P. Yang [17]; they used the Bethe ansatz to compute the energy eigenvalues in the sectors with given magnetization

$$B(M) = E \left(M + \frac{1}{N}, p_0 \right) - E(M, p_0). \quad (21)$$

If we now look at $\delta > 0$ [18]; we see the emergence of a plateau at $M = 1/4$. For this magnetization the wave number of the perturbing field $q = \pi/2$ coincides with the first soft mode:

$$q = \pi(1 - 2M) \Rightarrow M = \frac{1}{2} - \frac{q}{2\pi}. \quad (22)$$

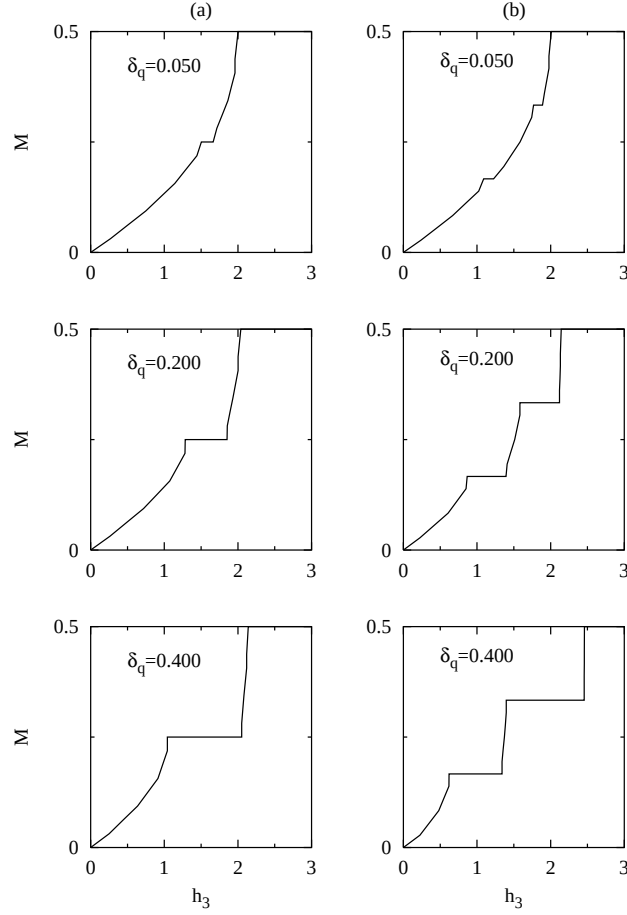


Figure 4. The magnetization curve for the nearest-neighbor Hamiltonian H_1 with a periodic perturbation: (a) of the external magnetic field with wave number $q = \pi/2$ (cf. Eq. (19)): Plateau at $M = 1/4$, and (b) of the nearest-neighbor coupling with wave numbers $q = \pi/3$ and $q = 2\pi/3$ (cf. Eq. (26)): Plateaus at $M = 1/6, 1/3$.

The connection between q and M is a special case of a more general rule established by Oshikawa, Yamanaka, and Affleck [19] on the position of possible magnetization plateaus. For small perturbations δ , the length Δ of the plateaus, i.e. the difference between the upper and lower critical fields, is

$$\Delta = B_U - B_L \sim \delta^\epsilon, \quad (23)$$

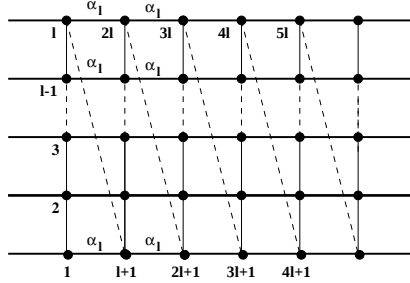


Figure 5. Mapping of the l -leg ladder on a one-dimensional system with modulated nearest-neighbor couplings.

i.e. scales with a power δ^ϵ , where ϵ is given by the critical η -exponent

$$\epsilon(M) = \frac{2}{4 - \eta(M)}. \quad (24)$$

Breaking of translation invariance can be achieved in many ways, e.g. instead of modulating the external magnetic field we can modulate as well the nearest-neighbor coupling

$$\bar{D}_1(q) = \frac{1}{2} (D_1(q) + D_1(-q)) = \sum_{i=1}^N \cos(q \cdot i) \vec{S}_i \vec{S}_{i+1}. \quad (25)$$

In Fig. 4(b) we can see the effect of a superposition of two modes:

$$\bar{D}_1\left(q = \frac{2\pi}{3}\right) + \bar{D}_1\left(q = \frac{\pi}{3}\right). \quad (26)$$

According to the rule (22) we observe two plateaus at

$$M = \frac{1}{6}, \quad M = \frac{1}{3}. \quad (27)$$

4 Gaps and Magnetization Plateaus in Spin Ladder Systems

A spin ladder system with l legs - as shown in Fig. 5 - can be viewed as a one-dimensional system

$$H = H_1 + \alpha_l H_l + D_1^{(l)}, \quad (28)$$

with nearest-neighbor coupling and couplings over l sites [15]. The translation invariant ring H_1 , however, also contains "diagonal" couplings $l \rightarrow l + 1$,

$2l \rightarrow 2l + 1$ (dashed lines in Fig. 5), which are absent in a normal ladder. We subtract these couplings using the term

$$D_1^{(l)} = \sum_{n=1}^N J_n^{(l)} \vec{S}_n \vec{S}_{n+1}, \quad (29)$$

$$J_n^{(l)} = \begin{cases} 0, & n = 1, \dots, l-1, \\ \delta, & n = l, \end{cases} \quad (30)$$

$$J_{n+l}^{(l)} = J_n^{(l)}. \quad (31)$$

The periodicity of the couplings,

$$J_n^{(l)} = \sum_{j=0}^{l/2} \cos\left(\frac{2\pi nj}{l}\right) \delta\left(q = \frac{2\pi j}{l}\right), \quad (32)$$

leads to a Fourier decomposition of the “unwanted” couplings:

$$D_1^{(l)} = \sum_q \delta_q^l \bar{D}_1(q). \quad (33)$$

The “unwanted” couplings induce a modulation of the nearest-neighbor couplings with certain wave numbers q , which again generate magnetization plateaus at $M = 1/2 - q/2\pi$ provided that the first soft mode is active. This consideration leads to the following prediction of magnetization plateaus for l leg ladders:

l	2	3	4	5	6
M	0	1/6	0; 1/4	1/10; 3/10	0; 1/6; 1/3

Note, that the even l ladders have a plateau at $M = 0$. This means they have a gap in the absence of a magnetic field. The odd l ladders do not have such a gap. (I mentioned this phenomenon [10] in the introduction.) The table means that one can associate a characteristic sequence of plateaus to each ladder.

We have tested this prediction by a numerical calculation of the magnetization curve in ladder systems with $l = 3, 4$ and 5 legs by means of the density matrix renormalization group (DMRG). The gap in the two leg ladder had been discussed before in the context of the compound CuGeO_3 , which shows a spin Peierls transition [20].

The plateaus at $M = 1/3$ in the three leg ladder have been discovered by Honecker and collaborators [21]. In Fig. 6 you see this magnetization

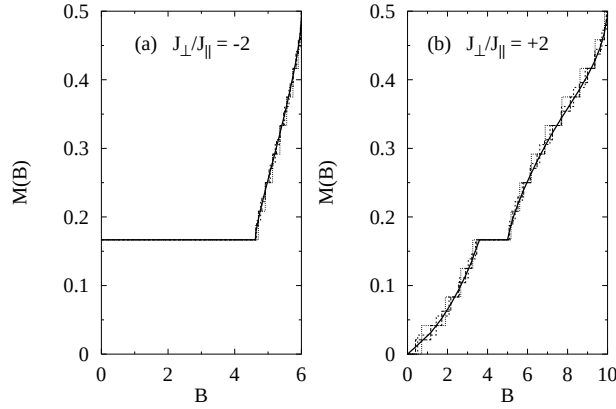


Figure 6. Magnetization curve of a three leg ladder with plateau at $M = 1/4$: (a) Couplings along the legs and rings are ferro- and antiferromagnetic respectively: ferrimagnetism. (b) Both couplings are antiferromagnetic.

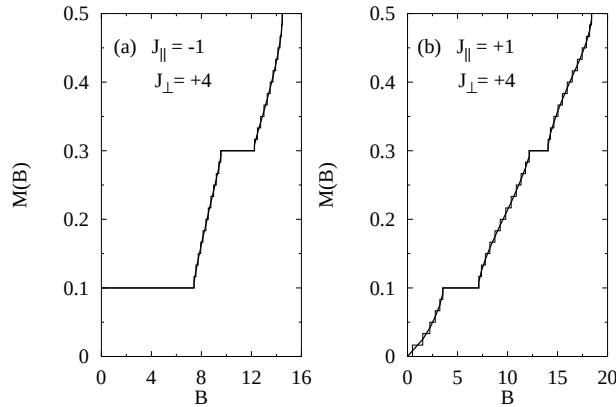


Figure 7. The same magnetization curve as Fig. 6, for a five leg ladder.

curve [15]. In the right picture both couplings $J_{\perp}$, $J_{\parallel}$ along the rings and the legs are antiferromagnetic (i.e. positive). In the left, the ring coupling is antiferromagnetic, the leg coupling ferromagnetic; here the plateau extends to a zero magnetic field, which means that the ground state at $B = 0$ is not a singlet, but a state with total spin $S = M \cdot N = N/3$. Such a phenomenon is often called “ferrimagnetism”. The five leg ladder is shown in Fig. 7. The two predicted plateaus at $M = 1/10$ and $M = 3/10$ are clearly visible. Switching

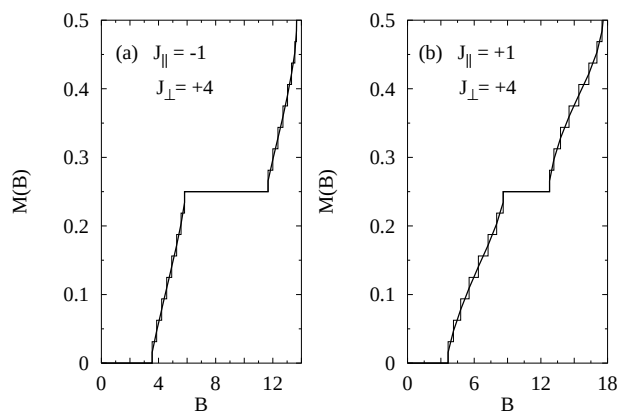


Figure 8. The same magnetization curve as Fig. 6 for a four leg ladder. But: No ferrimagnetism.

the leg coupling from antiferromagnetic to ferromagnetic, the phenomenon of ferrimagnetism appears again.

Finally, the four leg ladder is shown in Fig. 8. Two plateaus can be seen at $M = 0$ and $M = 1/4$, however the change from antiferromagnetic to ferromagnetic leg coupling does not change the magnetization curves substantially. In particular, there is no ferrimagnetism.

5 Experimental Evidence for Plateaus in Magnetization Curves

The following compounds have been synthesized and investigated by several groups:

(1) NH_4CuCl_3 :

The authors of Ref. [22] found plateaus at $M = 1/8, 3/8$. It is suggested that this compound forms a two-dimensional structure of coupled zig zag ladders with two legs [23]. However, two leg zig zag ladders alone cannot explain the observed plateaus.

(2) $CsCuCl_3$:

The authors of Ref. [24] found a plateau at $M = 1/6$. They consider the compound as a three-dimensional structure built up from coupled three leg ladders. This would easily explain the position of the plateau [25], if the coupling between the three leg ladders does not change the situation.

(3) $SrCu_2(BO_3)_2$:

The authors of Ref. [26] found plateaus at $M = 1/6, 1/8, 1/16$. They

suggest [27] that the compound has a two-dimensional coupling structure à la Shastry-Sutherland [28].

6 Perspectives

In this report I have restricted the discussion on the existence of magnetization plateaus in isotropic spin 1/2 systems. The considerations presented here can be extended to higher spin systems. Our next project is to consider a quite general class of systems with three states at each site. Spin 1 systems are included as well as lattice systems with spin 1/2 particles carrying charge, like the $t - J$ model. The $t - J$ model is a Hubbard model for electrons on a lattice with a constraint preventing two electrons, one with spin up, the other with spin down, to sit on the same site. The symmetry structure of three-state systems becomes apparent if we express the couplings between sites x and y in terms of the 8 generators of the $SU(3)$, the so called Gell-Mann matrices λ_A , $A = 1, \dots, 8$:

$$H(x, y) = \sum_{A=1}^8 \lambda_A(x) \lambda_A(y) J_A. \quad (34)$$

The Gell-Mann matrices are just the $SU(3)$ analogue of the Pauli matrices for $SU(2)$. For the $t - J$ model the coupling parameters J_A are related to the parameters t and J via

$$J_1 = J_2 = J_3 = \frac{J}{4}, \quad J_4 = J_5 = J_6 = J_7 = -\frac{t}{2}, \quad J_8 = -\frac{J}{12}. \quad (35)$$

The Cartan subalgebra of $SU(3)$ contains two elements. In case of the $t - J$ model we choose the 3-component of the spin (λ_3) and the charge (λ_8) of the state at site x . Total spin and total charge are conserved quantities in the $t - J$ model. In one-dimensional models the whole line of arguments developed by Lieb, Schultz, and Mattis [13] in the spin sector can be repeated as well as in the charge sector of the $t - J$ model. In other words there are soft modes with momenta $q(\rho) = 2\pi\rho$ which move with the charge density ρ as long as translation invariance and the finiteness of the range of the couplings is guaranteed. We expect that a breaking of translation invariance by means of a perturbation with wave number q will produce a plateau at $\rho = q/2\pi$ in the curve $\rho = \rho(\mu)$, which describes the dependence of the electronic density ρ (or filling factor) as function of the chemical potential μ for the electrons.

Moreover, the $t - J$ model on a ladder system with l legs should evolve a characteristic sequence of plateaus in $\rho(\mu)$. Plateaus in the filling factor $\rho(\mu)$ for the two-dimensional Hall system yield indeed the explanation for the integer and fractional quantum Hall effect. In these systems one can tune the chemical potential μ for the electrons by means of the perpendicular magnetic field.

Acknowledgments

The results I presented in this review were achieved in intensive collaborations over many years with Ender Aysal, Andreas Fledderjohann, Carsten Gerhardt, Michael Karbach, Andreas Schmitt, Roger Wiessner, and S.M. Yang. I am grateful for their help, many fruitful discussions and for their patience.

References

- [1] H.A. Bethe, *Z. Phys.* **71**, 205 (1931).
- [2] A. Klümper, *Eur. Phys. J. B* **5**, 677 (1998).
- [3] A.H. Bougourzi, M. Couture, and M. Kacir, *Phys. Rev. B* **54**, R12669 (1996); M. Karbach, G. Müller, A.H. Bougourzi, A. Fledderjohann, and K.-H. Mütter, *Phys. Rev. B* **55**, 12510 (1997).
- [4] J.L. Cardy, *Nucl. Phys. B* **270**, 186 (1986), and in *Phase Transitions and Critical Phenomena*, Vol. 11, Eds. C. Domb and J.L. Lebowitz (Academic Press, London, 1987), p. 55.
- [5] F.C. Alcaraz, M.N. Barber, and M.T. Batchelor, *Ann. Phys. (N.Y.)* **182**, 280 (1988).
- [6] N.M. Bogoliubov, A.G. Izergin, and N.Y. Reshetikin, *Nucl. Phys. B* **275**, 687 (1986).
- [7] S.R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [8] F.D.M. Haldane, *Phys. Rev. Lett.* **50**, 1153 (1983).
- [9] U. Schollwöck and T. Jolicoeur, *Eur. Phys. Lett.* **30**, 493 (1995); U. Schollwöck, T. Jolicoeur, and T. Garel, *Phys. Rev. B* **53**, 3304 (1996).
- [10] E. Dagotto and T.M. Rice, *Science* **271**, 618 (1996).
- [11] H.J. Schulz and T.A.L. Ziman, *Europhys. Lett.* **18**, 355 (1992); H.J. Schulz, T.A.L. Ziman, and D. Poilblanc, *J. Phys.* **6**, 675 (1996); U.J. Wiese and H.P. Ying, *Z. Phys. B* **93**, 147 (1994).

- [12] A. Fledderjohann, K.-H. Mütter, M.S. Yang, and M. Karbach, *Phys. Rev. B* **57**, 956 (1998).
- [13] E. Lieb, T. Schultz, and D. Mattis, *Ann. Phys. (N.Y.)* **16**, 407 (1961).
- [14] I. Affleck and E.H. Lieb, *Lett. Math. Phys.* **12**, 57 (1986).
- [15] A. Fledderjohann, C. Gerhardt, M. Karbach, K.-H. Mütter, and R. Wiessner, *Phys. Rev. B* **59**, 991 (1999).
- [16] A. Fledderjohann, C. Gerhardt, K.-H. Mütter, A. Schmitt, and M. Karbach, *Phys. Rev. B* **54**, 7168 (1996).
- [17] C.N. Yang and C.P. Yang, *Phys. Rev.* **150**, 321, 327 (1966); *ibid.* **151**, 258 (1966).
- [18] A. Fledderjohann, M. Karbach, and K.-H. Mütter, *Eur. Phys. J. B* **7**, 225 (1999).
- [19] M. Oshikawa, M. Yamanaka, and I. Affleck, *Phys. Rev. Lett.* **78**, 1984 (1997).
- [20] M. Hase, I. Terasaki, and K. Uchinokura, *Phys. Rev. Lett.* **70**, 3657 (1993); M.C. Cross and D.S. Fisher, *Phys. Rev. B* **19**, 402 (1979); M.C. Cross, *Phys. Rev. B* **20**, 4606 (1979).
- [21] D.C. Cabra, A. Honecker, and P. Pujol, *Phys. Rev. Lett.* **79**, 5126 (1997); *Phys. Rev. B* **58**, 6241 (1998); *Eur. Phys. J. B* **13**, 55 (2000).
- [22] W. Shiramura *et al.*, *J. Phys. Soc. Jpn.* **67**, 1548 (1998).
- [23] A.K. Kolezhuk, *Phys. Rev. B* **59**, 4181 (1999).
- [24] H. Nojiri, Y. Tabunaga, and M. Motokawa, *J. de Physique* **49**, 1459 (1988).
- [25] A. Honecker, M. Kaulke, and K.D. Schotte, *Eur. Phys. J. B* **15**, 423 (2000).
- [26] H. Kageyama *et al.*, *J. Phys. Soc. Jpn.* **69**, 1016 (2000).
- [27] S. Miyahara and K. Ueda, *Phys. Rev. Lett.* **82**, 3701 (1999); E. Müller-Hartmann, R.P. Singh, C. Knetter, and G. Uhrig, *Phys. Rev. Lett.* **84**, 1808 (2000).
- [28] B.S. Shastry and B. Sutherland, *Physica B* **108**, 1069 (1981).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

EFFECTIVE FREE ENERGY OF GINZBURG-LANDAU MODEL

A.M.J. SCHAKEL

*National Chiao Tung University, Department of Electrophysics,
Hsinchu, 30050, Taiwan, R.O.C.
E-mail: schakel@cc.nctu.edu.tw*

It is argued that the presence of a nonanalytic term in the effective potential of the Ginzburg-Landau model is immaterial as far as the order of the superconductor-normal phase transition is concerned. To achieve agreement with the renormalization group, the effective potential has to be extended to include derivative terms, providing the theory with a low momentum scale which can be varied to probe the (possible) fixed point.

1 Prelude

This paper is dedicated to Professor Kleinert on the occasion of his 60th birthday with special thanks for the exciting years I had the good fortune to be his *Wissenschaftlicher Assistent* at the Freie Universität Berlin. It is an implementation of one of the many neat ideas he shared with me over the years. This particular one he told me about four years ago in the late summer of 1996.

It is a common belief that the order of an equilibrium phase transition can be inferred from the form of the effective potential at criticality. For example, the presence of a cubic term in the order parameter is taken as signalling a discontinuous transition, while the absence of such a term is taken as indicating a continuous transition. At the mean-field level, where fluctuations in the order parameter are ignored, this is certainly true [1]. However, the question about the order of a transition is settled in the full theory, and fluctuations may well change the mean-field result.

In a renormalization group (RG) approach, a continuous phase transition is associated with an infrared-stable fixed point in the space of coupling parameters characterizing the theory. Sufficiently close to the transition, such a fixed point acts as an attractor to which the couplings flow when one passes to larger length scales by integrating out field fluctuations of smaller length scales. When no infrared-stable fixed point is detected in this process, the transition is discontinuous.

It is important for our purposes to note that a fixed point is probed by changing a scale. For this, the effective potential evaluated at criticality in itself is inadequate as it lacks a scale. In this contribution we show that a scale can be introduced by extending the effective potential to include derivative terms. The resulting effective free energy contains the same information as that obtained in RG. Along the way we are able to implement Kleinert's idea of finding a fixed point without calculating flow functions first.

2 Cubic Term

To be concrete we consider the superconductor-normal phase transition described by the $O(n)$ Ginzburg-Landau model, which has been one of Kleinert's research topics for many years, and to which a large part of the first volume of his textbook [2] *Gauge Fields in Condensed Matter* is devoted. The model is specified by the free energy density (in the notation of statistical physics)

$$\mathcal{E} = |(\partial_\mu - ieA_\mu)\phi|^2 + m^2|\phi|^2 + \lambda|\phi|^4 + \frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2, \quad (1)$$

with a complex order parameter ϕ having an even number n of real field components:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \vdots \\ \phi_{n-1} + i\phi_n \end{pmatrix}. \quad (2)$$

A conventional superconductor with Cooper pairing in the s-channel corresponds to $n = 2$. The parameters e and m are the electric charge and mass of the field, while λ is the coupling constant characterizing the 4th-order interaction term. We include a gauge-fixing term with parameter α , and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the (magnetic) field strength. For convenience we will work in the gauge $\partial_\mu A_\mu = 0$, which is implemented by taking the limit $\alpha \rightarrow 0$. The mass term depends on the temperature T , and changes sign at the critical

temperature T_c . In the Ginzburg-Landau model, $m^2 = \xi_0^{-2}(T/T_c - 1)$, where ξ_0 is the length scale of fluctuations in the amplitude of the order parameter.

As a first step to investigate the effect of fluctuations on the mean-field picture, let us, following Halperin, Lubensky, and Ma [3], consider those in the vector field A_μ . Since (at least in the normal phase) the corresponding mode is gapless, it can have an important impact on the infrared behavior of the theory. The functional integral over A_μ is a simple Gaussian and leads to a contribution to the effective potential energy density in d dimensions

$$\mathcal{V}_{\text{eff}} = \frac{d-1}{2} \int \frac{d^d k}{(2\pi)^d} \ln(k^2 + 2e^2|\Phi|^2), \quad (3)$$

where we assume that the order parameter is a nonfluctuating background field denoted by Φ . In deriving this, we used that the combination

$$P_{\mu\nu}(k) = \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}, \quad (4)$$

which shows up in intermediate steps, is a projection operator satisfying $P^2 = P$, and that its trace gives the number of transverse components: $\text{tr } P = d-1$. The momentum integral is easily carried out using dimensional regularization, with the result

$$\mathcal{V}_{\text{eff}} = \frac{1}{(2\pi)^{d/2}} \frac{d-1}{d} \Gamma(1-d/2) e^d |\Phi|^d, \quad (5)$$

where $\Gamma(x)$ is the Gamma function. For $d=3$, this leads to the cubic term in the effective potential mentioned in the introduction [3]. The 1-loop contribution is to be added to the mean-field potential $\mathcal{V}_0 = m|\Phi|^2 + \lambda|\Phi|^4$. It is often taken as indicating a discontinuous phase transition at a, what Kleiner [4] likes to call, precocious temperature T_1 above the mean-field critical temperature T_c ,

$$\frac{T_1}{T_c} = 1 + \frac{\xi_0^2}{18\pi^2} \frac{e^6}{\lambda}, \quad (6)$$

where the mass term is still positive, and the order parameter jumps from zero to the finite value $|\Phi|^2 = (1/18\pi^2)e^6/\lambda^2$ (see Fig. 1).

Since this result is obtained in perturbation theory, both e and λ are assumed to be small. This still leaves the ratio of the two, or the so-called Ginzburg-Landau parameter $\kappa_{\text{GL}}^2 = e^2/\lambda$, undetermined. This parameter separates type-II ($\kappa_{\text{GL}} > 1/\sqrt{2}$) superconductors, which have a Meissner phase where an applied magnetic field can penetrate the sample in the form

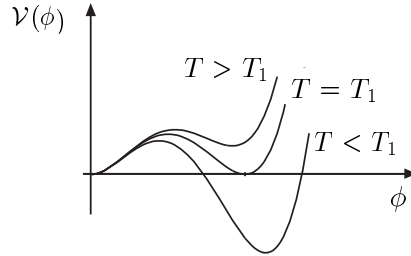


Figure 1. Sketch of the 1-loop effective potential.

of quantized flux tubes, and type-I ($\kappa_{GL} < 1/\sqrt{2}$) superconductors, for which these flux tubes, become unstable. Neglecting fluctuations in the order parameter, as is done here, is valid in the type-I regime, where κ_{GL} is small. (In the opposite limit of deep type-II superconductors where κ_{GL} is large, fluctuations in the vector field can be neglected instead.) This leads to the conclusion that type-I superconductors undergo a discontinuous phase transition [3]. It should, however, be noted that fluctuations in the order parameter produce also a cubic term in the effective potential at criticality, meaning that type-II superconductors, too, should undergo such a transition according to this argument.

As the result discussed in this section is independent of the number of field components, it should be valid for any n , including large numbers. This opens the possibility to check it using a $1/n$ expansion.

3 $1/n$ Expansion

The $1/n$ expansion can be used when the number n of field components is large, so that its inverse provides the theory with a small parameter. Contributions are then ordered not according to the number of loops, as in the loop expansion, but according to the number of factors $1/n$. The leading contribution in $1/n$ due to fluctuations in the vector field is obtained by dressing its correlation function with arbitrary many bubble insertions, and summing the entire set of Feynman diagrams [5]. The resulting series is a simple geometrical one, which leads to the following change in the correlation

function:

$$\frac{P_{\mu\nu}(k)}{k^2} \rightarrow \frac{P_{\mu\nu}(k)}{k^2 + bne^2 k^{d-2}}, \quad (7)$$

where $b = c/(d-1)$, with c the 1-loop integral

$$c(d) = \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2(k+q)^2} \Big|_{q^2=1} = \frac{\Gamma(2-d/2)\Gamma(d/2-1)}{(4\pi)^{d/2}\Gamma(d-2)}. \quad (8)$$

Below $d = 4$, the infrared behavior is now less singular. Instead of the effective potential (3), we now obtain

$$\mathcal{V}_{\text{eff}} = \frac{d-1}{2} \int \frac{d^d k}{(2\pi)^d} \ln(k^2 + be^2 k^{d-2} + 2e^2 |\Phi|^2). \quad (9)$$

For large fields, $|\Phi| > (be^{d-2})^{1/(4-d)}$, the main contribution to the integral comes from large momenta, where $k^2 + be^2 k^{d-4} \approx k^2$, so that in this limit we recover the result (5) obtained in the loop expansion: $\mathcal{V}_{\text{eff}} \propto |\Phi|^d$. In the limit of small fields, on the other hand, the main contribution comes from small momenta, where $k^2 + be^2 k^{d-4} \approx be^2 k^{d-4}$. We then find $\mathcal{V}_{\text{eff}} \propto |\Phi|^{2d}$ instead, implying that, in $d = 3$, the $1/n$ expansion no longer gives a cubic term in the effective potential. [This argument captures only those terms in the Taylor expansion of the logarithm in the effective potential (9) which diverge in the infrared. These are the important ones for our purposes as they can produce nonanalytic behavior. The first terms in the expansion may be infrared finite, depending on the dimensionality d , and have to be treated separately. But they always lead to analytic terms and are therefore of no concern to us here.]

The $1/n$ expansion calls into question the validity of the result obtained in the loop expansion as it corresponds to large field values. It is not clear that this is still in the realm of perturbation theory.

In the next section, we show that by extending the calculation of the effective potential to that of the effective free energy, which includes derivative terms, things become consistent and in agreement with RG.

4 Effective Free Energy

When computing the effective free energy, we not only consider fluctuations in the vector field, but also include those in the order parameter. To this end, we set

$$\phi(x) = \Phi(x) + \tilde{\phi}(x), \quad (10)$$

with $\Phi(x)$ being a nonfluctuating background field, and integrate out the fluctuating field $\tilde{\phi}(x)$. The main difference with the previous calculation of the effective potential is that the background field is no longer assumed to be constant, but can vary in space. Therefore, we also have to analyze the derivative terms in the effective theory.

Because of gauge invariance, the results will depend solely on the absolute value $|\Phi|$. Without loss of generality we can thus assume that the background field Φ has only one nonzero component, say the first one, $\Phi = (v, 0, \dots, 0)/\sqrt{2}$. For convenience we will work at criticality and set the mass parameter m to zero.

A straightforward calculation yields in the gauge $\partial_\mu A_\mu = 0$ the 1-loop effective free energy

$$F_{\text{eff}}[v] = \frac{1}{2} \text{Tr} \ln \left(1 + 3\lambda \frac{1}{p^2} v^2 \right) + \frac{n-2}{2} \text{Tr} \ln \left(1 + \lambda \frac{1}{p^2} v^2 \right) \quad (11)$$

$$+ \frac{1}{2} \text{Tr} \ln \left[\begin{pmatrix} 1 & 0 \\ 0 & \delta_{\mu\rho} \end{pmatrix} + \frac{1}{p^2} \begin{pmatrix} 1 & 0 \\ 0 & P_{\mu\nu}(p) \end{pmatrix} \begin{pmatrix} \lambda v^2 & e \partial_\nu v \\ e \partial_\rho v & e^2 v^2 \delta_{\nu\rho} \end{pmatrix} \right],$$

where we ignored an irrelevant constant. The trace Tr denotes the trace tr over discrete indices as well as the integration over momentum and space. More precisely,

$$\text{Tr} \ln[1 + K(x, p)] = \text{tr} \int d^d x \int \frac{d^d k}{(2\pi)^d} e^{-ik \cdot x} \ln[1 + K(x, p)] e^{ik \cdot x}, \quad (12)$$

with $p_\mu = -i\partial_\mu$ the derivative operating on everything to its right. Since the background field is space-dependent, the integrals in Eq. (11) cannot be evaluated in closed form, but only in a derivative expansion [6].

The first step in this scheme is to expand the logarithm in a Taylor series. Each term of the series contains powers of the derivative p_μ , which in the second step are shifted to the left using the identity

$$f(x) p_\mu g(x) = (p_\mu - i\partial_\mu) f(x) g(x), \quad (13)$$

where $f(x)$ and $g(x)$ are arbitrary functions. The symbol ∂_μ in the last term denotes the derivative which, in contrast to p_μ , operates *only* on the next object to its right. The next step is to repeatedly integrate by parts, until all p_μ derivatives operate to the left. They then simply produce factors of k_μ as only the function $\exp(-ik \cdot x)$ appears at the left. In shorthand, for an

arbitrary function $h(k)$, we have

$$\begin{aligned} \int \frac{d^d k}{(2\pi)^d} e^{-ik \cdot x} h(p) f(x) &= \int \frac{d^d k}{(2\pi)^d} e^{-ik \cdot x} h(-\overleftarrow{p}) f(x) \\ &= \int \frac{d^d k}{(2\pi)^d} e^{-ik \cdot x} h(k) f(x), \end{aligned} \quad (14)$$

ignoring total derivatives. In this way, all occurrences of the derivative operator p_μ are replaced with a mere integration variable k_μ . The function $\exp(ik \cdot x)$ at the right can now be moved to the left where it is annihilated by the function $\exp(-ik \cdot x)$. The momentum integration can then, in principle, be performed and the effective free energy takes the form of a spatial integral over a local density $F_{\text{eff}} = \int d^d x \mathcal{E}_{\text{eff}}$.

Applied to the formal expression (11), the derivative expansion yields the following quadratic terms in the effective free energy density:

$$\begin{aligned} \mathcal{E}_{\text{eff}}(|\Phi|) &= -e^4(d-1) \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2} \frac{1}{(k-i\partial)^2} |\partial_\mu \Phi|^2 \\ &\quad - (n+8)\lambda^2 \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2} \frac{1}{(k-i\partial)^2} |\Phi|^2 |\Phi|^2 \\ &\quad - e^2 \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2} P_{\mu\nu}(k) \frac{1}{(k-i\partial)^2} P_{\nu\mu}(k-i\partial) |\Phi|^2 |\Phi|^2. \end{aligned} \quad (15)$$

Assuming that the order parameter carries a momentum κ , we can replace all occurrences of the derivative $-i\partial_\mu$ with κ_μ .

It is important to note that infrared divergences are absent here, because κ_μ acts as an infrared cutoff. Without such a cutoff, as in the effective potential (3), the individual terms in the Taylor expansion cannot be integrated because of infrared divergences, and the entire series has to be summed, which can lead to nonanalytic contributions. From the perspective of RG, the presence of the momentum scale κ_μ is crucial as it allows studying the fixed point by letting that scale approach zero.

With the tree contribution $\mathcal{E}_{\text{tree}}$ added, we obtain after carrying out the integrals over the loop momentum

$$\begin{aligned} \mathcal{E}_{\text{tree}} + \mathcal{E}_{\text{eff}} &= [1 - c(d-1)\hat{e}^2] |\partial_\mu \Phi|^2 \\ &\quad + \left\{ \hat{\lambda} - c \left[(n+8)\hat{\lambda}^2 + d(d-1)\hat{e}^4 \right] \right\} \kappa^{4-d} |\Phi|^4, \end{aligned} \quad (16)$$

where $\hat{\lambda} = \lambda \kappa^{d-4}$ and $\hat{e}^2 = e^2 \kappa^{d-4}$ are the rescaled dimensionless coupling

constants. As an aside, by evaluating the integrals in fixed dimension $2 < d < 4$, and not in an ϵ expansion close to the upper critical dimension $d = 4$, we implement in effect Parisi's approach to critical phenomena [7]. The factor in front of the kinetic term in Eq. (16) amounts to a field renormalization. It can be absorbed by introducing the renormalized field $\Phi_r = Z_\phi^{-1/2}\Phi$, with Z_ϕ the field renormalization factor

$$Z_\phi = 1 + c(d-1)\hat{e}^2. \quad (17)$$

The effective free energy density then becomes

$$\mathcal{E}_{\text{tree}} + \mathcal{E}_{\text{eff}} = |\partial_\mu \Phi_r|^2 + \lambda_r |\Phi_r|^4, \quad (18)$$

where λ_r is the renormalized coupling

$$\frac{1}{\hat{\lambda}} = \frac{1}{\hat{\lambda}_r} - c \left[(n+8) - 2(d-1) \frac{\hat{e}_r^2}{\hat{\lambda}_r} + \frac{1}{4} d(d-1) \frac{\hat{e}_r^4}{\hat{\lambda}_r^2} \right]. \quad (19)$$

It was Kleinert's idea to put this equation in this form. The reason being that the critical point is approached by letting the momentum scale approach zero. Since the original coupling constant λ is fixed, the left side tends to zero when $\kappa \rightarrow 0$. The resulting quadratic equation then determines the value of $\hat{\lambda}_r$ at the critical point, provided we know the value of $\hat{e}_r^2$ there. This procedure is equivalent to finding the root of the flow equation for $\hat{\lambda}_r$ in conventional RG [8].

To obtain the value of $\hat{e}_r^2$ at criticality, we rescale the vector field $A_\mu \rightarrow A_\mu/e$ and consider it instead of the order parameter to be the background field. The 1-loop effective free energy, obtained after integrating out the scalar fields, reads:

$$F_{\text{eff}}[A] = \frac{n}{2} \text{Tr} \ln \left[1 + \frac{1}{p^2} (2p_\mu A_\mu - A_\mu^2) \right], \quad (20)$$

where we again ignored an irrelevant constant and used the gauge $\partial_\mu A_\mu = 0$. The second term in the expansion of the logarithm yields the first nonzero contribution

$$\mathcal{E}_{\text{eff}}(A) = n \int \frac{d^d k}{(2\pi)^d} \frac{k_\mu (k_\nu - i\partial_\nu)}{k^2 (k - i\partial)^2} A_\mu A_\nu. \quad (21)$$

If we assume that the vector field carries the same momentum κ_μ as does the order parameter, we obtain after carrying out the momentum integral for the

sum of the tree and the 1-loop contribution

$$\mathcal{E}_{\text{tree}} + \mathcal{E}_{\text{eff}} = -\frac{1}{2e_r^2} \partial^2 A_\mu A_\mu, \quad (22)$$

where e_r^2 is the renormalized coupling constant,

$$\frac{1}{\hat{e}_r^2} = \frac{1}{\hat{e}^2} + \frac{n}{2} \frac{c}{d-1}. \quad (23)$$

Keeping e fixed and letting $\kappa \rightarrow 0$ so as to approach the critical point, we see that the renormalized coupling tends to a constant value

$$\hat{e}_r^{*2} = \frac{2}{n} \frac{d-1}{c}. \quad (24)$$

When this constant is substituted in Eq. (19) with the left side set to zero, the resulting quadratic equation in $\hat{\lambda}_r$ has two real solutions

$$\hat{\lambda}_r^* = \frac{1}{2n(n+8)} \frac{1}{c} [n + 4(d-1)^2 \pm \Delta_d], \quad (25)$$

provided that the determinant

$$\Delta(d) := \sqrt{n^2 - 4(d-2)(d-1)^2(d+1)n - 16(d-1)^3(d+1)} \quad (26)$$

is real. For fixed dimensionality, this condition is satisfied only for a sufficient number of field components n . Specifically, the minimum number in $d = 2, 3$, and 4 is $n_c(2) = 4\sqrt{3} \approx 6.9$, $n_c(3) = 16(2 + \sqrt{6}) \approx 71.2$, and $[3]$ $n_c(4) = 12(15 + 4\sqrt{15}) \approx 365.9$. The plus sign in Eq. (25) corresponds to the infrared-stable fixed point, while the minus sign corresponds to the tricritical point, where the continuous phase transition changes to a discontinuous one. For a fixed number of field components n , this 1-loop result shows that continuous behavior is favored when $d \rightarrow 2$, while in the opposite limit, $d \rightarrow 4$, discontinuous behavior is favored.

5 Conclusions

For a conventional 3-dimensional superconductor corresponding to $n = 2$, the 1-loop result fails to give an infrared-stable fixed point. It should, however, be kept in mind that there are infinitely many loop diagrams, and that the values of the coupling constants at criticality are not particular small to justify low-order perturbation theory. For example, $\hat{e}_r^{*2} = 32/n$ according to Eq. (24)

with $d = 3$, which only for large n is small. Different approaches are therefore required to investigate the presence of a fixed point.

One of these is the dual approach [2,9,10], which is a formulation in terms of magnetic vortex loops. Monte Carlo simulations of the dual model led to the conclusion that 3-dimensional type-II superconductors undergo a continuous phase transition belonging to the XY universality class with an inverted temperature axis [11]. Using this approach, Kleinert [12] predicted the presence of a tricritical point at a value of the Ginzburg-Landau parameter $\kappa_{\text{tri}} \approx 0.8/\sqrt{2}$.

Another approach within the framework of the Ginzburg-Landau model itself was put forward in Ref. [13], where – in our language – the vector field was assumed to carry not the same momentum κ_μ as the order parameter, but κ_μ/x instead. The charged fixed points for $n = 2$ and $d = 3$ are then located at $\hat{e}_r^{*2} = 16/x$ which becomes small for large x , facilitating the existence of an infrared-stable fixed point. The free parameter x was determined by matching κ_{tri} with Kleinert's estimate or Monte Carlo simulations [14] which gave $\kappa_{\text{tri}} \approx 0.42/\sqrt{2}$. From Eqs. (24) and (25) one obtains, with the parameter x included [15],

$$\kappa_{\text{tri}}^2 = (8 + x + \sqrt{x^2 + 16x - 176})/40. \quad (27)$$

The resulting values for the critical exponents are consistent with the expected (inverted) XY universality class.

Also resummation techniques applied to results obtained to second order in the loop expansion [16] predict the existence of an infrared-stable fixed point in the Ginzburg-Landau model [17].

In conclusion, by extending the effective potential to include derivative terms, we achieved agreement with RG as they provide the theory with a low momentum scale which can be varied to probe the fixed point. The presence of a nonanalytic term in the effective potential at criticality is argued to be immaterial as far as the order of the phase transition is concerned.

Acknowledgments

I wish to thank B. Rosenstein for the kind hospitality at the Department of Electrophysics, National Chiao Tung University, Hsinchu, Taiwan, and for helpful discussions. This work was funded by the National Science Council (NCS) of Taiwan.

References

- [1] L.D. Landau and E.M. Lifshitz, *Statistical Physics*, Part 1 (Pergamon, New York, 1986).
- [2] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* (World Scientific, Singapore, 1989).
- [3] B.I. Halperin, T.C. Lubensky, and S.-K. Ma, *Phys. Rev. Lett.* **32**, 292 (1974).
- [4] H. Kleinert, *Phys. Lett. B* **128**, 69 (1983).
- [5] T. Appelquist and U. Heinz, *Phys. Rev. D* **25**, 2620 (1982).
- [6] C.M. Fraser, *Z. Phys. C* **28**, 101 (1985).
- [7] G. Parisi, Lectures presented at the Cargèse Summer School 1973 (unpublished); *J. Stat. Phys.* **23**, 23 (1980).
- [8] A.M.J. Schakel, *Boulevard of Broken Symmetries*, Habilitationsschrift, Freie Universität Berlin, eprint: cond-mat/9805152.
- [9] T. Banks, B. Meyerson, and J.B. Kogut, *Nucl. Phys. B* **129**, 493 (1977); M.E. Peskin, *Ann. Phys. (N.Y.)* **113**, 122 (1978); P.R. Thomas and M. Stone, *Nucl. Phys. B* **144**, 513 (1978).
- [10] M. Kiometzis and A.M.J. Schakel, *Int. J. Mod. Phys. B* **7**, 4271 (1993); M. Kiometzis, H. Kleinert, and A.M.J. Schakel, *Phys. Rev. Lett.* **73**, 1975 (1994); *Fortschr. Phys.* **43**, 697 (1995).
- [11] C. Dasgupta and B.I. Halperin, *Phys. Rev. Lett.* **47**, 1556 (1981).
- [12] H. Kleinert, *Lett. Nuovo Cim.* **35**, 405 (1982).
- [13] I.F. Herbut and Z. Tešanović, *Phys. Rev. Lett.* **76**, 4588 (1996).
- [14] J. Bartholomew, *Phys. Rev. B* **28**, 5378 (1983).
- [15] C. de Calan, A.P.C. Malbouisson, F.S. Nogueira, and N.F. Svaiter, *Phys. Rev. B* **59**, 554 (1999).
- [16] J. Tessmann, Diplomarbeit, Freie Universität Berlin (1984); S. Kolnberger and R. Folk, *Phys. Rev. B* **41**, 4083 (1990).
- [17] R. Folk and Yu. Holovatch, *J. Phys. A* **29**, 3409 (1996).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

SCALING AND DUALITY IN THE SUPERCONDUCTING PHASE TRANSITION

F.S. NOGUEIRA

*Institut für Theoretische Physik, Freie Universität Berlin,
Arnimallee 14, 14195 Berlin, Germany
E-mail: nogueira@physik.fu-berlin.de*

The field theoretical approach to duality in the superconducting phase transition is reviewed. Emphasis is given to the scaling behavior, and recent results are discussed.

1 Introduction

The renewed interest for critical fluctuations in superconductors is due to the enormous variety of interesting phenomena observed in the last decade. In a zero external field regime, non-classical values of the critical exponents and amplitude ratios were measured [1]. In the observed critical region we have $-0.03 < \alpha < 0$, $\nu \approx 0.67$ and $A_+/A_- \approx 1.065$. These values were measured for bulk samples of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ (YBCO). They indicate that the observed critical region corresponds to the XY universality class. In the non-zero field regime, the thermal fluctuations make the vortex lattice melt. For high fields this transition is known to be of first order [2].

Most thermal fluctuation effects in superconductors can be understood using the Ginzburg-Landau (GL) model. In zero field, the GL model is given by

$$L = \frac{1}{2}(\nabla \times \mathbf{A})^2 + |(\nabla - ie\mathbf{A})\phi|^2 + m^2|\phi|^2 + \frac{u}{2}|\phi|^4. \quad (1)$$

The different regimes of the GL model are controlled by the size of the Ginzburg parameter $\kappa = \sqrt{u/2e^2}$. In the weak-coupling regime ($\kappa \gg 1$) the

critical fluctuations in the GL model lead to a fluctuation induced first-order phase transition [3]. This scenario no longer holds in the strong-coupling regime ($\kappa \gg 1$), where a second-order phase transition takes place [4,5]. Usually it is very difficult to access the strong-coupling regime of the GL model through conventional perturbative methods. An alternative approach uses duality arguments. Duality allows us to transform a strong-coupling problem into a weak-coupling one. A well-known example is the 2d Ising model where such a transformation has an additional feature: Two-dimensionality makes the model self-dual, which leads to an exact determination of the critical temperature [6]. In three dimensions life is more complicated, but duality remains a powerful tool. In the case of the GL model, lattice duality studies [4] helped to conclude that the transition must be of second order in the strong-coupling regime. A deeper point of view was pioneered by Kleinert who developed a scaling (continuum) limit of the lattice dual model [5,7,8] which gives a field theoretic description of vortex lines. Using this field theoretical approach, Kleinert made the remarkable discovery that a tricritical point exists in the phase diagram of a superconductor. The tricritical point separates the first- and second-order phase transition regimes of the superconductor. We shall see in the next section that this discovery has far reaching consequences and is useful even in a nonzero field regime.

In zero field, the vortex lines are closed loops, and the corresponding field theory features a *disorder parameter* field ψ (as opposed to the *order parameter* field ϕ). The field ψ describes a grand canonical ensemble for vortex loops, and $|\psi|^2$ gives the vortex density. The duality transformation has transformed a field theory, where the basic objects are the Cooper pairs, into another one, where the basic objects are vortex lines. In the case of the GL model, currents interact through the electromagnetic vector potential $\mathbf{A}$, while in the disorder field theory they interact through a fluctuating field which is proportional to the *magnetic induction* field $\mathbf{h}$. The simplest example of duality is obtained in the London limit, where the amplitude fluctuations are frozen. There, the dual Lagrangian corresponds to the London model:

$$L_{\text{London}}^{\text{dual}} = \frac{1}{2}(\nabla \times \mathbf{h})^2 + \frac{m_A^2}{2}\mathbf{h}^2 + im_A \mathbf{J}_v \cdot \mathbf{h}, \quad (2)$$

where m_A is the photon mass and $\mathbf{J}_v$ is the vortex current. The full disorder field theory corresponds to a generalization of this model. Intuitively this can be done as follows: in the classical limit, thermal fluctuations are absent and, therefore, there are no vortex loops. The only way to create vortices

is by applying an external magnetic field. This creates vortex lines parallel to this field but no loops. The classical solution is well known in this case and corresponds to the Abrikosov vortex lattice [9]. In the context of the model (2), the external magnetic field couples linearly to the induction field $\mathbf{h}$. Thermal fluctuations create additional vortex loops. They are closed as a consequence of Ampere's law which gives $\nabla \cdot \mathbf{J}_v = 0$. In the disorder field theory of fluctuating vortex loops, the coupling $\mathbf{J}_v \cdot \mathbf{h}$ in (2) becomes a minimal coupling, and the result is Kleinert's dual Lagrangian [7,8]

$$L_d = \frac{1}{2}[(\nabla \times \mathbf{h})^2 + m_A^2 \mathbf{h}^2] + |(\nabla - ie_d \mathbf{h})\psi|^2 + m_\psi^2 |\psi|^2 + \frac{u_\psi}{2} |\psi|^4, \quad (3)$$

where $e_d = 2\pi m_A/e$ is the dual charge.

In the next sections we shall review recent results on Kleinert's model. We shall discuss the relation between the tricritical point as obtained from the dual model and the tricritical point in the original GL model. It will be shown that the tricritical point in the GL model is of a Lifshitz type [12]. In the GL model, the most important manifestation of the Lifshitz point is a negative sign of the anomalous dimension of the order parameter. Duality transforms the tricritical Lifshitz point into an ordinary tricritical point. As a consequence, the sign of the anomalous dimension of the disorder parameter will be positive.

In Section 3 we discuss the scaling behavior of the dual model, which allows for many possibilities of scaling due to the massiveness of the induction field [13]. Of course, only one scaling corresponds to the superconducting phase transition as obtained from the original GL model, that is, the transition which is governed by an infrared stable charged fixed point. The other scalings correspond to crossover regimes. In principle, all these regimes can be observed experimentally. Ironically, the true superconducting phase transition remains the most difficult to be observed experimentally. In fact, the critical region of this charge fluctuation regime is very small [7]. The regime very often probed is the XY regime [1] and we will see that there is a corresponding scaling in the dual Lagrangian that corresponds to it. A scaling that deserves some experimental attention is the Kleinert scaling [8] which is characterized by the *exact* value $\nu' = 1/2$ of penetration depth exponent. This scaling seems to be verified in two experiments using thin films of YBCO at optimal doping [14,15]. At optimal doping 3D fluctuations are still dominant even in thin films of YBCO and a three-dimensional model is still relevant.

2 Tricritical and Lifshitz Point

Let us briefly review Kleinert's discovery of the tricritical point. In order to make the discussion simple, we will take advantage of the intuitive point of view adopted in the introduction. The technical details can be found in the textbook of Kleinert [7] and in the seminal paper Ref. [5].

Let us integrate out the induction field $\mathbf{h}$ in Eq. (3). We obtain the following effective action:

$$S_{\text{eff}} = \frac{1}{2} \text{Tr} \ln [(-\partial^2 + m_A^2 + e_d^2 |\psi|^2) \delta_{\mu\nu} + (1 - 1/a) \partial_\mu \partial_\nu] \\ + \frac{e_d^2}{2} \int d^3 r \int d^3 r' j_\mu(\mathbf{r}) D_{\mu\nu}(\mathbf{r}, \mathbf{r}') j_\nu(\mathbf{r}') \\ + \int d^3 r \left(|\nabla \psi|^2 + m_\psi^2 |\psi|^2 + \frac{u_\psi}{2} |\psi|^4 \right), \quad (4)$$

where j_μ is the μ component of the current operator $\mathbf{j} = ie_d(\psi^\dagger \nabla \psi - \psi \nabla \psi^\dagger)$ and the limit $a \rightarrow 0$ must be taken at the end in order to enforce the constraint $\nabla \cdot \mathbf{h} = 0$. The kernel $D_{\mu\nu}(r, r')$ is the inverse of the operator inside the $\text{Tr} \ln$. Now we will perform a Landau expansion of the effective action (4). As usual, in this expansion we assume that ψ has no spatial variation. In this way we obtain the following free energy density:

$$\mathcal{F} = [m_\psi^2 + e_d^2 D_{0;\mu\mu}(0)] |\psi|^2 + \frac{1}{2} \left[u_\psi - e_d^4 \int \frac{d^3 k}{(2\pi)^3} \hat{D}_{0;\mu\nu}(k) \hat{D}_{0;\nu\mu}(k) \right] |\psi|^4 \\ + \frac{e_d^3}{3} \int \frac{d^3 k}{(2\pi)^3} \hat{D}_{0;\mu\lambda}(k) \hat{D}_{0;\lambda\delta}(k) \hat{D}_{0;\delta\mu}(k), \quad (5)$$

where $D_{0;\mu\nu}(\mathbf{r} - \mathbf{r}')$ is the kernel $D_{\mu\nu}(\mathbf{r}, \mathbf{r}')$ for $\psi = 0$ and $\hat{D}_{0;\mu\nu}(k)$ is its Fourier transform which is given by

$$\hat{D}_{0;\mu\nu}(k) = \frac{1}{k^2 + m_A^2} \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right). \quad (6)$$

After calculating the integral in the $|\psi|^4$ term in (5), we see that it will be negative if $u_\psi < 4\pi^3 m_A^3 / e^4$. When this happens, we have a first-order phase transition scenario. Thus, it is clear that the point $m_\psi^2 = 2\pi m_A^3 / e^2$, $u_\psi = 4\pi^3 m_A^3 / e^4$ corresponds to a tricritical point (we have redefined m_ψ^2 by absorbing a factor $e_d^2 \Lambda / \pi^2$, where Λ is the ultraviolet cutoff).

Now we can ask the following question: How does the tricritical point manifest itself in the GL model? From a RG point of view we have the

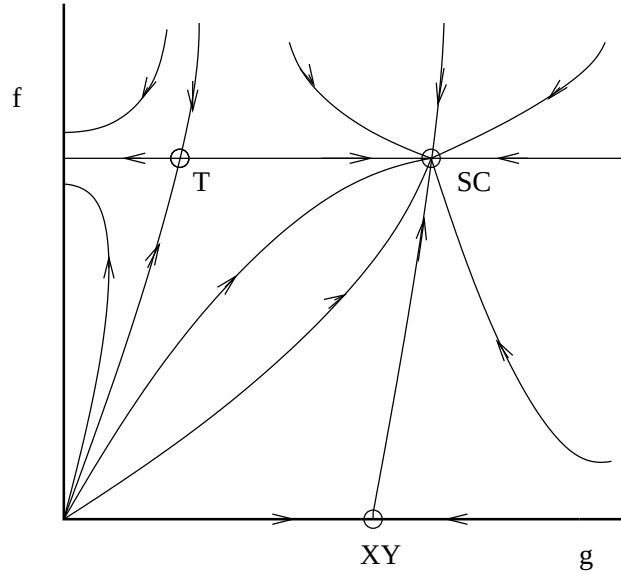


Figure 1. Flow diagram for the GL model. The labels T and SC are for tricritical and superconducting, respectively.

following scenario: let us define the renormalized dimensionless couplings $f = e_r^2/\mu$ and $g = u_r/\mu$, where e_r and u_r are the renormalized couplings and μ is a running scale. The fixed point structure in the gf -plane contains four fixed points [16–19]. Two of them are uncharged: the Gaussian fixed point corresponding to mean-field behavior of an uncharged superfluid, and the XY fixed point that governs the critical behavior of ^4He superfluid. The other two fixed points are charged: the infrared stable fixed point that governs the critical behavior of the superconductor and the tricritical fixed point, which is infrared stable along a line connecting the Gaussian and the tricritical fixed point, being unstable in the g -direction. The line connecting the Gaussian fixed point and the tricritical point is called the tricritical line. This line separates the regimes of first- and second-order phase transitions. The schematic flow diagram is shown in Fig. 1. The tricritical point obtained in Kleinert's duality map [5,7] corresponds to the tricritical fixed point of the RG picture. However, this tricritical point obtained directly in the GL model is of a different nature due to the local gauge symmetry. Indeed, we

have suggested recently [11] that the tricritical point of the GL model is of a Lifshitz type. This picture is founded on the behavior of the 2-point bare correlation function, which is given at 1-loop and $d = 3$ by

$$\tilde{W}^{(2)}(p) = \frac{1}{p^2 + m^2 + \Sigma(p)}, \quad (7)$$

with the self-energy

$$\Sigma(p) = -\frac{m}{2\pi}(u + e^2) - \frac{e^2}{4\pi|p|}(p^2 - m^2) \left[\frac{\pi}{2} + \arctan\left(\frac{p^2 - m^2}{2m|p|}\right) \right]. \quad (8)$$

In writing the above equations, we have absorbed in the bare mass a contribution with a linear dependence on the ultraviolet cutoff Λ . From Eq. (7) we see that $\tilde{W}^{(2)}(p)$ has a real pole in the critical regime ($m^2 = 0$) at a nonzero momentum, besides the usual pole at $p = 0$. In the above 1-loop calculation this pole is at $|p| = e^2/4$. This means that $\tilde{W}^{(2)}(p) < 0$ when $|p| < e^2/4$ and the bare 2-point correlation function violates the infrared bound $0 < \tilde{W}^{(2)}(p) \leq 1/p^2$, usually satisfied for pure scalar models [20]. This violation of the infrared bound explains the negative sign of the η exponent usually found in RG calculations [16–19]. The negativeness of the η exponent is also confirmed by recent Monte Carlo calculations [21]. Since the 2-point critical correlation function changes its sign, it follows that the same sign change happens in the 1-particle irreducible 2-point function $\tilde{\Gamma}^{(2)}(p)$, which is the coefficient of the quadratic term in the effective action Γ . This change of sign with momentum at the critical point is a behavior characteristic of a Lifshitz point [12]. It is worth to mention that in scalar models of Lifshitz points the sign of η is also negative for the dimension $d = d_c - 1$ where d_c is the corresponding critical dimension. For instance, a fixed-dimension calculation in a $1/N$ expansion gives, for the isotropic Lifshitz point in $d = 7$ ($d_c = 8$ in this case), $\eta_{I4} \approx -0.08/N$ [22].

The phase transition scenario that emerges is the following. The phase diagram in the $\kappa^2 - T$ plane contains three phases: the normal phase, the type I and the type II regime. The type I regime is separated from the type II regime by a line terminating at a tricritical Lifshitz point, the latter belonging to a line that separates the normal phase from the two other phases. The phase diagram is therefore quite similar to that of the so-called $R - S$ model [23]. In the $R - S$ model the phase diagram is drawn in the $X - T$ plane where $X = S/R$, S and R being the couplings of the model. The three phases of the $R - S$ model are paramagnetic, ferromagnetic and helical.

Thus, the $R - S$ model differs from ordinary magnets by the presence of a modulated regime for the order parameter, the helical phase. If κ^2 plays a role analogous to X , we see that the type II regime is analogous to the helical phase. Indeed, the type II regime should correspond to a modulated order parameter, as can be seen experimentally by applying a magnetic field, which leads to the formation of the Abrikosov vortex lattice. The type I regime, on the other hand, must be associated to the ferromagnetic phase since it corresponds to a uniform order parameter.

3 Kleinert's Scaling in the Dual Model

Let us discuss the renormalization of the dual model Eq. (3). An important feature of the dual model is the presence of the two mass scales m_ψ and m_A . This fact allows some freedom in the scaling of the model which is clarified in Ref. [13].

The scaling is defined by the behavior near T_c of the *bare* ratio

$$\kappa_d^2 = \frac{m_\psi^2}{m_A^2}. \quad (9)$$

It must be observed that $m_\psi^2 \sim t$, where t is the reduced temperature. Since the following argument is valid to all orders, we are assuming that the critical temperature contains already all the fluctuations. Thus, in the bare mass we are using a *renormalized* critical temperature.^a If we look at the scaling of the bare photon mass in the GL model, we see that we have also $m_A^2 \sim t$. This is the main motivation of what we will call Kleinert's scaling [8], where κ_d is constant, just like the Ginzburg constant κ in the GL model.

We define the renormalized fields as $\psi_r = Z_\psi^{-1/2} \psi$ and $\mathbf{h}_r = Z_h^{-1/2} \mathbf{h}$. From the Ward identities we conclude that the term $m_A^2 \mathbf{h}^2/2$ does not renormalize. Thus, we obtain $m_{A,r}^2 = Z_h m_A^2$. The renormalization of the remaining parameters is given by $m_{\psi,r}^2 = Z_m^{-1} Z_\psi m_\psi^2$, $u_{\psi,r} = Z_{u_\psi}^{-1} Z_\psi^2 u$ and $e_{d,r}^2 = Z_h e_d^2$. We observe from the renormalization of e_d^2 that the charge e is not renormalized in the dual model. Let us introduce the dimensionless renormalized couplings $f_d = e_d^2/m_\psi$ and $g_\psi = u_{\psi,r}$. In order to obtain the flow equations we need to differentiate the renormalized quantities with respect to $m_{\psi,r}$, keeping fixed the bare quantities that *do not depend on t* . The following flow equations are

^aFor instance, at 1-loop the critical temperature would be corrected by a term proportional to the ultraviolet cutoff.

easily obtained:

$$m_{\psi,r} \frac{\partial m_{A,r}^2}{\partial m_{\psi,r}} = (\eta_h + 2 + \eta_m - \eta_\psi) m_{A,r}^2, \quad (10)$$

$$m_{\psi,r} \frac{\partial f_d}{\partial m_{\psi,r}} = (\eta_h + 1 + \eta_m - \eta_\psi) f_d, \quad (11)$$

where the RG functions η_h , η_m and η_ψ are defined by

$$\eta_h = m_{\psi,r} \frac{\partial \ln Z_h}{\partial m_{\psi,r}}, \quad (12)$$

$$\eta_m = m_{\psi,r} \frac{\partial \ln Z_m}{\partial m_{\psi,r}}, \quad (13)$$

$$\eta_\psi = m_{\psi,r} \frac{\partial \ln Z_\psi}{\partial m_{\psi,r}}. \quad (14)$$

It is straightforward to see that the infrared stable fixed point corresponds to $f_d^* = 0$. Therefore, the correlation length exponent is simply given by $\nu \approx 0.67$. Near the infrared stable fixed point, Eq. (10) becomes

$$m_{\psi,r} \frac{\partial m_{A,r}^2}{\partial m_{\psi,r}} \approx \frac{1}{\nu} m_{A,r}^2. \quad (15)$$

This means that $m_{A,r}^2 \sim m_{\psi,r}^{1/\nu}$. Since $m_{A,r} = \lambda^{-1}$ we have the *exact* penetration depth exponent $\nu' = 1/2$. Thus, in Kleinert's scaling the exponent ν' remains classical to *all orders*.

Other scalings corresponding to different physical situations are also possible in the dual model. For example, the scaling that gives the *XY* universality class corresponds to taking m_A *fixed*, that is, with no dependence on t . In this scaling $\kappa_{d,r} \rightarrow 0$ as $t \rightarrow 0$ and the penetration depth exponent is given by $\nu' = \nu/2$. This scaling is well known experimentally [1].

Experimentally, Kleinert's scaling seems to have been probed by the authors of Ref. [15] and more recently in Ref. [14]. Both experiments used YBCO thin films at *optimal doping*. At optimal doping, 3D fluctuations are still more relevant, even in YBCO thin films, due to the strong coupling between the CuO planes.

4 Conclusion

Let us summarize the main lessons of this paper. First, the charged fixed point corresponds to a strong-coupling problem which is very difficult to be studied directly in the GL model where special perturbative or non-perturbative techniques must be employed. Duality gives in this case a powerful access towards the understanding of this problem since it provides a weak-coupling realization of the strong-coupling limit of the GL model. Second, the tricritical point of the superconductor is a Lifshitz point, and that is the reason why the anomalous dimension of the superconductor is negative.

An interesting perspective is the use of duality in a nonzero field problem. This approach is considerably more difficult since the phase diagram is richer. However, it can be hoped that the duality approach would also be helpful in this context.

Acknowledgments

The author would like to thank A. Pelster for his comments and suggestions. He acknowledges the Alexander-von-Humboldt foundation for financial support.

References

- [1] M.B. Salamon, J. Shi, N. Overend, and M.A. Howson, *Phys. Rev. B* **47**, 5520 (1993); N. Overend, M.A. Howson, and I.D. Lawrie, *Phys. Rev. Lett.* **72**, 3238 (1994); S. Kamal, D.A. Bonn, N. Goldenfeld, P.J. Hirschfeld, R. Liang, and W.N. Hardy, *Phys. Rev. Lett.* **73**, 1845 (1994).
- [2] E. Zeldov, E. Maier, M. Konczykowski, V.B. Geshkenbein, V.M. Vinokur, and H. Shtrikman, *Nature* (London) **375**, 373 (1995); M. Roulin, A. Junod, and E. Walker, *Science* **273**, 1210 (1996); for a review see G. Blatter, M.V. Feigel'man, V.B. Geshkenbein, A.I. Larkin, and V.M. Vinokur, *Rev. Mod. Phys.* **66**, 1125 (1994).
- [3] B.I. Halperin, T.C. Lubensky, and S.-K. Ma, *Phys. Rev. Lett.* **32**, 292 (1974); J.H. Chen, T.C. Lubensky, and D.R. Nelson, *Phys. Rev. B* **17**, 4274 (1978).
- [4] C. Dasgupta and B.I. Halperin, *Phys. Rev. Lett.* **47**, 1556 (1981).
- [5] H. Kleinert, *Lett. Nuovo Cim.* **35**, 405 (1982).
- [6] H.A. Kramers and G.H. Wannier, *Phys. Rev.* **60**, 252 (1941).

- [7] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [8] M. Kiometzis, H. Kleinert, and A.M.J. Schakel, *Phys. Rev. Lett.* **73**, 1975 (1994); *Fortschr. Phys.* **43**, 697 (1995), and references therein.
- [9] A.A. Abrikosov, *Soviet Phys. (JETP)* **5**, 1174 (1957).
- [10] Z. Tešanović, *Phys. Rev. B* **59**, 6449 (1999).
- [11] F.S. Nogueira, *Phys. Rev. B* **62**, 14559 (2000).
- [12] W. Selke, in *Phase Transitions and Critical Phenomena*, Vol. 15, Eds. C. Domb and J.L. Lebowitz (Academic Press, London, 1992), pp. 1-72, and references therein.
- [13] C. de Calan and F.S. Nogueira, *Phys. Rev. B* **60**, 4255 (1999).
- [14] K.M. Paget, B.R. Boyce, and T.R. Lemberger, *Phys. Rev. B* **59**, 6545 (1999).
- [15] Z.H. Lin *et al.*, *Europhys. Lett.* **32**, 573 (1995).
- [16] B. Bergerhoff *et al.*, *Phys. Rev. B* **53**, 5734 (1996).
- [17] R. Folk and Yu. Holovatch, *J. Phys. A* **29**, 3409 (1996).
- [18] I.F. Herbut and Z. Tešanović, *Phys. Rev. Lett.* **76**, 4588 (1996); I.D. Lawrie, *ibid.* **78**, 979 (1997); I.F. Herbut and Z. Tešanović, *ibid.* **78**, 980 (1997).
- [19] C. de Calan, A.P.C. Malbouisson, F.S. Nogueira, and N.F. Svaiter, *Phys. Rev. B* **59**, 554 (1999).
- [20] J. Fröhlich, B. Simon, and T. Spencer, *Commun. Math. Phys.* **50**, 79 (1976).
- [21] A.K. Nguyen and A. Sudbø, *Phys. Rev. B* **60**, 15307 (1999).
- [22] R.M. Hornreich, M. Luban, and S. Shtrikman, *Phys. Lett. A* **55**, 269 (1975).
- [23] S. Redner and H.E. Stanley, *Phys. Rev. B* **16**, 4901 (1977); *J. Phys. C* **10**, 4765 (1977).

THERMAL FLUCTUATIONS IN THE GROSS-NEVEU MODEL WITH $U(1)$ -SYMMETRY AT SMALL N

E. BABAEV

*Institute for Theoretical Physics, Uppsala University,
Box 803, S-75108 Uppsala, Sweden
E-mail: egor@teorfys.uu.se*

The chiral Gross-Neveu model is one of the most popular toy models for QCD. In the past, it has been studied in detail in the large- N limit. In this paper we study its small- N behavior at finite temperature in $2+1$ dimensions. We show that at small N the phase diagram of this model is *principally* different from its behavior at $N \rightarrow \infty$. For a small number N of fermions, the model possesses two characteristic temperatures T_{KT} and T^* . That is, at small N , along with a quasiordered phase $0 < T < T_{KT}$ the system possesses a very large region of precursor fluctuations $T_{KT} < T < T^*$ which disappear only at a temperature T^* , substantially higher than the temperature T_{KT} of the Kosterlitz-Thouless transition.

In this contribution we discuss the small- N behavior of the Gross-Neveu (GN) [1] model with $U(1)$ -symmetry in $2 + 1$ dimensions at finite temperature. The Gross-Neveu model is a field theoretic model of zero-mass fermions with quartic interaction, which provides us with considerable insight into the mechanisms of spontaneous symmetry breakdown and is considered to be an illuminating toy model for QCD. Our small- N study is motivated by recent results in the theory of superconductivity in the regimes, where BCS mean-field theory is not valid.

In the past years, remarkable progress was made in the theory of superconductivity in understanding mechanisms of symmetry breakdown in the regimes of strong interaction and low carrier density. It was observed that, in general, a Fermi system with attraction possesses two distinct characteristic temperatures corresponding to pair formation and pair condensation. That is, in a strong-coupling superconductor Cooper pairs are formed at a certain temperature T^* although there is no macroscopic occupation of zero momentum

level and thus no phase coherence and no symmetry breakdown. The temperature should be lowered to $T_c (\ll T^*)$ in order to make these pairs condense and establish phase coherence. The large region $T_c < T < T^*$ where there exist Cooper pairs but no continuous symmetry is broken is called *pseudogap phase* (see Ref. [2] and numerous references therein). Thus the symmetry breakdown in a strong-coupling superconductor resembles onset of long-range order in ^4He , where one can also introduce formally a characteristic temperature of thermal decomposition of a He atom. However it does not mean that this temperature is connected in any respect with the temperature of the onset of phase coherence. In fact, the BCS scenario, where the superconductive phase transition can approximately be described as a pair-formation transition, is very exceptional, since there is only one characteristic temperature T_c . That is, the BCS scenario holds true only at infinitesimally weak coupling strength or very high carrier density.

Since the original work by Bardeen, Cooper, and Schrieffer [3] in 1957, the BCS theory has been a source of inspiration in particle physics. In particular, it had direct influence on the construction of the Nambu-Jona-Lasinio and Gross-Neveu models [4,5]. In this spirit, the pseudogap concept was recently introduced to particle physics in Ref. [6], sparking recently many intensive discussions.

Let us now return to the Gross-Neveu model. In this article we would like to readdress the pseudogap properties of the Gross-Neveu model, extending the initial work of Ref. [6] to finite temperatures and higher dimensions. At finite temperatures in the limit of large N its behavior closely resembles a BCS superconductor [7]. We show that a very rich physical structure, similar to the phase diagram of a strong-coupling superconductor, emerges in the small- N limit in the chiral GN model.

The chiral GN model [1] has the following Lagrange density [8]

$$\mathcal{L} = \bar{\psi}_a i \not{\partial} \psi_a + \frac{g_0}{2N} \left[(\bar{\psi}_a \psi_a)^2 + (\bar{\psi}_a i \gamma_5 \psi_a)^2 \right],$$

where the index a runs from 1 to N . The fields $\psi(x)$ can be integrated out yielding the collective field action (for a detailed discussion see Ref. [9]):

$$\mathcal{A}_{\text{coll}}[\sigma, \pi] = N \left\{ -\frac{\sigma^2 + \pi^2}{2g_0} - i \text{Tr} \log [i \not{\partial} - \sigma(x) - i \gamma_5 \pi] \right\}.$$

This model is invariant under the continuous set of chiral $O(2)$ transformations which rotate σ and π fields into each other. In the large- N limit, the

model has a second-order phase transition at which fermions acquire mass. At zero temperature in 2+1 dimensions it is accompanied by an appearance of a massless composite Goldstone boson (for details see e.g. Ref. [9]). In the symmetry-broken phase the model is characterized by a typical "mexican hat" effective potential. The propagator of the massive σ fluctuations can be readily extracted and it coincides with the σ -propagator of the ordinary GN model [7,9]:

$$G_{\sigma'\sigma'} = -\frac{i}{N} \left[g_0^{-1} - i \operatorname{tr} \int \frac{d^3k}{(2\pi)^3} \frac{(\not{k} + M)(\not{k} - \not{q} + M)}{(k^2 - M^2)[(k - q)^2 - M^2]} \right]^{-1},$$

where M is the mass dynamically acquired by fermions. According to standard dimensional reduction arguments, the system is at finite temperature effectively two-dimensional and thus the Coleman theorem forbids the spontaneous breakdown of the $U(1)$ -symmetry. However, as it was shown by Witten [10], this does not preclude the system from generating a fermion mass. As it was shown in Ref. [10] by employing "modulus-phase" variables

$$\sigma + i\pi = |\Delta|e^{i\theta} \quad (1)$$

one can see that the system generates the fermion mass $M = |\Delta|$ that coincides with the modulus of the complex order parameter, but its phase remains incoherent and the correlators of the complex order parameter have algebraic decay. Existence of the *local* gap modulus Δ does not contradict the Coleman theorem since Δ is neutral under $U(1)$ transformations. Thus at low temperature in 2+1 dimensions there appears an "almost" Goldstone boson that becomes a real Goldstone boson at exactly zero temperature.

Let us here first study the effective potential of the model at finite temperature in the leading-order approximation and then take into account the next-to-leading-order corrections. Following Ref. [10], the fermion mass at finite temperature is given by the gap equation which coincides with the gap equation for the ordinary GN model with discrete symmetry (for detailed calculations see Refs. [7,6,9]):

$$1 = g_0 \operatorname{tr} (1) \int \frac{d^2k}{(2\pi)^2} \frac{1}{2E} \tanh \left(\frac{E}{2T} \right), \quad (2)$$

where E stands for $\sqrt{k^2 + \Delta^2}$. In the leading-order mean-field approximation we have a situation similar to the BCS superconductor: There is a gap that disappears at a certain temperature which we denote by T^* in the sequel. It

can be expressed via the gap function at zero temperature:

$$T^* = \frac{\Delta(0)}{2 \log 2}. \quad (3)$$

Near T^* the gap function has the following behavior in the mean-field approximation:

$$\Delta(T) = T^* 4 \sqrt{\log 2} \sqrt{1 - \frac{T}{T^*}}. \quad (4)$$

On the other hand, at low temperatures the gap function receives an exponentially small temperature correction:

$$\Delta(T) = \Delta(0) - 2T \exp \left[-\frac{\Delta(0)}{T} \right]. \quad (5)$$

Let us note once more that a straightforward calculation of next-to-leading-order corrections leads to the conclusion that the gap should be exactly zero at any finite temperature in 2+1 dimensions. However, as shown in Ref. [10], such a direct calculation of corrections misses the essential physics of a two-dimensional problem. The fluctuations can be made arbitrarily weak by decreasing temperature in 2+1 dimensions (or e.g. increasing N in 2D zero temperature calculations in Ref. [10]) and then the system possesses a very well-defined "mexican hat" effective potential that determines the fermion mass. Due to phase fluctuations in the degenerate valley of the potential, the average of the complex gap function is zero, however there exists a gap locally (i.e. in some sense the system in its low energy domain degenerates to a nonlinear sigma model). In 2+1 dimensions, as the temperature approaches zero the thermal fluctuations in the degenerate valley of the effective potential gradually vanish and at $T = 0$ a local gap becomes a real gap. The most interesting effect happens however when temperature is increased. It was anticipated before that there is only one characteristic temperature in such a system, namely the temperature of the Kosterlitz-Thouless (KT) transition which coincides with the temperature of the formation of the local gap. This scenario holds true only for a very large number of field components. In general, the model has two characteristic temperatures like in the case of a superconductor with a pseudogap. In order to show this we have to go beyond the mean-field approximation.

Let us make an expansion around a saddle-point solution and derive the propagator of the imaginary part of the field Δ that has a pole at $q^2 = 0$.

The procedure is standard and will not be reproduced here (for details see e.g. Refs. [6,9]):

$$G_{\Delta'_{\text{im}}\Delta'_{\text{im}}} = \frac{1}{N} \left[\frac{1}{8\pi\Delta(T)} \tanh\left(\frac{\Delta(T)}{2T}\right) \right]^{-1} \frac{i}{q^2}. \quad (6)$$

The propagator (6) characterizes the stiffness of the phase fluctuations in the degenerate minimum of the effective potential. It gives the following expression for the kinetic term of phase fluctuations for the chiral GN model:

$$E_{\text{kin}} = \int d^2x \frac{N}{8\pi} \Delta(T) \tanh\left(\frac{\Delta(T)}{2T}\right) [\nabla\theta]^2. \quad (7)$$

Now we have all the tools to find the position of the KT transition in the chiral GN model. It is well known that the KT transition cannot be found by straightforward perturbative methods. In order to find the KT transition point one should first observe that the system is described by a complex scalar field. The key feature is that the field θ is cyclic: $\theta = \theta + 2\pi n$. In two dimensions such a system possesses excitations of the form of vortices and antivortices that are attracted to each other by a Coulomb potential. At low temperatures, the vortices and antivortices form bound pairs. The grand-canonical ensemble of the pairs exhibits quasi-long-range correlations. At a certain temperature T_{KT} , the vortex pairs break up, which is the Kosterlitz-Thouless phase-disorder transition [11,12]. This transition was studied in detail in the field theory of a pure phase field $\theta(x)$, with a Hamiltonian

$$\mathcal{H} = \frac{\beta}{2} [\partial\theta(x)]^2, \quad (8)$$

where β is the stiffness of the phase fluctuations. In our case the coefficient β depends on the temperature and on the parameters of the GN models, namely the number of field components and Δ [see Eq. (7)].

The temperature of the Kosterlitz-Thouless phase transition in the system (8) is given by Ref. [11,12]:

$$T_{KT} = \frac{\pi}{2} \beta. \quad (9)$$

In order to study the phase-disorder transition in the chiral GN model, we should solve a *set* of equations, namely the equation for $T_{KT}(\Delta, N)$ that follows from the kinetic term, and Eq. (2) for the gap modulus $\Delta(T_{KT}, N)$ that follows from the effective potential. Thus in our case the phase transition

is a competition of two processes, the thermal excitation of directional fluctuations in the degenerate valley of the effective potential and the thermal depletion of the stiffness coefficient.

Let us first consider the case of small N . From expressions (3), (5), (7), and (9) we see that, in the regime of small N , $T_{KT} \ll T^*$. In this regime the temperature corrections to the phase stiffness are exponentially suppressed. Thus, at temperatures $T \ll T^*$, the asymptotic expression for the kinetic term (7) reads

$$H_{\text{kin}} = \int d^2x \frac{N}{8\pi} \Delta(0) [\nabla\theta]^2, \quad (10)$$

and the Kosterlitz-Thouless transition will take place at the temperature

$$T_{KT} = \frac{N}{8} \Delta(0). \quad (11)$$

This is significantly lower at small N than the temperature (3) at which the gap modulus disappears. For the ratio T_{KT}/T^* at small N we obtain:

$$\frac{T_{KT}}{T^*} = \frac{N \log(2)}{4}. \quad (12)$$

So with decreasing N , the separation of T_{KT} and T^* increases. Let us now turn to the regime where N is no longer small. Here we see from Eqs. (3), (4), (7), and (9) that T_{KT} tends to T^* from below. The Hamiltonian (7) reads in this limit near T^* :

$$H_{\text{kin}} = \int dx \frac{N}{16\pi} \frac{\Delta(T)^2}{T} [\nabla\theta]^2. \quad (13)$$

From Eqs. (3), (4), (13), and (9) we find the following expression for the behavior of T_{KT} at large N :

$$T_{KT} \simeq T^* \left(1 - \frac{1}{N \log(2)} \right). \quad (14)$$

This equation explicitly shows a merging of the temperatures T_{KT} and T^* in the limit of large N . This can be interpreted as the restoration of the "BCS-like" scenario for the quasi-condensation in the limit $N \rightarrow \infty$. The ratio T_{KT}/T^* is displayed in Fig. 1.

Thus the phase diagram of the model at *small* N consists of the following phases at non-zero temperature:

- (i) $0 < T < T_{KT}$: the low temperature quasi-ordered phase with bound

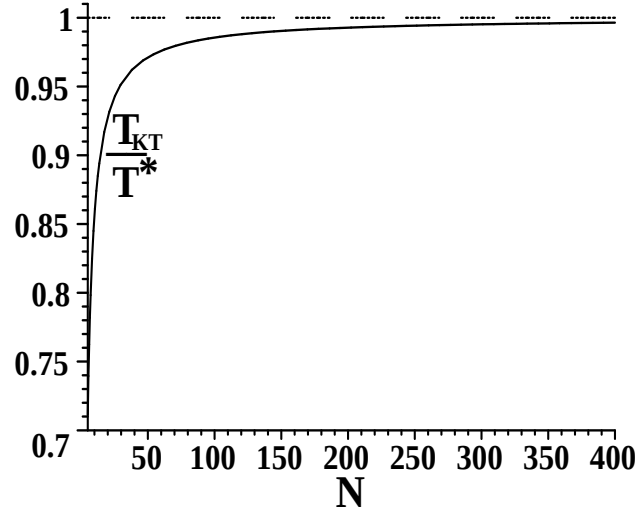


Figure 1. Recovery of a “BCS-like” scenario for quasicondensation at large N in the chiral GN model. The solid curve is the ratio of the temperature of the KT transition (T_{KT}) and the characteristic temperature of the formation of the effective potential (T^*). This ratio tends from below to unity (the horizontal dashed line) as N is increasing and the region of precursor fluctuations shrinks.

vortex-antivortex pairs,

(ii) $T_{KT} < T < T^*$: the phase analogous to the *pseudogap phase* of superconductors, i.e. the chirally symmetric phase with unbound vortex-antivortex pairs that exhibit violent precursor fluctuations and a nonzero local modulus of the complex gap function,

(iii) $T > T^*$: high temperature “normal” chirally symmetric phase.

The mechanism of the phase separation is very transparent with the key feature being the fact that the stiffness is proportional to N [see Eq. (7)]. At large N , the directional fluctuations are energetically extremely expensive and thus, the phase transition is controlled basically by the modulus of the order parameter. On the other hand, the stiffness is low at small N , and the thermal excitation of the fluctuations in the degenerate valley of the effective potential starts governing the phase transition in the system.

Now let us briefly discuss the physical meaning of T^* and what is expected to happen when the system reaches it at small N . At first, we can conclude from simple physical reasoning in analogy with superconductivity that the

appearance of the second characteristic temperature is very natural. Besides the fact that the phase analogous to the intermediate phase $T_{KT} < T < T^*$ is the dominating region on a phase diagram of strong-coupling and low carrier density superconductors, similar effects are known in a large variety of condensed matter systems such as excitonic condensates, Josephson junction arrays and many other systems. One of the most illuminating examples of the appearance of the pseudogap phase is the chiral GN model in $2 + \epsilon$ dimensions at zero temperature where this phenomenon is governed by quantum dynamical fluctuations at small N [6]. In $D = 2 + \epsilon$ the presence of two small parameters in the problem has allowed to prove the existence and the different physical origin of two characteristic values of a renormalized coupling constant and of the formation of an intermediate pseudogap phase [6]. We can also observe that the mean-field approximation gives a second-order phase transition at T^* which is certainly an artifact since much above T_{KT} there are violent thermal phase fluctuations. These fluctuations should wash out the phase transition at T^* which should degenerate to a smeared crossover as it happens in superconductors. Apparently, this crossover cannot be studied adequately in the framework of an $1/N$ -expansion. The best insight into the properties of the system in the region $T_{KT} < T < T^*$ can be obtained by numerical simulations. Although the KT transition is very hard to observe in simulations, the hint for the phase separation would be a gradual degradation of the transition at T^* with decreasing N .

Acknowledgments

The author gratefully acknowledges Prof. H. Kleinert for many discussions and for providing the access to the draft version of his forthcoming book [9].

References

- [1] D. Gross and A. Neveu, *Phys. Rev. D* **10**, 3235 (1974).
- [2] E. Babaev and H. Kleinert, *Phys. Rev. B* **59**, 12083 (1999).
- [3] J. Bardeen, L.N. Cooper, and J.R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).
- [4] Y. Nambu and G. Jona Lasinio, *Phys. Rev.* **122**, 345 (1961); **124**, 246 (1961); V.G. Vaks and A.I. Larkin, *JETP (Sov. Phys.)* **13**, 979 (1961); A.A. Anselm, *JETP (Sov. Phys.)* **9**, 608 (1959).
- [5] H. Kleinert, *On the Hadronization of Quark Theories*, Lectures presented

at the Erice Summer Institute 1976, in *Understanding the Fundamental Constituents of Matter*, Ed. A. Zichichi (Plenum Press, New York, 1978), pp. 289-390.

- [6] H. Kleinert and E. Babaev, *Phys. Lett. B* **438**, 311 (1998), eprint: hep-th/9809112.
- [7] B. Rosenstein, B.J. Warr, and S.H. Park, *Phys. Rep.* **205**, 59 (1991).
- [8] In 2 + 1 dimensions we choose γ -matrices as in the review Ref. [7]:
$$\gamma^\mu = \sigma^\mu \otimes \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \text{ and } i\gamma_5 = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$
- [9] H. Kleinert, *Particles and Quantum Fields*, Lecture notes at FU Berlin.
- [10] E. Witten, *Nucl. Phys. B* **145**, 110 (1978).
- [11] V.L. Berezinskii, *Zh. Eksp. Teor. Fiz.* **59**, 907 (1970); J. Kosterlitz and D. Thouless, *J. Phys. C* **6**, 1181 (1973).
- [12] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

FLEX-THEORY FOR HIGH- T_c SUPERCONDUCTIVITY DUE TO SPIN FLUCTUATIONS

D. MANSKE AND K.H. BENNEMANN

Institut für Theoretische Physik, Arnimallee 14, D-14195 Berlin, Germany

E-mails: dmanske@physik.fu-berlin.de, khb@physik.fu-berlin.de

Using the Hubbard model we develop a microscopic theory for high-temperature superconductivity due to the exchange of antiferromagnetic spin fluctuations. We treat the corresponding pairing mechanism self-consistently within the framework of the FLuctuation EXchange (FLEX) approximation and study some extensions. Solving the generalized Eliashberg equations for hole- and electron-doped superconductors we obtain both phase diagrams, respectively, and always a d -wave gap function. Furthermore, for hole-doped cuprates we find three characteristic temperature scales which are in qualitative agreement with the experimental situation: a pseudogap temperature T^* , below which a gap opens in the density of states, a mean-field transition temperature T_c^* for superconductivity below which we obtain Cooper-pairs without long-range phase coherence ("pre-formed Cooper-pairs"), and a critical temperature T_c , where these pairs become phase coherent.

1 Introduction

One of the most important and fascinating fields in condensed matter physics is the appearance of unconventional superconductivity, in particular high- T_c superconductivity, in which the underlying mechanism is still under debate, even 15 years after the discovery by Bednorz and Müller [1]. In hole-doped superconductors the highest transition temperature T_c (without applying pressure), namely $T_c = 133$ K, has been measured in $\text{HgBa}_2\text{Ca}_2\text{Cu}_3\text{O}_{8+\delta}$, followed by (to name just a few) $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ ($\delta = 0.15 \leftrightarrow T_c = 95$ K), $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ ($x = 0.93 \leftrightarrow T_c = 93$ K), and $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ where for a doping concentration $x = 0.15$, a maximum value of $T_c = 39$ K occurs.

While hole-doped superconductors have been studied intensively the analysis of electron-doped cuprates remains largely unclear. Of course, it is of high

interest to see whether the behavior of hole-doped and electron-doped cuprate superconductors can be explained within a unified physical picture using for example the exchange of antiferromagnetic spin fluctuations as the relevant pairing mechanism. If Cooper-pairing is controlled by antiferromagnetic spin fluctuation one expects on general physical grounds that *d*-wave symmetry pairing should also occur for electron-doped cuprates. Previous experiments have not clearly supported this, reporting mainly *s*-wave pairing [2,3]. Maybe as a result, so far electron-doped cuprates received much less attention than hole-doped cuprates. However, recently phase sensitive experiments [4] and magnetic penetration depth measurements [5,6] exhibited *d*-wave symmetry Cooper-pairing also for electron-doped cuprates.

In general, all high- T_c superconductors discovered so far contain CuO_2 -planes and various metallic elements. Hence, they are often called cuprates. Their crystal structure resembles that of the perovskites. It is now mainly established that the relevant physics related to superconductivity occurs in the CuO_2 -planes and that the other layers simply act as charge-reservoirs. Thus, the coupling in the *c*-direction provides a three-dimensional superconducting state but the main pairing interaction acts between carriers within a CuO_2 -plane. As mentioned above, T_c for hole-doped cuprates is of the order of 100 K and thus much larger than in conventional strong-coupling superconductors like lead ($T_c = 7.2$ K) or niobium ($T_c = 9.25$ K). The phenomenon of high- T_c superconductivity occurs for hole- and electron-doped cuprates in the vicinity of an antiferromagnetic phase transition. This suggests a purely electronic or magnetic mechanism in contrast to the conventional picture of electrons paired through the exchange of phonons. The simplest idea to explain such high critical temperature might be to introduce a higher cut-off energy ω_c due to electronic correlations in the system instead of integrating over an energy shell of ω_D (Debye frequency), i.e. $T_c \propto \omega_c \exp(-1/\lambda)$, where λ denotes the usual coupling strength for a given symmetry of the gap function. In the BCS theory [7] λ is equal to $N(0)V$, where $N(0)$ is the density of states (per spin) at the Fermi level and where $V = \text{const.}$ is the attractive pairing potential in $\mathbf{k}$ -space acting between electrons leading to the superconducting instability of the normal state. If the relevant energy cut-off ω_c of the problem is of the order of *electronic* degrees of freedom, e.g. $\omega_c \simeq 0.3 \text{ eV} \approx 250 \text{ K}$ [8], one can easily obtain a transition temperature of the order of 100 K. However, as we will discuss below, in a more realistic treatment the relation between T_c and λ is, of course, not that simple.

Many researchers believe that in order to find the origin of the high- T_c

superconductivity in the cuprates it is necessary to investigate their normal state as a function of the doping concentration. Therefore, phenomenological models like the Marginal-Fermi-Liquid (MFL) [9], the Nested-Fermi-Liquid (NFL) [10,11] and the Nearly Antiferromagnetic Fermi Liquid (NAFL) theory [12] have been developed in order to understand the unusual non-Fermi liquid properties in the normal state. We will see later that our FLEX-theory provides a microscopic justification for these theories. In particular, in the underdoped regime of high- T_c superconductors, i.e. closer to the antiferromagnetic transition than for optimal doping, a number of physical quantities exhibit quite unusual properties. Examples are the ^{63}Cu spin-lattice relaxation rate and the inelastic neutron scattering intensity: while in the overdoped regime $1/T_1T$ increases monotonously as T decreases down to T_c , one finds in the underdoped case that $1/T_1T$ passes through a maximum at a temperature T^* (for decreasing T) [13]. These results are fully corroborated by inelastic neutron scattering data, where in the underdoped regime $\text{Im } \chi(\mathbf{Q}, \omega)$ at fixed small ω ($\simeq 10\text{-}15$ meV) passes also through a maximum at T^* for decreasing T [14]. In addition, angle-resolved photoemission experiments [15,16] on underdoped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ indicate a presence of a gap with $d_{x^2-y^2}$ -wave symmetry above T_c also in the charge-excitation-spectrum even up to room temperature. Recently, several experiments including heat capacity [17], transport [18], Raman scattering [19], and, in particular, scanning tunneling microscopy [20,21] indicate the existence of a gap in the excitation spectrum of the single-particle properties. This gap in the charge-response of the system occurs at the same temperature where also the spin gap opens. Therefore, this gap is then called “(weak) pseudogap”. One believes both gaps might have the same origin. Due to the d -wave symmetry of the pseudogap many researchers believe that it is related to precursor effects of the superconducting state [22,23]. In connection with simple arguments on the quasi two-dimensional nature of the system (see e.g. the theorem of Hohenberg [24]) these ideas suggest a non-trivial mechanism of the unusual behavior of underdoped cuprates. Thus, one of the main theoretical questions is to explain the origin of this weak pseudogap in the normal state and its relation to the underlying pairing mechanism.

Another fundamental problem to solve is the theoretical determination of the superconducting transition temperature T_c itself. A schematic phase diagram for hole-doped superconductors is shown in Fig. 1. At around $x = 0.15$ one finds the highest T_c values. This region is called optimal doping. In the overdoped region, i.e. $x > 0.15$, many experimental data suggest that

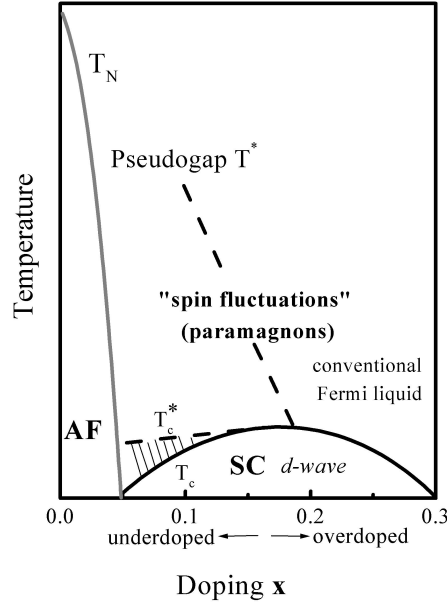


Figure 1. Schematic phase diagram of hole-doped cuprates [25]. High- T_c superconductivity occurs in the vicinity of an antiferromagnetic phase transition. The corresponding superconducting order parameter is of d -wave symmetry. In the overdoped region, i.e. $x > 0.15$, the system behaves like a conventional Fermi liquid, whereas in the underdoped regime below the pseudogap temperature T^* one finds strong antiferromagnetic (AF) correlations. As we will discuss below, Cooper-pairing can mainly be described by the exchange of AF spin fluctuations (often denoted as paramagnons) which are present everywhere in the system. The doping region between T_c and T_c^* (shaded region) is due to local Cooper-pair formation. Below T_c these pairs become phase coherent.

the system is a conventional Fermi liquid. On the underdoped side of the phase diagram in contrast it is believed that below a mean-field transition temperature T_c^* one finds pre-formed Cooper-pairs without long-range phase coherence. This part of the phase diagram is sometimes called the "strong pseudogap" region. Below T_c these pairs become phase coherent and a Meissner effect of the bulk material is observed. Furthermore, in the experiment $T_c \propto n_s$ is found only in underdoped superconductors [26]. So far, this has been mainly described in terms of the Ginzburg-Landau theory [22,23,27–29]. Recently, a microscopic calculation confirmed and clarified the picture [30].

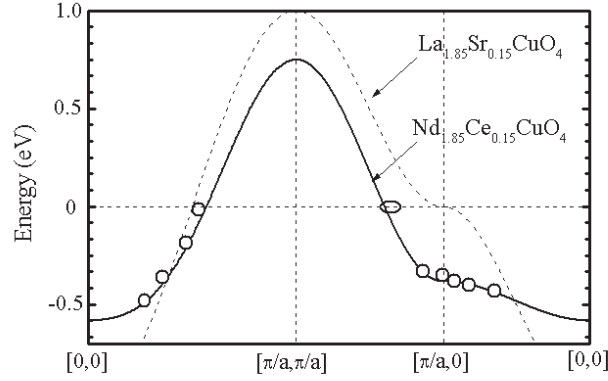


Figure 2. Results of the energy dispersion ϵ_k of optimally hole-doped $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ (dashed line) and of optimally electron-doped $\text{Nd}_{1.85}\text{Ce}_{0.15}\text{CuO}_4$ (NCCO). The solid curve refers to our tight-binding calculation choosing $t = 138$ meV and $t' = 0.3$. Data (open dots) are taken from Ref. [31]. The dashed curve corresponds to using $t = 250$ meV and $t' = 0$ and is typical for hole-doped cuprates.

2 Theory of Cooper-Pairing by Antiferromagnetic Spin Fluctuations

2.1 Hubbard Model

In order to obtain a unified theory for both hole-doped and electron-doped cuprates we use the same one-band Hubbard Hamiltonian taking into account the different dispersions for the carriers [31]. When doping the electrons, they occupy copper d -like states of the upper Hubbard band while the holes refer to oxygen-like p -states yielding different energy dispersion as used in our calculations. Thus, assuming similar itinerancy of the electrons and holes, the mapping on an effective one-band model seems to be justified. We consider U as an effective Coulomb interaction.

On a square lattice the Hamiltonian H reads in second quantization

$$H = - \sum_{\langle ij \rangle \sigma} t_{ij} (c_{i\sigma}^\dagger c_{j\sigma} + c_{j\sigma}^\dagger c_{i\sigma}) + U \sum_i n_{i\uparrow} n_{i\downarrow} - \mu t \sum_{i\sigma} n_{i\sigma}, \quad (1)$$

where $c_{i\sigma}^\dagger$ ($c_{i\sigma}$) creates (annihilates) an electron on site i with spin σ and t_{ij} is a hopping matrix element. The sum performed over nearest neighbors is denoted by $\langle ij \rangle$. Then, t_{ij} is equal to t . U is the intra-orbital (i.e. on-site) Coulomb repulsion and $n_{i\sigma}$ is equal to $c_{i\sigma}^\dagger c_{i\sigma}$. μ denotes the chemical

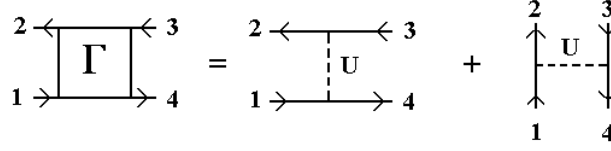


Figure 3. Antisymmetric four-point vertex-function $\Gamma(1, 2, 3, 4)$. The solid lines represent the electron propagators and the dashed line represents the on-site interaction U , respectively.

potential. Therefore, this model can be characterized by two dimensionless parameters, namely U/t and μ .

Using Bloch wave functions we rewrite Eq. (1) as

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma} + \frac{1}{2} \frac{U}{N} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}, \sigma} c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}', -\sigma}^{\dagger} c_{\mathbf{k}'+\mathbf{q}, -\sigma} c_{\mathbf{k}-\mathbf{q}, \sigma}, \quad (2)$$

where the one-band electron dispersion in the normal-state $\epsilon_{\mathbf{k}}$ reads for its nearest neighbor

$$\epsilon_{\mathbf{k}} = -2t [\cos k_x - \cos k_y + \mu/2] \quad (3)$$

and for next-nearest neighbor hopping

$$\epsilon_{\mathbf{k}} = -2t [\cos k_x - \cos k_y - 2t' \cos k_x \cos k_y + \mu/2], \quad (4)$$

respectively. Here, N is the number of lattice sites. As mentioned earlier, the bandstructure in Eq. (3) describes the Fermi surface of $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ [32], whereas for $B = 0.45$, Eq. (4) resembles the Fermi surface of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ [33].

In order to discuss the dispersion relation for electron-doped cuprates in more detail, we show experimental results in Fig. 2 as well as our tight-binding calculation. We choose the parameters $t = 138$ meV and $t' = 0.3$. For comparison we also show the results with $t = 250$ meV and $t' = 0$ which are often used to describe the hole-doped superconductor $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. One immediately sees the important difference: in the case of NCCO the flat band is approximately 300 meV *below* the Fermi level, whereas for the hole-doped case the flat band lies very close to it. Thus, as discussed later, using the resulting ϵ_k in a spin-fluctuation-induced pairing theory we get a smaller T_c for electron-doped cuprates than for the hole-doped ones.

2.2 Pairing Theory

In contrast to the usual Eliashberg theory of strong-coupling superconductors [34,35] in which phonons are involved one has to develop a theory for the exchange of AF spin fluctuations. One possibility is to use the phenomenological ansatz originally introduced by Millis, Monien, and Pines [12] where the effective pairing interaction $V_{eff}(\mathbf{q}, \omega)$ is of the Ornstein-Zernike form. However, we believe that a self-consistent description is required because electronic degrees of freedom not only condense into Cooper-pairs, but also create the pairing interaction. Thus, one has to generalize the Eliashberg equations on a microscopic basis.

As we have seen earlier, the nearness to a spin-density-wave instability corresponds to a simple physical picture, in which the spins have short-range antiferromagnetic order surviving from the long-range order of the insulating phase. In terms of the Hubbard Hamiltonian the exchange of longitudinal and transverse spin fluctuations gives rise to an effective electron-electron interaction (originally introduced by Berk and Schrieffer [36]) that provides a pairing interaction leading to d -wave superconductivity near half filling. The effective paramagnon-like interaction then reads

$$V_{eff}(\mathbf{q}, i\nu_m) = \frac{3}{2}U^2 \frac{\chi_0(\mathbf{q}, i\nu_m)}{1 - U\chi_0(\mathbf{q}, i\nu_m)} - \frac{1}{2}U^2 \frac{\chi_0(\mathbf{q}, i\nu_m)}{1 + U\chi_0(\mathbf{q}, i\nu_m)}. \quad (5)$$

In order to solve the generalized Eliashberg equations we will use a self-consistent theory called FLuctuation-EXchange (FLEX) approximation [37,38]. Remember that the Hubbard-Hamiltonian can be rewritten in the form $H = H_0 + H_{int}$, where H_0 describes the one-particle properties and H_{int} denotes a perturbation [39]. Let us start with introducing the antisymmetric four-point vertex-function $\Gamma(1, 2, 3, 4)$ (see Fig. 3) [40] as $\Gamma(1, 2, 3, 4) = V(1, 2, 3, 4) - V(1, 4, 3, 2)$. The antisymmetry corresponds to a sign change in Γ after permuting the creation (annihilation) operators 2 (1) and 4 (3) due to the fact that no distinction can be made between identical electrons. With the help of Γ it is now possible to write down the corresponding self-energies of the conducting holes or electrons:

$$\begin{aligned}\Sigma^G &= \text{diagram 1} + \text{diagram 2} + \text{diagram 3} + \dots \\ \Sigma^F &= \text{diagram 4} + \text{diagram 5} + \text{diagram 6} + \dots\end{aligned}$$

The diagrams represent various self-energy contributions. Diagram 1 is a single vertex Γ with a loop. Diagram 2 is two vertices Γ connected by a horizontal line with a loop above. Diagram 3 is three vertices Γ connected by two horizontal lines with a loop above. Diagram 4 is a single vertex Γ with a loop. Diagram 5 is two vertices Γ connected by a horizontal line with a loop below. Diagram 6 is three vertices Γ connected by two horizontal lines with a loop below.

Here, Σ^G (Σ^F) corresponds to the normal (anomalous) Green's function contribution. In order to solve these self-energy equations it is convenient to introduce the T -matrix [41] which is defined as

$$\text{diagram 7} = \text{diagram 8} + \text{diagram 9} + \text{diagram 10}$$

Diagram 7 is a box labeled T with four external lines (1, 2, 3, 4). Diagram 8 is a box labeled Γ with four external lines (1, 2, 3, 4). Diagram 9 is a box labeled Γ connected to a box labeled T by two horizontal lines, with four external lines (1, 2, 3, 4). Diagram 10 is a box labeled Γ connected to a box labeled Γ by two horizontal lines, with four external lines (1, 2, 3, 4).

Within the Nambu notation one arrives after a straightforward calculation at

$$\begin{aligned}\Sigma^G(\mathbf{k}, i\omega_n) &= \frac{1}{\beta N} \sum_{\mathbf{k}', i\omega'_n} \left[\frac{1}{2} U^2 \frac{\chi_{s0}(\mathbf{q}, i\nu_m)}{1 + U\chi_{s0}(\mathbf{q}, i\nu_m)} + \frac{3}{2} U^2 \frac{\chi_{s0}(\mathbf{q}, i\nu_m)}{1 - U\chi_{s0}(\mathbf{q}, i\nu_m)} \right. \\ &\quad \left. + U^2 \chi_G(\mathbf{q}, i\nu_m) + U \right] G(\mathbf{k}', i\omega'_n), \\ \Sigma^F(\mathbf{k}, i\omega_n) &= -\frac{1}{\beta N} \sum_{\mathbf{k}', i\omega'_n} \left[\frac{1}{2} U^2 \frac{\chi_{s0}(\mathbf{q}, i\nu_m)}{1 + U\chi_{s0}(\mathbf{q}, i\nu_m)} - \frac{3}{2} U^2 \frac{\chi_{s0}(\mathbf{q}, i\nu_m)}{1 - U\chi_{s0}(\mathbf{q}, i\nu_m)} \right. \\ &\quad \left. - U^2 \chi_F(\mathbf{q}, i\nu_m) - U \right] F(\mathbf{k}', i\omega'_n),\end{aligned}\quad (6)$$

where $i\nu_m = i\omega_n - i\omega'_n$. Note that the term $U^2 \chi_{G,F}$ on the right-hand side compensates double counting that occurs in the second order. χ_{s0} and χ_{c0} denote the irreducible spin- and charge-susceptibility, respectively, and are given by $\chi_{s0}(q) = -\sum_k [G(k+q)G(k)F(k+q)F^\dagger(-k)]/\beta N$ and $\chi_{c0}(q) = -\sum_k [G(k+q)G(k)F(k+q)F^\dagger(-k)]/\beta N$, where $q = (\mathbf{q}, i\nu_m)$ and $k = (\mathbf{k}, i\omega_n)$. The evaluation of the equations mentioned above are performed on the real axis [38,42]. In order to determine the superconducting transition temperature T_c we solve the linearized gap equation. Below T_c we find that the superconducting gap function has $d_{x^2-y^2}$ -wave symmetry. Vertex corrections for the two-particle correlation function which are not included have been discussed elsewhere [42].

3 Results and Discussion

Let us start with the normal state of hole-doped cuprates. In Fig. 4(a), we display the density of states for different doping concentrations $x = 1 - n$. We define $\rho(\omega) \equiv N(\omega) = \sum_{\mathbf{k}} N(\mathbf{k}, \omega)$, with the spectral function $N(\mathbf{k}, \omega) = -\text{Im } G(\mathbf{k}, \omega + i\delta)/\pi$. For a doping value of 20 percent we obtain a peak close to the Fermi energy ($\omega = 0$). However, if one gets closer to the spin-density-wave (AF) instability spectral weight at the same frequency is suppressed, leading to a dip. The temperature for which this phenomenon occurs defines the pseudogap temperature $T = T^*$. Recently, this has indeed been found by scanning tunneling microscopy experiments within the normal state of underdoped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. This pseudogap corresponds also to the gap measured in the charge-excitation-spectrum in the optical conductivity and is also seen in angle-resolved photoemission (ARPES) experiments by Shen and others as mentioned earlier (see, for example, Ref. [43]).

Next we want to discuss the superconducting gap function $\phi(\mathbf{k}, \omega)$ shown in Fig. 4(b) for a given temperature $T = 48\text{K}$ obtained for an intermediate coupling strength $U/t = 4$ and for various doping concentrations x . The corresponding T_c values are $T_c = 50\text{ K}$, $T_c = 63\text{ K}$, $T_c = 60\text{ K}$ (from top to bottom). In order to demonstrate the strong momentum dependence of the superconducting gap we show only the static part ($\omega = 0$) and, for simplicity, only 1/4 of the first Brillouin zone. Due to feedback-effects from the gap function on the dynamical spin susceptibility one obtains a solution for the superconducting order parameter which belongs to the $d_{x^2-y^2}$ wave representation but which is not the simple basis function $\psi(\mathbf{k}) = \cos(k_x) - \cos(k_y)$. In other words, the appearance of higher harmonics in the gap equation $\phi(\mathbf{k}, \omega)$ (see in particular $x = 0.07$) is a result of the self-consistent treatment of the effective pairing interaction which is dominated, of course, by the dynamical spin susceptibility $\chi(\mathbf{q}, \omega)$.

In order to calculate the phase diagram for hole-doped superconductors we must also calculate the superfluid density $n_s(x, T)/m$ self-consistently from the current-current correlation function and from the f-sum rule: the real part of the conductivity $\sigma_1(\omega)$, i.e. $\int_0^\infty \sigma_1(\omega) d\omega = \pi e^2 n / 2m$, where n is the 3D electron density and m denotes the effective band mass for the tight-binding band considered. $\sigma(\omega)$ is calculated in the normal and superconducting state using the Kubo formula [44]. Vertex corrections have been neglected. Physically speaking we are looking for the loss of spectral weight of the Drude peak at $\omega = 0$ that corresponds to excited quasiparticles above the superconduct-

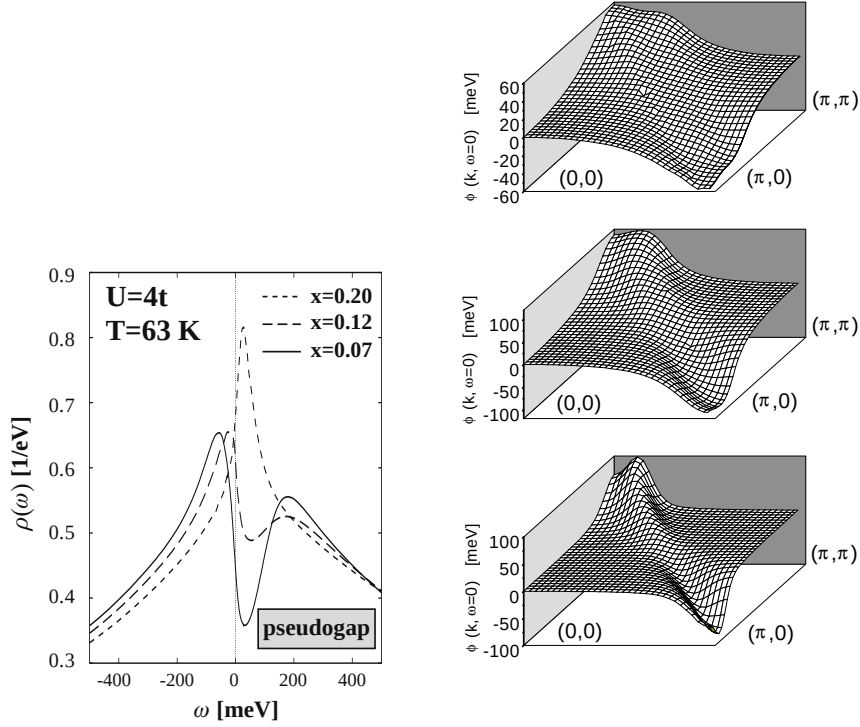


Figure 4. (a) Momentum averaged density of states $\rho(\omega)$ for various doping concentrations. For large doping $\rho(\omega)$ is similar to the uncorrelated case with a large van Hove singularity above the Fermi energy at $\omega = 0$. For small doping a pseudogap appears that is related to the antiferromagnetic correlations and a precursor of the spin density wave gap of a long-range ordered system. (b) Superconducting order parameter $\phi(\mathbf{k}, \omega = 0)$ for various doping concentrations ($x = 0.18$, $x = 0.13$, $x = 0.07$, from top to bottom). All calculations are performed at $T = 48$ K using $U = 4t$. Note, the d -wave symmetry of $\phi(\mathbf{k}, \omega = 0)$ and the appearance of higher harmonics in addition to the simple form $\phi_{\mathbf{k}}^d = \phi_0[\cos(k_x) - \cos(k_y)]$ for underdoped systems.

ing condensate for temperatures $T < T_c^*$. Most importantly, using our results for $n_s(x, T)$, we calculate the doping dependence of the Ginzburg-Landau-like free-energy change $\Delta F \equiv F_S - F_N$, where $\Delta F_{\text{cond}} \simeq \alpha(n_s/m)\Delta_0(x)$ is the condensation energy due to Cooper-pairing and $\Delta F_{\text{phase}} \simeq \hbar^2/2m^*n_s$ the loss in energy due to phase incoherence of the Cooper-pairs. The parameter α describes the available phase space for Cooper-pairs (normalized per unit volume) and can be estimated in the strongly overdoped regime. In the BCS-

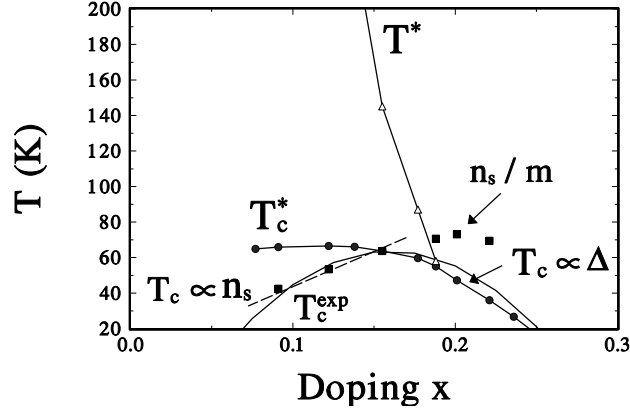


Figure 5. Phase diagram for hole-doped high- T_c superconductors resulting from a spin fluctuation induced Cooper-pairing including their phase fluctuations. The calculated values for $n_s(T=0)/m$ are in good agreement with muon-spin rotation experiments. T_c^* denotes the temperature below which Cooper-pairs are formed. The dashed curve gives $T_c \propto n_s(T=0, x)$. Below T^* we get a gap structure in the spectral density which is shown in Fig. 4(a).

limit one finds $\alpha \simeq 1/400$. Δ_0 is the superconducting order parameter at $T=0$. Note, T_c and in particular $T_c \propto n_s$ follows also from $\langle n_s \rangle = 0$, where one averages over the phase fluctuation time.

In Fig. 5 we show the resulting doping dependence of $T_c(x)$. We also display our results for the doping dependence of $n_s(0)/m$, which are in good agreement with experimental results. The curve T_c^{exp} describes many classes of cuprate material (it is taken from Loram and co-workers [45]). We would like to emphasize that, for the underdoped cuprates, $T_c \propto n_s$ yields indeed better agreement with experimental results than T_c^* obtained from $\Delta(x, T) = 0$ which mark the onset of Cooper-pairing within our mean-field theory. For the temperatures $T_c < T < T_c^*$, one finds pre-formed Cooper-pairs. For the overdoped cuprates, i.e. $x > 0.15$, we get largely BCS-type behavior and $T_c \simeq T_c^* \propto \Delta$. Hence, our electronic theory yields in fair agreement with experiment the non-monotonic doping dependence of $T_c(x)$. Note, we find similar results for the doping dependence of T_c from determining T_c using $n_s(x, T) = 0$. Here, one must include the coupling between Cooper-pairs and their phase fluctuations causing the reduction of $T_c^* \rightarrow T_c$ for the underdoped cuprates and $T_c \propto n_s$. In Fig. 5 results are also given for the characteristic temperature T^* at which a gap appears in the spectral density. Within

our FLEX-theory the occurrence of a pseudogap is due to inelastic electron-electron scattering which leads to a loss of spectral weight at the Fermi level. These are in qualitative agreement with experiments. Finally, we calculate also T_c for the underdoped cuprates with $n_s(T)$ and the Kosterlitz-Thouless theory [46], $k_B T_c(x) = \hbar^2 n_s(T_c)/4ma$ with $a = 2/\pi$, and have found similar T_c values.

We also calculate the phase diagram $T_c(x)$ and $T_N(x)$ of electron-doped cuprates. In order to obtain a unified theory for both hole-doped and electron-doped cuprates it is tempting to use the same Hubbard Hamiltonian taking of course into account the different dispersions for the carriers. Note, in the case of electron doping the electrons occupy copper d -like states of the upper Hubbard band while the holes refer to oxygen-like p -states yielding different energy dispersion as used in our calculations. Then, assuming similar itinerancy of the electrons and holes the mapping on an effective one-band model seems to be justified. We consider U as an effective Coulomb interaction. We find in comparison to hole-doped superconductors smaller T_c values and that superconductivity occurs in a narrower doping range as also observed in experiment. Responsible for this are poorer nesting properties of the Fermi surface and the flat band around $(\pi, 0)$ which lies well below the Fermi level. The narrow doping range for T_c is due to antiferromagnetism up to $x = 0.13$ and rapidly decreasing nesting properties for increasing x [47]. In order to understand the behavior of $T_c(x)$ in underdoped electron-doped cuprates we have calculated the Cooper-pair coherence length ξ_0 , i.e. the size of a Cooper-pair, and find similar values for electron-doped and hole-doped superconductors (from 6 Å to 9 Å). If the superfluid density n_s/n becomes small (for example due to strong coupling lifetime effects), the distance d between Cooper pairs increases. If for $0.15 > x > 0.13$ the Cooper-pairs do not overlap significantly, i.e. $d/\xi_0 > 1$, then Cooper-pair phase fluctuations get important. Thus we expect like for hole-doped superconductors $T_c \propto n_s$. Below T_c we find for all doping concentrations that the gap function has clearly $d_{x^2-y^2}$ -wave symmetry. This is in agreement with the reported linear and quadratic temperature dependence of the in-plane magnetic penetration depth for low temperatures in the clean and dirty limit, respectively, and with phase-sensitive measurements [4]. Previous experiments did not clearly support this and reported mainly s -wave pairing. Maybe as a result of this, so far electron-doped cuprates received much less attention than hole-doped cuprates.

4 Summary

We have used the Hubbard Hamiltonian and the self-consistent FLEX-theory as a model to calculate some basic properties of the hole-doped and electron-doped cuprate superconductors. For the hole-doped case we have discussed the superfluid density n_s/m , and the critical temperature T_c as a function of the doping concentration. We found a phase diagram with two different regions: on the overdoped side a mean-field-like transition and $T_c \propto \Delta(T=0)$, and on the underdoped regime $T_c \propto n_s(T=0)$. For temperatures $T_c < T < T_c^*$ there is a finite superfluid density, but no Meissner effect. This region may be attributed to pre-formed Cooper-pairs without long-range phase coherence. Above T_c^* one has a third energy scale, namely T^* with a gap below in the spectral density of states ("pseudogap").

Our unified model for cuprate superconductivity yields for electron-doped cuprates like for hole-doped ones pure $d_{x^2-y^2}$ symmetry pairing in a good agreement with recent experiments [4]. In contrast to hole-doped superconductors we find smaller T_c values for electron-doped cuprates due to a flat dispersion ϵ_k around $(\pi, 0)$ well below the Fermi level. Furthermore, superconductivity occurs only for a narrow doping range $0.18 > x > 0.13$, because of the onset of antiferromagnetism and, on the other side, due to poorer nesting conditions. We get $2\Delta/k_B T_c = 5.3$ for $x = 0.15$ for the electron-doped cuprates, whereas we obtain much larger values for the hole-doped ones, namely $2\Delta/k_B T_c = 10-12$. The overall agreement with experiments on hole- and electron-doped high- T_c superconductors is remarkably good and suggests spin-fluctuation exchange as the dominant pairing mechanism for superconductivity.

Acknowledgments

It is a pleasure to thank H. Kleinert, A. Pelster, I. Eremin, and C. Joas for helpful discussions. In particular, K.H. Bennemann likes to thank H. Kleinert for more than 20 years of friendly and inspiring colleagueship and many discussions.

References

- [1] J.G. Bednorz and K.A. Müller, *Z. Phys. B* **64**, 189 (1986).

- [2] B. Stadlober *et al.*, *Phys. Rev. Lett.* **74**, 4911 (1995).
- [3] S.M. Anlage *et al.*, *Phys. Rev. B* **50**, 523 (1994).
- [4] C.C. Tsuei and J.R. Kirtly, *Phys. Rev. Lett.* **85**, 182 (2000).
- [5] J.D. Kokales *et al.*, *Phys. Rev. Lett.* **85**, 3696 (2000).
- [6] R. Prozorov, R.W. Gianetta, P. Fournier, and R.L. Greene, *Phys. Rev. Lett.* **85**, 3700 (2000).
- [7] J. Bardeen, L.N. Cooper, and J.R. Schrieffer, *Phys. Rev. B* **108**, 1175 (1957).
- [8] J. Ruvalds *et al.*, *Phys. Rev. B* **51**, 3797 (1995).
- [9] C.M. Varma *et al.*, *Phys. Rev. Lett.* **63**, 1996 (1989).
- [10] A. Virosztek and J. Ruvalds, *Phys. Rev. B* **42**, 4064 (1990).
- [11] J. Ruvalds, C.T. Rieck, J. Zhang, and A. Virosztek, *Science* **256**, 1664 (1992).
- [12] A.J. Millis, H. Monien, and D. Pines, *Phys. Rev. B* **42**, 167 (1990).
- [13] For a review, see C.P. Slichter in *Strongly Correlated Electronic Systems*, Eds. K.S. Bedell, Z. Wang, and D.E. Meltzer (Addison Wesley, Reading, MA, 1994).
- [14] L.P. Regnault *et al.*, *Physica B* **213-214**, 48 (1995).
- [15] D.S. Marshall *et al.*, *Phys. Rev. Lett.* **76**, 4841 (1996).
- [16] J.M. Harris *et al.*, *Phys. Rev. B* **54**, R15655 (1996).
- [17] J.W. Loram, K.A. Mirza, J.R. Cooper, and W.Y. Liang, *Phys. Rev. Lett.* **71**, 1740 (1993); J.W. Loram *et al.*, *J. Superconductivity* **7**, 243 (1994).
- [18] B. Batlogg *et al.*, *Physica C* **235-240**, 130 (1994).
- [19] R. Nemetschek *et al.*, *Phys. Rev. Lett.* **78**, 4837 (1997); G. Blumberg *et al.*, *Science* **278**, 1427 (1997).
- [20] C. Renner *et al.*, *Phys. Rev. Lett.* **80**, 149 (1998).
- [21] A.K. Gupta and K.-W. Ng, *Phys. Rev. B* **58**, R8901 (1998).
- [22] B.K. Chakraverty, A. Taraphder, and M. Avignon, *Physica C* **235-240**, 2323 (1994); B.K. Chakraverty and T.V. Ramakrishnan, *Physica C* **282-287**, 290 (1997).
- [23] V.J. Emery and S.A. Kivelson, *Nature* **374**, 434 (1995).
- [24] P.C. Hohenberg, *Phys. Rev.* **158**, 383 (1967).
- [25] J.L. Tallon, *Phys. Rev. B* **51**, 12911 (1995).
- [26] Y.J. Uemura *et al.*, *Phys. Rev. Lett.* **62**, 2317 (1989).
- [27] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* (World Scientific, Singapore, 1989).

- [28] H. Kleinert and E. Babaev, *Phys. Rev. B* **59**, 12083 (1999).
- [29] A.K. Nguyen and A. Sudbø, *Phys. Rev. B* **60**, 15307 (1999).
- [30] D. Manske and K.H. Bennemann, *Physica C* **341-348**, 83 (2000).
- [31] D.M. King *et al.*, *Phys. Rev. Lett.* **70**, 3159 (1993).
- [32] Q. Si, Y. Zha, K. Levin, and J.P. Lu, *Phys. Rev. B* **47**, 9055 (1993).
- [33] J.C. Campuzano *et al.*, *Phys. Rev. Lett.* **64**, 2308 (1990).
- [34] G.M. Eliashberg, *Sov.-Phys.-JETP* **11**, 696 (1960).
- [35] P.B. Allen and B. Mitrovic, *Solid State Physics* **37**, 1 (1982).
- [36] N.F. Berk and J.R. Schrieffer, *Phys. Rev. Lett.* **17**, 433 (1966).
- [37] N.E. Bickers, D.J. Scalapino, and S.R. White, *Phys. Rev. Lett.* **62**, 961 (1989); N.E. Bickers and D. J. Scalapino, *Ann. Phys. (N.Y.)* **193**, 206 (1989); P. Monthoux and D.J. Scalapino, *Phys. Rev. Lett.* **72**, 1874 (1994); C.-H. Pao and N.E. Bickers, *Phys. Rev. Lett.* **72**, 1870 (1994).
- [38] T. Dahm and L. Tewordt, *Phys. Rev. Lett.* **74**, 793 (1995); *Phys. Rev. B* **52**, 1297 (1995).
- [39] D.J. Scalapino, *Phys. Rep.* **250**, 329 (1995).
- [40] A.A. Abrikosov, L. Gorkov, and I.E. Dzyaloshinskii, *Quantum Field Theoretical Methods in Statistical Physics* (Pergamon Press, Oxford, 1965).
- [41] L. Tewordt, D. Fay, P. Dörre, and D. Einzel, *J. Low Temp. Phys.* **21**, 645 (1975).
- [42] D. Manske, *Phonons, Electronic Correlations, and Self-Energy Effects in High- T_c Superconductors*, PhD thesis, Hamburg, Germany, 1997 (Shaker Verlag, Aachen, 1997).
- [43] Z.X. Shen and D.S. Dessau, *Phys. Rep.* **253**, 1 (1995).
- [44] S. Wernbter and L. Tewordt, *Physica C* **211**, 132 (1993).
- [45] M.R. Presland *et al.*, *Physica C* **176**, 95 (1991); J.R. Cooper and J.W. Loram, *J. Phys. (France) I* **6**, 1 (1996).
- [46] J.M. Kosterlitz and D.J. Thouless, *J. Phys. C* **6**, 1181 (1973); J.M. Kosterlitz, *J. Phys. C* **7**, 1046 (1974).
- [47] D. Manske, I. Eremin, and K.H. Bennemann, *Phys. Rev. B* **62**, 13922 (2000).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

FROM SUPERFLUID ^3He TO TRIPLET SUPERCONDUCTOR Sr_2RuO_4

K. MAKI

*Department of Physics and Astronomy, University of Southern California,
Los Angeles, CA 90089-0484, USA
E-mail: kmaki@usc.edu*

H. WON

*Department of Physics, Hallym University, Chunchon, 200-702, South Korea
E-mail: hkwon@sun.hallym.ac.kr*

The post BCS development of the field of superconductivity and its implication to physics are briefly reviewed. After superfluid ^3He , heavy fermion superconductors, organic superconductors, and high- T_c superconductors, the garden of superconductivity is inhabited by unconventional (i.e. non- s -wave) superconductors. In particular, aspects of d -wave superconductivity are tested semi-quantitatively on high quality single crystals of YBCO and Bi2212.

1 Prologue

*Per correr miglior acque alza le vele omai la navicella del mio
ingegno, ...*

Dante, *Purgatorio*

One of us (K.M.) met Hagen Kleinert for the first time in the spring of 1978. A young guy jumped into my office unannounced and started talking about his new ideas on topological defects in the superfluid ^3He . Superfluid ^3He was discovered in 1972 by Doug Osheroff, Bob Richardson, and Dave Lee, at Cornell University, Ithaca [1,2]. After the success of the BCS theory [3] in describing the superconductivity phenomenon in metals, many authors [4–6] considered possible superfluid ^3He based on the similar pairing of ^3He atoms.

However, there was almost insurmountable difficulty in predicting the superfluid transition temperature and the symmetry of the pairing, though the possibility of s -wave pairing had been rejected early in the game as very unlikely due to the strong repulsive interaction between two ^3He atoms. In the summer of 1972 at the 13th Low Temperature Conference (LT13) at Boulder, Colorado, a special session on this subject was announced. I was in euphoria as if a brave new world was unfolding in front of me and I realized that I was witnessing the almost unique event in my lifetime. After ingenious NMR experiments at Cornell [2] and a brilliant theory by Tony Leggett [7], it became clear that the superfluid ^3He is of spin triplet p -wave pairing and consists of at least two distinct phases: an A phase and a B phase. Unlike s -wave superconductors in metals, the superfluid ^3He possesses a large internal degree of freedom which manifests itself as several Nambu-Goldstone modes (zero sound, spin waves, orbital waves) and a large class of topological defects [8,9]. In order to catch up this rapid development I moved in 1974 from the Tohoku University, Sendai, Japan to the University of Southern California, which is situated at two hours driving distance from the University of California, San Diego (UCSD) in La Jolla. After the discovery of superfluid ^3He in 1972, John Wheatley at La Jolla had done a number of ingenious experiments on superfluid ^3He including magnetic ringing, fourth sound, and zero sound in uniform and non-uniform textures [10]. At that time, Tsuneto and I just got a striking confirmation of our theory of magnetic ringing in the superfluid ^3He -A [11,12]. Actually Hagen was visiting UCSD at La Jolla in 1978 on his sabbatical. When he tried to discuss his ideas with Wheatley, it was then very natural that Wheatley suggested that Hagen should talk with me.

For us many-body theoreticians, the superfluid ^3He provides a wonderful playground where exotic topological objects abound. Just before Hagen came to my office, Kumar and I had succeeded in interpreting a strange NMR satellite, first reported [13] at LT14 in Helsinki in 1975 and later explored in more detail by Gould and Lee [14] in terms of a "soliton" or a "domain wall" [15,16]. This was the first topological defect observed and identified in superfluid ^3He . In 1978, Yu Ren Lin-Liu and I were puzzling about the instability of the uniform texture (i.e. $\mathbf{l} \parallel \mathbf{d}$) and $\mathbf{l}$ and $\mathbf{d}$ constant all over the space, where $\mathbf{l}$ is the quantization axis of the orbital angular momentum and $\mathbf{d}$ is the spin vector in superfluid ^3He -A. At that time, P. Bhattacharya *et al.* [17] had just published an elegant and intriguing theory about the critical point where the uniform texture becomes unstable in the presence of superflow. The question was what happens to the system after this? We could write down a set of

coupled differential equations for $\mathbf{l}$ and $\mathbf{d}$. But the usual solution method led us to nowhere. This was where Hagen entered the discussion with a bright idea how to cook the equations, which enabled us to avoid unphysical singularities. Indeed we found the helical texture [18] which was reported at LT15 in Grenoble, France in the summer of 1978. The helical texture is characterized as the static texture where both $\mathbf{l}$ and $\mathbf{d} \parallel \mathbf{l}$ are winding around the superfluid velocity $\mathbf{v}_s$ [8]. A part of the result and the phase diagram [18,19] was confirmed through the measurement of the abrupt change in the sound attenuation [20]. After this successful collaboration, Hagen invited me to spend one month at the Freie Universität Berlin at Dahlem in the summer of 1980. At that time, the blue phases in liquid crystal were one of the hot topics in this field. Together, we set up a Ginzburg-Landau equation appropriate to the system and tried to find a stable crystalline-like solution. Besides usual tetrahedral and octahedral solutions, we considered the possibility of a dodecahedral phase and described in great detail the complicated icosahedral phase. This was before icosahedral phases were discovered in sputtered aluminium. Our paper [21] would have been completely forgotten if Wright and Mermin [22] had not kindly mentioned our work in their concise review on the blue phases.

Since then, our paths have diverged almost completely. While Hagen developed a disorder field theory for statistical mechanics of line-like vortices and defects [23], worked on membranes and strings [24–26], and found a variational perturbation theory of critical phenomena [27], I focussed attention on the exciting completely new classes of superconductors which appeared on the scene.

2 Day of Unconventional Superconductors

How beauteous mankind is! O brave new world that has such people in't!

Shakespeare, *The Tempest*

The superconducting CeCu_2Si_2 was the first heavy fermion superconductor discovered in 1979 [28]. Heavy fermion superconductors are found mostly in intermetallic compounds based on Ce or U [29,30]. We call them the heavy fermions since the mass of the quasi-particles involved in superconductivity is 100 to 1 000 times larger than the bare electron mass. First of all, the presence of superconductivity is very surprising since most of these systems are

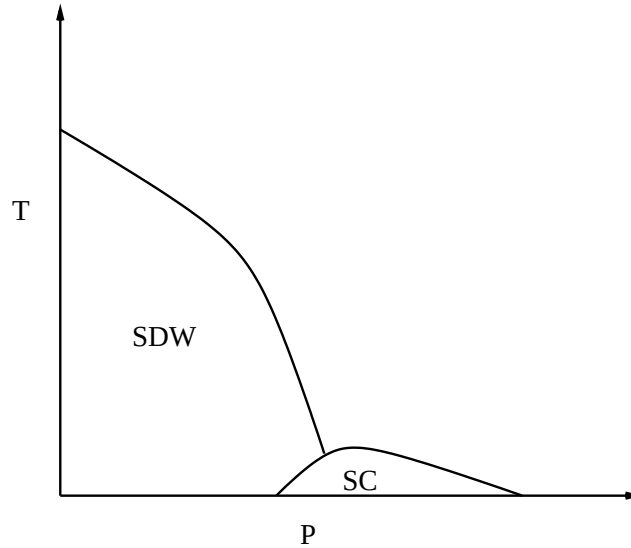


Figure 1. The schematic phase diagram of $(\text{TMTSF})_2\text{PF}_6$ where T and P are the temperature and the pressure, respectively. The spin density wave (SDW) is destroyed as the pressure increases and is replaced by the superconducting state (SC).

magnetic, implying the Coulomb dominance [31]; the Coulomb interaction is stronger than the phonon exchange interaction. Indeed a later experiment shows the nodal lines in the superconducting order parameter $\Delta(\mathbf{k})$ [29]; $\Delta(\mathbf{k}) = 0$ along lines on the Fermi surface, which is rather common in unconventional superconductors (non- s -wave). Here $\mathbf{k}$ is the quasi-particle wave vector.

Almost at the same time the superconductivity in an organic conductor $(\text{TMTSF})_2\text{PF}_6$, also called Bechgaard salts, was discovered by Jerome *et al.* [32] at Orsay. The Bechgaard salts is a quasi one-dimensional system with strong anisotropy in the electric conductivity. It has the particular phase diagram where the spin density wave (SDW) exists next to the superconducting state [32] and the absence of Hebel-Slichter peak in T_1^{-1} in NMR [33,34] indicates unconventional superconductors (see Fig. 1). After that, a variety of organic superconductors have been discovered [35,36]. It appears that most of them are unconventional. In particular there is evidence indicating that the superconductivity in Bechgaard salts is of p -wave while the one in κ -(ET) $_2$ salts is of d -wave [37–39]. This development culminated in 1986 in the discovery of the high- T_c cuprate superconductor $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ by Bednorz and

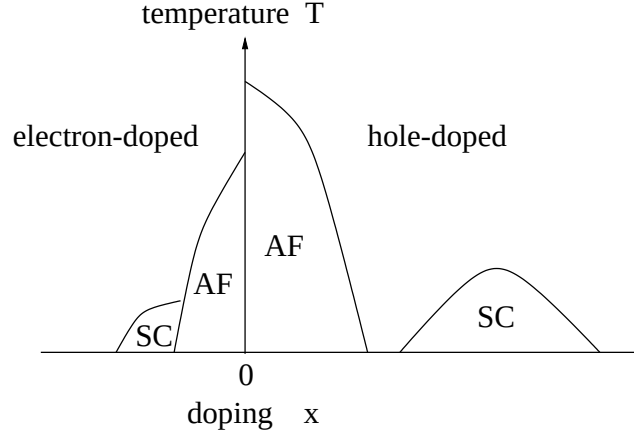


Figure 2. The schematic diagram of the hole-doped ($x > 0$) and the electron-doped ($x < 0$) high- T_c cuprates. Here AF means the antiferromagnetic state and SC the d -wave superconducting state.

Müller [40]. Within a few years the superconducting transition temperature T_c increased from 35 K to 165 K, though these temperatures are still much lower than room temperature. All of these high- T_c cuprates have a layered structure and have the Cu-O₂ planes as basic elements.

These cuprates are also insulating and in an antiferromagnetic state in the absence of electron or hole doping. As the carrier density increases, the antiferromagnetic state is destroyed such that the superconductivity arises with further increase in the carrier density. This behavior is sketched in Fig. 2. For $x > 0$ we present a typical phase diagram of the hole doped cuprates. For the hole concentration around 0.15–0.2, the superconducting transition temperature reaches the maximum value. Further the dependence of the superconducting transition temperature is well approximated by

$$T_c(x) = T_c^0 - a(x - x_0)^2, \quad (1)$$

where T_c^0 is the superconducting transition temperature at the optimal doping (i.e. $x = x_0$) and $a = T_c^0/x_0^2$. On the other hand, T_c of the electron-doped cuprates decreases monotonically with increasing electron density. In order to understand this phase diagram, P.W. Anderson [41] proposed his famous dogma, stating that

- a) all the actions take place in the Cu-O₂ plane;
- b) the phase diagram should be understood in terms of a single Hamiltonian. Here, the presence of the antiferromagnetic phase implies the Coulomb dominance;
- c) one should propose a two-dimensional one-band Hubbard model for high- T_c cuprate superconductor which is the consequence of a) and b).

Of course the Hubbard model is considered as the simplest model to describe magnetism. It may be surprising that a similar Hamiltonian can describe the high- T_c superconductors. Unfortunately the complete solution of the 2D one-band Hubbard model is still not available. However, it appears that the following is certain:

- a) The normal state is the Fermi liquid, though unlike the usual Fermi liquid some of the nesting channels (i.e. $\mathbf{q} = (\pi, \pi)$) play an important role. This point was clarified by Shankar [42] and others [43,44].
- b) There is strong spin fluctuation (antiparamagnon) in this model, which gives rise to d -wave pairing (i.e. $\Delta(\mathbf{k}) \sim \cos(2\phi)$ where $\phi = \tan^{-1}(k_y/k_x)$) [45–47].
- c) The superconductivity is well described by the BCS theory of d -wave superconductor. In general, the mean-field theory appears to apply for all unconventional superconductors [48].

We consider the period 1993–1994 to be the most important time for understanding high- T_c cuprate superconductors, where a number of ingenious methods of exploring d -wave order parameter were developed. Also the availability of high quality single crystals of high- T_c cuprate superconductors provides indispensable support for this success. Perhaps one of the most important experiments is the phase sensitive test using the Josephson interference effect. First Wollman *et al.* [49] constructed a SQUID configuration between YBCO and Pb and studied the interference pattern. They saw the shift, i.e. the peak position shifted by π [see Fig. 3(b)] and current I versus Φ/Φ_0 , where $\Phi_0 = hc/2e$ is the quantum flux. This shift reflects the fact that the order parameter $\Delta(\mathbf{k})$ has the opposite sign for $\mathbf{k} \parallel \mathbf{a}$ and $\mathbf{k} \parallel \mathbf{b}$. In another experiment the whole corner of YBCO was covered by Pb. There instead of a usual Fraunhofer pattern, they observed the anti-Fraunhofer pattern (see Fig. 3(d)) [50].

Also the tricrystal geometry was exploited by Tsuei and Kirtley [51]. They grew three crystals mutually oriented by 60° to each other epitaxially. If we are here dealing with d -wave superconductors, the order parameter has to

change the sign three times when coming back to the starting crystal. But this is unacceptable. In order to resolve this frustration, there appears a half quantum flux at the center of these three crystals, which is detected by an extremely sensitive micromagnet-meter with the diameter ~ 10 microns. With this special technique they have established d -wave superconductivity in YBCO, GdBCO, Bi2212, Tl2201, and more recently in two electron-doped high- T_c cuprates NCCO and PrCCO [51].

Also the $\mathbf{k}$ dependence of $\Delta(\mathbf{k})$ became accessible through the angular resolved photoemission spectrum (ARPES) done by Shen *et al.* [52,53].

Of course, the nodal lines in $\Delta(\mathbf{k})$ imply that the low temperature thermodynamic and transport properties are completely different from the ones in s -wave superconductors. For example, the magnetic penetration depth increases linearly in T while the specific heat like T^2 at low temperatures, which is observed in YBCO [54,55] and in LSCO [56]. Also the quasi-particle density of states, as seen by STM [57] and the electronic Raman scattering [58], exhibits clear d -wave signatures.

Further it is known that the impurity provides a fine proof for unconventional superconductors. Let us just indicate some papers on this subject [59,60].

Perhaps the vortex state will provide the better test of the BCS theory of d -wave superconductivity [61]. For $\mathbf{H} \parallel \mathbf{c}$ the most striking prediction is that the square vortex lattice tilted by 45° from the a - b axis is more stable than the usual hexagonal vortex lattice [62,63].

Though originally the prediction was made in the vicinity of the upper critical field, the square vortex lattice is seen in single crystals of YBCO at low temperatures ($T \simeq 4$ K) and at a low magnetic field ($H \simeq$ a few Tesla) by small angle neutron scattering [64] and in a scanning tunneling microscope [65]. Strictly speaking, the apex angle of the observed vortex lattice is 77° and not 90° . But this difference is easily understood in terms of the a - b anisotropy in YBCO. Due to the orthorhombic distortion in YBCO, the coherence lengths ξ_a and ξ_b are not equal but $\xi_b/\xi_a \simeq 1.5$, where the subscripts a and b mean the component parallel to the a - and b -axis, respectively.

Still controversial are also the thermodynamic and the transport properties of the vortex state. Concerning the weak magnetic field (i.e. $H/H_{c2} \ll 1$, where H_{c2} is the upper critical field), Volovik [66] has written the seminal paper that the effect of the magnetic field can be treated quasi-classically. This approach was generalized by a number of studies, including one by us [67–69]. More exciting is that the careful measurement of thermodynamics, NMR, and

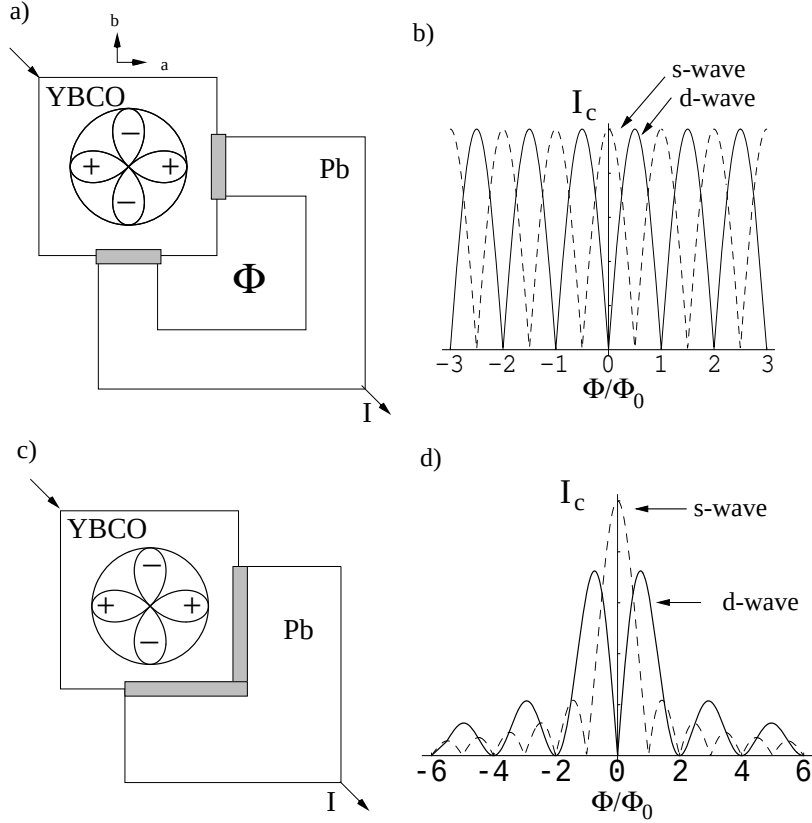


Figure 3. a) The Josephson interferometry between *d*-wave superconductor (YBCO) and *s*-wave superconductor (Pb). The critical current I_c depends on the magnetic flux Φ as shown in b). Therefore the observed π shift provides the test of the *d*-wave superconductor. c) The corner junction where the corner of YBCO is covered by Pb. Here the magnetic flux Φ is distributed along the boundary between YBCO and Pb. The *d*-wave superconductor then exhibits the anti-Fraunhofer pattern as shown in d).

thermal conductivity in the vortex state in YBCO and Bi2212 at low temperature have been reported, which confirms in general the predicted $\sqrt{H}$ -dependence of the specific heat [55] and the spin susceptibility seen through the Knight shift and the linear H -dependence of T_1^{-1} (the nuclear spin lattice relaxation rate) in NMR. In the future, the single crystals of YBCO and Bi2212 will provide an extremely useful testing ground for new ideas and concepts in unconventional superconductors.

Among heavy fermion superconductors, UPt_3 is the only system where the nature of the symmetry is well established [30]. UPt_3 is the hexagonal crystal with an axis parallel to the c -axis. At low temperatures ($T \ll T_c = 0.55$ K), both the electronic thermal conductivity κ_c and κ_b behave linear in T , where the subscripts c and b mean parallel to the c -axis and the b -axis [70]. The simplest possibility consistent with this is $\Delta(\mathbf{k}) \sim Y_{3,\pm 2}(\theta, \phi)$ or E_{2u} , where $Y_{3,\pm 2}(\theta, \phi)$ is the spherical harmonics [71,72]. The Knight shift seen by NMR almost at the same time confirmed the triplet spin pairing [73]. Later the details of the spin configuration in the A, B, and C phases were identified [74]. In 1999, a paper confirmed that both the upper critical field and the ultrasonic attenuation data of UPt_3 are consistent with E_{2u} [75]. Therefore, UPt_3 appears to provide another nice system to test new ideas on unconventional superconductors.

3 Story of Sr_2RuO_4

*Ch'ebbe l'origine nell'Alemagna, che poi s'è celebre là in Francia fu.
Mozart/da Ponte, Così fan tutte*

The superconductivity in Sr_2RuO_4 was discovered in 1994 by Maeno *et al.* [76]. This is an isocrystal to La_2CuO_4 , a mother system to $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$, the high temperature cuprates. But unlike HTSC it is already metallic while La_2CuO_4 is insulating and antiferromagnetic. Also, at low temperatures ($T < 20$ K) the electron in Sr_2RuO_4 behaves as the Fermi liquid. Further, the system becomes superconducting below $T \sim 1$ K.

From the analogy to superfluid ^3He it was then proposed [77] that the superconductor is of triplet p -wave with $\Delta(\mathbf{k}) \sim \hat{d}(k_1 \pm ik_2) = \hat{d}e^{i\phi}$ and the spin vector $\mathbf{d}$ parallel to the c -axis. Both the spontaneous spin polarization observed by the muon spin rotation experiment [78] and a flat Knight shift [79] in the superconducting state of Sr_2RuO_4 confirm the triplet nature of the pairing and also the sensitivity to the disorder points to the unconventional superconductivity [80].

But recently the situation changed dramatically by the availability of the high quality single crystals of Sr_2RuO_4 with $T \sim 1.5$ K. As shown in Figs. 4 and 5, both the specific heat [81] and the superfluid density [82] suggest there is no energy gap. Indeed, the overall behavior is much more consistent with d -wave superconductors [83]. So many people proposed possible f -wave superconductivity in Sr_2RuO_4 ($\Delta(\mathbf{k}) \sim \cos(2\phi)e^{\pm i\phi}$, $\Delta(\mathbf{k}) \sim \sin(2\phi)e^{\pm i\phi}$,

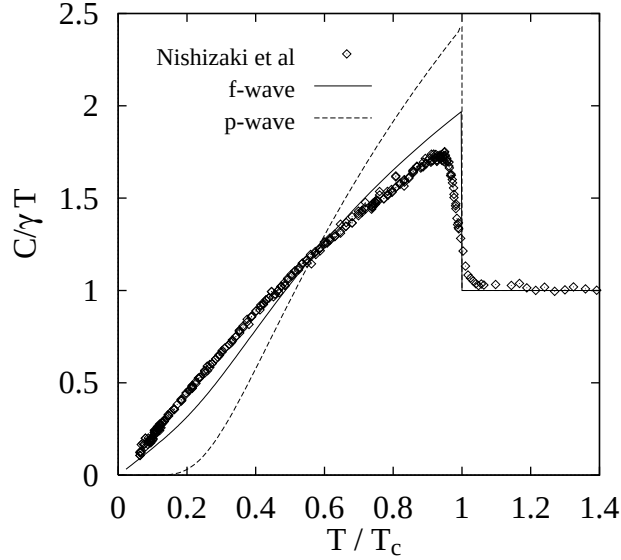


Figure 4. The specific-heat data [76] divided by γT , where γ is the Sommerfeld constant, is compared with the theoretical results for the p -wave [72] and f -wave [81] superconductors.

and $\cos(ck_3)e^{\pm i\phi}$ [83–87]. In particular, within the weak-coupling theory, these three f -states have the same thermodynamics as the one in d -wave superconductors [88]. Therefore the thermodynamic data cannot discriminate one from the other.

Another result is that the specific heat in the vortex state for $\mathbf{H} \parallel \mathbf{c}$ at $T = 0.1$ K exhibits clearly the $\sqrt{H}$ -behavior (see Fig. 6) [81,85]. The deviation from the $\sqrt{H}$ -behavior for $H < 0.01$ T is most likely due to the fact that the system is in the Meissner state. Additionally, the thermal conductivity in the vortex state at low temperature exhibits the H -linear dependence [85,89,90], which indicates not only the nodal structure in $\Delta(\mathbf{k})$ but also that the system is in the superclean limit ($\Gamma/\Delta \ll H/H_{c2} \ll 1$, where Γ is the quasi-particle scattering rate). Indeed, the quasi-particle mean free path of these systems is much longer than a few micrometers.

To choose the correct order parameter, we need experiments which are sensitive to the anisotropy within the a - b plane. A first experiment of such a kind was provided by the upper critical field in a planar magnetic field [91]. Indeed, they found a fourfold term in the upper critical field or $H_{c2}(\theta, T)$

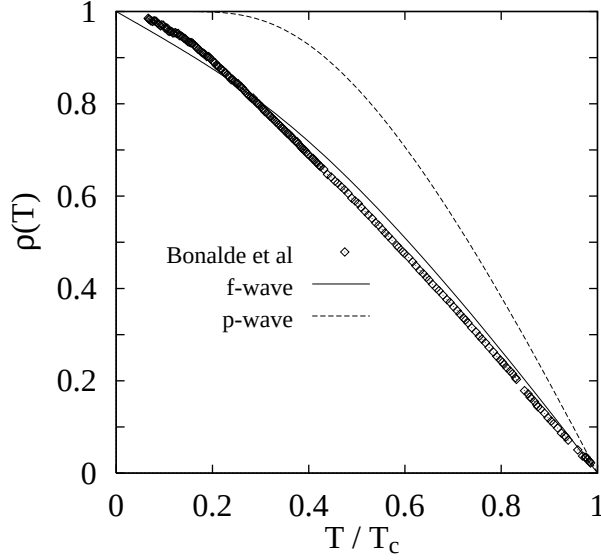


Figure 5. The superfluid density data [77] is compared with the p -wave and f -wave models.

which is approximated by

$$H_{c2}(\theta, T) = H_{c2}^0(T) - H_{c2}^1(T) \cos(4\theta), \quad (2)$$

where θ is the angle which the magnetic field makes with the a -axis. However, $H_{c2}^1(T)$ is rather small ($H_{c2}^1(T)/H_{c2}^0(T) \sim 3\%$). This suggests strongly that this anisotropy does not reflect the symmetry of $\Delta(\mathbf{k})$ but rather the band structure effect [92,93]. Indeed both $\cos(2\phi)e^{\pm i\phi}$ and $\sin(2\phi)e^{\pm i\phi}$ should exhibit large anisotropy ($\sim 30\%$) and therefore they are incompatible with the experiment. The thermal conductivity in the vortex state in a planar magnetic field has also been measured recently: it shows extremely small anisotropy [90,94]. Therefore, the thermal conductivity data are also incompatible with $\cos(2\phi)e^{\pm i\phi}$, and $\sin(2\phi)e^{\pm i\phi}$. This leaves us only with $\cos(ck_3)e^{\pm i\phi}$, though this state requires a strong interlayer spin coupling or interlayer Coulomb interaction [95].

Since the discovery of unconventional superconductors, the collective modes and the possible topological defects have been considered [29]. However, so far there is no evidence for the collective modes or topological defects, except for the usual Abrikosov vortex. Perhaps this situation may change

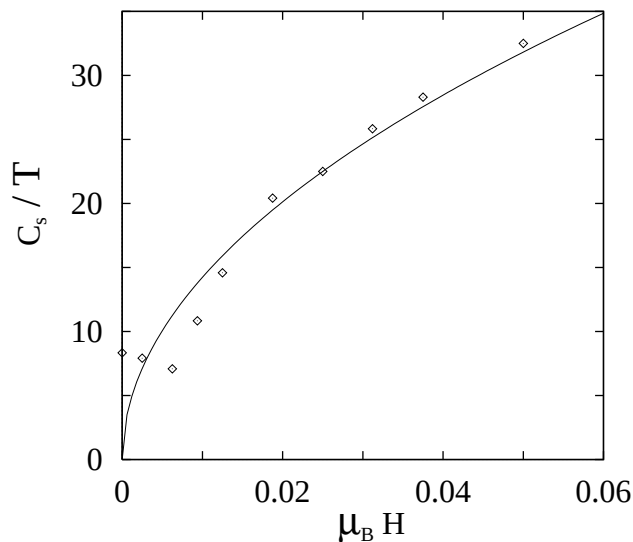


Figure 6. The specific-heat data of Sr_2RuO_4 in $\mathbf{H} \parallel \mathbf{c}$ and $T = 0.05$ K [76] is compared with the square-root of H . The deviation below $H \simeq 0.01$ T is due to the fact that $H_{c1} \simeq 0.01$ T.

drastically by the appearance of the triplet superconductors.

First of all, the triplet superconductors should have spin waves as collective modes [96,97]. Due to the two-dimensional representation there should also be the clapping mode [98,99], which couples to both the sound wave and the Raman photon. Unfortunately the coupling to the sound wave appears to be discouragingly small [99]. On the other hand, we believe that the Raman scattering is more promising if one can do the Raman scattering experiment below 100 mK. For this we clearly need the development of a new technology. The topological defect-like $\hat{\mathbf{l}}$ -soliton [100] and $\hat{\mathbf{d}}$ -soliton with half-quantum vortex [101] have also been predicted. So perhaps the single crystal of Sr_2RuO_4 will provide the unique laboratory to test these new concepts.

4 Outlook

*O glücklich, wer noch hoffen kann, aus diesem Meer des Irrtums
aufzutauchen!*

Goethe, *Faust*

We have seen that the field of superconductivity expanded enormously since 1979. Actually most of the new superconductors in heavy fermion systems, charge conjugated organic conductors and high- T_c cuprates are unconventional. Therefore, unconventional superconductors will play a central role in the 21st century. Compared with conventional superconductors, unconventional superconductors are more sensitive, subtle and delicate to the environment. This will require much more delicate control of the sample preparation and the crystal formation. Their response to the external perturbation is more subtle and delicate. In spite of this, it is surprising that the mean-field theory as embodied in the BCS theory and the Landau theory of Fermi liquid works very fine in describing a manifold of phenomena. Can we trust in this approach for a long time? Of course, there are now many people claiming that the mean-field theory is unreliable. But if we limit ourselves to unconventional superconductors, we have not seen any failure or sign of failure of the mean-field theory. Quite parallel to this development we may have now unconventional charge density wave and spin density wave as well [102,103]. We believe that the understanding of all subtleties of these new superconductors is also crucial for the real application of unconventional superconductors, including high- T_c cuprate superconductors. So we may approach to soft matter physics from solid state physics through a different passage.

5 Acknowledgments

We are very happy to dedicate this article to the 60th birthday of Hagen Kleinert, our friend and colleague. We wish him fruitful work in coming years. On our trajectory from superfluid ^3He to triplet superconductor Sr_2RuO_4 we have enjoyed many collaborations and support. We would like to thank in particular Thomas Dahm, Balazs Dora, Peter Fulde, Stephan Haas, Takehiko Ishiguro, Koichi Izawa, Hae-Young Kee, Yong-Baek Kim, Mahito Kohmoto, Yoshiteru Maeno, Yuji Matsuda, Yoshifumi Morita, Jun'ichi Shiraishi, Makaryi Tanatar, Silvia Tomić, and Attila Virostek for helpful collaborations. H.W. acknowledges the support from the Korea Research Foundation under the Professor Dispatching Scheme.

References

- [1] D.D. Osheroff, R.C. Richardson, and D.M. Lee, *Phys. Rev. Lett.* **28**, 885 (1972).

- [2] D.D. Osheroff, W.J. Gully, R.C. Richardson, and D.M. Lee, *Phys. Rev. Lett.* **29**, 920 (1972).
- [3] J. Bardeen, L.N. Cooper, and J.R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).
- [4] L.P. Pitaevskii, *Soviet Phys. JETP* **10**, 1267 (1960).
- [5] K.A. Brueckner, T. Soda, P.W. Anderson, and P. Morel, *Phys. Rev.* **118**, 1442 (1960).
- [6] V.J. Emery and A.M. Sessler, *Phys. Rev.* **119**, 43 (1960).
- [7] A.J. Leggett, *J. Phys. C* **6**, 3187 (1973); *Ann. Phys. (N.Y.)* **85**, 11 (1974).
- [8] D. Vollhardt and P. Wölfe, *The Superfluid Phases of Helium 3* (Taylor and Francis, London, 1990).
- [9] G.E. Volovik, *Exotic Properties of Superfluid ^3He* (World Scientific, Singapore, 1992).
- [10] J.C. Wheatley, *Rev. Mod. Phys.* **7**, 415 (1975); *Progress in Low Temperature Physics VII A*, Ed. D.F. Brewer (North-Holland, Amsterdam, 1978), pp. 77.
- [11] K. Maki and T. Tsuneto, *Prog. Theor. Phys.* **52**, 617 (1974).
- [12] R.A. Webb, R.L. Kleinberg, and J.C. Wheatley, *Phys. Rev. Lett.* **33**, 145 (1974).
- [13] O. Avenel, M.E. Bernier, E.J. Varoquaux, and C. Vibet, in *Proc. 14th Int. Conf. on Low Temperature Physics*, Otaniemi, Finland 1975, Vol. 5, Eds. M. Krusius and M. Vuorio (North-Holland, Amsterdam, 1976), p. 429.
- [14] C.M. Gould and D.M. Lee, *Phys. Rev. Lett.* **37**, 1223 (1976).
- [15] K. Maki and P. Kumar, *Phys. Rev. Lett.* **38**, 557 (1977); *Phys. Rev. B* **16**, 187 (1977).
- [16] K. Maki, in *Solitons*, Eds. S.E. Trullinger, V.E. Zakharov, and V.L. Pokrovsky (North-Holland, Amsterdam, 1986).
- [17] P. Bhattacharya, T.L. Ho, and N.D. Mermin, *Phys. Rev. Lett.* **39**, 1290 (1977).
- [18] H. Kleinert, Y.R. Lin-Liu, and K. Maki, *Journal de Physique* **36**, C6 (1978); *Phys. Lett. A* **70**, 27 (1979).
- [19] D. Vollhardt, Y.R. Lin-Liu, and K. Maki, *J. Low Temp. Phys.* **37**, 627 (1979).
- [20] R.L. Kleinberg, *Phys. Rev. Lett.* **42**, 182 (1979).
- [21] H. Kleinert and K. Maki, *Fortschr. Phys.* **29**, 219 (1981).

- [22] D.C. Wright and N.D. Mermin, *Rev. Mod. Phys.* **61**, 385 (1989).
- [23] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [24] H. Kleinert, *Phys. Lett. B* **174**, 335 (1986).
- [25] H. Kleinert, *Phys. Lett. A* **257**, 269 (1999).
- [26] M. Bachmann, H. Kleinert, and A. Pelster, *Phys. Lett. A* **261**, 127 (1999).
- [27] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [28] F. Steglich, J. Aarts, C.D. Bredl, W. Lieke, D. Meschede, W. Franz, and H. Schäfer, *Phys. Rev. Lett.* **43**, 1892 (1979).
- [29] M. Sigrist and K. Ueda, *Rev. Mod. Phys.* **63**, 239 (1991).
- [30] R.H. Heffner and M.R. Norman, *Comments on Cond. Matt. Physics* **17**, 361 (1996).
- [31] H. Won and K. Maki, *J. Phys. (France) I* **6**, 2317 (1996).
- [32] D. Jerome, A. Mazaud, M. Ribault, and K. Bechgaard, *J. Phys. Lett.* **41**, L95 (1980).
- [33] M. Takigawa, H. Yasuoka, and G. Saito, *J. Phys. Soc. Jpn.* **56**, 873 (1984).
- [34] Y. Hasegawa and H. Fukuyama, *J. Phys. Soc. Jpn.* **56**, 877 (1984).
- [35] T. Ishiguro and K. Yamaji, *Organic Superconductors* (Springer, Berlin, 1990).
- [36] M. Lang, *Superconductivity Review* **2**, 1 (1996).
- [37] K. Maki, E. Puchkaryov, and H. Won, *Synth. Metals* **703**, 1933 (1999).
- [38] A. Carrington, I.J. Bonalde, R. Prozorov, R.W. Giarmetta, A.M. Kini, J. Schlueter, H.H. Wang, V. Geiser, and J.M. Williams, *Phys. Rev. Lett.* **83**, 4172 (1999).
- [39] M. Pinteric, S. Tomić, M. Pester, D. Droba, O. Milat, K. Maki, D. Schweitzer, I. Heinen, and W. Strunz, *Phys. Rev. B* **61**, 7033 (2000).
- [40] J.G. Bednorz and K.A. Müller, *Z. Phys. B* **64**, 189 (1986).
- [41] P.W. Anderson, *Science* **235**, 1196 (1987); *The Theory of High T_c Superconductivity* (University Press, Princeton, 1998).
- [42] R. Shankar, *Rev. Mod. Phys.* **66**, 129 (1994).
- [43] W. Metzner, *Physica B* **197**, 452 (1994).
- [44] A. Houghton, H.J. Kwon, and J.B. Marston, *Advances in Physics* **49**, 141 (2000).

- [45] N.E. Bickers, D.J. Scalapino, and S.R. White, *Phys. Rev. Lett.* **62**, 961 (1989); C.-H. Pao and N.E. Bickers, *Phys. Rev. Lett.* **72**, 1870 (1994).
- [46] T. Moriya, Y. Takahashi, and K. Ueda, *J. Phys. Soc. Jpn.* **59**, 2905 (1990).
- [47] P. Monthoux, A.V. Balatsky, and D. Pines, *Phys. Rev. Lett.* **67**, 3448 (1992).
- [48] H. Won and K. Maki, in *Symmetry and Pairing in Superconductors*, Eds. M. Ausloos and S. Kruchinin (Kluwer Academic Press, Dordrecht, 1999).
- [49] D.A. Wollman, D.J. van Harlingen, W.C. Lee, D.M. Ginsberg, and A.J. Leggett, *Phys. Rev. Lett.* **71**, 2134 (1993).
- [50] D.A. Wollman, D.J. van Harlingen, J. Giapintzakis, and D.M. Ginsberg, *Phys. Rev. Lett.* **74**, 797 (1995).
- [51] C.C. Tsuei and J.R. Kirtley, *Physica C* **282-287**, 4 (1997); *Rev. Mod. Phys.* **72**, 969 (2000); *Phys. Rev. Lett.* **85**, 182 (2000).
- [52] Z.X. Shen, D.S. Dessau, B.O. Wells, D.M. King, W.E. Spicer, A.J. Arko, D.S. Marshall, L.W. Lombardo, A. Kapitulnik, P. Dickinson, S. Doniach, J. DiCarlo, A.G. Loeser, and C.H. Park, *Phys. Rev. Lett.* **70**, 1553 (1993).
- [53] H. Ding, M.R. Norman, J. Yokoya, T. Takeuchi, M. Randeira, J.C. Campuzano, T. Takahashi, T. Machida, and K. Kadowaki, *Phys. Rev. Lett.* **78**, 2628 (1997).
- [54] W.N. Hardy, D.A. Bonn, D.C. Morgan, R. Liang, and K. Zhang, *Phys. Rev. Lett.* **70**, 3989 (1993).
- [55] K.A. Moler, D.J. Baar, J.S. Urbach, R. Liang, W.N. Hardy, and A. Kapitulnik, *Phys. Rev. Lett.* **73**, 2744 (1994).
- [56] N. Momono, N. Ido, T. Nakano, M. Oda, Y. Okijima, and K. Yamaya, *Physica C* **233**, 396 (1994).
- [57] C. Renner and Ø. Fischer, *Phys. Rev. B* **51**, 9208 (1995).
- [58] T.P. Devereaux, D. Einzel, B. Stadlober, R. Hackl, D.H. Leach, and J.J. Neumeier, *Phys. Rev. Lett.* **72**, 396 (1994).
- [59] K. Maki and H. Won, *Ann. Phys. (Leipzig)* **5**, 320 (1996).
- [60] E. Puchkaryov and K. Maki, *Europhys. J. B* **4**, 191 (1998).
- [61] K. Maki, H. Won, M. Kohmoto, J. Shiraishi, Y. Morita, and G.F. Wang, *Physica C* **317-318**, 353 (1999).
- [62] H. Won and K. Maki, *Europhys. Lett.* **30**, 421 (1995); *Phys. Rev. B* **53**, 5927 (1996).

- [63] J. Shiraishi, M. Kohmoto, and K. Maki, *Europhys. Lett.* **44**, 367 (1998); *Phys. Rev. B* **50**, 4497 (1999).
- [64] B. Keimer, W.Y. Shin, R.W. Erwin, J.W. Lynn, F. Dagan, and I.A. Alesay, *Phys. Rev. Lett.* **73**, 3459 (1994).
- [65] I. Maggio-Aprile, C. Renner, A. Erb, E. Walker, and Ø. Fischer, *Phys. Rev. Lett.* **75**, 2754 (1995).
- [66] G.E. Volovik, *JETP Lett.* **58**, 469 (1993).
- [67] Yu.S. Barash, V.P. Mineev, and A.A. Svidzinskii, *JETP* **65**, 638 (1997); Yu.S. Barash and A.A. Svidzinskii, *Phys. Rev. B* **58**, 476 (1998).
- [68] C. Kübert and P.J. Hirschfeld, *Sol. Stat. Comm.* **105**, 439 (1998); *Phys. Rev. Lett.* **80**, 4963 (1998).
- [69] H. Won and K. Maki, *Europhys. Lett.* **52**, 427 (2000).
- [70] B. Lussier, B. Elleman, and L. Taillefer, *Phys. Rev. B* **51**, 5145 (1996).
- [71] M.R. Norman and P.J. Hirschfeld, *Phys. Rev. B* **53**, 5145 (1996).
- [72] M.J. Graf, S.K. Yip, and A.J. Sauls, *J. Low Temp. Phys.* **102**, 367 (1996).
- [73] H. Tou, Y. Kitaoka, K. Asayama, N. Kimura, Y. Juki, E. Yamamoto, and K. Maezawa, *Phys. Rev. Lett.* **77**, 1374 (1996).
- [74] H. Tou, Y. Kitaoka, K. Ishida, K. Asayama, N. Kimura, Y. Juki, E. Yamamoto, Y. Haga, and K. Maezawa, *Phys. Rev. Lett.* **80**, 3129 (1998).
- [75] G. Yang and K. Maki, *Europhys. Lett.* **48**, 208 (1999); K. Maki and G. Yang, *Fizika A (Zagreb)* **8**, 345 (1999).
- [76] Y. Maeno *et al.*, *Nature* **372**, 532 (1994).
- [77] T.M. Rice and M. Sigrist, *J. Phys. Cond. Matt.* **7**, L643 (1995); M. Sigrist *et al.*, *Physica C* **317-318**, 134 (1999).
- [78] G.M. Luke *et al.*, *Nature* **394**, 558 (1998).
- [79] K. Ishida *et al.*, *Nature* **396**, 658 (1998).
- [80] A.P. Mackenzie, R.K.W. Haselwimmer, A.W. Tyler, G.G. Lonzarich, Y. Mori, S. Nishizaki, and Y. Maeno, *Phys. Rev. Lett.* **80**, 161 (1998).
- [81] A. Nishizaki, Y. Maeno, and Z.Q. Mao, *J. Phys. Soc. Jpn.* **69**, 572 (2000); *J. Low Temp. Phys.* **117**, 1581 (1999).
- [82] I. Bonalde, B.D. Yanoff, M.H. Salannon, D.J. van Harlingen, E.M.E. Chia, Z.Q. Mao, and Y. Maeno, *Phys. Rev. Lett.* **85**, 4775 (2000).
- [83] T. Dahm, K. Maki, and H. Won, eprint: cond-mat/0006301.
- [84] Y. Hasegawa, K. Machida, and M. Ozaki, *J. Phys. Soc. Jpn.* **69**, 336 (2000).

- [85] H. Won and K. Maki, *Europhys. Lett.* **53**, in press (2001), eprint: cond-mat/0012336.
- [86] H. Won and K. Maki, *Europhys. Lett.* **52**, 427 (2000).
- [87] M.J. Graf and A.V. Balatsky, *Phys. Rev. B* **62**, 9697 (2000).
- [88] H. Won and K. Maki, *Phys. Rev. B* **49**, 1397 (1994).
- [89] H. Won and K. Maki, in *NATO Workshop on Borocarbides* (Dresden, June 2000).
- [90] K. Izawa *et al.*, *Phys. Rev. Lett.* **86**, 2653 (2001), eprint: cond-mat/0012137.
- [91] Z.Q. Mao, Y. Maeno, S. Nishizaki, T. Akima, and T. Ishiguro, *Phys. Rev. Lett.* **84**, 991 (2000).
- [92] K. Maki, G.J. Wang, and H. Won, *J. of Superconductivity* **12**, 551 (1999).
- [93] G. Yang and K. Maki, preprint.
- [94] M.A. Tanatar, S. Nagai, Z.Q. Mao, Y. Maeno, and T. Ishiguro, *Physica C* **341-348**, 1841 (2000).
- [95] M. Sato and M. Kohmoto, *J. Phys. Soc. Jpn.* **69**, 3505 (2000).
- [96] L. Tewordt, *Phys. Rev. Lett.* **83**, 1007 (1999).
- [97] H.Y. Kee, Y.B. Kim, and K. Maki, *Phys. Rev. B* **61**, 3584 (2000).
- [98] S. Higashitani and K. Nagai, *Physica B* **284-288**, 539 (2000).
- [99] H.Y. Kee, Y.B. Kim, and K. Maki, *Phys. Rev. B* **62**, 5877 (2000).
- [100] M. Sigrist and D.F. Agterberg, *Prog. Theor. Phys.* **102**, 965 (1999).
- [101] H.Y. Kee, Y.B. Kim, and K. Maki, *Phys. Rev. B* **62**, R9274 (2000).
- [102] B. Dora and A. Virosztek, *J. Phys. France IV* **9**, Pr10 (1999).
- [103] B. Dora, A. Virosztek, and K. Maki, preprint.

Part V

Topological Defects, Strings, and Membranes

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

**DESCRIPTION OF VORTICITY BY GRASSMANN
VARIABLES AND AN EXTENSION TO SUPERSYMMETRY**

R. JACKIW

*Center for Theoretical Physics,
Massachusetts Institute of Technology,
Cambridge, MA 02139-4307, USA
E-mail: jackiw@ralph2.mit.edu*

Hagen Kleinert's early interest in particle physics quantum field theory served him well for his subsequent researches on statistical physics and collective phenomena. Therefore, on the occasion of a significant birthday, I offer him this essay, in which particle physics concepts are blended into a field theory for macroscopic phenomena: Fluid mechanics is enhanced by anticommuting Grassmann variables to describe vorticity, while an additional interaction for the Grassmann variables leads to supersymmetric fluid mechanics.

1 Précis of Fluid Mechanics (With No Vorticity)

Let me begin with a précis of fluid mechanical equations [1]. An isentropic fluid is described by a matter density field ρ and a velocity field $\mathbf{v}$, which satisfy a continuity equation involving the current $\mathbf{j} = \rho\mathbf{v}$:

$$\dot{\rho} + \nabla \cdot (\rho\mathbf{v}) = 0 \quad (1)$$

and a force equation involving the pressure P :

$$\dot{\mathbf{v}} + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{1}{\rho} \nabla P \quad (2)$$

(over-dot denotes differentiation with respect to time). For isentropic fields, the pressure P is a function only of the density, and the right-hand side of Eq. (2) may also be written as $-\nabla V'(\rho)$, where $V'(\rho)$ is the enthalpy,

$P(\rho) = \rho V'(\rho) - V(\rho)$, and $\sqrt{\rho V''(\rho)} = \sqrt{P'(\rho)}$ is the sound speed (prime denotes differentiation with respect to the argument).

Equations (1) and (2) can be obtained by bracketing the dynamical variables ρ and $\mathbf{v}$ with the Hamiltonian $H(\rho, \mathbf{v})$,

$$H(\rho, \mathbf{v}) = \int d\mathbf{r} \left[\frac{1}{2} \rho v^2 + V(\rho) \right], \quad (3)$$

according to

$$\dot{\rho} = \{H, \rho\}, \quad (4a)$$

$$\dot{\mathbf{v}} = \{H, \mathbf{v}\}, \quad (4b)$$

provided the nonvanishing Poisson brackets of the fundamental variables $(\rho, \mathbf{v})$ are taken to be [2]

$$\{v^i(\mathbf{r}), \rho(\mathbf{r}')\} = \partial_i \delta(\mathbf{r} - \mathbf{r}'), \quad (5a)$$

$$\{v^i(\mathbf{r}), v^j(\mathbf{r}')\} = -\frac{\omega_{ij}(\mathbf{r})}{\rho(\mathbf{r})} \delta(\mathbf{r} - \mathbf{r}') \quad (5b)$$

(the fields in the brackets are at equal times, hence the time argument is suppressed). Here ω_{ij} is the vorticity, defined as the curl of v^i :

$$\omega_{ij} = \partial_i v^j - \partial_j v^i. \quad (6)$$

One naturally asks whether there is a canonical 1-form that leads to the symplectic structure (5); that is, one seeks a Lagrangian whose canonical variables can be used to derive (5) from canonical brackets. When the velocity is irrotational, the vorticity vanishes, $\mathbf{v}$ can be written as the gradient of a velocity potential θ , $\mathbf{v} = \nabla\theta$, and (5) is satisfied by postulating that

$$\{\theta(\mathbf{r}), \rho(\mathbf{r}')\} = \delta(\mathbf{r} - \mathbf{r}'), \quad (7)$$

that is, the velocity potential is conjugate to the density, so that the Lagrangian can be taken as

$$L|_{\text{irrotational}} = \int d\mathbf{r} \theta \dot{\rho} - H|_{\mathbf{v}=\nabla\theta}, \quad (8)$$

where H is given by (3) with $\mathbf{v} = \nabla\theta$.

2 Extending the Formalism to Include Vorticity

The traditional method of including vorticity in a Lagrangian formalism [3] involves writing the velocity in a more elaborate potential representation, the so-called Clebsch parameterization [4],

$$\mathbf{v} = \nabla\theta + \alpha\nabla\beta \quad (9)$$

which supports nonvanishing vorticity

$$\omega_{ij} = \partial_i\alpha\partial_j\beta - \partial_j\alpha\partial_i\beta. \quad (10)$$

The Lagrangian

$$L = - \int d\mathbf{r} \rho(\dot{\theta} + \alpha\dot{\beta}) - H|_{\mathbf{v}=\nabla\theta+\alpha\nabla\beta} \quad (11)$$

identifies canonical pairs to be $\{\theta, \rho\}$ (as in the irrotational case) and also $\{\beta, \alpha\rho\}$. It then follows that the algebra (5) is satisfied, provided $\mathbf{v}$ is given by (9). The quantities (α, β) are called the "Gauss potentials".

The situation here is similar to the electromagnetic force law: The Lorentz equation can be presented in terms of the electric and magnetic field strengths, but a Lagrangian for the motion requires describing the fields in terms of potentials.

3 Some Further Observations on the Clebsch Decomposition of the Vector Field $\mathbf{v}$

In three dimensions, Eq. (9) involves the same number of functions on the left and right sides of the equality, i.e. three. The total number of dynamical variables $(\rho, \mathbf{v})$ is even, it is four, so an appropriate phase space can be constructed from ρ and the three potentials (θ, α, β) . Nevertheless the Gauss potentials are not uniquely determined by $\mathbf{v}$. The following is the reason why a canonical formulation of (5) requires using the Clebsch decomposition (9). Although the algebra (5) is consistent in that the Jacobi identity is satisfied, it is degenerate in that the kinematic helicity

$$h \equiv \frac{1}{2} \int d^3r \mathbf{v} \cdot (\nabla \times \mathbf{v}) = \frac{1}{2} \int d^3r \mathbf{v} \cdot \boldsymbol{\omega}, \quad (12)$$

with $\omega^i = \varepsilon^{ijk}\omega_{jk}/2$, has vanishing bracket with ρ and $\mathbf{v}$ (note that h is just the Abelian Chern-Simons term of $\mathbf{v}$ [5]). Consequently, a canonical

formulation requires eliminating the kernel of the algebra, that is, neutralizing h . This is achieved by the Clebsch decomposition: $\mathbf{v} = \nabla\theta + \alpha\nabla\beta$, $\boldsymbol{\omega} = \nabla\alpha \times \nabla\beta$, $\mathbf{v} \cdot \boldsymbol{\omega} = \nabla\theta \cdot (\nabla\alpha \times \nabla\beta) = \nabla \cdot (\theta\nabla\alpha \times \nabla\beta)$. Thus in the Clebsch parameterization the helicity is given by a surface integral $h = \int d\mathbf{S} \cdot \theta(\nabla\alpha \times \nabla\beta)/2$ – it possesses no bulk contribution, and the obstruction to a canonical realization of (5) is removed [6].

In two spatial dimensions, the Clebsch parameterization is redundant, involving three functions to express the two velocity components. Moreover, the kernel of (5) in two dimensions comprises an infinite number of quantities

$$k_n = \int d^2r \rho \left(\frac{\omega}{\rho} \right)^n \quad (13)$$

for which the Clebsch parameterization offers no simplification (here ω is the two-dimensional vorticity $\omega_{ij} = \varepsilon_{ij}\omega$). Nevertheless, a canonical formulation in two dimensions also uses Clebsch variables to obtain an even-dimensional phase space.

4 Kinematical Grassmann Variables for Vorticity

Rather than using the Gauss potentials (α, β) of the Clebsch parameterization (9) in the description of vorticity (10), we propose an alternative that makes use of Grassmann variables [7]. We write

$$\mathbf{v} = \nabla\theta - \frac{1}{2}\psi_a \nabla\psi_a, \quad (14)$$

where ψ_a is a multicomponent, real Grassmann spinor $\psi_a^* = \psi_a$, $(\psi_a\psi_b)^* = \psi_a^*\psi_b^*$ (the number of components depends on spatial dimensionality). Evidently the nonvanishing vorticity is

$$\omega_{ij} = -\partial_i\psi_a \partial_j\psi_a. \quad (15)$$

Moreover, the canonical 1-form in the Lagrangian that replaces (11) reads

$$L = - \int dr \rho \left(\dot{\theta} - \frac{1}{2}\dot{\psi}_a\dot{\psi}_a \right) - H|_{\mathbf{v}=\nabla\theta-\frac{1}{2}\psi\nabla\psi}. \quad (16)$$

The Hamiltonian retains its (bosonic) form (3), but the Grassmann variables are hidden in the formula for the velocity. From the canonical 1-form, we deduce that (θ, ρ) remains a conjugate pair [see Eq. (7)] and that the canonically independent Grassmann variables are $\sqrt{\rho}\psi$. Thus we postulate, in addition

to the Poisson bracket (7), satisfied by (θ, ρ) , a Poisson antibracket for the Grassmann variables

$$\{\psi_a(\mathbf{r}), \psi_b(\mathbf{r}')\} = -\frac{\delta_{ab}}{\rho(\mathbf{r})} \delta(\mathbf{r} - \mathbf{r}') \quad (17)$$

and this, together with (7), has the further consequence that the following brackets hold:

$$\{\theta(\mathbf{r}), \psi(\mathbf{r}')\} = -\frac{\psi(\mathbf{r})}{2\rho(\mathbf{r})} \delta(\mathbf{r} - \mathbf{r}'), \quad (18)$$

$$\{\mathbf{v}(\mathbf{r}), \psi(\mathbf{r}')\} = -\frac{\nabla\psi(\mathbf{r})}{\rho(\mathbf{r})} \delta(\mathbf{r} - \mathbf{r}') . \quad (19)$$

Then the algebra (5) follows.

One may state that it is natural to describe vorticity by Grassmann variables: vortex motion is associated with spin, and the Grassmann description of spin within classical physics is well known. In the model developed so far, the Grassmann variables have no role beyond the kinematical one of parameterizing vorticity (15) and providing the correct bracket structure. They do not contribute to the equations of motion (1) and (2) for ρ and $\mathbf{v}$, even though they are hidden in the formula (14) for $\mathbf{v}$. Moreover, they satisfy a free equation: from (16) it follows that

$$\dot{\psi} + \mathbf{v} \cdot \nabla \psi = 0 . \quad (20)$$

5 Dynamical Grassmann Variables for Supersymmetry

Thus far the Grassmann variables' only role has been to parameterize the velocity/vorticity (14), (15) and to provide canonical variables for the symplectic structure (5). The equations for the fluid (1), (2) are not polluted by them and they do not appear in the Hamiltonian, beyond their hidden contribution to $\mathbf{v}$. Thus the equation (20) for the Grassmann fields is free.

But now we enquire whether we can add a Grassmann term to the Hamiltonian so that the Grassmann variables enter the dynamics and the entire model enjoys supersymmetry.

We have succeeded for a specific form of the potential $V(\rho)$:

$$V(\rho) = \frac{\lambda}{\rho} \quad (21)$$

and for the specific dimensionalities of space-time, i.e. (2+1) and (1+1). The reason for these specificities will be explained in the next section.

The potential (21), with $\lambda > 0$, leads to negative pressure

$$P(\rho) = \rho V'(\rho) - V(\rho) = -2\lambda/\rho \quad (22)$$

and sound speed

$$s = \sqrt{P'(\rho)} = \sqrt{2\lambda}/\rho \quad (23)$$

(hence $\lambda > 0$). This model is called the "Chaplygin gas".

Chaplygin introduced his equation of state as a mathematical approximation to the physically relevant adiabatic expressions $V(\rho) \propto \rho^n$ with $n > 0$ [8] (constants are arranged so that the Chaplygin formula is tangent at one point to the adiabatic profile). Also it was realized that certain deformable solids can be described by the Chaplygin equation of state [9]. These days negative pressure is recognized as a possible physical effect: exchange forces in atoms give rise to negative pressure; stripe states in the quantum Hall effect may be a consequence of negative pressure; the recently discovered cosmological constant may be exerting negative pressure on the cosmos, thereby accelerating expansion.

5.1 Planar Model

In (2+1) dimensions the Grassmann variables possess 2-components and two real 2×2 Dirac " α "-matrices act on them: $\alpha^1 = \sigma^1$, $\alpha^2 = \sigma^3$. The supersymmetric Hamiltonian is

$$H = \int d^2r \left\{ \frac{1}{2} \rho v^2 + \frac{\lambda}{\rho} + \frac{\sqrt{2\lambda}}{2} \psi \alpha \cdot \nabla \psi \right\}, \quad (24)$$

where it is understood that $\mathbf{v} = \nabla \theta - \psi \nabla \psi / 2$ [7,10]. While the continuity equation retains its form (1), the force equation acquires a contribution from the Grassmann variables

$$\dot{\mathbf{v}} + \mathbf{v} \cdot \nabla \mathbf{v} = \nabla \frac{\lambda}{\rho^2} + \frac{\sqrt{2\lambda}}{\rho} (\nabla \psi) \alpha \cdot \nabla \psi \quad (25)$$

and ψ is no longer free:

$$\dot{\psi} + \mathbf{v} \cdot \nabla \psi = \frac{\sqrt{2\lambda}}{\rho} \alpha \cdot \nabla \psi. \quad (26)$$

These equations of motion, together with (1), ensure that the following supercharges are time independent:

$$Q = \int d^2r \{ \rho \mathbf{v} \cdot (\boldsymbol{\alpha} \psi) + \sqrt{2\lambda} \psi \}, \quad (27a)$$

$$\tilde{Q} = \int d^2r \rho \psi. \quad (27b)$$

They generate the following transformations:

$$\delta \rho = -\nabla \cdot [\rho(\eta \boldsymbol{\alpha} \psi)], \quad \tilde{\delta} \rho = 0, \quad (28a)$$

$$\delta \psi = -(\eta \boldsymbol{\alpha} \psi) \cdot \nabla \psi - \mathbf{v} \cdot \boldsymbol{\alpha} \eta - \frac{\sqrt{2\lambda}}{\rho} \eta, \quad \tilde{\delta} \psi = -\eta, \quad (28b)$$

$$\delta \mathbf{v} = -(\eta \boldsymbol{\alpha} \psi) \cdot \nabla \mathbf{v} + \frac{\sqrt{2\lambda}}{\rho} (\eta \nabla \psi), \quad \tilde{\delta} \mathbf{v} = 0, \quad (28c)$$

where η is a two-component constant Grassmann spinor. The antibrackets of the supercharges produce other conserved quantities:

$$\{Q_a, Q_b\} = -2\delta_{ab}H, \quad (29a)$$

$$\{\tilde{Q}_a, \tilde{Q}_b\} = -\delta_{ab}N, \quad (29b)$$

$$\{\tilde{Q}_a, Q_b\} = \boldsymbol{\alpha}_{ab} \cdot \mathbf{P} + \sqrt{2\lambda} \delta_{ab} \Omega. \quad (29c)$$

Here N is the conserved number $\int d^2r \rho$, $\mathbf{P}$ is the conserved momentum $\int d^2r \rho \mathbf{v}$, and Ω is a center given by the volume of space $\int d^2r$.

5.2 Lineal Model

In (1+1) dimensions the Chaplygin gas equation can be written in compact form in terms of the Riemann coordinates

$$R_{\pm} = v \pm \sqrt{2\lambda}/\rho. \quad (30)$$

Both Eqs. (1) and (2) are equivalent to

$$\dot{R}_{\pm} = -R_{\mp} \frac{\partial}{\partial x} R_{\pm}. \quad (31)$$

It is known that this system is completely integrable [1,11]. One hint for this is the existence of an infinite number of constants of motion

$$I_{\pm}^n = \int dx \rho (R_{\pm})^n, \quad (32)$$

which are time-independent by virtue of (31).

The supersymmetric Hamiltonian makes use of a real, 1-component Grassmann field ψ [12]:

$$H = \int dx \left(\frac{1}{2} \rho v^2 + \frac{\lambda}{\rho} + \frac{\sqrt{2\lambda}}{2} \psi \frac{\partial}{\partial x} \psi \right). \quad (33)$$

The velocity is given by

$$v = \frac{\partial}{\partial x} \theta - \frac{1}{2} \psi \frac{\partial}{\partial x} \psi$$

and the equations of motion for the bosonic variables retain the same form as the absence of ψ , that is, (1), (2) continue to hold. The Grassmann field satisfies

$$\dot{\psi} + R_- \frac{\partial}{\partial x} \psi = 0 \quad (34)$$

and a general solution follows immediately with the help of (31): ψ is an arbitrary function of R_+ :

$$\psi = \Psi(R_+). \quad (35)$$

Thus the system remains completely integrable.

The supersymmetry charges and transformation laws are obvious dimensional reductions of (27)–(28):

$$Q = \int dx \rho R_+ \psi, \quad (36a)$$

$$\tilde{Q} = \int dx \rho \psi, \quad (36b)$$

$$\delta \rho = -\eta \frac{\partial}{\partial x} (\rho \psi), \quad \tilde{\delta} \rho = 0, \quad (37a)$$

$$\delta \psi = -\eta \psi \psi' - \eta R_+, \quad \tilde{\delta} \psi = -\eta, \quad (37b)$$

$$\delta v = -\eta (\psi v)' + \eta R_+ \psi', \quad \tilde{\delta} v = 0. \quad (37c)$$

The algebra of these is

$$\{Q, Q\} = -2H, \quad (38a)$$

$$\{\tilde{Q}, \tilde{Q}\} = -N, \quad (38b)$$

$$\{\tilde{Q}, Q\} = -P - \sqrt{2\lambda} \Omega. \quad (38c)$$

In view of (35), we see that evaluating the supercharges Q and $\tilde{Q}$ on the solution gives expressions of the same form as the bosonic conserved charges (33).

Indeed, we recognize that two charges in (36) are the first two in an infinite tower of conserved supercharges, which generalizes the infinite number of bosonic conserved quantities (32):

$$Q_n = \int dx \rho R_+^n \psi . \quad (39)$$

6 The Origins of Our Models

We have succeeded in supersymmetrizing a specific model – the Chaplygin gas – in specific dimensionalities – the 2-dimensional plane and the 1-dimensional line – leading to nonrelativistic, supersymmetric fluid mechanics in (2+1)- and (1+1)-dimensional space-time. The reason for these specificities is that both models descend from Nambu-Goto models for extended systems in a target space of one dimension higher than the world volume of the extended object. Specifically, a membrane in three spatial dimensions and a string in two spatial dimensions, when gauge-fixed in a light-cone gauge, can be shown to devolve to a bosonic Chaplygin gas in two and one spatial dimensions, respectively [13]. The fluid velocity potential arises from the single dynamical variable in the gauge-fixed Nambu-Goto theory, namely, the transverse direction variable for the membrane in space and the string on a plane. Although purely bosonic Chaplygin gas models in other dimensions can devolve from appropriate Nambu-Goto models for extended objects, for the supersymmetric case we need a superextended object, and these exist only in specific dimensionalities. In our case it is the light-cone parameterized supermembrane in (3+1)-dimensional space-time [14] and the superstring in (2+1)-dimensional space-time [15] that give rise to our planar and lineal supersymmetric fluid models.

One naturally wonders whether an arbitrary bosonic potential $V(\rho)$ has a supersymmetric partner in arbitrary dimensions, and this problem is under further investigation. One promising approach is to consider parameterizations of extended objects other than the light-cone one. It is known that in the purely bosonic case, other parameterizations of the Nambu-Goto actions lead to other fluid mechanical models, and this should carry over to a supersymmetric generalization.

Incidentally, the existence of Nambu-Goto antecedents of the fluid models that we have discussed allows one to understand some of their remarkable properties: complete integrability in the lineal case; existence of further symmetries (which we have not discussed here) and relation to other models (which devolve from the same extended system, but are parameterized differently from the light-cone method) [16].

References

- [1] See, for example, L.D. Landau and E.M. Lifshitz, *Fluid Mechanics*, 2nd ed. (Pergamon, Oxford, 1987).
- [2] This algebra implies the familiar translation algebra for the momentum density $\mathcal{P} = \rho \mathbf{v}$ (= current density). In the fluid mechanical context, the brackets (5) were posited by L.D. Landau, *Zh. Eks. Teor. Fiz.* **11**, 592 (1941) [English translation: *J. Phys. USSR* **5**, 71 (1941)]; see also P.J. Morrison and J.M. Greene, *Phys. Rev. Lett.* **45**, 790 (1980); (E) **48**, 569 (1982).
- [3] C.C. Lin, *International School of Physics E. Fermi (XXI)*, Ed. G. Careri (Academic Press, New York, 1963).
- [4] A. Clebsch, *J. Reine Angew. Math.* **56**, 1 (1859); H. Lamb, *Hydrodynamics* (Cambridge University Press, Cambridge, 1932), p. 248.
- [5] For a discussion of Abelian and non-Abelian Chern-Simons terms, see S. Deser, R. Jackiw, and S. Templeton, *Ann. Phys. (N.Y.)* **140**, 372 (1982); (E) **185**, 406 (1985). In fluid mechanics and magnetohydrodynamics the Abelian Chern-Simons term is known as the fluid or magnetic helicity. It was introduced by L. Woltier, *Proc. Nat. Acad. Sci.* **44**, 489 (1958).
- [6] Some further peculiarities of the Clebsch parameterization and the Chern-Simons term are discussed by S. Deser, R. Jackiw, and A.P. Polychronakos, *Phys. Lett. A* (in press), eprint: physics/0006056. A non-Abelian generalization of the Clebsch parameterization is in R. Jackiw, V.P. Nair, and S.-Y. Pi, *Phys. Rev. D* **62**, 085018 (2000).
- [7] R. Jackiw and A.P. Polychronakos, *Phys. Rev. D* **62**, 085019 (2000).
- [8] S. Chaplygin, *Sci. Mem. Moscow Univ. Math. Phys.* **21**, 1 (1904). [Chaplygin was a colleague of fellow USSR Academician N. Luzin. Although accused by Stalinist authorities of succumbing excessively to foreign influences, unaccountably both managed to escape the fatal consequences of their alleged actions; see N. Krementsov, *Stalinist Science*

- (Princeton University Press, Princeton, 1997).] The same model (21) was later put forward by H.-S. Tsien, *J. Aeron. Sci.* **6**, 399 (1939) and T. von Karman, *J. Aeron. Sci.* **8**, 337 (1941).
- [9] K. Stanyukovich, in *Unsteady Motion of Continuous Media* (Pergamon, Oxford, 1960), p. 128.
- [10] Some of these results are described in unpublished papers by J. Hoppe, Karlsruhe preprints KA-THEP-6-93, KA-THEP-9-93, eprint: hep-th/9311059.
- [11] Y. Nutku, *J. Math. Phys.* **28**, 2579 (1987); P. Olver and Y. Nutku, *J. Math. Phys.* **29**, 1610 (1988); M. Arik, F. Neyzi, Y. Nutku, P. Olver, and J. Verosky, *J. Math. Phys.* **30**, 1338 (1989); J. Brunelli and A. Das, *Phys. Lett. A* **235**, 597 (1997).
- [12] Y. Bergner, in preparation.
- [13] J. Goldstone (unpublished); M. Bordemann and J. Hoppe, *Phys. Lett. B* **317**, 315 (1993); **329**, 10 (1994); R. Jackiw and A.P. Polychronakos, *Proc. Steklov Inst. Math.* **226**, 193 (1999); *Comm. Math. Phys.* **207**, 107 (1999).
- [14] B.S. de Witt, J. Hoppe, and H. Nicolai, *Nucl. Phys. B* **305** [FS23], 525 (1988).
- [15] J. Gauntlett, *Phys. Lett. B* **228**, 188 (1989).
- [16] For more details on these and related topics, see R. Jackiw, eprint: physics/0010042.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

NON-EQUILIBRIUM WORLDLINE DUALITY IN CONDENSED MATTER

R.J. RIVERS

*Blackett Laboratory, Imperial College, Prince Consort Rd.,
London SW7 2BZ, England
E-mail: r.rivers@ic.ac.uk*

The dual “worldline” description of adiabatic Ginzburg-Landau field theory near a phase transition in terms of quasi-Brownian strings and loops is well understood, particularly through the work of Hagen Kleinert and his coworkers. In reality, the implementation of a transition is intrinsically non-equilibrium. We sketch how time-dependent Ginzburg-Landau theory leads to a modified dual string picture. A causal bound on the growth of unstable string prevents the uncontrolled proliferation of string (the Shockley-Hagedorn transition) suggested by the adiabatic approximation.

1 Introduction

The dual worldline approach to time-independent Ginzburg-Landau (TIGL) theory is well understood, significantly through the work of Hagen Kleinert and his coworkers. From the early days of lattice models [1] it has been appreciated that time-independent Ginzburg-Landau (TIGL) theory has a dual representation in terms of Brownian strings. A detailed discussion is given in the textbook by Kleinert [2] and, more recently, in lecture notes by Schakel [3]. An intrinsic ingredient of all these calculations is that they are performed for systems in *equilibrium*, in which the temperature T is fixed at values closer and closer to the critical temperature T_c . For the case that interests us, that of continuous transitions, the instabilities that characterise the transitions are, in this adiabatic dual picture, a consequence of the uncontrolled proliferation of string.

In practice, transitions occur in a finite, often short, time in which the temperature T crosses T_c from its initial value. The adiabatic approximation is only valid away from the transition, when the field can adjust to the changing environment. Close to the transition the adiabatic regime is replaced by an impulse regime, in which the field falls out of step with what its equilibrium behavior would be. In particular, the correlation length cannot grow faster than the relevant speed (of sound) at which the field can order itself and, as a result, is unable to diverge [4,5]. We would expect a realistic dual worldline description in terms of strings to show a similar causal bound that will provide a limit to their proliferation as we pass through a transition.

The question that we shall begin to address here is how the intrinsically non-equilibrium theory has a dual representation in terms of strings and loops, and how the transition is manifest.

2 Equilibrium Ginzburg-Landau Duality

We begin with a brief recapitulation of equilibrium dual theory. The simplest condensed matter TIGL theory is that of a single complex field ϕ , with the Ginzburg-Landau free energy

$$F(T) = \int d^3x \left(\frac{\hbar^2}{2m} |\nabla\phi|^2 + \alpha(T)|\phi|^2 + \beta|\phi|^4 \right), \quad (1)$$

in which the chemical potential $\alpha(T) = \alpha_0\epsilon(T)$, where $\epsilon(T) = (T/T_c - 1)$, vanishes at T_c . We envisage changing $\alpha(T)$ through an external cooling of the system or a change in the pressure of the system that leads to a change in T_c . Such an energy provides a reasonable description of superfluid ^4He , a simplified model for ^3He and a good description of the scalar sector of low- T_c superconductors.

It is convenient to work in spatial units of $\xi_0 = \sqrt{\hbar^2/2m\alpha_0}$ when, on rescaling the field, the self-coupling and the temperature (but leaving ϵ unchanged),

$$\bar{F}(T) = \int d^3x \left(|\nabla\phi|^2 + \epsilon(T)|\phi|^2 + \bar{\beta}|\phi|^4 \right). \quad (2)$$

At temperature $T_0 > T_c$ the *free-field* correlation function that follows from

$$\bar{F}_0(T_0) = \int d^3x \left(|\nabla\phi|^2 + \epsilon(T_0)|\phi|^2 \right) \quad (3)$$

is

$$\langle \phi(\mathbf{x}) \phi^*(\mathbf{0}) \rangle = G_0(r) = \int d^3k e^{i\mathbf{k} \cdot \mathbf{x}} P(k), \quad (4)$$

($r = |\mathbf{x}|$) in which the power spectrum

$$P(k) = \frac{1}{\mathbf{k}^2 + \epsilon(T_0)} = \int_0^\infty d\tau e^{-\tau k^2} e^{-\tau \epsilon(T_0)} \quad (5)$$

has the usual representation in terms of the Schwinger proper-time. In turn, this gives

$$G_0(r) = \int_0^\infty d\tau \left(\frac{1}{4\pi\tau} \right)^{3/2} e^{-r^2/4\tau} e^{-\tau \epsilon(T_0)}. \quad (6)$$

$G_0(0)$, necessary for loops, diverges from the UV singularities at $\tau = 0$, but these can be regulated with a fixed cutoff. Typically, we choose a cutoff of order unity in units of ξ_0 , which will be implicit throughout.

The dual picture is obtained by observing that $(1/4\pi\tau)^{3/2} e^{-r^2/4\tau}$ is the probability distribution for a Brownian “worldline” or, more usefully, a “polymer” path $\mathbf{x}(\tau)$ beginning at the origin $\mathbf{x}(0) = 0$ and ending at $\mathbf{x}(\tau) = \mathbf{x}$,

$$\left(\frac{1}{4\pi\tau} \right)^{3/2} e^{-r^2/4\tau} = \int_{\mathbf{x}(0)=0}^{\mathbf{x}(\tau)=\mathbf{x}} \mathcal{D}\mathbf{x} \exp \left[- \int_0^\tau d\tau' \frac{1}{4} \left(\frac{d\mathbf{x}}{d\tau'} \right)^2 \right]. \quad (7)$$

If we think of the path as having step length unity (in units of ξ_0) then τ is proportional to the length of the path [3] and we shall use τ and path length synonymously. Then $G_0(r)$ is the sum over string paths of all lengths,

$$G_0(r) = \int_0^\infty d\tau \int_{\mathbf{x}(0)=0}^{\mathbf{x}(\tau)=\mathbf{x}} \mathcal{D}\mathbf{x} e^{-S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(T_0)]}, \quad (8)$$

where $S_{\text{eq}}[\mathbf{x}; \tau, \epsilon(T_0)]$ is the equilibrium Euclidean action

$$S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(T_0)] = \int_0^\tau d\tau' \left[\frac{1}{4} \left(\frac{d\mathbf{x}}{d\tau'} \right)^2 + \epsilon(T_0) \right]. \quad (9)$$

The free-field partition function

$$\mathcal{Z} = \int \mathcal{D}\phi \mathcal{D}\phi^* e^{-\bar{F}_0} = \exp[-\text{tr} \ln(-\nabla^2 + \epsilon(T_0))] \quad (10)$$

is just that of a gas of free orientable polymer or string *loops*, whose lengths are labelled by τ , with the one-loop partition function

$$\ln \mathcal{Z} = \int_0^\infty \frac{d\tau}{\tau} \left(\frac{1}{4\pi\tau} \right)^{3/2} e^{-\tau\epsilon(T_0)}. \quad (11)$$

$\epsilon(T_0)$ is understood as the energy/unit length or tension of the string. The factor $\propto \tau^{-5/2}$ in (11) is the length distribution for Brownian strings of length τ . As we drive $T_0 \rightarrow T_c +$ and $\epsilon(T_0)$ to zero it is easier to create strings. If, formally, we take T_0 below T_c then the onset of *negative* tension makes long loops of this unstable string overwhelmingly favoured and the condensation of loops that follows is understood by condensed matter physicists as a Shockley-Feynman transition, and by quantum field theorists as a Hagedorn transition.

The average loop length is

$$\langle \tau \rangle = -\frac{d \ln \mathcal{Z}}{d \epsilon(T_0)} = \frac{\int_0^\infty d\tau (4\pi\tau)^{-3/2} e^{-\tau\epsilon(T_0)}}{\int_0^\infty d\tau \tau^{-1} (4\pi\tau)^{-3/2} e^{-\tau\epsilon(T_0)}}. \quad (12)$$

Yet again, forcing T below T_c makes $\langle \tau \rangle$ diverge from the IR divergence of the integrals at large τ . A formal cutoff at $\tau = \tau_{\max}$ leads to $\langle \tau \rangle = O(\tau_{\max})$ once this happens.

The incorporation of the $\bar{\beta}|\phi|^4$ field self-interaction into equilibrium dual string theory is implemented by the introduction of a repulsive steric string interaction at the points where paths or strings cross, to give an additional action of the form [2]

$$S_{\text{eq}}^{\text{int}}[\mathbf{x}, \tau] = -\bar{\beta} \int_0^\tau d\tau_k d\tau_l \delta[\mathbf{x}(\tau_k) - \mathbf{x}(\tau_l)] \quad (13)$$

for strings parameterized by τ_k and τ_l . Although this changes the weight of string configurations, qualitatively the transition still occurs in this adiabatic picture because of the proliferation of strings.

3 Non-Equilibrium Ginzburg-Landau Duality

In practice, a change of phase is enforced by letting the temperature T change with time, or by varying the critical temperature T_c . Experiments with superconductors and (the neutron bombardment of) ^3He do the former, while

pressure quenches of ^4He do the latter. If we write

$$\epsilon(t) = \epsilon(T(t)) = \frac{T(t)}{T_c(t)} - 1, \quad (14)$$

the adiabatic approximation for the "free" correlation function $G_0(r, t)$ at time t is

$$G_0^{\text{ad}}(r, t) = \int_0^\infty d\tau \int_{\mathbf{x}(0)=\mathbf{0}}^{\mathbf{x}(\tau)=\mathbf{x}} \mathcal{D}\mathbf{x} e^{-S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(t)]}, \quad (15)$$

in which we just make a straightforward substitution of the equilibrium $\epsilon(T_0)$ with $\epsilon(t)$ in $S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(T_0)]$. This is treating the equilibrium pictures as a series of snapshots that can be run together as a continuous film. In this film string lengths increase uncontrollably as we cross the transition. In particular, $G_0^{\text{ad}}(r, t)$ of (6) is simply calculable as the Yukawa correlator

$$G_0^{\text{ad}}(r, t) = \frac{1}{4\pi r} e^{-r/\xi_{\text{ad}}(t)}, \quad (16)$$

where, on rescaling, $\xi_{\text{ad}}(t) = \xi_0/\sqrt{\epsilon(t)}$ diverges as $T(t) \rightarrow T_c$. As we noted earlier, this cannot be the case, since causality alone prevents the correlation length diverging in a finite time. In terms of the dual picture this implies that the production of an infinity of strings in a finite time is equally prohibited.

To see how, we adopt the time-dependent Landau-Ginzburg (TDLG) equation for $\bar{F}$,

$$\frac{1}{\Gamma} \frac{\partial \phi}{\partial t} = -\frac{\delta \bar{F}}{\delta \phi} + \eta, \quad (17)$$

where η is Gaussian thermal noise, satisfying

$$\langle \eta(\mathbf{x}, t) \eta^*(\mathbf{y}, t') \rangle = 2T(t) \Gamma \delta(\mathbf{x} - \mathbf{y}) \delta(t - t'). \quad (18)$$

Let us continue to consider the free-field case, $\bar{F} = \bar{F}_0$. In (17) the natural unit of time is $\tau_0 = 1/\alpha_0 \Gamma$ and, in units of τ_0 and ξ_0 , Eq. (18) becomes

$$\dot{\phi}(\mathbf{x}, t) = -[-\nabla^2 + \epsilon(t)]\phi(\mathbf{x}, t) + \bar{\eta}(\mathbf{x}, t), \quad (19)$$

where $\bar{\eta}$ is the renormalized noise. The equal-time correlation function is constructed from the solution

$$\phi(\mathbf{k}, t) = \int_{-\infty}^t dt' e^{-\int_{t'}^t dt'' (\mathbf{k}^2 + \epsilon(t''))} \bar{\eta}(\mathbf{k}, t') \quad (20)$$

for the spatial Fourier transform of ϕ as

$$\langle \phi(\mathbf{x}, t) \phi^*(\mathbf{0}, t) \rangle = G_0(r, t) = \int d^3k e^{i\mathbf{k} \cdot \mathbf{x}} P(k, t). \quad (21)$$

$P(k, t)$ has a representation in terms of the Schwinger proper-time τ as

$$P(k, t) = \int_0^\infty d\tau \bar{T}(t - \tau/2) e^{-\tau k^2} e^{-\int_0^\tau d\tau' \epsilon(t - \tau'/2)}, \quad (22)$$

where $\bar{T}$ is the renormalized temperature. Assuming that, for sufficiently negative $t < 0$ in the past, $T(t) = T_0$, then^a $\bar{T}(t) = T(t)/T_0$. In turn, this gives

$$G_0(r, t) = \int_0^\infty d\tau \bar{T}(t - \tau/2) \left(\frac{1}{4\pi\tau} \right)^{3/2} e^{-r^2/4\tau} e^{-\int_0^\tau d\tau' \epsilon(t - \tau'/2)}. \quad (23)$$

However, $G_0(r, t)$ can still be expressed in terms of paths, as

$$G_0(r, t) = \int_0^\infty d\tau \bar{T}(t - \tau/2) \int_{\mathbf{x}(0)=\mathbf{0}}^{\mathbf{x}(\tau)=\mathbf{x}} \mathcal{D}\mathbf{x} e^{-S[\mathbf{x}, \tau, t]}, \quad (24)$$

where $S[\mathbf{x}, \tau, t]$ is not the equilibrium action $S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(t)]$, but

$$S[\mathbf{x}, \tau, t] = \int_0^\tau d\tau' \left[\frac{1}{4} \left(\frac{d\mathbf{x}}{d\tau'} \right)^2 + \epsilon(t - \tau'/2) \right]. \quad (25)$$

This differs significantly from (15), which would replace $\epsilon(t - \tau'/2)$ in (25) with $\epsilon(t)$. The integrated tension $\int_0^\tau d\tau' \epsilon(t - \tau'/2)$ does *not* vanish at the transition time. Since it is the uniform vanishing of tension which triggers the avalanche of string production we see already that it will not happen in this case. Also, because of $\bar{T}(t - \tau/2)$, the paths are no longer Brownian in their length distribution.

4 Examples

4.1 The Instantaneous Quench

As a simple, if unrealistic, demonstration that transitions implemented in a finite time do not display singular dual behavior we consider the case of an *instantaneous* quench at time $t = 0$ from a temperature T_0 above the

^aThis slightly cumbersome formalism hides the fact that, conventionally, the fields are rescaled so that the correlation function contains an additional factor of temperature.

critical temperature T_c to absolute zero. That is, $\epsilon(t) = \epsilon(T_0)\theta(-t)$. Simple calculation shows that $G_0(r, t)$ takes the form of (6) for $t < 0$, whereas for $t > 0$

$$\begin{aligned} G_0(r, t) &= e^{2tT_0/T_c} \int_{2t}^{\infty} d\tau \int_{\mathbf{x}(0)=\mathbf{0}}^{\mathbf{x}(\tau)=\mathbf{x}} \mathcal{D}\mathbf{x} e^{-S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(T_0)]} \\ &= e^{2tT_0/T_c} \int_{2t}^{\infty} d\tau \left(\frac{1}{4\pi\tau} \right)^{3/2} e^{-r^2/4\tau} e^{-\tau\epsilon(T_0)}. \end{aligned} \quad (26)$$

After the transition we have a representation in terms of positive tension paths at the *initial* temperature T_0 . The unstable (negative tension) paths, for which $\tau < 2t$, are *totally excluded*. Instead, the instabilities are encoded in the non-singular exponential prefactor.

This lack of IR singular behavior is made even more explicit in a saddle-point approximation for $G_0(r, t)$. For $r^2 > 4\epsilon(T_0)t$ we find

$$G_0(r, t) \sim e^{2tT_0/T_c} \frac{1}{4\pi r} e^{-r/\xi}, \quad (27)$$

where $\xi = \xi(T_0)$ remains frozen in at its initial equilibrium value. Unlike $\xi_{\text{ad}}(t)$ there is no divergence of ξ as we cross $T = T_c$ in this abrupt way.

For the equilibrium theory, $\langle \tau \rangle$ of (12) can also be expressed as

$$\langle \tau \rangle = \frac{G_0(0)}{2G_0''(0)}, \quad (28)$$

where the prime denotes differentiation with respect to r . If we adopt the same definition out of equilibrium, the benign effect of the quench to $T_f = 0$ is again apparent in that $\langle \tau \rangle_t$ agrees with (12) for $t < 0$, but for $t > 0$ we have

$$\langle \tau \rangle_t = \frac{\int_{2t}^{\infty} d\tau (4\pi\tau)^{-3/2} e^{-\tau\epsilon(T_0)}}{\int_{2t}^{\infty} d\tau \tau^{-1} (4\pi\tau)^{-3/2} e^{-\tau\epsilon(T_0)}}. \quad (29)$$

The exponential prefactors have cancelled to reproduce the equilibrium result (12), again at the initial temperature $T = T_0$, but for the absence of unstable loops with a length less than $2t$, to which there is now no reference. Because of the exponential damping of a long string, the dominant string length is $\tau = 2t$.

There is no question of an IR divergence of loop length as naively suggested by the equilibrium theory. We understand the stability of loops with length $\tau > 2t$ in (29) as a causal bound. In our units, negative tension can only

propagate at speed $c = 1$, which happens to be the cold speed of sound in the ϕ field.

This is reinforced by an extension to non-zero final temperature T_f

$$\epsilon(t) = \begin{cases} \epsilon(T_0) > 0, & t < 0, \\ \epsilon(T_f) < 0, & t > 0. \end{cases} \quad (30)$$

For $t > 0$ we now find

$$\begin{aligned} G_0(r, t) = & e^{2t(T_0 - T_f)/T_c} \int_{2t}^{\infty} d\tau \int_{\mathbf{x}(0)=\mathbf{0}}^{\mathbf{x}(\tau)=\mathbf{x}} \mathcal{D}\mathbf{x} e^{-S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(T_0)]} \\ & + \frac{T_f}{T_0} \int_0^{2t} d\tau \int_{\mathbf{x}(0)=\mathbf{0}}^{\mathbf{x}(\tau)=\mathbf{x}} \mathcal{D}\mathbf{x} e^{-S_{\text{eq}}[\mathbf{x}, \tau, \epsilon(T_f)]}. \end{aligned} \quad (31)$$

We see explicitly how the limited length $\tau < 2t$ of string with negative tension $\epsilon(T_f) < 0$ prevents such string from giving a divergent contribution.

Small unstable loops are no longer precluded, but their contribution is finite. The dominant length remains $\tau = 2t$.

4.2 Slower Quenches

Although an instantaneous quench is impossible, more general quenches show similar qualitative behavior.

Specifically, suppose that $\epsilon(t)$ decreases monotonically, with a single zero $\epsilon(0) = 0$ at time $t = 0$. Then strings of limited length $\tau < 2t$, for which $\epsilon(t - \tau/2) < 0$, have “negative” tension. However, for $\tau > 2t$, for which $\epsilon(t - \tau/2) > 0$, strings have segments with negative and positive tension. There is a causal bound $c = 1$ on the speed at which instability can propagate along a string. The string with negative tension, with a contribution that is independent of τ , gives a prefactor growing at least exponentially in time. The effect is to leave only stable string of length $\tau > 2t$ in the τ integral for $\tau > 2t$. This boundedness on unstable string prevents the unlimited production of string suggested by the adiabatic approximation.

Examples of linear quenches are given in our earlier work [6-8], but the framework of duality was not developed in them (although it was identified in the last of them). Rather than go into any detail, we can generalise further by attempting to accommodate the self-interaction which, in itself, leads to a modified mass-behavior $\epsilon(t)$.

5 Back-Reaction

The exponential growth of $G_0(r, t)$ with time in (27) can only be accommodated for a very short period, since $\langle |\phi|^2 \rangle = G_0(0, t)$ must be constrained by the value of the order parameter $\langle |\phi| \rangle$ after the transition, equal to $\sqrt{\bar{\beta}^{-1}/2}$ in the absence of corrections. A rough guide to the maximum time $\bar{t}$, for which the free-field approximation is valid, is that $G_0(0, \bar{t}) = \bar{\beta}^{-1}/2$. As t approaches $\bar{t}$, and thereafter, the reaction of the field with itself will cut off the exponential growth.

It is difficult to see how the dual string picture will survive at later times without further approximation. One indication is through a mean-field (or large- N) approach in which the linear nature of the theory is maintained self-consistently [9]. In this approach (19) is replaced by

$$\dot{\phi}(\mathbf{x}, t) = -[-\nabla^2 + \epsilon_{\text{eff}}(t)]\phi(\mathbf{x}, t) + \bar{\eta}(\mathbf{x}, t), \quad (32)$$

where

$$\epsilon_{\text{eff}}(t) = \epsilon(t) + p\bar{\beta}\langle |\phi|^2 \rangle = \epsilon(t) + p\bar{\beta}G(0, t), \quad (33)$$

and $G(0, t)$ is determined self-consistently^b from (32).

The assumed single zero of $\epsilon(t)$ at $t = 0$ will lead to a zero of $\epsilon_{\text{eff}}(t)$ at $t \approx 0$, and the previous analysis applies. With only a finite length of string with negative tension, there is no singular behavior as we cross the transition.

To have a quantitative estimate of the effect of back-reaction we make a Gaussian approximation for $G(0, t)$ by expanding about the zero of $\epsilon_{\text{eff}}(t)$. Assuming that the quench is not too rapid, we find that, for $t \leq \bar{t}$, the average loop length is

$$\langle \tau \rangle_t \approx \frac{\int_0^\infty d\tau \bar{T}(t - \tau/2)(4\pi\tau)^{-3/2} e^{-(\tau-2t)^2|\epsilon'(0)|/4}}{\int_0^\infty d\tau \tau^{-1} \bar{T}(t - \tau/2)(4\pi\tau)^{-3/2} e^{-(\tau-2t)^2|\epsilon'(0)|/4}}. \quad (34)$$

In this approximation the prefactors encoding the unstable strings cancel, and we need no information about the self-consistent mass, but for the fact that $\epsilon_{\text{eff}}(t) \rightarrow 0$ at large times to stop $G(0, t)$ from growing. We see the peak at $\tau = 2t$ in the length distribution moving clear of the UV endpoint behavior. This continues for later times and $\langle \tau \rangle_t = O(t)$ once the temperature

^bThe coefficient p depends on whether we adopt a mean-field (Hartree) approximation or a large- N limit for $N = 2$.

has become low enough that $\bar{T}(t)$ has suppressed the UV singular behavior. Details are given in our earlier work [7,8].

We stress that this discussion has been restricted to the "first quantised" dual representation of the Ginzburg-Landau theory. However, it is a familiar result from a different viewpoint. For a linear system like (32) it can be shown [10,11] that the mean loop length is

$$\langle \tau \rangle_t = \frac{1}{4\pi n(t)}, \quad (35)$$

where $n(t)$ is the density of *line zeroes* of the complex field ϕ . The relevance of this is that, at later times, the global *vortices* of this $U(1)$ scalar theory can be identified by the line zeroes of their cores [12]. The linear growth of *dual* loop lengths with time corresponds to a $t^{1/2}$ behavior for ϕ -*field line-zero* separation and, when vortices become well-defined, vortex separation. This scaling behavior can be justified [13] for the decay of vortices of ${}^4\text{He}$, and is used to determine initial vortex densities at the ${}^4\text{He}$ transition [14].

6 Conclusions

Our conclusions are simple. In the adiabatic approximation for the (world-line) string representation of scalar field theory, the transition is signaled by the proliferation of string at a Shockley-Feynman/Hagedorn transition, as it becomes unstable (negative tension). In reality, transitions implemented in a finite time are intrinsically non-adiabatic and this picture breaks down. In a more realistic dual string representation at time t after the transition begins, there is a causal constraint at the rate at which instability can propagate along strings. With only short strings and short segments of strings unstable, there is no Shockley-Feynman/Hagedorn transition, since this relies on the instability of arbitrarily long strings.

Acknowledgments

This work is the result of a network supported by the European Science Foundation. I thank Adriaan Schakel for helpful comments.

References

- [1] P.R. Thomas and M. Stone, *Nucl. Phys. B* **144**, 513 (1978).

- [2] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* (World Scientific, Singapore, 1989).
- [3] A.M.J. Schakel, in *Topological Defects and the Non-Equilibrium Dynamics of Symmetry Breaking Phase Transitions*, Proceedings of Les Houches, 1999 NATO ASI (Vol. 549), Eds. Y.M. Bunkov and H. Godfrin (Kluwer Academic Publishers, Dordrecht, 2000), p. 213.
- [4] W.H. Zurek, *Nature* **317**, 505 (1985); *ibid.* **382**, 297 (1996); *Acta Phys. Pol. B* **24**, 1301 (1993).
- [5] W.H. Zurek, *Phys. Rep.* **276**, 177 (1996).
- [6] G. Karra and R.J. Rivers, *Phys. Rev. Lett.* **81**, 3707 (1998).
- [7] R.J. Rivers, *Phys. Rev. Lett.* **84**, 1248 (2000).
- [8] E. Kavoussanaki, R.J. Rivers, and G. Karra, *Cond. Matt. Phys.* **3**, 133 (2000).
- [9] D. Boyanovsky, H.J. de Vega, and R. Holman, in *Topological Defects and the Non-Equilibrium Dynamics of Symmetry Breaking Phase Transitions*, Proceedings of Les Houches, 1999 NATO ASI (Vol. 549), Eds. Y.M. Bunkov and H. Godfrin (Kluwer Academic Publishers, Dordrecht, 2000), p. 139.
- [10] B.I. Halperin, in *Physics of Defects*, Proceedings of Les Houches, Session XXXV 1980 NATO ASI, Eds. R. Balian, M. Kléman, and J.-P. Poirier (North-Holland Press, Amsterdam, 1981), p. 816.
- [11] F. Liu and G.F. Mazenko, *Phys. Rev. B* **46**, 5963 (1992).
- [12] N.D. Antunes, L.M.A. Bettencourt, and W.H. Zurek, *Phys. Rev. Lett.* **82**, 2824 (1999).
- [13] W.F. Vinen, *Proc. Roy. Soc. London A* **242**, 493 (1957).
- [14] M.E. Dodd *et al.*, *Phys. Rev. Lett.* **81**, 3703 (1998); *J. Low Temp. Phys.* **15**, 89 (1999).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

FIELD THEORIES AND THE PROBLEM OF TOPOLOGICAL ENTANGLEMENT IN POLYMER PHYSICS

F. FERRARI

*Physics Department, University of Szczecin, Wielkopolska 15,
70-451 Szczecin, Poland*

E-mail: fferrari@univ.szczecin.pl

*Physics Departments, University of Trento, Povo, and INFN,
Gruppo Collegato di Trento, Italy*

In this contribution some recent advances in understanding the statistical mechanics of topologically linked polymers will be reviewed. This is an interdisciplinary subject in which polymer physics, knot theory, and field theories meet together. Hopefully, we will convince more polymer physicists that field theories and knot theory are useful tools for their research and, on the other side, attract more theoretical physicists and mathematicians to polymer physics.

1 Foreword

Professor Hagen Kleinert is certainly one of the most brilliant contemporary scientists. One can hardly find a sector of theoretical physics in which he has not contributed with outstanding results. His books, constantly updated, are an invaluable source of information for many researchers. Coworking with him is a pleasure, even if sometimes it is not easy to cope with the amount of brilliant ideas and quick solutions of difficult problems which, miraculously, he is able to produce in a short time.

His achievements in polymer physics and in the theory of random walks are too many to be listed here (see for instance Refs. [1,2]). In my opinion, the most important of them has been the formulation of the problem of topological entanglement in terms of Chern-Simons (CS) field theories [2]. Here I will talk about a joint work on the statistical mechanics of topological polymers. This has led to the first quantitative physical prediction that could be extracted from a topological field theory.

2 A Brief Introduction to the Physics of Polymers

2.1 Introduction to Polymers

In our everyday experience we are constantly in contact with synthetic polymers like plastics, rubbers etc., without mentioning the fundamental role for the existence of all living beings played by biopolymers like DNA, proteins, and viruses. The marvelous properties of polymer materials and their outstanding performances attract considerable interest from both chemists and physicists. A starting point to the vast realm of polymers from a physicist point of view is given in Ref. [3]. At a more specialized level, there are other excellent books (see Refs. [1,2,4]). A clear and concise explanation of advanced techniques of polymer physics can also be found in recent review articles, as for example Ref. [5].

Polymers are macromolecules composed of many units consisting of particular molecules called monomers. The latter are able to join together via covalent bonds forming very long chains. Artificial polymers may contain up to 16 000 monomers, while a biopolymer arrives to the astounding number of 10^{10} (ten billions!) monomers. The chains have macroscopic lengths which can be as long as a few meters in the case of biopolymers. On the contrary, in the remaining two dimensions their section is of the same size as the monomers, i.e. it amounts to a few Ångströms. Thus, polymers may be treated as nearly one-dimensional objects.

Polymerized materials appear in a variety of forms with strikingly different features. For instance, plastics and rubbers are both solids, but certainly they do not have the same behavior under stretching. For this reason, in order to classify polymers, it is better to introduce the concept of phases instead of using the more traditional division in gas, liquid, and solid states. Four possible phases of polymers are distinguished [3]: viscous, elastic, semi-crystalline, and glassy phases. In the following, the viscous phase will be mainly discussed, which is liquid and contains the physically interesting cases of polymer melts and solutions.

2.2 Polymers and Physics

What opens the way to physics is the fact that, to a large extent, the differences in the macroscopical properties of polymer systems are not due to

the chemical composition of the monomers.^a This is the so-called universal behavior of polymers. The main reason for such a universality is that polymers are almost macroscopical objects in one direction. At scales which are much larger than the monomer size, they can be considered for all practical purposes as very long and flexible tubes. Incidentally, this makes the study of polymer liquids easier than that of normal liquids, where the motion of a molecule depends on the motion of its nearest neighbors and on that of its next-to-nearest neighbors etc. Instead, the action of each monomer is "averaged" over all the chain length.

The flexibility of polymers is negligible at short distances, but after some monomer lengths it starts to show up. Thus, there should be a critical length a such that any segment shorter than a can be regarded as rigid. One usually calls a the length of the Kuhn segment, after its discoverer; it is a parameter which depends on the temperature. Typical experimental values of a are 1nm for the simplest synthetic chains and about 100 nm for DNA. It can also be shown that the memory of the orientation of a given monomer gets completely lost after a distance greater than $a/2$. As a consequence, it is possible to describe a vast class of polymers using a model of freely joint segments of equal length a . In this picture polymers are treated as random chains subjected to thermal fluctuations. Of course, there are a few different mechanisms of flexibility, which make some macromolecules more rigid than others. Some methods to take into account the rigidity of the joints are reported in Refs. [2,6].

Still an important question has been left unanswered. What is driving the behavior of polymers at large scales if not their chemical compositions? The answer is: the entropy of the system. In fact, the computation of the entropy and of the free energy is the fundamental problem of the statistical mechanics of macromolecules. To understand how entropy is relevant for describing polymer systems, it is sufficient to consider the example of a single chain. The trajectory of an unstretched chain is able to entangle in an infinite number of different configurations like the path of a particle subjected to random walk. When it starts to get stretched, however, the entanglement freedom decreases, so that this simple system moves from a more probable state (realized in more different ways) to a less probable one (realized in fewer

^aOf course, this claim should not be taken in a too strict sense. For instance polyelectrolytes, i.e. polymers in which the monomers are charged, have a different behavior with respect to uncharged polymers.

ways). In the limit in which the chain is completely stretched, it becomes a straight line which has no freedom at all. Clearly, a stretched chain will attempt to return to the most probable state, reacting to the stress with an opposite elastic force. Indeed, the elasticity of polymer systems can be explained in terms of entropy alone.

2.3 Polymers and Field Theories

Let us assume that our system of polymers is at equilibrium and at constant temperature T . To compute the entropy and free energy, it is necessary to sum over all possible trajectories of polymers using numerical or path-integral techniques.^b In either case, we apply the above large-scale picture of polymers as random chains. The chain length L , the number n of freely joint segments and the Kuhn length a are related together by $L = na$.

Let us remember at this point that we are talking about polymers in the liquid state. In the standard theory of fluids, it is better to deal with densities and currents instead of following the trajectory of each molecule. In the case of polymers it is also more convenient to work with monomer densities and currents. Besides, the density of monomers is easily controlled in the laboratory, while it is more difficult to measure the average polymer conformations in space. The passage from trajectories to monomer densities is analogous to the passage from first to second quantization in quantum mechanics. This procedure will be explained later.

One advantage of having a field theoretic model of polymer liquids is that it is possible to apply sophisticated techniques already developed in high energy and condensed matter physics to study critical phenomena [1]. In fact, even at constant temperature polymers have still a non-trivial critical behavior with respect to physical parameters like polymer length and monomer density. Field theories have played a crucial role in the achievement of a satisfactory microscopical model of the physics of linear (unentangled) open chains [7]. Some computations of the critical exponent of polymers with field theoretical methods are performed in Ref. [2].

Last but not least, we may expect the existence of something fundamental like a gauge principle behind the universality of the behavior of polymers. The first gauge field theory of polymers has been constructed in Ref. [8]. In the following we will see that topologically linked polymers are deeply related to

^bPath-integral techniques imply a continuous limit in which the finite monomer size is neglected because it is very small compared with the total length of the polymer.

topological gauge field theories.

3 The Problem of Topological Entanglement in Polymer Physics

3.1 Polymers and Knot Theory

As mentioned in the previous section, the present understanding of the statistical mechanics of linear open chains is relatively satisfactory. However, with increasing monomer densities, polymers find it convenient to entangle for entropy reasons. In principle, higher densities can be tackled in field theory using scaling arguments. However, in this way one may easily overlook the appearance of new effects like for instance the topological ones. In fact, both in natural and artificial substances there is an abundance of polymer rings, in chemistry also called catenanes, which often are linked together to form non-trivial topological configurations. In vivo, for example, the shape of DNA is usually that of a ring. On the other side, in industrial processes the formation of knots is controlled in order to obtain materials with desired viscoelasticity properties. Since trajectories of polymers are self-avoiding and cannot penetrate each other, once a system of polymers is created in a given topological state, this cannot be changed. Thus there are real topological constraints which remain stable in time. Again, entropy considerations explain the differences in the behavior of open and closed chains. These are due to the fact that the topological constraints reduce the possible configurations of the system.

3.2 Statistical Mechanics of Topologically Linked Polymers

Most attempts to derive a microscopical model of topological entanglement in polymer physics are based on the so-called Edwards' approach in which the polymers are considered as fluctuating random chains or rings and where the entropy is computed via path-integral techniques [9]. Moreover, one starts with open chains having fixed ends at the points $\mathbf{x}_1, \mathbf{x}'_1, \dots, \mathbf{x}_N, \mathbf{x}'_N$, respectively. The case of closed polymers is recovered at the end in the limit of coinciding end points $\mathbf{x}_1 = \mathbf{x}'_1, \dots, \mathbf{x}_N = \mathbf{x}'_N$.^c

Since polymers are almost one-dimensional objects linked together, the language of knot theory is the most adequate to discuss their topological

^cThe presence of fixed points is certainly a limitation because it is unphysical. Indeed, in the laboratory polymers fluctuate freely. However, this shortcoming may be eliminated by performing an average over all possible positions of the fixed points [10].

states. Accordingly, we define a link as a collection of N circles in $3-d$ space. Two links are said to be equivalent if they can be deformed one in another without breaking any line. It is possible to classify inequivalent links with the help of topological invariants. The latter are numbers remaining constant under any continuous transformation which maps a given link to an equivalent one. Clearly, topological invariants which explicitly depend on the trajectories of the polymers are needed to distinguish the topological configurations of a system of polymers. For physical purposes it is enough to have a finite set of sufficiently powerful topological invariants $\{\chi\} = \chi_1, \chi_2, \dots$

At this point we consider N linked polymers with trajectories $P_1, \dots, P_N$ of lengths $L_1, \dots, L_N$, respectively. The topological constraints are imposed by requiring that the topological invariants $\{\chi\}$ take given values $\{m\} = m_1, m_2, \dots$. The fundamental problem of computing the entropy of the above system is solved once the following configurational probability is known:

$$G_m(x; L) = \sum_{\substack{\text{all paths } P_1, \dots, P_N \\ \text{of lengths } L_1, \dots, L_N \\ \text{and ends in } \mathbf{x}_1, \dots, \mathbf{x}_N}} \prod_{\sigma} \delta(\chi_{\sigma}(P_1 \dots P_N) - m_{\sigma}) e^{-\frac{V}{kT}}. \quad (1)$$

Here, k is the Boltzmann constant and V is the potential energy of the system, which will be specified later. The topological relations are fixed by Dirac δ -functions.

In principle, all mathematical tools to compute Eq. (1) can be borrowed from the theory of knots, but in practice the link invariants $\{\chi\}$ are given in the form of polynomials of one, two or three variables which have no evident connection to the physical conformations of the polymers. For this reason, the possibility of constructing knot invariants from Wilson loop amplitudes of field theories [11] has been welcomed with excitement. In fact, these amplitudes are correlation functions of gauge invariant and metric independent operators which explicitly contain the trajectories of closed curves in $3-d$. Alas, in mathematics it is sufficient to deal with static trajectories, but in the present context one has also to sum over all polymer conformations. To this purpose, the topological invariants obtained from field theories are too complicated to allow the evaluation of the configurational probability (1) in any closed form.

4 The Gaussian Approximation

4.1 The Gaussian Linking Number

Until now, the only topological invariant which has been successfully incorporated in the Edwards' approach is the Gaussian linking number (GLN) which is also called intersection number. For any couple of non-intersecting paths P_1 and P_2 it is defined as follows:

$$\chi(P_1, P_2) = \frac{1}{4\pi} \int_0^{L_1} ds_1 \int_0^{L_2} ds_2 \dot{\mathbf{x}}^1(s_1) \cdot \left[\dot{\mathbf{x}}^2(s_2) \times \frac{\mathbf{x}^1(s_1) - \mathbf{x}^2(s_2)}{|\mathbf{x}^1(s_1) - \mathbf{x}^2(s_2)|^3} \right]. \quad (2)$$

Here P_1 and P_2 have been parameterized in the standard way as curves in space. The significance of the GLN has been discussed for instance in Chap. 16 of Ref. [2].

The GLN has the advantage to depend explicitly on the polymer trajectories, but it is relatively weak in distinguishing different topological configurations. Strictly speaking, it can be used with high accuracy only when the monomer density is so low that the polymers cannot be linked together in a too complex way.

4.2 The Interactions Acting on the Monomers

Here we are rather lucky, because only the so-called repulsive steric interactions act on the monomers and thus contribute to the potential energy V of Eq. (1). These short range forces are responsible for the fact that the polymers cannot penetrate each other. When two monomers get too close, they experience an infinite potential barrier which causes the repulsion. Thus, the potential energy of the polymers can be written in terms of Dirac δ -functions:^d

$$V = \sum_{i,j=1}^N \frac{v_{ij}^0}{2a^2} \int_0^{L_i} ds_i \int_0^{L_j} ds_j \delta^{(3)}(\mathbf{x}^i(s_i) - \mathbf{x}^j(s_j)). \quad (3)$$

This completes the definition of the configurational probability given in Eq. (1). Are there really no other relevant interactions? The correct answer is yes, there are still the topological interactions necessary to maintain the system in a given topological state. To see how topological forces work

^dWe suppose that the energy involved in the thermal fluctuations is so high that the monomers do not see the finest details of the two-body potentials.

on monomers, we rewrite the product of δ -functions in Eq. (1) in the Fourier representation as follows:

$$\prod_{\sigma} \delta(\chi_{\sigma}(P_1..P_N) - m_{\sigma}) = \int \frac{d\lambda}{2\pi} \exp \left\{ \sum_{\sigma} i\lambda_{\sigma} [m_{\sigma} - \chi_{\sigma}(P_1..P_N)] \right\}. \quad (4)$$

Physically, λ_{σ} denotes the chemical potential corresponding to the topological number m_{σ} . At this point we assume that the topological constraints are imposed in the Gaussian approximation i.e. by means of the GLN's $\chi(P_i, P_j)$, $i \neq j = 1, \dots, N$. We see that now, in the exponent of Eq. (4), the $\chi(P_i, P_j)$ appear as two-body potentials given by Eq. (2).

4.3 A Path Integral Model of Polymers

For simplicity, only the case of two polymers will be treated here. The generalization to a system of N polymers can be found in Ref. [12]. We work in the Gaussian approximation so that there is only one topological number m corresponding to the GLN $\chi(P_1, P_2)$. Now it is possible to write explicitly the configurational probability of Eq. (1) in terms of path-integral sums over the trajectories P_1 and P_2 :^e

$$G_m(x; L) = \lim_{\substack{\mathbf{x}^1 \rightarrow \mathbf{x}'^1 \\ \mathbf{x}^2 \rightarrow \mathbf{x}'^2}} \int_{\mathbf{x}'^1}^{\mathbf{x}^1} \mathcal{D}\mathbf{x}^1(s_1) \int_{\mathbf{x}'^2}^{\mathbf{x}^2} \mathcal{D}\mathbf{x}^2(s_2) e^{-(\mathcal{A}_0 + V)} \delta(\chi(P_1, P_2) - m), \quad (5)$$

where $\mathcal{A}_0$ is given by:

$$\mathcal{A}_0 = \frac{3}{2a} \sum_{i=1}^2 \int_0^{L_i} \dot{\mathbf{x}}^i{}^2 ds_i. \quad (6)$$

We notice that, following Edwards' approach, the configurational probability has been defined here starting from a system of open polymers^f and then taking the limit of coincident extrema (see Section 3.2).

4.4 From Polymers to Fields

First of all, one observes that the trajectory of a randomly fluctuating chain and the trajectory of a particle subjected to a Brownian motion have many

^eHereafter, the temperature will be incorporated in other constant numerical factors and will thus be ignored.

^fIn this case $\chi(P_1, P_2)$ is no longer a topological invariant.

similarities. Indeed, the action of Eq. (6) describes the free random walk of two particles. This duality between polymers and particles is explained in details in Ref. [1]. In this way, the statistical problem of polymers becomes equivalent to that of particles subjected to self-avoiding random walks which are constrained to satisfy given topological relations. At this point, the desired final mapping to field theories could be performed in principle exploiting the wave-particle duality present in statistical mechanics as in quantum mechanics. However, a first technical difficulty arises: In this “second quantization” procedure one encounters expressions of the kind $\log Z$ or Z^{-1} , where Z is the partition function of a Landau-Ginzburg field theory interacting with topological fields. Such nonlinear terms can be simplified exploiting the well known limits (method of replicas): $\log Z = \lim_{n \rightarrow 0} (Z^n - 1)/n$ and $1/Z = \lim_{n \rightarrow 0} Z^{n-1}$.

The presence of a Dirac δ -function in the configuration probability of Eq. (5) is not a terrible obstacle, since it may always be put in the more convenient form of Eq. (4) after a Fourier transformation. However, after doing that, one finds that the resulting particle action (i.e. the former polymer action) is non-Markovian. Moreover, even in the Gaussian approximation the topological interactions give rise to a potential which couples the trajectories in a nonlinear and complicated way (see Eq. (2)). For this reason, the “second quantization” of such a system has been a difficult long-standing problem for more than twenty years, which could be solved only recently with the help of CS field theories [13].

5 A Field Theoretical Model of Polymers

Skipping all details, which can be found in Refs. [12,13], the above “second quantization” procedure arrives at a model of linked polymers which consists of N Landau-Ginzburg field theories of the kind

$$\mathcal{A}_{\text{L-G}} = \sum_{i=1}^N \int d^3x \left[|\nabla \psi_i|^2 + m_i^2 |\psi_i|^2 \right], \quad (7)$$

with $i = 1, \dots, N$, which are minimally coupled to Abelian CS fields.[§] The latter “propagate” the long-range topological interactions. In the language of

[§]Actually, things are a little bit more complicated than that: According to the method of replicas, anyone of the above N systems must be replicated a number n_i of times. The physical configurational probability is then obtained in the limit of zero replicas.

fields, the monomer density of the i -th polymer is $|\psi_i|^2$, while the masses m_i are Boltzmann factors controlling the polymer lengths. The polymer lengths approach infinity in the limit of zero masses. If also the repulsive steric interactions are considered, one should add suitable quartic interaction terms for the complex fields ψ_i, ψ_i^* .

6 Applications and Conclusions

Field theories provide an exact description of a system of N fluctuating polymers in terms of a topological Landau-Ginzburg model. Previous difficulties of the Edwards' approach have been solved by decoupling the nonlinear interactions between the trajectories due to the topological potential (2) with the help of Abelian CS fields. CS field theories are topological and enjoy many special properties. For instance, they are supposed to have only finite radiative corrections and their Hamiltonian is zero in the absence of couplings with matter fields. As a consequence, it is licit to suppose that the topological interactions do not change the critical properties of polymers. Indeed, this has been explicitly checked in a semi-classical approximation, valid in case the monomer density is sufficiently high to be considered as uniform [14]. In the same approximation, it has been shown that topological forces are attractive in agreement with experimental observations.

Moreover, the square average GLN $\langle m^2 \rangle$ of two fluctuating polymers may be exactly computed by field theoretical methods [10]. With some approximations, this result has been applied to estimate the average square number of intersections $\langle m^2 \rangle$ formed in a solution by a test polymer P of length L with the other polymers [10]:

$$\langle m^2 \rangle \sim \frac{a^2 \rho L}{18 m_a \pi^{\frac{1}{2}}}. \quad (8)$$

Here ρ represents the average mass density of the polymers per unit volume and m_a is the mass of a single segment. Eq. (8), which is valid in the limit $L \gg 1$, provides an approximate formula for the probability of knot formation inside polymer solutions. The relation (8) has been obtained after averaging over all possible fixed points (see Section 3.2). Curiously, if a segment of P is anchored at a given location $\mathbf{x}$, $\langle m^2 \rangle$ does no longer depend on L ! This difference can be explained by the fact that topological interactions are attractive. Since knotted polymers get closer than unknotted ones, the chance increases that they become more and more topologically entangled. On the

contrary, if their trajectories are fixed at some points, the attractive forces become irrelevant.

References

- [1] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [2] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [3] A.Yu. Grosberg and A.R. Khokhlov, *Giant Molecules* (Academic Press, San Diego, 1997).
- [4] P.G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca NY, 1979); M. Doi and S.F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, 1986).
- [5] A.L. Kholodenko and T.A. Vilgis, *Phys. Rep.* **298**, 251 (1998); S. Nechaev, eprint: cond-mat/9812105.
- [6] A.L. Kholodenko, *Ann. Phys. (N.Y.)* **202**, 186 (1990).
- [7] P.G. de Gennes, *Phys. Lett. A* **38**, 339 (1972); J. des Cloiseaux, *Phys. Rev. A* **10**, 1665 (1974); V.J. Emery, *Phys. Rev. B* **11**, 239 (1975).
- [8] M.G. Brereton and S. Shah, *J. Phys. A: Math. Gen.* **13**, 2751 (1980).
- [9] S.F. Edwards, *Proc. Phys. Soc. London* **91**, 613 (1967); *J. Phys. A* **1**, 15 (1968).
- [10] F. Ferrari, H. Kleinert, and I. Lazzizzera, *Eur. Phys. J. B* **18**, 645 (2000); eprint: cond-mat/0003355.
- [11] E. Witten, *Commun. Math. Phys.* **121**, 351 (1989).
- [12] F. Ferrari, H. Kleinert, and I. Lazzizzera, eprint: cond-mat/0005300.
- [13] F. Ferrari and I. Lazzizzera, *Phys. Lett. B* **444**, 167 (1998).
- [14] F. Ferrari and I. Lazzizzera, *Nucl. Phys. B* **559**, 673 (1999).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

THE ROLE OF TOPOLOGICAL EXCITATIONS AT SECOND-ORDER TRANSITIONS

L.M.A. BETTENCOURT

*Theoretical Division, Los Alamos National Laboratory,
Los Alamos, NM 87545, USA*

E-mail: lmbett@lanl.gov

The identification of the collective degrees of freedom relevant for the description of a given macroscopic (thermo)dynamic behavior is a broad objective across branches of physics. Statistical models and field theories describing critical phenomena lead to several different kinds of collective excitations. We discuss the role of topological excitations in continuous phase transitions, including recent developments in 3D. The latter allow in principle for the construction of dual descriptions of the phase transitions in terms of these degrees of freedom and help make contact with new experiments aimed at measuring topological defect formation and evolution.

1 Introduction

Topological excitations are features of the spectrum of models describing critical (thermo)dynamics of many important systems. Examples are superfluids and superconductors, (liquid) crystals, magnets and models of high energy particle physics, which predict phase transitions in the early Universe.

Topological defects are collective excitations, largely independent of the microscopic details of the theory. Their nature is instead determined by broad characteristics of the underlying model like its symmetries and dimensionality. Thus they give a convenient mesoscopic description of the system. Qualitatively their importance is that they are in many known cases thermodynamically inexpensive vehicles of long-range disorder. For this reason they are typically associated with order-disorder transitions.

Topological excitations are known to be important in low spatial dimensions, preventing long-range order to set in. In three spatial dimensions they

were also suggested by Onsager [1] and Feynman [2] as the vehicles of disorder responsible for destroying the superfluid state of liquid ^4He . In three spatial dimensions order-disorder transitions in $O(N)$ models are of second order and are well described by renormalization group methods. A description of the critical phenomenon in terms of topological excitations seems therefore unnecessary.

Independently of these considerations topological defects were proposed as a means to solve several cosmological problems [3,4]. In this context, the problem of determining densities and other properties of topological defects lead to the design of experiments that search directly for these quantities. Such experiments have now been performed in a very large range of materials and conditions, including superfluids (^4He [5] and ^3He [6,7]), high- T_c superconductors [8], liquid crystals [9], and hopefully in the near future nonlinear optical systems [10] and atomic Bose-Einstein condensates.

Motivated by these questions we have been seeking to understand the behavior of topological excitations at second-order transitions. This paper will describe some of our findings together with their relevance for a dual description of the critical phenomena in terms of topological excitation creation and/or proliferation. The dual formulation of critical phenomena in terms of topological excitations received many important contributions by Hagen Kleinert. These issues are reviewed in his book [11], to which the reader is referred.

2 Low-Dimensional Cases: Defects and Long-Range Order

The simplest example of a defect is a domain wall in the Ising model (or in a real $\lambda\phi^4$ -field theory). The domain wall divides regions of space where the system is in one of two energetically equivalent minima. The domain wall costs a finite amount of energy per unit area (its tension). In the Ising model there is an exact equivalence between a description of the system in terms of domain walls (i.e. only the sites where neighboring spins anti-align) and directly in terms of the spins themselves. The model has a second-order phase transition for $D \geq 2$.

In 1D, domain walls are so likely thermodynamical that they subsist in the system all the way down to $T = 0$ in the infinite-volume limit. For this reason the system never displays long-range order at finite temperature and a second-order transition never occurs. This is the essence of the Mermin-Wagner theorem.

In 2D the Ising model can be exactly solved. The system has a second-order transition which is associated with domain wall proliferation. The existence of walls implies local disorder in the sense that at either side of the wall both values of the spin are realized. If a wall is small and closed onto itself this disorder is local and the long-range order of the state will subsist - this situation is therefore characteristic of the symmetry broken phase, below T_c . Conversely, if a wall can be produced that crosses the volume, independently of the size of the system, then long-range order has been destroyed. This is what happens at and above T_c .

Complicating the model so that the magnitude of the spins is also a degree of freedom (resulting in e.g. a $\lambda\phi^4$ -model) results in a phase transition of the same universality class - only dimensionful quantities like the value of T_c or of correlation length change but not their functional variation of temperature. The change in critical temperature can be traced back to the change of the wall tension in the new model.

As we have seen in the Ising model the description of the critical phenomenon in terms of the proliferation of walls or directly in terms of spins is exact. This ceases to be true in more general circumstances.

The next important example is the Kosterlitz-Thouless (KT) transition. It deals with the $O(2)$ model in 2D. As we have seen above this model allows for vortex solutions, in addition to other collective degrees of freedom such as spin waves and quasi-particles. Kosterlitz and Thouless [12] and Berezinskii [13] suggested that the transition in this model (between a disordered state at high temperature and a state with algebraic order at low temperature) proceeds by vortex pair separation, which, as we discussed above, leads to long-range disorder.

At long distances, vortex solutions in 2D behave as unscreened point charges, i.e. they have a $\log(r/a)$ potential. By mapping the vortex charge to the electric charge one can establish the equivalence between a gas of vortices and the Coulomb gas in 2D. In this process the remaining excitations of the model were neglected. Thus the statement that the Coulomb gas describes the transition in the $O(2)$ model is equivalent to the statement that other degrees of freedom are irrelevant in the critical region. This is a much stronger statement than the equivalence between the domain wall and the spin description in the Ising model.

The Coulomb gas has a well-known transition between an insulator and a conductor state. Conduction is associated with the presence of free charges in the plasma whereas the insulator state is characterized by dipole bound

states.

At low temperatures there is an energetic suppression of vortices. These can only occur as bound pairs, which cost a vanishing amount of energy with separation. As the temperature is increased pairs become bigger and more frequent. As a result, screening of charges becomes more and more effective. This process continues as the temperature is increased until a critical temperature is reached such that pairs of any size can be created and the system becomes conducting.

This effect can be seen by estimating the free energy of vortices. Start below the KT transition and estimate the free energetic cost of adding a new charge dipole of a given separation R to the system. The bare interaction can be written as

$$V(R) = 2\pi\rho_s \log(R/a). \quad (1)$$

On the other hand, the number of states Ω available to the pair is $\Omega = 2\pi VR/a$, where the volume V accounts for all places where the first charge of the pair can be placed and the perimeter factor arises for all the sites the second charge can take, at a distance R from the first in 2D. Thus the free energy F for the new pair is estimated to be

$$F(R) \simeq V(R) - k_B T \log(\Omega). \quad (2)$$

There is always a temperature $T_{KT} = 2\pi\rho_s/k_B$ at which it becomes possible to create a pair of arbitrary separation. Of course, in rigorously computing T_{KT} one needs to take into account the interaction between the charges of the pair and those of the existing plasma. This can be done rigorously via the Kosterlitz-Thouless recursion relations which account for the screening of the interaction of the new pair due to the polarizable medium. This effect makes the potential weaker than the bare one, as can be expected on general grounds. Thus the Coulomb gas together with the original models connected to it by duality can in fact be solved analytically. Their predictions have been confirmed with great success, e.g. in ^4He films and 2D superconductors [14].

We end this section by noting the fundamental role of the topological excitations in the examples considered above in bringing about the critical phenomenon. Moreover the description in terms of these excitations is simpler, which made it possible to solve the resulting dual models analytically.

In what follows we will discuss whether these advantages generalize to 3D. In 3D, $O(N)$ models always display second-order transitions, at which

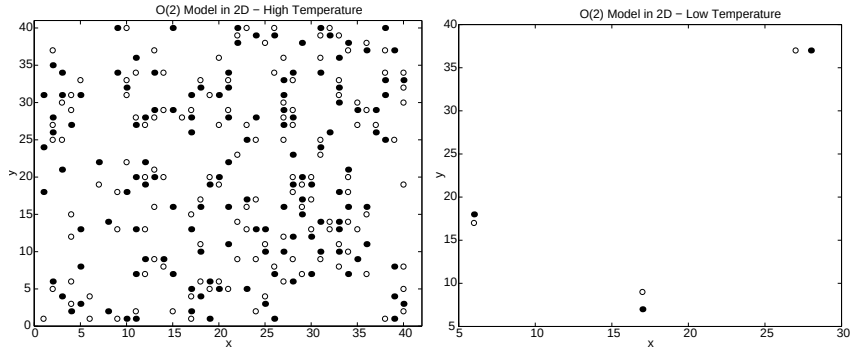


Figure 1. Typical vortex (full circles) and antivortex (open circles) configurations at high and low temperatures in the 2D XY model. At low temperatures all vortices are in pairs and the material is an insulator. The transition to a conductor proceeds by the nucleation of unpaired vortices.

the characteristic length scale diverges. Then a scaling hypothesis together with renormalization group methods allows us to compute critical exponents, which stand in excellent agreement with experiments. The description of the transition in terms of topological excitations seems therefore superfluous unless it can generate new information not easily describable in terms of field correlators.

Recently it has become clear that the existence of a description of the critical phenomena in terms of topological excitations can indeed be investigated quantitatively.

3 What Happens in Three Spatial Dimensions?

The exploration of the role of topological excitations in 3D has recently received much attention [15-18]. While it is still too early to form the complete picture, many properties of the model for $N = 2$ are now known. This section especially comprises a cursory description of this recent progress.

Very much independent of the details of the theory of second-order transitions questions arose in cosmology, first formulated by Kibble [3,4], about the possibility of forming topological defects at cosmological phase transitions. Motivated by these questions several experiments in condensed matter systems were developed to test topological defect formation in the laboratory. Defect densities can be related to the size of the correlation length in the vicinity of the critical point. But where do topological defects come from

and why is it possible that almost no defects are formed if the system is cooled from $T < T_c$?

It is actually quite easy to answer these questions, at least in the context of specific models. Consider for example ^4He , one of the systems where defect formation experiments have been performed [5]. The superfluid transition is excellently described by an $O(2)$ model in 3D. This model has a simple partition function which can be readily sampled. One can then identify and characterize topological excitations in this model as a function of, say, temperature. The results are shown below. The model is also relevant for extreme type II superconductors such as all high- T_c materials.

To gauge our expectations let us follow the simple argument [19] used above to derive the conductor-insulator transition in the Kosterlitz-Thouless case. We will do so with a twist: it is extremely difficult to account for the detailed interaction between string segments because of the many configurations a string of a given size can take. We will therefore neglect them, but consider a simple string that has a given finite energy per unit length (a tension), which we call σ . The partition function is

$$Z = \mathcal{N} \int dE \Omega(E) e^{-\beta E}, \quad (3)$$

where $\Omega(E)$ is the number of configurations with a given energy E . Since we are considering non-interacting strings we only need to consider the number of configurations of a string of length $l = E/\sigma$.

To proceed we observe that a free string is equivalent to a gas of Brownian random walks. Furthermore we require that all string loops are closed. We regularize the problem by taking the step size to be of a given length a , which should be at least of the order of the string's core size. Then the number of configurations $\Omega(l)$ is

$$\Omega(l) = \frac{V}{a^3} A \left(\frac{l}{a} \right)^{-3/2} \left(\frac{l}{a} \right)^{-1} z^{l/a}, \quad (4)$$

where we purposefully separated several different contributions. From left to right, the first factor accounts for the number of possible starting points, the second for the probability that the string will return to its origin in l/a steps, the third removes overcounting since any point along the string is a legitimate starting point, and the forth and final accounts for the number of configurations of a walk of l/a steps, given that each step has access to z states (z is, for example, the coordination number of the lattice).

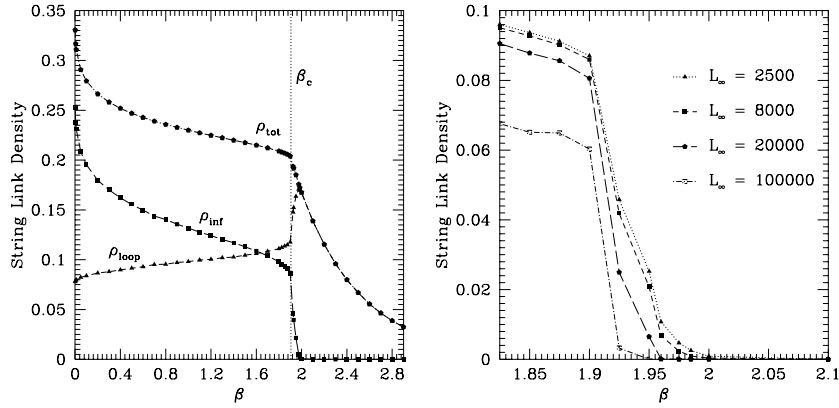


Figure 2. String densities as a function of inverse temperature β (left), for long strings ρ_{inf} , short loops ρ_{loop} , and total ρ_{tot} . The transition proceeds by the creation of long strings. The long string density jumps (right) at $T = T_c$, as the length scale is taken to be larger.

The present exercise is also valid for domain walls in 2D. Walls unlike most strings experience short range interactions and for them the Brownian walk approximation is actually much better.

Notice that the number of configurations is enormous: it grows exponentially with the string's length. It is the extraordinary configurational entropy of extended topological excitations that makes them likely thermodynamically. The partition function can therefore be written as

$$Z = \mathcal{N} \sum_{L/a=1}^{\infty} \Omega(l/a) e^{-\beta l} = A' \sum_{L/a=1} \left(\frac{l}{a}\right)^{-5/2} e^{-\beta L \sigma_{\text{eff}}}, \quad (5)$$

where

$$\sigma_{\text{eff}} = \sigma - \frac{T \log(z)}{a} \equiv \sigma \frac{T_H - T}{T_H} \quad (6)$$

is the effective tension or free energy of the string per unit length. Notice that strings are exponentially suppressed at low temperatures because of their energetic cost. Thanks to the entropic contribution, however, there is a critical (Hagedorn) temperature $T_H = \sigma a / \log(z)$, at which the string length distribution becomes scale invariant and long strings are no longer exponentially suppressed.

This transition is analogous to both the conductor insulator transition in the Kosterlitz-Thouless picture (the interacting nature of strings will be discussed below) and the wall percolation transition in the 2D Ising model.

Notice that the argument for a defect transition *per se* does not tell us the order of the transition (between analytic or second order). If a defect percolates the volume, it is clear that fluctuations of the field modulus can occur over arbitrarily large length scales - the transition must therefore be associated with a diverging length scale and be second order.

Non-interacting strings are rare. They exist e.g. at critical coupling in the Abelian Higgs model. What happens when strings do interact? The simplest case to discuss is again the vortex string in the $O(2)$ model. As discussed above, these interact with long-range $\log(R/a)$ potential between any two segments. The interaction is attractive if segments have opposite orientations (the analog of a vortex-antivortex pair) and repulsive otherwise. Consequently strings will try to minimize their energetic cost by aligning themselves in the most favorable configuration. Strings in the $O(2)$ scalar model are therefore in general self-seeking as we shall see below in more detail. The analog of a conductor insulator transition in 3D must therefore occur when string segments become free. Then the resulting strings lose their self-seeking character and should become Brownian random walks.

These qualitative expectations are realized for vortex string excitations in a complex $\lambda\phi^4$ -theory. To see this we have sampled the model's partition function and characterized its string vortex excitations at different temperatures. The results are shown in Figs. 2 and 3.

Figure 2 shows the string densities as a function of $\beta = 1/T$. Strings cease to be exponentially suppressed at $T = T_c$ (with T_c measured via field correlators). In that case the derivative of the string densities has a discontinuity. At this same point long strings appear in the system.

What is then the quantitative character of individual strings as the temperature is changed? To proceed we will assume and confirm *a posteriori* that the length distribution of strings assumes the same functional form as Eq. (5) but with temperature dependent coefficients. Thus we take the loop length distribution to be

$$n(l) = Al^{-\gamma}e^{-\beta\sigma_{\text{eff}}l}, \quad (7)$$

where l is in units of a and A , γ , and σ_{eff} will be computed numerically. The behavior of γ and σ_{eff} is shown in Fig. 3, together with the distance

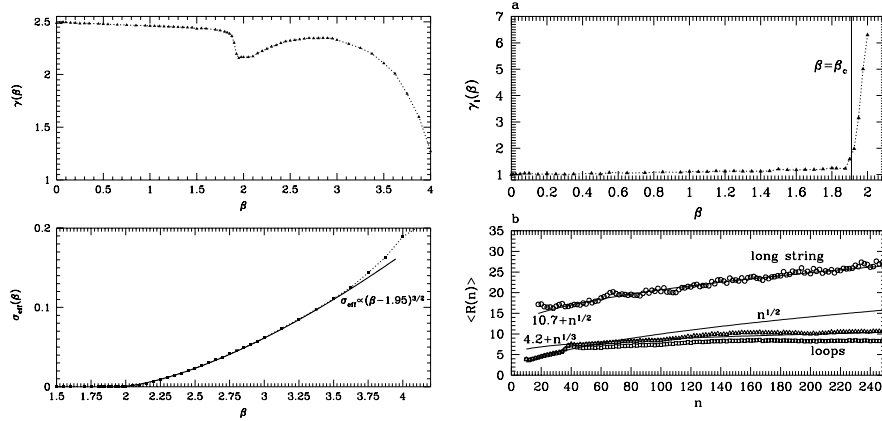


Figure 3. The critical behavior of the string tension (bottom left) and the exponent γ (top left). The long string formed at the transition is approximately Brownian. The exponent $\gamma \simeq 1$ (top right) for long strings is the correct Brownian value in a periodic domain. The Brownian character of long strings can be seen directly by plotting the distance *vs.* number of steps (bottom right), while short string loops remain self-seeking.

between two string segments as a function of length. Detailed study of these quantities indicates that strings below the transition are self-seeking, but that long strings, when produced at $T = T_c$, appear essentially as random walks.

Critical exponents measured directly in terms of strings can be related in turn to those of the field correlators [20]. This suggests that there is a dual model in terms of interacting strings that describes the system just as well as the original one based on the fields. At present it is of course still more cumbersome to compute most quantities based on such a model, in contrast to the canonical resolution of the problem in terms of scaling and the renormalization group.

Thus we have shown that for a $\lambda\phi^4$ $U(1)$ model the second-order transition is accompanied by the proliferation of strings. Results by Nguyen and Sudbø [17] and by Kajantie *et al.* [18] have confirmed the same picture for the type II Abelian Higgs model and the XY magnet, which are all in the same universality class.

The remaining question is whether the phase transition in terms of strings can somehow be seen to be more or less fundamental than that in terms of fields directly. This is certainly a difficult question. To answer it one would need to construct a statistical model in the same universality class, where one

of the degrees of freedom (the strings or the fields) would be absent.

4 Conclusions and Discussion

We discussed how topological excitations are linked to critical phenomena in several important models. In particular the numerical investigation of models with a proven track record in terms of critical exponents has recently revealed that the phase transition there is accompanied by critical behavior in terms of topological excitations.

In 3D for $O(N)$ models, the renormalization group description of the transition is sufficient to yield thermodynamic predictions but, as we have shown, it is a generic feature of the transition that it is accompanied by topological excitation percolation, at least in the best studied cases for $N = 1, 2$. It remains unclear if either of these two perspectives for the transition is more fundamental than the other. Depending on the experimental observable either can become more advantageous.

The advantage of topological defects and other collective excitations is that they provide us with a mesoscopic description of the relevant degrees of freedom involved in the critical phenomenon. In doing so we are allowed to disregard microscopic details of the underlying theory and obtain a simpler effective description.

Acknowledgments

It is a great pleasure to thank my collaborators N.D. Antunes and M. Hindmarsh for many useful discussions. Numerical work was carried out at the T-Division/CNLS Avalon Beowulf cluster. This work was supported by DOE.

References

- [1] L. Onsager, *Nuovo Cim. Suppl.* **6**, 249 (1949).
- [2] R.P. Feynman, in *Progress in Low Temperature Physics*, Vol. 1, Ed. C.J. Gorter (North-Holland, Amsterdam, 1955), p. 17.
- [3] T.W.B. Kibble, *J. Phys. A* **9**, 1387 (1976); *Phys. Rep.* **67**, 183 (1980).
- [4] A. Vilenkin and E.P.S. Shellard, *Cosmic Strings and Other Topological Defects* (Cambridge University Press, Cambridge, 1994).
- [5] M.E. Dodd *et al.*, *Phys. Rev. Lett.* **81**, 3703 (1998).
- [6] C. Bäuerle *et al.*, *Nature* **382**, 332 (1996).

- [7] V.M.H. Ruutu *et al.*, *Nature* **382**, 334 (1996); V.M.H. Ruutu *et al.*, *Phys. Rev. Lett.* **80**, 1465 (1998).
- [8] R. Carmi and E. Polturak, *Phys. Rev. B* **60** 7595 (1999).
- [9] I. Chuang, R. Durrer, N. Turok, and B. Yurke, *Science* **251**, 1336 (1991); M.J. Bowick, L. Chandar, E.A. Schiff, and A.M. Srivastava, *Science* **263**, 943 (1994).
- [10] R.Y. Chiao, *Opt. Commun.* **179**, 157 (2000).
- [11] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [12] J.M. Kosterlitz and D.J. Thouless, *J. Phys. C* **5**, L124 (1972), *ibid.* **6**, 1181 (1973); J.M. Kosterlitz, *J. Phys. C* **7**, 1046 (1974).
- [13] V.L. Berezinskii, *Sov. Phys. JETP* **34**, 610 (1972).
- [14] D.J. Bishop and J.D. Reppy, *Phys. Rev. Lett.* **40**, 1727 (1978).
- [15] N.D. Antunes, L.M.A. Bettencourt, and M. Hindmarsh, *Phys. Rev. Lett.* **80**, 908 (1998).
- [16] N.D. Antunes and L.M.A. Bettencourt, *Phys. Rev. Lett.* **81**, 3083 (1998).
- [17] A.K. Nguyen and A. Sudbø, *Phys. Rev. B* **58**, 2802 (1998).
- [18] K. Kajantie *et al.*, *Phys. Lett. B* **428**, 334 (1998).
- [19] E.J. Copeland *et al.*, *Physica A* **179**, 507 (1991).
- [20] A.M.J. Schakel, eprint: cond-mat/0008443 and references therein.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

TOPOLOGICAL SINGULARITIES, DEFECT FORMATION, AND PHASE TRANSITIONS IN QUANTUM FIELD THEORY

G. VITIELLO

*Dipartimento di Fisica, Università di Salerno, INFN and INFM,
84100 Salerno, Italia*

E-mail: vitiello@sa.infn.it

By resorting to results in quantum field theory some light is shed on the mechanism of the formation of topological defects in the process of phase transitions.

1 Symmetry Breaking Phase Transitions and Topological Singularities

I am pleased to dedicate this article to Hagen Kleinert on the occasion of his sixtieth birthday. I met Hagen for the first time about 25 years ago, at one of the Karpacz Winter Schools in Theoretical Physics in Poland. Since then a sincere friendship has been established. Besides the reciprocal sympathy, such a friendship finds perhaps its roots in the sense of humor with which both of us look at life and theoretical physics.

In this article I discuss, from the standpoint of quantum field theory (QFT), why topological defects, such as vortices, are formed in the process of symmetry breaking phase transitions and which effects boundaries and temperature have on defect formation [1,2]. Topological defects are described in QFT as extended objects created by the non-homogeneous condensation of Nambu-Goldstone (NG) modes carrying a topological singularity [3,4]. Therefore my attention is focused on the non-homogeneous boson condensation. Topological defects appear in many systems in a wide range of energy scales [5,6], from condensed matter physics to cosmology. I hope that from my discussion will emerge the unified view of collective phenomena in

many physically different systems which Hagen Kleinert has pursued during the many years of his dense research activity. The results have been obtained by the use of tools such as group theory and path integrals which have been and are Hagen's preferred fields of research.

I begin by recalling the mechanism of dynamical rearrangement of symmetry by which boson field translations are introduced in the theory. Then I discuss how macroscopic fields and currents are generated by boson condensation. By using such results, I will show that phase transitions in a gauge theory involve non-homogeneous boson condensation with a topological singularity and hence that topological defects appear in the process of symmetry breaking phase transitions. Finally, I will discuss finite temperature and finite volume effects.

To be specific I start by considering a complex scalar Heisenberg field $\phi_H(x)$ interacting with a gauge field $A_{H,\mu}(x)$ [7,8]. The Lagrangian density $\mathcal{L}[\phi_H(x), \phi_H^*(x), A_{H,\mu}(x)]$ is assumed to be invariant under global and local gauge transformations:

$$\phi_H(x) \rightarrow e^{i\theta} \phi_H(x), \quad A_{H,\mu}(x) \rightarrow A_{H,\mu}(x), \quad (1)$$

$$\phi_H(x) \rightarrow e^{ie_0\lambda(x)} \phi_H(x), \quad A_{H,\mu}(x) \rightarrow A_{H,\mu}(x) + \partial_\mu \lambda(x), \quad (2)$$

respectively. I assume $\lambda(x) \rightarrow 0$ for $|x_0| \rightarrow \infty$ and/or $|\mathbf{x}| \rightarrow \infty$. I use the Lorentz gauge $\partial^\mu A_{H,\mu}(x) = 0$ and set $\phi_H(x) = [\psi_H(x) + i\chi_H(x)]/\sqrt{2}$. I also assume that spontaneous breakdown of symmetry (SBS) can occur: $\langle 0|\phi_H(x)|0\rangle \equiv \tilde{v} \neq 0$, with constant $\tilde{v}$ and with $\rho_H(x) \equiv \psi_H(x) - \tilde{v}$. The generating functional is [9]

$$W[J, K] = \frac{1}{N} \int [dA_\mu][d\phi][d\phi^*][dB] \exp \left[i \int d^4x (\mathcal{L}(x) + B(x) \partial^\mu A_\mu(x) + K^* \phi + K \phi^* + J^\mu(x) A_\mu(x) + i\epsilon |\phi(x) - \tilde{v}|^2) \right], \quad (3)$$

where N is a convenient normalization. In the gauge constraint term, $B(x)$ is an auxiliary field. The ϵ -term specifies the condition of breakdown of symmetry under which we want to compute the path integral. It may represent the small external field triggering the symmetry breakdown. The limit $\epsilon \rightarrow 0$ must be performed at the end of the computations.

The LSZ maps (dynamical maps) between Heisenberg field and asymptotic

(also called physical) in- (or out-) fields are:

$$\phi_H(x) = : \exp \left\{ i \frac{Z_\chi^{\frac{1}{2}}}{\tilde{v}} \chi_{\text{in}}(x) \right\} \left[\tilde{v} + Z_\rho^{\frac{1}{2}} \rho_{\text{in}}(x) + F \right] :, \quad (4)$$

$$A_H^\mu(x) = Z_3^{\frac{1}{2}} U_{\text{in}}^\mu(x) + \frac{Z_\chi^{\frac{1}{2}}}{e_0 \tilde{v}} \partial^\mu b_{\text{in}}(x) + : F^\mu : . \quad (5)$$

The functionals $F \equiv F[\rho_{\text{in}}, U_{\text{in}}^\mu, \partial(\chi_{\text{in}} - b_{\text{in}})]$ and $F^\mu \equiv F^\mu[\rho_{\text{in}}, U_{\text{in}}^\mu, \partial(\chi_{\text{in}} - b_{\text{in}})]$ are determined within a particular model. The S -matrix is given by $S = : S[\rho_{\text{in}}, U_{\text{in}}^\mu, \partial(\chi_{\text{in}} - b_{\text{in}})] :$, and I use $A_H^{0\mu}(x) \equiv A_H^\mu(x) - e_0 \tilde{v} : \partial^\mu b_{\text{in}}(x) :$. The field χ_{in} denotes the NG mode, b_{in} the ghost mode, ρ_{in} the massive matter field, and U_{in}^μ the massive vector field. Their respective equations are

$$\partial^2 \chi_{\text{in}}(x) = 0, \quad \partial^2 b_{\text{in}}(x) = 0, \quad (\partial^2 + m_\rho^2) \rho_{\text{in}}(x) = 0, \quad (6)$$

$$(\partial^2 + m_V^2) U_{\text{in}}^\mu(x) = 0, \quad \partial^\mu U_{\text{in}}^\mu(x) = 0, \quad (7)$$

with $m_V^2 = Z_3(e_0 \tilde{v})^2 / Z_\chi$. One also has

$$B_H(x) = \frac{e_0 \tilde{v}}{Z_\chi^{\frac{1}{2}}} [b_{\text{in}}(x) - \chi_{\text{in}}(x)], \quad (8)$$

$$\partial^2 B_H(x) = 0, \quad -\partial^2 A_{H\mu}(x) = j_{H\mu}(x) - \partial_\mu B_H(x), \quad (9)$$

with $j_{H\mu}(x) = \delta \mathcal{L}(x) / \delta A_H^\mu(x)$. If one requires that the current $j_{H\mu}$ is the only source of the gauge field $A_{H\mu}$ in any observable process, one has to impose the condition ${}_p \langle b | \partial_\mu B_H(x) | a \rangle_p = 0$, i.e.

$$(-\partial^2) {}_p \langle b | A_{H\mu}^0(x) | a \rangle_p = {}_p \langle b | j_{H\mu}(x) | a \rangle_p. \quad (10)$$

Here $|a\rangle_p$ and $|b\rangle_p$ denote two generic physical states. The equations (10) are the classical Maxwell equations. The condition ${}_p \langle b | \partial_\mu B_H(x) | a \rangle_p = 0$ leads to the Gupta-Bleuler-like condition

$$[\chi_{\text{in}}^{(-)}(x) - b_{\text{in}}^{(-)}(x)] | a \rangle_p = 0, \quad (11)$$

where $\chi_{\text{in}}^{(-)}$ and $b_{\text{in}}^{(-)}$ are the positive-frequency parts of the corresponding fields χ_{in} and b_{in} which do not participate in any observable reaction. However, I stress that NG bosons do not disappear from the theory. As we will see, their condensation in the vacuum can have observable effects.

Finally, one finds that the $U(1)$ local and global gauge transformations [Eqs. (2) and (1)] of the Heisenberg fields are induced by the in-field transformations

$$\chi_{\text{in}}(x) \rightarrow \chi_{\text{in}}(x) + \frac{e_0 \tilde{v}}{Z_\chi^{\frac{1}{2}}} \lambda(x), \quad b_{\text{in}}(x) \rightarrow b_{\text{in}}(x) + \frac{e_0 \tilde{v}}{Z_\chi^{\frac{1}{2}}} \lambda(x), \quad (12)$$

$$\rho_{\text{in}}(x) \rightarrow \rho_{\text{in}}(x), \quad U_{\text{in}}^\mu(x) \rightarrow U_{\text{in}}^\mu(x), \quad (13)$$

and by

$$\chi_{\text{in}}(x) \rightarrow \chi_{\text{in}}(x) + \frac{\tilde{v}}{Z_\chi^{\frac{1}{2}}} \theta f(x), \quad (14)$$

$$b_{\text{in}}(x) \rightarrow b_{\text{in}}(x), \quad \rho_{\text{in}}(x) \rightarrow \rho_{\text{in}}(x), \quad U_{\text{in}}^\mu(x) \rightarrow U_{\text{in}}^\mu(x), \quad (15)$$

with $\partial^2 f(x) = 0$, respectively. Eq. (14) with $f(x) = 1$, which describes the homogeneous boson condensation, is not unitarily implementable. It induces transitions among unitarily inequivalent Fock spaces. The function $f(x)$ makes the generator of such a transformation well defined. The limit $f(x) \rightarrow 1$ is to be performed at the end of the computation. The fact that the Heisenberg field transformations are induced by Eqs. (12)-(15) is named the *dynamical rearrangement of symmetry* [3,4]. The in-field equations and the S -matrix are invariant under the above in-field transformations, and B_H is changed by an irrelevant c-number under (14) and (15) (in the limit $f \rightarrow 1$). It can be shown that the group of the transformations under which the in-field equations are invariant is the group contraction of the symmetry group for the Heisenberg field equations [10].

1.1 How Macroscopic Field and Current are Generated by Boson Condensation

Translations of bosonic physical fields (not necessarily massless) by space-time dependent functions, say $\alpha(x)$, satisfying the same field equation of the translated physical field, are called *boson transformations* [3]. Eqs. (12) and (14) (with $\partial^2 \lambda(x) = 0$ and $\partial^2 f(x) = 0$) are examples of boson transformations. Consider the transformation $\phi_H[x; \chi_{\text{in}}(x)] \rightarrow \phi'_H \equiv \phi_H[x; \chi_{\text{in}}(x) + \alpha(x)]$. The *boson transformation theorem* states that ϕ'_H is also a solution of the Heisenberg field equation for ϕ_H [3,4].

It can be shown that, in the presence of a gauge field, the boson transformation with regular (i.e. Fourier transformable) $\alpha(x)$ is equivalent to a gauge transformation. The only effect is the appearance of a phase factor in the

order parameter: $\tilde{v}(x) = e^{ica(x)}\tilde{v}$, with a constant c . Observable quantities are thus not affected in such a case. Note that in a theory which has only global gauge invariance non-singular boson transformations of the NG fields can produce non-trivial physical effects (like linear flow in superfluidity).

The proof of the boson transformation theorem relies on the fact that $\alpha(x)$ is a regular function. If $\alpha(x)$ carries some singularity (divergence or topological singularity) the singular region must be excluded when integrating on space and/or time. For example, if $\alpha(x)$ is singular on the axis of a cylinder (at $r = 0$), the singular line $r = 0$ must be excluded by a cylindrical surface of infinitesimal radius. The phase of the order parameter will be singular on that line. This means that SBS does not occur in that region (the core): there one has the "normal" state rather than the ordered one.

Since translation of a boson field describes boson condensation, we see that boson transformations describe *non-homogeneous* boson condensation. The boson theorem then shows that the same dynamics (the same field equations) may describe homogeneous and non-homogeneous phenomena. This leads us directly to the mechanism of formation of extended objects, which are in fact created by the non-homogeneous boson condensation [3,4,11]. Also, since different phases (described in QFT by unitarily inequivalent representations) are associated to different NG boson condensation densities, we see that, by inducing variations of NG boson condensation, boson transformations represent transitions through physically different phases of the system. This establishes a connection between the formation of extended objects and the process of phase transitions.

Boson transformations must be also compatible with the physical state condition (11). Under the boson transformation $\chi_{\text{in}}(x) \rightarrow \chi_{\text{in}}(x) + (\tilde{v}/Z_\chi^{\frac{1}{2}})f(x)$, B_H changes as

$$B_H(x) \rightarrow B_H(x) - \frac{e_0 \tilde{v}^2}{Z_\chi} f(x) . \quad (16)$$

Eq. (10) is violated upon imposing the Gupta-Bleuler-like condition. In order to restore it, one must compensate the shift in B_H by means of the transformation of U_{in} :

$$U_{\text{in}}^\mu(x) \rightarrow U_{\text{in}}^\mu(x) + Z_3^{-\frac{1}{2}} a^\mu(x) , \quad \partial_\mu a^\mu(x) = 0 , \quad (17)$$

with a convenient c-number function $a^\mu(x)$. The dynamical maps of the

various Heisenberg operators are not affected by (17) provided

$$(\partial^2 + m_V^2)a_\mu(x) = \frac{m_V^2}{e_0}\partial_\mu f(x) . \quad (18)$$

This is the Maxwell equation for the vector potential a_μ [9,12]. The classical ground-state current j_μ is

$$j_\mu(x) \equiv \langle 0 | j_{H\mu}(x) | 0 \rangle = m_V^2 \left[a_\mu(x) - \frac{1}{e_0} \partial_\mu f(x) \right] , \quad (19)$$

where $m_V^2 a_\mu(x)$ is the *Meissner current* and $m_V^2 \partial_\mu f(x)/e_0$ the *boson current*. Note that the classical current is given in terms of variations $\partial_\mu f$ of the non-homogeneous boson condensate density.

Summarizing, the macroscopic field and current are expressed in terms of the boson condensation function.

1.2 Why Do Topological Defects Exist Only in the Presence of Massless Bosons?

Let us now show that boson transformation functions carrying topological singularities are allowed only for massless bosons [3,4].

Consider the boson transformation $\chi_{\text{in}}(x) \rightarrow \chi_{\text{in}}(x) + f(x)$. Let $f(x)$ carry a topological singularity making it path-dependent such that it fails to satisfy the integrability condition of Schwarz:

$$G_{\mu\nu}^\dagger(x) \equiv [\partial_\mu, \partial_\nu] f(x) \neq 0 \quad (20)$$

for certain μ, ν, x . We have seen that $\partial_\mu f$ is related to observables and therefore it is single-valued, i.e. $[\partial_\rho, \partial_\nu] \partial_\mu f(x) = 0$. Recall that $f(x)$ is a solution of the χ_{in} equation, and suppose that χ_{in} is massive: $(\partial^2 + m^2)f(x) = 0$. It follows from the regularity of $\partial_\mu f(x)$ that

$$\partial_\mu f(x) = \frac{1}{\partial^2 + m^2} \partial^\lambda G_{\lambda\mu}^\dagger(x) , \quad (21)$$

which leads to $\partial^2 f(x) = 0$, which in turn implies $m = 0$. Thus (20) is compatible only with a *massless* equation for χ_{in} . This explains why topological defects are observed always in the presence of NG bosons, namely of ordering induced by NG condensate. The topological charge is defined as

$$N_T = \int_C dl^\mu \partial_\mu f = \int_S dS_\mu \epsilon^{\mu\nu\sigma} \partial_\nu \partial_\sigma f = \frac{1}{2} \int_S dS^{\mu\nu} G_{\mu\nu}^\dagger , \quad (22)$$

where C is a contour enclosing the singularity and S a surface with C as a boundary. The charge N_T does not depend on the path C provided it does not cross the singularity. The tensor $G^{\mu\nu}$ is $G^{\mu\nu}(x) \equiv -\frac{1}{2}\epsilon^{\mu\nu\lambda\rho}G_{\lambda\rho}^\dagger(x)$. It satisfies the continuity equation

$$\partial_\mu G^{\mu\nu}(x) = 0 \quad \Leftrightarrow \quad \partial_\mu G_{\lambda\rho}^\dagger + \partial_\rho G_{\mu\lambda}^\dagger + \partial_\lambda G_{\rho\mu}^\dagger = 0. \quad (23)$$

This completely characterizes the topological charge of the extended object [4].

On the other hand, the macroscopic ground-state effects do not occur for regular $f(x)$ ($G_{\mu\nu}^\dagger = 0$). In fact, from Eq. (18) we obtain $a_\mu(x) = \partial_\mu f(x)/e_0$ for regular f which implies a zero classical current ($j_\mu = 0$) and a zero classical field ($F_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$) (the Meissner and the boson current cancel each other).

In conclusion, the vacuum current appears only when $f(x)$ has topological singularities and this is allowed only for the condensation of massless bosons, i.e. when SBS occurs. We thus see that, in a gauge theory, the symmetry breaking phase transitions characterized by macroscopic ground-state effects, such as the vacuum current and field (as in superconductors), can occur only when there are non-zero gradients of topologically non trivial, non-homogeneous condensation of NG bosons. *Since these are also the conditions for the formation of topological defects, we also see why topological defects are observed in the process of phase transitions.*

Notice that the appearance of a space-time order parameter is no guarantee that persistent ground-state currents (and fields) will exist: if f is a regular function, the space-time dependence of $\tilde{v}$ can be gauged away by an appropriate gauge transformation.

Since the boson transformation with regular f does not affect observable quantities, the S matrix is actually given by

$$S = : S[\rho_{\text{in}}, U_{\text{in}}^\mu - \frac{1}{m_V} \partial(\chi_{\text{in}} - b_{\text{in}})] :. \quad (24)$$

This is in fact independent of the boson transformation with regular f :

$$S \rightarrow S' = : S[\rho_{\text{in}}, U_{\text{in}}^\mu - \frac{1}{m_V} \partial(\chi_{\text{in}} - b_{\text{in}}) + Z_3^{-\frac{1}{2}}(a^\mu - \frac{1}{e_0} \partial^\mu f)] : , \quad (25)$$

since $a_\mu(x) = \partial_\mu f(x)/e_0$ for regular f . However, $S' \neq S$ for singular f : S' includes the interaction of the quanta U_{in}^μ and ϕ_{in} with the classical field and current. Thus we see how quantum fluctuations may interact and have effects

on classically behaving macroscopic defects: *our picture includes interaction of quanta with macroscopic objects.*

2 Temperature and Volume Effects

The condition of breakdown of symmetry at finite temperature in the case of non-homogeneous condensation is [11]

$$\langle 0(\beta) | \phi(x) | 0(\beta) \rangle = \frac{1}{\sqrt{2}} \sigma(x, \beta) , \quad (26)$$

where $\beta \equiv 1/k_B T$. Here, $|0(\beta)\rangle$ denotes the temperature-dependent vacuum state in Thermo Field Dynamics [3,4,13]. Note that the statistical average of some operator A is given by $\langle A \rangle_0 = \langle 0(\beta) | \phi(x) | 0(\beta) \rangle$.

The fields $\phi \equiv \rho + \sigma(x, \beta)/\sqrt{2}$, χ and A_μ may undergo translation transformations by c-number functions, say σ , κ and α_μ , respectively, controlling the respective condensate structures. Usual gauge transformations are induced by using $\sigma = 0$, $\kappa = \alpha(x)$ and $\alpha_\mu(x) = \partial_\mu \alpha(x)$. The parameters m and λ will denote the Higgs field mass and self-coupling, respectively; $\tilde{v} \equiv \langle 0 | \phi | 0 \rangle$ at $T = 0$ is assumed to be non-zero, M is the gauge field mass, e the coupling between A_μ and ϕ .

The vortex solution can be studied by introducing cylindrical coordinates. The asymptotic gauge field configuration is imposed by considering the pure angular function as a gauge function at infinity $\kappa(x) = n\theta/e$, where $\kappa(x) = 0$ at $r = 0$:

$$\alpha_{as}^i = -\frac{n}{er} \mathbf{e}_\theta^i . \quad (27)$$

Here n is the winding number. One can show [11] that the masses are given by

$$m^2(x) = 2\lambda\sigma_0^2 f^2(x) , \quad M^2(x) = e^2(\sigma_0^2 f^2(x) + \langle : \bar{\rho}^2 : \rangle_0) , \quad (28)$$

with σ_0 the Higgs field condensate which goes to zero at the critical temperature T_C , and to $\tilde{v}$ at $T = 0$; $\bar{\rho}$ denotes the physical field. These masses act as potential terms in the field equations and only at spatial infinity ($r \rightarrow \infty$, $f(x) \rightarrow 1$) ordinary mass interpretation is recovered. In fact one has the asymptotic behavior

$$K(r) \simeq e^{-Mr} = e^{-r/R_0} , \quad f(r) \simeq 1 - f_0 e^{-mr} = 1 - f_0 e^{-r/r_0} . \quad (29)$$

$R_0 \equiv 1/M$ gives the size of the gauge field core and $r_0 \equiv 1/m$ the Higgs field core. $K(r)$ is related to the α function. Symmetry is restored above T_C .

In the case of the kink solution, the $\rho_\beta^{\text{in}}(x)$ condensation is induced by the boson transformation with $f_\beta(x) = \text{const.} \cdot e^{-\mu_0(\beta)x_1}$ playing the role of "form factor". The number of condensed bosons is proportional to $|f_\beta(x)|^2 = e^{-2\mu_0(\beta)(x_1-a)}$, which is maximal near the kink center $x_1 = a$ and decreases over a size $\xi_\beta = 2/\mu_0(\beta)$. The boson translation by f_β breaks the homogeneity of the order parameter $v(\beta)$ which is otherwise constant in space.

The mass $\mu_0 = (2\lambda)^{1/2}v(\beta)$ of the "constituent" fields ρ^{in} fixes the kink size $\xi_\beta \propto 2/\mu_0 = \sqrt{2}/\sqrt{\lambda}v(\beta)$ which thus increases as $T \rightarrow T_C$ (say $T \neq T_C$ but near T_C). In the $T \rightarrow 0$ limit the kink size is $\xi_0 \propto \sqrt{2}/\sqrt{\lambda}\tilde{v} < \sqrt{2}/\sqrt{\lambda}v(\beta) = \xi_\beta$, since [11] $v^2(\beta) = \tilde{v}^2 - 3\langle \rho^2 \rangle_0 < \tilde{v}^2$.

For T different from zero, the thermal Bose condensate $\langle \rho^2 \rangle_0$ develops which acts as a potential term for the kink field. Such a potential term controls the "size" and the number of the kinks. Only in the limit $v(x, \beta) \rightarrow \text{const.}$ the $\rho_\beta^{\text{in}}(x)$ field may be considered as a free field, e.g. far from the kink core.

Finally, let me consider finite-volume effects. Suppose we have homogeneous boson condensation. For large but finite volume, one expects that the condition of symmetry breakdown is satisfied "inside the bulk" far from the boundaries. However, "near" the boundaries, one might expect "distortions" in the order parameter: $\tilde{v} = \tilde{v}(x)$ (or even $\tilde{v} \rightarrow 0$): "near" the system boundaries we may have a non-homogeneous order parameter. Non-homogeneities in the boson condensation will "smooth out" in the $V \rightarrow \infty$ limit. Let me put $V \equiv \eta^{-3}$. Then, the NG mode acquires an effective gap (mass) of the order of η ($\omega_\eta^2 = \eta^2$). The Goldstone theorem (existence of gapless modes) is recovered in the infinite volume limit ($\eta \rightarrow 0$). The effect of the boundaries ($\eta \neq 0$) is to give an "effective mass" $m_{\text{eff}} \equiv \omega_\eta$ to the NG bosons. These will then propagate over a range of the order of $\xi \equiv 1/\eta$, which is the system linear size.

Note that only if $\epsilon \neq 0$ (cf. Eq. (3)), the order parameter can be kept different from zero, i.e. if $\eta \neq 0$, then ϵ must be non-zero in order to have $\tilde{v} \neq 0$. In such a case the symmetry breakdown is maintained thanks to the non-zero ϵ which acts as an external field acting as a pump providing energy which is required in order to condense modes of non-zero lowest energy ω_η . Boundary effects are thus in competition with the breakdown of symmetry [1,2]. They may preclude its occurrence or, if symmetry is already broken, they

may reduce the order parameter to zero.

Temperature may have similar effects on the order parameter (at T_C symmetry may be restored). Since the order parameter goes to zero in the absence of the external supply of energy, when NG modes acquire non-zero effective mass, one may represent the effect of thermalization in terms of finite-volume effects and get $\eta \propto \sqrt{|T - T_C|/T_C}$. In this way, temperature changes near T_C may be discussed as variations of the condensate domain size ξ . For example, in the presence of an external driving field ($\epsilon \neq 0$), for $T > T_C$ (but near to T_C) one may have the formation of ordered domains of size $\xi \propto (\sqrt{|T - T_C|/T_C})^{-1}$, before the transition to the fully ordered phase is achieved (as $T \rightarrow T_C$). As far as $\eta \neq 0$, the ordered domains are unstable; they disappear as the external field coupling $\epsilon \rightarrow 0$.

Of course, if ordered domains are still present at $T < T_C$, they also disappear as $\epsilon \rightarrow 0$. The possibility to maintain such ordered domains below T_C depends on the speed at which T is lowered, compared to the speed at which the system is able to become homogeneously ordered. Notice that the speed of $T \rightarrow T_C$ is related to the speed of $\eta \rightarrow 0$.

The order parameters $v(x, \beta)$ and $\sigma(x, \beta)$ provide a mapping between the variation domains of (x, β) and the *space of the unitarily inequivalent representations* of the canonical commutation relations, i.e. the set of Hilbert spaces where the operator field ϕ is realized for different values of the order parameter. As is well known, one has the mapping π of S^1 in the vortex case, surrounding the $r = 0$ singularity, to the group manifold of $U(1)$ which is topologically characterized by the winding number $n \in Z \in \pi_1(S^1)$. It is such a singularity which is carried by the boson condensation function of the NG modes. In the monopole case [11], the mapping π is the one of the sphere S^2 , surrounding the singularity $r = 0$, to $SO(3)/SO(2)$ group manifold, with homotopy classes of $\pi_2(S^2) = Z$. Again, the singularity is carried by the NG boson condensation function. The same situation occurs in the sphaleron case [11], provided one replaces $SO(3)$ and $SO(2)$ with $SU(2)$ and $U(1)$, respectively.

To conclude let me state that phase transitions imply "moving" over unitarily inequivalent representations, and this implies in general a non-trivial homotopy mapping between the (x, β) variability domain and the group manifold. The invariance of the theory under the involved symmetry group then leads to NG boson condensation with topological singularities. The conditions for the formation of topological defects are then satisfied. *This explains why we observe topological defect formation in the process of phase transitions.*

In the case of the kink there are no NG modes. Nevertheless the topologically non-trivial kink solution requires the boson condensation function to carry divergence singularity at spatial infinity.

Acknowledgments

This work has been partially supported by INFN, INFM, MURST, and the ESF Network on Topological Defect Formation in Phase Transitions.

References

- [1] G. Vitiello, in *Topological Defects and the Non-Equilibrium Dynamics of Symmetry Breaking Phase Transitions*, Eds. Y.M. Bunkov and H. Godfrin, NATO Science Series C 549 (Kluwer Acad. Publ., Dordrecht, 2000), p. 171.
- [2] E. Alfinito, O. Romei, and G. Vitiello, *Why Do Topological Defects Appear in the Process of Symmetry Breaking Phase Transitions?*, preprint.
- [3] H. Umezawa, *Advanced Field Theory: Micro, Macro, and Thermal Physics* (A.I.P., New York, 1993).
- [4] H. Umezawa, H. Matsumoto, and M. Tachiki, *Thermo Field Dynamics and Condensed States* (North-Holland, Amsterdam, 1982).
- [5] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [6] H. Kleinert, in *Proceedings of a NATO Advanced Study Institute on Formation and Interactions of Topological Defects*, Eds. A.C. Davis and R. Brandenburger (Plenum Press, New York, 1995).
- [7] P. Higgs, *Phys. Rev.* **45**, 1156 (1960).
- [8] T.W.B. Kibble, *Phys. Rev.* **155**, 1554 (1967).
- [9] H. Matsumoto, N.J. Papastamatiou, H. Umezawa, and G. Vitiello, *Nucl. Phys. B* **97**, 61 (1975).
- [10] C. De Concini and G. Vitiello, *Nucl. Phys. B* **116**, 141 (1976).
- [11] R. Manka and G. Vitiello, *Ann. Phys. (N.Y.)* **199**, 61 (1990).
- [12] H. Matsumoto, N.J. Papastamatiou, and H. Umezawa, *Nucl. Phys. B* **97**, 90 (1975).
- [13] Y. Takahashi and H. Umezawa, *Collective Phenomena* **2**, 55 (1975), reprinted in *Int. J. Mod. Phys. B* **10**, 1755 (1996).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

CONFINEMENT IN THE ENSEMBLES OF MONOPOLES

D. ANTONOV

*INFN-Sezione di Pisa, Università degli studi di Pisa, Dipartimento di Fisica,
Via Buonarroti, 2 - Ed. B - I-56127 Pisa, Italy*

and

*Institute of Theoretical and Experimental Physics, B. Cheremushkinskaya 25,
RU-117 218 Moscow, Russia*

E-mail: antonov@df.unipi.it

String representations of the Wilson loop in three-dimensional gases of $SU(2)$ and $SU(3)$ Abelian-projected monopoles are discussed. It is demonstrated that the summation over world sheets bounded by the contour of the Wilson loop is realized by summing over branches of a certain effective multi-valued potential of monopole densities. Finally, by virtue of the so-constructed representation of the Wilson loop, this quantity is evaluated in the $SU(2)$ -inspired case within the approximation of a dilute monopole gas, which makes confinement in the model under study manifest.

1 Introduction

On the way of constructing the string representation of QCD by means of the method of Abelian projections [1], the main results have been obtained under the assumption of the monopole condensation. Such condensation can be described by demanding that fluctuating monopole trajectories, forming the grand canonical ensemble, possess several natural properties. These properties, which can be elegantly formulated in the path-integral language [2,3], are the presence of the kinetic and mass terms of a trajectory, as well as the short-range interaction of the trajectories.

Another way to model the monopole condensation is to consider the grand canonical ensemble of monopoles as a Coulomb gas [2,4]. The respective $SU(2)$ -inspired theory then turns out to be compact QED, i.e. electrodynamics with monopoles. The novel type of gauge invariance appearing in this

theory (the so-called monopole gauge invariance) and its condensed matter analogues have been discovered by Professor Hagen Kleinert in Ref. [5]. The consistent local quantum field theory of electrically charged particles and monopoles has for the first time been constructed in Refs. [5,6] and discussed in details in Ref. [7]. A very important result of these investigations, which has then been used many times in the literature on the dual models of confinement, is that the Wilson loop in this theory remains invariant under the duality transformation.

In what follows, we shall just consider the grand canonical ensembles of $SU(2)$ and $SU(3)$ Abelian-projected monopoles in 2+1 dimensions and string representations of the Wilson loop in the respective disorder field theories. In Section 2, we shall consider the simplest $SU(2)$ -inspired case (i.e. compact QED) and then, in Section 3, we will extend this analysis to the $SU(3)$ -inspired theory. The main results of our study will be summarized in the conclusions.

2 String Representation of the Wilson Loop in Compact QED

The action of the Coulomb gas of monopoles in 3D compact QED has the form

$$S_{\text{mon}} = g_m^2 \sum_{a < b} q_a q_b (\Delta^{-1})(\mathbf{z}_a, \mathbf{z}_b) + S_0 \sum_a q_a^2. \quad (1)$$

Here, Δ is the 3D Laplace operator, and S_0 is the action of a single monopole, $S_0 = \text{const.}/e^2$. We have also adopted the standard Dirac notations, where $eg_m = 2\pi n$, restricting ourselves to the monopoles of the minimal charge, i.e. setting $n = 1$. Then, the partition function of the grand canonical ensemble of monopoles corresponding to the action (1) reads

$$\begin{aligned} \mathcal{Z}_{\text{mon}} = 1 + \sum_{N=1}^{\infty} \sum_{q_a=\pm 1} \frac{\zeta^N}{N!} \prod_{i=1}^N \int d^3 z_i \\ \times \exp \left[-\frac{g_m^2}{8\pi} \int d^3 x d^3 y \rho_{\text{gas}}(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{y}|} \rho_{\text{gas}}(\mathbf{y}) \right], \end{aligned} \quad (2)$$

where

$$\rho_{\text{gas}}(\mathbf{x}) = \sum_{a=1}^N q_a \delta(\mathbf{x} - \mathbf{z}_a)$$

is the density of the monopole gas. Here, a single monopole weight $\zeta \propto \exp(-S_0)$ has the dimension of $(\text{mass})^3$. It is usually referred to as the fugacity. Notice also that we have restricted ourselves to the values $q_a = \pm 1$, since monopoles with $|q| > 1$ turn out to be unstable and tend to dissociate into those with $|q| = 1$. That is because the energy of a single monopole is a quadratic function of its flux. Thus it is energetically more favourable for the vacuum to support a configuration of two monopoles of the unit magnetic charge than one monopole of the double charge.

Next, Coulomb interaction can be made local, albeit a nonlinear one, by introducing an auxiliary scalar field:

$$\mathcal{Z}_{\text{mon}} = \int D\chi \exp \left\{ - \int d^3x \left[\frac{1}{2} (\nabla\chi)^2 - 2\zeta \cos(g_m\chi) \right] \right\}. \quad (3)$$

The magnetic mass $m = g_m \sqrt{2\zeta}$ of the dual boson χ , following from the quadratic term in the expansion of the cosine on the r.h.s. of Eq. (3), is due to the Debye screening of this boson in the monopole gas.

Let us now cast the partition function (2) into the form of an integral over monopole densities. This can be done by introducing into Eq. (2) a unity of the form

$$1 = \int D\rho \delta(\rho(\mathbf{x}) - \rho_{\text{gas}}(\mathbf{x})) = \int D\rho D\lambda \exp \left\{ i \left[\sum_{a=1}^N q_a \lambda(\mathbf{z}_a) - \int d^3x \lambda \rho \right] \right\}. \quad (4)$$

Then, performing the summation over the monopole ensemble in the same way as it has been done in a derivation of the representation (3), we get

$$\begin{aligned} \mathcal{Z}_{\text{mon}} = \int D\rho D\lambda \exp \left\{ - \frac{g_m^2}{8\pi} \int d^3x d^3y \rho(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{y}|} \rho(\mathbf{y}) + \right. \\ \left. + \int d^3x (2\zeta \cos \lambda - i\lambda \rho) \right\}. \end{aligned} \quad (5)$$

Finally, integrating over the Lagrange multiplier λ by resolving the corresponding saddle-point equation,

$$\sin \lambda = - \frac{i\rho}{2\zeta}, \quad (6)$$

we arrive at the following expression for the partition function

$$\mathcal{Z}_{\text{mon}} = \int D\rho \exp \left\{ - \left[\frac{g_m^2}{8\pi} \int d^3x d^3y \rho(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{y}|} \rho(\mathbf{y}) + V[\rho] \right] \right\}. \quad (7)$$

Here,

$$V[\rho] = \int d^3x \left[\rho \operatorname{arcsinh} \left(\frac{\rho}{2\zeta} \right) - 2\zeta \sqrt{1 + \left(\frac{\rho}{2\zeta} \right)^2} \right] \quad (8)$$

is the effective multi-valued potential of monopole densities. Due to Eq. (4), the obtained representation (7) is natural to be called as a representation for the partition function in terms of the monopole densities.

We are now in the position to discuss the string representation of the Wilson loop in 3D compact QED. Such a loop is nothing else, but an averaged phase factor of an electrically charged test particle propagating along a closed trajectory C . Since the Wilson loop depends only on this trajectory, the desired string representation should be understood as some mechanism providing the independence of the loop from a certain string world sheet Σ bounded by C . By virtue of the Stokes theorem, the Wilson loop can be rewritten in the following form

$$\begin{aligned} \langle W(C) \rangle &= \left\langle \exp \left(\frac{i}{2} \int_{\Sigma} d\sigma_{\mu\nu} \mathcal{F}_{\mu\nu} \right) \right\rangle_{A_\mu, \rho} \\ &= \langle W(C) \rangle_{A_\mu} \left\langle \exp \left(\frac{i}{2} \int d^3x \rho(\mathbf{x}) \eta(\mathbf{x}) \right) \right\rangle_{\rho}. \end{aligned} \quad (9)$$

Here, $\langle W(C) \rangle_{A_\mu}$ stands for the standard free-photon contribution, whereas the average over monopole densities is defined by the partition function (7). We have also defined by $\mathcal{F}_{\mu\nu} \equiv \mathcal{F}_{\mu\nu}[\rho]$ the full electromagnetic field strength tensor $F_{\mu\nu} + F_{\mu\nu}^M$, which includes the monopole part $F_{\mu\nu}^M[\rho]$ obeying the modified Bianchi identities

$$\frac{1}{2} \varepsilon_{\mu\nu\lambda} \partial_\mu F_{\nu\lambda}^M = 2\pi\rho.$$

Also, in Eq. (9),

$$\eta(\mathbf{x}) = \partial_\mu^x \int_{\Sigma} d\sigma_\mu(\mathbf{y}) \frac{1}{|\mathbf{x} - \mathbf{y}|}$$

stands for the solid angle under which the surface Σ shows up to an observer located at point $\mathbf{x}$.

Equation (9) seems to contain some discrepancy, since its l.h.s. depends only on the contour C , whereas the r.h.s. depends on an arbitrary surface

Σ bounded by C . However, this actually occurs to be not a discrepancy, but a key point in the construction of the desired string representation. The resolution of the apparent paradox lies in the observation that the surface independence is realized by summing over all complex-valued branches of the monopole potential (8) at every space point $\mathbf{x}$.

It is worth noting that the so-obtained string representation (9) has been derived for a first time in another, more indirect, way in Ref. [8]. It is therefore instructive to establish a correspondence between our approach and the one of that paper.

The main idea of Ref. [8] was to calculate the Wilson loop starting with the direct definition of this average in the sense of the partition function (2) of the monopole gas. The respective expression has the form

$$\begin{aligned} \langle W(C) \rangle_{\text{mon}} &= 1 + \sum_{N=1}^{\infty} \sum_{q_a=\pm 1} \frac{\zeta^N}{N!} \prod_{i=1}^N \int d^3 z_i \exp \left[-\frac{g_m^2}{8\pi} \int d^3 x d^3 y \right. \\ &\quad \left. \times \rho_{\text{gas}}(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{y}|} \rho_{\text{gas}}(\mathbf{y}) + \frac{i}{2} \int d^3 x \rho_{\text{gas}}(\mathbf{x}) \eta(\mathbf{x}) \right] \\ &= \int D\varphi \exp \left\{ - \int d^3 x \left[\frac{e^2}{8\pi^2} \left(\partial_\mu \varphi - \frac{1}{2} \partial_\mu \eta \right)^2 - 2\zeta \cos \varphi \right] \right\}, \quad (10) \end{aligned}$$

where $\varphi \equiv g_m \chi + \eta/2$.

Next, one can prove the following equality

$$\begin{aligned} \exp \left[-\frac{e^2}{8\pi} \oint_C dx_\mu \oint_C dy_\mu \frac{1}{|\mathbf{x} - \mathbf{y}|} - \frac{e^2}{8\pi^2} \int d^3 x \left(\partial_\mu \varphi - \frac{1}{2} \partial_\mu \eta \right)^2 \right] \\ = \int Dh_{\mu\nu} \exp \left[- \int d^3 x \left(i\varphi \varepsilon_{\mu\nu\lambda} \partial_\mu h_{\nu\lambda} + g_m^2 h_{\mu\nu}^2 - 2\pi i h_{\mu\nu} \Sigma_{\mu\nu} \right) \right], \quad (11) \end{aligned}$$

which makes it possible to represent the contribution of the kinetic term on the r.h.s. of Eq. (10) and the free-photon contribution to the Wilson loop as an integral over the Kalb-Ramond field $h_{\mu\nu}$. The only nontrivial point necessary to prove the equality (11) is an expression for the derivative of the solid angle [2],

$$\partial_\lambda \eta(\mathbf{x}) = \varepsilon_{\lambda\mu\nu} \partial_\mu^x \oint_C dy_\nu \frac{1}{|\mathbf{x} - \mathbf{y}|} - 4\pi \int_\Sigma d\sigma_\lambda(\mathbf{y}) \delta(\mathbf{x} - \mathbf{y}).$$

Making use of this result and carrying out the Gaussian integral over the field

$h_{\mu\nu}$, one can demonstrate that both sides of Eq. (11) are equal to

$$\exp \left\{ -\frac{e^2}{2} \left[\frac{1}{4\pi^2} \int d^3x (\partial_\mu \varphi)^2 + \frac{1}{\pi} \int_\Sigma d\sigma_\mu \partial_\mu \varphi + \int_\Sigma d\sigma_\mu(\mathbf{x}) \int_\Sigma d\sigma_\mu(\mathbf{y}) \delta(\mathbf{x} - \mathbf{y}) \right] \right\},$$

thus proving the validity of this equation.

Substituting now Eq. (11) into Eq. (10), it is straightforward to carry out the integral over the field φ . Since this field has no more kinetic term, such an integration can be performed in the saddle-point approximation. The respective saddle-point equation has the same form as Eq. (6) with the replacement $\rho \rightarrow \varepsilon_{\mu\nu\lambda} \partial_\mu h_{\nu\lambda}$. As a result, we obtain the following expression for the full Wilson loop

$$\langle W(C) \rangle = \langle W(C) \rangle_{A_\mu} \langle W(C) \rangle_{\text{mon}} = \int Dh_{\mu\nu} \times \exp \left\{ - \int d^3x (g_m^2 h_{\mu\nu}^2 + V[\varepsilon_{\mu\nu\lambda} \partial_\mu h_{\nu\lambda}]) + 2\pi i \int_\Sigma d\sigma_{\mu\nu} h_{\mu\nu} \right\}, \quad (12)$$

where the world-sheet independence of the r.h.s. is again provided by summing over branches of the multi-valued action, which is now the action of the Kalb-Ramond field $h_{\mu\nu}$.

Comparing Eqs. (9) and (12), we see that the Kalb-Ramond field is indeed related to the monopole density via the equation $\varepsilon_{\mu\nu\lambda} \partial_\mu h_{\nu\lambda} = \rho$. Thus, we see that the same Legendre transformation which made out of $\rho_{\text{gas}}(\mathbf{x})$ the dynamical field $\rho(\mathbf{x})$ transforms the field $\mathcal{F}_{\mu\nu}[\rho_{\text{gas}}]/(4\pi)$ to the dynamical Kalb-Ramond field $h_{\mu\nu}$. In the formal language, such a decomposition of the Kalb-Ramond field is just the essence of the Hodge decomposition theorem.

Let us now consider the case of a very dilute monopole gas $|\rho| \ll \zeta$, and restrict ourselves to the real branch of the effective potential. This yields the following expression for the Wilson loop

$$\langle W(C) \rangle_{\text{dil. gas}} = \int Dh_{\mu\nu} \exp \left\{ - \int d^3x \left[\frac{1}{6\zeta} H_{\mu\nu\lambda}^2 + g_m^2 h_{\mu\nu}^2 - 2\pi i h_{\mu\nu} \Sigma_{\mu\nu} \right] \right\}, \quad (13)$$

where $H_{\mu\nu\lambda} = \partial_\mu h_{\nu\lambda} + \partial_\lambda h_{\mu\nu} + \partial_\nu h_{\lambda\mu}$ is the Kalb-Ramond field strength tensor. Notice that the mass of the Kalb-Ramond field stemming from this

equation is equal to the Debye mass m of the field χ from Eq. (3).

Clearly as we now see by restricting ourselves to the real branch of the potential we have violated the surface independence of the r.h.s. of Eq. (13). This independence can be restored by replacing Σ by the surface of the minimal area, $\Sigma_{\min} \equiv \Sigma_{\min}[C]$. After that, the quantity

$$S_{\text{str}} = -\ln \langle W(C) \rangle_{\text{dil. gas}} \Big|_{\Sigma \rightarrow \Sigma_{\min}} \quad (14)$$

can be considered as a string effective action of 3D compact QED in the dilute monopole gas approximation.

Integrating over the Kalb-Ramond field, we get (apart from the boundary term)

$$S_{\text{str}} = \frac{\pi\zeta}{2} \int_{\Sigma_{\min}} d\sigma_{\mu\nu}(\mathbf{x}) \int_{\Sigma_{\min}} d\sigma_{\mu\nu}(\mathbf{y}) \frac{e^{-m|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}.$$

This non-local action can be further expanded in powers of the derivatives with respect to the world-sheet coordinates $\xi = (\xi^1, \xi^2)$, which is equivalent to the $1/m$ -expansion. Then, as the first two non-trivial terms, we get the Nambu-Goto and the Polyakov-Kleinert terms [9,10]:

$$S_{\text{str}} \simeq \sigma \int d^2\xi \sqrt{g} + \frac{1}{\alpha} \int d^2\xi \sqrt{g} g^{ab} (\partial_a t_{\mu\nu}) (\partial_b t_{\mu\nu}). \quad (15)$$

Here, $g^{ab} = (\partial^a x_\mu)(\partial^b x_\mu)$ is the induced metric tensor corresponding to the world sheet $\Sigma(C)$ parameterized by the vector $x_\mu(\xi)$, g is the determinant of this tensor, and $t_{\mu\nu} = \varepsilon^{ab}(\partial_a x_\mu)(\partial_b x_\nu)/\sqrt{g}$ is the extrinsic curvature tensor corresponding to the same world sheet. The Polyakov-Kleinert term describes the stiffness of the string and makes the obtained local string action much more suitable for modelling the QCD string than the pure Nambu-Goto action [11,12]. The string tension and the inverse coupling constant of the Polyakov-Kleinert term read $\sigma = \pi^2 \sqrt{2\zeta}/g_m$ and $1/\alpha = -\pi^2/(8\sqrt{2\zeta}g_m^3)$, respectively. Both of them are non-analytic in the electric coupling constant, which manifests the non-perturbative nature of confinement in the model under study. Notice also that the negative sign of α is important for providing the stability of the string [10,12].

3 String Representation of the Wilson Loop in the Gas of $SU(3)$ Abelian-Projected Monopoles

The partition function describing the grand canonical ensemble of $SU(3)$ Abelian-projected monopoles has the form [13]

$$\mathcal{Z}_{\text{mon}} = 1 + \sum_{N=1}^{\infty} \frac{\zeta^N}{N!} \left(\prod_{a=1}^N \int d^3 z_a \sum_{\alpha_a=\pm 1, \pm 2, \pm 3} \right) \exp \left[-\frac{g_m^2}{4\pi} \sum_{a<b} \frac{\vec{q}_{\alpha_a} \vec{q}_{\alpha_b}}{|\mathbf{z}_a - \mathbf{z}_b|} \right]. \quad (16)$$

Here, the magnetic coupling constant g_m is related to the QCD coupling constant g according to the equation $gg_m = 4\pi$, $\zeta \propto \exp(-\text{const.}/g^2)$ is again the fugacity of a single monopole, and $\vec{q}_{\alpha_a}$'s are the nonzero weights of the zero triality adjoint representation of $SU(3)$. These weights are defined as $\vec{q}_1 = (1/2, \sqrt{3}/2)$, $\vec{q}_2 = (-1, 0)$, $\vec{q}_3 = (1/2, -\sqrt{3}/2)$, $\vec{q}_{-\alpha} = -\vec{q}_{\alpha}$. Notice that for every $\alpha = \pm 1, \pm 2, \pm 3$ and $\vec{\lambda} = (\lambda_3, \lambda_8)$ [where in the Gell-Mann basis $\lambda_3 = \text{diag}(1, -1, 0)$, $\lambda_8 = \text{diag}(1/\sqrt{3}, 1/\sqrt{3}, -2/\sqrt{3})$], the following relation holds: $\vec{q}_{\alpha} \vec{\lambda} = \hat{n}$. Here, $\hat{n}$ is some traceless diagonal matrix with the elements $0, \pm 1$. This matrix can thus be written as $\hat{n} = w\lambda_3 w^{-1}$, where w is any of the six elements of the permutation group S_3 , which is the reason for the same amount of vectors $\vec{q}_{\pm\alpha}$'s.

Equation (16) can be represented as

$$\mathcal{Z}_{\text{mon}} = \int D\vec{\chi} \exp \left\{ - \int d^3 x \left[\frac{1}{2} (\nabla \vec{\chi})^2 - 2\zeta \sum_{\alpha=1}^3 \cos(g_m \vec{q}_{\alpha} \vec{\chi}) \right] \right\}, \quad (17)$$

and we see that the property $\vec{q}_{-\alpha} = -\vec{q}_{\alpha}$ yields in this case the cosine in the action rigorously. Let us now define the Wilson loop as

$$\langle W(C) \rangle \equiv \frac{1}{3} \left\langle \text{tr} P \exp \left(\frac{i}{2} \oint_C dx_{\mu} \vec{A}_{\mu} \vec{\lambda} \right) \right\rangle \quad (18)$$

with $\vec{A}_{\mu} \equiv (A_{\mu}^3, A_{\mu}^8)$ and consider the monopole contribution to this quantity in the theory (17). Then, it turns out that we have the following string representation [14] [compare with Eqs. (7)-(9)]:

$$\langle W(C) \rangle_{\text{mon}} = \frac{1}{3\mathcal{Z}_{\text{mon}}} \sum_{\alpha=1}^3 \int D\vec{\rho} \exp \left\{ - \left[\frac{g_m^2}{8\pi} \int d^3 x d^3 y \vec{\rho}(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{y}|} \vec{\rho}(\mathbf{y}) \right. \right.$$

$$\left. + \sum_{\beta=1}^3 V[\rho_\beta] - i \int d^3x \vec{\rho} \vec{\eta}_\alpha \eta \right\}. \quad (19)$$

Here, we have introduced the vectors

$$\vec{\eta}_1 = \left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \quad \vec{\eta}_2 = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \quad \vec{\eta}_3 = \left(0, -\frac{1}{\sqrt{3}}\right),$$

which are just the weights of the representation **3** of $SU(3)$. These vectors thus determine the charges of quarks of three colours with respect to the diagonal gluons $\vec{A}_\mu$. Besides that we have denoted

$$\rho_1 = \frac{1}{\sqrt{3}} \left(\frac{1}{\sqrt{3}} \rho^1 + \rho^2 \right), \quad \rho_2 = -\frac{2}{3} \rho^1, \quad \rho_3 = \frac{1}{\sqrt{3}} \left(\frac{1}{\sqrt{3}} \rho^1 - \rho^2 \right),$$

where $\rho^{1,2}$ are the components of the monopole density, $\vec{\rho} \equiv (\rho^1, \rho^2)$.

One can see that it is again the sum over branches of the multi-valued monopole potential (8) which provides the surface independence of the r.h.s. of Eq. (19). Besides that, we see that the string representation of the Wilson loop in the $SU(3)$ monopole gas differs from that of the $SU(2)$ -case. This disproves the conjecture put forward in Ref. [8], according to which the $SU(2)$ -inspired action should be universal for any gauge group $SU(N)$. Finally, it is straightforward to integrate over monopole densities in the dilute gas approximation, which yields again the string effective action of the form (15).

4 Conclusions

In the present paper, we have derived the string representations of the Wilson loop in 2+1-dimensional gases of $SU(2)$ and $SU(3)$ Abelian-projected monopoles. Contrary to photons, monopoles do not interact with the contour of the Wilson loop, but rather with the string world sheet spanned by this contour. Therefore in both cases, the resulting string representation was understood as a certain mechanism providing the independence of the Wilson loop of the choice of such a world sheet. It has been demonstrated that such a mechanism can be based on the summation over branches of a certain multi-valued potential of monopole densities. Substituting for such densities the Kalb-Ramond field unambiguously related to them via the modified Bianchi identities, one arrives in the $SU(2)$ -inspired case at the theory of confining strings, whereas in the $SU(3)$ -inspired case one gets a different action.

Although such a reformulation of the functional integral in terms of the Kalb-Ramond field allows one to account automatically also for the non-compact (photon) part of the gauge fields, it is physically less transparent than the proposed representation in terms of the monopole densities.

In the approximation of a very dilute monopole gas, the real branch of the monopole potential becomes a quadratic functional, and one can explicitly integrate monopoles out. This produces the non-local string effective action, whose gradient expansion yields the Nambu-Goto and the Polyakov-Kleinert terms as the leading ones. Those make confinement in the models under study manifest and ensure the stability of strings.

However, it remains unclear within the monopole-gas models how to derive the monopole fugacity itself from the QCD Lagrangian. Some attempts in this direction have been done in Ref. [15], but the explicit answer is unfortunately not yet found. The still pending question is: What is the proportionality coefficient between the string tension in QCD and Λ_{QCD}^2 ? Since the answer to this question is very important for the understanding of the connection between the perturbative and non-perturbative phenomena in QCD, it requires further investigations.

References

- [1] G. 't Hooft, *Nucl. Phys. B* **190**, 455 (1981).
- [2] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* (World Scientific, Singapore, 1989).
- [3] M. Kiometzis, H. Kleinert, and A.M.J. Schakel, *Fortschr. Phys.* **43**, 697 (1995).
- [4] A.M. Polyakov, *Gauge Fields and Strings* (Harwood Academic Publishers, Chur, 1987).
- [5] H. Kleinert, *Int. Journ. Mod. Phys. A* **7**, 4693 (1992).
- [6] H. Kleinert, *Phys. Lett. B* **246**, 127 (1990); *ibid.* **293**, 168 (1992).
- [7] H. Kleinert, in *Proceedings of a NATO Advanced Study Institute on Formation and Interactions of Topological Defects at the University of Cambridge, England* (Plenum Press, New York, 1995), eprint: cond-mat/9503030.
- [8] A.M. Polyakov, *Nucl. Phys. B* **486**, 23 (1997), eprint: hep-th/9607049.
- [9] A.M. Polyakov, *Nucl. Phys. B* **268**, 406 (1986).
- [10] H. Kleinert, *Phys. Lett. B* **174**, 335 (1986).
- [11] H. Kleinert, *Phys. Rev. Lett.* **58**, 1915 (1987); *Phys. Lett. B* **211**, 151

- (1988); *Phys. Rev. D* **40**, 473 (1989); G. German and H. Kleinert, *Phys. Rev. D* **40**, 1108 (1989); *Phys. Lett. B* **220**, 133 (1989); *Phys. Lett. B* **225**, 107 (1989); M.C. Diamantini, H. Kleinert, and C.A. Trugenberger, *Phys. Rev. Lett.* **82**, 267 (1999), eprint: hep-th/9810171.
- [12] H. Kleinert and A. Chervyakov, *Phys. Lett. B* **381**, 286 (1996).
- [13] S.R. Wadia and S.R. Das, *Phys. Lett. B* **106**, 386 (1981).
- [14] D. Antonov, *Europhys. Lett.* **52**, 54 (2000), eprint: hep-th/0003043.
- [15] S.R. Das and S.R. Wadia, *Phys. Rev. D* **53**, 5856 (1996), eprint: hep-th/9503184.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

NAMBU-GOTO STRING WITHOUT TACHYONS BETWEEN A HEAVY AND A LIGHT QUARK

G. LAMBIASE

*Dipartimento di Scienze Fisiche E.R. Caianiello, Università di Salerno,
84081 Baronissi (Sa), Italy
E-mail: lambiase@sa.infn.it*

V.V. NESTERENKO

*Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research,
Dubna, 141980, Russia
E-mail: nestr@thsun1.jinr.ru*

We point out that the singularity in the interquark potential at small distances disappears in infinite space-time dimensions if a Nambu-Goto string is anchored at one end to an infinitely heavy quark and at the other to an infinitely light quark. This suggests that, if such quarks are placed at the ends of the string, some unphysical features such as tachyon states in the string spectrum are absent also in finite dimensions.

1 Introduction

We consider it a great honour presenting a contribution to the *Festschrift* dedicated to Professor Hagen Kleinert on the occasion of his 60th birthday. We were lucky having a very fruitful and instructive collaboration with him in our studies of string dynamics, when we had the opportunity to profit from Kleinert's intuition and his nonstandard approaches to complicated problems of theoretical physics.

Here we present the results of such a collaboration when we studied the effect of the boundary conditions imposed on dynamical variables (string coordinates) on the Nambu-Goto string spectrum. It is difficult to overestimate

the role of Professor Kleinert in obtaining these results which are in main part due to his skill of deducing far going physical inferences from the mathematical facts derived.

It is generally believed that some modification of the Nambu-Goto string model will eventually become a fundamental theory capable of explaining the forces between quarks in a simpler way than quantum chromodynamics (QCD). Indeed, the correct large-distance confinement behavior [1–3] is automatically obtained, by construction, whereas that in QCD can only be found by arduous lattice simulations [4,5]. Also the first quantum correction to this behavior, the universal Lüscher term [1,6], is found immediately. It is a one-loop contribution to the string energy and corresponds to the zero point energy of the small oscillations, coinciding with the Casimir energy at $T = 0$.

Certainly, it cannot be hoped that the Nambu-Goto string is anywhere close to the real color-electric flux tube between quarks since it is incapable of reproducing the $1/R$ -singularity at small R caused by the asymptotic freedom of gluons. Some essential modification accounting for the finite diameter of the flux tube will be necessary, in particular its transition into a spherical bag at small quark separations. A first attempt in this direction was taken some time ago by adding an asymptotically-free curvature stiffness term [7], but this term introduced other problems. In particular, the true stiffness constant of the flux tube appears with the opposite sign [8].

In spite of the essential differences between a Nambu-Goto string and a flux tube connecting quarks, the question arises how the unphysical properties of a Nambu-Goto string change if quarks are placed at its ends. The purpose of this note is to point out that in one extremal configuration, at least the singularity of the string potential disappears, indicating the absence of tachyons in that case [9].

To obtain a first idea about all properties of a fully fluctuating string, it is useful to investigate the limit of infinite space-time dimension D , where a saddle-point approximation to the functional integral yields exact results via one-loop calculations. This limit exhibits immediately an important unphysical feature of the Nambu-Goto string model [1]. There a complex string potential appears at distances smaller than a critical radius R_c , where the interquark potential vanishes. The existence of such a critical radius is attributed to tachyonic states in the string spectrum [10].

This and other unphysical properties are found in the so-called *static* interquark potential, where the string is anchored to immobile infinitely heavy

quarks. In this limit, the eigenfrequencies of the string are $\omega_n = n\pi/R$, $n = 1, 2, \dots$, where R is the distance between the quarks. The associated Casimir energy reads

$$E_C = \frac{D-2}{2} \sum_{n=1}^{\infty} \omega_n = -\frac{\pi(D-2)}{24R}, \quad (1)$$

yielding the well-known Lüscher term. Here Riemann's zeta function has been used [11].

We shall see in Section 3 that the Casimir energy determines the interquark potential completely in the limit $D \rightarrow \infty$, yielding

$$V(R) = M_0^2 R \sqrt{1 + \frac{\bar{R}^2}{R^2}}, \quad \bar{R}^2 = \frac{2RE_C}{M_0^2}. \quad (2)$$

Inserting (1), the potential calculated by Alvarez follows [1]

$$V_{\text{Alvarez}} = M_0^2 R \sqrt{1 - \frac{R_c^2}{R^2}}, \quad R_c^2 = \frac{\pi(D-2)}{12M_0^2}. \quad (3)$$

The quantity M_0^2 is the string tension. The same potential is found for strings with free ends due to the same Casimir energy (1).

2 Interquark Potential Generated by a String with Massive Ends

If a Nambu-Goto string has point-like quarks of masses m_1, m_2 at the ends moving along the worldlines $C_a, a = 1, 2$, the action reads [12] ($\hbar = c = 1$):

$$\mathcal{A} = -M_0^2 \iint d^2\xi \sqrt{g} - \sum_{a=1}^2 m_a \int_{C_a} ds_a, \quad (4)$$

where $g = \det(g_{\alpha\beta})$ is the determinant of the string metric. If the string coordinates are parameterized by $x^\mu(\xi)$, then $g_{\alpha\beta} = \partial_\alpha x^\mu \partial_\beta x_\mu$, $\alpha, \beta = 0, 1$. For calculating the interquark potential from such an action one conveniently uses the Gauss parameterization:

$$x^\mu(t, r) = (t, r, \mathbf{u}(t, r)), \quad 0 \leq r \leq R, \quad \mu = 0, 1, \dots, D-1, \quad (5)$$

where the vector field $\mathbf{u}(t, r) = (x^2(t, r), \dots, x^{D-1}(t, r))$ describes the transverse displacements of the string in D dimensions. Then $g_{\alpha\beta} = \delta_{\alpha\beta} +$

$\partial_\alpha \mathbf{u} \partial_\beta \mathbf{u}$, with $\mathbf{u} \mathbf{u} \equiv \sum_{j=2}^{D-1} u^j u^j$. The fluctuation spectrum is found from the linearized equations of motion and boundary conditions

$$\square \mathbf{u} = 0, \quad (6)$$

$$m_1 \ddot{\mathbf{u}} = M_0^2 \mathbf{u}', \quad r = 0, \quad (7)$$

$$m_2 \ddot{\mathbf{u}} = -M_0^2 \mathbf{u}', \quad r = R. \quad (8)$$

Here dots and primes denote the derivatives with respect to t and r , respectively, and $\square \equiv \partial^2/\partial t^2 - \partial^2/\partial r^2$. The general solution to these equations has the form

$$u^j(t, r) = i \frac{\sqrt{2}}{M_0^2} \sum_{n \neq 0} e^{-i\omega_n t} \frac{\alpha_n^j}{\omega_n} u_n(r), \quad j = 2, \dots, D-1, \quad (9)$$

where the amplitudes α_n^j satisfy the usual rule of the complex conjugation, $\alpha_n = \alpha_{-n}^*$. The unnormalized eigenfunctions $u_n(r)$ are

$$u_n(r) = \cos \omega_n r - \omega_n \frac{m_1}{M_0^2} \sin \omega_n r, \quad (10)$$

and the eigenfrequencies ω_n satisfy the secular equation

$$\tan \omega R = \frac{M_0^2 (m_1 + m_2) \omega}{m_1 m_2 \omega^2 - M_0^4}. \quad (11)$$

The Hamiltonian operator reads

$$H = \sum_n \sum_{j=2}^{D-1} \omega_n a_n^{j\dagger} a_n^j + E_C, \quad (12)$$

where E_C is the Casimir energy

$$E_C = \frac{D-2}{2} \sum_n \omega_n. \quad (13)$$

The creation and annihilation operators satisfy the usual commutation rules

$$[a_n^i, a_m^{j\dagger}] = \delta^{ij} \delta_{nm}. \quad (14)$$

The Casimir energy [13,14] renders the Lüscher correction to the interquark potential [6].

As in all field theories [14], the Casimir energy E_C of a string diverges for large n , and a renormalization is necessary to obtain physical results. If both masses are infinite or zero, the roots in Eq. (11) are $n\pi/R$ with integer n , and

the sum over eigenvalues is made finite with the help of the zeta function [11] in Eq. (1).

The interesting alternative situation is the limiting case, $m_1 = \infty$ and $m_2 = 0$, in which one end is fixed, the other free. Such a string approximates mesons consisting of one heavy and one light quark bound together by a color-electric flux tube. In this limit, the boundary conditions (7) and (8) simplify to

$$\mathbf{u}(t, 0) = 0, \quad \mathbf{u}'(t, R) = 0, \quad (15)$$

and the secular equation (11) assumes the form

$$\cos \omega R = 0, \quad (16)$$

which is solved by string eigenfrequencies ω_n being half-integer multiples of π/R : $\omega_n = (n + 1/2)\pi/R$ for $n = 0, 1, \dots$. Then, the Casimir energy is

$$E_C = \frac{\pi(D-2)}{48 R}, \quad (17)$$

where again the Riemann zeta function has been used [11,15]. The Casimir energy has now a *positive* sign, and half the magnitude, and Eq. (2) yields the interquark potential

$$V^{(R)} = M_0^2 R \sqrt{1 + \frac{1}{2} \frac{R_c^2}{R^2}}, \quad R_c^2 = \frac{\pi(D-2)}{12 M_0^2}. \quad (18)$$

This is an important result because the interquark potential (18) is real for all distances R in the limit $D \rightarrow \infty$. This implies that a string with these boundary conditions is a physical model for all distances R (certainly, again in the limit $D \rightarrow \infty$). Figure 1 compares the new string potential which is physical for all distances R with the Alvarez potential which is real only for $R > R_c$. This observation raises the question whether there might be an entire regime of asymmetric quark mass configurations for which the potential remains physical. We intend to study the general case of both masses being finite. Then the roots in Eq. (11) have the large- n behavior

$$\omega_n \simeq n \frac{\pi}{R} + \frac{M_0^2(m_1 + m_2)}{m_1 m_2} \frac{1}{n\pi} + O(n^{-3}). \quad (19)$$

The formal zeta function regularization can no longer be applied (since $\sum_{n=1}^{\infty} n^{-1} = \zeta(1) = \infty$), calling for a different and more physical subtraction procedure.

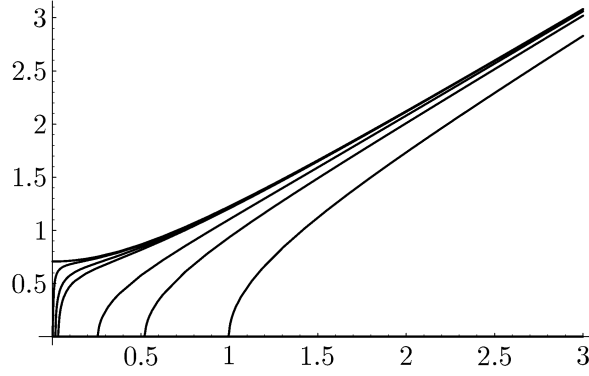


Figure 1. Dependence of the dimensionless interquark potential on the boundary conditions in the string model. The upper curve shows the potential (18) corresponding to the extremely asymmetric boundary condition of one string end being fixed, and the other free ($m_1 = \infty$, $m_2 = 0$). The lowest curve presents Alvarez' result (3) for strings with both ends fixed or free ($m_1 = \infty$, $m_2 = \infty$). The lengths are measured in units of Alvarez' critical radius R_c . The remaining curves show the potential for one infinite and one finite quark mass m_2 corresponding to the reduced mass parameter $\rho_2 = 0, 1/5, 1, 2, 10, 100, \infty$ (from lowest to highest curves).

There exists a simple analytic expression for the subtracted Casimir energy [16,17]. To find it we introduce the dimensionless frequency sum $S \equiv (12R/\pi) \sum_n \omega_n$ and rewrite it as

$$S = \frac{6Ri}{\pi^2} \int d\omega \omega \frac{d}{d\omega} \log[\cos(\omega R) M_0^2(m_1 + m_2) \omega - \sin(\omega R)(m_1 m_2 \omega^2 - M_0^4)] - (R \rightarrow \infty). \quad (20)$$

The derivative of the logarithm contains the solutions of the secular equation (11) as poles with unit residue. The contour of integration encloses the positive ω -axis in the clockwise sense. After opening up the contour and integrating along the imaginary frequency axis $\omega = iy$, a partial integration leads to

$$S = \frac{6R}{\pi^2} \int_{-\infty}^{\infty} dy \log[\cosh(yR) M_0^2(m_1 + m_2) y + \sinh(yR)(m_1 m_2 y^2 + M_0^4)] - (R \rightarrow \infty). \quad (21)$$

For a comparison of the behavior of the quark potential for various quark mass

Nambu-Goto String without Tachyons Between a Heavy and a Light Quark 631

configurations it is useful to go over to the dimensionless distance variable $\rho \equiv R/R_c$ and to reduced quantities $\rho_{1,2} \equiv R_{1,2}/R_c$, where $R_{1,2}$ are length parameters associated with the quark masses defined by

$$R_{1,2} \equiv \frac{\pi(D-2)}{12m_{1,2}}. \quad (22)$$

With the integration variable $z = yR$, we can rewrite S as

$$\begin{aligned} S(\rho) &= \frac{12}{\pi^2} \int_0^\infty dz \log [1 - e^{-2z} h(z, \rho)], \\ h(z, \rho) &= \frac{z^2 - (\rho_1 + \rho_2)\rho z + \rho_1 \rho_2 \rho^2}{z^2 + (\rho_1 + \rho_2)\rho z + \rho_1 \rho_2 \rho^2}. \end{aligned} \quad (23)$$

For $m_1 = \infty$, i.e. $\rho_1 = 0$, $S(\rho)$ is a simple function of $\rho_2 \rho$ which runs from $S = -1$ for $\rho_2 \rho = 0$ to $S = 1/2$ for $\rho_2 \rho = \infty$. In terms of $S(\rho)$, the interquark potential acquires the general form [16,18,19]

$$\frac{V}{M_0^2 R_c} = \rho \sqrt{1 + \frac{S(\rho)}{\rho^2}}. \quad (24)$$

In Fig. 1 we have plotted the potential for $\rho_1 = 0$ and different $\rho_2 = 0, 1/5, 1, 2, 10, 100, \infty$. The plot shows that only the limit $m_2 = 0$ is associated with a real interquark potential for all R . For a small but finite m_2 , the function $S(\rho)$ always becomes negative if the radius R is much smaller than m_2/M_0^2 .

3 Functional Integral for String Potential

Let us verify that the interquark potential is indeed determined by the Casimir energy as stated in Eq. (2). The potential $V(R)$ between massive quarks separated by a distance R is defined by the functional integral [6,20,21]

$$e^{-TV(R)} = \int [D\mathbf{u}] e^{-\mathcal{A}_E[\mathbf{u}]}, \quad T \rightarrow \infty, \quad (25)$$

where $\mathcal{A}_E$ is the Euclidean action

$$\mathcal{A}_E = M_0^2 \int_0^T dt \int_0^R dr \sqrt{\det(\delta_{\alpha\beta} + \partial_\alpha \mathbf{u} \partial_\beta \mathbf{u})} +$$

$$+ \sum_{a=1}^2 m_a \int_0^T dt \sqrt{1 + \dot{\mathbf{u}}^2(t, r_a)}. \quad (26)$$

We want to calculate the leading term for $D \rightarrow \infty$. As usual, we make the action harmonic in the string positions by introducing auxiliary composite fields $\sigma_{\alpha\beta}$ and by constraining these to be equal to $\partial_\alpha \mathbf{u} \partial_\beta \mathbf{u}$ by means of a Lagrange multiplier $\alpha^{\alpha\beta}$. By a similar manipulation, also the end-point actions can be made harmonic. After some manipulations, the functional integral (25) becomes Gaussian in $\mathbf{u}$ and can be performed with the result

$$e^{-TV(R)} = \int [D\alpha][D\sigma] e^{-\mathcal{A}_E[\alpha, \sigma]}, \quad T \rightarrow \infty, \quad (27)$$

where

$$\begin{aligned} \mathcal{A}_E = & M_0^2 \int_0^T dt \int_0^R dr \left[\sqrt{\det(\delta_{\alpha\beta} + \sigma_{\alpha\beta})} - \frac{1}{2} \alpha^{\alpha\beta} \sigma_{\alpha\beta} \right] + \\ & + \frac{D-2}{2} \text{Tr} \ln(-\partial_\alpha \alpha^{\alpha\beta} \partial_\beta). \end{aligned} \quad (28)$$

The boundary term in (26) is taken into account via the eigenvalues of the differential operator $-\partial_\alpha \alpha^{\alpha\beta} \partial_\beta$ in the action (28). As in Ref. [1], the functional integral is determined by the stationary point of (28) at which the matrices α and σ are diagonal. This simplifies the functional trace in (28) which becomes

$$\frac{D-2}{2} \text{Tr} \ln(-\partial_\alpha \alpha^{\alpha\beta} \partial_\beta) = T \sqrt{\frac{\alpha^{11}}{\alpha^{00}}} E_C. \quad (29)$$

Extremizing (28) with respect to $\sigma_{00}, \sigma_{11}, \alpha^{00}, \alpha^{11}$ yields indeed the string potential (2), as stated above.

4 Closing Remarks

It will be interesting to see whether the results derived in this note are present also in a finite dimension D . If this is so, then at least the limiting asymmetric quark mass configuration may be free of some of the unphysical features of present-day string models.

Finally we note that a dependence of the interquark potential on the quark masses at the ends was observed before in different ways [4,5]. In quantum field theory, the influence of different boundary conditions upon the Casimir effect has also been explored [14] resulting in energies of opposite signs.

Acknowledgments

The authors are indebted to the Editorial Board for inviting us to contribute to this *Festschrift*.

References

- [1] O. Alvarez, *Phys. Rev. D* **24**, 440 (1981).
- [2] H. Kleinert, *Phys. Rev. D* **40**, 473 (1989).
- [3] G. German, *Mod. Phys. Lett. A* **6**, 1815 (1991).
- [4] M. Flensburg and C. Peterson, *Nucl. Phys. B* **283**, 141 (1987).
- [5] K.D. Born *et al.*, *Phys. Lett. B* **239**, 325 (1994).
- [6] M. Lüscher, *Nucl. Phys. B* **180**, 317 (1981).
- [7] H. Kleinert, *Phys. Lett. B* **174**, 335 (1986); A.M. Polyakov, *Nucl. Phys. B* **268**, 406 (1986).
- [8] M.C. Diamantini, H. Kleinert, and C.A. Trugenberger, *Phys. Lett. B* **457**, 97 (1999); *Phys. Rev. Lett.* **82**, 267 (1999).
- [9] H. Kleinert, G. Lambiase, and V.V. Nesterenko, *Phys. Lett. B* **384**, 213 (1996).
- [10] P. Olesen, *Phys. Lett. B* **160**, 144, 408 (1985); *Phys. Lett. B* **168**, 220 (1986); *Nucl. Phys. B* **267**, 539 (1986).
- [11] E. Elizalde, S.D. Odintsov, A. Romeo, A.A. Bytsenko, and S. Zerbini, *Zeta Regularization Technique with Applications* (World Scientific, Singapore, 1994).
- [12] See, for instance, B.M. Barbashov and V.V. Nesterenko, *Introduction to the Relativistic String Theory* (World Scientific, Singapore, 1990).
- [13] V.M. Mostepanenko and N.N. Trunov, *The Casimir Effect and Its Applications* (Clarendon Press, Oxford, 1997).
- [14] G. Plunien, B. Müller, and W. Greiner, *Phys. Rep.* **134**, 87 (1986).
- [15] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products* (New York, Academic Press, 1980), p. 1074.
- [16] G. Lambiase and V.V. Nesterenko, *Phys. Rev. D* **54**, 6387 (1996).

- [17] H. Kleinert, *Mass Dependence of $1/R$ -Term in String Potential*, Preprint FU Berlin, 1986 (unpublished).
- [18] V.V. Nesterenko and I.G. Pirozhenko, *Phys. Rev. D* **55**, 6603 (1997).
- [19] G. German and Yu. Jiang, *Eur. Phys. J. C* **8**, 689 (1999).
- [20] M. Lüscher, K. Symanzik, and P. Weisz, *Nucl. Phys. B* **173**, 365 (1980).
- [21] K.G. Wilson, *Phys. Rev. D* **10**, 2445 (1974).

DECONFINEMENT OF QUARKS IN THE NAMBU-GOTO STRING WITH MASSIVE ENDS

V.V. NESTERENKO

*Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research,
Dubna, 141980, Russia*

E-mail: nestr@thsun1.jinr.ru

G. LAMBIASE

*Dipartimento di Scienze Fisiche E.R. Caianiello, Università di Salerno,
84081 - Baronissi (Sa), Italy.*

E-mail: lambiase@sa.infn.it

It is shown that the critical (or deconfinement) temperature for the Nambu-Goto string connecting the point-like masses (quarks) does not depend on the value of these masses and it is the same as that in the case of the string with fixed ends (infinitely heavy immobile quarks).

1 Introduction

We were lucky having collaborated with Professor Hagen Kleinert during our studies in string models. Prominent scientific achievements reached by him are well known but their presentation calls for a separate article. Here we would like to note briefly our impression from the personal contacts with him. The high incontestable authority of Professor Kleinert is harmoniously supplemented with his personal fascination. He literally "charges" colleagues by his creative energy. He is always full of ideas and new approaches to the problems under consideration and he readily shares those with his collaborators. To any question addressed to Professor H. Kleinert, one receives a clear-cut answer that testifies a profound understanding of the problem at hand. Kleinert's merry and slightly ironical temper renders the personal contacts with him easy and nice without any hint of pressure of his authority.

We have decided to present in this book the results of a direct continuation of our joint work with Professor H. Kleinert devoted to string dynamics [1]. In Refs. [2-4], the dependence of the interquark potential on the quark masses has been investigated in the model of the relativistic string with ends loaded by point-like masses. It was shown that allowing for finite values of quark masses leads to considerable corrections to the string potential in comparison with calculations when the string ends are fixed (immobile quarks with infinitely heavy masses). The critical radius^a in the string potential turns out to depend on the quark masses especially in the case of the asymmetric configurations when the string ties together light and heavy quarks [1].

Along with the critical radius of the interquark potential, an important parameter of the string model of hadrons is the temperature of deconfinement or critical temperature [6]. This prediction of the string model is directly compared with the lattice simulations in the framework of gauge theories [7]. In this connection the investigation of the quark mass contribution to the critical string temperature is of certain interest. It is this problem that will be treated in this paper.

Usually one believes that in string models the critical radius R_c of the potential and the critical temperature T_c are related. For example, in the Nambu-Goto string with fixed ends the relation $2 R_c T_c = 1$ holds. As the critical radius R_c depends on the quark masses [2], one could expect that this will also be valid for the critical temperature T_c . However, this is not the case [8]. The deconfinement temperature in the Nambu-Goto string with point-like masses (quarks) at its ends turns out to be independent of the quark masses, this temperature being the same as that in the string with fixed ends.

The layout of the paper is the following. In Section 2, we will recall the definition of the critical temperature in string models. A new approach to calculate the Casimir energy at finite temperature in the Nambu-Goto string of finite length is then suggested in Section 3. The temperature dependence of this energy determines the deconfinement or critical temperature in the string model under consideration. We show in Section 4 that the critical temperature in the Nambu-Goto string model with fixed ends and with point-like masses attached to the string ends is the same. Finally, in the conclusions,

^aAs is known [5], the potential generated by the string is not determined for any distances between quarks no sooner than at $R > R_c$, where R_c is the critical radius in the Nambu-Goto string with fixed ends which is determined by the string tension M_0^2 : $R_c^2 = \pi(D-2)/(12 M_0^2)$.

the obtained results are briefly discussed.

2 Critical Temperature in String Models

The critical temperature (or the temperature of deconfinement) in string models is defined in the following way. Let $V(R, T)$ be the free energy of the string (or string potential) calculated at finite temperature T . The asymptotics of this energy at large distances is

$$V(R, T) \rightarrow \sigma(T) R, \quad R \rightarrow \infty, \quad (1)$$

where $\sigma(T)$ is an effective string tension depending on the temperature T . At the critical value of T ($T = T_c$), the string tension vanishes

$$\sigma(T_c) = \lim_{R \rightarrow \infty} R^{-1} V(R, T_c) = 0. \quad (2)$$

The string potential at finite temperature is directly calculated in the same way as at $T = 0$ [2]. As a result, one obtains the well known square-root expression

$$V(R, T, m) = M_0^2 R \sqrt{1 + \frac{2(D-2)}{M_0^2 R} E_C(R, T, m)}, \quad (3)$$

where M_0^2 is the string tension at zero temperature, i.e. $\sigma(T = 0) = M_0^2$; m is the quark mass, $E_C(R, T, m)$ is the renormalized Casimir free energy for one transverse degree of freedom in the string model, and D is the dimension of the space-time. Usually one puts $D = 4$. Hence, determining the critical temperature requires the calculation of the Casimir free energy at finite temperature in the string model under investigation. For simplicity, we shall call the Casimir free energy briefly Casimir energy.

3 Casimir Energy at Finite Temperature

The Casimir energy in the Nambu-Goto string at finite temperature is given by [2]

$$E_C(R, T) = \frac{T}{2} \sum_{n=-\infty}^{+\infty} \sum_{k=1}^{\infty} \ln(\Omega_n^2 + \omega_k^2), \quad (4)$$

where Ω_n are the Matsubara frequencies $\Omega_n = 2\pi nT$, $n = 0, \pm 1, \dots$. Here T is the temperature, and ω_k are the eigenfrequencies of the string determined

by the boundary conditions for the string coordinates that, for fixed string ends (immobile quarks), assume the values

$$\omega_k = \frac{k\pi}{R}, \quad k = 1, 2, \dots \quad (5)$$

In the case of the string with masses at its ends, the frequencies ω_k are the positive roots of the following equation [2]

$$\tan(\omega R) = \frac{2\omega m}{\omega^2 - M_0^2}. \quad (6)$$

For simplicity the symmetric quark configuration is considered, i.e. $m_1 = m_2 = m$. Since the critical temperature is determined by the value of $E_C(R, T)$ in the limit $R \rightarrow \infty$ (see Eqs. (2) and (3)), it is convenient, in the case of string frequencies (5), to take this limit directly in Eq. (4) substituting the summation over k by integration

$$E_C^{\text{fixed}}(R \rightarrow \infty, T) = \frac{T}{4} \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} dk \ln \left[\Omega_n^2 + \left(\frac{k\pi}{R} \right)^2 \right] = -\frac{\pi R T^2}{6}. \quad (7)$$

Here the divergent integral and the divergent sum over n are calculated by using the analytical regularization and the Riemann zeta function, respectively. Substituting (7) in (3) we find the critical temperature in the Nambu-Goto string with fixed ends

$$\frac{T_c}{M_0} = \sqrt{\frac{3}{\pi(D-2)}}. \quad (8)$$

However this method cannot be applied to string frequencies determined by Eq. (6). Investigating the double sum [9] in (5) without introducing the integration over dk is again based on the fact that frequencies ω_k are multiples of π/R .

Here we use a new method [10] for calculating the Casimir energy at finite temperature that works equally well both with frequencies (5) and with string frequencies determined by Eq. (6). The idea of the method is the following. At first we represent the renormalized Casimir energy at zero temperature in terms of the integral over string eigenfrequencies. In other words, we obtain

the spectral representation

$$E_C(R, T = 0) = \int_0^{\infty} d\omega \mathcal{E}_C(R, \omega). \quad (9)$$

Then we pass in a standard way [11,12] from integration to summation over the Matsubara frequencies $\Omega_n = 2\pi nT$, $n = 0, \pm 1, \pm 2, \dots$. Practically this can be done by the following substitution in Eq. (9)

$$d\omega \rightarrow 2\pi T d\omega \sum_{n=0}^{\infty} \delta(\omega - \Omega_n), \quad (10)$$

with the result

$$E_C(R, T) = 2\pi T \sum_{n=0}^{\infty} \mathcal{E}_C(R, \Omega_n). \quad (11)$$

The prime in the summation symbol means that the term with $n = 0$ should be taken with a multiplier $1/2$.

It is worth noting here that the formal substitution (10) can lead to the free energy or to the internal energy of the quantum system under consideration. Therefore one has to be careful when applying this procedure [10].

The integral representation for the Casimir energy in the Nambu-Goto string with fixed ends is given by [2]

$$E_C^{\text{fixed}} = \frac{1}{2\pi} \int_0^{\infty} d\omega \ln(1 - \exp(-2\omega R)) = -\frac{R}{\pi} \int_0^{\infty} \frac{\omega d\omega}{\exp(2\omega R) - 1}. \quad (12)$$

The last expression in this formula is obtained by integration by parts. Omitting the minus sign in (12), we see that the spectral density of energy in this formula is of Planckian form for the one-dimensional black body, the effective temperature of the string vacuum being equal to $(2R)^{-1}$. Applying the algorithm explained above, we find

$$E_C^{\text{fixed}}(R, T) = -4\pi T^2 R \sum_{n=0}^{\infty} \frac{n}{\exp(4\pi nRT) - 1}. \quad (13)$$

Integration by parts in (12) was required for obtaining the term with $n = 0$ in the sum (13) without divergencies. In order to overcome the stated difficulty in analogous calculations in statistical physics [11,12], the Casimir force is

calculated at first and then the corresponding potential is recovered on this basis.

The sum in (13) can be evaluated in the two limiting cases of large and small temperatures. At large T , the main contribution in (13) is due to the term with $n = 0$:

$$E_C^{\text{fixed}}(R, T \rightarrow \infty) = -\frac{T}{2}. \quad (14)$$

Using the Euler-Maclaurin formula for small T ,

$$\sum_{n=0}^{\infty} F(n) = \int_0^{\infty} F(x) dx - \frac{1}{12} F'(0) + \dots,$$

Eq. (13) reduces to the form

$$E_C^{\text{fixed}}(R, T) = -\frac{\pi}{24R} - \frac{\pi RT^2}{6}. \quad (15)$$

Hence, unlike Eq. (7), we also preserve here the T -independent term $-\pi/24R$ which vanishes when $R \rightarrow \infty$. Proceeding from the physical consideration, it is clear that the string picture of quark confinement inside hadrons is applicable only at low temperatures. In string models the temperature scale is determined by the string tension $M_0 \sim 0.4 \text{ GeV}$. Hence when finding the critical temperature in string models, the region of small temperatures should be considered [6]. After allowing all this, we have to substitute in (1)–(3) the expression for the Casimir energy at small T (Eq. (15)). After taking the limit $R \rightarrow \infty$ in (2), the contribution of the first term in (15) to the effective string tension vanishes and we obtain the critical temperature (8).

Equation (15) allows one to introduce a critical temperature for the string of a finite length $\bar{T}_c(R)$. To this end, $\lim_{R \rightarrow \infty}$ should be removed from definition (2). At this temperature, the effective tension of the string of finite length R vanishes. It is obvious that the string critical temperature defined in this way will be dependent on the string length R . By making use of Eqs. (2), (3), and (15), we obtain

$$\bar{\tau}_c^2(\rho) = \tau_c^2 - \frac{1}{4\rho^2}, \quad (16)$$

where the following dimensionless variables are introduced

$$\tau_c = T_c/M_0, \quad \rho = M_0 R. \quad (17)$$

When $\rho \rightarrow \infty$, the critical temperature $\bar{\tau}_c(\rho)$, dependent on the string length ρ , tends to its limiting value τ_c from below, because the existence of longer strings requires a greater work for their splitting (phase transition) than it takes place in the case of short strings.

In another approach, the critical temperature for the Nambu-Goto string of finite length has been considered in Ref. [13]. The results obtained there probably imply that $\tau_c(\rho)$ tends to τ_c from above in the limit $\rho \rightarrow \infty$.

4 Critical Temperature for the Nambu-Goto String with Massive Ends

The method for calculating the Casimir energy at finite temperature presented in the preceding section can be directly applied to the Nambu-Goto string with massive ends. The integral representation of the Casimir energy at zero temperature in this string model is [2]

$$E_C(R, T=0, m) = \frac{1}{2\pi} \int_0^\infty d\omega \ln \left[1 - e^{-2\omega R} \left(\frac{\omega m - M_0^2}{\omega m + M_0^2} \right)^2 \right], \quad (18)$$

where m is the quark mass. We consider, for simplicity, the quarks with equal masses (the Casimir energy and the string potential for different quark masses $m_1 \neq m_2$ have been investigated in Ref. [1]). Integrating by parts in (18), one obtains

$$E_C(R, T=0, m) = -\frac{R}{\pi} \int_0^\infty \frac{\omega d\omega}{g(\omega) - 1} \left(1 - \frac{2mM_0^2}{R(\omega^2 m^2 - M_0^4)} \right), \quad (19)$$

where we have

$$g(\omega) = e^{2\omega R} \left(\frac{\omega m + M_0^2}{\omega m - M_0^2} \right)^2. \quad (20)$$

Notice that now the spectral density of energy in (19) is not Planckian. Then, the Casimir energy at finite temperature is

$$E_C(R, T, m) = -\sum_{n=0}^{\infty} \frac{4\pi T^2 R n}{g(2\pi n T) - 1} \left(1 - \frac{2mM_0^2}{R(4\pi^2 n^2 T^2 m^2 - M_0^4)} \right), \quad (21)$$

where $g(x)$ is defined in (20). When $m \rightarrow \infty$ or $m \rightarrow 0$, Eq. (21) reduces to (13). It is important to note that the Casimir energies at finite temper-

ature given by the sums (13) and (21) are free of any divergencies. It is a direct consequence of using the integral representations (12) and (18) for the *renormalized* Casimir energies at zero temperature.

In the limit $T \rightarrow \infty$, the Casimir energy (21) tends to the value

$$E_C(R, T \rightarrow \infty, m) = -\frac{T}{2} \left(1 + \frac{2m}{M_0^2 R} \right). \quad (22)$$

At low temperature the estimation of the sum in (21) can be done by making again use of the Euler-MacLaurin formula, where the quantity $F'(0)$ should be calculated. Here, the function $F(x)$ is given by

$$F(x) = \frac{x(1 - \eta x)^2}{e^{2Rx}(1 + \eta x)^2 - (1 - \eta x)^2} \left(1 - \frac{2\eta}{R(\eta^2 x^2 - 1)} \right), \quad (23)$$

where η is the ratio m/M_0^2 . Expanding $F(x)$ in a Taylor series gives

$$F(x) \simeq \frac{[1 - (R + 2\eta)x + \mathcal{O}(x^2)]}{2(R + 2\eta)} \left[1 + \frac{2\eta}{R}(1 + \eta^2 x^2) + \mathcal{O}(x^4) \right]. \quad (24)$$

Then, it follows that

$$F'(0) = -\frac{1}{2} \left(1 + \frac{2\eta}{R} \right) \rightarrow -\frac{1}{2} \quad \text{when} \quad R \rightarrow \infty. \quad (25)$$

In the limit $R \rightarrow \infty$ the integral term in the Euler-McLaurin formula vanishes. Hence, the critical temperature in the model under investigation turns out to be the same as that in the Nambu-Goto string with fixed ends

$$\tau_c^2 = \left(\frac{T_c}{M_0} \right)^2 = \frac{3}{\pi(D-2)}. \quad (26)$$

If we consider again the critical temperature dependent on the string length (see the end of the preceding section), then this quantity proves to be dependent on the quark mass. Indeed, now we have, instead of (16),

$$\bar{\tau}_c^2(\rho, \mu) = \left(1 + \frac{2\mu}{\rho} \right)^{-1} \left(\tau_c^2 - \frac{I(q)}{4\rho^2} \right), \quad (27)$$

where $\mu = m/M_0$, $\rho = M_0 R$, $q = M_0^2 R/m = \rho/\mu$, and where the function $I(q)$, generated by the integral term in the Euler-McLaurin formula, is given

by

$$I(q) = \frac{24}{\pi^2} \int_0^\infty dz \frac{z(z-q)}{e^{2z}(z+q)^2 - (z-q)^2} \frac{z^2 - q^2 - 2q}{z+q}. \quad (28)$$

When $m \rightarrow \infty$, the function $I(q)$ tends to 1. In order to obtain Eq. (16) from (27) in this limit, one has to send ρ to ∞ in the first multiplier in (27).

5 Conclusions

The method for calculating the Casimir energy at finite temperature proposed here enables us to find the critical temperature in the Nambu-Goto string with massive ends. Before it was not obvious that this temperature is the same as that in the string model with fixed string ends. Probably this is due to the fact that in string models, like in statistical models, the boundary effects appear to be unessential for implementing phase transitions.

Our consideration is restricted to the one-string approximation. Certainly, it is important to calculate the critical temperature for a gas of interacting strings with massive ends and to take in this way into account the processes of splitting and joining such strings which are followed by creation and annihilation of quark-antiquark pairs. Through this mechanism the quark masses will be involved in determining the deconfinement temperature in a more direct way. Unfortunately, the theory of interacting hadronic strings is still unknown. Therefore, new nonstandard methods should be developed for treating this problem.

References

- [1] H. Kleinert, G. Lambiase, and V.V. Nesterenko, *Phys. Lett. B* **384**, 213 (1996).
- [2] G. Lambiase and V.V. Nesterenko, *Phys. Rev. D* **54**, 6387 (1996).
- [3] V.V. Nesterenko and I.G. Pirozhenko, *Phys. Rev. D* **55**, 6603 (1997).
- [4] G. German and Yu. Jiang, *Eur. Phys. J. C* **8**, 689 (1999).
- [5] O. Alvarez, *Phys. Rev. D* **24**, 440 (1981).
- [6] R.D. Pisarski and O. Alvarez, *Phys. Rev. D* **26**, 3735 (1982).
- [7] K.D. Born, E. Laermann, R. Sommer, P.M. Zerwas, and T.F. Walsh, *Phys. Lett. B* **329**, 325 (1994).
- [8] G. Lambiase and V.V. Nesterenko, *Phys. Lett. B* **398**, 335 (1997).

- [9] M. Flensburg and C. Peterson, *Nucl. Phys. B* **283**, 141 (1987).
- [10] V.V. Nesterenko and I.G. Pirozhenko, *J. Math. Phys.* **38**, 6265 (1997).
- [11] I.E. Dzyaloshinskii, E.M. Lifshitz, and L.P. Pitaevskii, *Advances in Physics* **10**, No. 38, 165 (1961).
- [12] E.M. Lifshitz and L.P. Pitaevskii, *Statistical Physics*, Part 2 (Pergamon, Oxford, 1980).
- [13] A. Antillón and G. German, *Phys. Rev. D* **47**, 4567 (1993).

PRINCIPLES OF NON-LOCAL FIELD THEORIES AND THEIR APPLICATION TO POLYMERIZED MEMBRANES

K.J. WIESE

*Institute for Theoretical Physics, University of California at Santa Barbara,
Santa Barbara, CA 93106-4030, USA*

E-mail: wiese@itp.ucsb.edu

We give an overview about non-local field theories and their application to polymerized membranes, i.e. membranes with a fixed internal connectivity. The main technical tool is the multi-local operator product expansion (MOPE), generalizing ideas from local field theories to the multi-local situation [1,2].

1 Polymerized Membranes

Membranes have attracted much interest during the last years, especially due to their relevance for biological systems. Most of the membranes encountered there are fluid. I shall concentrate here on another fascinating class of membranes, namely polymerized tethered also called solid membranes. These have a fixed and constant internal metric. They are realized in experiments (e.g. the spectrin network of red blood cells [3,4], or sheets of graphite oxide [5,6]).

A microscopic model is given by the so-called "spring and bead model" which consists of balls (beads) which are connected by springs and form a regular lattice. The difficulty is to incorporate the self-avoidance of the beads, i.e. the fact that beads cannot intersect.

2 Representation of Membranes by a Non-Local Field Theory

We start from the continuous model for a D -dimensional flexible polymerized membrane introduced in Refs. [7,8], and further studied in Refs. [9-12]. This

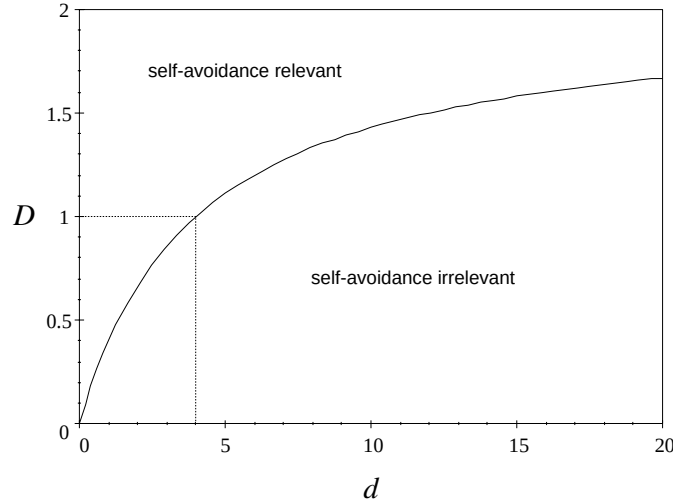


Figure 1. The critical curve $\varepsilon(D, d) = 0$. The dashed line corresponds to the standard polymer perturbation theory, critical in $d = 4$.

model is a simple extension of the well-known Edwards' model for continuous chains [13]. The membrane fluctuates in d -dimensional space. Points in the membrane are labeled by coordinates $x \in \mathbb{R}^D$, and the configuration of the membrane in physical space is described by the field $r : x \in \mathbb{R}^D \rightarrow r(x) \in \mathbb{R}^d$. At high temperatures, the free energy for a configuration is given by the (properly scaled) Hamiltonian

$$\mathcal{H}[r] = \frac{1}{2-D} \int_x \frac{1}{2} [\nabla r(x)]^2 + b_0 \int_x \int_y \tilde{\delta}^d(r(x) - r(y)) . \quad (1)$$

The integral $\int_x$ runs over D -dimensional space and ∇ is the usual gradient operator. The normalizations are $\int_x := \int d^D x / S_D$ with $S_D = 2\pi^{D/2} / \Gamma(D/2)$ and $\tilde{\delta}^d(r(x) - r(y)) = (4\pi)^{d/2} \delta^d(r(x) - r(y))$. The latter term is normally used in Fourier-representation $\tilde{\delta}^d(r(x) - r(y)) = \int_p e^{ip[r(x) - r(y)]}$, where the normalization of $\int_p$ is given by $\int_p = \pi^{-d/2} \int d^d p$ to have $\int_p e^{-p^2 a} = a^{-d/2}$. All normalizations are chosen in order to simplify the calculations. The first term in the Hamiltonian is a Gaussian elastic energy which is known to describe the free "phantom" surface. The interaction term corresponds (for $b_0 > 0$) to a weak repulsive interaction upon contact. The expectation values of physical observables are obtained by performing the average over all field

configurations $r(x)$ with the Boltzmann weight $e^{-\mathcal{H}[r]}$. This average can not be calculated exactly, but one can expand about the configurations of a phantom, i.e. a non-interacting surface.

Such a perturbation theory is constructed by performing the series expansion in powers of the coupling constant b . This expansion suffers from ultraviolet (UV) divergences which have to be removed by renormalization and which are treated by dimensional regularization, i.e. analytical continuation in D and d . Long-range infrared (IR) divergences also appear. They can be eliminated by using a finite membrane, or by studying translationally invariant observables, whose perturbative expansion is also IR-finite in the thermodynamic limit (infinite membrane). Such observables are "neutral" products of vertex operators

$$\mathcal{O} = \prod_{a=1}^N e^{ik_a r(x_a)} , \quad \sum_{a=1}^N k_a = 0 . \quad (2)$$

Let us now analyze the theory by power-counting. We use internal units $\mu \sim 1/x$, and note $[x]_x = 1$, and $[\mu]_x = -[\mu]_\mu = -1$. The dimension of the field and of the coupling constant are:

$$\nu := [r]_x = \frac{2-D}{2} , \quad \varepsilon := [b_0]_\mu = 2D - \nu d . \quad (3)$$

The interaction is relevant for $\varepsilon > 0$ (see Fig. 1). Perturbation theory is then expected to be UV-finite, except for subtractions associated to relevant operators. We shall come back to this point later.

For clarity, we represent graphically the different interaction terms which have to be considered. The local operators are

$$1 = \mathbf{1} , \quad \frac{1}{2}(\nabla r(x))^2 = \star . \quad (4)$$

The bi-local operator, the dipole, is

$$\tilde{\delta}^d(r(x) - r(y)) = \bullet \text{---} \bullet . \quad (5)$$

The expectation-value of an observable is

$$\langle \mathcal{O}[r] \rangle_b = \text{Norm} \int \mathcal{D}[r] \mathcal{O}[r] e^{-\mathcal{H}[r]} , \quad (6)$$

with Norm chosen such that $\langle \mathbf{1} \rangle_b = 1$. Perturbatively, all expectation-values are taken with respect to the free theory (again normalized s.t. $\langle \mathbf{1} \rangle_0 = 1$):

$$\langle \mathcal{O}[r] \rangle_0 = \text{Norm}' \int \mathcal{D}[r] \mathcal{O}[r] \exp \left\{ -\frac{1}{2-D} \int_x \frac{1}{2} (\nabla r(x))^2 \right\} . \quad (7)$$

A typical term in the expansion of Eq. (6) is

$$(-bZ_b\mu^\varepsilon)^n \iint \dots \iint \langle \mathcal{O} \bullet \dots \bullet \rangle_0^c , \quad (8)$$

where the integral runs over the positions of all dipole-endpoints.

3 Locality of Divergences and the Multilocal Operator Product Expansion (MOPE)

The next step to show is that all divergences are short distance divergences. Note that even for massless theories and in the absence of IR-divergences, this is not trivial. Divergences could as well appear, when some of the distances involved become equal, or a multiple of each other. A simple counter-example is the integral of $||a| - |b||^{-\nu d}$, where a and b are two of the distances involved. The proof is given in Ref. [2].

As in local field theory, divergences can be analyzed via an operator product expansion. Intuitively, in the context of multilocal theories – the interaction depends on more than one point – we also expect multilocal operators to appear in such an operator product expansion, which therefore will be called “multi-local operator product expansion” (MOPE) [14,15]. An example for such an operator product expansion is

$$= \sum_i C_i \left(\begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \\ \bullet \end{array} \right), \left(\begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \right), \left(\begin{array}{c} \bullet \\ \bullet \end{array} \right) \cdot \text{diagram with shaded circle } i \text{ and two external lines} . \quad (9)$$

Let us explain the formula. We consider n dipoles (here $n = 5$) and we separate the $2n$ end-points into m subsets (here $m = 3$) delimited by the dashed lines. The MOPE describes how the product of these n dipoles behaves when the points inside each of the m subsets are contracted towards a single point z_j . The result is a sum over *multilocal* operators $\Phi_i(z_1, \dots, z_m)$, depending

on the m points $z_1, \dots, z_m$, of the form

$$\sum_i C_i(x_1 - z_1, \dots) \Phi_i(z_1, z_2, \dots, z_m) , \quad (10)$$

where the MOPE-coefficients $C_i(x_1 - z_1, \dots)$ depend only on the distances $x_l - z_l$ *inside each subset*. This expansion is valid as an operator identity, i.e. it is inserted into any expectation value and in the limit of small distances between contracted points. No other operator should appear at the points $z_1, \dots, z_m$ towards which the operators are contracted. As the Hamiltonian (1) does not contain a mass scale, the MOPE-coefficients are homogeneous functions of the distances between the contracted points, with the degree of homogeneity given by dimensional analysis. In the case considered here, where n dipoles are contracted to an operator Φ_i , this degree is $-nvd - [\Phi_i]_x$. This means that

$$C_i(\lambda(x_1 - z_1), \dots) = \lambda^{-n\frac{2-D}{2}d - [\Phi_i]_x} C_i(x_1 - z_1, \dots) , \quad (11)$$

where $[\Phi_i]_x$ is the canonical dimension of the operator Φ_i and $-d(2-D)/2$ is the canonical dimension of the dipole.

In order to evaluate the associated singularity, one finally has to integrate over all relative distances inside each subset. This gives an additional scale factor with degree $D(2n - m)$. A singular configuration, such as in Eq. (9), will be UV-divergent if the degree of divergence

$$D(2n - m) - n\frac{2-D}{2}d - [\Phi_i]_x \quad (12)$$

is negative. It is superficially divergent if the degree is zero and convergent otherwise. The idea of renormalization is *to remove exactly these superficially divergent contributions recursively*.

4 Evaluation of the MOPE-Coefficients

The MOPE gives a convenient and powerful tool to calculate the dominant and all subdominant contributions from singular configurations. Let us explain how to calculate the MOPE-coefficients on some explicit examples which are the necessary diagrams at 1-loop order.

In the following we shall use the notion of normal-ordering. First let us state that

$$:e^{ikr(x)}: = e^{ikr(x)} . \quad (13)$$

Explicitly, tadpole-like contributions being powers of $\int d^D p/p^2$ are omitted. This is done via a finite part prescription (analytic continuation, dimensional regularization), valid for infinite membranes, for which the normal-order prescription is defined. Let us stress that this is a pure technical trick, which can be circumvented at the expense of more cumbersome calculations.

The key formula for all further manipulations is

$$:e^{ikr(x)}::e^{ipr(y)}: = e^{kpC(x-y)} :e^{ikr(x)}e^{ipr(y)}: . \quad (14)$$

This can be proven as follows: Consider the (free) expectation value of any observable $\mathcal{O}$ times the operators of Eq. (14). Then the left- and right-hand sides of the above equation read $\mathcal{L} = \langle \mathcal{O} :e^{ikr(x)}::e^{ipr(y)}: \rangle_0$ and $\mathcal{R} = e^{kpC(x-y)} \langle \mathcal{O} :e^{ikr(x)}e^{ipr(y)}: \rangle_0$. First of all, for $\mathcal{O} = 1$, the desired equality of $\mathcal{L} = \mathcal{R}$ holds, because $\langle :e^{ikr(x)}e^{ipr(y)}: \rangle_0 = 1$ and $\langle :e^{ikr(x)}::e^{ipr(y)}: \rangle_0 = e^{kpC(x-y)}$. Now consider a non-trivial observable $\mathcal{O}$, and contract all its fields r with $e^{ikr(x)}$ or $e^{ipr(y)}$, before contracting any of the fields $r(x)$ with $r(y)$. The result is a product of correlation functions between the points in $\mathcal{O}$ and x or y , and these are equivalent for both $\mathcal{L}$ and $\mathcal{R}$. However, contracting an arbitrary number of times $e^{ikr(x)}$, leaves the exponential $e^{ikr(x)}$ invariant. Completing the contractions for $\mathcal{L}$ therefore yields a factor of $e^{kpC(x-y)}$, and the latter one also appears in $\mathcal{R}$. Thus, the equality of $\mathcal{L}$ and $\mathcal{R}$ holds for all $\mathcal{O}$ and this proves Eq. (14).

We proceed to the first explicit example, the contraction of a single dipole with endpoints x and y :

$$x \text{---}\bullet\text{---}\bullet\text{---}y = \int_k :e^{ikr(x)}::e^{-ikr(y)}: . \quad (15)$$

This configuration may have divergences when x and y come close together. Let us stress that in contrast to ϕ^4 -theory, these divergences are not obtained as a finite sum of products of correlators: Since $C(x-y) = |x-y|^{2-D}$, the latter is always well-behaved at $x = y$. The singularity only appears when summing an infinite series of diagrams. To this purpose, we first normal-order the two exponentials using Eq. (14)

$$\int_k :e^{ik[r(x)-r(y)]}: e^{-k^2|x-y|^{2\nu}} . \quad (16)$$

Note that the operators $e^{ikr(x)}$ and $e^{-ikr(y)}$ are free of divergences upon approaching each other, since no more contractions can be made. The divergence is captured in the factor $e^{-k^2|x-y|^{2\nu}}$. Therefore we can expand the

exponential : $e^{ik[r(x)-r(y)]}$: for small $x - y$ and consequently in powers of $[r(x) - r(y)]$. This expansion is

$$\int_k : \left\{ \mathbf{1} + ik[r(x) - r(y)] - \frac{1}{2} (k[r(x) - r(y)])^2 + \dots \right\} : e^{-k^2|x-y|^{2\nu}} . \quad (17)$$

The expansion is truncated after the third term. It will turn out later that this is sufficient, since subsequent terms in the expansion are proportional to irrelevant operators for which the integral over the MOPE-coefficient is UV-convergent.

Due to the symmetry of the integration over k the term linear in k vanishes. Also due to symmetry, the next term can be simplified with the result

$$\int_k \left[\mathbf{1} - \frac{k^2}{2d} : [r(x) - r(y)]^2 : + \dots \right] e^{-k^2|x-y|^{2\nu}} . \quad (18)$$

Finally, the integration over k can be performed. Recall that normalizations were chosen such that $\int_k e^{-sk^2} = s^{-d/2}$. We obtain

$$\left(\text{dipole} \right) = \left(\text{dipole} \middle| \mathbf{1} \right) \mathbf{1} + \left(\text{dipole} \middle| \alpha \dagger \beta \right) \alpha \dagger \beta + \dots , \quad (19)$$

where we have used the notation $\alpha \dagger \beta = \frac{1}{2}(\partial_\alpha r)(\partial_\beta r)$ and the MOPE-coefficients (reminding Feynman's bra and ket notation)

$$\left(\text{dipole} \middle| \mathbf{1} \right) = |x - y|^{-\nu d} , \quad (20)$$

$$\left(\text{dipole} \middle| \alpha \dagger \beta \right) = -\frac{1}{2} (x - y)_\alpha (x - y)_\beta |x - y|^{-\nu(d+2)} . \quad (21)$$

As long as the angular average is taken (and this will be the case when integrating the MOPE-coefficient to obtain the divergence), we can replace, in Eq. (19), $\alpha \dagger \beta$ by $\dagger := \frac{1}{2}(\nabla r)^2$, and Eq. (21) by

$$\left(\text{dipole} \middle| \dagger \right) = -\frac{1}{2D} |x - y|^{D-\nu d} . \quad (22)$$

Next consider a real multi-local example of an operator-product expansion, namely the contraction of two dipoles towards a single dipole:

$$\begin{array}{c} x+u/2 \\ x-u/2 \end{array} \left(\text{dipole} \right) \begin{array}{c} y+v/2 \\ y-v/2 \end{array} = \int_k e^{ik[r(x+u/2)-r(y+v/2)]} \int_p e^{ip[r(x-u/2)-r(y-v/2)]} . \quad (23)$$

This has to be analyzed for small u and v , in order to control the divergences in the latter distances. As above, we normal-order operators which are approached, yielding

$$\begin{aligned} e^{ikr(x+u/2)} e^{ipr(x-u/2)} &= :e^{ikr(x+u/2)} : :e^{ipr(x-u/2)} : \\ &= :e^{ikr(x+u/2)} e^{ipr(x-u/2)} : e^{kpC(u)} . \end{aligned} \quad (24)$$

A similar formula holds when approaching $e^{-ikr(y+v/2)}$ and $e^{-ipr(y-v/2)}$. Eq. (23) then becomes

$$\int_k \int_p :e^{ikr(x+u/2)+ipr(x-u/2)} : :e^{-ikr(y+v/2)-ipr(y-v/2)} : e^{kp[C(u)+C(v)]} . \quad (25)$$

In order to keep things as simple as possible, let us first extract the leading contribution before analyzing subleading corrections. This leading contribution is obtained when expanding the exponential operators (here exemplified for the second one) as

$$:e^{-ikr(y+v/2)} e^{-ipr(y-v/2)} : = :e^{-i(k+p)r(y)} (1 + O(\nabla r)) : \quad (26)$$

and dropping terms of order ∇r . This simplifies Eq. (25) to

$$\int_k \int_p :e^{i(k+p)r(x)} : :e^{-i(k+p)r(y)} : e^{kp[C(u)+C(v)]} . \quad (27)$$

In the next step, first k and second p are shifted, $k \rightarrow k-p$ then $p \rightarrow p+k/2$. The result is (dropping the normal-ordering according to Eq. (13))

$$\int_k e^{ik[r(x)-r(y)]} \int_p e^{(\frac{1}{4}k^2-p^2)[C(u)+C(v)]} . \quad (28)$$

The factor of $\int_k e^{ik[r(x)-r(y)]}$ is again a δ -distribution, and the leading term of the short-distance expansion of Eq. (28). Derivatives of the δ -distribution appear when expanding $e^{(\frac{1}{4}k^2-p^2)[C(u)+C(v)]}$ in k^2 ; these are less relevant and only the first sub-leading term will be displayed for illustration:

$$\begin{aligned} \int_k e^{ik[r(x)-r(y)]} \int_p e^{-p^2[C(u)+C(v)]} \left(1 + \frac{k^2}{4} [C(u) + C(v)] + \dots \right) \\ = (\text{diagram 1}) + (\text{diagram 2}) + \dots , \end{aligned} \quad (29)$$

where in analogy to Eqs. (19) to (21)

$$(\text{diagram 1}) = [C(u) + C(v)]^{-d/2} ,$$

$$(\text{diagram}) = \frac{1}{4} [C(u) + C(v)]^{1-d/2}, \quad (30)$$

with $\longrightarrow = \tilde{\delta}^d(r(x) - r(y))$ and $\longleftrightarrow = (-\Delta_r) \tilde{\delta}^d(r(x) - r(y))$.

Let us mention that the leading contribution proportional to the δ -distribution will renormalize the coupling constant, and that the next-to-leading term is irrelevant and can be neglected. The same holds for the additional term proportional to (∇r) which was dropped in Eq. (26).

There is one more possible divergent contribution at the 1-loop level, namely diagram . We now show that the leading term of its expansion, which is expected to be proportional to $\longrightarrow$, is trivial. To this aim consider

$$\begin{aligned} \text{diagram} &= \int_{k,p} :e^{ikr(u)}: :e^{-ikr(x)}: :e^{ipr(y)}: :e^{-ipr(z)}: \\ &= \int_{k,p} :e^{ikr(u)}: :e^{-ikr(x)}: e^{ipr(y)} e^{-ipr(z)} :e^{-p^2 C(y-z)} e^{kp[C(x-z)-C(x-y)]} . \end{aligned}$$

In order to study the contraction of x , y , and z , and to look for all contributions which are proportional to $\longrightarrow = \int_k :e^{ikr(u)}: :e^{-ikr((x+y+z)/3)}:$, we observe that, in Eq. (31), the leading term is obtained by approximating $e^{kp[C(x-z)-C(x-y)]} \approx 1$. All subsequent terms yield factors of k , which after integration over k give derivatives of the $\tilde{\delta}^d$ -distribution. The result is

$$(\text{diagram}) - (\text{diagram}) = 0. \quad (31)$$

This means that divergences of diagram are already taken into account by a proper treatment of the divergences in diagram , analyzed in Eq. (19).

5 Renormalization at 1-Loop Order

In order to renormalize the theory, we shall identify the divergences through the MOPE, and then introduce counter terms which subtract these divergences. These counter terms are nothing else than integrals over the MOPE-coefficients, properly regularized, i.e. cut off. We introduce two renormalization group factors Z (renormalizing the field r) and Z_b (renormalizing the coupling b). Recalling Eq. (1), this becomes

$$\mathcal{H}[r] = \frac{Z}{2-D} \int_x \frac{1}{2} (\nabla r(x))^2 + b Z_b \mu^\varepsilon \int_x \int_y \tilde{\delta}^d(r(x) - r(y)) , \quad (32)$$

where r and b are the renormalized field and renormalized dimensionless coupling constant, and $\mu = L^{-1}$ is the renormalization momentum scale. (To be precise: The field r in Eq. (1) is the bare field and should be noted r_0 .)

Let us start to eliminate the divergences in the case, where the end points (x, y) of a single dipole are contracted towards a point (taken here to be the center-of-mass $z = (x + y)/2$). The MOPE was given in Eq. (19) and following. We have to distinguish between counter terms for relevant operators and those for marginal operators. The former can be defined by analytic continuation, while the latter require a subtraction scale. Indeed, the divergence proportional to $\mathbf{1}$ is given by the integral

$$\int_{\Lambda^{-1} < |x-y| < L} \left(x \text{---} \text{---} y \mid \mathbf{1} \right) = \int_{\Lambda^{-1}}^L \frac{dx}{x} x^{D-\nu d} = \frac{1}{D-\varepsilon} (\Lambda^{D-\varepsilon} - L^{\varepsilon-D}) , \quad (33)$$

where Λ is a high-momentum UV-regulator and L a large-distance regulator. For $\varepsilon \approx 0$ this is UV-divergent but IR-convergent. The simplest way to subtract this divergence is therefore to replace the dipole operator by

$$x \text{---} \text{---} y \longrightarrow x \text{---} \text{---} y - x \text{---} \text{---} y , \quad (34)$$

where $x \text{---} \text{---} y = |x - y|^{-\nu d}$. This amounts to adding to the bare Hamiltonian (1) the UV-divergent counter-term

$$\Delta \mathcal{H}_1 = -b Z_b \mu^\varepsilon \int_x \int_y |x - y|^{-\nu d} , \quad (35)$$

which is a pure number and thus does not change the expectation value of any physical observable.

We next consider marginal operators: In the MOPE of Eq. (21), the integral over the relative distance of $\int_{x-y} (x \text{---} \text{---} y \mid \alpha \dagger \beta)_{\alpha \dagger \beta}$ is logarithmically divergent at $\varepsilon = 0$. In order to find the appropriate counter term, we use dimensional regularization, i.e. set $\varepsilon > 0$. An IR-cutoff L , or equivalently a subtraction momentum scale $\mu = L^{-1}$, has to be introduced in order to define the subtraction operation. *As a general rule, let us integrate over all distances appearing in the MOPE-coefficient, bounded by the subtraction scale $L = \mu^{-1}$.* Defining

$$\left\langle \text{---} \text{---} \mid \alpha \dagger \beta \right\rangle_L := \int_{|x-y| < L} \left(x \text{---} \text{---} y \mid \alpha \dagger \beta \right) , \quad (36)$$

we need the following counter term in the Hamiltonian

$$\Delta\mathcal{H}_+ = -b\mu^\varepsilon \left\langle \text{Diagram} \middle| \alpha \star \beta \right\rangle_L \int_x \alpha \star \beta_x, \quad (37)$$

subtracting explicitly the divergence in the integrals. The reader is invited to verify this explicitly in the example of the expectation value of $\mathcal{O} = e^{ik[r(s)-r(t)]}$. The solution is given in Appendix H of Ref. [2].

Since the angular integration in Eq. (36) reduces $\alpha \star \beta$ to $\star$, we can replace Eq. (37) by the equivalent expression

$$\Delta\mathcal{H}_+ = -b\mu^\varepsilon \left\langle \text{Diagram} \middle| \star \right\rangle_L \int_x \star_x, \quad (38)$$

for which the numerical value of the diagram is calculated as

$$\left\langle \text{Diagram} \middle| \star \right\rangle_L = \int_{|x-y|<L} \left(\text{Diagram} \middle| \star \right) = -\frac{1}{2D} \int_0^L \frac{dx}{x} x^{2D-\nu d} = -\frac{1}{2D} \frac{L^\varepsilon}{\varepsilon}. \quad (39)$$

We can now subtract this term in a minimal subtraction scheme (MS). The internal dimension of the membrane D is kept fixed and (39) is expanded as a Laurent series in ε , which starts here at ε^{-1} . The residue of the pole in Eq. (39) is

$$\left\langle \text{Diagram} \middle| \star \right\rangle_\varepsilon = -\frac{1}{2D}. \quad (40)$$

It is this pole that is subtracted in the MS-scheme by adding to the Hamiltonian a counter term

$$\Delta\mathcal{H}_+ = -\frac{b}{\varepsilon} \left\langle \text{Diagram} \middle| \star \right\rangle_\varepsilon \int_x \star_x. \quad (41)$$

Similarly, the divergence arising from the contraction of two dipoles to a single dipole is subtracted by a counter term

$$\Delta\mathcal{H}_\leftarrow = b^2\mu^{2\varepsilon} \left\langle \text{Diagram} \middle| \longrightarrow \right\rangle_L \int_x \int_y \text{Diagram}, \quad (42)$$

with

$$\left\langle \text{Diagram} \middle| \longrightarrow \right\rangle_L = \int_{|x|<L} \int_{|y|<L} \left(\text{Diagram} \middle| \longrightarrow \right). \quad (43)$$

Reducing this integral counter term to a number, we subtract the residue of the single pole of

$$\langle \text{diagram} \rangle_L = \int_{|x|<L} \int_{|y|<L} (\text{diagram}) = \int_{|x|<L} \int_{|y|<L} (|x|^{2\nu} + |y|^{2\nu})^{-d/2}. \quad (44)$$

Note that the regulator L cuts off both integrations. One can now either utilize some simple algebra or show, by the methods of conformal mapping (see Ref. [2]), that the residue is obtained by fixing one distance to 1 and by freely integrating over the remaining one:

$$\langle \text{diagram} \rangle_\varepsilon = \int_0^\infty \frac{dx}{x} x^D (1 + x^{2-D})^{-\frac{2D}{2-D}} = \frac{1}{2-D} \frac{\Gamma\left(\frac{D}{2-D}\right)^2}{\Gamma\left(\frac{2D}{2-D}\right)}. \quad (45)$$

As a result, the model is UV-finite at 1-loop order, if we use, in the renormalized Hamiltonian (32), the renormalization factors

$$Z = 1 - (2-D) \langle \text{diagram} \rangle_\varepsilon \frac{b}{\varepsilon} + O(b^2), \quad (46)$$

$$Z_b = 1 + \langle \text{diagram} \rangle_\varepsilon \frac{b}{\varepsilon} + O(b^2). \quad (47)$$

Note that due to Eq. (31) no counter term for diagram is necessary.

The renormalized field and coupling are re-expressed in terms of their bare counterparts through

$$r_0(x) = Z^{1/2} r(x), \quad b_0 = b Z_b Z^{d/2} \mu^\varepsilon. \quad (48)$$

Finally, the renormalization group functions are obtained from the variation of the coupling constant and the field with respect to the renormalization scale μ , keeping the bare coupling fixed. The flow of the coupling is written in terms of Z and Z_b with the diagrams given in Eqs. (40) and (45) as

$$\begin{aligned} \beta(b) &:= \mu \frac{\partial}{\partial \mu} \bigg|_{b_0} b = \frac{-\varepsilon b}{1 + b \frac{\partial}{\partial b} \ln Z_b + \frac{d}{2} b \frac{\partial}{\partial b} \ln Z} \\ &= -\varepsilon b + \left(\langle \text{diagram} \rangle_\varepsilon - \nu d \langle \text{diagram} \rangle_\varepsilon \right) b^2 + O(b^3). \end{aligned} \quad (49)$$

Similarly, the full dimension of the field (the exponent entering into the cor-

relation function) is obtained as

$$\begin{aligned}\nu(b) &:= \frac{2-D}{2} - \frac{1}{2}\mu \frac{\partial}{\partial \mu} \bigg|_{b_0} \ln Z = \frac{2-D}{2} - \frac{1}{2}\beta(b) \frac{\partial}{\partial b} \ln Z \\ &= \frac{2-D}{2} \left[1 - b \left\langle \text{diagram} \right\rangle_\varepsilon \right] + O(b^2) = \frac{2-D}{2} \left[1 + b \frac{1}{2D} \right] + O(b^2). \quad (50)\end{aligned}$$

Note that minimal subtraction is used on the level of counter terms or equivalently Z -factors. Since Z enters as Z^d into the β -function, the latter also contains a factor of d in the 1-loop approximation, i.e. Z^d is not minimally renormalized. In order to calculate the leading order in ε , the factor of d can be replaced by $d_c = 4D/(2-D)$.

The β -function has a non-trivial fixed point with $\beta(b^*) = 0$, which has a positive slope and thus describes the behavior of the model at large distances. The anomalous dimension $\nu^* := \nu(b^*)$ becomes to first order in ε

$$\nu^* = \frac{2-D}{2} \left[1 + \frac{\varepsilon}{2D} \left(\frac{1}{2-D} \frac{\Gamma(\frac{D}{2-D})^2}{\Gamma(\frac{2D}{2-D})} + 1 \right)^{-1} \right] + O(\varepsilon^2). \quad (51)$$

For polymers, this result reduces to the well-known formula

$$\nu^*(D=1) = \frac{1}{2} + \frac{4-d}{16} + O((4-d)^2).$$

6 Results for Self-Avoiding Membranes from 2-Loop Calculations

Two-loop results are obtained by an explicit analytic and numeric integration of the following combination of diagrams (which are all of order $1/\varepsilon$):

$$\begin{aligned}\mathcal{C}_1 &= -\frac{2}{3} \left\langle \text{diagram} \right\rangle_L + \left(\left\langle \text{diagram} \right\rangle_L \right)^2, \\ \mathcal{C}_2 &= -2 \left\langle \text{diagram} \right\rangle_L + \left\langle \text{diagram} \right\rangle_L \left\langle \text{diagram} \right\rangle_L, \\ \mathcal{C}_3 &= \left\langle \text{diagram} \right\rangle_{\varepsilon^{-1}} \left(\left\langle \text{diagram} \right\rangle_L + \frac{(2-D)d}{4} \left\langle \text{diagram} \right\rangle_{\varepsilon^{-1}} \right), \\ \mathcal{F}_1 &= \frac{1}{2} \left\langle \text{diagram} \right\rangle_L - \frac{1}{2} \left\langle \text{diagram} \right\rangle_L \left\langle \text{diagram} \right\rangle_L - \frac{1}{2} \left\langle \text{diagram} \right\rangle_L \left\langle \text{diagram} \right\rangle_L, \\ \mathcal{F}_2 &= -\frac{1}{2} \left\langle \text{diagram} \right\rangle_{\varepsilon^{-1}} \times \\ &\quad \times \left(\left\langle \text{diagram} \right\rangle_L + \frac{(2-D)(d+2)}{2} \left\langle \text{diagram} \right\rangle_L - \left\langle \text{diagram} \right\rangle_L \left\langle \text{diagram} \right\rangle_L \right),\end{aligned}$$

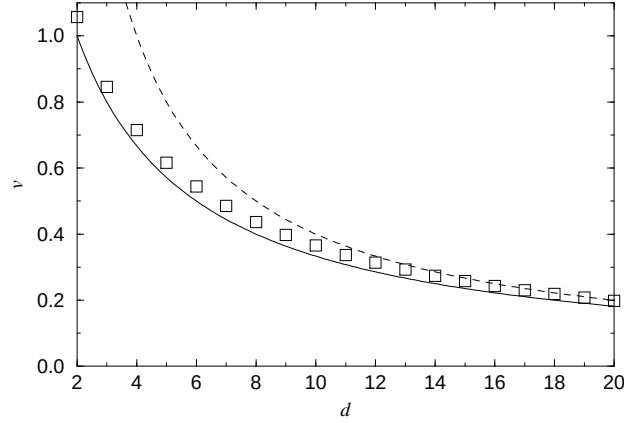


Figure 2. Extrapolation of the 2-loop results in d and ε for membranes ($D = 2$) in d dimensions, using the expansion of $\nu^*(d+2)$ (squares). The solid line is the prediction made by Flory's approximation, the dashed line by the variational ansatz.

$$\mathcal{F}_3 = \frac{1}{2} \left\langle \left(\text{diagram} \right) \right\rangle_{\varepsilon^{-1}} \left(\left\langle \left(\text{diagram} \right) \right\rangle_L - \left\langle \left(\text{diagram} \right) \right\rangle_{\varepsilon^{-1}} \right). \quad (52)$$

One then has the freedom to extrapolate about any point on the critical curve $\varepsilon(D, d) = 0$ (see Fig. 1). This freedom – or ambiguity – can be used to optimize the results. The results of such a procedure for the 2-loop calculations of ν^* are given in Fig. 2 for membranes ($D = 2$) in d dimensions ($2 \leq d \leq 20$). We see that for $d \rightarrow \infty$ the prediction of a Gaussian variational method $\nu_{\text{var}} = 2D/d$ becomes exact. For small d , the prediction made by Flory's argument $\nu_{\text{Flory}} = (2 + D)/(2 + d)$ is close to our results. This is a non-trivial result, since the membrane case corresponds to $\varepsilon = 4$, and in comparison with polymers in $d = 3$, where $\varepsilon = 1/2$, the 2-loop corrections were expected to be large. In fact, they are small when one expands around the critical curve $\varepsilon = 0$ for an adequate range of $D \sim 1.5$ (depending slightly on d and on the choice of variables) and a suitable choice of extrapolation variables. In this case the 2-loop corrections are even smaller than the 1-loop corrections and allow for more reliable extrapolations to $\varepsilon = 4$. This can be understood from the large-order behavior [16].

Let us now turn to the physically relevant case of membranes in three dimensions ($D = 2$, $d = 3$). Our calculations predict an exponent $\nu^* \approx 0.85$ or equivalently a fractal dimension of $d_f \approx 2.4$, which is in agreement with those experiments and simulations finding a fractal phase. However, this is

still under debate, and a lot of evidence has been collected that the flat phase is generic.

7 Outlook

We have here only discussed the simplest applications of non-local field theories. Much more has been achieved during the last years.

First of all, it has been demonstrated that the dynamics, described by a Langevin equation

$$\frac{d}{dt}r(x, t) = -\frac{\partial \mathcal{H}[r]}{\partial r(x, t)} + \zeta(x, t), \quad \langle \zeta(x, t) \zeta(x', t') \rangle = \lambda \delta^D(x - x') \delta(t - t') \quad (53)$$

also leads to a renormalizable field theory [17], where the dynamic exponent z , defined by the decay of the auto-correlation function $\langle [r(x, t) - r(x, t')]^2 \rangle \sim |t - t'|^{2/z}$ is given by $z = 2 + D/\nu^* = 2 + d_f$ ($d_f = D/\nu^*$ is the fractal dimension of the membrane). This result had been suggested long-time ago by De Gennes [18] for polymers and by Kardar *et al.* for membranes [19], but only with the methods discussed above, a proof of that conjectures could be given [17]. When hydrodynamics is included, the dynamical exponent z changes to $z = d$ [20].

Interestingly, also disorder can be treated via the same methods, since averaging over disorder leads to interactions very similar to self-avoidance. The above methods have been applied to the motion of a polymer or membrane in a static disordered force field with both potential and non-conserved parts – leading to new universal physics [21,22].

The advection of a polymer in a turbulent flow has also been analyzed [23], paralleling the discussion of the passive advection of particles, which in the turbulence community is known as the passive scalar problem.

Another interesting generalization concerns anisotropically tethered membranes [24-26]. In these models, the membrane is more rigid in one direction, forming a tubular phase.

One of the most useful tricks for self-avoiding polymers is the mapping onto a massive scalar field theory, i.e. a ϕ^4 -theory [27]. In the limit, where the number of components $n \rightarrow 0$, results for self-avoiding polymers are obtained. Stated differently, ϕ^4 -theory is a generalization of self-avoiding polymers. Another generalization has been discussed above: The generalization to membranes with internal dimension $D \neq 1$. The question arises, whether a common generalization of both the ϕ^4 -theory and polymerized membranes

is possible. Such a model has indeed been constructed [28,29] and leads to interesting new physics.

Another still open question is the analysis of the spectrum of subdominant operators [30]. This may give a hint of why polymerized membranes are generically seen to be flat in simulations.

Acknowledgments

It is a pleasure to thank the organizers of the workshop at Heidelberg, Hans Jürgen Pirner and Franz Wegner, for the opportunity to give these lectures. I wish the honoree of this volume 60 more years of fruitful scientific work.

References

- [1] This is a summary of lecture notes held for the Graduiertenkolleg in Heidelberg, September 25-27, 2000. For a more detailed version, see Ref. [2].
- [2] K.J. Wiese, *Polymerized Membranes, a Review in Phase Transitions and Critical Phenomena*, Vol. 19, Eds. C. Domb and J.L. Lebowitz (Academic Press, London, 2000).
- [3] A. Elsgaeter, B.T. Stokke, A. Mikkelsen, and D. Branton, *Science* **234**, 1217 (1986).
- [4] C.F. Schmidt, K. Svoboda, N. Lei, I.B. Petsche, L.E. Berman, C.R. Safinya, and G.S. Grest, *Science* **259**, 952 (1993).
- [5] T. Hwa, E. Kokufuta, and T. Tanaka, *Phys. Rev. A* **44**, 2235 (1991).
- [6] M.S. Spector, E. Naranjo, S. Chiruvolu, and J.A. Zasadzinski, *Phys. Rev. Lett.* **73**, 2867 (1994).
- [7] J.A. Aronovitz and T.C. Lubensky, *Europhys. Lett.* **4**, 395 (1987).
- [8] M. Kardar and D.R. Nelson, *Phys. Rev. Lett.* **58**, 1289, 2280E (1987).
- [9] F. David, B. Duplantier, and E. Guitter, *Phys. Rev. Lett.* **70**, 2205 (1993).
- [10] F. David, B. Duplantier, and E. Guitter, *Nucl. Phys. B* **394**, 555 (1993).
- [11] B. Duplantier, *Phys. Rev. Lett.* **58**, 2733 (1987).
- [12] B. Duplantier, T. Hwa, and M. Kardar, *Phys. Rev. Lett.* **64**, 2022 (1990).
- [13] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).

- [14] F. David, B. Duplantier, and E. Guitter, *Phys. Rev. Lett.* **72**, 311 (1994).
- [15] F. David, B. Duplantier, and E. Guitter, eprint: cond-mat/9702136.
- [16] F. David and K.J. Wiese, *Nucl. Phys. B* **535**, 555 (1998).
- [17] K.J. Wiese, *Eur. Phys. J. B* **1**, 269 (1998).
- [18] P.G. De Gennes, *Macromolécules* **9**, 587 (1976).
- [19] Y. Kantor, M. Kardar, and D.R. Nelson, *Phys. Rev. A* **35**, 3056 (1987).
- [20] K.J. Wiese, *Eur. Phys. J. B* **1**, 273 (1998).
- [21] P. Le Doussal and K.J. Wiese, *Phys. Rev. Lett.* **80**, 2362 (1998).
- [22] K.J. Wiese and P. Le Doussal, *Nucl. Phys. B* **552**, 529 (1999).
- [23] K.J. Wiese, eprint: chao-dyn/9911005.
- [24] M. Bowick, M. Falcioni, and G. Thorleifsson, *Phys. Rev. Lett.* **79**, 885 (1997).
- [25] M. Bowick and E. Guitter, *Phys. Rev. E* **56**, 7023 (1997).
- [26] M. Bowick and A. Travesset, *Phys. Rev. E* **59**, 5659 (1999).
- [27] P.G. De Gennes, *Phys. Lett. A* **38**, 339 (1972).
- [28] K.J. Wiese and M. Kardar, *Nucl. Phys. B* **528**, 469 (1998).
- [29] K.J. Wiese and M. Kardar, *Eur. Phys. J. B* **7**, 187 (1998).
- [30] K.J. Wiese and M. Shpot, in preparation.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

RANDOM PATHS AND SURFACES WITH RIGIDITY

B. DURHUUS

*Department of Mathematics, University of Copenhagen, Universitetsparken 5,
DK-2100 Copenhagen Ø, Denmark*

E-mail: durhuus@math.ku.dk

We discuss discretized models of random paths and surfaces with particular emphasis on a rigorous analysis of the influence of rigidity terms in the action upon the continuum properties.

1 Introduction

In the present article we study a class of models of fluctuating paths and surfaces, whose statistical weight depends on geometric properties of the fluctuating object. In particular, a rigidity dependence may be introduced in the form of curvature dependent terms in the action functional. Our main purpose is to outline results from a rigorous analysis of the scaling properties of suitably regularized models. For random paths it turns out to be possible to analyze in detail the scaling behavior of rigidity terms in the action. We show that such terms are irrelevant perturbations (essentially as a consequence of the central limit theorem), and have only an effect on the scaling limit when the associated bare coupling constant is fine tuned to infinity. In this case, there are non-vanishing correlations between tangents to the paths, as opposed to ordinary Brownian paths. It is nevertheless possible to obtain an explicit expression for the corresponding propagator.

The situation for fluctuating surfaces is more complicated. A general framework, analogous to that for paths, can however be set up in a straightforward way, and a number of general results can be obtained. Thus, there is substantial evidence that for small values of the curvature coupling the only possible scaling limit equals that of a simple random walk, as we explain in Section 3. This result can be viewed as some sort of generalization

of the central limit theorem to random surfaces, although as a surface theory the limit is degenerate. This phenomenon, in fact, makes the investigation of rigidity enhancing mechanisms for surfaces an important, albeit difficult issue. In particular, the question poses itself whether there exists a critical curvature coupling at which a non-degenerate scaling limit can be obtained. We also comment on this problem, but the results obtained so far are only fragmentary.

It is important to note that the models we consider are without any self-avoidance constraints, i.e. the paths and surfaces are allowed to self-intersect and overlap arbitrarily. As such, the models are not directly interpretable as representing fluctuating polymers or membranes, except in dimensions high enough to render the self-avoidance constraint irrelevant. Rather, we shall consider the models from a field- or string-theoretic point of view and regard the discretizations as regularizations of appropriate functional integrals to be explained below.

2 Random Paths with Rigidity

For definiteness we consider models of piecewise linear random paths, also called random flight models. The paths will be parameterized on the interval $[0, 1]$ and, if the path $\omega : [0, 1] \rightarrow R^d$ has N linear steps, the i 'th step is parameterized linearly on the interval $[(i-1)/N, i/N]$ such that the path and its parameterization are uniquely fixed by the points $x_i = \omega(i/N)$, $i = 0, \dots, N$. Letting θ_i denote the angle between the i 'th and the $(i+1)$ 'th step, i.e.

$$\theta_i = \arccos \frac{r_i \cdot r_{i+1}}{|r_i||r_{i+1}|},$$

where

$$r_i = x_{i+1} - x_i,$$

we define a propagator (or two-point function) by

$$G_{\mu, \lambda}(x, y) = \sum_{N=1}^{\infty} \int \prod_{i=1}^{N-1} dx_i \exp \left\{ \sum_{i=1}^N \phi(|x_i - x_{i-1}|) + \lambda \sum_{i=1}^{N-1} \psi(\theta_i) + \mu N \right\}, \quad (1)$$

where we have set $\omega(0) = x$ and $\omega(1) = y$. Here ϕ and ψ are suitable continuous and non-negative functions to be specified more closely below,

and μ, λ are coupling constants.

The right hand side of Eq. (1) should be viewed as an integral over all piecewise linear paths, where each path contributes with a weight $\exp[-S(\omega)]$, and the action $S(\omega)$ is given by

$$S(\omega) = \sum_{i=1}^N \phi(|x_i - x_{i-1}|) + \lambda \sum_{i=1}^{N-1} \psi(\theta_i) + \mu N. \quad (2)$$

This action is actually a gauge fixed and discretized form of the continuum action

$$S(\omega) = \int_a^b dt e(t) \left\{ \phi_0 \left[\frac{|\dot{\omega}(t)|}{e(t)} \right] + \lambda_0 \psi_0 \left[\frac{k(t)}{e(t)} \right] + \mu_0 \right\}, \quad (3)$$

where $\omega : [a, b] \rightarrow R^d$ is now a (piecewise) smooth path and $e : [a, b] \rightarrow R_+$ is an intrinsic metric on the one-dimensional manifold ω . Moreover, k denotes the *extrinsic curvature* of ω defined as

$$k(t) = \frac{\{|\dot{\omega}|^2 |\ddot{\omega}|^2 - (\dot{\omega} \cdot \ddot{\omega})^2\}^{\frac{1}{2}}}{|\dot{\omega}|^2}.$$

Indeed, S is invariant under the reparametrizations

$$t' = \varphi(t), \quad \omega'(t') = \omega(t), \quad e'(t') = \frac{e(t)}{\dot{\varphi}(t)},$$

where φ is an increasing diffeomorphism between intervals. A unique parameterization is determined by fixing the parameter interval to be $[0, 1]$ and the metric $e(t)$ to be constant and equal to its total volume T . With this parameterization and ω piecewise linear the two actions (2) and (3) coincide [1], if the identification

$$T = a^2 N$$

is made, where a is a parameter with dimension of a length, to be viewed as a short distance cutoff, and ϕ and ψ are dimensionless analogs of ϕ_0 and ψ_0 .

Specifying ϕ_0 and ψ_0 suitably, one obtains the most familiar form [2] of the action

$$S(\omega) = \int_a^b dt \left(\frac{1}{2} \frac{|\dot{\omega}(t)|^2}{e(t)} + \lambda_0 k(t) + \mu_0 e(t) \right). \quad (4)$$

As will be seen, however, the scaling limit is essentially independent of the exact form of the functions ϕ and ψ . We shall demand that the function ψ

vanishes at 0, but is otherwise positive in the interval $]0, \pi]$, such that non-straight paths get increasingly suppressed with increasing λ . For $\lambda = 0$, it is well known that generic paths in the continuum limit are far from being smooth, the set of such paths having Wiener measure 0. We shall see that this still holds for any fixed value of λ , whereas for $\lambda \rightarrow \infty$ we can obtain different continuum limits.

First, let us discuss the domain of convergence of Eq. (1). Because of translation invariance of $G_{\mu,\lambda}$, it is convenient to work with its Fourier transform

$$\hat{G}_{\mu,\lambda}(p) = \int dx G_{\mu,\lambda}(0, x) e^{-ip \cdot x} \quad (5)$$

and to introduce the two functions

$$F(s) = \int_0^\infty dr r^{d-1} e^{-\phi(r)} e^{-isr} \quad (6)$$

and

$$K_\lambda(\hat{r}, \hat{r}') = e^{-\lambda\psi(\theta(\hat{r}, \hat{r}'))},$$

where $s \in R$, $\hat{r}, \hat{r}' \in S^{d-1}$ and $\theta(\hat{r}, \hat{r}')$ denotes the angle between $\hat{r}$ and $\hat{r}'$. We consider $K_\lambda(\hat{r}, \hat{r}')$ as the kernel of an integral operator K_λ on $L^2(S^{d-1})$, and for fixed $p \in R^d$ we let F_p denote multiplication by the function $F(p \cdot \hat{r})$ acting on $L^2(S^{d-1})$. One can then rewrite (1) as

$$\hat{G}_{\mu,\lambda}(p) = \langle 1 | (1 - e^{-\mu} F_p K_\lambda)^{-1} e^{-\mu} F_p | 1 \rangle, \quad (7)$$

where $\langle \cdot | \cdot \rangle$ is the inner product in $L^2(S^{d-1})$ with respect to the uniform measure $d\Omega$ and where $|1\rangle$ is the constant function 1 on S^{d-1} . Henceforth, we assume that ϕ increases sufficiently fast such that F is finite and, moreover, twice continuously differentiable.

Using that $|1\rangle$ is an eigenfunction of K_λ with eigenvalue

$$\|K_\lambda\| = \int d\Omega(\hat{r}') e^{-\lambda\psi(\theta(\hat{r}, \hat{r}'))},$$

one finds that the *susceptibility*

$$\chi(\mu, \lambda) = \hat{G}(\mu, \lambda)(0)$$

is given by

$$\chi(\mu, \lambda) = \frac{\omega_d e^{-\mu} F_0}{1 - e^{-\mu} \|K_\lambda\| F_0}, \quad (8)$$

where now $F_0 = F(0)$ and ω_d is the volume of S^{d-1} .

It follows that $\chi(\mu, \lambda)$ and hence also $\hat{G}_{\mu, \lambda}(p)$ is finite for all $p \in R^d$ exactly, if $e^\mu > \|K_\lambda\| F_0$, i.e. the critical line in the (μ, λ) -plane is given by

$$e^{\mu_0} = \|K_\lambda\| F_0, \quad (9)$$

which defines the critical coupling $\mu_0(\lambda)$ as a function of the extrinsic curvature coupling λ .

Now, let us first fix λ and consider $\hat{G}_{\mu, \lambda}(ap)$ for μ close to but larger than $\mu_0(\lambda)$, where we have replaced p by ap so that p , from now on, is a dimensionful (physical) momentum variable. Expanding

$$F(ap \cdot \hat{r}) = F_0(1 + ic_1(p \cdot \hat{r})a - c_2(p \cdot \hat{r})^2 a^2) + O(a^3) \quad (10)$$

and

$$K_\lambda = \|K_\lambda\|(P + R_\lambda),$$

where c_1, c_2 are positive constants and P denotes the projection onto the constant modes on S^{d-1} , one shows that $\|R_\lambda\| < 1$, and hence, by Eq. (7), that the singularity of $\hat{G}_{\mu, \lambda}(ap)$ as a function of μ close to $\mu_0(\lambda)$ is solely due to the largest eigenvalue of the leading term in the operator $e^{-\mu} F_p K_\lambda$ approaching 1. Moreover, since $\langle 1|(p \cdot \hat{r})1 \rangle = 0$ one finds that only the second-order term in Eq. (10) contributes in Eq. (7) for small a . To obtain a non-trivial limit, this forces us to set (to leading order)

$$\mu(a) = \mu_0(\lambda) + m_1^2 a^2 \quad (11)$$

for some positive constant m_1^2 . Then the denominator in Eq. (7) is of order a^2 and it follows [1,3] that

$$\lim_{a \rightarrow 0} a^2 \hat{G}_{\mu, \lambda}(ap) = \frac{c}{m^2 + p^2}, \quad (12)$$

where c and m^2 are positive constants.

This proves our first claim that the scaling limit of $G_{\mu, \lambda}(x, y)$ equals the free scalar propagator, regardless of the value of λ .

Next, we consider the possibility of letting $\lambda \rightarrow \infty$ and thus forcing $\mu \rightarrow \infty$ according to Eq. (9). In this limit one finds that all eigenvalues of $\|K_\lambda\|^{-1} K_\lambda$

tend to 1, and hence they all contribute in Eq. (7). More precisely, one finds [1]

$$K_\lambda = \|K_\lambda\| [1 + C(\lambda)L + o(C(\lambda))] ,$$

where L is the Laplace-Beltrami operator on $L^2(S^{d-1})$ and $C(\lambda)$ is a positive function tending to 0 as $\lambda \rightarrow \infty$. Since all eigenvalues of K_λ contribute on the right hand side of Eq. (7), the contribution from F_p is from the first-order term in Eq. (10). Hence, we set

$$\mu(a) = \mu_0(a) + c_1 \mu_R a \quad (13)$$

and choose $\lambda(a)$ such that

$$C(\lambda(a)) = c_1 \lambda_R a ,$$

where $\mu_R, \lambda_R > 0$ are renormalized coupling constants. With this choice one gets

$$\lim_{a \rightarrow 0} a \hat{G}_{\mu, \lambda}(ap) = \langle 1 | (\mu_R + \lambda_R L - ip \cdot \hat{r})^{-1} | 1 \rangle .$$

In fact, more generally, if the directions of the first and last steps are held fixed at $\hat{r}'$ and $\hat{r}''$, respectively, one gets for the scaling limit of the corresponding two-point function $\hat{G}_{\mu, \lambda}(p; \hat{r}', \hat{r}'')$ the result

$$\lim_{a \rightarrow 0} a \hat{G}_{\mu, \lambda}(ap; \hat{r}', \hat{r}'') = \langle \hat{r}' | (\mu_R + \lambda_R L - ip \cdot \hat{r})^{-1} | \hat{r}'' \rangle . \quad (14)$$

This fact explicitly entails the non-vanishing correlation between tangents to the paths, and hence also implicitly shows that the paths have acquired a rigidity represented by the coupling constant λ_R . It is an interesting open problem to identify the measure on the space of paths associated with this limit and to analyze in more detail the smoothness properties of generic paths. It can be argued [1,3] that their Hausdorff dimension is 1, whereas for standard Wiener paths it is known to be 2. In fact, this difference is intimately connected to the different scalings used in Eqs. (11) and (13) via scaling relations.

It is worth noting that the operator method sketched above can be applied to other types of models of fluctuating paths as well, such as the Ornstein-Uhlenbeck process [3].

3 Random Surfaces with Rigidity

In order to extend the framework of the preceding section to random surfaces we first describe the analogue of the action functional (3).

Let $X : D \rightarrow R^d$ be a parameterized surface in R^d defined on some fixed oriented parameter domain D . By $n_1, \dots, n_{d-2}$ we denote an oriented orthonormal basis in the normal bundle to X (defined locally), and we let

$$D_{a;ij} = \delta_{ij} \partial_a + n_i \cdot \partial_a n_j$$

be the covariant derivative in the normal bundle. Then the *extrinsic curvature* H of X is defined by

$$H^2 = \frac{1}{4} h^{ab} \sum_i D_a n_i \cdot D_b n_i ,$$

where h_{ab} is the first fundamental form of X .

A natural generalization of Eq. (3) is therefore

$$\begin{aligned} S(X, g_{ab}) = \int_D d^2 \xi \sqrt{g} \Big\{ & \phi_0 (g^{ab} \partial_a X \cdot \partial_b X) \\ & + \lambda_0 \psi_0 \left(g^{ab} \sum_i D_a n_i \cdot D_b n_i \right) + \mu_0 \Big\} , \end{aligned} \quad (15)$$

where g_{ab} is an intrinsic metric defined on D , and g is its determinant. Obviously, S is invariant under reparametrizations

$$\xi' = \varphi(\xi), \quad X'(\xi') = X(\xi), \quad g'_{ab}(\xi') \partial_c \varphi^a(\xi) \partial_d \varphi^b(\xi) = g_{cd}(\xi) ,$$

where φ is an orientation preserving diffeomorphism between parameter domains.

Specifying ϕ_0 and ψ_0 suitably, one obtains the most familiar form [4–6] of the action

$$S(X, g_{ab}) = \int_D d^2 \xi \sqrt{g} \left(\frac{1}{2} g^{ab} \partial_a X \cdot \partial_b X + \lambda_0 g^{ab} \sum_i D_a n_i \cdot D_b n_i + \mu_0 \right) . \quad (16)$$

To obtain a discretized form of S , we use the general idea [1] to consider piecewise linear surfaces defined as maps from the vertices of an (oriented) triangulation of D into R^d , and to assign to each such triangulation T the

metric that assigns equal length a to all links and is Euclidean on each triangle. This leads us to set [1,7]

$$S_T(X) = \sum_{(ij)} [\phi(|x_i - x_j|) + \lambda\psi(\theta_{(ij)})] + \mu|T|, \quad (17)$$

where the sum is over nearest-neighbouring pairs of vertices (i.e. links) in T , x_i is the image of the vertex i in R^d , $\theta_{(ij)}$ is the angle between the normals to the triangles sharing (ij) in the three-dimensional space spanned by the two triangles, and $|T|$ is the number of triangles in T .

The discretized model is then defined in terms of its *loop correlation functions*

$$G_{\mu,\lambda}(\gamma_1, \dots, \gamma_n) = \sum_{T \in \mathcal{T}(m_1, \dots, m_n)} \int \prod_{i \in T \setminus \partial T} dx_i \exp[-S_T(X)], \quad (18)$$

where $\gamma_1, \dots, \gamma_n$ are polygonal loops in R^d , having $m_1, \dots, m_n$ vertices, respectively, and $\mathcal{T}(m_1, \dots, m_n)$ denotes the set of abstract triangulations with n boundary components, having $m_1, \dots, m_n$ vertices, respectively, that are assumed to be mapped onto $\gamma_1, \dots, \gamma_n$ in the obvious way, and, of course, integration is over interior vertices only. Interpreting the vertices in T as locations of molecules and the first term in the action as representing pair interactions between nearest-neighbouring molecules, the randomness of T means that a given molecule does not have fixed neighbours. Therefore the model is usually referred to as a *fluid membrane* model as opposed to *crystalline membranes* where T is fixed. Both types of models have a broad variety of potential application ranging from biophysics to high-energy physics [8].

It is important to note first that the triangulations T have to be topologically constrained, since otherwise the sum in (18) is manifestly divergent [1]. Thus we assume in the following that T is homeomorphic to a sphere with n holes. Under similar assumptions on the functions ϕ and ψ as in the previous section one shows [1] that the loop correlation functions $G_{\mu,\lambda}$ are well defined and finite in a convex domain $\mathcal{B}$ in the (μ, λ) -plane, independent of the number of loops and their shape.

The boundary of $\mathcal{B}$ defines the critical line $\mu = \mu_0(\lambda)$ above which the loop correlation functions are finite. To discuss the critical behavior, it is useful to introduce the following three quantities: the *susceptibility*

$$\chi(\mu, \lambda) = \int_{R^d} dx G_{\mu,\lambda}(\gamma_0, \gamma_x),$$

where γ_x denotes the degenerate loop consisting of the single point x . The *mass gap*

$$m(\mu, \lambda) = - \lim_{|x| \rightarrow \infty} \frac{1}{|x|} \log G_{\mu, \lambda}(\gamma_0, \gamma_x) ,$$

and the *string tension*

$$\tau(\mu, \lambda) = - \lim_{R \rightarrow \infty} \frac{1}{R} \log G_{\mu, \lambda}(\gamma_{R \times R}) ,$$

where $\gamma_{R \times R}$ is a square loop with sides of length R and with a number of equidistant vertices proportional to R .

Of course, the domain $\mathcal{B}$ and the functions $G_{\mu, \gamma}$ depend on the functions ϕ and ψ , but it is expected that the critical behavior of the loop correlation functions is essentially independent of the details of those functions. As it turns out, however, the concrete analysis of the model is for many purposes most conveniently carried out for the discretized version of (16),

$$S_T(X) = \sum_{(ij)} \left(\frac{1}{2} |x_i - x_j|^2 + \lambda(1 - \cos \theta_{(ij)}) \right) + \mu |T| . \quad (19)$$

Thus, in this case, the extrinsic curvature contribution from the link shared by the triangles Δ, Δ' is simply

$$\frac{1}{2} \lambda (n_\Delta - n_{\Delta'})^2 = \lambda (1 - n_\Delta \cdot n_{\Delta'}) ,$$

where n_Δ is the unit normal vector to Δ in the space spanned by Δ and Δ' , and similarly for $n_{\Delta'}$. The results alluded to in the following mostly refer to this particular model.

For $\lambda = 0$ there is ample evidence [1] that the scaling limit is a free scalar field. To explain this statement in more detail, we first recall that a necessary condition for the existence of a scaling limit is that $m(\mu, \lambda) \rightarrow 0$ as $\mu \rightarrow \mu_0(\lambda)$, and in this case the physical mass m_0 and the scaling of μ (for fixed λ) is defined by

$$m(\mu, \lambda) = m_0 a . \quad (20)$$

Moreover, for a scaling limit of genuine continuum surfaces one would expect that $\tau(\mu, \lambda) \rightarrow 0$ as $\mu \rightarrow \mu_0(\lambda)$ such that the physical string tension τ_0 , defined by

$$\tau(\mu(a), \lambda) = \tau_0 a^2 , \quad (21)$$

is finite. It turns out [1], however, for $\lambda = 0$, that $\tau(\mu, 0)$ approaches a positive constant as $\mu \rightarrow \mu_0(0)$, implying an infinite continuum string tension. As a consequence, the dominant surfaces are collapsed ones with spiky outgrowths, effectively branched polymer-like structures. Furthermore, the scaling limit of the two-point function $G_{\mu,0}(\gamma_0, \gamma_x)$, defined as in Eq. (12), exists and is again proportional to the free scalar propagator. A thorough discussion of these results and in particular the role of branched polymers can be found in Ref. [1].

The question arises whether this so-called branched polymer phase or crumpled phase of the model persists for all values of λ , as was shown for the random walk model of the previous section. This is still an open problem, at least from a rigorous point of view within the setup sketched above. However, the framework allows for efficient numerical simulations [1]. Most of these seem to indicate the existence of a critical value $\lambda_c > 0$ at which both the string tension and the mass scale tend to zero, and the possibility of the existence of a trajectory approaching $(\mu_0(\lambda_c), \lambda_c)$ along which the ratio m^2/τ stays constant, as demanded by (20) and (21), is not excluded.

On the other hand, perturbative expansions in $1/\lambda$ have been carried out [9–13]. To lowest order it is found that the β -function is negative for small $1/\lambda$, and it is conceivable that there is no further zero, but this has so far not been established.

Of course, it is of considerable interest to decide firmly which of the two scenarios occurs, and, in case of the latter, whether it is possible by fine tuning λ to ∞ to obtain a non-trivial scaling limit corresponding to (14). This is a problem left for future research.

4 Conclusions

We have given an outline of results on the scaling properties of a class of discretized models of fluctuating paths and surfaces with rigidity. The discretizations considered here are, of course, by no means unique, and a variety of others have been considered. In particular, hypercubic models have played an important role and yield results complementary to the ones discussed in the text. We refer to the literature [1] for details.

Clearly, this article is by no means intended to be an exhaustive overview of the theory of random surfaces, which is a field in rapid development. In particular, the whole subject of crystalline membranes, which have been investigated by both analytical and numerical methods in recent years, has

been left out, and so have self-avoiding membranes. We refer the reader to a recent review [8] for an account of its fascinating aspects.

Acknowledgments

This research was supported in part by MaPhySto, funded by the Danish National Research Foundation.

References

- [1] J. Ambjørn, B. Durhuus, and T. Jonsson, *Quantum Geometry. A Statistical Field Theory Approach* (Cambridge University Press, 1997).
- [2] L. Brink, P. DiVecchia, and P. Howe, *Nucl. Phys. B* **118**, 76 (1977).
- [3] J. Ambjørn, B. Durhuus, and T. Jonsson, *J. Phys. A* **21**, 981 (1988).
- [4] L. Brink, P. DiVecchia, and P. Howe, *Phys. Lett. B* **65**, 471 (1976).
- [5] P.B. Canham, *J. Theor. Biol.* **26**, 61 (1970).
- [6] W. Helfrich, *Z. f. Naturforsch. C* **28**, 693 (1973).
- [7] J. Ambjørn, B. Durhuus, and T. Jonsson, *Phys. Rev. Lett.* **58**, 2619 (1987).
- [8] M.J. Bowick and A. Travesset, eprint: cond-mat/0002038.
- [9] D. Förster, *Phys. Lett. A* **114**, 115 (1986).
- [10] W. Helfrich, *J. Physique* **47**, 321 (1986).
- [11] H. Kleinert, *Phys. Lett. A* **114**, 263 (1986).
- [12] H. Kleinert, *Phys. Lett. B* **174**, 335 (1986).
- [13] A.M. Polyakov, *Nucl. Phys. B* **268**, 406 (1986).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

Part VI

Gravitation, Cosmology, and Astrophysics

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

STRING, SCALAR FIELD, AND TORSION INTERACTIONS

R.T. HAMMOND

*North Dakota State University, Physics Department,
Fargo, North Dakota 58105, USA
E-mail: rich_hammond@ndsu.nodak.edu*

It is shown how the scalar field is a necessary part of torsion. Its effects of being incorporated into the torsion tensor are examined.

1 Introduction

Before a great dam collapses, small rivulets form deep holes within its mass, allowing water to trickle through, and then gather enough momentum to erode the mighty structure and reduce it to rubble. Like the water and air separated by the dam, general relativity and quantum mechanics have been kept apart, and despite the countless years and desperate attempts to bring them together, the great wall of ignorance, stronger than any concrete, has held as firm barrier between the two. In recent years, however, we may have witnessed the first trickle cutting through the dam. These tiny rivulets are strings, and if they succeed in gathering sufficient momentum, they may carve out the new river bed for a quantum theory of gravity. It is a pleasure to write this article for the occasion of Hagen Kleinert's sixtieth birthday, whose prolific work in so many areas has surely worn down the barrier, and has established many inroads to developing a more complete picture of physics.

When it was discovered that gravity was incorporated in string theory the landscape began to change in several ways. The string theorists welcomed its honored guest with open arms, and argued that its natural inclusion indicates that string theory is a physical theory, and not simply ornate mathematics. However, string theory did not only invite the gravitational field on stage,

but an antisymmetric and a scalar field made their appearance as well, and could not be denied by their audience. Scherk and Schwarz [1] made a major breakthrough, creating the first rivulet, showing that the low energy effective Lagrangian of string theory is simply the curvature scalar R of space-time with torsion. This amazing result used more than string to tie together hitherto disparate theories. It gave a rich and compelling low energy Lagrangian that announced the presence of gravity quite effectively, while giving a geometrical interpretation to all three fields.

These results were planted decades ago, but their full harvest has still not been reaped. Many general relativists have remained unwilling to consider gravitation with torsion. String theory, aside the ground breaking work of Kibble [2], Hehl [3], and others, who used torsion in formulating a local Poincaré gauge theory, was unable to provide a rich enough soil to host the full growth of these ideas. In recent years, though, the combined theoretical predictions of string theory and general relativity with torsion have produced a synergism that has breathed new life into both the foundation and physical interpretations of the new fields.

One of the recent advances from the gravitational side showed that the torsion arises from intrinsic spin. This was not only known from the local Poincaré approach, but was also verified by explicit calculations of the equations of motion [4]. In fact, the use of general relativity provided for the detailed interaction of intrinsic spin with gravity or spin. In developing the phenomenological source of spin, it was found that an intrinsic vector, ξ^μ , had to be introduced. At first this vector posed as a mysterious stranger, and only pointed to general hints at physical interpretations. For example, it indicated that the source could not be represented by a point particle but must be at least one-dimensional, but the issue was not resolved until the incorporation of a string source [5].

Not only has the antisymmetric field been given a physical interpretation and home, but the scalar field has been adopted as well. Under my original program torsion was defined as $S_{\mu\nu\sigma} = \psi_{[\mu\nu,\sigma]}$. However, when the source was taken to be given by the Dirac Lagrangian, several more interesting results sprouted. One of these is that the low energy limit of the resulting Dirac equation gave the same interaction as the phenomenological result, and although this shed no light on ξ^μ , it justified its introduction. Another result was the germination of the scalar field. With Dirac coupling, if torsion of the kind given above is included, then a scalar field must be present. In fact, it may be interpreted that the scalar field must arise due to the non-conservation

of the axial current. In order to incorporate the scalar field in the original seeding, it was shown that the generalization of the torsion to [6]

$$S_{\mu\nu\sigma} = \psi_{[\mu\nu],\sigma} + \frac{2}{3}\epsilon_{\mu\nu\sigma\gamma}\phi^{\cdot\gamma} \quad (1)$$

yields the correct self-consistent results.

With this generalization it was found that the curvature scalar R of gravitation with torsion not only gives the correct low energy effective limit of string theory, it provides a physical interpretation of the origin of the scalar field [7]. It turned out that the source is related to the pseudoscalar invariant, a result that gives a definite direction in which to search for scalar field effects. There was also an improvement over the original results of Ref. [1]: in the current case there are no longer undefined fields, and the covariant derivative of the metric tensor vanishes, which insures constancy of length upon parallel transport.

The ties that bind string theory to gravitation discussed so far, are the antisymmetric field and its definition as torsion, the resulting equivalence of the low energy effective string theory Lagrangian to R , and the fact that gravitation with this torsion would not allow zero-dimensional sources. In recent years the final tie was to formulate gravitation with a string source [8]. This not only provided a natural origin for the antisymmetric field, and gave a very satisfying interpretation to the conservation laws, it provided a physical interpretation of ξ^μ [9]. Shedding its cloak of obscurity, this vector was finally recognized as the tangent vector to the string (see Ref. [9]). In fact, if strings had not been invented already, this line of growth would have led naturally to their doorstep.

Like the nourishment a flood can provide to a barren soil, these new fields plant the possibilities of providing new effects that may be measured. Below, I will establish an interaction between the scalar field and the spin of a particle. This is a new prediction and grows from a very simple foundation. Jumping ahead, the result I will show is that the interaction is given by the scalar product of the spin of a particle and the scalar field gradient.

2 Torsion and the Action

The variation principle is taken to be

$$\delta(I_g + I) = 0, \quad (2)$$

where

$$I_g = \int \sqrt{-g} \frac{R}{2k} d^4x \quad (3)$$

and I is the material action. The curvature scalar R is that of U_4 space-time. The energy momentum tensor is defined by

$$\delta I_m = \frac{1}{2} \int d^4x \sqrt{-g} T^{\mu\nu} \delta \phi_{\mu\nu}, \quad (4)$$

where $\phi_{\mu\nu} = g_{\mu\nu} + \psi_{\mu\nu}$, which gives an antisymmetric energy momentum tensor that has been discussed in Ref. [9]. The field equations are obtained by taking variations with respect to $\phi_{\mu\nu}$, and are

$$G^{\mu\nu} - 3S^{\mu\nu\sigma}{}_{;\sigma} - 2S^\mu{}_{\alpha\beta} S^{\nu\alpha\beta} = kT^{\mu\nu}, \quad (5)$$

where

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \quad (6)$$

and $R^{\mu\nu}$ is the (asymmetric) Ricci tensor in U_4 space-time. The field equations for torsion are given by

$$S^{\mu\nu\sigma}{}_{;\sigma} = -kj^{\mu\nu}, \quad (7)$$

where $j^{\mu\nu} \equiv (1/2)T^{[\mu\nu]}$.

3 The Source, and the Essential Scalar Field

The equations defined above carry little content until $T^{\mu\nu}$ is specified. One should not be swayed by (4), which carries as much content as vacuum: it is merely a definition. All things that are measured come from the equations of motion, or the interaction energies predicted by the theory, but none of these can be obtained until an explicit form of $T^{\mu\nu}$ is given. This is part of the beauty and simplicity of the Einstein equations, or their generalizations, when they are derived from an action principle.

For now, we would like to investigate the effects on the scalar field of using the Dirac Lagrangian as the source. Thus, assume that the material action is given by

$$L = -\frac{i}{2} [(D_a \bar{\psi}) \gamma^a \psi - \bar{\psi} \gamma^a D_a \psi - 2im \bar{\psi} \psi], \quad (8)$$

with

$$D_a \psi = \psi_{,a} - \frac{1}{4} \Gamma_{abc} \gamma^b \gamma^c \psi, \quad (9)$$

where Γ_{abc} is the non-holonomic spin connection and contains (1). Thus, we have the variational principle,

$$\delta \int e \left(\frac{R}{2k} + L \right) d^4x = 0, \quad (10)$$

where e is $\sqrt{-g}$.

A comforting surprise reveals itself when R is expanded into ${}^oR$, the curvature scalar of Riemann space-time (no torsion), plus everything else. Ignoring terms that give no contributions to any of the field equations, the action becomes

$$\delta \int \sqrt{-g} d^4x \left(\frac{{}^oR - H_{\mu\nu\sigma} H^{\mu\nu\sigma} + 4\phi_{,\sigma} \phi^{,\sigma}}{2k} + L \right) = 0. \quad (11)$$

If we take $k = 8\pi G e^{2\phi}$, which does not affect the derivation of (11), the low energy effective Lagrangian of string theory results. Our goal is to investigate interactions and possible measurements of the scalar field, so let us consider the weak ϕ limit in the Einstein frame, which yields

$$\delta \int \sqrt{-g} d^4x \left(\frac{{}^oR - S_{\mu\nu\sigma} S^{\mu\nu\sigma} - \frac{1}{2} \phi^{,\sigma} \phi_{,\sigma}}{2k_0} + L \right) = 0, \quad (12)$$

where $k_0 = 8\pi G$ is the standard constant.

This can be generalized somewhat by introducing non-minimal coupling, in which case the undetermined coupling constant κ is introduced (we get minimal coupling results by letting $\kappa = 1$) [10]. Proceeding in this most general case, the Dirac equation, which is obtained by performing variations with respect to $\bar{\psi}$, is

$$\gamma^a \left(\psi_{,a} - \frac{\kappa}{4} S_{abc} \gamma^b \gamma^c \right) \psi + im\psi = 0, \quad (13)$$

where

$$S_{abc} = e_a^\alpha e_b^\beta e_c^\gamma S_{\alpha\beta\gamma} \quad (14)$$

is the torsion in the anholonomic basis.

One result immediately shows itself from (13): the scalar field interactions with the Dirac particle. Since our current mission is to investigate the effects

of the scalar field, we will ignore the contribution from $H_{\alpha\beta\gamma}$. The Dirac equation may be separated, in the low energy limit, in terms of its “large” and “small” two-spinors, as usual. Calling Ψ the large two-component spinor, the Dirac equation yields

$$i\hbar \frac{\partial}{\partial t} \Psi = + \frac{p^2}{2m} \Psi + \frac{\sqrt{24}\kappa c}{2} \mathbf{S} \cdot \nabla \phi, \quad (15)$$

where $\mathbf{S}$ is the spin, and cgs units have been restored.

This is part of the main result, and gives the explicit form of the interaction with the scalar field. It shows that the interaction is given by the scalar product of the intrinsic spin and the gradient of the scalar field, but there is more to come. The field equation for the scalar field, obtained by considering variations with respect to ϕ , may be put in the form

$$\square \phi = -16\pi\sqrt{6}i \frac{Gm}{c^2} \bar{\psi} \gamma_5 \psi. \quad (16)$$

4 Conclusion

The last two equations, (15) and (16), constitute the main result. In words, Eq. (16) states that the scalar field arises from the pseudoscalar invariant, a quantity that vanishes for free particles but not in general. Once this is used in order to find ϕ , Eq. (15) can be taken to find the interaction. Thus, I must leave you hanging, a bit. The next steps in this research program are to evaluate Eq. (16) for any systems that seem propitious, and then to calculate the interaction strength. These results will be used to look for such an interaction, or apply the results to bound the non-minimal coupling constant.

References

- [1] J. Scherk and J.H. Schwarz, *Phys. Lett. B* **52**, 347 (1974).
- [2] T.W.B. Kibble, *J. Math. Phys.* **2**, 212 (1961).
- [3] F.W. Hehl, P. von der Heyde, G.D. Kerlick, and J.M. Nestor, *Rev. Mod. Phys.* **15**, 393 (1976).
- [4] R.T. Hammond, *Gen. Rel. Grav.* **26**, 247 (1994).

- [5] R.T. Hammond, *Gen. Rel. Grav.* **28**, 749 (1996).
- [6] R.T. Hammond, *Gen. Rel. Grav.* **30**, 1803 (1998).
- [7] R.T. Hammond, *Class. Quantum Grav.* **13**, L73 (1996).
- [8] R.T. Hammond, *Gen. Rel. Grav.* **31**, 347 (1999).
- [9] R.T. Hammond, *Gen. Rel. Grav.* **31**, 233 (1999).
- [10] R.T. Hammond, *Phys. Rev. D* **52**, 6918 (1995).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

METRIC AND CONNECTION: KINEMATIC AND DYNAMIC SOLUTIONS OF THE SPACE PROBLEM

H.-H. VON BORZESZKOWSKI

*Institute for Theoretical Physics, Technical University Berlin,
Hardenbergstraße 36, D-10623 Berlin, Germany
E-mail: borzeszk@itp.physik.tu-berlin.de*

H.-J. TREDER

Rosa-Luxemburg-Str. 17a, D-14482 Potsdam, Germany

We compare metric theories to bi-metric theories, to theories with teleparallelism, and to metric-affine (Einstein-Cartan) theories of gravitation in order to discuss the implications arising from the equivalence principle for the related space and energy problems.

1 Introduction

All classical local geometric theories of gravitation are based on the assumption that the space-time manifold is primarily or can secondarily be endowed with a Lorentz metric, i.e. with a symmetric tensor field of signature $(+, -, -, -)$. This is even true for purely affine theories such as the Einstein-Schrödinger theory [1,2], where the geometry is specified by a connection as basic variable, while the metric is only of secondary importance. In the final analysis, a metric structure is always needed in order to arrive at a theory which is physically interpretable. Only such a structure allows us to introduce the notions of spatial distance, time interval, angle, and relativistic velocity.

As a consequence of the principle of equivalence, however, the metric has to be related to the gravitational field. In contrast to the special-relativistic approach, it cannot be introduced *a priori*, but has to be specified in accordance with the other unrenouncable geometric structure, the connection. In

other words, a theory satisfying the principle of equivalence has to be established in such a way that it solves the so-called space problem (see Refs. [3,4]). First of all, this means to find a gravitational dynamics that does not clash with the "kinematic" conditions ensuring the compatibility between metric and connection.

As far as the kinematic part of the Weyl-Cartan space problem is concerned, it was shown by Schrödinger [2], that the relationship

$$\nabla_l g_{ik} = 0 \quad (1)$$

between the metric g_{ik} and the connection Γ_{kl}^i defined by ∇_l is a sufficient condition for their compatibility.^a

In view of the dynamic aspect of the space problem, there exists a further argument in favor of Eq. (1). It says that the coupling of spinorial matter to geometric structures which are related to gravitation requires to assume the validity of Eq. (1) [6].^b Indeed, it follows from (1) that

$$\Gamma_{kl}^i = \{^i_{kl}\} + \Gamma_{[kl]}^i + g^{im} g_{kr} \Gamma_{[lm]}^r + g^{im} g_{lr} \Gamma_{[km]}^r = \{^i_{kl}\} + K_{kl}^i, \quad (2)$$

where $\{^i_{kl}\}$ is the Christoffel connection and K_{kl}^i the contorsion which is antisymmetric in the first two indices, $K_{ikl} = -K_{kil}$. In the anholonomic version, Eq. (2) reads

$$\Lambda^A_{Bl} = \gamma^A_{Bl} + K^A_{Bl}, \quad (3)$$

$$\gamma_{ikl} = -\gamma_{kil}, K_{ikl} = -K_{kil}, \quad (4)$$

where γ^A_{Bl} are the Ricci rotation coefficients, $\gamma^A_{Bl} = h^A_k h_B{}^l{}_{|l}{}^k$, and the last expression in Eq. (3) is defined as $K^A_{Bl} = h^A_i h_B{}^k K^i_{kl}$; both are antisymmetric in the first two indices. As a consequence, the anholonomic components of the internal connection are antisymmetric in the first two indices, too:

$$\Lambda_{ikl} = -\Lambda_{kil}. \quad (5)$$

Exploiting the one-to-two correspondence of the Lorentz group $O(3,1)$ to the unimodular group $SL(2,C)$, this recovers the usual spinor formalism, where

^aFor physical reasons, Einstein [5] considered Eq. (1) even as a necessary condition.

^bBy means of another line of arguments, this was also shown in Refs. [7,8]. Even more, it was demonstrated there how the usual spinor formalism can be generalized for the case that the nonmetricity is of the Weyl form. We shall confine ourselves to spaces with a vanishing nonmetricity. But all our arguments given in this paper can be generalized to spaces with a nonvanishing Weyl nonmetricity, i.e. to semi-metric spaces.

$\gamma_{\alpha\beta}$ and $\gamma_{\dot{\alpha}\dot{\beta}}$ are the "metrics" in the spin spaces:

$$\gamma_{\alpha\beta} = -\gamma_{\beta\alpha}, \gamma_{\dot{\alpha}\dot{\beta}} = -\gamma_{\dot{\beta}\dot{\alpha}}, \gamma_{\alpha\beta,l} = \gamma_{\dot{\alpha}\dot{\beta},l} = 0, \Lambda_{\alpha\beta l} = \Lambda_{\beta\alpha l}, \Lambda_{\dot{\alpha}\dot{\beta}l} = \Lambda_{\dot{\beta}\dot{\alpha}l}. \quad (6)$$

These conditions, here justified with regard to a possible gravitational dynamics, must of course be recovered or, at least, not be violated by the full gravitational dynamics.

2 Metric Theories

In metric theories the space problem is solved *a priori* by assuming a Riemannian space-time, i.e. by assuming a four-dimensional manifold with a Lorentz metric, where the connection given by the metric and their first derivatives satisfies (1) identically. In this case, the field equations for the metric cannot conflict with (1). Then the free gravitational Lagrange density is formed from g_{ik} and $\partial_l g_{ik}$

$$\mathbf{L}_f = \mathbf{L}_f(g_{ik}, \partial_l g_{ik}, \partial_l \partial_m g_{ik}), \quad (7)$$

while the matter Lagrange density depends on both the gravitational and matter variables, g_{ik} and $\phi^{(A)}$; imposing, as an implication of the principle of equivalence, the principle of minimal coupling, one generally has

$$\mathbf{L}_m = \mathbf{L}_m(g_{ik}, \partial_l g_{ik}, \phi^{(A)}, \partial_l \phi^{(A)}). \quad (8)$$

The Euler variation of the action integral

$$I = \int (\mathbf{L}_f + 2\kappa \mathbf{L}_m) d^4x, \quad (9)$$

by the matter field provides the matter field equations

$$\frac{\delta \mathbf{L}_m}{\delta \phi^{(A)}} = 0 \quad (10)$$

and by the metric the gravitational equations

$$G_{ik} = -\kappa P_{ik}, \quad (11)$$

where the Einstein tensor is defined by

$$G_{ik} = \frac{\delta \mathbf{L}_f}{\delta g^{ik}} \quad (12)$$

and the metric energy-momentum tensor of matter reads

$$P_{ik} = \frac{2}{\sqrt{-g}} \frac{\delta \mathbf{L}_m}{\delta g^{ik}}. \quad (13)$$

Furthermore, assuming that all field equations are satisfied, one obtains [9],

$$\delta^* \mathbf{L} + (\mathbf{L} \xi^m)_{,m} = 0 \quad (14)$$

with $\delta x^i := \xi^i$, the Lie differential $\delta^* g^{ik}$ of g^{ik} ,

$$\delta^* g^{ik} = g^{sk} \xi^i_{,s} + g^{is} \xi^k_{,s} \quad (15)$$

and the corresponding Lie differential $\delta^* \phi^{(A)}$ of $\phi^{(A)}$. For $\xi^i = \alpha^i = \text{const.}$ Eq. (14) yields the Noether identity

$$\left(P_i^k + \frac{1}{2\kappa} t_i^k \right)_{,k} \cong \nabla_k P_i^k = 0, \quad (16)$$

where t_i^k is the gravitational energy-momentum pseudo-tensor. This equation is equal to the contracted Bianchi identities when calculated via the field equations (16).

Following Lorentz [10], one can read the field equations (11) as the statement that the total metric energy-momentum density of matter and gravity is equal to zero. That such a reading of (11) has a physical meaning was shown for Einstein's GRT, where $\mathbf{L}_f = \sqrt{-g}R$, and for fourth-order gravitational equations, where $\mathbf{L}_f = \sqrt{-g}(R + \alpha R_{ik} R^{ik} + \beta R^2)$ [11,12]. For the recent development of the topic see Ref. [13] and the papers quoted therein.

Purely metric theories are insofar physically completely satisfying as they guarantee the following three points: (i) they solve the space problem; (ii) they are in agreement with the principles of equivalence and general relativity; (iii) their "conservation" laws given by Noether's theorem follow, via the field equations, automatically from the Bianchi identities. Otherwise, and that is their disadvantage, as a matter of these principles, they do not allow for the existence of a genuine law of energy-momentum conservation.

3 Bi-Metric Theories

Bi-metric theories have in addition to the Riemann-Einstein metric g_{ik} , a second metric $\bar{g}_{ik}$. Such theories go back to ideas of Rosen [14], Band [15], and Papapetrou [16], but their satisfying elaboration was only given by

Kohler [17,18].^c In this theory, the Weyl-Cartan space problem is solved in the same manner as in the uni-metric theories discussed above (both metrics g_{ik} and $\bar{g}_{ik}$ are presupposed to satisfy Eq. (1)). And, even more, since the second metric $\bar{g}_{ik}$ is assumed to be a pseudo-Euclidean one, one has 10 Killing vectors such that also the Helmholtz-Lie space problem is solved.

Corresponding to the two metrics, there exist the two (Riemann-Christoffel) connections Γ_{kl}^i and $\bar{\Gamma}_{kl}^i$, where the latter vanishes in pseudo-Cartesian coordinates. Their difference is a tensor ρ_{kl}^i describing the gravitational field such that the connection Γ_{kl}^i is the sum of the inertial field $\bar{\Gamma}_{kl}^i$ and the gravitational field ρ_{kl}^i :

$$\Gamma_{kl}^i = \bar{\Gamma}_{kl}^i + \rho_{kl}^i. \quad (17)$$

Kohler specified the free Lagrangian as

$$\mathbf{L}_\kappa = \mathbf{L}_f(g_{ik}, \partial_l g_{ik}, \bar{g}_{ik}, \partial_l \bar{g}_{ik}) + \mathbf{L}_m(g_{ik}, \partial_l g_{ik}, \phi^{(A)}, \partial_l \phi^{(A)}), \quad (18)$$

such that, as in the above uni-metric theories, the metric-energy-momentum tensor fulfills again the dynamical equation with respect to Γ_{kl}^i :

$$\nabla_k P_i^k = 0. \quad (19)$$

But now, due to the existence of the second metric, this can be rewritten as

$$\bar{\nabla}_k \left(\sqrt{\frac{\bar{g}}{g}} (P_i^k + t_i^k) \right) = 0, \quad (20)$$

where $\bar{\nabla}$ is the covariant derivative with respect to $\bar{\Gamma}_{kl}^i$, g is the determinant of g_{ik} , and $\bar{g}$ the determinant of $\bar{g}_{ik}$. In contrast to t_i^k in Eq. (16), the expression t_i^k arising in Eq. (20) is a tensor under general coordinate transformations.

The variation of Eq. (18) by g_{ik} provides 10 field equations, while its variation by $\bar{g}_{ik}$ leads to four additional equations. The latter ones can be interpreted as coordinate conditions soldering the two metrics; they are equivalent to Eq. (19). Thus, in this theory the dynamical Eq. (19), representing a conservation law according to Eq. (20), is no implication of the Bianchi identities and the field equations. This is the price one has to pay for the formulation of differential laws of conservation [19]. Except for this point, one finds the same situation as in the uni-metric theories considered above.

^cSee also Refs. [19,20].

4 Theories with Teleparallelism (Einstein-Mayer-Type Theories)

Theories with teleparallelism consider the tetrad field $h^A_i(x^l)$ given by the coframe field $h^A = h^A_i dx^i$ as basic quantity; the metric is a secondary concept defined as

$$g_{ik} = \eta_{AB} h^A_i h^B_k, \quad (21)$$

with $\eta = \text{diag}(1, -1, -1, -1)$. Since the h^A_i are assumed to be fixed by the gravitational field equations, they define a teleparallelism in the Riemannian space. Thus the Weyl-Cartan space problem is solved in the same way as in GRT.

Let us start now from a Weitzenboeck Lagrange density which is a scalar density with respect to coordinate transformations but not invariant under local Lorentz transformations rule. One finds here a restricted invariance, namely an invariance with respect to global Lorentz transformations rule. To discuss the points in this paper under consideration we shall confine ourselves to the Einstein-Mayer class of Lagrangians [21-26]

$$\mathbf{L}_{EM} = \sqrt{-g}R + ahF_{Aik}F^{Aik} + bh\phi_A\phi^A, \quad (22)$$

where $h = \det(h^A_i) = \sqrt{-g}$, $R = g^{ik}R_{ik}$ is the Ricci scalar, a and b are numerical constants, the ϕ_A are defined as $\phi_A = h^i_A \gamma^m_{im}$, and the F_{Aik} are Cartan's anholonomy objects

$$F_{Aik} = h_{Ai,k} - h_{Ak,i} = h_A^l (\gamma_{lik} - \gamma_{lki}), \quad (23)$$

with the Ricci rotation coefficients $\gamma_{ikl} = h^A_i h_{Ak;l}$. The Lagrangian (22) lies on the basis of Møller's tetrad theory of gravitation [27].^d (From the viewpoint of a unified gravito-electromagnetic theory in the sense of Einstein's program of 1929, it is also discussed in Ref. [29]; however, in contrast to Einstein's approach, there, like in Møller's theory, an additional matter Lagrangian is introduced.) For historical reasons, we shall call the Lagrangian (22) "Einstein-Mayer Lagrangian", while the tetrad equivalent of the Einstein-Hilbert Lagrangian will be called "Møller Lagrangian". (For other metric-teleparallel theories, see, e.g. Kopczyński [30], Hayashi and Shirafuji [31]).

Now, in the anholonomic (Einstein-Cartan) representation of the space-time structure the reference systems, i.e. the "tetrads" $h^A_i(x^l)$, are the grav-

^dIn order to remove singularities from the theory of gravitation, later Møller [28] also introduced Lagrangians quadratic in the Weitzenboeck invariants.

itational field variables. From the standpoint of the theories here under consideration, Einstein's GRT is a degenerate case, now following from (22) for $a = b = 0$.

In Refs. [26,27] one assumes for matter a Lagrange density depending only on the tetrads via the algebraic combination (21). This leads to the field equations

$$\frac{1}{h} \frac{\delta \mathbf{L}_{EM}}{\delta h^{Ai}} h^A{}_k = -\kappa T_{ik}, \quad (24)$$

with

$$\frac{\delta \mathbf{L}_{EM}}{\delta h^{Ai}} h^A{}_k - \frac{\delta \mathbf{L}_{EM}}{\delta h^{Ak}} h^A{}_i = 0. \quad (25)$$

The 10+6 equations (24) and (25) determine the 16 components of $h^A{}_i$ up to constant Lorentz rotations of the tetrads. Together with the Noether identities

$$H^i{}_{k,i} = F^I{}_{k;i} - F^{il} \gamma_{ilk}, \quad (26)$$

where $H_{ik} \equiv E_{ik} + \Theta_{(ik)}$ and $F_{ik} = \Theta_{[ik]}$, and together with Eq. (25) this leads again to the dynamical equations (16).

As argued in Refs. [32,33], in a Riemannian theory with Einstein-Cartan teleparallelism one finds quite a satisfying situation for the gravitational energy. One has a Lagrangian of canonical structure that is coordinate-covariant but, like Møller's Lagrangian, not covariant with respect to local Lorentz transformations rule. Thus the Lagrangian leads to field equations which do not satisfy the general principle of relativity. These are equations fixing the 16 components of the tetrads $h^A{}_i$ instead of their 10 combinations $g_{ik} = h^A{}_i h_{Ak}$, and the Hamiltonian $\mathbf{H}$ is given by the 00-component of the energy-momentum complex (which is a tensor with respect to the group of global Lorentz transformations rule lying on the basis of this theory).^e

Similarly to the Kohler case, the dynamical equations of the theory with teleparallelism are not simply an implication of the Bianchi identities and the field equations, now given by Eq. (24). They result from the additional conditions (25) which specify the reference tetrads instead of the coordinates. Again, this is the price that one has to pay for having differential laws of

^eIn Refs. [32,33] such theories are considered as models realizing the local version of the Mach principle, according to which the cosmic matter distribution fixes the reference system.

conservation. Or better, this is a part of the price, for, the other part is that one has even to sacrifice the local Lorentz covariance and, thus, the general principle of relativity.

5 Einstein-Cartan Theories

The more general case is considered in Ref. [34], where the matter Lagrange density depends independently on the 16 tetrad and the 64 connection components. Such a theory with teleparallelism can be considered as a constrained Einstein-Cartan theory which generally is formulated in spaces with nonvanishing curvature and torsion. Therefore, Einstein-Cartan theories can be regarded as a straightforward generalization of the Einstein-Mayer-type theories. Thus, let us turn to Einstein-Cartan theories.

As to the general principle of relativity, such theories resemble the purely metric theories: they are covariant under local Lorentz rule and general coordinate transformations. A great similarity is also found as to the energy-momentum problem. Indeed, starting from a Lagrange density of the form

$$\mathbf{L}_{EC} = \mathbf{L}_f(h^A_i, \Lambda^{AB}_i) + \kappa \mathbf{L}_m(h^A_i, \Lambda^{AB}_i, \Phi^{(A)}, \partial_l \Phi^{(A)}), \quad (27)$$

where the tetrad and the connection fields, h and Γ , are regarded as independent variables, the Euler variation of $\Phi^{(A)}$, h^A_i , and Γ^{AB}_i yields the matter field equations and the following gravitational equations (h denotes the determinant of h^A_i),

$$-\frac{\delta \mathbf{L}_f}{\delta h^A_i} = \kappa h T_A^i := \kappa \frac{\delta \mathbf{L}_m}{\delta h^A_i} \quad (28)$$

and

$$-\frac{\delta \mathbf{L}_f}{\delta \Lambda^{AB}_i} = \kappa h \tau_{AB}^i := \kappa \frac{\delta \mathbf{L}_m}{\delta \Lambda^{AB}_i}. \quad (29)$$

The source term in Eq. (28), T_A^i , is the canonical energy-momentum tensor, and the source term in Eq. (29), τ_{AB}^i , represents the canonical spin-angular momentum tensor of the matter field. Therefore, as in GRT and its unimetric generalizations, Eq. (28) can be read as the statement that the total energy-momentum density is equal to zero. (However, this statement does not concern the total metric, but the canonical tensor.)

As to the fact that one has the same symmetries as in GRT, there does not exist a differential law of conservation. There exist dynamical equations

which are a generalization of Eq. (19), and, for the class of theories under consideration in Refs. [35-38], it turns out that, provided the field equations are fulfilled, the generalized dynamical equations are again an automatic by-product of the Bianchi identities which hold in Riemann-Cartan space [39].

For spinless matter these generalized dynamical equations reduce to Eq. (19). This implies that spinless point particles move along geodesics of the Riemann-Christoffel connection, but not along autoparallels. Hagen Kleinert [40,41] considers this fact as an objection to the proposed class of Einstein-Cartan theories.

After mentioning a series of similarities between GRT and Einstein-Cartan theories, it should finally be stressed that there is a great difference as to the Weyl-Cartan space problem. The point is that in Einstein-Cartan theories the validity of Eq. (1) cannot kinematically be guaranteed. To solve the space problem now means to solve it by using an appropriate dynamic starting point. That is, one has to look for a Lagrangian such that the solutions of the corresponding gravitational field equations (28) and (29) satisfy Eq. (1), i.e. they yield solutions satisfying the constraints $\Lambda_{ABl} = -\Lambda_{BA l}$ (or $\Lambda_{\alpha\beta l} = \Lambda_{\beta\alpha l}$). From this point of view, for example, the Einstein-Hilbert Lagrangian $\sqrt{-g}R$, viewed as a functional of the metric and an arbitrarily generalized connection, has to be excluded from the consideration [42]; it leads to

$$\nabla_l g_{ik} = -2g_{is}\Gamma_{[kl]}^s + \frac{2}{3}(\Gamma_i g_{kl} + \Gamma_k g_{il}) \quad (30)$$

with the torsion vector $\Gamma_i = \Gamma_{is}^s$.

6 Conclusion

Physically interpretable theories need a metric. However, due to the principle of equivalence, the metric cannot be presupposed a priori. Taking this fact into consideration and rejecting any restriction of the principle of general relativity, one is necessarily led to purely metric theories like GRT or to Einstein-Cartan theories. Both suffer from the fact that the only energetic statement one can make consists in that the (metric and canonical, respectively) total energy-momentum density vanishes. As to the space problem, metric theories seem to be less problematic since for them it can be solved kinematically, while for Einstein-Cartan theories it has to be solved dynamically.

References

- [1] A. Einstein, *The Meaning of Relativity*, 4th and 5th ed. (Princeton University Press, Princeton, 1950 and 1955).
- [2] E. Schrödinger, *Space-Time-Structure* (Cambridge University Press, Cambridge, 1950).
- [3] H. Weyl, *Mathematische Analyse des Raumproblems* (Springer, Berlin, 1923).
- [4] E. Cartan, *Jour. d. Math. Pura et Appl.* **2**, 167 (1923); *Œuvres Complètes* (Gauthier-Villars, Paris, 1955), p. 633.
- [5] A. Einstein, *Remark to Weyl*, in: H. Weyl, *Gravitation und Elektrizität*, Gesammelte Abhandlungen, Vol. 2 (Springer, Berlin, 1968), p. 40.
- [6] H.-H. v. Borzeszkowski, and H.-J. Treder, *Gen. Rel. Grav.* **33**, in press (2001).
- [7] K. Hayashi and T. Shirafuji, *Prog. Theor. Phys.* **57**, 302 (1977).
- [8] K. Hayashi, *Phys. Lett. B* **65**, 437 (1976).
- [9] P.G. Bergmann, *General Relativity of Theory*, in: *Encyclopedia of Physics*, Vol. II, Ed. S. Flügge (Springer, Berlin, 1962).
- [10] H.A. Lorentz, *Proc. K. Akad. Wet. Amsterdam* **19**, 751 (1917).
- [11] H.-H. v. Borzeszkowski, H.-J. Treder, and W. Yourgrau, *Ann. Physik (Leipzig)* **35**, 471 (1978).
- [12] W. Yourgrau, H.-H. v. Borzeszkowski, and H.-J. Treder, *Astron. Nachr.* **300**, 57 (1979).
- [13] H. Kleinert and H.-J. Schmidt, eprint: gr-qc/0006074.
- [14] N. Rosen, *Phys. Rev.* **57**, 147 and 150 (1940).
- [15] W. Band, *Phys. Rev.* **61**, 698 (1942).
- [16] A. Papapetrou, *Proc. Irish Acad.* **52**, 11 (1948).
- [17] M. Kohler, *Z. Phys.* **131**, 571 (1952).
- [18] M. Kohler, *Z. Phys.* **134**, 286 and 306 (1953).
- [19] M. v. Laue, *Die Relativitätstheorie*, Zweiter Band: *Die Allgemeine Relativitätstheorie*, 5th ed. (Friedrich Vieweg und Sohn, Braunschweig, 1965).
- [20] H.-H. v. Borzeszkowski, U. Kasper, E. Kreisel, D.-E. Liebscher, and H.-J. Treder, in *Gravitationstheorie und Äquivalenzprinzip*, Ed. H.-J. Treder (Akademie-Verlag, Berlin, 1971).
- [21] A. Einstein, in *Berliner Berichte*^f 1928, p. 219.

^fThis is a colloquial abbreviation of *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*.

- [22] A. Einstein, in *Berliner Berichte* 1929, p. 124.
- [23] A. Einstein and W. Mayer, in *Berliner Berichte* 1931, p. 3.
- [24] A. Einstein, in *Berliner Berichte* 1928, p. 224.
- [25] A. Einstein, in *Berliner Berichte* 1929, p. 2.
- [26] C. Pellegrini, and J. Plebański, *Math.-Fys. Skr. Dan. Vid. Selskab* **2**, No. 4 (1963).
- [27] C. Møller, *Math.-Fys. Skr. Dan. Vid. Selskab* **39**, No. 13 (1978).
- [28] C. Møller, in *Einstein-Centenary 1979*, Ed. H.-J. Treder (Akademie-Verlag, Berlin, 1979).
- [29] H.-J. Treder, *Ann. Phys. (Leipzig)* **35**, 371 (1978).
- [30] W. Kopczyński, *J. Phys. A* **15**, 493 (1982).
- [31] K. Hayashi and T. Shirafuji, *Phys. Rev. D* **19**, 3524 (1979).
- [32] H.-H. v. Borzeszkowski and H.-J. Treder, in *Classical and Quantum Non-locality*, Eds. P.G. Bergmann, V. de Sabbata, and N. Goldberg (World Scientific, Singapore, 2000).
- [33] H.-H. v. Borzeszkowski and H.-J. Treder, *Found. Phys.* **25**, 291 (1993); **26**, 929 (1996); **27**, 595 (1997); **28**, 273 (1998).
- [34] F. Müller-Hoissen and J. Nitsch, *Phys. Rev. D* **28**, 718 (1983).
- [35] F.W. Hehl, in *Cosmology and Gravitation*, Eds. P.G. Bergmann and V. de Sabbata (Plenum, New York, 1980).
- [36] H. Bauer, *Phys. Zeitschr.* **19**, 163 (1918).
- [37] P.G. Bergmann, *Phys. Rev.* **75**, 680 (1949).
- [38] P.G. Bergmann, *Rev. Mod. Phys.* **33**, 510 (1961).
- [39] F.W. Hehl and J.D. McCrea, *Found. Phys.* **16**, 267 (1986).
- [40] H. Kleinert, *Gen. Rel. Grav.* **32**, 769 (2000).
- [41] H. Kleinert, *Gen. Rel. Grav.* **32**, 1271 (2000).
- [42] H.-H. v. Borzeszkowski and H.-J. Treder, in *Quantum Gravity*, Eds. P.G. Bergmann, V. de Sabbata, and H.-J. Treder (World Scientific, Singapore, 1996).

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

ASYMPTOTIC FREEDOM IN CURVATURE-SATURATED GRAVITY

S. CAPOZZIELLO, G. LAMBIASE

*Dipartimento di Scienze Fisiche "E.R. Caianiello",
Università di Salerno, 84081 Baronissi (Sa), Italy and
Istituto Nazionale di Fisica Nucleare, Sez. di Napoli, Italy
E-mails: capozziello@sa.infn.it, lambiase@sa.infn.it*

H.-J. SCHMIDT

*Freie Universität Berlin, Institut für Theoretische Physik,
Arnimallee 14, D-14195 Berlin, Germany and
Institut für Mathematik, Universität Potsdam,
Am Neuen Palais 10, D-14469 Potsdam, Germany
E-mail: hjschmi@rz.uni-potsdam.de*

For a spatially flat Friedmann model with line element $ds^2 = a^2[da^2/B(a) - dx^2 - dy^2 - dz^2]$, the 00-component of the Einstein field equation reads $8\pi GT_{00} = 3/a^2$ and contains no derivative. For a nonlinear Lagrangian $\mathcal{L}(R)$, we obtain a second-order differential equation for B instead of the expected fourth-order equation. We discuss this equation for the curvature-saturated model proposed by Kleinert and Schmidt [1]. Finally, we argue that asymptotic freedom $G_{\text{eff}}^{-1} \rightarrow 0$ is fulfilled in curvature-saturated gravity.

1 Introduction

In the past decades, several extended theories of gravity have been proposed, whose effective actions are more general than the Einstein-Hilbert action. This approach is motivated by unification schemes which consider gravity at the same level as the other interactions of elementary particles. In such theories, we define an effective gravitational coupling G_{eff} and an effective cosmological constant Λ_{eff} which give the observed values $G_{\text{eff}} \rightarrow G$ and $\Lambda_{\text{eff}} \rightarrow \Lambda$ in the weak-energy limit.

These extended theories introduce new features into gravitational physics which general relativity does not possess, in particular higher-order terms in curvature invariants or non-minimal couplings between geometry and scalar fields. For example, *asymptotic freedom* can be related to singularity-free cosmological models. Its emergence could be the result of the fact that gravity is "induced" by an average effect of the other interactions [2,3].

However, the concept of gravitational asymptotic freedom is not completely analogous to that in non-abelian gauge theories of strong interaction, since a full quantum theory of gravity does not yet exist. So far, asymptotic freedom has been shown for several classes of gravitational Lagrangians which are not of physical interest.

Another interesting feature is that, by extended theories of gravity, cosmological singularities can be avoided introducing a *limiting curvature hypothesis* for values of curvature near the Planck scale [4]. This fact is particularly relevant for string-dilaton gravity since, by duality, it allows to recover large classes of cosmological solutions for $t \rightarrow -\infty$ [5].

In Ref. [1], it was argued that the Lagrangian for gravity should remain bounded at large curvature. One example for such a behavior is the curvature-saturated Lagrangian

$$\mathcal{L}_{CS} = \frac{1}{16\pi G} \frac{R}{\sqrt{1 + l^4 R^2}}. \quad (1)$$

It has been discussed in Ref. [1] with methods developed for the study of fourth-order gravity developed in Refs. [6,7] and the publications cited therein. In Eq. (1), l is a length parameter, and for $l = 0$, $\mathcal{L}_{CS}$ reduces to the Einstein-Hilbert Lagrangian. Earlier models with a nonlinear Lagrangian $\mathcal{L}(R)$ made only a polynomial approximation like

$$\mathcal{L} = \frac{R}{16\pi G} + \sum_{k=2}^n c_k R^k, \quad c_n \neq 0, \quad (2)$$

sometimes accompanied by a logarithmic or a R^m -term with non-integer m . In all such cases, one had $d\mathcal{L}/dR \rightarrow \pm\infty$ for $|R| \rightarrow \infty$. In contrast to this, $\mathcal{L}_{CS}$ of Eq. (1) has the behavior $d\mathcal{L}/dR \rightarrow 0$ for $R \rightarrow \pm\infty$.

This limiting property can be reformulated in terms of the effective gravitational coupling

$$G_{\text{eff}} = \left[16\pi \frac{d\mathcal{L}}{dR} \right]^{-1} \quad (3)$$

as follows: In the previously extended models (2), G_{eff} is bounded (including zero), in contrast to the curvature-saturated model (1) where one gets $|G_{\text{eff}}| \rightarrow \infty$ as $|R| \rightarrow \infty$; of course, for $l = 0$ we recover $G_{\text{eff}} \equiv G$.

Other limiting behaviors have been discussed in Refs. [2,3]. In Ref. [3], a scalar-tensor theory with Lagrangian

$$\mathcal{L}_\varphi = F(\varphi)R + \frac{1}{2}g^{ij}\varphi_{,i}\varphi_{,j} - V(\varphi) \quad (4)$$

has been used where $i, j = 0, 1, 2, 3$ and the effective gravitational constant is

$$G_{\text{eff}} = \frac{1}{16\pi F(\varphi)} \quad (5)$$

instead of our Eq. (3). In Ref. [2], G_{eff} was assumed to depend on the matter density in such a way that it vanishes for high density, thus leading to asymptotic freedom at high energies.

In Ref. [8], possible finite-size *Casimir* effects to the free energy have been calculated for massive and massless scalar fields, which can produce a quantum-effected effective gravitational constant.

From a geometric point of view, one may argue as follows: For large curvatures, the gravitational action $I = \int \mathcal{L}_C \sqrt{-g} d^4x$ shall not depend on any scale; the simplest Lagrangian leading to such a behavior is

$$\mathcal{L}_C = C_{ijkl}C^{ijkl}, \quad (6)$$

the Lagrangian of the conformally invariant Weyl gravity (see Ref. [9] and the references therein).

Here, we extend the discussion of Ref. [1] to more general types of curvature-saturated Lagrangians than Eq. (1), and we give more details about the set of spatially flat Friedmann models solving the corresponding field equations. As it will become clear below, the concept of asymptotic freedom is meaningful also for curvature-saturated theories and it is necessary to define more specifically the notion of an effective cosmological constant.

2 Deduction of the Field Equations

For the action $I = \int \mathcal{L}(R) \sqrt{-g} d^4x$, it is useful to write the metric of the expanding spatially flat Friedmann model as

$$ds^2 = a^2 \left[\frac{da^2}{B(a)} - dx^2 - dy^2 - dz^2 \right], \quad (7)$$

with $a > 0$ and $B(a) > 0$ as shown in Ref. [1]. In these coordinates, the parameter a is called *curvature time* because the curvature scalar has the simple form

$$R = -\frac{3}{a^3} \frac{dB}{da}, \quad (8)$$

which is linear in the only unknown function $B(a)$, and it does not contain second derivatives. The non-vanishing components of the Christoffel affinity are as follows

$$\Gamma_{00}^0 = \frac{1}{a} - \frac{B'}{2B}, \quad \Gamma_{0\alpha}^\beta = \frac{1}{a} \delta_\alpha^\beta, \quad \Gamma_{\alpha\beta}^0 = \frac{B}{a} \delta_{\alpha\beta}, \quad (9)$$

where the dash denotes d/da . The greek indices assume the spatial values $\alpha, \beta = 1, 2, 3$, and $\sqrt{-g} = a^4/\sqrt{B}$. The Ricci tensor has the components

$$R_{\alpha\beta} = \left(\frac{B}{a^2} + \frac{B'}{2a} \right) \delta_{\alpha\beta}, \quad R_{00} = \frac{3}{a^2} - \frac{3B'}{2aB}. \quad (10)$$

Together with the metric (7), its mixed-variant version can be calculated to

$$R_\alpha^\beta = -\left(\frac{B}{a^4} + \frac{B'}{2a^3} \right) \delta_\alpha^\beta, \quad R_0^0 = \frac{3B}{a^4} - \frac{3B'}{2a^3}. \quad (11)$$

The most often used direct way to deduce the field equation is to insert these expressions into the fourth-order field equation $\delta I/\delta g_{ij} = 0$ following from the variation of the action I .

Here we present a shorter and more direct derivation: We first use the fact that the field equation also implies $\delta I/\delta B = 0$, and in a second step, verify that in spite of this simplified variation, no spurious solutions appear.

Let us denote $d\mathcal{L}/dR$ by $h(R)$, and $d^2\mathcal{L}/dR^2$ by $k(R)$. The vacuum field equation reads

$$0 = \frac{\delta(\mathcal{L}\sqrt{-g})}{\delta B} = \frac{\partial(\mathcal{L}\sqrt{-g})}{\partial B} - \left(\frac{\partial(\mathcal{L}\sqrt{-g})}{\partial B'} \right)'. \quad (12)$$

After multiplying this by a^7 to avoid negative a -powers, we obtain

$$0 = -a^7 \mathcal{L}(-3B'/a^3) + 3a^3(2B - aB')h(-3B'/a^3) + 18B(3B' - aB'')k(-3B'/a^3). \quad (12)$$

It is remarkable that this equation is of second order for one function B only, but nevertheless, it is equivalent to the whole fourth-order field equation for

the metric (7). One order reduction follows from Eq. (8), the other from the fact that Eq. (12) is a constraint and not the full dynamical equation.

To exclude the existence of spurious solutions it now suffices to see that exactly two free initial conditions can be put: $B(a_0)$ and $B'(a_0)$.

Example: Let $\mathcal{L} = R + 2\Lambda$, where Λ is a constant. Then $h = 1$ and $k = 0$, and Eq. (12) reads

$$B(a) = \frac{a^4 \Lambda}{3}. \quad (13)$$

The Hubble parameter H is related to B via

$$B(a) = H^2 a^4 \quad (14)$$

(see Ref. [1]). So, with Eq. (13) we get $\Lambda = 3H^2 = \text{const.}$, consistent with the usual de Sitter space-time calculation.

3 Solutions of the Field Equations

Now we consider some more general cases for $\mathcal{L}(R)$ and look for the corresponding solutions of Eq. (12).

3.1 Solution for Einstein's Theory

To get a feeling for the coordinates (7) we look for the 00-component of the Einstein field equation

$$8\pi G T_{ij} = R_{ij} - \frac{1}{2} R g_{ij}.$$

Using Eqs. (8) and (10), we obtain

$$8\pi G T_{00} = \frac{3}{a^2}. \quad (15)$$

The energy density is always non-negative for spatially flat Friedmann models in Einstein's theory. Eq. (15) provides us with another justification for calling this a curvature-coordinate: a is chosen such that T_{00} does not depend at all on the function $B(a)$.

Raising one index of Eq. (15) we find

$$8\pi G \rho \equiv 8\pi G T_0^0 = \frac{3B}{a^4}. \quad (16)$$

Together with Eq. (14), this represents a good consistency test: it yields the Friedmann equation

$$8\pi G\rho = 3H^2, \quad (17)$$

i.e. curvature time and synchronous time give identical results.

For the equation of state $p = \alpha\rho$ we have

$$\rho a^{3(1+\alpha)} = \text{const.} \quad (18)$$

Together with Eq. (16), this yields

$$B = B_0 a^{1-3\alpha}, \quad (19)$$

with a positive constant B_0 . This is consistent with Eq. (13) for $\alpha = -1$.

3.2 Solution for $\mathcal{L} = R^m$

For the Lagrangian $\mathcal{L} = R^m$, with a constant m ($m \neq 0, 1$), Eq. (12) simplifies to

$$0 = 2m(3m-4)BB' + (m-1)aB'^2 - 2m(m-1)aBB''. \quad (20)$$

We insert

$$B = \exp\left(\int \beta da\right). \quad (21)$$

The integration constant in Eq. (21) need not be specified, because Eq. (20) implies that with $B(a)$ also $CB(a)$, with any positive constant C , represents a solution.

Eq. (21) implies $B' = \beta B$ and $B'' = (\beta' + \beta^2)B$. Then Eq. (20) becomes

$$0 = 2m(3m-4)\beta - (m-1)(2m-1)a\beta^2 - 2m(m-1)a\beta'. \quad (22)$$

Here we insert $\beta = \gamma/a$, $\beta' = \gamma'/a - \gamma/a^2$ and get

$$0 = 2m(4m-5)\gamma - (m-1)(2m-1)a\gamma^2 - 2m(m-1)a\gamma'. \quad (23)$$

By putting $x = \ln a$, Eq. (23) can be rewritten in the form

$$\frac{d\gamma}{dx} = \gamma \left[\frac{4m-5}{m-1} - \frac{2m-1}{2m}\gamma \right]. \quad (24)$$

The de Sitter space-time is represented by $\gamma \equiv 4$. From Eq. (24) it becomes clear that it is a solution for $m = 2$ only, i.e. for $\mathcal{L} = R^2$.

All solutions of (24) can be given in closed form. For $m = 5/4$, we get with any constant x_0 :

$$\gamma = \frac{5}{3(x - x_0)},$$

and, together with $a_0 = e^{x_0}$ and Eq. (21), finally

$$B(a) = \exp \left[\frac{5}{3} \int \frac{da}{a \ln(a/a_0)} \right] = B_0 \left(\ln \frac{a}{a_0} \right)^{5/3}. \quad (25)$$

For $m = 1/2$ we find $\gamma = \pm e^{6(x-x_0)}$ and

$$B(a) = \exp[\pm(a/a_0)^6]. \quad (26)$$

Let us study the stability of these solutions. Instead of directly comparing with neighboring functions $B(a)$ it is easier to consider the neighborhood of the solution: $\gamma = \text{const.}$ in Eq. (24). Up to the uninteresting solution $\gamma = 0$ representing $R \equiv 0$ we find

$$\gamma = \gamma_0 \equiv \frac{4m-5}{m-1} \frac{2m}{2m-1},$$

i.e.

$$\frac{d\gamma}{dx} = \gamma(\gamma_0 - \gamma) \frac{2m-1}{2m}. \quad (27)$$

For $m > 1/2$ this implies stability. For $m = 2$ this proves again the attractor property of the de Sitter space-time.

In synchronized time, power-law inflation is described by $a(t) \sim t^n$, $n \geq 1$. In our coordinates, such solutions correspond to a constant γ within the interval $2 \leq \gamma < 4$.

3.3 Solutions for Curvature-Saturated Lagrangians

As another model, consider the high-curvature ansatz

$$\mathcal{L} = \Lambda + \frac{C}{R}, \quad (28)$$

in which $h(R) = -C/R^2$, $k(R) = 2C/R^3$, with constants Λ and C . The ansatz (28) is supposed to approximate the regions $R \rightarrow \pm\infty$, and it will be matched at small curvature values to a polynomial Lagrangian of type Eq. (2). Here we are only interested in the high-curvature regions. The concrete values

of Λ and C are not yet fixed, and they may be different for $R \rightarrow +\infty$ and $R \rightarrow -\infty$, respectively.

For $\Lambda = 0$ we can directly apply Eq. (27) with $m = -1$, i.e. $\gamma_0 = 3$, and

$$\frac{d\gamma}{dx} = \frac{3}{2}\gamma(3 - \gamma). \quad (29)$$

This equation can be integrated in closed form, but we only need that $\gamma = 3$ represents power-law inflation and that this represents an attractor solution. In synchronized coordinates, it reads

$$ds^2 = dt^2 - t^4(dx^2 + dy^2 + dz^2), \quad (30)$$

giving $R \rightarrow 0$ as $t \rightarrow \infty$. Thus, for sufficiently large t , the development leads to our Universe today.

For $\Lambda \neq 0$, we insert (28) into (12), and have to solve

$$3\Lambda(B')^3 = 2a^2C[5BB' + aB'^2 - 2aBB'']. \quad (31)$$

This equation will be further integrated in future work.

4 Curvature-Saturated Gravity with Matter

To include matter, we have to replace the l.h.s. of Eq. (12) by $a^7\rho$. For an ideal fluid, we can use Eq. (18) as an equation of state. As a test of this procedure, we can insert $\mathcal{L} = R/16\pi G$, and get $a^7\rho = 3a^3B/8\pi G$, which is consistent with Eq. (16).

Therefore, matter with the equation of state $p = \alpha\rho$ can be described by $\rho = \rho_0 a^{-3(1+\alpha)}$, and the equations of motion are

$$\begin{aligned} a^7\rho &= -a^7\mathcal{L}(R) + 3a^3(2B - aB')h(R) + 18B(3B' - aB'')k(R), \\ R &= -3B'a^{-3}. \end{aligned} \quad (32)$$

We stress that such equations hold for general $\mathcal{L}(R)$ -Lagrangians and for any perfect fluid matter.

5 Conclusions

In this paper, we continued the discussion on properties of the curvature-saturated cosmological models proposed in Ref. [1]. Especially, we deduced the main equation *with* matter, Eq. (32), for a spatially flat Friedmann model in the curvature coordinates (7).

We have shown that, from the singularity, the Universe expands via power-law inflation (which represents a transient attractor) to the actual state.

In order to find the Wheeler-DeWitt equation for fourth-order gravity models (see Refs. [6,10,11] and the papers cited there), one usually has to introduce, in a more or less natural way, more degrees of freedom to reduce the equation down to second order. Here we presented a version where the Lagrangian can be directly used: $\mathcal{L}(R)\sqrt{-g}$ depends on a , B and B' only; of course, different from the use of the synchronized time coordinate, we have now an explicit dependence of the metric coefficients on the time-like coordinate a .

As a final remark, we see that asymptotic freedom $G_{\text{eff}}^{-1} = 16\pi d\mathcal{L}/dR \rightarrow 0$ can easily be incorporated also in curvature-saturated gravity but its meaning is different from that in Ref. [3]. There G_{eff} is a function of matter density which regulates the gravitational coupling, here it is the scalar curvature (i.e. the form of the gravitational Lagrangian) which leads towards the saturation and then towards asymptotic freedom. Future studies will be devoted to more physically motivated effective Lagrangians.

Acknowledgments

We thank Professor Hagen Kleinert for his constant encouragement over the years. He combines horizontal broadness achieved by treating a large number of topics, with vertical deepness shown by his ability to criticize weaknesses in manuscripts even specialists in the field had not noticed. We heartily congratulate him to his 60th birthday, and we wish him a fruitful continuation of his active life.

H.-J.S. thanks the colleagues of Salerno University where this work has been completed, especially Professor Gaetano Scarpetta, for kind hospitality. The researches of S.C. and G.L. have been supported by fund MURST PRIN 99. H.-J.S. is supported by the HSP III-program Potsdam.

References

- [1] H. Kleinert and H.-J. Schmidt, e-print: gr-qc/0006074.
- [2] M. Markov, *Physics-Uspekhi* **37**, 57 (1994).
- [3] S. Capozziello and R. de Ritis, *Phys. Lett. A* **208**, 181 (1995); S. Capozziello, R. de Ritis, and A. Marino, *Phys. Lett. A* **249**, 395 (1998).
- [4] Y. Anini, *Current Topics in Mathematical Cosmology*, Eds. M. Rainer

and H.-J. Schmidt (World Scientific, Singapore, 1998), p. 183; R. Brandenberger, V. Mukhanov, and A. Sornborger, *Phys. Rev. D* **48**, 1629 (1993).

- [5] A.A. Tseytlin and C. Vafa, *Nucl. Phys. B* **372**, 443 (1992).
- [6] H.-J. Schmidt, *Phys. Rev. D* **49**, 6354 (1994), e-print: gr-qc/9404038.
- [7] S. Capozziello, G. Lambiase, and H.-J. Schmidt, *Ann. Phys. (Leipzig)* **9**, 39 (2000).
- [8] A. Zhuk and H. Kleinert, *Theor. and Math. Phys.* **109**, 1483 (1996); H. Kleinert and A. Zhuk: unpublished preprint 1993: *Finite-Size and Temperature Properties of Matter and Radiation Fluctuations in Closed Friedmann Universe*; H. Kleinert and A. Zhuk: *Casimir Effect at Nonzero Temperature in Closed Friedmann Universe*; the files can be downloaded from <http://www.physik.fu-berlin.de/~kleinert/> as No. 218 and 251, respectively.
- [9] V. Dzhunushaliev and H.-J. Schmidt, *J. Math. Phys.* **41**, 3007 (2000), e-print: gr-qc/9908049.
- [10] J. Fabris and S. Reuter, *Gen. Rel. Grav.* **32**, 1345 (2000).
- [11] M. Bachmann and H.-J. Schmidt, *Phys. Rev. D* **62**, 043515 (2000), e-print: gr-qc/9912068.

WHAT CAN ISING SPINS TEACH US ABOUT QUANTUM GRAVITY?

C. HOLM

*Max-Planck-Institut für Polymerforschung,
Ackermannweg 10, D-55128 Mainz, Germany
E-mail: holm@mpip-mainz.mpg.de*

We investigate numerically an Ising spin model coupled to two-dimensional Euclidean quantum gravity. We employ Regge calculus to discretize the gravitational interaction. We study this system on a toroidal and a spherical manifold, with two different local path integral measures, and an added R^2 interaction term. We find in all cases that the critical exponents of the Ising transition are consistent with the Onsager values, and that the KPZ exponents are definitely excluded.

1 Introduction

The study of two-dimensional (2D) models has often proved to be an important first step in developing fundamental ideas concerning higher-dimensional physics. Such has happened for 2D Euclidean quantum gravity where we have by now independent analytic results from conformal field theory [1] and matrix models [2], for the critical exponents of a simple toy matter field, namely a spin model coupled to a fluctuating geometry. Historically, the interest in 2D quantum gravity was inspired by string theory. The time development of strings leads to 2D world surfaces, which are reparametrization invariant and quantized, hence describe 2D quantum gravity. Kazakov [3] suggested a model of Ising spins living on the vertices of so-called planar ϕ^4 graphs. An exact expression of the partition function was given for Z_n in the thermodynamic limit $n \rightarrow \infty$ by relating this model to an exactly solvable Hamiltonian model of two Hermitian matrices [2,4]. Shortly afterwards [5], the same model was solved also on ϕ^3 graphs, and some time later Knizhnik, Polyakov, and

Zamolodchikov (KPZ) [1] found the same set of critical exponents in a continuum model of 2D quantum gravity using methods of conformal field theory for matter of central charge $c = 1/2$. Since continuum field theory as well as lattice models agree on the set of critical exponents, one felt comfortable with the idea that Kazakov's model was really a model of quantum gravity, and that matter fields can be strongly influenced, when they are coupled to a quantum geometry. These new critical exponents were called KPZ exponents, and turn out to be quite different from the original Ising exponents. The critical exponents change from the flat-space Onsager values $\alpha = 0$, $\beta = 0.125$, $\gamma = 1.75$, and $\nu = 1$ to the values $\alpha = -1$, $\beta = 0.5$, $\gamma = 2$, and $D\nu = 3$, where D is the internal fractal dimension of the manifold.

In the following work we will entirely remain in two dimensions and investigate if the analytic KPZ results obtained for the Ising system ($c = 1/2$) can be obtained with Regge's method transcribed to the quantum domain. Although this question poses itself rather naturally, only few people have actually investigated this subject. The first to look at this problem were Gross and Hamber [6] who found the classical flat-space Ising critical exponents. This came as a surprise since a different method, that was termed dynamical triangulated random surfaces (DTRS) and that is more or less a Monte Carlo version of the Boulatov model, gave KPZ results. We have put considerably effort in modifying the global topology, the local path integral measure, and added an R^2 interaction term, in order to see if one observes any effect on the critical exponents. The results we will present here, have been obtained over the course of the past seven years [7,8]. We will first review the method, then the simulation technique, followed by the results, and end with some conclusion.

2 Regge Calculus

Regge calculus [9] is a discretization approach to gravity which reduces the infinite degrees of freedom of Riemannian manifolds to a finite number of parameters by working with piecewise linear spaces. Regge calculus has found numerous applications in classical and quantum physics. An introduction with an extensive list of references can be found in Ref. [10]. It can undoubtedly be regarded as the best-understood method to discretize classical gravity. Historically this method was used in the first numerical attempts to study quantum gravity non-perturbatively [11,12], but it can also be used as a regularized version of quantum gravity in which one can perform analytic

calculations [13]. It has been mostly employed in four dimensions where extensive simulations have been carried out [14]. The Regge approach in two dimensions consists of choosing a triangulation of the manifold under consideration, which means that the topology stays fixed from the beginning. One then assigns link lengths to each triangle (or simplex, in higher dimensions), which play the role of the dynamical variables. This is, incidentally, the opposite procedure to what one does in the so-called DTRS method, where one has fixed edge lengths, and a fluctuating connectivity. The values of its squared edge length $q_{ij} = l_{ij}^2$ induce a constant metric in the interior of each simplex, because q_{ij} can linearly be related to the three components of the metric $\eta_{\mu\nu}$. Local curvature can be described by the rotation experienced by a vector when it is parallel transported in a closed curve around a vertex, where several triangles meet. The angle of rotation is measured by the deficit angle δ_i at the vertex i that can intrinsically be defined as

$$\delta_i = 2\pi - \sum_{\text{all } t \text{ sharing } i} \theta_i(t), \quad (1)$$

and $\theta_i(t)$ is the dihedral angle associated with the triangle t . Curvature is therefore distributed delta-function-like with support on the vertices. The dihedral angles can be computed purely out of the link length information. Defining the barycentric area connected to the site i ,

$$A_i = \sum_{t \supset i} \frac{1}{3} A_t, \quad (2)$$

where A_t denotes the area of the triangle t , one can then write down the simplicial analogue of continuum integrals like

$$\int d^2x \sqrt{g(x)} \longrightarrow \sum_i A_i, \quad (3)$$

$$\frac{1}{2} \int d^2x \sqrt{g(x)} R(x) \longrightarrow \sum_i \delta_i = 2\pi \chi(\mathcal{M}), \quad (4)$$

$$\int d^2x \sqrt{g(x)} R^2(x) \longrightarrow 4 \sum_i \frac{\delta_i^2}{A_i}. \quad (5)$$

Eq. (4) is the simplicial analogue of the Gauss-Bonnet theorem which relates the differential-geometric integral on the left hand side to a topological invariant, namely the Euler characteristic χ . This makes pure 2D gravity, based on the action corresponding to Eq. (4), dynamically trivial, but leaves still

open the possibility that matter fields can be influenced by the fluctuating geometry. In two dimensions it can be shown that Eq. (5) gives the exact continuum result on any regular triangulation of the sphere [11].

The Euler characteristic of a two-dimensional manifold $\mathcal{M}$ can also be written as $\chi = 2(1-g)$, where g is the gender of the surface, which counts how many holes there are in $\mathcal{M}$. For a simplicial complex, the Euler characteristic can also be computed as

$$N_0 - N_1 + N_2 = 2(1 - g), \quad (6)$$

where N_0 , N_1 , and N_2 denote the number of sites, links, and triangles, respectively. For a compact complex we also know that a link is shared by two triangles, resulting in the relation $N_1/3 = N_2/2$. From these two relations one can derive two more, namely $N_0 - 2(1-g) = N_2/2$ and $N_0 - 2(1-g) = N_1/3$, which will become useful later. The sphere has $g = 0$ and the torus has $g = 1$.

3 Model and Simulation Techniques

We simulated the partition function

$$Z = \sum_{\{s\}} \int D\mu(l) \exp[-I(l) - KE(l, s)] \delta\left(\sum_i A_i - A\right), \quad (7)$$

where $\{s\}$ denotes the set of all spin configurations of the Ising spins $s_i = \pm 1$, and the gravitational action is defined as

$$I(l) = \sum_i \left(\lambda A_i + a \frac{\delta_i^2}{A_i} \right). \quad (8)$$

The energy of Ising spins, which are located at the vertices i of the lattice, is denoted by

$$E(l, s) = \frac{1}{2} \sum_{\text{edges } l_{ij}} A_{ij} \left(\frac{s_i - s_j}{l_{ij}} \right)^2, \quad (9)$$

and the barycentric area A_{ij} associated with a link l_{ij} is defined as

$$A_{ij} = \sum_{\text{triangles } t \supset l_{ij}} \frac{1}{3} A_t. \quad (10)$$

The energy is the discretized analogue of the continuum action for a scalar field φ , i.e. $\int d^2x \sqrt{g(x)} g^{\mu\nu}(x) \partial_\mu \varphi \partial_\nu \varphi$.

The delta function in Eq. (7) ensures that the total area is kept fixed at a given value A , thus rendering the cosmological constant irrelevant, and classically, gravitational dynamics can only arise from an R^2 -interaction term. Such a term was used in two and higher dimensions to cure the unboundedness problem of the pure gravitational action [11]. From dimensional arguments, one would expect such a term to be irrelevant for the Ising transition. An inclusion of this term enabled us to test these arguments.

The quantization procedure of the classical Regge action rests on the path integral formulation [15] of Eq. (7). For the path-integral measure $D\mu(l)$ we mostly used a simple scale-invariant measure of the form [7,16]

$$\mathcal{D}\mu(q) = \prod_{\langle ij \rangle} dq_{ij}/q_{ij} F_{\epsilon}(\{q_{ij}\}). \quad (11)$$

The function $F_{\epsilon}(\{q_{ij}\})$ ensures Euclidean geometry. If the triangle inequalities are obeyed it assumes the value one, otherwise it vanishes. The parameter ϵ serves to suppress very thin triangles by generalizing the triangle inequalities to a (still scale-invariant) form $(l_1 + l_2) \geq (1 - \epsilon)l_3$ and $(l_1 - l_2) \leq (1 + \epsilon)l_3$. This makes the algorithm faster, because many proposed new values for l that would get discarded with high probability, are thrown out at some earlier program steps. This is not necessary for convergence, unlike in higher dimensions. For our simulations, ϵ was of the order 10^{-4} . We checked that a different value of ϵ did not change the outcome of our measurements. The attractiveness of the measure (11) lies in the fact that it is local, scale invariant, and the integral runs directly over the variables $\xi_{ij} = \ln l_{ij}$, which makes it well-adapted for computer simulations, hence the term "computer measure". Sometimes we will also use the abbreviation " dl/l -measure" for the measure of Eq. (11).

Most continuum measures differ by the power of the determinant of the metric g , which stands in front of the integration measure. In a naive transcription of the volume element $\sqrt{g}$ to the Regge formalism one identifies $\sqrt{g} \rightarrow A_{ij}$, where A_{ij} is a volume associated with a link l_{ij} . Keeping the freedom of having some power of l_{ij} appearing in the measure, one is led to consider the following two-parameter class of measures [6]:

$$\mathcal{D}\mu(q) = \prod_{\langle ij \rangle} A_{ij}^{2\sigma} q_{ij}^{(\alpha-1)} dq_{ij} F_{\epsilon}(\{q_{ij}\}). \quad (12)$$

In terms of the two parameters α and σ the computer measure corresponds to $\alpha = 0$, $\sigma = 0$. An analog of the scale invariant Misner measure would be

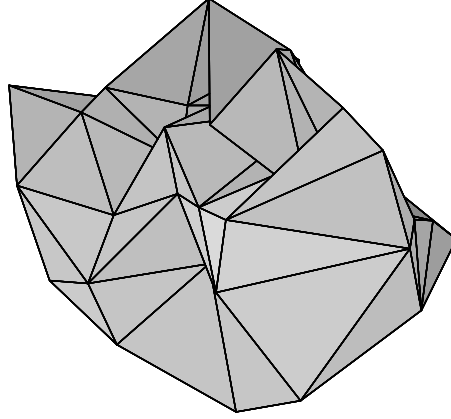


Figure 1. dl/l measure configuration for the sphere with $d = 4$, corresponding to $N_0 = 56$.

$\alpha = 1, \sigma = -1/2$, and the lattice analog of the popular DeWitt measure is given by $\alpha = 1, \sigma = -1/4$. We will present here also some preliminary data with the DeWitt measure, which is not scale-invariant [8].

We have simulated the gravitational action using a standard multi-hit Metropolis update. As Ising update we used the single-cluster (Wolff) algorithm [17] which prevents the critical slowing down near the phase transition. One update in the single-cluster variant consists of choosing a random mirror plane and a random site, which is the starting point for growing a cluster of reflected spins. The size and shape of the cluster is controlled by a Metropolis like accept/reject criterion satisfying detailed balance.

Between measurements we performed $n = 2, \dots, 4$ Monte Carlo steps consisting of one lattice sweep to update the link lengths l_{ij} followed by a single-cluster flip to update (a fraction of) the spins s_i . We tested in some cases that varying the relative frequency of link and spin updates does not change the results within error bars.

We used two different global topologies for our simulations to check for a possible influence of topology on the phase transition. The main investigation was performed on regular triangulated tori of size $N_0 = L^2$ with fixed coordination number $q = 6$. This triangulation gives rise to $2N_0$ triangles and $3N_0$ link variables. The principal simulations were performed at a curvature squared coupling value of $a = 0.001$ and the couplings $K = 1$ and $K = 1.025$ for $L = 6, 8, 10, 12, 16, 32, 64, 100, 128, 200, 256$, and 512 . Additional simulations were performed with $a = 0$ and 0.1 at $K = 1.025$, using lattices of

size $L = 8, 16, 32, 64, 100, 128$, and 256 . Because of the scale invariance of the measure we could rescale each link when proposing a link update such that the total area was kept fixed to its initial value $A = N_0$. The difference of the model defined by (7) and the Ising model on a static triangular lattice is that the spins are coupled by geometric weight factors $w_{ij} = A_{ij}/l_{ij}^2$ which can fluctuate around the static value $w_{ij} = \sqrt{3}/6$.

For the spherical lattice topology we used the triangulated surface of a three-dimensional cube of edge length d . This provides us with an almost regular triangulation of the sphere where six vertices have coordination number four, and all others have coordination number six. In terms of the linear length d of the cube the number of vertices is $N_0 = 6(d-1)^2 + 2$. For further reference the number of links and triangles in terms of N_0 are given by $N_1 = 3N_0 - 6$, and $N_2 = 2N_0 - 4$, respectively. We studied ten system sizes ranging from $d = 10$ ($N_0 = 488$) up to $d = 55$ ($N_0 = 17498$). The area was kept fixed to its initial value $A = N_2/2$, and we used no coupling to the R^2 term, i.e. $a = 0$. As simulation point we have chosen a value of $K_0 = 1.025$, already anticipating that this value is close to the critical coupling K_c on the torus. To compare previous values we set as our linear length scale $L = \sqrt{N_0}$. A typical configuration can be viewed at in Fig. 1, which was produced using the computer measure of Eq. (11)

For each run we recorded the time series of the energy density $e = E/N_0$, the magnetization density $m = \sum_i A_i s_i / N_0$ and the Liouville field density $\varphi = \sum_i \ln A_i / N_0$. After an initial equilibration time, we performed for each lattice size about 50 000 measurements. From an analysis of the time series we found integrated autocorrelation times for the energy and the magnetization of about 1–7 (in units of measurements) for all lattice sizes. To obtain results for the various observables $\mathcal{O}$ at K values in an interval around the simulation point K_0 , we applied the reweighting method [18]. Since we recorded the time series this amounts to computing

$$\langle \mathcal{O} \rangle_K = \frac{\langle \mathcal{O} e^{-\Delta KE} \rangle_{K_0}}{\langle e^{-\Delta KE} \rangle_{K_0}}, \quad (13)$$

with $\Delta K = K - K_0$. To obtain errors we divided each run into 20 blocks and computed standard Jackknife errors. At $a = 0.001$ where we had two simulations at different K values, we combined the results according to their errors [19].

From the time series we computed the Binder parameter [20],

$$U_L(K) = 1 - \frac{1}{3} \frac{\langle m^4 \rangle}{\langle m^2 \rangle^2}. \quad (14)$$

It is well known that the $U_L(K)$ curves for different L cross around (K_c, U^*) with slopes $\propto L^{1/\nu}$, apart from confluent corrections explaining small systematic deviations. This allows an almost unbiased estimate of the critical coupling K_c , the critical correlation length exponent ν , and the renormalized charge U^* . We further analyzed the (finite lattice) susceptibility,

$$\chi(K) = N_0(\langle m^2 \rangle - \langle |m| \rangle^2), \quad (15)$$

the susceptibility in the disordered phase,

$$\chi'(K) = N_0(\langle m^2 \rangle), \quad (16)$$

the specific heat,

$$C(K) = K^2 N_0(\langle e^2 \rangle - \langle e \rangle^2), \quad (17)$$

and studied the (finite lattice) magnetization at its point of inflection, $\langle |m| \rangle|_{\text{inf}}$. The inflection point can be obtained from the maximum of $d\langle |m| \rangle/dK$. Further useful quantities are the logarithmic derivatives $d\ln\langle |m| \rangle/dK$ and $d\ln\langle m^2 \rangle/dK$. Another gravitational quantity of interest is the Liouville field $\varphi(x) = \ln \sqrt{g(x)}$. In the discretized version its lattice average reads as $\varphi = 1/N_0 \sum_i \ln A_i$, and the associated lattice Liouville susceptibility is defined as $\chi_\varphi(L) = N_0(\langle \varphi^2 \rangle - \langle \varphi \rangle^2)$.

4 Results

By applying the outlined reweighting techniques we first determined the maxima of χ , C , $d\langle |m| \rangle/dK$, $d\ln\langle |m| \rangle/dK$, and $d\ln\langle m^2 \rangle/dK$. The location of the maxima provided us with five sequences of pseudo-transition points $K_{\text{max}}(L)$ for which the scaling variable $x = (K_{\text{max}}(L) - K_c)L^{1/\nu}$ should be constant. Using this information we then have several possibilities to extract the critical exponent ν from (linear) least square fits of the finite-size scaling (FSS) ansatz $dU_L/dK \cong L^{1/\nu} f_0(x)$ or $d\ln\langle |m|^p \rangle/dK \cong L^{1/\nu} f_p(x)$ to the data at the various $K_{\text{max}}(L)$.

For the torus, the very extensive simulations at $a = 0.001$ show good agreement with the Onsager value $\nu = 1$ at a 2 % level. For the other

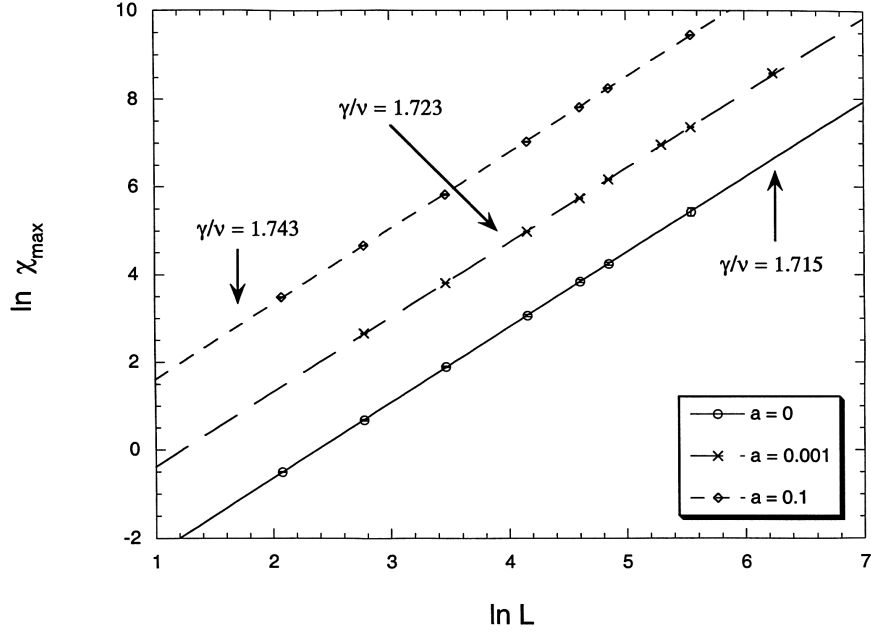


Figure 2. Double logarithmic finite-size scaling plot of the susceptibility maxima $\chi_{\max}$ for $a = 0.0, 0.001$, and 0.1 on the torus. To disentangle the curves we added an offset of -2 (2) to the data for $a = 0.0$ ($a = 0.1$). The slopes are in all three cases compatible with the Onsager value $\gamma/\nu = 1.75$ for regular static lattices.

two couplings, $a = 0.1$ and $a = 0$, the data scatter a bit more but are still compatible with $\nu = 1$. For the sphere the values are slightly smaller, but still compatible with one. Combining this information we use in further analyses

$$K_c = 1.0234 \pm 0.0002 \quad (a = 0.0), \text{ Torus}, \quad (18)$$

$$K_c = 1.0230 \pm 0.0010 \quad (a = 0.0), \text{ Sphere}, \quad (19)$$

$$K_c = 1.0265 \pm 0.0001 \quad (a = 0.001), \text{ Torus}, \quad (20)$$

$$K_c = 1.0295 \pm 0.0001 \quad (a = 0.1), \text{ Torus}. \quad (21)$$

In particular we can now test the consistency of our data and extract ν also from the scaling of dU/dK and $d \ln \langle |m|^p \rangle / dK$ at K_c , see Table 1.

To extract the critical exponent ratio γ/ν we used the scaling $\chi \cong L^{\gamma/\nu} f_3(x)$ at the previously discussed points of constant x , as well as the

Table 1. Comparison of our Monte Carlo results on the torus and on the sphere with the exact results for the Ising model on static lattices (Onsager) and the KPZ exponents. The values marked with a star were computed from hyperscaling relations with $D = 2$, thereby neglecting possible scaling effects due to the internal fractal dimension in the DTRS approach.

	α	β	γ	δ	η	ν
KPZ	-1	0.5	2	5	2/3*	1.5*
Onsager	0	0.125	1.75	15	0.25	1
Torus- dl/l	≈ 0	0.126(2)	1.75(2)	14.9(3)	0.272(3)	1.01(1)
Sphere- dl/l	≈ 0	0.130(26)	1.61(12)	14(3)	0.256(6)	0.93(5)
Sphere-DeWitt	≈ 0	0.12(1)	1.75(1)	15(2)	0.25(1)	1.00(1)

scaling of χ' at K_c . The values for γ/ν for the different values of a are compatible with each other, but are all slightly below the Onsager value of $\gamma/\nu = 1.75$. Due to their respective error range, however, they are still consistent with the flat space exponent ratio. The quality of the fits for $\chi_{\max}$ on the torus can be inspected in Fig. 2, and the final values for γ , inserting our previously determined value for ν , can be found also in Table 1. For the sphere a weighted fit over all values gave $\gamma/\nu = 1.744(6)$.

To extract the magnetical critical exponent ratio β/ν we used that $\langle |m| \rangle \cong L^{-\beta/\nu} f_4(x)$ at all constant x -values. Another method is to look at the scaling of $d\langle |m| \rangle/dK \cong L^{(1-\beta)/\nu} f_5(x)$. Because the errors on the different estimates turned out to vary over a large range we chose to compute error-weighted averages. Using our average values for ν in Table 1 we obtain the final estimates of $\beta/\nu = 0.127(3)$ (torus, $a = 0.001$), $\beta/\nu = 0.123(2)$ (torus, $a = 0.1$), $\beta/\nu = 0.123(4)$ (torus, $a = 0.0$), and $\beta/\nu = 0.14(2)$ (sphere, $a = 0.0$). Again we see little influence of the curvature square term, and the results are again in agreement with the Onsager result $\beta/\nu = 0.125$.

For a specific-heat exponent α of zero we expect a logarithmic divergence like

$$C(x, L) = A(x) + B(x) \ln L. \quad (22)$$

Indeed the data at the different fixed values of x could all be fitted nicely with this ansatz, supporting again the Onsager value. We also performed simulations using a lattice transcription of the DeWitt measure according to Eq. (12) with $\alpha = 1$, $\sigma = -1/4$, which did not change the results at all [8], and some results from a preliminary data analysis can be found in Table 1 as well,

where we also used the scaling relations $\eta = 2 - \gamma/\nu$, and $\delta = 1 + \gamma/\beta$. The short summary of our findings is that the critical exponents still agree with the Onsager exponents for regular static lattices to a high degree of accuracy, and KPZ exponents are definitely excluded.

For the sphere, the non-regular triangulation seems to affect the finite-size behavior in a negative way, and one could probably obtain more accurate results comparable to those on the torus, by using a random triangulation of the sphere [21,22]. Unlike in the pure gravity case, where the global lattice topology enters in the formula for the string susceptibility exponent [23], it does not affect the critical exponents of the Ising phase transition.

5 Conclusions

Using a highly efficient cluster update algorithm and advanced reweighting techniques, our results have shown that the model defined by (7), describing Ising spins coupled to quantum gravity, remained in the Onsager universality class. This statement still holds if the global lattice topology was changed from the torus to the sphere. We conclude therefore that the global topology does not play any role for the Ising critical exponents, as it does for example for the string susceptibility exponent γ_{str} . We have also tested two local lattice measures, the dl/l and DeWitt measure, and saw no effect on the Ising transition. We have also found no influence of an added curvature square term R^2 , as one would expect from dimensional arguments. Overall we can conclude that neither the global topology nor the change of the local measure or the added R^2 interaction term change the critical exponents of the Ising system coupled to gravity via Eq. (9). The KPZ exponents are definitely excluded. We have further shown that one can use the Ising system ($c = 1/2$) as a probe to test KPZ scaling because there the FSS analyses are standard and give very accurate results.

Unfortunately the present situation is unsatisfactory, because one needs to explain why different discrete approaches to a "simple" 2D quantum gravity model lead to different results. After all, both models were supposed to describe the same continuum physics. There are still various possibilities to explain this:

- one needs to use a discrete nonlocal measure,
- the spin coupling to gravity is not correct,
- KPZ results are due to the fluctuating incidence matrix of the lattice,

- KPZ results are due to Euclidean gravity.

The first point was advocated strongly by Menotti and others [24], however, up to today nobody has shown that with a non-local measure one is able to get the KPZ exponents using the Regge calculus formulation. There are also calculations which show [25] that the proposed non-local lattice measures fail to agree with their continuum counterparts already in the weak field, low momentum limit, hence are not acceptable discrete functional measures. The second point relies on the fact that the coupling of the Ising spins to gravity in the Regge formulation was only heuristically written down in Eq. (9). A more profound investigation might lead to a different lattice implementation, which in turn could change the critical behavior. However, nobody has yet investigated this possibility. The third point suggests that the Regge method, because it is based on a fixed incidence matrix, cannot capture the necessary randomness in the coordination number, to induce a change of universality class. This randomness might, however, be of such an importance only in two dimensions, because here the Einstein-Hilbert term is trivial, and all what is left is the freedom of the incidence matrix to rearrange. In higher dimensions one can assume that the propagation modes of the Einstein-Hilbert term will be dominant so that this possible drawback of Regge calculus becomes unimportant. It is nevertheless unclear, why a purely field theoretic continuum Lagrangian like the one in Ref. [1] should give different results from the Regge method which is designed to approximate just the same model.

The last point goes back to suggestions by Ambjørn *et al.* [26] who showed that even for the DTRS method the critical exponents actually remain in the Onsager universality class if one uses a Lorentzian gravity formulation instead of an Euclidean one. The Lorentzian space-time structure and its causal requirement seems to be so stringent that it smoothes the possible randomness, thereby leaving the Ising system in its flat-space class. Regge calculus was actually designed to work in a Lorentzian geometry, but this still does not explain why it apparently is not able to describe well Euclidean geometry. This might again be related to the lack of the necessary randomness due to the fixed incidence matrix.

We still think that the question why the Regge method shows no effect on the Ising transition has to be resolved completely, because in one way or the other, we will learn lessons for future studies of a more realistic theory of quantum gravity. Even though two-dimensional gravity is, classically spoken, rather trivial, these questions await still their final answer.

Acknowledgments

The work described here was performed in a pleasant and close collaboration with W. Janke. I would like to thank H. Kleinert for enjoyable discussions and for providing a stimulating environment during my time in Berlin. I sincerely wish him many more fruitful years to come.

References

- [1] V.G. Knizhnik, A.M. Polyakov, and A.B. Zamolodchikov, *Mod. Phys. Lett.* **A3**, 819 (1988).
- [2] P. Di Francesco, P. Ginsparg, and J. Zinn-Justin, *Phys. Rep.* **254**, 1 (1995); P. Ginsparg, in *Proceedings of the 1991 Summer School in High Energy Physics and Cosmology*, ICTP Series in Theoretical Physics, Vol. 8, Ed. E. Gava *et al.* (World Scientific, Singapore, 1992).
- [3] V.A. Kazakov, *JETP Lett.* **44**, 133 (1986); *Phys. Lett. A* **119**, 140 (1986).
- [4] C. Itzykson and J.B. Zuber, *J. Math. Phys.* **21**, 411 (1980); M.L. Metha, *Comm. Math. Phys.* **79**, 327 (1981).
- [5] D.V. Boulatov and V.A. Kazakov, *Phys. Lett. B* **186**, 379 (1987).
- [6] M. Gross and H.W. Hamber, *Nucl. Phys. B* **364**, 703 (1991).
- [7] C. Holm and W. Janke, *Phys. Lett. B* **335**, 143 (1994); *Nucl. Phys. B* **42** (Proc. Suppl.), 725 (1995); *Phys. Lett. B* **375**, 69 (1996).
- [8] C. Holm and W. Janke, preprint.
- [9] T. Regge, *Nuovo Cim.* **19**, 558 (1961).
- [10] R.M. Williams and P.A. Tuckey, *Class. Quantum Grav.* **9**, 1409 (1992). This bibliographic review contains an almost complete list of references about Regge calculus up to 1991.
- [11] H.W. Hamber and R.M. Williams, *Nucl. Phys. B* **248**, 392 (1984); *Phys. Lett. B* **157**, 368 (1985); *Nucl. Phys. B* **267**, 482 (1986).
- [12] B.A. Berg, *Phys. Rev. Lett.* **55**, 904 (1985); *Phys. Lett. B* **176**, 39 (1986).
- [13] M. Rocek and R.M. Williams, *Phys. Lett. B* **104**, 31 (1981); *Z. Phys. C* **21**, 371 (1984).
- [14] H.W. Hamber, *Nucl. Phys. B* **400**, 347 (1993); H.W. Hamber and R.M. Williams, *Nucl. Phys. B* **435**, 361 (1995); W. Beierl, E. Gerstenmayer, H. Markum, and J. Riedler, *Phys. Rev. D* **49**, 5231 (1994); *Nucl. Phys. B* **30** (Proc. Suppl.), 764 (1993); W. Beierl, B.A. Berg, B. Krishnan, H.

- Markum, and J. Riedler, *Phys. Lett. B* **348**, 355 (1995); B.A. Berg and B. Krishnan, *Phys. Lett. B* **318**, 59 (1993).
- [15] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [16] H.W. Hamber, in *Proceedings of the 1984 Les Houches Summer School*, Session XLIII, Eds. K. Osterwalder and R. Stora (North Holland, Amsterdam, 1986).
- [17] U. Wolff, *Phys. Rev. Lett.* **62**, 361 (1989); *Nucl. Phys. B* **322**, 759 (1989).
- [18] A. Ferrenberg and R.H. Swendsen, *Phys. Rev. Lett.* **61**, 2635 (1988); *ibid.* **63**, 1658 (1989) (Erratum).
- [19] C. Holm and W. Janke, *Phys. Lett. A* **173**, 8 (1993); *Phys. Rev. B* **48**, 936 (1993).
- [20] K. Binder, *Z. Phys. B* **43**, 119 (1981).
- [21] W. Janke *et al.*, *Phys. Lett. B* **315**, 412 (1993); *Phys. Rev. B* **49**, 9644 (1994).
- [22] C. Holm and W. Janke, *Phys. Lett. B* **390**, 59 (1997); *Int. J. Mod. Phys. A* **14**, 3885 (1999).
- [23] C. Holm and W. Janke, *Nucl. Phys. B* **47** (Proc. Suppl.), 465, 621 (1996); *Nucl. Phys. B* **477**, 465 (1996).
- [24] P. Menotti and P.P. Peirano, *Phys. Lett. B* **353**, 444 (1995); J. Ambjørn, J.L. Nielsen, and J. Rolf, *Class. Quantum Grav.* **14**, 3225 (1997).
- [25] H.W. Hamber and R.M. Williams, *Phys. Rev. D* **59**, 064014 (1999); H.W. Hamber, *Phys. Rev. D* **61**, 124008 (2000).
- [26] J. Ambjørn and R. Loll, *Nucl. Phys. B* **536**, 407 (1998); J. Ambjørn, J. Correia, C. Kristjansen, and R. Loll, *Phys. Lett. B* **475**, 24 (2000).

GRAVITATIONAL EXCITONS – FLUCTUATING PARTICLES FROM EXTRA DIMENSIONS

U. GÜNTHER

*Gravitationsprojekt, Mathematische Physik I,
Institut für Mathematik, Universität Potsdam,
Am Neuen Palais 10, PF 601553, D-14415 Potsdam, Germany
E-mail: u.guenther@htw-zittau.de*

A. ZHUK

*Department of Physics, University of Odessa,
2 Petra Velikogo St., Odessa 270100, Ukraine
E-mail: zhuk@paco.odessa.ua*

We show that for warped product space-times the conformal (geometrical moduli) excitations of the internal compactified factor spaces should be observable as massive scalar fields in the external space-time. These scalar fields (gravitational excitons) describe weakly interacting particles and can be considered as dark matter component. On the other hand, they provide possible values for the effective cosmological constant.

1 Introduction

The multi-dimensionality of our Universe is one of the basic assumptions in modern theories beyond the $SU(3) \times SU(2) \times U(1)$ Standard Model of electroweak and strong interactions. Superstring theory and M-theory use this concept as basic assumptions and have a consistent formulation in space-times with total dimension $D = 10$ and $D = 11$.

The fundamental constants in these theories are related to the vacuum expectation values of the dilaton and moduli fields, and variations of these fields would result in variations of the constants. In the context of standard

Kaluza-Klein models, the moduli are defined by the shape and size of the internal spaces (geometrical moduli). Up to now, there are no experiments which show a time variation of fundamental constants. This means, that according to observations, the internal spaces should be static or nearly static at least from the time of recombination (in some papers, arguments are given in favor of the assumption that a variation of the fundamental constants is absent from the time of primordial nucleosynthesis). Therefore, a mechanism for moduli stabilization should be part of any realistic multi-dimensional model.

Within multi-dimensional cosmological models of the standard Kaluza-Klein type^a such a stabilization is achieved, e.g. via trapping of the geometrical moduli fields by effective potentials of dimensionally reduced models. On the other hand, it is important to consider possible observable consequences of various stabilization mechanisms. We show that fluctuations of the multi-dimensional geometry near minima of corresponding effective potentials should be observable as fluctuating scalar fields in our Universe, i.e. as scalar particles [1].

2 Conformal Fluctuations of Internal Spaces

We consider a cosmological model with the metric

$$g = g^{(0)} + \sum_{i=1}^n e^{2\beta^i(x)} g^{(i)}, \quad (1)$$

which is defined on a manifold with warped product topology

$$M = M_0 \times M_1 \times \cdots \times M_n, \quad (2)$$

where x are some coordinates of the $D_0 = (d_0 + 1)$ -dimensional external space-time manifold M_0 and

$$g^{(0)} = g_{\mu\nu}^{(0)}(x) dx^\mu \otimes dx^\nu. \quad (3)$$

Let the manifolds M_i be d_i -dimensional Einstein spaces with metric $g^{(i)}$, i.e.

$$R_{mn} [g^{(i)}] = \lambda^i g_{mn}^{(i)}, \quad m, n = 1, \dots, d_i, \quad R [g^{(i)}] = \lambda^i d_i \equiv R_i. \quad (4)$$

^aHere, it is assumed that the internal spaces are compactified at sizes somewhere between the Planck scale $L_{P1} \sim 10^{-33}$ cm and the Fermi scale $L_F \sim 10^{-17}$ cm to make them unobservable.

In the case of constant curvature spaces, the parameters λ^i are normalized as $\lambda^i = k_i(d_i - 1)$ with $k_i = \pm 1, 0$. Later on, we shall not specify the structure of the spaces M_i . We require only that M_i are compact spaces with arbitrary sign of curvature.

With total dimension $D = D_0 + \sum_{i=1}^n d_i$, κ_D^2 being a D -dimensional gravitational constant and Λ a D -dimensional cosmological constant, we consider an action functional of the form

$$S = \frac{1}{2\kappa_D^2} \int_M d^D x \sqrt{|g|} \{R[g] - 2\Lambda\} + S_m, \quad (5)$$

where S_m is a non-specified action term which takes into account additional matter fields. To illustrate the natural origin of gravitational excitons (gravexcitons for short) we shall consider a pure geometrical model: $S_m \equiv 0$. The generalization to models with explicit matter-terms is obvious. As an illustration we consider such a term (resulting from the Casimir effect) at the end of this section.

Let β_0^i be the scale of compactification of the internal spaces at the present time. Instead of β^i it is convenient to introduce a shifted quantity: $\tilde{\beta}^i = \beta^i - \beta_0^i$.

Then, after dimensional reduction and conformal transformation

$$g_{\mu\nu}^{(0)} = \Omega^2 \tilde{g}_{\mu\nu}^{(0)} := \left(\prod_{i=1}^n e^{d_i \tilde{\beta}^i} \right)^{-\frac{2}{D_0-2}} \tilde{g}_{\mu\nu}^{(0)} \quad (6)$$

action (5) reads

$$S = \frac{1}{2\kappa_0^2} \int_{M_0} d^{D_0} x \sqrt{|\tilde{g}^{(0)}|} \left\{ \tilde{R}[\tilde{g}^{(0)}] - \bar{G}_{ij} \tilde{g}^{(0)\mu\nu} \partial_\mu \tilde{\beta}^i \partial_\nu \tilde{\beta}^j - 2U_{\text{eff}} \right\}, \quad (7)$$

where $\tilde{R}_i := R_i e^{-2\beta_0^i}$, $\bar{G}_{ij} = d_i \delta_{ij} + d_i d_j / (D_0 - 2)$ is the midisuper-space metric and $\kappa_0^2 := \kappa_D^2 / V_{D'}$ denotes the D_0 -dimensional gravitational constant ($V_{D'}$ is the total volume of the internal space). If we take the TeV scale [2,3] M_{TeV} and the Planck scale M_{Pl} as fundamental ones for D -dimensional space-time and the 4-dimensional large-scale space-time, respectively: $\kappa_D^2 = 8\pi / M_{\text{TeV}}^{2+D'}$, and $\kappa_0^2 = 8\pi / M_{\text{Pl}}^2$, then we reproduce the well-known relation [2,3]: $M_{\text{Pl}}^2 = V_{D'} M_{\text{TeV}}^{(2+D')}$. This implies that the scale of the internal space compactification is fixed and of the order

$$a \sim V_{D'}^{1/D'} \sim 10^{32/D' - 17} \text{cm}. \quad (8)$$

The effective potential in (7) reads

$$U_{\text{eff}}[\tilde{\beta}] = \left(\prod_{i=1}^n e^{d_i \tilde{\beta}^i} \right)^{-\frac{2}{D_0-2}} \left[-\frac{1}{2} \sum_{i=1}^n \tilde{R}_i e^{-2\tilde{\beta}^i} + \Lambda \right]. \quad (9)$$

With the help of a regular coordinate transformation $\varphi = Q\tilde{\beta}$, $\tilde{\beta} = Q^{-1}\varphi$, the midisuperspace metric (target space metric) $\bar{G}$ can be transformed to a pure Euclidean form: $\bar{G}_{ij} d\tilde{\beta}^i \otimes d\tilde{\beta}^j = \sigma_{ij} d\varphi^i \otimes d\varphi^j = \sum_{i=1}^n d\varphi^i \otimes d\varphi^i$, $\sigma = \text{diag}(+1, +1, \dots, +1)$ (see e.g. Ref. [1]).

It is clear that a stabilization of the internal spaces can be achieved, if the effective potential U_{eff} has a minimum with respect to fields $\tilde{\beta}^i$ (or fields φ^i). In general, it is possible that the potential U_{eff} has more than one extremum. But it can be easily seen that, for the pure geometrical model under consideration, we can get one extremum only. This corresponds to $\tilde{\beta}^i = 0$. For the masses of the normal mode excitations of the internal spaces (gravitational excitons) around the minimum position we obtain:

$$m_1^2 = \dots = m_n^2 = -\frac{4\Lambda_{\text{eff}}}{D_0-2} = -2\frac{\tilde{R}_k}{d_k} > 0, \quad (10)$$

where

$$\Lambda_{\text{eff}} := U_{\text{eff}} \Big|_{\tilde{\beta}^i=0} \quad (11)$$

plays the role of an effective cosmological constant in the external space-time. These equations show that a global minimum can only exist for our specific model in the case of compact internal spaces with negative curvature $R_k < 0$ ($k = 1, \dots, n$). The effective cosmological constant is also negative: $\Lambda_{\text{eff}} < 0$. Models which include matter can have minima for internal spaces of positive curvature. Usually, the effective cosmological constant is positive in this case.

For small fluctuations of the normal modes in the vicinity of the minima of the effective potential action Eq. (7) reads

$$S = \frac{1}{2\kappa_0^2} \int_{M_0} d^{D_0}x \sqrt{|\tilde{g}^{(0)}|} \left\{ \tilde{R}[\tilde{g}^{(0)}] - 2\Lambda_{\text{eff}} \right\} - \frac{1}{2} \int_{M_0} d^{D_0}x \sqrt{|\tilde{g}^{(0)}|} \left\{ \sum_{i=1}^n \left(\tilde{g}^{(0)\mu\nu} \psi_{,\mu}^i \psi_{,\nu}^i + m_i^2 \psi^i \psi^i \right) \right\} \quad (12)$$

(for convenience we use here the normalizations: $\kappa_0^{-1}\tilde{\beta} \rightarrow \tilde{\beta}$). Thus, conformal excitations of the metric of the internal spaces behave as massive scalar fields developing on the background of the external space-time. In analogy to excitons in solid state physics, where they are excitations of the electronic subsystem of a crystal, we called the excitations of the subsystem of internal spaces gravitational excitons [1]. Later these particles have also become known as radions [2,3].

From Eq. (10) follows that

$$|\Lambda_{\text{eff}}| \sim m_i^2 \sim a_{(0)i}^{-2}, \quad (13)$$

where $a_{(0)i} = \exp \beta_0^i$ are the scale factors of the stabilized internal spaces.

The calculations above were performed for a model with the TeV scale M_{TeV} as fundamental scale of the D -dimensional theory (see Eq. (8)). Clearly, it is also possible to choose the Planck scale as the fundamental scale.

For this purpose we do not fix the compactification scale of the internal spaces at the present time. We consider them as free model parameters and demand only that $L_{\text{Pl}} < a_{(0)i} = e^{\beta_0^i} < L_F \sim 10^{-17}\text{cm}$. So, we shall not transform β^i to $\tilde{\beta}^i$. In this case, $\kappa_D^2 \sim M_{\text{Pl}}^{-(2+D')}$, so that the Planck scale becomes the fundamental scale of the D -dimensional theory. In this approach, Eqs. (6), (7), and (9) preserve their form with the only substitutions $\tilde{\beta} \rightarrow \beta$ and $\tilde{R}_i \rightarrow R_i$. The Einstein frame metrics of the external space-time in both approaches are equivalent to each other up to a numerical prefactor:

$$\tilde{g}_{\mu\nu}^{(0)} \Big|_{\text{TeV}} = v_0^{-2/(D_0-2)} \tilde{g}_{\mu\nu}^{(0)} \Big|_{\text{Pl}}, \quad (14)$$

where $v_0 = \prod_{i=1}^n \exp(d_1 \beta_0^i)$. Obviously, the same rescaling takes place for the squared masses of the gravitational excitons and the effective cosmological constant: $m_i^2 \rightarrow (v_0)^{-2/(D_0-2)} m_i^2$ and $\Lambda_{\text{eff}} \rightarrow (v_0)^{-2/(D_0-2)} \Lambda_{\text{eff}}$. Thus, in the latter approach we get, instead of (13), the relation:

$$|\Lambda_{\text{eff}}| \sim m_i^2 \sim (a_{(0)i})^{-(D-2)}, \quad (15)$$

where we set $D_0 = 4$. This relation shows that, due to the power $(2-D)$, the effective cosmological constant and the masses of the gravitational excitons can be very far from Planckian values even for scales of compactification of the internal spaces close to the Planck length.

Let us return to the comparison of the TeV scale and the Planck scale approaches. If we set, e.g. $6 \leq D < \infty$ in the TeV-scale approach, then the internal space scale factors, the gravexciton masses and the effective cosmological constant run correspondingly as: $10^{-1}\text{cm} \leq a_{(0)i} < 10^{-17}\text{cm}$, $10^{-4}\text{eV} \leq m_i < 1\text{TeV}$ and $10^{-64}\Lambda_{\text{Pl}} \leq |\Lambda_{\text{eff}}| \leq 10^{-32}\Lambda_{\text{Pl}}$. For this approach, the internal space scale factors are defined by Eq. (8) due to the demand that the D -dimensional gravitational constant is of order of the TeV scale. In the Planck scale approach such a condition is absent and the scale factors $a_{(0)i}$ are free parameters. Let us take, e.g. $a_{(0)i} \sim 10^{-18}\text{cm}$. Then, in the Planck scale approach for $6 \leq D \leq 10$, the gravexciton masses and effective cosmological constant run correspondingly as $10^{-2}\text{eV} \leq m_i \leq 10^{-32}\text{eV}$ and $10^{-60}\Lambda_{\text{Pl}} \leq |\Lambda_{\text{eff}}| \leq 10^{-120}\Lambda_{\text{Pl}}$.

These estimates show that, within the TeV scale approach, the effective cosmological constant is much greater than the present day observable limit $\Lambda \leq 10^{-122}\Lambda_{\text{Pl}} \sim 10^{-57}\text{cm}^{-2}$ (for our model $|\Lambda_{\text{eff}}|_{\text{TeV}} \geq 10^2\text{cm}^{-2}$), whereas in the Planck scale approach we can satisfy this limit even for very small compactification scales. For example, if we require $|\Lambda_{\text{eff}}| \sim 10^{-122}\Lambda_{\text{Pl}}$ in accordance with observations, then Eq. (15) gives a compactification scale $a_{(0)1} \sim 10^{122/(D-2)}L_{\text{Pl}}$. Thus, $a_{(0)1} \sim 10^{15}L_{\text{Pl}} \sim 10^{-18}\text{cm}$ for $D = 10$ and $a_{(0)1} \sim 10^5L_{\text{Pl}} \sim 10^{-28}\text{cm}$ for $D = 26$, which does not contradict observations because, for this approach, the scales of compactification should be $a_{(0)1} \leq 10^{-17}\text{cm}$. Assuming an estimate $\Lambda_{\text{eff}} \sim 10^{-122}L_{\text{Pl}}$, we automatically get from Eq. (15) the value of the gravitational exciton mass: $m_1 \sim 10^{-61}M_{\text{Pl}} \sim 10^{-33}\text{eV} \sim 10^{-66}\text{g}$, i.e. the gravexcitons are in this case extremely light particles. Nevertheless, such light particles are not in contradiction with the observable Universe, because they do not overclose the Universe [4].

As shown above, the effective cosmological constant is negative in the pure geometrical case. However, according to the modern observation $\Lambda_{\text{eff}} > 0$. This problem can be solved if we take into account matter fields. In many important cases matter can be described phenomenologically via a perfect fluid ansatz. In this case the effective potential reads [1]

$$U_{\text{eff}}[\tilde{\beta}] = \left(\prod_{i=1}^n e^{d_i \tilde{\beta}^i} \right)^{-\frac{2}{D_0-2}} \left[-\frac{1}{2} \sum_{i=1}^n \tilde{R}_i e^{-2\tilde{\beta}^i} + \Lambda + \kappa_D^2 \rho \right], \quad (16)$$

where ρ is the energy density of the perfect fluid. A number of effective potentials of this type ensuring stability were described in Refs. [1,4–6]. Among

them, the Casimir potential is one of the most important [5]. The Casimir effect is connected with the vacuum polarization of quantized fields due to the non-trivial topology of the background space or the presence of boundaries in the space. As a result, one obtains a nonvanishing energy density of the quantized fields in the vacuum state. In our case, this phenomenon should take place due to the compactness of the internal spaces. For one compact internal space ($n = 1$) with compactification scale $a_1 \ll a_0$, the Casimir energy density reads

$$\rho = C \exp(-D\beta^1) \equiv \tilde{C} \exp(-D\tilde{\beta}^1). \quad (17)$$

Here C is a constant which strongly depends on the topology of the model. The equations of state in the external and internal spaces read, respectively [7]:

$$P_0 = -\rho, \quad P_1 = \frac{D_0}{d_1} \rho. \quad (18)$$

It can easily be seen that the effective potential (16) with the Casimir energy density (17) can have a non-negative minimum (non-negative effective cosmological constant) [5].

Conclusion

From the geometrical point of view it is clear that gravitational excitons are an inevitable consequence of the existence of extra dimensions. For any theory with compactified internal spaces, conformal excitations (fluctuations) of the internal space metric will result in gravitational excitons in the external space-time. The form of the effective potential as well as masses of gravitational excitons and the value of the effective cosmological constant are strongly model dependent. Gravexcitons may play an important role in cosmology. On the one hand, they are connected with the observed cosmological constant via the effective potential. On the other hand, the interaction between gravexcitons and usual matter is Planck-scale suppressed and they can give a significant contribution to Dark Matter [8].

References

- [1] U. Günther and A. Zhuk, *Phys. Rev. D* **56**, 6391 (1997), eprint: gr-qc/9706050.

- [2] N. Arkani-Hamed, S. Dimopoulos, and J. March-Russell, eprint: hep-th/9809124.
- [3] N. Arkani-Hamed, S. Dimopoulos, N. Kaloper, and J. March-Russell, *Nucl. Phys. B* **567**, 189 (2000), eprint: hep-ph/9903224.
- [4] U. Günther and A. Zhuk, *Phys. Rev. D* **61**, 124001 (2000), eprint: hep-ph/0002009.
- [5] U. Günther and A. Zhuk, in Proc. Workshop *Modern Modified Theories of Gravitation and Cosmology*, Beer Sheva, Israel, June 29 - 30, 1997; *Hadronic Journal* **21**, 279 (1998), eprint: gr-qc/9710086; U. Günther, S. Krisikiv, and A. Zhuk, *Grav. and Cosm.* **4**, 1 (1998), eprint: gr-qc/9801013.
- [6] U. Günther and A. Zhuk, *Class. Quant. Grav.* **15**, 2025 (1998), eprint: gr-qc/9804018.
- [7] H. Kleinert and A. Zhuk, *Theor. Math. Phys.* **108**, 1236 (1996); *Theor. Math. Phys.* **109**, 1483 (1996).
- [8] U. Günther and A. Zhuk, in Proc. Memorial International Conference (GMIC'99) *The Universe of Gamow: Original Ideas in Astrophysics and Cosmology*, Odessa, Ukraine, August 16-22, 1999, *Odessa Astr. Pubs.* **12**, 37 (1999), <http://oap12.webjump.com>; U. Günther, A. Starobinsky, and A. Zhuk, *Interacting Gravitational Excitons from Extra Dimensions*, preprint.

PATH INTEGRALS IN QUANTUM COSMOLOGY

C. KIEFER

*Fakultät für Physik, Universität Freiburg,
Hermann-Herder-Str. 3, 79104 Freiburg, Germany
E-mail: claus.kiefer@physik.uni-freiburg.de*

Path integrals in quantum cosmology differ both in formulation and interpretation from ordinary path integrals. After reviewing the highlights of these differences, I shall discuss in detail a particular model with a cosmological constant. This example contains some interesting features connected with space-time transformations in the path integral.

1 Introduction

Path integrals play a prominent role in modern physics [1]. Their application ranges from quantum mechanics, statistical mechanics up to quantum field theory. It thus seems natural to apply path integration also to quantum gravity, a theory being still in the process of construction.

In this context, interest focuses mainly on two issues. First, path integrals are applied to the full gravitational field in order to find a nonperturbative formulation of quantum gravity. This is also a convenient starting point for lattice formulations. Second, path integrals are frequently formulated for finite-dimensional models of quantum cosmology in order to study features of the early Universe, in particular in connection with an inflationary phase. Here, applications are usually made for an energy regime somewhat lower than the Planck scale. This gives rise to the hope that the results are independent of the unknown behavior of the full theory.

This contribution to Hagen Kleinert's *Festschrift* focuses on the second application. In the next section I will review the main properties of a quantum cosmological path integral and discuss some results for a simple model – the indefinite harmonic oscillator. Interesting consequences are drawn for the

meaning of the *no-boundary condition* in quantum cosmology. The third section contains a model that has not yet been studied in the present context – a Friedmann model with a cosmological constant. In the evaluation of the path integral, it is necessary to take into account the curved nature of the configuration space. Some technical details are relegated to the Appendix.

2 General Properties of Quantum Cosmological Path Integrals

In quantum cosmology one focuses on finite-dimensional models whose Hamiltonian is given by

$$H = \frac{1}{2} G^{ab}(q) p_a p_b + V(q) , \quad (1)$$

where q is a shorthand notation for n degrees of freedom. These can be the scale factor(s) of a cosmological model plus homogeneous matter variables. Classically, this Hamiltonian is constrained to vanish as a consequence of the reparametrization invariance of general relativity. The coefficients G^{ab} denote the (inverse) metric on configuration space (DeWitt metric or superspace metric). For a discussion of its features see e.g. Ref. [2]. In the canonical version of quantum gravity, the classical constraint is turned into a condition on allowed wave functions, the so-called Wheeler-DeWitt equation

$$\hat{H}\psi(q) = 0 . \quad (2)$$

The corresponding path integral has to be formulated within the theory of constrained systems, i.e. including gauge fixing and Faddeev-Popov ghosts. In the present finite-dimensional case, the procedure can be highly simplified and leads to the following path integral (for a review see e.g. Ref. [3])

$$\begin{aligned} G(q', q'') &= \int dT \int Dp_a Dq^a \exp \left(i \int_0^T dt (p_a \dot{q}^a - H) \right) \\ &= \int dT \langle q'', T | q', 0 \rangle , \end{aligned} \quad (3)$$

where $\langle q'', T | q', 0 \rangle$ is an ordinary quantum mechanical propagator which satisfies the Schrödinger equation. Since quantum gravity does not contain any external time parameter, time must be “integrated out” to get the correct path integral. The range of the T -integration in Eq. (3) is not fixed a priori. Since integration along the real axis leads in general to a divergent result, the idea is to look for contours in the *complex* T -plane that render the integral

convergent [3,4]. In order to prevent a violation of gauge invariance, attention is restricted to either infinite contours, half-infinite contours starting at zero, or closed loops around the origin [3]. Infinite contours and closed loops should then lead to *solutions* of Eq. (2), while half-infinite contours should yield *Green functions*.

Due to the T -integration in Eq. (3), path integrals in quantum cosmology have different properties than quantum mechanical path integrals [5]. In fact, they behave more like *energy Green functions* [1] rather than propagators, as was first emphasised by Hájíček [6]. For this reason, they do not obey any composition law, which means that they are not intimately connected with an external time variable. A simpler but similar analogy for the path integral in Eq. (3) is provided by the proper-time representation for the relativistic particle [5]. Due to these properties of the path integral, the central interest focused on the relation between *boundary conditions* and the choice of metrics to be integrated over [7]. The famous “no-boundary proposal” by Hartle and Hawking [8] aimed at finding a *unique* solution to the Wheeler-DeWitt equation (2) by integrating only over *compact* (“finite”) geometries. This should have yielded the “wave function of the Universe” (for a critical assessment concerning physical interpretations of the no-boundary wave function see, e.g. Ref. [10]), but it did not. Even when integrating over complex metrics corresponding to the complex T -integration above, one gets a whole *class* of solutions to Eq. (2).

Still, it is of interest to investigate this distinguished class of solutions in more detail. The question posed in Refs. [5,9], for example, was: Can one get special cases which reflect the properties of classical solutions from the class of no-boundary wave functions, i.e. can one construct wave packets that follow a classical trajectory in configuration space?

The simplest nontrivial model to investigate this question is the *indefinite oscillator* which has the advantage of being exactly soluble. The Wheeler-DeWitt equation (2) then reads

$$\hat{H}\psi \equiv \frac{1}{2} \left(\frac{\partial^2}{\partial a^2} - \frac{\partial^2}{\partial \chi^2} - a^2 + \chi^2 \right) \psi(a, \chi) = 0, \quad (4)$$

where a denotes the scale factor of a closed Friedmann Universe, and χ a (rescaled) conformally coupled field. The classical solutions of this model would consist of trajectories (“Lissajous ellipses”) that are confined to a rectangle around the origin. The interesting point is that the path integral (3) can be exactly evaluated for this model. The path integration over Da and

$D\chi$ gives the usual result for the harmonic oscillator [1], but with the relative sign between the two oscillators reversed. The remaining T -integration then reads

$$G(a'', \chi''; a', \chi') = \frac{1}{2\pi} \int \frac{dT}{\sin T} \exp \left(\frac{i}{2 \sin T} (Q_1 \cos T + Q_2) \right), \quad (5)$$

with

$$Q_1 \equiv \chi''^2 + \chi'^2 - (a''^2 + a'^2), \quad (6)$$

$$Q_2 \equiv -2\chi'\chi'' + 2a'a''. \quad (7)$$

One can interpret the no-boundary proposal as imposing the values $a' = \chi' = 0$ and looking for contours in the T -plane that render the integral convergent. It turns out [5] that for half-infinite contours the result of Eq. (5) is

$$G_1(a, \chi) = \frac{1}{2\pi} K_0 \left(\frac{|a^2 - \chi^2|}{2} \right), \quad (8)$$

where K_0 denotes the modified Bessel function, while for an infinite contour the result is

$$G_2(a, \chi) = \frac{1}{2} I_0 \left(\frac{a^2 - \chi^2}{2} \right), \quad (9)$$

with the Bessel function I_0 . While K_0 is a *fundamental solution* to Eq. (4), I_0 is its associated *Riemann function*. Other convergent contours only yield linear combinations of these functions. Inspecting the asymptotic behavior of G_1 and G_2 , one recognises that G_1 diverges along the “lightcone” in the (a, χ) -space, while G_2 diverges exponentially for large arguments. Therefore, neither of these two solutions can be used to construct wave packets following the classical solutions. This is an important result, since it demonstrates that the relation between no-boundary solutions and the classical theory is very loose. My conjecture is that these properties hold in general. The asymptotic behavior does not change if the condition $a' = \chi' = 0$ is relaxed.

The path integral, when evaluated exactly, does therefore exhibit properties that cannot necessarily be seen in a semiclassical approximation. This has consequences, for example, for the discussion of the arrow of time in a recollapsing quantum Universe [11].

3 Quantum Cosmological Path Integral for a Model with a Cosmological Constant

The purpose of this section is to discuss a quantum cosmological model which exhibits certain nontrivial features with respect to path integration in curvilinear coordinates. Configuration spaces with curvature (and torsion) find fruitful application, for example, in quantum mechanics with a Coulomb potential [1,12]. For our purpose it is sufficient to consider a one-dimensional model in which the Wheeler-DeWitt equation (2) is given by

$$\hat{H}\psi(\alpha) = \frac{1}{2} \left(\frac{d^2}{d\alpha^2} - e^{4\alpha} + \lambda^2 e^{8\alpha} \right) \psi(\alpha) = 0. \quad (10)$$

Here, $\alpha = \ln a$, where a is the scale factor, and λ^2 denotes a positive cosmological constant. It should be noted that the term arising from general relativity would read $\lambda^2 e^{6\alpha}$; however, the choice in Eq. (10) enables an exact evaluation of the path integral, without changing the qualitative features of the realistic model. It is also possible to obtain the potential used in Eq. (10) from general relativity, but this would require a very unrealistic equation of state [13].^a

Due to the presence of the complicated potential in Eq. (10), a direct evaluation of the path integral (3) is impossible. One can, however, try to perform a space-time transformation in the path integral to cope with this problem. Such a procedure has turned out to be very useful for evaluating path integrals in situations where a direct calculation is unfeasible. This has been elaborated in Ref. [12], where the hydrogen atom has been mapped into a harmonic-oscillator system, which is why it is now called the Duru-Kleinert transformation. Several other systems have been investigated by this method [1,14,15]. A particular example is the case of the Morse potential which can be mapped into the system of a harmonic oscillator with a centrifugal barrier [1,15]. In fact, the potential in our model (10) is similar to the Morse potential, which is why an analogous transformation can be applied.

Because Eq. (10) resembles the case of Liouville quantum mechanics, the transformation

$$q = e^{2\alpha} \Leftrightarrow \alpha = \frac{1}{2} \ln q \equiv F(q) \quad (11)$$

^aI am grateful to Alexander Zhuk for pointing this out to me.

is convenient. Thus one finds for the Hamiltonian

$$\hat{H} = 2 \left(q^2 \frac{d^2}{dq^2} + q \frac{d}{dq} \right) - \frac{q^2}{2} + \frac{\lambda^2 q^4}{2} . \quad (12)$$

The next step is to introduce a time-transformation in the $D\alpha$ part of the path integral (which is an ordinary quantum mechanical path integral) according to

$$dt = [F'(q)]^2 d\tau = \frac{d\tau}{4q^2(\tau)} , \quad (13)$$

with $\tau(T) \equiv s$. This simplifies both the kinetic and the potential term. One arrives at the new Hamiltonian

$$\tilde{H} = [F'(q)]^2 \hat{H} = \frac{1}{2} \left(\frac{d^2}{dq^2} + \frac{1}{q} \frac{d}{dq} \right) - \frac{1}{8} + \frac{\lambda^2 q^2}{8} . \quad (14)$$

It is shown in the Appendix that the implementation of all these transformations in the path integral leads to the appearance of a quantum correction

$$\Delta V = \frac{1}{8q^2} \quad (15)$$

to the potential. It is called *quantum* because it is proportional to $\hbar^2$ (here set equal to one). The correction appears because the lattice definition of the path integral requires the use of a certain ordering prescription for coordinates and momenta [1], which is given in the present case by Weyl ordering and midpoint prescription (see the Appendix). It should be emphasised, however, that this quantum correction to the potential is a *formal* correction which is needed for a correct evaluation of the path integral; there is *no* correction whatsoever to the potential in the Wheeler-DeWitt equation (10). In fact, the path integral is used to gain solutions (or Green functions) to Eq. (10).

The path integral can then be written as

$$G(q'', q') = \frac{1}{2\sqrt{q'q''}} \int ds \langle q'', s; q', 0 \rangle = \frac{1}{2\sqrt{q'q''}} \times \\ \int ds \int_{q(0)=q'}^{q(s)=q''} Dq(\tau) \exp \left(\frac{is}{8} + i \int_0^s \left[-\frac{\dot{q}^2}{2} - \frac{1}{8q^2} - \frac{\lambda^2 q^2}{8} \right] d\tau \right) . \quad (16)$$

The origin of the factor $1/2\sqrt{q'q''}$ lies in the demand for covariance with respect to point canonical transformations [1]. An exact expression for

$\langle q'', s; q', 0 \rangle$ can be given if one recalls that the path integral corresponding to the Lagrangian

$$L(x, \dot{x}) = \frac{\dot{x}^2}{2} - \frac{\omega^2 x^2}{2} - \frac{g}{x^2} \quad (17)$$

can be expressed in closed form [16]. The present case corresponds to the choice $g = -1/8$ and $\omega^2 = \lambda^2/4$, remembering that the kinetic term in our model is negative definite. It is interesting that this value for g corresponds to the limiting case in quantum mechanics, where a particle can fall into the centre under the influence of a radial potential. We thus have

$$\langle q'', s; q', 0 \rangle = \frac{i\lambda\sqrt{q'q''}}{2\sinh(\lambda s/2)} \exp\left(\frac{is}{8} - \frac{i\lambda}{4}[q'^2 + q''^2]\coth\frac{\lambda s}{2}\right) J_0\left(\frac{-\lambda q'q''}{2\sinh(\lambda s/2)}\right), \quad (18)$$

where J_0 denotes a Bessel function.

For $\lambda \neq 0$, the classical solutions are constrained to regions $\alpha > \alpha_0$, where $\alpha_0 = -(1/2)\ln\lambda$ is the zero point of the potential (there exists also the possibility of the classical system to sit at $\alpha = -\infty$).

Turning to the evaluation of the path integral (16), the integral to be evaluated reads, taking into account Eq. (18),

$$G(q'', q') = i \int \frac{d\tau}{2\sinh\tau} \exp\left(\frac{i\tau}{4\lambda} - \frac{i\lambda}{4}[q'^2 + q''^2]\coth\tau\right) J_0\left(\frac{-\lambda q'q''}{2\sinh\tau}\right), \quad (19)$$

where $\tau = \lambda s/2$ is introduced as the new integration variable. This integral can easily be evaluated by a steepest-descent approximation, similar to the discussion in Ref. [5]. One thereby finds saddle points contributing either a factor $e^{-c/\lambda}$ (saddle points in the upper half plane) or a factor $e^{c/\lambda}$ (saddle points in the lower half plane), where c is positive. If λ were dynamical, this would lead to a "peak" of the wave function either at zero or infinite cosmological constant. Consider now, for the case $\lambda q > 1$, the contour parallel to the real axis of Fig. 1. It can be deformed into a contour (shown as a dashed line) which receives its dominant contributions from the steepest-descent paths through the saddle points C and D. The integral (19) can, however, be evaluated *exactly* along this contour for any value of q' (not necessarily no-boundary conditions). Writing $\tau = \rho + i\pi/2$, one finds

$$G(q'', q') = \frac{1}{2} \exp\left(-\frac{\pi}{8\lambda}\right) \times$$

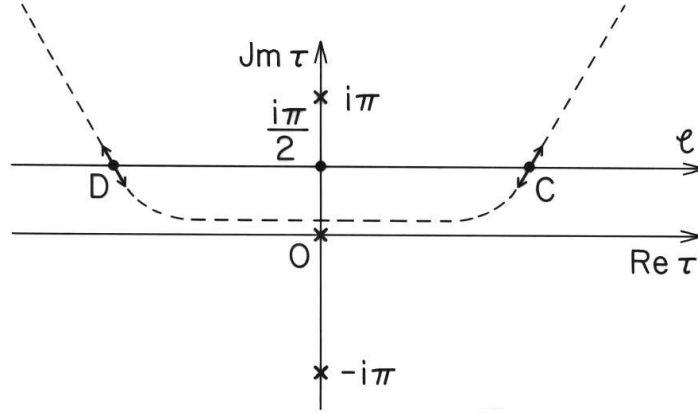


Figure 1. The exact integration contour (solid line) can be deformed into a contour (dashed line) which receives its dominant contributions from steepest-descent contours through the saddle points C and D.

$$\int_{-\infty}^{\infty} \frac{d\rho}{\cosh \rho} \exp\left(\frac{i\rho}{4\lambda} - \frac{i\lambda}{4}[q'^2 + q''^2] \tanh \rho\right) J_0\left(\frac{i\lambda q' q''}{2 \cosh \rho}\right), \quad (20)$$

which yields the result [17]

$$G(q'', q') = \frac{-i\pi \exp(-\pi/8\lambda)}{\lambda q' q'' \cosh(\pi/8\lambda)} M_{i/8\lambda, 0}\left(\frac{i\lambda q'^2}{2}\right) M_{i/8\lambda, 0}\left(\frac{i\lambda q''^2}{2}\right), \quad (21)$$

where M denotes the Whittaker functions. One can easily see that Eq. (21) satisfies the Wheeler-DeWitt equation (10) for both arguments. The expression in Eq. (21) separates in q' and q'' , so there is not much loss of generality in discussing only the case $q' = 0$. In this limit, Eq. (21) becomes

$$G(q) = \frac{\pi \exp(-\pi/8\lambda)}{2 \cosh(\pi/8\lambda)} \exp(-i\lambda q^2/4) M\left(\frac{1}{2} - \frac{i}{8\lambda}, 1, \frac{i\lambda q^2}{2}\right), \quad (22)$$

with

$$M_{\kappa, \mu}(z) = e^{-z/2} z^{1/2+\mu} M\left(\frac{1}{2} + \mu - \kappa, 1 + 2\mu, z\right), \quad (23)$$

which relates Whittaker's function to Kummer's function. It is not surprising that the above expressions resemble the expressions for the energy Green

function in the quantum mechanical case of the Morse potential [15].

Both the semiclassical approximation to Eq. (22) and the steepest-descent approximation to Eq. (19) yield the same result, namely

$$G(q) \stackrel{\lambda q \gg 1}{\sim} 2 \sqrt{\frac{\pi}{\lambda q^2}} \exp\left(-\frac{\pi}{8\lambda}\right) \cos\left(-\frac{\ln 2\lambda q}{4\lambda} + \frac{\lambda q^2}{4} - \frac{1}{8\lambda} - \frac{\pi}{4}\right). \quad (24)$$

The factor $\exp(-\pi/8\lambda)$ is the usual WKB penetration factor that *would* correspond to a particle tunnelling from $q = 1/\lambda$ to $q = 0$. The concept "tunnelling" must, however, be used with great care, since no tunnelling process is happening – due to the absence of any external time. This is a major difference between quantum mechanics and quantum cosmology [10,11].

Appendix

The origin of the problem with path integration on curved manifolds (or in curvilinear coordinates) lies in the very definition of the path integral, which is only valid in Cartesian coordinates. The occurrence of a quantum correction ΔV to the potential is directly connected with the ordering prescription of the lattice coordinates and momenta in the action, which in turn is related to operator ordering in the Hamiltonian [1].

For a quantum mechanical system whose classical counterpart is given by the Lagrangian

$$L(q, \dot{q}) = \frac{m}{2} g_{ab}(q) \dot{q}^a \dot{q}^b - V(q), \quad (25)$$

the corresponding Hamilton operator is Weyl ordered, if the *midpoint* prescription is used in the path integral [18], i.e. if the lattice action is taken to be of the form

$$S = \sum_{k=1}^N \left(\frac{m}{2\epsilon} g_{ab}(\bar{q}_k) (q_k^a - q_{k-1}^a) (q_k^b - q_{k-1}^b) - \epsilon V(\bar{q}_k) \right), \quad (26)$$

where

$$\bar{q}_k \equiv \frac{q_k + q_{k-1}}{2}. \quad (27)$$

The Hamilton operator is Weyl ordered, if it is of the form

$$H_W = \frac{1}{8m} (g^{ab} p_a p_b + 2p_a g^{ab} p_b + p_a p_b g^{ab}) + V(q), \quad (28)$$

and it is obtained by randomly ordering all coordinates and momenta, counting all different orderings once and forming their arithmetic mean.

The natural ordering for the Hamiltonian, however, consists in the choice of the Laplace-Beltrami operator Δ_{LB} for the kinetic term, because this preserves the covariance of the theory. In fact, this is what one arrives at, when performing a transformation from Cartesian to curvilinear coordinates. We thus consider the Hamiltonian

$$H = -\frac{1}{2}\Delta_{LB} + V(q) = \frac{1}{2}g^{-1/4}p_a g^{ab} g^{1/2}p_b g^{-1/4} + V(q) = H_W + \Delta V, \quad (29)$$

where g is the determinant of the metric g_{ab} . To evaluate the matrix element $\langle q' | \exp(-iHt) | q \rangle$ with H according to Eq. (29), we thus have to take into account the correction ΔV in the action (26). The canonical momenta are given by

$$p_a = \frac{1}{i} \left(\frac{\partial}{\partial q^a} + \frac{1}{2} \frac{\partial \ln \sqrt{g}}{\partial q^a} \right), \quad (30)$$

because they are in this form – as well as the Hamiltonian (29) – self-adjoint with respect to the inner product

$$\langle \varphi | \psi \rangle = \int dq \sqrt{g} \varphi^*(q) \psi(q). \quad (31)$$

By comparing Eq. (28) and Eq. (29), one finds an explicit expression for ΔV ,

$$\Delta V = \frac{1}{8m} (g^{ac} \Gamma_{ac}^d \Gamma_{bd}^b - R), \quad (32)$$

where Γ_{ab}^c and R are the Christoffel symbols and the Ricci scalar of the configuration space, respectively.

In the quantum cosmological model discussed in Section 3, we can easily find ΔV directly. Starting from Eq. (14), we first recognise that due to the time transformation (13) the kinetic term of $\tilde{H}$ is not the Laplace-Beltrami operator with respect to any metric. The procedure described above has therefore to be slightly modified. The Hamiltonian $\tilde{H}$ is self-adjoint with respect to an inner product of the form (31), where $\sqrt{g}$, however, must be changed into

$$J(q) = \exp \left(\int \frac{dq}{q} \right) = q, \quad (33)$$

arising from the first-derivative term in Eq. (14).^b The canonical momentum then reads according to Eqs. (30) and (33)

$$p_q = -i \left(\frac{d}{dq} + \frac{1}{2q} \right), \quad (34)$$

leading to

$$\tilde{H} = -\frac{p_q^2}{2} + \frac{1}{8q^2} - \frac{1}{8} + \frac{\lambda^2 q^2}{8}. \quad (35)$$

Due to the trivial form of the second-derivative term in Eq. (14), this is already Weyl ordered, and ΔV can be directly read off, leading to the expression (15). Finally, I would like to mention that all these transformations can be implemented consistently into the measure of the path integral, leading in fact to the expression (16) with the correct prefactor.

Acknowledgments

I am grateful to Axel Pelster for critical comments.

References

- [1] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [2] D. Giulini, *Phys. Rev. D* **51**, 5630 (1995); D. Giulini and C. Kiefer, *Phys. Lett. A* **193**, 21 (1994); H.-J. Schmidt, in *Differential Geometry and Its Application*, Eds. J. Janyška and D. Krupka (World Scientific, Singapore, 1990), p. 405.
- [3] J.J. Halliwell, in *Conceptual Problems of Quantum Gravity*, Eds. A. Ashtekar and J. Stachel (Birkhäuser, Boston, 1991).
- [4] J.J. Halliwell and J. Louko, *Phys. Rev. D* **39**, 2206 (1989).
- [5] C. Kiefer, *Ann. Phys. (NY)* **207**, 53 (1991).
- [6] P. Hájíček, *J. Math. Phys.* **27**, 1800 (1986).
- [7] C. Kiefer, in *The Evolution of the Universe*, Eds. G. Börner and S. Gottlöber (John Wiley & Sons, Chichester, 1997).
- [8] J.B. Hartle and S.W. Hawking, *Phys. Rev. D* **28**, 2960 (1983).
- [9] C. Kiefer, *Class. Quantum Grav.* **10**, S233 (1993).

^bAlternatively, one can perform a transformation of the wave function in order to arrive directly at a Hamiltonian with standard form, cf. Ref. [19].

- [10] J. Butterfield and C. Isham, in *The Arguments of Time*, Ed. J. Butterfield (Oxford University Press, Oxford, 1999).
- [11] C. Kiefer and H.D. Zeh, *Phys. Rev. D* **51**, 4145 (1995).
- [12] I.H. Duru and H. Kleinert, *Phys. Lett. B* **84**, 30 (1979).
- [13] C. Stornaiolo, *Phys. Lett. A* **189**, 351 (1994).
- [14] C. Grosche and F. Steiner, *Handbook of Feynman Path Integrals* (Springer-Verlag, Berlin, 1998).
- [15] A. Pelster, *Zur Theorie und Anwendung nichtintegrabler Raum-Zeit-Transformationen in der klassischen Mechanik und in der Quantenmechanik* (Shaker-Verlag, Aachen, 1996).
- [16] D.C. Khandekar and S.V. Lawande, *J. Math. Phys.* **16**, 384 (1975).
- [17] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, Orlando, 1980).
- [18] T.D. Lee, *Particle Physics and Introduction to Field Theory* (Harwood, Chur, 1981).
- [19] A. Pelster and A. Wunderlin, *Z. Phys. B* **89**, 373 (1992).

VARYING LIGHT VELOCITY AS A SOLUTION TO THE PROBLEMS IN COSMOLOGY

J.W. MOFFAT

*Department of Physics, University of Toronto,
Toronto, Ontario M5S 1A7, Canada
E-mail: moffat@medb.physics.utoronto.ca*

In earlier published work, it was proposed that light speed was larger in the early Universe by 30 orders of magnitude compared to its presently observed value. This change in the speed of light is associated with a spontaneous breaking of local Lorentz invariance in the early Universe, associated with a first-order phase transition at a critical time $t = t_c$. This solves the horizon problem, leads to a mechanism of monopole suppression in cosmology, and can resolve the flatness problem. It also offers the potential of solving the cosmological constant problem. After the critical time t_c , local Lorentz invariance is restored and light travels at its presently measured speed. We investigate the field equations in the spontaneously broken phase and study further the flatness problem, and the cosmological constant problem. The entropy is shown to undergo a large increase as the light velocity goes through a phase transition. A scale invariant prediction for microwave background fluctuations originating at the time of the phase transition is discussed.

1 Introduction

The idea that the velocity of light varies in the early Universe and several of its consequences for cosmology were published some time ago [1–3]. The idea originated with the hypothesis that there is a phase of spontaneously broken, local Lorentz invariance and diffeomorphism invariance due to a non-vanishing vacuum expectation value (vev) of a field, ϕ , shortly after the beginning of the Universe [3]. The local Lorentz and diffeomorphism symmetries of Einstein's gravitational theory are spontaneously broken by the symmetry breaking pattern: $SO(3,1) \rightarrow O(3)$ at a critical temperature T_c , below which the symmetry is restored. It was shown that this new scenario is capable of

solving the horizon problem, the excess relic particle, and flatness problems, and that this leads to predictions for small scale inhomogeneities.

Recently, a series of papers, beginning with the paper by Albrecht and Maguiejo [4–6], has appeared in which the velocity of light (and possibly the gravitational constant G and other fundamental constants) was postulated to vary in the early Universe.

In the following, we shall investigate further the physical consequences of the spontaneous breaking of the symmetries of space-time in the early Universe, and the predictions for cosmology when the velocity of light goes through a first-order phase transition at a critical time $t \sim t_c$.

2 Spontaneous Breaking of Space-Time Symmetries

We assumed earlier [1–3] that local Lorentz vacuum symmetry is spontaneously broken by a Higgs mechanism. There we postulated the existence of four scalar fields, ϕ_a , where a labels the flat tangent space coordinates, and we assumed that the vev of the scalar fields, $\langle \phi_a \rangle_0$, vanishes for some temperature T less than a critical temperature T_c , when the local Lorentz symmetry is restored. Above T_c the non-zero vev will break the symmetry of the ground state of the Universe from $SO(3, 1)$ down to $O(3)$. The domain formed by the direction of the vev of the field ϕ will produce an arrow of time pointing in the direction of increasing entropy and the expansion of the Universe.

Let us now introduce the four real fields $\phi^a(x)$ ($a = 0, 1, 2, 3$), where $a, b, \dots$ denote flat tangent space coordinates. The fields ϕ^a are scalars in coordinate space with the coordinates x^μ , and they are invariant under Lorentz transformations

$$\phi'^a(x) = L_b^a(x) \phi^b(x). \quad (1)$$

We can use a vierbein e_μ^a to convert ϕ^a into a 4-vector in coordinate space: $\phi^\mu = e_\mu^a \phi^a$. The e_μ^a satisfy

$$e_\mu^a e_b^\mu = \delta_b^a, \quad e_a^\mu e_\nu^\mu = \delta_a^\nu, \quad (2)$$

and they obey the Lorentz transformation rule

$$e'^a_\mu(x) = L_b^a(x) e^b_\mu(x). \quad (3)$$

The covariant derivative operator acting on ϕ^a is defined by

$$D_\mu \phi^a = [\partial_\mu \delta_b^a + (\Omega_\mu)_b^a] \phi^b, \quad (4)$$

where $(\Omega_\mu)_b^a$ denotes the gauge spin connection.

Consider the infinitesimal Lorentz transformation rule

$$L_b^a(x) = \delta_b^a + \omega_b^a(x) \quad (5)$$

with

$$\omega_{ab}(x) = -\omega_{ba}(x). \quad (6)$$

The matrix $D(L)$ in the transformation rule for a general field $f_n(x)$,

$$f_n(x) \rightarrow \sum_m [D(L)(x)]_{nm} f_m(x), \quad (7)$$

takes the form

$$D[1 + \omega(x)] = 1 + \frac{1}{2} \omega^{ab}(x) \sigma_{ab}, \quad (8)$$

where the σ_{ab} are the six generators of the Lorentz group which satisfy $\sigma_{ab} = -\sigma_{ba}$ and the commutation rules

$$[\sigma_{ab}, \sigma_{cd}] = \eta_{cb} \sigma_{ad} - \eta_{ca} \sigma_{bd} + \eta_{db} \sigma_{ca} - \eta_{da} \sigma_{cb}. \quad (9)$$

The set of scalar fields ϕ transforms as

$$\phi'(x) = \phi(x) + \omega^{ab}(x) \sigma_{ab} \phi(x). \quad (10)$$

The gauge spin connection which satisfies the transformation law

$$(\Omega_\sigma)_b^a \rightarrow [L \Omega_\sigma L^{-1} - (\partial_\sigma L) L^{-1}]_b^a, \quad (11)$$

is given by

$$\Omega_\mu = \frac{1}{2} \sigma^{ab} e_a^\nu e_{b\nu;\mu}, \quad (12)$$

where the sign “;” denotes covariant differentiation with respect to the Christoffel symbol $\Gamma_{\mu\nu}^\lambda$:

$$\Gamma_{\mu\nu}^\lambda = g^{\lambda\rho} \eta_{ab} (D_\mu e_\nu^a) e_\rho^b. \quad (13)$$

A Higgs sector is included in the Lagrangian density such that the gravitational vacuum symmetry, which is set equal to the Lagrangian symmetry at low temperatures, breaks to a smaller symmetry at high temperature. The pattern of vacuum phase transition that emerges contains a symmetry anti-restoration. This vacuum symmetry breaking leads to the interesting possibility that exact zero temperature conservation laws, e.g. electric charge

and baryon number, are broken in the early Universe. It was shown that the spontaneous breaking of the Lorentz symmetry of the vacuum leads to a violation of the exact zero temperature conservation of energy in the early Universe, which can explain the origin of matter in the big bang.

Let us consider the Lorentz invariant potential:

$$V(\phi) = -\frac{1}{2}\mu^2 \sum_{a=0}^3 \phi_a \phi^a + \lambda \sum_{a=0}^3 (\phi_a \phi^a)^2, \quad (14)$$

where we choose ϕ^a to be a timelike Lorentz vector, $\phi_a \phi^a > 0$, and $\lambda > 0$, so that the potential is bounded from below. If V has a minimum at $\phi_a = v_a$, then the spontaneously broken solution is given by $v_a^2 = \mu^2/4\lambda$ and an expansion of V around the minimum yields the mass matrix:

$$(\mu^2)_{ab} = \frac{1}{2} \left(\frac{\partial^2 V}{\partial \phi_a \partial \phi_b} \right)_{\phi_a = v_a}. \quad (15)$$

We can choose ϕ_a to be of the form

$$\phi_a = \begin{pmatrix} 0 \\ 0 \\ 0 \\ v \end{pmatrix} = \delta_{a0}(\mu^2/4\lambda)^{1/2}. \quad (16)$$

All the other solutions of ϕ_a are related to this one by a Lorentz transformation rule. Then, the homogeneous Lorentz group $SO(3,1)$ is broken down to the spatial rotation group $O(3)$. The three rotation generators $J_i (i = 1, 2, 3)$ leave the vacuum invariant,

$$J_i v_i = 0, \quad (17)$$

while the three Lorentz-boost generators K_i break the vacuum symmetry

$$K_i v_i \neq 0. \quad (18)$$

The mass matrix $(\mu^2)_{ab}$ can be calculated from (15):

$$(\mu^2)_{ab} = \left(-\frac{1}{2}\mu^2 + 2\lambda v^2 \right) \delta_{ab} + 4\lambda v_a v_b = \mu^2 \delta_{a0} \delta_{b0}, \quad (19)$$

where v denotes the magnitude of v_a . There are three zero-mass Goldstone bosons, the same as the number of massive vector bosons, $V_\mu^i = (\Omega_\mu)_{0i} = -(\Omega_\mu)_{i0}$, and there are three massless vector bosons, $U_\mu^n = (\Omega_\mu)_{mn} =$

$-(\Omega_\mu)_{nm}$, corresponding to the unbroken $O(3)$ symmetry. In addition to these particles, one massive physical boson particle h remains, after the spontaneous breaking of the vacuum.

A phase transition is assumed to occur at the critical temperature T_c , when $v_a \neq 0$ and the Lorentz symmetry is broken and the three gauge fields $(\Omega_\mu)_{i0}$ become massive vector bosons. Below T_c the Lorentz symmetry is restored, and we regain the usual classical gravitational field with massless gauge fields Ω_μ . The symmetry breaking will extend to the singularity or the possible singularity-free initial state of the big bang, and since quantum effects associated with gravity do not become important before $T \sim 10^{19}$ GeV, we expect that $T_c \leq 10^{19}$ GeV.

The total action for the theory is

$$S = S_G + S_M + S_\phi, \quad (20)$$

where the action for Einstein gravity is

$$S_G = -\frac{c^4}{16\pi G} \int d^4x e(R + 2\Lambda). \quad (21)$$

The determinant of the metric is abbreviated as

$$e \equiv \sqrt{-g} = \det(e_\mu^a e_{a\nu})^{1/2}, \quad (22)$$

Λ is the cosmological constant, and S_M is the matter action for gravity. Moreover,

$$S_\phi = \int d^4x \sqrt{-g} \left[\frac{1}{2} D_\mu \phi_a D^\mu \phi^a - V(\phi) \right]. \quad (23)$$

By choosing ϕ^a to be a Lorentz timelike vector, we ensure that the kinetic energy term $D_\mu \phi_a D^\mu \phi^a > 0$ for all events in the past and future light cones of the flat tangent space. Since the kinetic energy term is positive definite within the light cone and $\phi_a \phi^a > 0$, we guarantee that local Lorentz invariance is broken within the past and future light cones where space-time events are causally connected.

Let us consider small oscillations about the true minimum and define a shifted field

$$\phi'_a = \phi_a - v_a. \quad (24)$$

We perform a Lorentz transformation rule on ϕ^a , so that we obtain

$$\phi^0 = h, \quad \phi^1 = \phi^2 = \phi^3 = 0. \quad (25)$$

In this special coordinate frame, the remaining component h is the physical Higgs particle that survives after the three Goldstone modes have been removed. This corresponds to choosing the "unitary gauge" in the standard electroweak model.

The total action for the theory in the broken symmetry phase, $T > T_c$, is

$$S = S_G + S_M + S_h + S_V. \quad (26)$$

In our specially chosen coordinate frame in which (25) holds, we have

$$S_h = \int d^4x \sqrt{-g} \left[\frac{1}{2} \partial_\mu h \partial^\mu h - V(h) \right], \quad (27)$$

where

$$V(h) = 4\lambda v^2 h^2 + 4\lambda v h^3 + \lambda h^4 - \frac{1}{2} V_\mu^2 h^2 - v V_\mu^2 h, \quad (28)$$

and we have for convenience suppressed the index i on V_μ^i . Moreover,

$$S_V = \frac{1}{2} m^2 \int d^4x \sqrt{-g} g^{\mu\nu} V_\mu V_\nu, \quad (29)$$

where the mass $m \propto \langle h \rangle_0$.

The field equations are of the form:

$$G^{\mu\nu} \equiv R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = \frac{8\pi G}{c^4} (T^{\mu\nu} + K^{\mu\nu} + H^{\mu\nu}) + \Lambda g^{\mu\nu}, \quad (30)$$

where $K^{\mu\nu}$ is given by

$$K^{\mu\nu} = m^2 \left(V^\mu V^\nu - \frac{1}{2} g^{\mu\nu} V^\beta V_\beta \right). \quad (31)$$

Moreover, the h field energy-momentum tensor is of the usual form:

$$H^{\mu\nu} = \partial^\mu h \partial^\nu h - \mathcal{L}_h g^{\mu\nu}. \quad (32)$$

Since $G^{\mu\nu}$ satisfies the Bianchi identities:

$$G^{\mu\nu}{}_{;\nu} = 0, \quad (33)$$

we find from (30) that

$$T^{\mu\nu}{}_{;\nu} = -(K^{\mu\nu} + H^{\mu\nu})_{;\nu}, \quad (34)$$

where we have used $g^{\mu\nu}{}_{;\nu} = 0$. In the unbroken phase of space-time, we regain the standard energy-momentum conservation laws ($K^{\mu\nu} = 0$, $H^{\mu\nu}{}_{;\nu} = 0$):

$$T^{\mu\nu}{}_{;\nu} = 0, \quad (35)$$

and the spin connection corresponds to a massless graviton gauge field.

3 Field Equations in the Broken Symmetry Phase

The space-time manifold in the broken phase has the symmetry $R \times O(3)$. The three-dimensional space with $O(3)$ symmetry is assumed to be homogeneous and isotropic, and yields the usual maximally symmetric three-dimensional space:

$$d\sigma^2 = R^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right], \quad (36)$$

where t is the external time variable. This is the Robertson-Walker theorem for our ordered phase of the vacuum. It has the correct subspace structure for the FRW Universe with the metric:

$$ds^2 = dt^2 c^2 - R^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right]. \quad (37)$$

In the broken symmetry phase, the "time" t is the *absolute physical time* measured by standard clocks. In contrast to GR, while $\langle\phi\rangle_0$ is non-zero, we no longer have reparametrization invariance and time is no longer an arbitrary label.

Let us consider Einstein's field equations in the broken symmetry phase $T > T_c$. We have $g_{00}(t) = c_0^2$ and $g_{ik}(t) = -R^2(t)\gamma_{ik}$ and the energy-momentum tensor for a perfect fluid takes the form:

$$T^{\mu\nu} = \left(\rho + \frac{p}{c_0^2} \right) u^\mu u^\nu - p g^{\mu\nu}, \quad (38)$$

where c_0 labels the velocity of light in this epoch, $u^\mu = dx^\mu/ds$, $u^\mu u_\mu = c_0^2$, ρ is the density of matter and radiation and p is the pressure. In our homogeneous space the spatial part of the massive vector field is $V^i = 0$. Let us use the

notation: $V^0(t) = \chi(t)$. We obtain the field equations in the broken symmetry phase

$$\frac{\dot{R}^2}{c_0^2 R^2} + \frac{k}{R^2} = \frac{8\pi G}{3c_0^2} \rho + \frac{8\pi G}{c_0^4} \left[\frac{1}{2} m^2 \chi^2 + \frac{1}{2} (\dot{h})^2 + V(h) \right] + \frac{\Lambda}{3}, \quad (39)$$

$$2 \left(\frac{\ddot{R}}{c_0^2 R} \right) + \frac{\dot{R}^2}{c_0^2 R^2} + \frac{k}{R^2} = -\frac{8\pi G}{c_0^4} \left[p + \frac{1}{2} m^2 \chi^2 + \frac{1}{2} (\dot{h})^2 - V(h) \right] + \Lambda, \quad (40)$$

$$\frac{1}{R^3} \frac{\partial}{\partial t} \left[R^3 c_0^2 \left(\rho + \frac{p}{c_0^2} \right) \right] - \dot{p} = -W, \quad (41)$$

where $\dot{R} = dR/dt$ and where W is given by

$$W = m^2 \chi \left[\dot{\chi} + 3 \left(\frac{\dot{R}}{R} \right) \chi \right] + (\dot{h})^2 \left[1 + 3 \left(\frac{\dot{R}}{R} \right) \right] + V'(h) \dot{h}. \quad (42)$$

4 Superluminary Universe

The horizon scale is determined by

$$d_H(t) = c_0 R(t) \int_0^t \frac{dt'}{R(t')}. \quad (43)$$

For $t > t_c$, this will have the usual value: $d_H(t) = 2ct$, since $R(t) \propto t^{1/2}$ for a radiation dominated Universe. Let us assume that for $t \leq t_c$ the speed of light is very large. During a first-order phase transition, the velocity of light is assumed to undergo a discontinuous change from the value

$$c_0 \sim ac \quad (44)$$

for $t \leq t_c$ to $c_0 = c$ (c is the present value of the velocity of light and a is a constant) for $t > t_c$. Then, we get for $t \leq t_c$:

$$d_H(t) \approx c_0 g(t), \quad (45)$$

where $g(t)$ is the dynamical time dependence arising from $R(t)$ in (43). Thus, for a fraction of time near the beginning of the Universe, and for $a \rightarrow \infty$, all points of the expanding space will have been in communication with one another solving the horizon problem.

Suppose the field ϕ that breaks the space-time symmetries is characterized by a correlation length ξ . Then, the monopole density is approximately given by

$$n_M \approx \xi^{-3}. \quad (46)$$

In the superluminary model, the bound on the length ξ is given by

$$\xi < d_H(t) \approx c_0 g(t), \quad (47)$$

so that the bound on the number density of monopoles is exponentially weakened. This solves the relic particle (monopole) problem.

The present observational data restrict $\Omega_0 = \rho_{\text{crit}}/\rho_0$ to lie in the interval $[0.01, \text{few}]$, which implies that $R_{\text{curv}} \sim c/H_0$ and $\rho_0 \sim \rho_{\text{crit}}$. From Eq. (39) in the broken phase, we can derive the expression

$$\Omega(t) = 1/[1 - x(t)], \quad (48)$$

where, in the radiation dominated superluminary era,

$$x(t) = \frac{c_0^2 k}{R^2 H^2} \sim \frac{c_0^2 k/R^2}{8\pi G \rho_r/3}, \quad (49)$$

and where ρ_r denotes the radiation density. Moreover we have

$$\Omega = \Omega_{\rho_r} + \Omega_F + \Omega_\Lambda, \quad (50)$$

in which

$$\Omega_{\rho_r} = \frac{8\pi G \rho_r}{3H^2}, \quad \Omega_F = \left(\frac{8\pi G}{c_0^2 H^2} \right) \left[\frac{1}{2} m^2 \chi^2 + \frac{1}{2} \dot{h}^2 + V(h) \right], \quad (51)$$

$$\Omega_\Lambda = \frac{c_0^2 \Lambda}{3H^2}. \quad (52)$$

In the radiation dominated era, we have $\rho_r = \rho_{0r}(R_0/R)^4$, so that

$$x \sim \frac{c_0^2 k R^2}{R^*}, \quad (53)$$

where $R^* = 8\pi G \rho_{0r} R_0^4/3$. This yields

$$|\Omega(10^{-43} \text{sec}) - 1| \sim O(a^2 10^{-60}). \quad (54)$$

Thus, in the short time that the Universe is superluminary with $a \sim 5 \times 10^{29}$, we get

$$|\Omega(10^{-43}\text{sec}) - 1| \sim O(1), \quad (55)$$

which implies much less fine tuning than the standard FRW model.

We observe that in contrast to the inflationary model (which predicts that $\Omega = 1$) [7,8], the prediction for the value of Ω in the superluminary model depends on the detailed dynamics of the theory [1,4-6]. Indeed, if we were to assume the equation of state: $\rho = \text{const.}$, that χ and V are uniform in the broken phase and $k = 0$, then $R(t)$ has the inflationary solution:

$$R(t) \propto \exp \left[t c_0 \left(\Lambda/3 \right)^{1/2} \right]. \quad (56)$$

Thus we would regain the standard inflationary prediction $\Omega = 1$. Clearly, the superluminary model *does not lead automatically to the generic prediction* $\Omega = 1$. The possibility of obtaining an open Universe version of the inflationary scenario has been the subject of much controversy recently [9-11]. The fact that obtaining an open Universe in the superluminary model is not a problem is a positive feature in favour of the model.

Let us now consider the cosmological constant problem. We ignore the effects of the h and V fields, since they will not play an important role in the present discussion. Then, using the equation of state, $p = \rho/3$, we can derive at some instant of time $\bar{t}$:

$$H^2(\bar{t})(q(\bar{t}) - 1) = \frac{c_0^2 k}{R^2(\bar{t})} - \frac{2\Lambda c_0^2}{3}, \quad (57)$$

where q is the deceleration parameter

$$q = -\ddot{R}R/\dot{R}^2. \quad (58)$$

We also have

$$\frac{c_0^2 k}{R^2(\bar{t})H^2(\bar{t})} = \Omega(\bar{t}) - 1 + \frac{c_0^2 \Lambda}{3H^2(\bar{t})}. \quad (59)$$

It follows that

$$|q(\bar{t}) - \Omega(\bar{t})| = \left| \frac{c_0^2 \Lambda}{3H^2(\bar{t})} \right|. \quad (60)$$

Assuming that the radiation dominant solution of the field equations holds near the phase transition, then $\Omega(\bar{t}) \sim 0.1 - 1$, $q(\bar{t}) \sim 1$, $H(\bar{t}) \sim 1/2\bar{t}$ and

$$R(\bar{t}) = \left(\frac{32\pi G \rho_{0r}}{3} \right)^{1/4} R_0 \bar{t}^{1/2}. \quad (61)$$

We obtain

$$|\Lambda| \approx \frac{1}{c_0^2 \bar{t}^2}. \quad (62)$$

For a rapid phase transition in the velocity of light and for $c_0 = ac$ and $\log_{10} a \geq 60$, we have for $\bar{t} \sim 10^{-43}$ sec:

$$|\Lambda| < 10^{-54} \text{ cm}^{-2}. \quad (63)$$

From the critical density $\rho_{\text{crit}} \sim \rho_0$ this bound corresponds to $\Lambda/8\pi G \leq 8 \times 10^{-47} h^2 \text{ GeV}^4$, where $0.4 \leq h \leq 1$. This would solve the cosmological constant problem [2,4,5]. The observational bound in (63) is obtained by using $H_0 \sim 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

However, there is a serious conflict between the value $\log_{10} a \leq 30$ required to solve the flatness problem and the value $\log_{10} a > 60$ required to solve the cosmological constant problem. Indeed, the latter value is far too large to accomodate a reasonable evolution of the FRW Universe just after the occurrence of the phase transition in the velocity of light. However, if we assume that another phase transition in the velocity of light occurs, before the one that solves the flatness problem, with $\log_{10} \sim 60$, then this could solve the cosmological constant problem and be followed by a phase transition with a lowering of the speed of light to a value with $\log_{10} \sim 30$, which could solve the flatness problem and allow the Universe to expand to its present day value.

We have assumed the radiation dominant solution for $R(t)$ in this derivation. Perhaps another dynamical solution of the field equations would accomodate a solution to the horizon, flatness and cosmological constant problems in the presence of a phase transition in the velocity of light. Barrow [5] has considered the time dependent solutions of Brans-Jordan-Dicke theories with a field $\psi = c^4$, but such theories are severely restricted dynamically and can easily lead to consistency problems. In any eventuality, the horizon, flatness, and monopole problems can be resolved by the superluminary model. The problem with the cosmological constant is not resolved in inflationary models. It is indeed exacerbated by the enormous vacuum energy density

required to drive the initial inflation. The potential for the superluminary model to solve the cosmological constant problem could provide the model with another significant advantage.

One important aspect of the superluminary model is that it is not sensitive to the choice of an equation of state. No exotic forms of matter with negative pressure are required to resolve cosmological problems. This is an advantage over the standard inflationary scenarios which require vacuum energy with $p = -\rho$ and unusual forms of potentials for the inflaton field that often require fine-tuning to implement the inflationary period.

5 Black-Body Radiation and Entropy in the Superluminary Phase

In the Lorentz symmetry broken phase of the Universe, the total black-body radiation energy at temperature T is given by

$$u = \left(\frac{8\pi h\nu^3}{c_0^3} \right) \int_0^\infty \left[\exp\left(\frac{h\nu}{k_B T} \right) - 1 \right]^{-1} d\nu = \frac{8\pi^5 (k_B T)^4}{15 h^3 c_0^3}. \quad (64)$$

The number density of photons is

$$n_\gamma = \frac{60.42198 (k_B T)^3}{(h c_0)^3}. \quad (65)$$

The energy densities for photons and neutrinos are

$$\epsilon_\gamma = \sigma_B T^4, \quad \epsilon_\nu = \frac{7}{16} \sigma_B T^4, \quad (66)$$

where σ_B is the Stefan-Boltzmann constant

$$\sigma_B = \frac{\pi^2 k_B^4}{60 h^3 c_0^3} \quad (67)$$

and k_B is Boltzmann's constant. For $\log_{10} a \leq 30$ Stefan-Boltzmann's constant is $\sigma_B \sim 1.7 \times 10^{-84} \text{ erg cm}^{-3} K^{-4}$ in the superluminary phase.

The Universe has negligible thermal and neutrino energy during the short period of the spontaneously broken symmetry phase. After the superluminary phase ends when $c_0 = c$, both ϵ_γ and ϵ_ν regain their standard values in an FRW Universe. In contrast to the inflationary model, there is no problem with a reheating epoch necessary in the inflationary scenario to replenish the matter and radiation in an "empty" de-Sitter Universe. The velocity of light phase transition in the spontaneous symmetry breaking process automatically takes care of the creation of matter when $c_0 = c$ [12,13].

The entropy of thermal photons is

$$S(T) = \frac{4\pi^2 V}{45} \left(\frac{T}{\hbar c_0} \right)^3. \quad (68)$$

Thus, at the phase transition at $t \sim t_c$, the entropy increases enormously in the direction of the expanding Universe when $c_0 = c$. The arrow of time connected with the increase of the entropy is determined by the domain arrow produced by the non-zero $\langle \phi \rangle_0$ in the spontaneously broken phase [1–3].

We see that the large value of c in the early Universe, in the broken symmetry phase, changes radically the thermal physics at the beginning of the Universe. This is in accord with the known observation that the entropy of a system increases rapidly as the system undergoes a first-order or second-order phase transition from an ordered to a more disordered state.

6 Quantum Fluctuations and Density Perturbations

Let us consider the possibility in our model of generating the seed perturbations that can grow to form the large-scale structures. During the superluminary phase for $t < t_c$, the fluctuation wavelengths grow as $\lambda \propto R(t)$. However, the horizon grows rapidly, $d_H \approx c_0 g(t)$, where c_0 is given by (44), and it will become equal to the physical wavelength at some time $t = t_{\text{exit}}$. This means that it becomes larger than $\lambda(k)$ for a mode labeled by a wave vector $\vec{k}$.

After the symmetry is restored at $t > t_c$, the proper length $R(t)$ grows as $t^{1/2}$, whereas the horizon will increase as $cH(t)^{-1} \sim ct$. Therefore, the wavelength will be completely within the Hubble radius for an interval of time Δt . Thus, in the superluminary model the fluctuations are in microcausal connection very early in the Universe ($t \sim 10^{-35}$ s) and have time to grow into physical modes sufficiently large to form galaxy structures. These fluctuations will have a Gaussian form, provided any self-couplings of the matter fields are small.

The fluctuations associated with the Higgs field, h , could be a candidate for seed perturbations. The h field satisfies

$$\ddot{h} + 3H\dot{h} + V'(h) = 0, \quad (69)$$

where $V'(h) = dV(h)/dh$.

When the velocity of light undergoes a discontinuous change, during a first-order phase transition at $t \sim t_c$, to the value c_0 given by (44) with $a \geq$

10^{30} , then the horizon, $d_H(t)$, determined by (43) will also have a discontinuity in its first derivative with respect to t , and $d_H(t)$ for $t < t_c$ can be matched to $d_H(t)$ for $t \geq t_c$ in such a way that λ crosses $d_H(t)$ twice. The fluctuations are "frozen in" and leave an imprint on the metric tensor.

Fluctuations in h give rise to perturbations in the density

$$\delta\rho_h = \delta h \left(\frac{\partial V}{\partial h} \right). \quad (70)$$

At horizon crossings, $\lambda_{\text{phys}} \sim cH^{-1}$, the gauge invariant quantity ζ takes the simple form $\zeta = \delta\rho/(\rho + p/c^2)$ [14]. In the radiation dominated era and in the matter dominated era, ζ at horizon crossing is, up to a factor of order unity, equal to $\delta\rho/\rho$. Equating the values of ζ at the two horizon crossings, we find

$$\left(\frac{\delta\rho}{\rho} \right)_{\text{Hor}} \sim \frac{\delta h V'}{\dot{h}^2} \sim \frac{H V'}{2\pi \dot{h}^2}, \quad (71)$$

where we have used the fact that $\delta h \sim H/2\pi$. We must now model V' and $\dot{h}$ at the phase transition, in order to estimate the density fluctuation, $\delta\rho/\rho$. Clearly, H is rapidly varying at the phase transition. We have

$$\dot{h} \propto \frac{H}{2\pi\delta t}. \quad (72)$$

A natural time scale for the duration of the phase transition is given by

$$\delta t \propto \left(\frac{H}{2\pi\dot{h}^3} \right)^{1/2}. \quad (73)$$

Thus, if we choose $1/H \sim 10^{-34}$ s and $h \sim M_P \propto 10^{43} \text{s}^{-1}$, then the duration of the phase transition is $\delta t \sim 10^{-48}$ s.

By assuming that $V(h)$ is dominated by $V(h) \sim \lambda/4h^4$, we obtain from Eqs. (71), (72), and (73) the scale invariant prediction for the amplitude

$$\left(\frac{\delta\rho}{\rho} \right)_{\text{Hor}} \sim \lambda. \quad (74)$$

We can fit the data measured by the Cosmic Background Explorer (COBE), which is consistent with a Gaussian, scale invariant spectrum [15,16], by choosing the coupling constant $\lambda \sim 10^{-5}$, and using $\Delta T/T \sim \delta\rho/3\rho$. The measurements are quoted in terms of a spectral index n , with $n = 1.1 \pm 0.5$. These measurements are also consistent with the predictions of inflationary

models, and with other mechanisms of inhomogeneity generation, such as cosmic strings.

7 Conclusions

The results obtained above suggest that the superluminary model could be an attractive alternative to inflation as a solution to the initial value problem in cosmology. Moreover, our picture of the period immediately following the big bang is radically altered from the standard big bang model. The violation of the conservation of energy in the spontaneously broken symmetry phase can provide an explanation for the creation of matter in the beginning of the Universe, an explanation which is not available in the standard FRW model or the inflationary scenario. The model also provides a possible solution for the cosmological constant problem.

The quantum fluctuations of the Higgs field, near the phase transition in the velocity of light, produce microwave background density fluctuations, which are frozen in at the horizon. The spectrum of the fluctuations is Gaussian and scale invariant with a scalar field coupling constant, $\lambda \sim 10^{-5}$, which is more reasonable in size than the standard value, $\lambda \sim 10^{-14}$, predicted by generic inflationary models [7,8]. Further work is necessary to investigate in more detail the predictions of the density fluctuation spectrum in the superluminary model.

The superluminary model of the early Universe [1-3] was introduced to provide an interesting alternative to the inflationary model, at a time when the latter model was enjoying a popular revival. It is remarkable that there has been no serious alternative model considered besides the superluminary model (or varying light speed model, as it is called by Albrecht and Maguiejo [4]). With the advent of a new generation of accurate satellite measurements of the microwave background, it may be possible to distinguish the specific predictions of the two models and the predictions of other alternative models that may be forthcoming in the future.

An important feature of the superluminary model is the unavoidable breaking of local Lorentz invariance in the early Universe. In view of the significant changes that will occur in the fundamental physics, it is necessary to have a well-defined model for this symmetry breaking. Such a model is provided by the spontaneous symmetry breaking of the Lagrangian in the scenario presented in earlier publications [1-3] and in the present exposition. The "hidden" symmetry of the gravitational vacuum has the advantage, en-

joyed in the standard model of particle physics, of retaining the vital features of gauge symmetries, such as Ward identities, in a future theory of quantum gravity.

In recent work [17,18], gravitational theories based on a bimetric structure formed from a metric and a vector field or the gradient of a scalar field have been proposed. These theories begin with a Lorentz and diffeomorphism invariant formulation and provide an alternative picture to the one described here, based on a rapid change in the velocity of light associated with a phase transition at a critical time $t = t_c$ in the early Universe when local Lorentz invariance symmetry is broken.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

- [1] J.W. Moffat, *Int. J. Mod. Phys. D* **2**, 351 (1993), eprint: gr-qc/9211020.
- [2] J.W. Moffat, Review Talk given at the *International Conference on Problems in Gravitation and Particle Physics*, Institute for Theoretical Physics, Protvino, Russia, June 1992.
- [3] J.W. Moffat, *Found. of Phys.* **23**, 411 (1993), eprint: gr-qc/9209001.
- [4] A. Albrecht and J. Maguiejo, eprint: astro-ph/9811018.
- [5] J.D. Barrow, eprint: astro-ph/9811022.
- [6] J.D. Barrow and J. Maguiejo, eprint: astro-ph/9811072.
- [7] A.H. Guth, *Phys. Rev. D* **23**, 347 (1981).
- [8] For reviews of inflationary cosmologies, see: E.W. Kolb and M.S. Turner, *The Early Universe* (Addison-Wesley, California, 1990); A.D. Linde, *Particle Physics and Inflationary Cosmology* (Harwood, Chur, 1990).
- [9] N. Turok and S.W. Hawking, *Phys. Lett. B* **425**, 25 (1998); *Phys. Lett. B* **432**, 271 (1998).
- [10] A.D. Linde, *Phys. Rev. D* **58**, 083514 (1998).
- [11] A. Vilenkin, *Int. J. Theor. Phys.* **38**, 3135 (1999).
- [12] T. Harko and M.K. Mak, *Class. Quantum Grav.* **16**, 2741 (1999).
- [13] M.K. Mak and T. Harko, *Class. Quantum Grav.* **16** (1999), to be published.

- [14] J.M. Bardeen, P.J. Steinhardt, and M.S. Turner, *Phys. Rev. D* **28**, 679 (1983).
- [15] E.R. Harrison, *Phys. Rev. D* **1**, 2726 (1970); Ya.B. Zel'dovich, *Mon. Not. Roy. Astron. Soc.* **160**, 1p (1972).
- [16] G.F. Smoot *et al.*, *Ap. J. Letts.* **396**, L1 (1992); C. Bennett *et al.*, *ibid.* **396**, L7 (1992).
- [17] M.A. Clayton and J.W. Moffat, *Phys. Lett. B* **460**, 263 (1999), eprint: astro-ph/9812481.
- [18] M.A. Clayton and J.W. Moffat, *Phys. Lett. B* **477**, 269 (2000), eprint: gr-qc/9910112.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

RENORMALIZATION GROUP METHOD AND INHOMOGENEOUS UNIVERSE

Y. NAMBU

*Department of Physics, Graduate School of Science,
Nagoya University, Nagoya 464-8602, Japan
E-mail: nambu@allegro.phys.nagoya-u.ac.jp*

Using the renormalization group method, we improve the first-order solution of the long-wavelength expansion of the Einstein equation and obtain the renormalization group equation. The solution of the renormalization group equation shows that the renormalized metric describes the behavior of the gravitational collapse in the expanding Universe qualitatively well and is suitable for modeling an inhomogeneous Universe.

1 Introduction

Our Universe seems to be very close to a Friedmann-Robertson-Walker (FRW) space-time at a length scale of the order of the Hubble radius, but the metric and matter content appear to be highly inhomogeneous at smaller scales. The conventional cosmological perturbative approach [1] treats such a situation as the homogeneous isotropic background plus a small perturbation, and investigates the evolution of linear fluctuations. We must go beyond the linear approximation to treat the nonlinear structure and to construct a suitable model of an inhomogeneous Universe which is close to a FRW Universe on a large scale. The spatial gradient expansion [2] of the Einstein equation is a nonlinear approximation method which describes the long-wavelength inhomogeneity in the Universe. This approximation scheme expands the Einstein equation with respect to the order of the spatial gradient. As a background solution, we solve Einstein's equation by neglecting all spatial gradient terms. The resulting solution has the same form as that for the spatially flat FRW

Universe, but the three-metric can have a spatial dependence. It is possible to include the effect of spatial gradient terms by calculating the next order. This method can describe a long-wavelength nonlinear perturbation without imposing any symmetry for a space, and is suitable for analyzing the global structure of an inhomogeneous Universe.

However, this scheme is valid only for a perturbation whose wavelength is larger than the Hubble horizon scale. For a matter field which satisfies the energy conditions, the perturbation terms induced by the spatial gradient terms grow in time and finally dominate the background solution. This occurs when the wavelength of the perturbation equals the Hubble horizon scale. After this time, the wavelength of the perturbation becomes shorter than the horizon scale and the result of the gradient expansion becomes unreliable.

A similar situation occurs in the field of nonlinear dynamical systems. To obtain the temporal evolution of the solution of a nonlinear differential equation, we usually apply a perturbative expansion. But naive perturbation often yields secular terms due to resonance phenomena. The secular terms prevent us from getting approximate but global solutions. There are many techniques to circumvent the problem, for example, the averaging method, the multi-time scale method, the WKB method and so on [3]. Although these methods yield globally valid solutions, they provide no systematic procedure for general dynamical systems because we must select a suitable assumption on the structure of the perturbation series.

The renormalization group method [4] as a tool for a global asymptotic analysis of the solution to differential equations unifies the techniques listed above, and can treat many systems irrespective of their features. Starting from a naive perturbative expansion, the secular divergence is absorbed in the constants of integration contained in the zeroth-order solution by the renormalization procedure. The renormalized constants obey the renormalization group equation. This method can be viewed as a tool of system reduction. The renormalization group equation corresponds to the amplitude equation which describes slow motion dynamics in the original system. We can describe complicated dynamics contained in the original equation by extracting a simpler representation using the renormalization group method.

In this article, we apply the renormalization group method to the gradient expansion of Einstein's equation. Our purpose is to obtain the renormalized long-wavelength solution of Einstein's equation which is also valid for later times. Through the procedure of renormalization, we extract slow motion from the Einstein equation [5].

2 Renormalization Group Method

The renormalization group method [4] improves the long-time behavior of a naive perturbative expansion. We explain the basic concept of the renormalization group method using two examples. The first one is a harmonic oscillator. The equation of motion is

$$\ddot{x} + x = -\epsilon x, \quad (1)$$

where ϵ is a small parameter. We solve this equation perturbatively by expanding the solution with respect to ϵ :

$$x = x_0 + \epsilon x_1 + \cdots. \quad (2)$$

The solution up to $O(\epsilon)$ becomes

$$x = B_0 \cos t + C_0 \sin t + \frac{\epsilon}{2} (t - t_0)(C_0 \cos t - B_0 \sin t) + O(\epsilon^2), \quad (3)$$

where B_0 and C_0 are constants of integration determined by the initial condition at arbitrary time $t = t_0$. This naive perturbation breaks down when $\epsilon(t - t_0) > 1$ because of the secular term. To regularize the perturbation series, we introduce an arbitrary time μ , split $t - t_0$ as $t - \mu + \mu - t_0$, and absorb the divergent term containing $\mu - t_0$ into the renormalized counterparts B and C of B_0 and C_0 , respectively.

We introduce renormalized constants as follows:

$$B_0 = B(\mu) + \epsilon \delta B(\mu, t_0), \quad C_0 = C(\mu) + \epsilon \delta C(\mu, t_0), \quad (4)$$

where δB and δC are counter terms that absorb the terms containing $\mu - t_0$ in a naive solution. Inserting Eq. (4) in Eq. (3), we have

$$x = B(\mu) \cos t + C(\mu) \sin t + \epsilon \left\{ \delta B \cos t + \delta C \sin t + \frac{1}{2} (t - \mu + \mu - t_0)(C(\mu) \cos t - B(\mu) \sin t) \right\}. \quad (5)$$

We choose δB and δC as

$$\delta B(\mu, t_0) + \frac{1}{2}(\mu - t_0)C(\mu) = 0, \quad \delta C(\mu, t_0) - \frac{1}{2}(\mu - t_0)B(\mu) = 0. \quad (6)$$

Using the relations $\epsilon \delta B = B_0 - B(\mu)$, and $\epsilon \delta C = C_0 - C(\mu)$, we obtain

$$B(t_0) = B(\mu) - \frac{\epsilon}{2}(\mu - t_0)C(\mu), \quad C(t_0) = C(\mu) + \frac{\epsilon}{2}(\mu - t_0)B(\mu). \quad (7)$$

These equations define the transformation up to $O(\epsilon)$:

$$\mathcal{R}_{\mu-t_0} : (B(t_0), C(t_0)) \rightarrow (B(\mu), C(\mu)),$$

and the transformation forms a Lie group up to $O(\epsilon)$. Assuming the properties of a Lie group, we can extend the locally valid expression (7) to a global one, which is valid for arbitrary large $\mu - t_0$. We apply this transformation to get $(B(\mu), C(\mu))$ at arbitrary large μ . By differentiating Eq. (7) with respect to μ and setting $t_0 = \mu$, we have the renormalization group equations

$$\frac{\partial B}{\partial \mu} = \frac{\epsilon}{2} C(\mu), \quad \frac{\partial C}{\partial \mu} = -\frac{\epsilon}{2} B(\mu). \quad (8)$$

The renormalized solution becomes

$$x = B(\mu) \cos t + C(\mu) \sin t + \frac{\epsilon}{2} (t - \mu) [C(\mu) \cos t - B(\mu) \sin t]. \quad (9)$$

Solving the renormalization group equations (8) and equating μ and t in (9) eliminates the secular term and we get a uniformly valid result

$$x = B(0) \cos \left(1 + \frac{\epsilon}{2}\right) t + C(0) \sin \left(1 + \frac{\epsilon}{2}\right) t. \quad (10)$$

The second example is the Einstein equation for a FRW Universe with dust. The spatial component is

$$\ddot{\alpha} + \frac{3}{2} \dot{\alpha}^2 = -\frac{\epsilon}{2} e^{-2\alpha}, \quad (11)$$

where $\alpha(t)$ is the logarithm of the scale factor of the Universe $a(t)$ and ϵ is the sign of the spatial curvature. The exact solution is given by

$$a(t) = e^{\alpha(t)} = \begin{cases} a_0 (1 - \cos \eta), & t = a_0 (\eta - \sin \eta) & \text{for } \epsilon = 1, \\ a_0 \eta^2/2, & t = a_0 \eta^3/6 & \text{for } \epsilon = 0, \\ a_0 (\cosh \eta - 1), & t = a_0 (\sinh \eta - \eta) & \text{for } \epsilon = -1. \end{cases} \quad (12)$$

We solve Eq. (11) perturbatively by assuming that the right hand side is small. This represents an expansion with respect to a small spatial curvature around the flat Universe. By substituting $\alpha = \alpha_0 + \epsilon \alpha_1 + \dots$ in Eq. (11), we find the naive solution

$$\alpha = \ln \tau + C_0 - \epsilon \frac{9}{20} e^{-2C_0} (\tau - \tau_0) + O(\epsilon^2), \quad (13)$$

where $\tau = t^{2/3}$ is a new time variable and C_0 a constant of integration determined by the initial condition at $\tau = \tau_0$. The $O(\epsilon)$ term is secular and we

regularize this term by introducing the arbitrary time μ and the renormalized constant $C_0 = C(\mu) + \epsilon \delta C(\mu, \tau_0)$:

$$\alpha = \ln \tau + C(\mu) + \epsilon \delta C(\mu, \tau_0) - \epsilon \frac{9}{20} e^{-2C(\mu)} (\tau - \mu + \mu - \tau_0). \quad (14)$$

The counter term δC is determined in such a way that it absorbs the term depending on $\mu - \tau_0$:

$$\delta C(\mu, \tau_0) - \frac{9}{20} e^{-2C(\mu)} (\mu - \tau_0) = 0. \quad (15)$$

This defines the renormalization group transformation

$$\mathcal{R}_{\mu-\tau_0} : C(\tau_0) \rightarrow C(\mu)$$

according to

$$C(\mu) = C(\tau_0) - \epsilon \frac{9}{20} e^{-2C(\mu)} (\mu - \tau_0), \quad (16)$$

and this transformation forms a Lie group up to $O(\epsilon)$. So we can have $C(\mu)$ for arbitrary large values of $\mu - \tau_0$ by assuming the property of a Lie group. This makes it possible to produce a globally uniform approximative solution of the original equation. The renormalization group equation reads

$$\frac{\partial C(\mu)}{\partial \mu} = -\epsilon \frac{9}{20} e^{-2C(\mu)}, \quad (17)$$

and its solution is

$$C(\mu) = \frac{1}{2} \ln \left(c - \frac{9\epsilon}{10} \mu \right), \quad (18)$$

where c is a constant of integration. The renormalized scale factor is given by

$$a(t) = e^{\alpha(t)} = \tau e^{C(\tau)} = t^{2/3} \left(c - \frac{9\epsilon}{10} t^{2/3} \right)^{1/2}. \quad (19)$$

As the zeroth-order solution, it is possible to include another integration constant t_0 which defines the origin of the cosmic time t . By requiring that the renormalization group transformation forms a Lie group, it can be shown that t_0 is not renormalized. Hence it is sufficient to consider the solution with the boundary condition $a(t=0) = 0$ which fixes the value of t_0 to zero. This point is different from the example of the harmonic oscillator, in which case two integration constants B and C are renormalized.

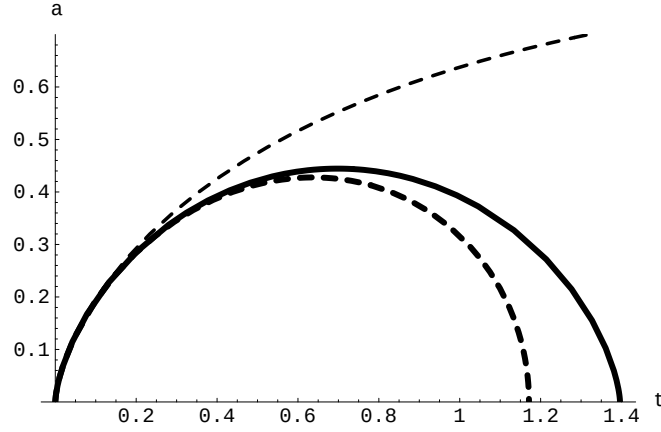


Figure 1. The evolution of the scale factor for a closed FRW Universe with dust. The solid curve is the exact solution, the thin dashed curve is the naive solution and the thick dashed curve is the renormalized solution.

We compare the renormalized solution (19) with the exact solution (12) and the naive solution (13) for the case of a closed Universe ($\epsilon = 1$). We choose $a_0 = 2/9$ and $c = 1$. The scale factor of the exact solution has a maximum at $t = 2\pi/9$ and goes to zero at $t = 4\pi/9$. The naive solution does not show this behavior. The renormalized solution improves the naive solution and reproduces the expanding and contracting feature of the exact solution (Fig. 1).

3 Application of Renormalization Group Method to Gradient Expansion

For the dust dominated Universe, the long-wavelength solution up to the second order of the spatial gradient becomes

$$\begin{aligned}\gamma_{ij} &= t^{4/3} \left[h_{ij} - \frac{9}{5} t^{2/3} \left(R_{ij}(h) - \frac{1}{4} R(h) h_{ij} \right) \right], \\ \rho &= \frac{4}{3t^2} \left(1 - \frac{9}{20} t^{2/3} R(h) \right), \quad u_i = 0,\end{aligned}\tag{20}$$

where h_{ij} is the seed metric which is an arbitrary function of the spatial coordinate. We can see that the perturbation term, which originated from the spatial gradient of the seed metric h_{ij} , grows as the Universe expands

and finally has the same amplitude as the background term at $t \sim H^{-1}$ when the wavelength of the perturbation equals the Hubble horizon scale. After this time, the wavelength of perturbation becomes smaller than the horizon scale and the long-wavelength expansion breaks down. To make the gradient expansion applicable to the perturbation whose wavelength is smaller than the horizon scale, we use the renormalization group method.

We renormalize the secular behavior of the three-metric γ_{ij} . By introducing a new time variable $\tau = t^{2/3}$ and the initial time τ_0 by redefining the seed metric h_{ij} , we define the renormalized metric and the counter term $h_{ij}(x) = h_{ij}(x, \mu) + \delta h_{ij}(x, \mu, \tau_0)$ as

$$\begin{aligned} h_{ij}(x) - \frac{9}{5}(\tau - \tau_0) \left(R_{ij} - \frac{1}{4}R(h)h_{ij} \right) \\ = h_{ij}(x, \mu) + \delta h_{ij} - \frac{9}{5}(\tau - \mu + \mu - \tau_0) \left(R_{ij} - \frac{1}{4}R(h)h_{ij} \right). \end{aligned} \quad (21)$$

By determining the counter term in such a way that it absorbs terms containing $\mu - \tau_0$, we have

$$h_{ij}(x, \mu) = h_{ij}(x, \tau_0) - \frac{9}{5}(\mu - \tau_0) \left(R_{ij}(h(\mu)) - \frac{1}{4}R(h(\mu))h_{ij}(\mu) \right). \quad (22)$$

This equation defines the renormalization group transformation

$$\mathcal{R}_{\mu-\tau_0} : h_{ij}(\tau_0) \rightarrow h_{ij}(\mu),$$

which is a Lie group up to $O(\epsilon)$. We can therefore get the value of $h_{ij}(\mu)$ for arbitrary μ using relation (22) by assuming the property of a Lie group. The renormalization group equation is obtained by differentiating Eq. (22) with respect to μ and setting $\tau_0 = \mu$,

$$\frac{\partial}{\partial \mu} h_{ij}(x, \mu) = -\frac{9}{5} \left[R_{ij}(h(\mu)) - \frac{1}{4}R(h(\mu))h_{ij}(\mu) \right], \quad (23)$$

such that the renormalized solution is

$$\gamma_{ij} = t^{4/3} h_{ij}(x, t), \quad \rho = \frac{4}{3t^2} + \frac{3t^{-4/3}}{10} R(h(x, t)). \quad (24)$$

We solve the renormalization group equation (23) for some special cases and see how the renormalization group method improves the behavior of the long-wavelength solution.

3.1 FRW Case

The metric is

$$h_{ij}(x, t) = \Omega^2(t) \sigma_{ij}(x), \quad R_{ij}(\sigma) = \frac{1}{3} \sigma_{ij} R(\sigma), \quad (25)$$

where $\sigma_{ij}(x)$ is the metric of the three-dimensional maximally symmetric space. In this case, the renormalization group equation (23) reduces to

$$\frac{\partial}{\partial \tau} \Omega^2(\tau) = -\frac{9}{10} k, \quad (k = \pm 1, 0), \quad (26)$$

and the renormalized solution is

$$\gamma_{ij} = t^{4/3} \left(c - \frac{9k}{10} t^{2/3} \right) \sigma_{ij}(x), \quad \rho = \frac{4}{3t^2} \left(1 + \frac{27}{20} \frac{k t^{2/3}}{c - \frac{9k}{10} t^{2/3}} \right). \quad (27)$$

Thus the scale factor of the Universe is given by

$$a(t) = t^{2/3} \sqrt{c - \frac{9k}{10} t^{2/3}}, \quad (28)$$

which is the same as the solution (19).

3.2 Spherically Symmetric Case

In spherical coordinates (r, θ, ϕ) , the metric reads

$$h_{ij} = \text{diag} \left(A^2(\tau, r), B^2(\tau, r), B^2(\tau, r) \sin^2 \phi \right). \quad (29)$$

The renormalization group equation (23) becomes for the (rr) component and the $(\theta\theta)$ component

$$2A \frac{\partial A}{\partial \tau} = \frac{9}{5} \left(\frac{A^2}{2B^2} - \frac{A_{,r} B_{,r}}{AB} - \frac{(B_{,r})^2}{2B^2} + \frac{B_{,rr}}{B} \right), \quad (30)$$

$$2B \frac{\partial B}{\partial \tau} = \frac{9}{5} \left(-\frac{1}{2} + \frac{(B_{,r})^2}{2A^2} \right), \quad (31)$$

respectively. The solution is given by

$$B = \left(1 - \frac{9\alpha(r)}{10} \tau \right)^{1/2} \beta(r), \quad A = \frac{B_{,r}}{\sqrt{1 - \alpha(r)\beta^2(r)}}, \quad (32)$$

where α and β are arbitrary functions of r . The renormalized metric and density are

$$ds^2 = -dt^2 + \frac{(t^{2/3}B_{,r})^2}{1 - \alpha\beta^2} dr^2 + (t^{2/3}B)^2 d\Omega_2^2, \quad (33)$$

$$\rho = \frac{4}{3t^2} + \frac{3}{5t^{4/3}} \frac{3\alpha\beta_{,r}(1 - \frac{9\alpha}{10}\tau) + \alpha_{,r}\beta(1 - \frac{27\alpha}{20}\tau)}{(1 - \frac{9\alpha}{10}\tau)[(1 - \frac{9\alpha}{10}\tau)\beta_{,r} - \frac{9}{20}\alpha_{,r}\beta\tau]}. \quad (34)$$

This solution corresponds to the Toleman-Bondi solution [6]. The renormalized solution reproduces very well the feature of the metric of the spherically symmetric gravitational collapse of dust.

3.3 Szekeres Solution

In Cartesian coordinates (x, y, z) , the metric is assumed to be

$$h_{ij} = \text{diag}(1, 1, A^2(\tau, x, y, z)). \quad (35)$$

The renormalization group equation (23) reduces to the following three equations for the xy , xx , and zz component

$$\begin{aligned} 0 &= -\frac{A_{,xy}}{A}, \\ 0 &= \frac{A_{,yy} - A_{,xx}}{2A}, \\ \frac{\partial A^2}{\partial \tau} &= -\frac{1}{2}A(A_{,xx} + A_{,yy}). \end{aligned} \quad (36)$$

The renormalized metric is

$$ds^2 = -dt^2 + t^{4/3} \left[dx^2 + dy^2 + \left(g(z)(x^2 + y^2) + \frac{9g(z)}{5}t^{2/3} + c(z) \right)^2 dz^2 \right]. \quad (37)$$

This is the exact solution [7] of Szekeres, which represents a one-dimensional gravitational collapse. It is known that the "naive" gradient expansion reproduces this solution by including the fourth-order spatial gradient [2]. We obtained the solution using the second-order spatial gradient with renormalization. In this case, the renormalization procedure strongly improves the naive solution.

4 Summary

After applying the renormalization group method to improve the long-time behavior of the solution of the gradient expansion, we obtained the solutions of the renormalization group equation for FRW, spherically symmetric and Szekeres cases. The behavior of the renormalized solution indicates that they describe the collapsing phase of the system qualitatively well. The renormalization group method is regarded as the procedure of system reduction. This means the renormalization group Eq. (23) is the reduced version of the original Einstein equation and describes the slow motion dynamics of the original equation. We expect that the renormalization group equation (23) has physically interesting properties and solutions which are contained in the original Einstein equation.

We can consider the cosmological back reaction problem [8,9] from the point of view of the renormalization of the fluctuation. The naive solution represents the evolution of the perturbation with the fixed background metric. By renormalizing the naive solution, the constants contained in the background solution become time dependent due to the spatial inhomogeneity. Therefore we can investigate how the spatial inhomogeneity affects the "background" metric by solving the renormalization group equation. Based on the conventional cosmological perturbation, we can describe the back reaction effect by renormalizing the second-order zero mode perturbation [10].

For the quantum dynamics of an inflationary Universe, it is possible to derive the basic equation of the stochastic approach by using the renormalization group method. In this case, the long-wavelength quantum fluctuation is renormalized to the slowly varying background classical field. We believe that the renormalization group method gives us further understanding of the inhomogeneous cosmology.

Acknowledgments

This work was supported in part by the Grant-In-Aid for Scientific Research of the Ministry of Education, Science, Sports and Culture of Japan (11640270).

References

- [1] H. Kodama and M. Sasaki, *Prog. Theor. Phys. Suppl.* **78**, 1 (1984); V.F. Mukhanov, H.A. Feldman, and R.H. Brandenberger, *Phys. Rep.* **215**,

- 203 (1992).
- [2] K. Tomita, *Prog. Theor. Phys.* **54**, 730 (1975); D.S. Salopek and J.M. Stewart, *Class. Quantum Grav.* **9**, 1943 (1992); J. Parry, D.S. Salopek, and J.M. Stewart, *Phys. Rev. D* **46**, 2872 (1994); G.L. Comer, N. Deruelle, and J. Parry, *Phys. Rev. D* **49**, 2759 (1994).
 - [3] A.H. Nayfeh, *Introduction to Perturbation Techniques* (Wiley, New York, 1993).
 - [4] C.L. Chen, N. Goldenfeld, and Y. Oono, *Phys. Rev. E* **54**, 376 (1996); T. Kunihiro, *Prog. Theor. Phys.* **94**, 503 (1995); S. Sasa, *Physica D* **108**, 45 (1997); K. Matsuba and K. Nozaki, *Phys. Rev. E* **56**, R4926 (1997); Y.Y. Yamaguchi and Y. Nambu, *Prog. Theor. Phys.* **100**, 199 (1998).
 - [5] Y. Nambu and Y.Y. Yamaguchi, *Phys. Rev. D* **60**, 104011 (1999).
 - [6] R.C. Tolman, *Proc. Natl. Acad. Sci.* **20**, 169 (1934); H. Bondi, *Mon. Not. R. Astron. Soc.* **410**, 107 (1947).
 - [7] P. Szekeres, *Comm. Math. Phys.* **41**, 55 (1975).
 - [8] R.A. Isaacson, *Phys. Rev.* **166**, 1263 (1968); T. Futamase, *Mon. Not. R. Astr. Soc.* **273**, 187 (1989); H. Russ, M.H. Soffel, M. Kasai, and G. Börner, *Phys. Rev. D* **56**, 2044 (1997).
 - [9] V.M. Mukhanov, L.R. Abramo, and R.H. Brandenberger, *Phys. Rev. Lett.* **78**, 1624 (1997); L.R. Abramo, eprint: gr-qc/9709049; L.R. Abramo, R.H. Brandenberger, and V.M. Mukhanov, *Phys. Rev. D* **56**, 3248 (1997); L.R. Abramo, *Phys. Rev. D* **60**, 064004 (1999).
 - [10] Y. Nambu, *Phys. Rev. D* **62**, 104010 (2000); Y. Nambu, eprint: gr-qc/0009005, to appear in *Phys. Rev. D*.

May 7, 2001

11:52

WorldScientific/M.B. style "mybook"

reorderedFPAF

ANALOGIES, NEW PARADIGMS, AND OBSERVATIONAL DATA AS GROWING FACTORS OF RELATIVISTIC ASTROPHYSICS

R.J. RUFFINI

*Physics Department and International Center for Relativistic Astrophysics,
University of Rome, I-00185 Roma, Italy*

E-mail: ruffini@icra.it

Patterns in the scientific development of relativistic astrophysics are analyzed with special attention to the physics and astrophysics of black holes and gamma-ray bursts.

1 Introduction

Hagen Kleinert has been a pioneer in establishing analogies among widely separated fields of theoretical physics, applying relativistic quantum field theoretical techniques, notoriously his classical work with Richard Feynman [1], to the treatment of a variety of research topics ranging from condensed matter to crystal melting, polymer physics, phase transitions, differential geometry etc. I have always been very impressed by his profound knowledge of physics and his courage in approaching so vast a research program, as testified also in his classical books [2–4]. On the occasion of his sixtieth birthday, I am dedicating to him a few brief considerations based on my work, of how the role of analogies among different fields, the establishment of new paradigms, as well as the crucial and timely arrival of observational data have marked the development of relativistic astrophysics.

2 The Birth of Relativistic Astrophysics

In 1931, in analogy with the development made in the field of atomic physics by Enrico Fermi and his school [5], Lev D. Landau [6] was led to introduce

a new paradigm in the approach of astrophysical problems. This was not based on a hypothesis chosen merely for mathematical convenience, but was founded on new concepts developed in theoretical physics. Using the Fermi-Dirac statistics to study the equilibrium of a star paved the way to analyze the latest phases of the evolution of a star. This approach has developed into an entirely new theory for understanding white dwarfs in the classical book by Chandrasekhar [7]. Subsequently, new concepts were introduced by Julius Robert Oppenheimer, using the general relativistic techniques by Tolman and the seminal ideas of George Gamow (with whom Hagen Kleinert studied from 1965 to 1967 while working on his Ph.D. thesis). I recall here the treatment of neutron stars in the classic work with Oppenheimer's student Volkoff [8]. This entire research field reached full maturity with the discovery of pulsars in 1968, especially with the discovery of the crab nebula pulsar. The observation of the period of that pulsar and his rate of slowing-down gave clear evidence for identifying the first neutron star in the galaxy. It also confirmed our understanding that the energy source of pulsars was simply the rotational energy of the neutron star. The year 1968 can definitely be considered as the date of birth of relativistic astrophysics.

I was in Princeton in those days, initially at the university as a postdoctoral fellow in the group of John Archibald Wheeler, later as a member of the Institute for Advanced Study, and finally as an instructor and assistant professor at the university. The excitement about the neutron star discovery boldly led us directly to a yet unexplored classic paper by Oppenheimer and Snyder: *On Continued Gravitational Contraction* [9]. This opened up an entire new research field to which I have dedicated all the rest of my life and which is giving, still today, some distinctively important results. I will comment in the following on a few crucial moments, the way I remember them, that influenced very much the development of relativistic astrophysics, with particular emphasis on the establishment of analogies, new paradigms and crucial observational data.

3 Analogies Between Trajectories of Cosmic Rays and Trajectories in General Relativity

An "effective potential" technique had been used very successfully by Carl Størmer in the 1930s for studying the trajectories of cosmic rays in the earth's magnetic field [10]. In the fall of 1967, Brandon Carter visited Princeton and presented his remarkable mathematical work leading to the separability

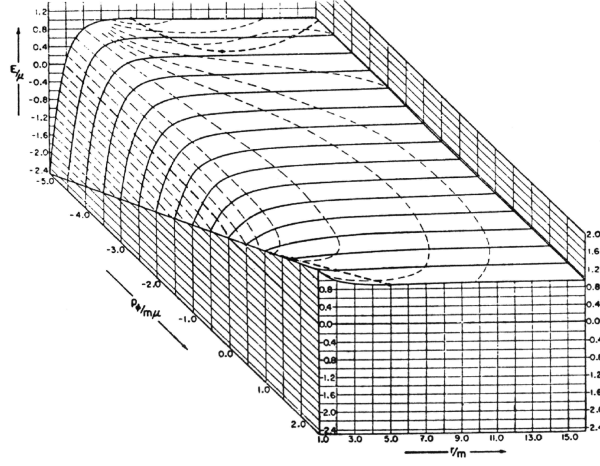


Figure 1. “Effective potential” around a Kerr black hole (reproduced from Ref. [13]).

of the Hamilton-Jacobi equations for the trajectories of charged particles in the field of a Kerr-Newmann geometry [11]. Carter’s visit had a profound impact on our small group working with John Wheeler on the physics of gravitational collapse. Indeed, it was Johnny who had the idea to exploit the analogy between the trajectories of cosmic rays and the trajectories in general relativity, using Størmer’s “effective potential” technique in order to obtain physical consequences from the set of first-order differential equations obtained by Carter. I still remember my preparing the $2m \times 2m$ grid plot of the effective potential for particles around a Kerr metric which finally appeared in print [12] (see Fig. 1). From this work came the celebrated result of the maximum binding energy of $1 - 1/\sqrt{3} \sim 42\%$ for corotating orbits, and $1 - 5/3\sqrt{3} \sim 3.78\%$ for counterrotating orbits in the Kerr geometry. We were very pleased to be associated with Brandon Carter in a “gold medal” award for this work presented by Yevgeny Lifshitz: in the last edition of Vol. 2 of the Landau and Lifshitz series, entitled *The Classical Theory of Fields*, both Brandon’s and my work with Wheeler are mentioned as exercises for bright students! This was certainly a simple and fruitful analogy which led to a successful theoretical accomplishment, opening up a new window on completely unexpected general relativistic effects. It is interesting that it has become clear in recent years that the difference in the binding energies

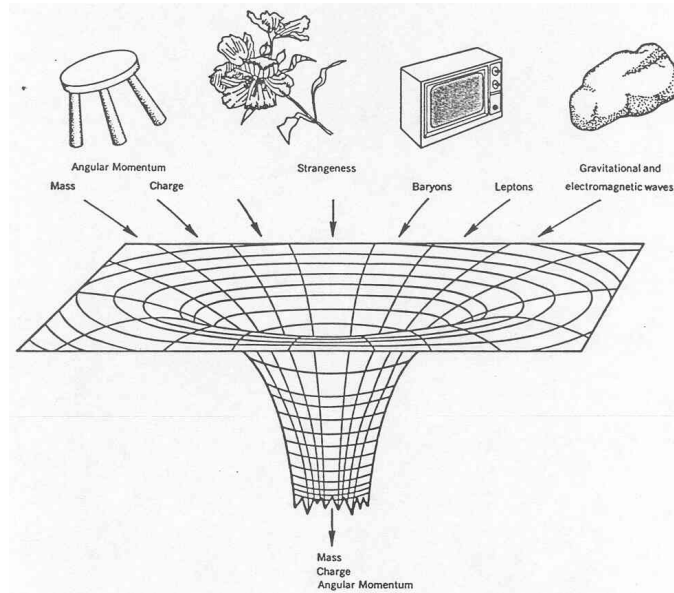


Figure 2. The black hole uniqueness theorem.

of the corotating and counterrotating orbits in a Kerr-Newmann geometry have become the object of direct astrophysical observations in binary X-ray sources.

4 Analogy of “Black Hole” with Elementary Physical System

In my 1971 article in *Physics Today* with Wheeler, entitled *Introducing the Black Hole*, we first proposed the famous “uniqueness theorem” stating that black holes can be completely characterized by their mass-energy E , charge Q , and angular momentum L [13]. This analogy between a black hole and a most elementary physical system was magnificently represented by Johnny in an unconventional figure in which TV sets, bread, flowers, and other objects loose their characteristic features and merge in the process of gravitational collapse into the three fundamental parameters of a black hole (see Fig. 2). That picture became the object of a great deal of light-hearted discussion in the physics community. A proof of this uniqueness theorem, satisfactory for some astrophysical cases, has been obtained after twenty five years of metic-

ulous mathematical work (see e.g. Regge and Wheeler [14], Zerilli [15,16], Teukolsky [17], Lee [18], and Chandrasekhar [19]). However, the proof still presents some outstanding difficulties in its most general form. Possibly, some progress will be reached in the near future with the help of computer algebraic manipulation techniques to overcome the extremely difficult mathematical calculations (see e.g. Cruciani [20], Cherubini and Ruffini [21], and Bini *et al.* [22,23]).

It is interesting that this proposed analogy, which appeared at first to be almost trivial, turned out to be one of the most difficult to be proven, implying a monumental work unsurpassed in difficulty, both in mathematical physics and relativistic field theories. An extremely good example of the difference between general relativity and classical physics is offered by the fact that this analogy is still unproven and that the most general perturbation of a black hole endowed with electromagnetic structure (EMBH) and rotation is still far from being solved. The solution of this problem from a mathematical physics point of view may have profound implications on our understanding of the fundamental physical laws.

5 Analogy Between Pulsars and Black Hole Physics: The Extraction of Rotational Energy

We were still under the sobering effects of the pulsar discovery and the very clear explanation by Tommy Gold and Arrigo Finzi that the rotational energy of the neutron star had to be the energy source of the pulsar phenomenon, when the first meeting of the European Physical Society took place in Florence in 1969. In a splendid talk, Roger Penrose [24] advanced the possibility that, much like in pulsars, the rotational energy could be in principle extracted in black holes. The first specific example of such an energy extraction process was given in a gedanken experiment using the above mentioned effective potential technique by Ruffini and Wheeler [25] (see Fig. 3), and later by Floyd and Penrose [26]. The reason for showing this figure here is a) to recall the first explicit computation and b) to remind the introduction of the "ergosphere" which is the region between the horizon of a Kerr-Newmann metric and the surface of an infinite redshift where the energy extraction process can occur, and finally c) to emphasize how contrived, difficult and conceptually novel such a mechanism of energy extraction can be. It is a phenomenon which is not localized at a point but can occur in an entire region: a global effect which relies essentially on the concept of a field. However, it can only

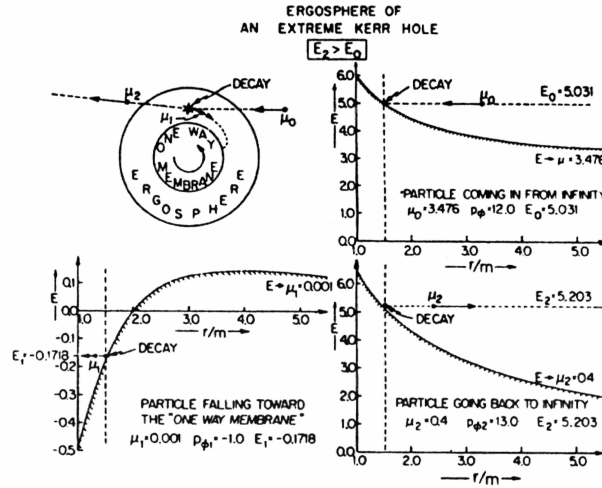


Figure 3. Decay of a particle of rest-plus-kinetic energy E_0 into a particle which is captured by the black hole with positive energy as judged locally, but negative energy E_1 as judged from infinity, together with a particle of rest-plus-kinetic energy $E_2 > E_0$ which escapes to infinity. The cross-hatched curves give the effective potential (gravitational plus centrifugal) defined by the energy E for constant values of p_ϕ and μ (reproduced from Ref. [27] with the kind permission of Ruffini and Wheeler).

work for very special parameters, and is in general associated with a reduction of the rest mass of the particle involved in the process. While it is almost trivial to slow down the rotation of a black hole and to increase its horizon by accretion of counterrotating particles, it is extremely difficult to extract the rotational energy from a black hole by a slowing-down process, as also clearly pointed out by the example in Fig. 3.

The establishment of this analogy offered us the opportunity to appreciate once more the profound difference of seemingly similar effects in general relativity and classical field theories. In addition we had the first glimpse to the existence of totally new phenomena, such as the dragging of the inertial frames around a rotating black hole, and thus to an entire new field of theoretical physics implied by the field equations of general relativity. The deep discussions of these problems with Demetrios Christodoulou, who was my first graduate student at the time in Princeton at the age of 17, led us to the discovery of the existence of "reversible and irreversible transformations" in black holes physics.

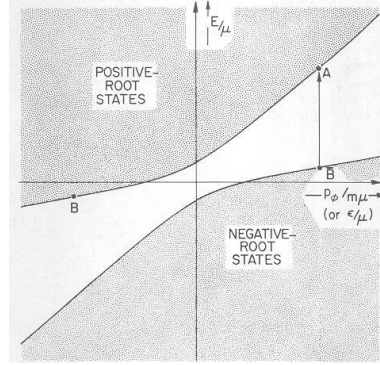


Figure 4. Reversing the effect of having added one particle A to the black hole by adding another particle B of the same rest mass but opposite angular momentum and charge in a “positive-root negative-energy state”. The addition of B is equivalent to the subtraction of B^- . Thus the combined effect of the capture of particles A and B is an increase in the mass of the black hole given by the vector B^-A . This vector vanishes and reversibility is achieved when and only when the separation between positive root states and negative root states is zero, in this case the hyperbolas coalesce to a straight line (reproduced from Ref. [28]).

6 The First Analogy Between Thermodynamics and Black Hole Physics: Reversible and Irreversible Transformations

It was by analyzing the capture of test particles by an electromagnetic black hole that we identified a set of limiting transformations which did not affect the surface area. These special transformations had to be performed very slowly, with a limiting value of zero kinetic energy on the horizon of the EMBH (see Fig. 4). It then became immediately clear that the total energy of an EMBH could in principle be expressed as a sum of “rest energy”, “Coulomb energy”, and “rotational energy”. The rest energy is “irreducible”, the other two are variable by the process of adding and extracting energy, respectively.

While Wheeler was mainly addressing the issue of the thermodynamical analogy, I and Demetrios were interested in the fundamental issue of the energetics of EMBH using the tools of the reversible and irreversible transformations. We finally obtained the general mass-energy formula for black holes [28]:

$$E^2 = M^2 c^4 = \left(M_{\text{ir}} c^2 + \frac{Q^2}{2\rho_+} \right)^2 + \frac{L^2 c^2}{\rho_+^2}, \quad (1)$$

$$S = 4\pi\rho_+^2 = 4\pi\left(r_+^2 + \frac{L^2}{c^2 M^2}\right) = 16\pi\left(\frac{G^2}{c^4}\right) M_{\text{ir}}^2, \quad (2)$$

with the condition

$$\frac{1}{\rho_+^4} \left(\frac{G^2}{c^8}\right) (Q^4 + 4L^2 c^2) \leq 1, \quad (3)$$

where M_{ir} is the irreducible mass, r_+ is the horizon radius, ρ_+ is the quasi-spheroidal cylindrical coordinate of the horizon evaluated at the equatorial plane, and S is the horizon surface area. Extreme black holes satisfy the equality in Eq. (3). The crucial point is that transformations at constant surface areas of the black hole, namely reversible transformations, can release an energy up to 29% of the mass-energy of an extremal rotating black hole and up to 50% of the mass-energy of an extremely magnetized and charged black hole. Since my Les Houches lectures *On the Energetics of Black Holes* [29], one of my main research goals has been to identify an astrophysical setting where the extractable mass-energy of the black hole could manifest itself. As we will see in the following, I propose that this extractable energy of an EMBH is the same as the energy source of gamma-ray bursts.

The thermodynamic analogy was further developed by Wheeler. By this time I had become convinced that the establishment of a one-to-one analogy between general relativistic effects and classical results was totally hopeless. It appeared to me that analogies could only be used as a very important tool to explore, possibly to be rationalized in a formula, the extremely vast theoretical world of the space-time structures contained in Einstein's theory of general relativity. A good example is our mass formula. Trying to enforce a perfect analogy is too risky and reductive.

7 Paradigm for Identifying the First "Black Hole" in Our Galaxy, and the Development of X-Ray Astronomy

The launch of the "Uhuru" satellite for the purpose of a first examination of the Universe in the X-ray spectrum, by a group directed by Riccardo Giacconi at the American Science and Engineering Department, marked a fundamental progress and generated a tremendous momentum in the field of relativistic astrophysics. The very fortunate work established simultaneous observations in the optical and in the radio wavelengths and thus allowed to have high-quality data on binary star systems composed of a normal star stripped off

matter by a compact massive companion star: either a neutron star or a black hole.

The *Maximum Mass of a Neutron Star* was the subject of the thesis of Clifford Rhoades, my second graduate student at Princeton. A criteria was found there to overcome fundamental ignorance about the behaviour of matter at supranuclear densities by establishing an absolute upper limit to the neutron star mass. Our result was only based on general relativity, causality, and the behaviour of matter at nuclear and subnuclear densities. This work, presented at the 1972 Les Houches summer school, only appeared after a prolonged debate (see the reception and publication dates!) [30]. Its results were the following:

- the “black hole uniqueness theorem”, implying the axial symmetry of the configuration and the absence of regular pulsations from black holes,
- the “effective potential technique”, determining the efficiency of the energy emission in the accretion process, and
- the “upper limit on the maximum mass of a neutron star”, discriminating between an unmagnetized neutron star and a black hole.

These three essential components established the paradigm for identifying the first black hole in Cygnus X1. The results were also presented in a widely attended session chaired by John Wheeler at the 1972 Texas Symposium in New York, extensively reported by the *New York Times*. The New York Academy of Sciences, which hosted the symposium, had just awarded me their prestigious Cressy Morrison Award for my work on neutron stars and black holes. Much to their dismay I never handed in the manuscript for the proceedings, since it coincided with a paper submitted for publication [31].

The definition of the paradigm did not come easily but matured slowly after innumerable discussions, mainly on the phone, both with Riccardo Giacconi and Herb Gursky. I still remember an irate professor of the Physics Department at Princeton pointing publicly to my outrageous phone bill of \$274 for one month, at the time considered scandalous, due to my frequent calls to the Smithsonian, and a much more relaxed and sympathetic attitude about this situation by the department chairman Murph Goldberger. This work was finally summarized in two books: one with Herbert Gursky [32], following the 1973 AAAS Annual Meeting in San Francisco, and the second with Riccardo Giacconi [33] following the 1975 LXV Enrico Fermi Summer School (see also the proceedings of the 1973 Solvay Conference).

8 More Analogies Between Thermodynamics and Black Hole Physics

The analogy between thermodynamics and general relativity became in 1971 the topic of the Ph.D. thesis of Wheeler's student Jacob Bekenstein. Through a profound set of gedanken experiments, Jacob pushed further the analogy between thermodynamics and black hole physics. Demetrios and I had formally established the existence of reversible transformations in black hole physics as well as the monotonic increase, as occurs for entropy in thermodynamics, of the irreducible mass M_{irr} of a black hole (from which the word "irreducible" arose). This was formally established independently by Steven Hawking in his area theorem [34]. The complete equivalence between the two results immediately follows from the identity relating the surface area $S = 16\pi M_{\text{irr}}^2$ of the black hole to the irreducible mass M_{irr} , as suggested by Bryce De Witt and confirmed in a quick calculation by Demetrios and myself. Jacob went one step further proposing that the area of the black hole S measured in Planck-Wheeler units could indeed be identified with the entropy [35]. He did this by formulating a statistical interpretation of the black hole entropy, and introducing the first generalized form of the first law of thermodynamics in physical processes involving black holes. Even today, thirty years later, these topics inspire lively debates. Jacob's proposal was extremely interesting and very intriguing at the time and remains so, for me, in some ways even today. It certainly was not contradictory, but I could not find a necessity for transforming it into an identity. The entire matter became the subject of even more intense discussions after Stephen Hawking proposed a physical process which, if true (in the sense used by Wigner), would transform all these theoretical conjectures into physical reality: the black hole quantum evaporation process [36]. Some three decades later, this topic also inspires lively debates. It is likely that these issues will be clarified once there is a theory encompassing both general relativity and relativistic quantum field theories. The basic properties of describing the Hawking radiation process from a black hole can simply be summarized in the following three formulae:

$$\begin{aligned} \text{radiation temperature :} & \quad T \simeq 0.62 \times 10^{-7} (M_{\odot}/M) \text{K}, \\ \text{evaporation time :} & \quad \tau \simeq E/(dE/dt) \sim 2 \times 10^{63} (M/M_{\odot})^3 \text{years}, (4) \\ \text{energy flux :} & \quad dE/dt \simeq 10^{-22} (M_{\odot}/M)^2 \text{erg/sec}. \end{aligned}$$

9 Analogy Between the Electrodynamics of Black Hole and Perfect Conducting Sphere

Before closing this set of analogies I would like to recall a gedanken experiment Richard Hanni and I made by integrating the general relativistic equations back in 1974. The idea was to compute and draw the lines of force of a test charge in the field of a Schwarzschild black hole. This was the topic of the senior thesis of Rick, then my undergraduate student at Princeton. In order to solve this problem we decided to introduce the concept of an “induced charge” on a black hole [37]. By doing so, we proposed the analogy between a black hole and a classical system with a surface equal to the horizon and endowed with an appropriate conductivity (see Figs. 5 and 6). This gedanken experiment opened the way to the field of black hole electrodynamics which was further expanded in important contributions by Damour, Hanni, Ruffini, and Wilson [38] and references therein, and by Thibault Damour. Thibault came to Princeton in 1974 to work with me on his state doctorate thesis to be defended in Paris at the École Normale Supérieure. This beautiful thesis incorporated a detailed discussion of the general relativistic effects of electrodynamics of EMBH, including generalized Ohm’s, Joule’s, Ampere’s, Navier Stoke’s laws using as a tool the above-mentioned analogy. Interestingly, this analogy was taken very seriously and written up in the book by Kip Thorne and collaborators [39] as a final theory of black holes (see, however, the subtle relativistic effects and the different conclusions reached by Brian Punsly in his recent interesting book [40]).

The analogy was reconsidered in a widely publicized article on a full page in the science section of the *Frankfurter Allgemeine Zeitung* of January 31, 2001, shown to me by Hagen Kleinert one day after its appearance, praising “new results” in an article by Wilczek and a collaborator [41]. In that paper, Wilczek, without giving any reference, purports to extend our results with Hanni on the induced charge on a Schwarzschild black hole to the case of a Reissner-Nordström geometry with charge Q . They obtained the simple result for the effective potential

$$V(r) = q \left(\frac{Q - q}{r} - \frac{qR}{r^2 - R^2} \right), \quad (5)$$

derived from the technique of introducing the image charge q . Unfortunately, they overlooked that in a very interesting set of papers, Bičák and his students Dvořák [42] and Ledvinka [43] had proven a totally new situation in the lines

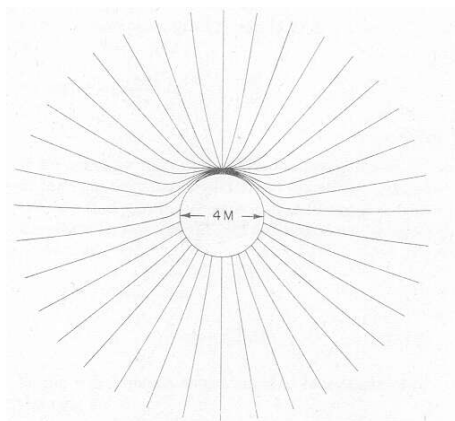


Figure 5. Lines of force of a charge near a black hole (reproduced from Ref. [37]).

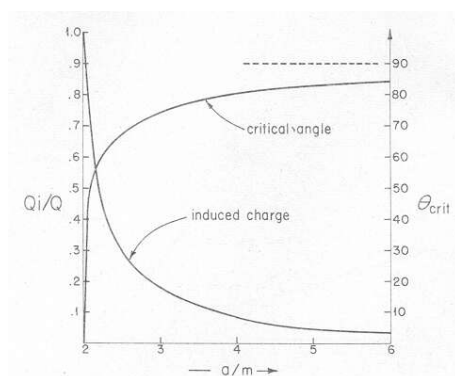


Figure 6. Induced charge on a Schwarzschild black hole (reproduced from Ref. [37]).

of force of a charge in the presence of a Reissner-Nordström geometry (see Fig. 7). In other words, in the limit $Q/M = 1$ of an extreme black hole, no line of force from the test particle crosses the horizon. We are confronted with a totally new effect for the electric field due to general relativity reminiscent of the Meissner effect in classical magnetic fields in the presence of superconductors. Such a general relativistic electric Meissner effect (GREME) does not appear to be contemplated in the Wilczek solution developed in analogy

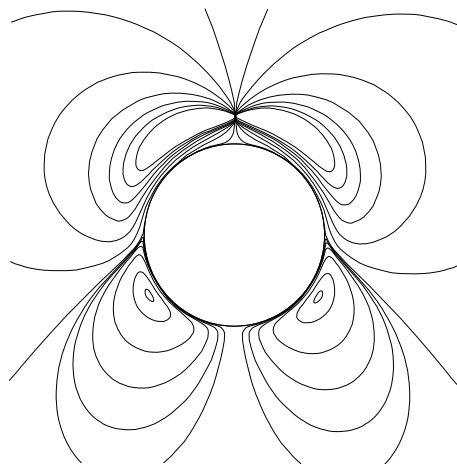


Figure 7. Lines of force of a point charge near an extreme EMBH with $Q = M$ (reproduced from Refs. [42,43]).

to my solution with Hanni. This case can be considered a propedeutic example of the incomparable richness of physical regimes present in the general relativity which, once again, transcends the direct analogies with classical field theories.

This series of events offer a clear pedagogical example of how the enforcement of direct and unproven analogies in general relativity can be dangerous and lead to incorrect conclusions.

10 Discovery of Gamma-Ray Bursts

In 1975 Herbert Gursky and myself had been invited by the AAAS to organize a session on neutron stars, black holes, and binary X-ray sources for the annual meeting in San Francisco. During the preparation of the meeting, we heard that some observations, made from the military Vela satellites and conceived in order to monitor the Limited Test Ban Treaty of 1963 on atomic bomb explosions, had just been unclassified. We asked Ian B. Strong to report, for the first time in a public meeting, on the observation of the gamma-ray bursts [32] (see Fig. 8). From the start it was clear that these signals were not coming either from the earth nor the planetary system. By 1991, a great improvement on the distribution of the gamma-ray bursts came with the

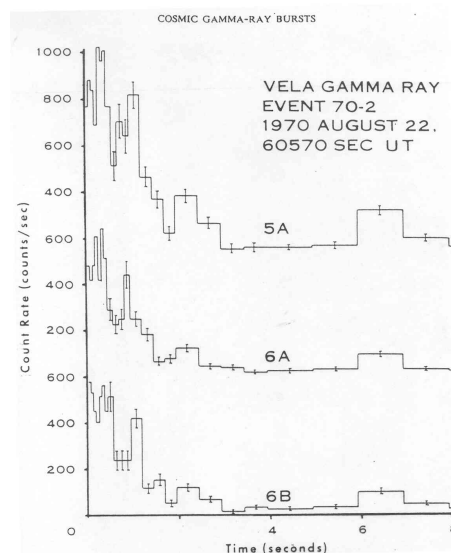


Figure 8. One of the first gamma-ray bursts observed by the Vela satellite (reproduced from an article of Strong in the book [32]).

NASA launch of the Compton Gamma-Ray Observatory which, in ten years of observation, gave a beautiful evidence for the perfect isotropy of the angular distribution of the gamma-ray burst sources in the sky (see Fig. 9). The sources had to be either at cosmological distances or very close to the solar system because one cannot detect the galactic anisotropical distribution there. In the meantime, the number of theories grew exponentially but without any clear conclusion. Quite prominent among these theories were some relating the gamma-ray burst phenomenon to the Hawking radiation process.

11 Analogy of Heisenberg-Euler Critical Capacitor and Vacuum Polarization Around Macroscopic Black Hole

In 1975, following the work of Christodoulou and Ruffini on the energetics of black holes [28], we pointed out [44] the relevance of the vacuum polarization process à la Heisenberg-Euler-Schwinger [45,46] around EMBHs. Such a process can only occur for EMBHs of mass smaller than $7.2 \times 10^6 M_\odot$. The basic energetics implications were contained in Table 1 of that paper [44], where it

2704 BATSE Gamma-Ray Bursts

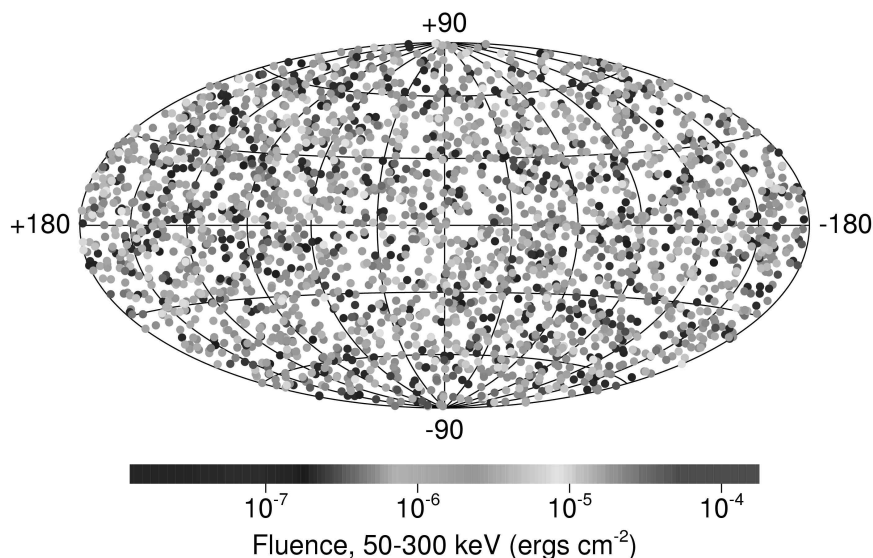


Figure 9. Angular distribution of gamma-ray bursts in galactic coordinates from the Compton GRO satellite.

was also shown that this process is almost reversible in the sense explained by Christodoulou and Ruffini [28], and that it extracts the mass energy of an EMBH very efficiently. We also pointed out that this vacuum polarization process around an EMBH offered a natural mechanism for explaining gamma-ray bursts, in particular their large characteristic energy of 10^{54} ergs (see Fig. 10).

12 Beppo-SAX Satellite and Instantaneous Destruction of More Than 135 Theoretical Models

It was only with the very unexpected and fortuitous observations of the Beppo-SAX satellite that the existence of a long-lasting afterglow of these sources was identified. This led to determining a much more accurate position for these sources in the sky, which permitted for the first time their optical and radio identification. The optical identification has led to the de-

Heisenberg, Euler, 1935 Schwinger, 1951

$$+ \left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \right] - \left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \right] \quad E_c = \frac{m^2 c^3}{\hbar e}, \quad Z_c \sim \frac{\hbar c}{e^2} \sim 137; \quad \Delta t \sim \frac{\hbar}{m_e c^2} \sim 10^{-18} s$$

$\underbrace{\hspace{10em}}_{\frac{\hbar}{m_e c}}$

Damour & Ruffini 1974, "... naturally leads to a most simple model for the explanation of the recently discovered γ -rays bursts"

$$\text{Kerr-Newmann black hole } (M, Q, a = L/M), \text{ with } a^2 + Q^2 \leq M^2$$

$$E = Q \Sigma^{-2} (r^2 - a^2 \cos^2 \vartheta) \quad B = Q \Sigma^{-2} 2ar \cos \vartheta \quad \Sigma = r^2 + a^2 \cos^2 \vartheta$$

Vacuum polarization process occurs if $3.2M_* \leq M_{BH} \leq 7.2 \cdot 10^6 M_*$

Maximum energy extractable $1.8 \cdot 10^{54} (M_{BH}/M_*) \text{ ergs}$

Figure 10. Summary of the EMBH vacuum polarization process (see Ref. [44] for details).

termination of their cosmological distances and to their paramount energetic requirements, in some cases $\geq 10^{54}$ ergs (see Ref. [47]).

The very fortunate interaction and resonance between X-ray, optical, and radio astronomy which, in the seventies, allowed the maturing of the physics and astrophysics of neutron stars and black holes (see e.g. Ref. [33]), promises to be active again in these days in unravelling the physics and astrophysics of the gamma-ray burst sources.

The observations of the Beppo-SAX satellite had a sobering effect on the theoretical development on gamma-ray burst models. Practically all existing theories (see a partial list in Fig. 11) were at once wiped out, not being able to fit the stringent energy requirements imposed by observations. They were particularly constraining for the models based on the Hawking radiation process (see Table 1 for details). These are by far the theoretical predictions furthest from any observational data in the entire history of the homo sapiens and possibly in the entire Universe.

13 Analogy Between Ergosphere and Dyadosphere of Black Holes

The enormous energy requirements of gamma-ray bursts, which are very similar to those predicted by Damour and Ruffini in 1975 [44], have convinced us to return to our earlier work in studying more accurately the process of vac-

Model #	Author	Year Pub	Reference	Main Body	2nd Body	Place	Description
1.	Colgate	1968	CfPys, 46, 5476	ST		COS	SN shocks stellar surface in distant galaxy
2.	Colgate	1974	AJ, 187, 333	ST		COS	Type II SN shock brems. env. Comp. scal. at stellar surface
3.	Snodder et al.	1977	Nature, 245, P570	ST		DISK	Stellar superflare from nearby star
4.	Harwit et al.	1973	AJ, 186, 137	WD		DISK	Superfame from nearby WD
5.	Lamb et al.	1973	Nature, 244, P552	NS		DISK	Helic. comet perturbed to collide with old galactic NS
6.	Lamb et al.	1973	Nature, 244, P552	NS		DISK	Accretion onto WD from flare in companion
7.	Lamb et al.	1973	Nature, 244, P552	NS		DISK	Accretion onto NS from flare in companion
8.	Zwicki	1974	ApJ, 187, 133	NS		DISK	Accretion onto BH from flare in companion
9.	Brecher et al.	1974	ApJ, 187, 187	NS		DISK	NS shock, contained by external pressure, escapes, explodes
10.	Schlosser	1974	SolarWind, 18, 390	ST		DISK	Helic. rot. disk grain up-batters solar radiation
11.	Brecher et al.	1974	ApJ, 187, 187	NS		DISK	Precised stellar flares on nearby stars
12.	Schlosser	1974	SolarWind, 18, 390	NS		DISK	Comet from system's cloud strikes NS
13.	Binowany et al.	1975	ApJ, 255, 29	ST		COS	Thermal emission when small star heated by SN shock wave
14.	Binowany et al.	1975	ApJ, 255, 29	ST		COS	Expected matter from NS explosion
15.	Binowany et al.	1975	ApJ, 255, 29	ST		COS	Thermal emission when small star heated by SN shock wave
16.	Binowany et al.	1975	ApJ, 255, 29	ST		COS	Expected matter from NS explosion
17.	Pacini et al.	1974	Nature, 251, 590	NS		DISK	NS crustal starquake glitch should time coincide with GRB
18.	Narkar et al.	1974	Nature, 251, 590	NS		DISK	NS crustal starquake glitch should time coincide with GRB
19.	Byrgen	1975	AAA, 44, 21	NS		HALO	NS corequake excites vibrations, changing E & B fields
20.	Channugan	1975	ApJ, 193, 175	WD		DISK	Convection inside WD with high B field produces flare
21.	Phutai et al.	1975	ApJ, 255, 395	AGN	ST	COS	Collapse of supermassive body in nucleus of active galaxy
22.	Narkar et al.	1975	ApJ, 255, 395	NS		DISK	WH accretes synchrotron emission, inverse Compton scattering
23.	Phutai et al.	1975	ApJ, 255, 395	NS		DISK	Inv. Com. scat. deep in ergosphere of fast rotating, accreting BH
24.	Channugan	1975	ApJ, 255, 395	NS		DISK	NS crustal starquake strikes NS surface
25.	Mulan	1976	ApJ, 256, 189	WD		DISK	Magnetic WD suffers MHD instabilities, flares
26.	Phutai et al.	1976	ApJ, 256, 189	WD		DISK	Thermal radiation from flare near magnetic WD
27.	Phutai et al.	1976	ApJ, 256, 189	WD		DISK	Carbon denatation from accreted matter onto NS
28.	Lamb et al.	1977	AJ, 217, 187	NS		DISK	Mag. gating of accret disk around NS causes sudden accretion
29.	Phutai et al.	1977	AJ, 217, 187	NS		DISK	Instability in accretion onto rapidly rotating BH
30.	Dagoussa	1978	ApJ, 255, 517	NS		DISK	Changed intergal. dust grain enters sol. sys. breaks up
31.	Tyrgen	1980	AAA, 87, 224	NS		DISK	WD surface nuclear burst causes chromospheric flares
32.	Tyrgen	1980	AAA, 87, 224	NS		DISK	NS surface nuclear burst causes chromospheric flares
33.	Rayner et al.	1981	ApJ, 249, 207	NS		HALO	NS neutrinos heat with to pair produce, annihilate, synch. cool
34.	Newman et al.	1980	AJ, 242, 318	NS		AST	Asieroid from interstellar medium hits NS
35.	Howard et al.	1981	AJ, 249, 302	NS		AST	Asieroid hits NS, B-field confines mass, creates high lamp
36.	Howard et al.	1981	AJ, 249, 302	NS		AST	Asieroid hits NS, B-field confines mass, creates high lamp
37.	Mendelsohn et al.	1981	AJ, 249, 302	NS		AST	Asieroid hits NS, B-field confines mass, creates high lamp
38.	Colgate	1981	AJ, 249, 302	NS		AST	Asieroid hits NS, B-field confines mass, creates high lamp
39.	van Buren	1981	AJ, 249, 302	NS		AST	Asieroid hits NS, B-field confines mass, creates high lamp
40.	Kozminsky	1982	Cosmos, 20, 72	NS		AST	Asieroid hits NS, B-field confines mass, creates high lamp
41.	Katz	1982	Cosmos, 20, 72	NS		AST	Asieroid hits NS, B-field confines mass, creates high lamp
42.	Woolsey et al.	1982	AJ, 250, 371	NS		DISK	Magnetic reconnection at heliopause
43.	Frail et al.	1982	AJ, 250, 371	NS		DISK	NS flare pair plasma confined in NS magnetosphere
44.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	Magnetic reconnection after NS surface He flash
45.	Mendelsohn et al.	1982	AJ, 250, 371	NS		DISK	He flash fusion on NS B-pole helium flash
46.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
47.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
48.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
49.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
50.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
51.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
52.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
53.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
54.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
55.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
56.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
57.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
58.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
59.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
60.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
61.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
62.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
63.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
64.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
65.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
66.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
67.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
68.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
69.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
70.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
71.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
72.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
73.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
74.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
75.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
76.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
77.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
78.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
79.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
80.	Hamary et al.	1982	AJ, 250, 371	NS		DISK	NS surface nuclear burst causes chromospheric flares
81.	Trofimov et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
82.	Sturrock et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
83.	Fennell et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
84.	Rodrigues	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
85.	Prasad et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
86.	Mata et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
87.	Trofimov	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
88.	Eichler et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
89.	Wong et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
90.	Alexander et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
91.	Mata	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
92.	Ho et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
93.	Alexander et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
94.	Demar	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
95.	Blas et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
96.	Paczynski	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
97.	Zdziarski et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
98.	Prasad	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
99.	Uffink	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
100.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
101.	Horvath et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
102.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
103.	Blas et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
104.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
105.	Woolsey et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
106.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
107.	Dar et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
108.	Thompson et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
109.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
110.	Mazares et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
111.	Eichler et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
112.	Eichler et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
113.	Carier	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
114.	Uffink	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
115.	Blas et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
116.	Narayan et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
117.	Narayan et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
118.	Blas et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
119.	Smith et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
120.	Mazares et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
121.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
122.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
123.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
124.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
125.	Hamary et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
126.	Woolsey et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
127.	Mata et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
128.	Ries et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
129.	Korpi et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
130.	Mazares et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
131.	Chang et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
132.	Mata et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
133.	Phutai et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
134.	Fabian et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares
135.	Fabian et al.	1989	ApJ, 343, 839	NS		DISK	NS surface nuclear burst causes chromospheric flares

Figure 11. Partial list of theories before the Beppo-SAX, from a talk presented at MGIXMM [48].

Table 1. Hawking radiation process versus gamma-ray burst observations. The comparison is performed for a) the energetics $E_{\text{tot}} \simeq 10M_{\odot} \simeq 10^{55}\text{erg}$, b) the time scale $\tau \simeq 1\text{ sec}$, and c) the spectrum energy $T = 10^8\text{K}$.

	theoretical value	observed value	discrepancy
a)	$T = 6.2 \times 10^{-9}\text{K}$	$T \simeq 10^8\text{K}$	$\sim 10^{-17}$
	$\tau = 10^{73}\text{sec}$	$\tau \simeq 1\text{sec}$	$\sim 10^{73}$
	$\left(\frac{dE}{dt}\right) = 10^{-24}\text{ erg/sec}$	$\left(\frac{dE}{dt}\right) \simeq 10^{54}\text{ erg/sec}$	$\sim 10^{-78}$

	theoretical value	observed value	discrepancy
b)	$E_{\text{tot}} \simeq 10^{-24}M_{\odot} \simeq 10^{30}\text{erg}$	$E_{\text{tot}} \simeq 10^{55}\text{erg}$	$\sim 10^{-25}$
	$T = 10^{17}\text{K}$	$T \simeq 10^8\text{K}$	$\sim 10^9$
	$\left(\frac{dE}{dt}\right) = 10^{26}\text{ erg/sec}$	$\left(\frac{dE}{dt}\right) \simeq 10^{54}\text{ erg/sec}$	$\sim 10^{-28}$

	theoretical value	observed value	discrepancy
c)	$E_{\text{tot}} \simeq 10^{-9}M_{\odot} \simeq 10^{45}\text{erg}$	$E_{\text{tot}} \simeq 10^{55}\text{erg}$	$\sim 10^{-10}$
	$\tau = 10^{43}\text{sec}$	$\tau \simeq 1\text{sec}$	$\sim 10^{43}$
	$\left(\frac{dE}{dt}\right) = 10^{-4}\text{ erg/sec}$	$\left(\frac{dE}{dt}\right) \simeq 10^{54}\text{ erg/sec}$	$\sim 10^{-58}$

uum polarization and the region of pair creation around an EMBH. This has led a) to the new concept of the dyadosphere of an EMBH (“dyado” from the Greek word for “pair”), b) to the concept of a plasma-electromagnetic (PEM) pulse, and c) to the analysis of its temporal evolution generating signals with the characteristic features of a gamma-ray burst.

In our theoretical approach, we claim that, by the observations of gamma-ray bursts, we are witnessing the formation of an EMBH, and that we thus follow the process of a gravitational collapse in real time. Even more important: the tremendous energies involved in the energetics of these processes

coincide with the extractable energy of black holes which is described by the Eqs. (1)–(3).

Various models have been proposed in order to extract the rotational energy of black holes by processes of relativistic magnetohydrodynamics (see e.g. Ref. [49]). It should be expected, however, that these processes are relevant over the long time scales characteristic of accretion processes.

In the present case of gamma-ray bursts, a sudden mechanism appears to be at work on time scales of the order of few seconds or shorter, and they are naturally explained by the vacuum polarization process introduced by Damour and Ruffini [44].

The fundamental new points we have found reexamining our previous work can be summarized as follows (see Ref. [50] for details):

- The vacuum polarization process can occur in an extended region around the black hole called the dyadosphere, extending from the horizon radius r_+ out to the dyadosphere radius r_{ds} . Only black holes with a mass larger than the upper limit of a neutron star and up to a maximum mass of $7.2 \times 10^6 M_\odot$ can have a dyadosphere.
- The efficiency of transforming the mass-energy of a black hole into particle-antiparticle pairs outside the horizon can approach 100% for black holes in the above mass range.
- The pairs created are mainly positron-electron pairs and their number is much larger than the quantity Q/e one would have naively expected on the grounds of qualitative considerations. It is actually given by $N_{\text{pairs}} \sim Q/e \times r_{ds}/(\hbar/mc)$, where m and e are the electron mass and charge, respectively. The energy of the pairs, and consequently the emission of the associated electromagnetic radiation, peaks in the gamma-X-ray region, as a function of the black hole mass.

Let us now recall the main results on the dyadosphere obtained by Preparata, Ruffini, and Xue [50]. Although the general considerations presented by Damour and Ruffini [44] did refer to a Kerr-Newmann field with axial symmetry around the rotation axis, we have there considered, for simplicity, the case of a non-rotating Reissner-Nordström EMBH to illustrate the basic gravitational-electrodynamical process. The dyadosphere then lies between the radius

$$r_{ds} = \left(\frac{\hbar}{mc} \right)^{\frac{1}{2}} \left(\frac{GM}{c^2} \right)^{\frac{1}{2}} \left(\frac{m_p}{m} \right)^{\frac{1}{2}} \left(\frac{e}{q_p} \right)^{\frac{1}{2}} \left(\frac{Q}{\sqrt{GM}} \right)^{\frac{1}{2}} \quad (6)$$

and the horizon radius

$$r_+ = \frac{GM}{c^2} \left(1 + \sqrt{1 - \frac{Q^2}{GM^2}} \right). \quad (7)$$

The number density of pairs created in the dyadosphere is

$$N_{e^+e^-} \simeq \frac{Q - Q_c}{e} \left[1 + \frac{(r_{ds} - r_+)}{\hbar/mc} \right], \quad (8)$$

where $Q_c = 4\pi r_+^2 m^2 c^3 / \hbar e$. The total energy of pairs converted from the static electric energy and deposited within the dyadosphere is then

$$E_{e^+e^-}^{\text{tot}} = \frac{1}{2} \frac{Q^2}{r_+} \left(1 - \frac{r_+}{r_{ds}} \right) \left[1 - \left(\frac{r_+}{r_{ds}} \right)^2 \right]. \quad (9)$$

There are many interesting analogies between the ergosphere and the dyadosphere:

- Both of them are extended regions around the black hole.
- In both regions the energy of the black hole can be extracted, approaching the limiting case of reversibility from Christodoulou and Ruffini [28].
- The electromagnetic energy extraction by the pair creation process in the dyadosphere is much simpler and less contrived than the corresponding process of extracting rotational energy from the ergosphere.

14 Analogies Between EM Pulse of Atomic Explosion and PEM Pulse of Black Holes

The analysis of the radially resolved evolution of the energy deposited within the e^+e^- -pair and photon-plasma fluid created in the dyadosphere of an EMBH is much more complex than we had initially anticipated. In this respect, the collaboration with Jim Wilson and his group at Livermore Radiation Laboratory was very important for us. Recently, we decided to join forces in a new collaboration, renewing our previous successful collaboration of 1974 [49]. We proceeded in parallel: in Rome with simple, almost analytic models, in Livermore by confirming these models with computer codes [51,52].

For the evolution we assumed the relativistic hydrodynamic equations (for details see Refs. [53,54]). We assumed the plasma fluid of e^+e^- -pairs,

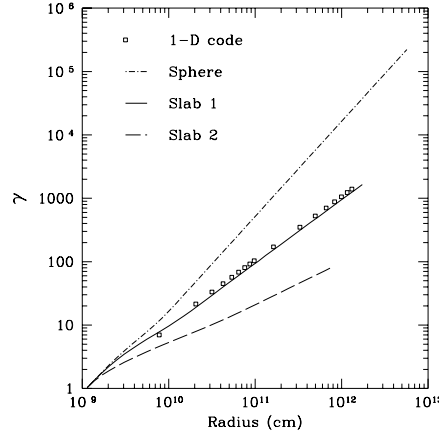


Figure 12. Lorentz factor γ as a function of radius. Three models for the expansion pattern of the e^+e^- pair plasma are compared with the results of the one-dimensional hydrodynamic code for a $1000M_\odot$ black hole with charge $Q = 0.1Q_{\text{max}}$. The 1-D code has an expansion pattern that strongly resembles that of a shell with constant coordinate thickness (reproduced from Ref. [54]).

photons and baryons to be a simple perfect fluid in curved space-time. The baryon-number and energy-momentum conservation laws are

$$(n_B U^\mu)_{;\mu} = (n_B U^t)_{,t} + \frac{1}{r^2} (r^2 n_B U^r)_{,r} = 0, \quad (10)$$

$$(T^\sigma_\mu)_{;\sigma} = 0, \quad (11)$$

with the rate equation

$$(n_{e^\pm} U^\mu)_{;\mu} = \overline{\sigma v} [n_{e^-}(T) n_{e^+}(T) - n_{e^-} n_{e^+}] , \quad (12)$$

where U^μ is the four-velocity of the plasma fluid, n_B is the proper baryon-number density, $n_{e^\pm}$ are the proper densities of electrons and positrons $e^\pm$, σ is the mean pair annihilation-creation cross-section, v is the thermal velocity of $e^\pm$, and $n_{e^\pm}(T)$ are the proper number-densities of $e^\pm$ at an appropriate equilibrium temperature T . The calculations are continued until the plasma fluid expands, cools, and the e^+e^- pairs recombine, making the system optically thin.

The results of the Livermore computer work are compared and contrasted with three almost analytical models:

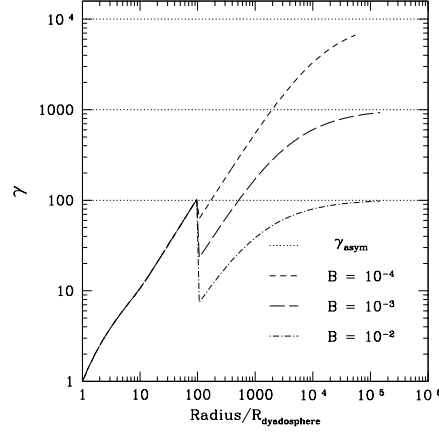


Figure 13. Lorentz factor γ as a function of radius for the PEM pulse interacting with the baryonic matter of the remnant (PEMB pulse) for selected values of the baryonic matter (reproduced from Ref. [55]).

- spherical model where the radial component of four-velocity is of the form $U(r) = Ur/\mathcal{R}$ with U being the four-velocity at the surface $\mathcal{R}$ of the plasma, similar to a portion of a Friedmann model,
- slab 1 where $U(r) = U_r = \text{const.}$, which is an expanding slab with constant width $\mathcal{D} = R_o$ in the coordinate frame in which the plasma is moving,
- slab 2 which is an expanding slab with constant width $R_2 - R_1 = R_o$ in the comoving frame of the plasma.

We compute the relativistic Lorentz factor γ of the expanding e^+e^- pair and photon plasma. In Fig. 12 we see a comparison of the Lorentz factor of the expanding fluid as a function of radius for all these models. We can see that the one-dimensional computation, of which only a few significant points are plotted, matches the expansion pattern of a shell of constant coordinate thickness.

In analogy with the notorious electromagnetic radiation EM pulse of some explosive events, we called this relativistic counterpart of an expanding pair electromagnetic radiation shell a PEM pulse. In recent works we have computed the interaction of the expanding plasma with the surrounding baryonic matter [55] (see Fig. 13). We have also been able to follow the expansion

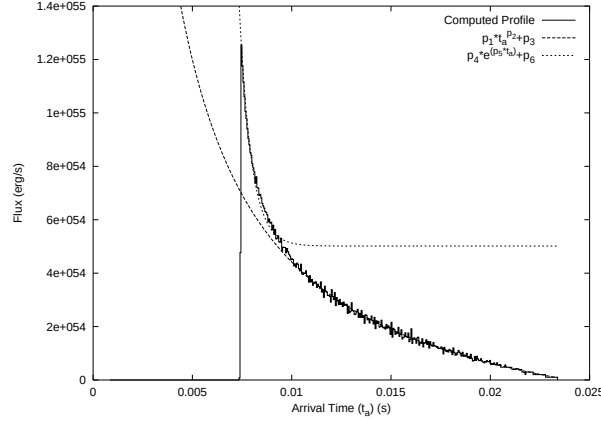


Figure 14. Proper gamma-ray burst from an EMBH with $M = 100M_{\odot}$ and $Q = 0.1Q_{\max}$ (reproduced from Ref. [56]).

process all the way to the point where the transparency condition is reached and what we have defined the “proper gamma-ray burst” is emitted [56] (see Fig. 14). These results of our theoretical model have reached the point where they can be subjected to a direct comparison with observational data.

15 Three New Paradigms for Interpreting Gamma-RAY Bursts

Starting from this theoretical background, we have moved ahead to fit, for the first time, the observational data on the ground of the EMBH model. We have used, as a prototype, the object GRB 991216, both for its very high energetics, which we have estimated in the range of $E_{\text{dya}} \sim 9.57 \times 10^{52}$ ergs, and for the superb data obtained by the Chandra and RXTE satellites. We have found a necessity to formulate in our novel approach three new paradigms, in order to understand the gamma-ray burst phenomenon:

- the relative space-time transformations (RSTT) paradigm (see Ref. [57]),
- the interpretation of the burst structure (IBS) paradigm (see Ref. [58]),
- the multiple-collapse time sequence (MCTS) paradigm (see Ref. [59]).

These results are currently under refereeing process in *Astrophysical Journal Letters* since 28/11/2000.

16 Conclusions

From the above experience, I can venture to formulate some conclusions, which may be of general validity.

16.1 On Analogies

- Analogies have been extremely helpful in establishing similarities and deepening our physical understanding, if applied to two circumstances both derived within a general relativistic framework. The analogies between dyadosphere and ergosphere are good examples.
- Analogies between classical regimes and general relativistic regimes have sometimes been helpful in giving a glance at the enormous richness of new physical processes contained in Einstein's theory of space-time structure. In some cases they have allowed us to reach new knowledge and formalize new physical laws, the derivation of Eqs. (1)–(3) being a good example. Such analogies have also dramatically evidenced the enormous differences in depth and physical complexity between classical physics and general relativistic effects. The extraction of rotational energy from a neutron star and a rotating black hole are good examples.
- In no way an analogy based on classical physics can be enforced for general relativistic regimes. Such an analogy is too constraining, and the relativistic theory shows systematically a wealth of novel physical circumstances and conceptual subtleties which are unreachable within a classical theory. The analogies in classical electrodynamics we just outlined are good examples.

16.2 On New Paradigms

The establishment of new paradigms is essential to the scientific process, and certainly not easy to obtain. Such paradigms are important in order to guide a meaningful comparison between theories and observations, and much attention should be given to their development and inner conceptual consistency.

16.3 On Observational Data

The major factors in driving the progress of scientific knowledge are always the confrontation of theoretical predictions with observational data. In recent

years, the evolution of new technologies has allowed to dramatically improve the sensitivity of observational apparatus. It is gratifying that in this process of learning the structure of our Universe, observational data do not intervene in a marginal way, but with clear and unequivocal results: they confirm by impressive agreement the correct theories and disprove, by equally impressive disagreement, the wrong ones. Our examples are also significant features of this process.

Acknowledgments

It is a pleasure to thank the many students who have, through the years, collaborated with me to reach these conclusions.

References

- [1] R.P. Feynman and H. Kleinert, *Phys. Rev. A* **34**, 5080 (1986).
- [2] H. Kleinert, *Gauge Fields in Condensed Matter*, Vol. I: *Superflow and Vortex Lines* and Vol. II: *Stresses and Defects* (World Scientific, Singapore, 1989).
- [3] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed. (World Scientific, Singapore, 1995).
- [4] H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 -Theories* (World Scientific, Singapore, 2001).
- [5] E. Fermi, *Rend. Acc. Lincei* **6**, 602 (1928).
- [6] L.D. Landau, *Phys. Zeitschr. Sov.* **1**, 285 (1932).
- [7] S. Chandrasekhar, *Introduction to the Theory of the Stellar Structure* (University of Chicago Press, Chicago, 1931).
- [8] J.R. Oppenheimer and G.M. Volkoff, *Phys. Rev.* **55**, 374 (1939).
- [9] J.R. Oppenheimer and H. Snyder, *Phys. Rev.* **56**, 455 (1939).
- [10] C. Størmer, *Astrophysica Norvegica* **1**, 1 (1934).
- [11] B. Carter, *Phys. Rev.* **174**, 1559 (1968).
- [12] M. Rees, R. Ruffini, and J.A. Wheeler, *Black Holes, Gravitational Waves, and Cosmology* (Gordon and Breach, New York, 1974); also in Russian (MIR, Moscow, 1973).
- [13] R. Ruffini and J.A. Wheeler, *Physics Today* **24**, 30 (1971).
- [14] T. Regge and J.A. Wheeler, *Phys. Rev.* **108**, 1063 (1957).
- [15] F.J. Zerilli, *Phys. Rev. D* **2**, 2141 (1970).
- [16] F.J. Zerilli, *Phys. Rev. D* **9**, 860 (1974).

- [17] S.A. Teukolsky, *Astrophys. J.* **185**, 635 (1973).
- [18] C.H. Lee, *J. Math. Phys.* **17**, 1226 (1976); *Prog. Theor. Phys.* **66**, 180 (1981).
- [19] S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Clarendon Press, Oxford, 1983); *Proc. R. Soc. Lond. A* **349**, 571 (1976).
- [20] G.L. Cruciani, in *Proceedings of the Third Icri Network Workshop*, Eds. C. Cherubini and R. Ruffini, *Nuovo Cim. B* **115**, 693 (2000).
- [21] C. Cherubini and R. Ruffini, in *Proceedings of the Third Icri Network Workshop*, Eds. C. Cherubini and R. Ruffini, *Nuovo Cim. B* **115**, 699 (2000).
- [22] D. Bini, C. Cherubini, R.T. Jantzen, and R. Ruffini, *Wave Equation for Tensor Valued p-Forms: Application to the Teukolsky Master Equation*, preprint 2001.
- [23] D. Bini, C. Cherubini, R.T. Jantzen, and R. Ruffini, *De Rham Wave Equation for the Gravitational and Electromagnetic Fields in Vacuum*, preprint 2001.
- [24] R. Penrose, *Nuovo Cim. Rivista* **1**, 252 (1969).
- [25] R. Ruffini and J.A. Wheeler, in *Proceedings of the Conference on Space Physics*, Eds. V. Hardy and H. Moore (E.S.R.O., Paris, 1971). The preparation of this report took more than one year and the authors were unwilling to publish parts of it before the final publication. In order to avoid delays, the results of the energy extraction process from a Kerr black hole, as well as the definition of the "ergosphere", were inserted as Fig. 2 in the article of Christodoulou [27] which was published on November 30, 1970.
- [26] R.M. Floyd and R. Penrose, *Nature* **229**, 177 (1971), submitted December 16, 1970.
- [27] D. Christodoulou, *Phys. Rev. Lett.* **25**, 1596 (1970).
- [28] D. Christodoulou and R. Ruffini, *Phys. Rev. D* **4**, 3552 (1971).
- [29] B. de Witt and C. de Witt (Eds.), *Black Holes* (Gordon and Breach, New York, 1973).
- [30] C. Rhoades and R. Ruffini *Phys. Rev. Lett.* **32**, 324 (1974).
- [31] R.W. Leach and R. Ruffini, *Astrophys. J.* **180**, L15 (1973).
- [32] H. Gursky and R. Ruffini, Eds. and coauthors, *Neutron Stars, Black Holes and Binary X-Ray Sources* (D. Reidel, Dordrecht, 1975).
- [33] R. Giacconi and R. Ruffini, Eds. and coauthors, *Physics and Astrophysics of Neutron Stars and Black Holes* (North Holland, Amsterdam, 1978).

- [34] S.W. Hawking, *Phys. Rev. Lett.* **26**, 1344 (1971).
- [35] J.D. Bekenstein, *Phys. Rev. D* **7**, 2333 (1973).
- [36] S.W. Hawking, *Nature* **248**, 30 (1974).
- [37] R.S. Hanni and R. Ruffini, *Phys. Rev. D* **8**, 3259 (1973).
- [38] T. Damour, R.S. Hanni, R. Ruffini, and J. Wilson, *Phys. Rev. D* **17**, 1518 (1978).
- [39] K.S. Thorne, R.H. Price, and D.A. Macdonald, *Black Holes: The Membrane Paradigm* (Yale University Press, New Haven, 1986).
- [40] B. Punsly, *Black Hole Gravitohydro-Magnetics* (Springer, Berlin, 2001).
- [41] F. Wilczek and M.K. Parikh, *Phys. Rev. Lett.* **85**, 5042 (2000).
- [42] J. Bičák and L. Dvořák, *Phys. Rev. D* **22**, 2933 (1980).
- [43] J. Bičák and T. Ledvinka, in *Proceedings of the Third Icara Network Workshop*, Eds. C. Cherubini and R. Ruffini, *Nuovo Cim. B* **115**, 739 (2000).
- [44] T. Damour and R. Ruffini, *Phys. Rev. Lett.* **35**, 463 (1975).
- [45] W. Heisenberg and H. Euler, *Z. Phys.* **98**, 714 (1935).
- [46] J. Schwinger, *Phys. Rev.* **98**, 714 (1951).
- [47] E. Costa, invited talk at *IX Marcel Grossmann Meeting on General Relativity*, Eds. V. Gurzadyan, R.T. Jantzen, and R. Ruffini (World Scientific, Singapore, 2001).
- [48] V. Gurzadyan, R.T. Jantzen, and R. Ruffini (Eds.), *Proceedings of IX Marcel Grossmann Meeting on General Relativity* (World Scientific, Singapore, 2001).
- [49] R. Ruffini and J.R. Wilson, *Phys. Rev. D* **12**, 2959 (1975).
- [50] G. Preparata, R. Ruffini, and S.-S. Xue, preprint (1998); *Astronomy and Astrophysics* **338**, L87 (1998).
- [51] J.R. Wilson, J.D. Salmonson, and G.J. Mathews, in *Gamma-Ray Bursts: 4th Huntsville Symposium*, Eds. C.A. Meegan, R.D. Preece, and T.M. Koshut (A.I.P., New York, 1998).
- [52] J.R. Wilson, J.D. Salmonson, and G.J. Mathews, in *2nd Oak Ridge Symposium on Atomic and Nuclear Astrophysics*, Ed. A. Mezzacappa (IOP Publishing, London, 1998).
- [53] R. Ruffini, J.D. Salmonson, J.R. Wilson, and S.-S. Xue, *Astronomy and Astrophysics Suppl. Ser.* **138**, 511 (1998).
- [54] R. Ruffini, J.D. Salmonson, J.R. Wilson, and S.-S. Xue, *Astronomy and Astrophysics* **350**, 334 (1999).
- [55] R. Ruffini, J.D. Salmonson, J.R. Wilson, and S.-S. Xue (2000), *Astronomy and Astrophysics* **359**, 855 (2000).

-
- [56] C.L. Bianco, R. Ruffini, and S.-S. Xue, *Astronomy and Astrophysics* **368**, 377 (2001).
 - [57] R. Ruffini, C.L. Bianco, P. Chardonnet, F. Fraschetti, and S.-S. Xue, submitted to *Astrophys. J. Lett.*
 - [58] R. Ruffini, C.L. Bianco, P. Chardonnet, F. Fraschetti, and S.-S. Xue, submitted to *Astrophys. J. Lett.*
 - [59] R. Ruffini, C.L. Bianco, P. Chardonnet, F. Fraschetti, and S.-S. Xue submitted to *Astrophys. J. Lett.*

Index

- 'T HOOFT, G., vii, xiii, 45, 55, 68, 224, 396, 622
- AARTS, J., 547
- ABDURASHITOV, J.N., 205
- ABE, H., 185, 195
- ABRAMO, L.R., 769
- Abrikosov vortex, 543
- ABRIKOSOV, A.A., 506, 531
- ABUD, M., 206
- ACHARYA, R., 197, 198, 200
- action, 3, 5, 7, 16, 18, 20, 23, 119, 120, 125, 136, 297, 435, 567, 620, 621, 665, 681, 723, 738, 745
- Aharonov-Bohm, 87
- broken symmetry phase, 746
- Chern-Simon, 161
- collective field, 508
- curvature dependence, 663
- effective, 289, 500, 502, 724
- eikonal, 60
- equilibrium, 570
- Euclidean, 44, 84
- external source, 147
- forward-backward, 21–23
- free random walk, 585
- gravitational, 699, 700, 709, 710, 712, 745
- increment, 23
- interaction, 44, 46, 47
- Jacobi, 44, 46, 47
- Kalb-Ramond field, 618
- Kepler problem, 74
- lattice, 425
- many-fermion, 286
- material, 680
- matter, 745
- monopole, 614
- multivalued, 618
- Nambu-Goto, 619, 622, 627
- non-local, 619
- permutation invariance, 287
- polaron, 284
- Polyakov-Kleinert, xi, 619, 622
- polymer, 579, 585
- principle, 680
- Regge, 711
- retarded, 285, 291
- rigidity, 663
- scalar field, 710
- string, 619, 621, 622
- trial, 284, 288, 291
- variation, 146
- adiabatic approximation, 565, 566, 569, 572, 574
- Adler conjecture, 197
- ADLER, S., 198–200
- Adler-Weisberger sum rule, 215
- AFFLECK, I., 472, 476, 483
- AGTERBERG, D.F., 550
- Aharonov-Bohm effect, 83, 84, 87–89
- AHARONY, A., 422, 462, 465, 467
- AIZENMANN, M., 431
- AKIMA, T., 550
- ALBEN, R., 466
- ALBRECHT, A., 742, 755, 756
- ALCARAZ, F.C., 482
- ALESAY, I.A., 549
- Alfaro-Fubini-Furlen-Jackiw transformation, 120

- ALFINITO, E., 611
- algebra
- chiral, 210, 211, 215
 - Grassmann, 176
 - Lie, 160–165, 169, 176
 - structure function, 124, 125
- algorithm
- ϵ , 305, 308
 - integration-by-parts, 389
 - Lanczos, 470
 - Metropolis, 32, 35, 37
 - staging, 33, 38
 - Wolff, 712
- ALLEN, P.B., 531
- α -decay, 13
- ALTARELLI, G., 206
- Alvarez potential, 629
- ALVAREZ, O., 627, 633, 643
- AMALDI, U., 205
- AMBJØRN, J., 673, 718, 720
- AMIT, D.J., 279, 466
- AMLAGE, S.M., 530
- amplitude
- fluctuation, 487, 498
 - ratio, 247, 377, 402, 414, 497
 - scattering, 260, 261
- analytic confinement, 225, 226, 232
- ANDERSON, P.W., 537, 546, 547
- anharmonic oscillator, 109, 315, 316, 322, 333, 335, 347, 349, 352, 353, 362, 363, 371, 372, 391
- anharmonicity, 337, 338, 340, 344, 345
- ANINI, Y., 705
- ANISIMOV, M.A., 411
- anisotropic vibration, 338, 341
- anomalous
- dimension, 387–389, 393, 421, 434, 436, 437, 441, 442, 499, 505, 657
 - magnetic moment, 301
- anomaly
- chiral, 175
 - dilatation, 175
 - gluon, 223
 - neutrinos, 201
 - QCD, 219
 - trace, 198, 199
 - triangle, 198
- ANSELM, A.A., 514
- ANSELMANN, P., 205
- antiferromagnetic
- correlation, 520, 526
 - order, 471, 523
 - phase, 538
 - phase transition, 518–520
 - spin fluctuation, 517, 518, 521
 - state, 537
- antiferromagnetism, 528, 529
- ANTILLÓN, A., 644
- ANTONENKO, S.A., 364, 442, 467
- ANTONI, M., 456
- ANTONOV, D., 613, 623
- ANTUNES, N.D., 575, 598, 599
- APPELQUIST, T., 495
- approach
- constituent quark, 215
 - cosmological perturbative, 759
 - dual, 494, 505, 565
 - Edward, 583, 586
 - Feynman diagram, 260
 - Parisi, 492
 - path integral, 11, 101, 117, 283–285
 - Planck scale, 726
 - Regge, 709
 - renormalization group, 458, 464
 - to scaling, 348, 355, 365, 366, 381, 384, 391, 393, 394
 - variational, 283
- approximation
- adiabatic, 565, 566, 569, 572, 574
 - dilute monopole gas, 613, 619, 621
 - FLEX, 517, 519, 523, 528, 529
 - Flory, 658
 - Gauss, 573, 583–585
 - mean-field, 509, 510, 514, 573
 - narrow resonance, 210
 - one-particle exchange, 227
 - one-string, 643

- saddle-point, 369, 571, 618, 626
semiclassical, 4, 5, 11, 16, 18, 22, 25,
98, 136, 586, 732, 737
stationary phase, 6, 10, 16–19, 23,
24
steepest-descent, 735, 737
strong-coupling, 394
Wigner, 23, 24
area
surface, 777, 778, 780
theorem, 780
ARIK, M., 563
ARKANI-HAMED, N., 195, 728
ARKO, A.J., 548
ARNEODO, M., 206
ARONOVITZ, J.A., 660
ARVANITIS, C., 313
ARVANITIS, D., 346
ASAYAMA, K., 549
ASHMAN, J., 206
ASHTEKAR, A., 739
ASPRAY, W., 40
associated memory, 271
astrophysics, xi
relativistic, 771, 772, 778
asymptotic
analysis, 760
behavior, 4, 215, 259–262, 348, 390,
417, 426, 427, 608, 732
expansion, 259, 262, 264, 366, 371
freedom, 175, 626, 697–699, 705
regime, 428
series, 301, 302, 328, 371, 390, 457,
459, 460
ATIYAH, M.F., 170
attractor, 703
AUSLOOS, M., 548
AVENEL, O., 546
AVIGNON, M., 530
axial charge, 207, 210–212, 215
AYSAL, E., 482
BAAR, D.J., 548
BABAEV, E., 507, 514, 515, 531
BABERSCHKE, K., 346
BABU, K.S., 205
BACHMANN, M., xii, 55, 57, 68, 245,
257, 298, 325, 326, 547, 706
back reaction problem, 768
background fluctuations
microwave, 741, 755
BAGNULS, C., 257, 401, 410–412, 422
BAHCALL, J.N., 205
BAKER JR., G.A., 364, 411, 422, 431,
432, 442, 467
BAKER, M., 198, 200
BALATSKY, A.V., 548, 550
BALDIT, A., 206
BALIAN, R., 575
BALLESTEROS, H.G., 422
BALOG, J., 423
Banach manifold, 149
BANAVAR, J.R., 467
BAND, W., 688, 694
BANKS, T., 495
BAR-NATAN, D., 157, 169, 170
BARASH, YU.S., 549
BARBASHOV, B.M., 633
BARBER, M.N., 482
BARDEEN, J., 508, 514, 530, 546
BARDEEN, J.M., 757
BARDEEN, W., 432
BARKER, J.A., 34, 40
barrier
penetration, 4, 5, 7, 13
potential, 4, 38, 583
property, 9
transmission, 11
BARROW, J.D., 756
BARTHOLOMEW, J., 495
BARUT, A.O., vii, 72
baryons, 208, 212, 214
BASKS, M.I., 346
BASTANIELLI, F., 68
BATCHELOR, M.T., 482
BATISTA, V.S., 26
BATLOGG, B., 530
BAUER, H., 695

- BÄUERLE, C., 598
 BAYER, P., 346
 BCS
 limit, 527
 scenario, 508, 512, 513
 superconductor, 508, 509
 theory, 469, 507, 508, 518, 533, 538, 539, 545
 BECHGAARD, K., 547
 BECKER, J.J., 466
 BEDELL, K.S., 530
 BEDNORZ, J.G., 517, 529, 536, 547
 behavior
 asymptotic, 4, 215, 259–262, 348, 390, 417, 426, 427, 608, 732
 convergence, 371, 393
 critical, 259, 264, 378, 390, 401–403, 405, 406, 410, 421, 426, 459, 464, 465, 469, 501, 580, 597, 598
 dual, 570
 finite-size, 470, 717
 infrared, 259, 260, 262, 264, 487, 489
 large-order, 301, 305, 307, 365, 366, 368, 369, 371, 658
 logarithmic, 428, 430
 mean-field, 501
 power-law, 353, 392, 449, 452–454
 Regge, 225, 229, 231
 scaling, 415, 433, 446, 448, 449, 455, 497, 499, 574, 663
 singular, 571, 573
 strong-coupling, 355, 373
 ultraviolet, 573, 574
 universal, 579
 BEIERL, W., 719
 BEKENSTEIN, J.D., 780, 797
 BELANGER, D.P., 466
 Bell inequality, xiii
 BELOHOREC, P., 423
 BENDER, C.M., 105, 109, 115, 308, 313, 322, 326, 432
 BENDER, I., 41
 Bender-Wu recursion relation, 322
 BENI, G., 346
 BENNEMANN, K.H., 517, 529, 531
 BENNETT, C., 757
 BENOBY, L.P., 432
 BEREZINSKII, V.L., 515, 591, 599
 BERG, B.A., 719, 720
 BERG, M., 139, 155
 BERGÈRE, M.C., 270, 412
 BERGERHOFF, B., 506
 BERGMANN, P.G., 694, 695
 BERGNER, Y., 563
 BERK, N.F., 523, 531
 BERMAN, L.E., 660
 BERNE, B.J., 41
 BERNIDO, C., 103
 BERNIER, M.E., 546
 BERNREUTHER, W., 432
 BERRY, M.V., 13
 BERVILLIER, C., 257, 401, 410–412, 422
 β -function, 242, 387–389, 416, 433, 434, 436, 437, 439, 441, 657, 672
 Bethe ansatz, 470, 473–475
 BETHE, H.A., 470, 482
 Bethe-Salpeter equation, 225, 228
 BETTENCOURT, L.M.A., 575, 589, 599
 BHATTACHARYA, P., 534, 546
 BHAUMIK, D., 81
 bi-metric theories of gravitation, 685, 688
 Bianchi identity, 616, 621, 688, 689, 691, 693, 746
 BIANCO, C.L., 798
 BIČÁK, J., 781, 797
 BICKERS, N.E., 531, 548
 BIEDENHARN, L.C., 124, 128
 big bang, 744, 745, 755
 binary
 fluid, 378
 star, 778
 Binder parameter, 714
 BINDER, K., 411, 720
 BINI, D., 775, 796

- bipolaron stability, 284
 BIRGENEAU, R.J., 411
 BIRMAN, J., 170
 BISHOP, D.J., 599
 black hole, 771, 773–784, 786, 789–791, 794
 BLATTER, G., 505
 Bloch
 equation, 34
 wave function, 522
 BLOCH, F., 260, 269
 blue phase, x
 BLUMBERG, G., 530
 BOETTCHER, S., 105, 115
 Bogoliubov inequality, 284, 285
 BOGOLIUBOV, N.M., 482
 BOGOLIUBOV, N.N., 396
 BOLDRIGHINI, C., 204
 Boltzmann
 distribution, 38, 272
 limit, 204
 weight, 29, 59, 238, 647
 BONALDE, I.J., 547, 549
 BONDI, H., 769
 BONN, D.A., 505, 548
 BORDEMANN, M., 563
 Borel
 resummation, 302, 367, 384
 transformation, 418
 BORN, K.D., 633, 643
 BÖRNER, G., 739, 769
 Bose
 condensate, 469, 609
 statistics, 31
 Bose-Einstein condensation, 590
 boson, 744
 composite, 469
 condensation, 601, 602, 604–607, 609–611
 field, 604
 gauge, 186, 194
 Goldstone, 208–210, 215, 509
 transformation, 604–607, 609
 BOUGOURZI, A.H., 470, 482
 Boulatov model, 708
 BOULATOV, D.V., 719
 bound states, 225–227, 229, 231, 232
 boundary
 condition, 3, 19, 34, 106, 114, 140, 285, 287, 445, 446, 625, 628–630, 633, 638, 731, 763
 dimension model, 436
 effect, 601, 608, 609, 643
 phase, 276
 value problem, 16
 BOURBAKI, N., 155
 BOURRELY, C., 206
 BOWICK, M., 661
 BOWICK, M.J., 599, 673
 BOYANOVSKY, D., 575
 BOYCE, B.R., 506
 BRANDENBERGER, R.H., 706, 768, 769
 BRANDENBURGER, R., 611
 BRANTON, D., 660
 BRAY, A.J., 456, 467
 BREDL, C.D., 547
 BRERETON, M.G., 587
 BREWER, D.F., 546
 BREWER, M.L., 26
 BRÉZIN, E., 280, 396, 423, 431, 443
 BREZINSKI, C., 385
 BRILLIANTOV, N.V., 410
 BRILLOUIN, L., 13
 BRINK, L., 673
 BROADHURST, D.J., 397, 443
 BRODSKY, S.J., 206
 BROSENS, F., 285, 297–299
 BROUGHTON, J.Q., 41
 BROWN, L.S., 81
 BROWN, R.C., 27
 Brownian
 motion, 129–131, 584
 path, 142, 145, 570, 663
 random walk, 594–596
 string, 565, 568, 597
 BRUECKNER, K.A., 546
 BRUNELLI, J., 563

- BRUSKIN, E.J., 41
 BUCCELLA, F., viii, 201, 204–207, 214, 216, 217
 BUCKLEY, I.R.C., 313
 bulk
 fermions, 185–187, 189, 190, 194
 melting, 344, 345
 BUNKOV, Y.M., 575, 611
 BURDANOV, J.A.V., 234
 BURNETT, S.S.C., 411
 BUTERA, P., 422, 425, 432
 BUTTERFIELD, J., 740
 BYTSENKO, A.A., 633
- CABIBBO, N., viii
 CABRA, C., 483
 CAFLISCH, R.G., 467
 CAI, J.M., 128
 CAI, P.Y., 128
 CAIANI, L., 456
 CALABRESE, P., 423
 CALLAN, C., 200
 Callan-Symanzik beta function, 197, 198
 CALLAWAY, D., 431
 CALLEN, E., 466
 Cameron-Martin formula, 140–144
 CAMPOSTRINI, M., 422
 CAMPUZANO, J.C., 531, 548
 CANHAM, P.B., 673
 CANNELL, D.S., 410
 canonical transformation, 96, 332, 734
 CAO, J.S., 41
 CAPOZZIELLO, S., 697, 705, 706
 CARACCILO, S., 411, 423, 432
 CARDY, J.L., 482
 CARERI, G., 562
 CARLITZ, R.D., 13
 CARLSON, C.E., 205
 CARMI, R., 599
 CARMONA, J.M., 423, 466
 carrier density, 507, 508, 514, 537
 CARRINGTON, A., 547
 Cartan formula, 151
- CARTAN, E., 694
 CARTER, B., 772, 773, 795
 CARTIER, P., 139, 155, 158, 170
 CASELLE, M., 412, 413, 423
 CASETTI, L., 456
 Casimir
 effect, 633, 699, 723, 725, 727
 energy, 626–631, 636–643, 727
 force, 639
 free energy, 637
 potential, 727
 CASTRO, E.A., 335
 Cauchy theorem, 263
 causal bound, 565, 566, 571, 572
 CAUSO, M.S., 411, 423
 CEA, P., 431
 CELEGHINI, E., 204, 216
 CELI, A., 423
 central limit theorem, 663, 664
 CEPERLEY, D.M., 37, 41, 42
 CHAICHIAN, M., 128
 CHAKRABATI, J., 467
 CHAKRAVARTY, C., 41, 42
 CHAKRAVERTY, B.K., 530
 CHANDAR, L., 599
 CHANDLER, D., 26, 38, 41
 CHANDRASEKHAR, S., 772, 775, 795, 796
 CHANG, D., 205
 CHANG, S., 189, 195
 Chaplygin gas, 558, 559, 561
 CHAPLYGIN, S., 562
 CHARDONNET, P., 798
 charge
 axial, 207, 210–212, 215
 density, 481
 density wave, 545
 dilatation, 198
 dual, 499
 fluctuation, 499
 fractional, 179
 isospin, 211, 215
 light-like, 213
 magnetic, 615

- minimal, 614
- renormalization, 503
- screening, 592
- supersymmetry, 560
- topological, 606, 607
- vortex, 591
- CHAWLA, M.S., 298
- Chebyshev process, 124, 125
- chemical potential, 481, 482, 522, 566, 584
- CHEN, C.L., 769
- CHEN, J.H., 422, 467, 505
- CHENG, B.K., 313
- CHENG, H.C., 195
- Chern-Simons
 - Abelian, 555, 562
 - field theory, 577
 - Lagrangian, 163, 165, 166
 - non-Abelian, 562
 - three-form, 161
- CHERUBINI, C., 775, 796, 797
- CHERVYAKOV, A., 43, 54, 55, 57, 68, 623
- CHETYRKIN, K.G., 245, 364, 385, 396, 397, 411, 412, 442, 443
- CHI, M.C., 466
- CHIA, E.M.E., 549
- CHIAO, R.Y., 599
- CHILD, M.S., 26
- CHIONH, E.H., xii
- chiral
 - algebra, 210, 211, 215
 - anomaly, 175
 - dynamics, 219
 - group, 207
 - limit, 221
 - quark model, 215
 - rotation, 190
 - symmetry, 207–210, 215, 216
 - transformation, 508
- CHIRUVOLU, S., 660
- CHITRE, D.M., 72
- CHOQUARD, P., 103
- CHRISTODOULOU, D., 776, 777, 780, 784–786, 790, 796
- Christoffel connection, 152, 686, 689, 693, 700, 738, 743
- CHUANG, I., 599
- CHUI, T.C.P., 364, 375, 385, 422
- ČÍŽEK, J., 314, 364
- classification group, 209, 214
- CLAYTON, M.A., 757
- Clebsch
 - decomposition, 555, 556
 - parameterization, 555, 556, 562
- CLEBSCH, A., 562
- Clebsch-Gordan coefficients, 181, 213
- CLEMENTI, C., 456
- CLOSE, F.E., 217
- clusters, 15
- COALSON, R.D., 41
- COCCO, L., 205
- coefficients
 - Clebsch-Gordan, 181, 213
 - large-order, 301
 - Ricci rotation, 686, 690
 - strong-coupling, 348, 349, 352, 366, 369, 371, 372
 - transmission and reflection, 12
 - weak-coupling, 348, 349, 351, 352, 355, 356, 360
- coherence
 - length, 528, 539
 - phase, 508, 517, 520, 529
- coherent state, 17, 18, 21, 73, 74, 76, 78, 80
- Coleman theorem, 509
- COLEMAN, S., 6, 13
- collapse
 - gravitational, 759, 767, 773, 774, 788
 - scaling, 451–453
 - state-vector, 174
- collective
 - excitation, 589, 598
 - field action, 508
 - mode, 543, 544
 - phenomena, 601

- COLLINS, P.D.B., 233
 COMELLI, G., 346
 COMER, G.L., 769
 COMI, M., 422, 425, 432
 compactification scale, 725–727
 complex
 deformations, 107
 potential, 111
 composite
 boson, 469
 fermion, 469
 field, ix, 188, 632
 operator, 435
 Compton wave length, 84, 226, 232
 condensation
 Bose-Einstein, 590
 boson, 601, 602, 604–607, 609–611
 Cooper pair, 526
 fermion pair, 507
 loop, 568
 monopole, 613
 Nambu-Goldstone mode, 601, 603
 confinement
 analytic, 225, 226, 232
 dual models, 614
 ensembles of monopoles, 613
 frequency, 295
 large-distance behavior, 626
 model, 226
 non-perturbative nature, 619
 parabolic, 293, 295, 296
 phase, 271, 276–278
 potential, 294, 296
 property, 216
 quark, 209, 216, 640
 region, 226, 232
 stability of strings, 622
 conformal
 coordinate transformation, 118
 excitation, 721, 725, 727
 field theory, 421, 469, 470, 473, 475, 707, 708
 fluctuations, 722, 727
 invariance, 699
 mapping, 368, 656
 transformation, 723
 conjecture
 Adler, 197
 Haldane, 470
 connection
 Christoffel, 152, 686, 689, 693, 700, 738, 743
 gauge, 161, 271
 spin, 743
 CONNES, A., 179, 183
 conservation
 energy, 744, 755
 energy-momentum, 688
 law, 274, 679, 688, 689, 692, 743, 747
 momentum, 210, 261
 CONSOLI, M., 431
 constant
 cosmological, 95, 96, 102, 558, 697, 699, 711, 721, 723–727, 729, 730, 733, 735, 741, 745, 750–752, 755
 effective gravitational, 699
 fine structure, 197
 Ginzburg, 503
 renormalization, 243, 353, 354, 373, 379, 380, 388, 390, 392, 437
 Sommerfeld, 542
 constituent quark, viii, 201, 203, 207–209, 214–216
 convergence
 behavior, 371, 393
 radius, 347, 366–369, 391
 convergent series, 327, 329
 Cooper pair condensation, 526
 COOPER, F., 432
 COOPER, J.R., 530
 COOPER, L.N., 508, 514, 530, 546
 coordinate
 independence, 45, 47, 53, 54
 transformation, 43, 46, 54, 118, 123, 127, 689, 690, 692, 724
 COPELAND, E.J., 599

- COQUEREAUX, R., 180, 183
 CORNISH, F.H.J., 81
 CORREIA, J., 720
 correlation
 antiferromagnetic, 520, 526
 electronic, 518
 function, 17, 18, 22, 24, 25, 33, 35, 44, 45, 47–49, 51–53, 60, 61, 236, 237, 242–244, 259, 260, 390, 393, 408, 426, 448, 450–452, 455, 462, 488, 502, 524, 525, 566, 569, 570, 582, 650, 657, 659, 670, 671
 length, 37, 390, 393, 402, 408, 409, 426, 462, 504, 566, 569, 591, 593, 714, 749
 long-range, 259, 260, 279, 413, 511
 time, 37, 38
 CORWIN, L., 183
 COSMAI, L., 431
 cosmological
 back reaction problem, 768
 constant, 95, 96, 102, 558, 697, 699, 711, 721, 723–727, 729, 730, 733, 735, 741, 745, 750–752, 755
 model, 698, 704, 722, 730, 733, 738
 path integral, 729, 730, 733
 perturbative approach, 759
 phase transition, 593
 problem, 590, 752
 singularity, 70, 698
 solution, 698
 cosmology, x, xi, 593, 601, 727, 729, 741, 742, 755, 768
 quantum, 729–731, 737
 Robertson-Walker, 70
 COSTA, E., 797
 COTTA-RAMUSINO, P., 170
 Coulomb
 dominance, 536, 538
 gas, 591, 592, 613, 614
 interaction, 521, 528, 536, 543, 615
 potential, 89, 511, 733
 problem, 34, 117, 127
 propagator, 117
 relativistic spectra, 92
 repulsion, 521
 system, 89, 301
 counterterm, 54, 389, 442
 coupling
 gauge, 202
 minimal, 499, 585, 687
 non-minimal, 681, 682, 698
 COUTURE, M., 482
 covariance
 Lorentz, 190, 692
 matrix, 148
 point canonical transformations, 734
 CRAMER, H., 335
 CREMMER, E., 175
 CREUTZ, M., 34, 35, 40
 critical
 bare mass, 408, 409
 behavior, 259, 264, 378, 390, 401–403, 405, 406, 410, 421, 426, 459, 464, 465, 469, 501, 580, 597, 598
 crossover, 402
 domain, 401
 dynamics, 449, 452–455, 589
 exponent, x, 235, 236, 243, 245, 301, 347, 348, 353, 355, 357–359, 361, 377, 387–396, 414, 416, 434, 437, 440, 457, 461, 462, 464, 465, 467, 470, 473, 474, 477, 708
 fluctuation, 497, 498
 length, 579
 phenomena, x, 237, 259, 260, 348, 353, 365, 413, 414, 433, 589–593, 598
 radius, 188, 194, 626, 630, 636
 slowing down, 37
 criticality, 268, 269, 401, 433, 435, 436, 446, 485, 486, 488, 490, 492–494
 CROSS, M.C., 483
 CRUCIANI, G.L., 775, 796

-
- crystalline membrane, 670, 672
 CUCCOLI, A., 346
 CULLEN, J.R., 466
 cumulant, 337–340, 342–344
 expansion, 322, 327, 328, 330, 331
 current
 algebra, 209, 211
 density, 562
 dilatation, 198, 199
 operator, 500
 quark, viii, 201, 203, 207, 208, 210, 213, 214, 216
 vacuum, 607
 curvature, 129, 164, 709
 scalar, 133, 700
 tensor, 164–166, 619
 time, 700
 curvature-saturated model, 697, 699
 curved space, 43

 D-brane theory, xiii
 DAGAN, F., 549
 DAGOTTO, E., 482
 DAHM, T., 531, 545, 549
 DAMOUR, T., 781, 789, 797
 Darboux theorem, 154
 dark matter, 721, 727
 DAS, A., 195, 563
 DAS, S.R., 623
 DASGUPTA, C., 495, 505
 DASHEN, R., 13
 DAVID, F., 412, 660, 661
 DAVIS, A.C., 611
 DAVIS, H.L., 26
 DAVIS JR., R., 205
 DAW, M.S., 346
 DE ALFARO, V., 118, 128
 DE BOER, J., 55, 68
 DE BOER, W., 205
 de Broglie wave length, 4
 Debye
 frequency, 518
 mass, 619
 model, 343
 screening, 615
 temperature, 343
 Debye-Waller factor, 337
 DE CALAN, C., 270, 495, 506
 DE CARVALHO, C.A., 432
 decay
 algebraic, 509
 α -, 13
 meson, 210, 212
 power-law, 446, 449, 453
 vacuum, 71
 vortex, 574
 DE CONCINI, C., 611
 deconfinement, 635–637, 643
 defects
 topological, x, 533, 534, 543, 589, 590, 593, 598, 601, 602, 606, 607, 610
 deformation parameter, 122, 127
 DE GENNES, P.G., 587, 659, 661
 DELABAERE, E., 109, 115
 DELLA SELVA, A., 204
 DE MARIA, M., 204
 density, 293, 553, 554
 boson condensate, 606
 carrier, 507, 508, 514, 537
 charge, 481
 current, 562
 electron, 481, 525, 537
 energy, 446–448, 487, 701, 713, 726, 727
 energy-momentum, 688, 692, 693
 fluctuation, 754, 755
 free energy, 47, 486, 491, 492, 500
 magnetization, 713
 mass, 586
 matrix, 18, 30, 31, 33, 36, 57–60, 315, 316, 320, 321, 327, 329
 reduced, 17, 20
 renormalization group, 470, 478
 matter, 553, 692, 699, 705, 747
 momentum, 562
 monomer, 580, 583, 586

- monopole, 613, 615, 616, 618, 621, 622, 749
- of states, 517, 518, 525, 526, 529, 539
- operator, 30
- perturbation, 754
- photon, 752
- probability, 34, 38
- radiation, 749
- scalar, 690
- spectral, 527, 639, 641
- spin wave, 523, 525, 526, 536
- string, 595
- superfluid, 525, 528, 529, 541, 543
- vacuum energy, 242, 751
- vortex, 498, 574
- wave
 - charge, 545
 - spin, 545
- dephasing effect, 25
- DE RAEDT, B., 41
- DE RAEDT, H., 41
- DE RITIS, R., 705
- derivative
 - expansion, 490, 491
 - Jackson q -, 122
 - Lie, 149
 - logarithmic, 352, 354, 355, 357, 358, 381, 382, 392, 393, 714
 - operator, 491, 742
 - Radon-Nikodym, 140
- DERRIDA, B., 466
- DERUELLE, N., 769
- DE SABBATA, V., 695
- DESCLOISEAUX, J., 587
- DESER, S., 562
- de Sitter
 - space-time, 701–703
 - Universe, 752
- DESSAU, D.S., 548
- determinant
 - Van Vleck, 6, 16, 17
- DE VEGA, H.J., 575
- DEVEREAUX, T.P., 548
- DEVREESE, J.T., 283, 285, 298, 299
- DeWitt
 - measure, 712
 - metric, 730
- DEWITT, B.S., 138, 146, 147, 175, 563, 780, 796
- DEWITT-MORETTE, C., 139, 155, 156, 158, 170
- diagram
 - chain, 66
 - chord, 160, 164, 169
 - one-particle irreducible, 236, 239–245, 247, 248
 - tadpole, 235, 236, 242–245
 - three-loop bubbles, 47, 48, 53
 - two-bubbles, 47, 52
 - vacuum, 235, 238–240, 256, 721
 - watermelon, 47, 48, 53
 - weight, 17, 160, 235, 236, 238–240, 243, 245, 409
- diagrammatic identity, 52, 53
- DIAMANTINI, M.C., 623, 633
- DICARLO, J., 548
- DICKINSON, P., 548
- DIENES, K.R., 195
- differential
 - calculus, 149
 - form, 161, 177, 178
 - Gateaux, 142
 - geometry, 149
 - Lie, 688
- DI FRANCESCO, P., 719
- dilatation
 - anomaly, 175
 - charge, 198
 - current, 198, 199
- DIMAKIS, A., 128
- dimensional regularization, 45, 46, 48, 54, 57, 387, 388, 406, 409, 435, 438, 487, 647, 650, 654
- DIMOPOULOS, S., 195, 728
- DING, H., 548
- Dirac neutrino, 202
- DIRAC, P.A.M., 100, 103

- DI SALVO, E., 233
 discontinuity, 306
 disorder, 589–591
 field theory, x, 498, 499, 535, 614
 parameter, 499
 quenched, x
 random, x
 dispersion relation, 305
 dissipation, 23
 dissipative bath, 16
 distribution, 44, 45, 49, 51
 Boltzmann, 38, 272
 DITTES, F.M., 443
 DIVECCHIA, P., 673
 divergence
 infrared, 491, 647
 logarithmic, 407, 716
 superficial, 266, 649
 ultraviolet, 46, 52–54
 DOBRESCU, B.A., 195
 DODD, M.E., 575, 598
 DOEBNER, H.D., 183
 DOHM, V., 410, 411
 DOI, M., 587
 DOKSHITZER, YU.L., 206
 DOLL, J.D., 25, 36, 41
 DOMB, C., 314, 396, 411, 431, 482, 506, 660
 DOMICHEV, A., 128
 DONIACH, S., 548
 DORA, B., 545, 550
 DÖRRE, P., 531
 DORSNER, I., 206
 double-well potential, 35, 38, 315
 DOWKER, J.S., 71, 138
 DROBA, D., 547
 DROUFFE, J.M., 280
 DRUMMOND, I.T., 432
 dual
 approach, 494, 565
 behavior, 570
 boson, 615
 charge, 499
 description, 565, 566, 589, 590
 Lagrangian, 498, 499
 model, 494, 499, 503, 504, 592, 597, 614
 picture, 567, 573
 representation, 565, 566, 574
 theory, 566, 568
 topological, 140
 duality, 497–499, 505, 572, 592, 698
 approach, 505
 equilibrium, 566
 lattice, 498
 map, Kleinert, 501
 non-equilibrium, 565, 568
 polymer and particles, 585
 transformation, 498, 614
 wave-particle, 585
 DUANE, S., 432
 DUDAS, E., 195
 DUDKA, M., 457, 467
 DUNCAN, A., 313
 DUPLANTIER, B., 660, 661
 DURHUUS, B., 663, 673
 DURRER, R., 599
 DURU, I.H., vii, 34, 40, 69, 71–73, 81, 93, 95, 102, 103, 117, 128, 301, 307, 314, 740
 Duru-Kleinert transformation, 301, 733
 DUTTA-ROY, B., 81
 DVALI, G., 195
 DVORÁK, L., 781, 797
 dyadosphere, 786, 788–790, 794
 DYAKONOV, D.I., 206
 dynamic
 exponent, 451–455
 fermion mass generation, 185, 186, 188, 193, 194
 rearrangement of symmetry, 602, 604
 scaling, 445, 446, 448
 dynamical group, vii
 dynamics
 chiral, 219
 critical, 449, 452–455, 589
 gravitational, 686, 687, 711

- Hamiltonian, 446, 448, 452, 454–456
 model *A*, 446
 molecular, 24
 Monte Carlo, 445, 446, 452
 ordering, 447–449, 452, 453, 455
 phase ordering, 445–447, 452, 454, 455
 pion, 223
 quantum brain, 279
 semiclassical, viii
 forward-backward, 15, 16, 19
 slow motion, 760, 768
 stochastic, 445, 448, 452–455
 string, 625, 636
 vibrational, 24
 Dyson equation, 241, 243–245
 DZHUNUSHALIEV, V., 706
 DZYALOSHINSKII, I.E., 531, 644
 ECKMANN, J.P., 431
 Edward
 approach, 583, 586
 model, 646
 EDWARDS, S.F., 587
 effect
 Aharonov-Bohm, 83, 84, 87–89, 92
 boundary, 601, 608, 609, 643
 Casimir, 633, 699, 723, 725, 727
 dephasing, 25
 finite-size, 450, 451, 472, 699
 Hall, 469, 482, 558
 Meissner, 520, 529, 782
 Pauli blocking, 204
 renormalization, 176
 Stark, 301
 temperature, 601, 608, 610
 topological, 83, 92, 581
 Zeeman, 301
 effective
 action, 46, 500, 502, 724
 classical
 Hamiltonian, 57–59, 61, 62, 64, 65, 67
 potential, 337, 342, 347
 cosmological constant, 697
 energy, 247, 249–251
 free energy, 489–492
 gravitational
 constant, 699
 coupling, 697
 potential, 23, 37, 187, 188, 191, 485–491, 494, 509–513, 613, 616, 618, 722, 724, 726, 727, 772, 773, 775, 776, 779, 781
 EFIMOV, G.V., 225, 234
 EGAMI, T., 466
 eigenvalue problem, 113, 402
 Einstein field equation, 697
 EINSTEIN, A., 686, 694, 695
 Einstein-Cartan theory, 685, 692, 693
 Einstein-Hilbert action, 697
 Einstein-Mayer Lagrangian, 690
 EINZEL, D., 531, 548
 electron-doped cuprates, 517, 518, 521, 522, 528, 529
 electroweak symmetry breaking, 194
 Eliashberg
 equation, 517, 523
 theory, 523
 ELIASHBERG, G.M., 531
 ELIZALDE, E., 633
 ELLEMAN, B., 549
 ELSGAETER, A., 660
 ELWORTHY, D.K., 137
 EMERY, V.J., 530, 546, 587
 ENDO, O., 346
 ENDRIAS, S., xii
 energy
 density, 446–448, 487, 701, 713, 726, 727
 estimation, 38
 ground-state, 34, 43, 45, 47, 54, 107, 109, 284, 286, 291, 292, 294–296, 301, 302, 305, 307, 309–312, 315, 330, 347, 349, 352, 353, 363, 370–372
 problem, 685
 energy-momentum

- density, 688, 692, 693
- problem, 692
- entanglement, 577
- entropy, 465, 579–582, 595, 741, 742, 752, 753, 780
- ϵ -algorithm, 305
- ϵ -expansion, 301, 379, 388–390, 393, 458, 459, 462, 463, 465
- EPSTEIN, R., 431
- equation
 - Bethe-Salpeter, 225, 228
 - Bloch, 34
 - Dyson, 241, 243–245
 - Eliashberg, 517, 523
 - Faddeev, 227
 - Friedmann, 702
 - Ginzburg-Landau, 535
 - Hamilton-Jacobi, 98, 136, 773
 - Langevin, 130, 137, 445, 447, 659
 - Lorentz, 555
 - Maxwell, 3, 603, 606
 - of motion, 447
 - renormalization group, 433, 434, 437, 465, 759, 760, 762, 763, 765–768
 - Schwinger-Dyson, 147
 - Stratonovich, 131
 - vacuum field, 700
 - variational, 332
 - Wheeler-DeWitt, 705, 730, 731, 733, 734, 736
 - Yang-Baxter, 470
- equilibrium duality, 566
- equivalence principle, 133, 685–688, 693
- ERB, A., 549
- EREMIN, I., 529
- ergodicity, 446
- ergosphere, 775, 786, 790, 794, 796
- error
 - standard Jackknife, 713
 - statistical, 32, 35, 38, 452, 454
 - systematic, 31, 32, 35, 36, 413, 415
- ERWIN, R.W., 549
- ESCANDE, D., 456
- ESPOSITO-FARESE, G., 183
- EULER, H., 797
- Everett-Wheeler many-worlds, 174
- evolution operator, 15, 16, 18, 19, 22, 77, 97, 99, 101
- EXAFS, 337
- exchange
 - identity, 159
 - spin fluctuation, 529
- excitation
 - collective, 589, 598
 - conformal, 721, 725, 727
 - electronic, 725
 - low-lying, 470
 - mode, 724
 - number operator, 330, 332
 - spectrum, 519
 - thermal, 512, 513
 - topological, 589, 590, 592–595, 598
 - vortex, 511, 596
 - zero energy, 469
- exciton
 - gravitational, 721, 723–727
- exclusion principle, 203
- expansion
 - asymptotic, 259, 262, 264, 366, 371
 - cumulant, 322, 327, 328, 330, 331
 - derivative, 490, 491
 - five-loop, 387, 388, 394, 396
 - gradient, 764
 - high-temperature, 57, 62, 64, 65, 67, 414, 425
 - Landau, 500
 - large N , 187, 434, 439, 488, 489, 502, 514
 - loop, 488, 489, 494
 - operator product, 198, 645, 648, 651
 - perturbation, 43, 45–47, 54, 301–303, 365, 370–373, 388, 391, 396
 - perturbative, 413–415, 417, 418
 - semiclassical, 130, 136

- strong-coupling, 301, 303, 305–312,
 352, 355, 366–369, 371, 391,
 425, 427, 430
 thermal, 342, 345
 Universe, 705, 742, 751, 753, 764
 variational perturbation, 390, 391
 weak-coupling, 62, 301, 304, 307,
 308, 348, 355, 357, 358, 360,
 363, 426
 Wegner, 378
 Wigner-Kirkwood, 36
 exponential map, 130
 extra dimension, 721, 727
 extrinsic curvature, 665, 667, 669, 671

 FABRIS, J., 706
 Faddeev equations, 227
 FADDEEV, L.D., 175
 Faddeev-Popov ghost, 730
 FAIRLIE, D.B., 176, 182
 FALCIONI, M., 205, 661
 FALCONE, D., 206
 FAY, D., 531
 FEDERBUSH, P., 200
 Federbush-Johnson theorem, 199
 FEIGEL'MAN, M.V., 505
 FELDMAN, D.E., 467
 FELDMAN, H.A., 768
 FERANCHUK, I.D., 327, 335
 Fermi
 energy, 292, 293, 526
 liquid, 519, 520, 538, 541, 545
 scale, 722
 statistics, 31, 772
 surface, 522, 536
 wave number, 292
 FERMI, E., 456, 771, 795
 fermion
 composite, 469
 Kaluza-Klein, 185
 pair condensation, 507
 system, 285
 FERNANDEZ, A., 431
 FERNÁNDEZ, F.M., 314, 335

 FERNANDEZ, L., 422
 FERRARI, F., 577, 587
 FERRENBURG, A., 720
 ferrimagnetism, 479, 480
 ferromagnetic
 order, 458, 459, 465
 phase, 272, 273, 426, 457, 458
 Feynman
 amplitude, 259–261, 263–265, 267,
 269
 diagram, 43, 44, 46–48, 52, 53, 59,
 65–67, 235–240, 242, 245,
 260, 269, 319, 321, 388, 389
 formula, 129, 135
 graphs, 247, 251, 256, 261, 262
 integral, 30, 44, 45, 47–51, 53, 54,
 129–134, 136, 139, 157, 161,
 259, 261
 measure, 85, 132
 path, 29, 30, 33, 39
 rules, 64, 248
 sum, 146
 volume form, 146, 149
 FEYNMAN, R.P., viii, ix, 25, 27, 29,
 40, 71, 92, 117, 128, 175, 203, 206,
 224, 297, 313, 347, 364, 377, 384,
 590, 598, 771, 795
 Feynman-Kac
 formula, 85
 functional, 285
 fibre bundle, 174
 field
 composite, ix, 188, 632
 equation, 49, 50, 52
 fluctuation, 486
 gauge, ix, 57, 161–165, 173, 179,
 275, 602–604, 608, 609, 622,
 745, 747
 ghost, 175, 176
 Higgs, 173, 178, 179, 181, 182, 252,
 253, 378, 753, 755
 operator, 147
 renormalization, 492
 theory

- disorder, x, 498, 499, 535, 614
 - non-local, 645
 - topological, 577, 581
- virton, 226
- FIELD, R.D., 203, 206
- Fierz transformation, 187
- filling factor, 481, 482
- fine structure
 - constant, 197
 - spectroscopy, 337
- finite-size
 - behavior, 470, 717
 - effect, 450, 451, 472, 699
 - scaling, 714, 715
- FINZI, A., 775
- FISCHER, Ø, 548, 549
- FISCHER, K.M., 467
- FISH, R., 467
- FISHER, D.S., 456, 483
- FISHER, M.E., 396, 410, 411, 422
- five-loop expansion, 387, 388, 394, 396
- fixed point, 457, 460, 462, 463, 465, 486
 - Gauss, 417, 501
 - ultraviolet, 197, 198
- flatness problem, 741, 742, 751
- FLEDDERJOHANN, A., 482, 483
- FLENSBURG, M., 633, 644
- FLEX approximation, 517, 519, 523, 528, 529
- Flory approximation, 658
- FLOYD, R.M., 775, 796
- fluctuation, 499, 510, 513, 514
 - amplitude, 487, 498
 - charge, 499
 - correlation length, 393
 - critical, 497, 498
 - density, 754, 755
 - field, 486
 - long wavelength, 38
 - order parameter, 485
 - phase, 510, 511, 514, 527, 528
 - polymer, 664
 - precursor, 507, 513
 - quadratic, 16
 - quantum, 341, 755, 768
 - renormalization, 768
 - sigma, 509
 - spin, 517, 518, 520–523, 527, 538
 - thermal, 236, 497, 498, 507, 579, 583
 - wave length, 753
 - zero-frequency, 59
- FLÜGGE, S., 694
- fluid
 - binary, 378
 - mechanics, 553, 561, 562
 - membrane model, 670
- flux
 - color-electric, 626, 629
 - magnetic, 89, 92, 540, 615
 - neutrino, 202
 - quantization, 92, 538, 539
 - tube, 488
- Fock space, 124, 144, 146, 147, 604
- FOILES, S.M., 346
- FOLACCI, A., 170
- FOLK, R., 457, 466, 467, 495, 506
- FOMIN, V.M., 297, 299
- FORD, J., 456
- form calculus, 175
- formula
 - Cameron-Martin, 140–144
 - Cartan, 151
 - Feynman, 129, 135
 - Feynman-Kac, 85
 - Malliavin, 142–146, 155
 - Poisson, 89
 - strong-coupling, 392
 - transmission and reflection, 4
 - Trotter, 36, 135
- FÖRSTER, D., 673
- forward-backward
 - propagator, 15
 - semiclassical dynamics, 15, 16, 19
- FOSDICK, L.D., 33, 35, 40
- FOULON, S., 298
- four-fermion interactions, 186, 187, 190, 194

-
- four-point function
 - one-particle irreducible, 235, 236, 240–243
 - Fourier path integral Monte Carlo, 36
 - fourth-order gravity, 698
 - FOX, I.A., 432
 - fractal
 - decomposition, 36
 - dimension, 658, 659, 708, 716
 - phase, 658
 - fractional charge, 179
 - FRADKIN, E., 280
 - FRANZ, W., 547
 - FRASCHETTI, F., 798
 - FRASER, C.M., 495
 - FRAUTSHI, S.C., 270
 - free
 - energy, 235, 238, 283, 284, 288, 291–293, 327–330, 332, 333, 408, 409, 416, 426, 579, 580, 592, 646, 699
 - Casimir, 637
 - density, 47, 486, 500
 - effective, 489–492
 - Ginzburg-Landau, 485, 526, 566
 - string, 595, 637
 - vortex, 592
 - particle, 57, 58, 69, 78, 120–127, 682
 - FREEDMAN, B., 34, 35, 40, 432
 - freedom
 - asymptotic, 175, 626, 697–699, 705
 - gauge, 95
 - FREEMAN, D.L., 25, 36, 41
 - FRENKEL, A.I., 346
 - FRICK, C., 432
 - Friedmann
 - equation, 702
 - model, 697, 699, 701, 704, 730
 - Universe, 731
 - Friedmann-Robertson-Walker
 - space-time, 759
 - Universe, 747, 751, 752, 759, 760, 762, 764
 - FRITZSCH, H., 205, 219, 224
 - FRÖHLICH, J., 157, 170, 431, 432, 506
 - Fröhlich polaron, 283, 285
 - frustration, 279, 539
 - FUBINI, S., 118, 128
 - FUJIKAWA, K., 95, 103
 - FUJIKAWA, T., 346
 - FUKUDA, Y., 205
 - FUKUYAMA, H., 547
 - FULDE, P., 545
 - function
 - Wigner, 23
 - zeta, 66, 67, 109, 438, 627, 629, 638
 - functional
 - differential equation, 235, 236, 242, 244
 - Feynman-Kac, 285
 - influence, 20, 21
 - integral, Witten, 157, 160–162, 166, 169
 - integration, 139
 - fundamental scale, 725
 - FURLAN, G., 118, 128
 - FURNIER, P., 530
 - FURSTENAU, H., 205
 - FUTAMASE, T., 769
 - GALLAVOTTI, G., 155, 431
 - gamma-ray burst, 771, 778, 783–786, 788, 789, 793
 - GAMOW, G., 728, 772
 - GARCIA, P.L., 183
 - GAREL, T., 482
 - GARLAND, C.W., 411
 - GARRABOS, Y., 411
 - gas
 - Coulomb, 591, 592, 613
 - Gateaux differential, 142
 - GATTO, R., 456
 - gauge
 - boson, 186, 194
 - condition, 97, 99
 - connection, 161, 271

-
- coupling, 202
 - field, ix, 57, 161–165, 173, 179, 275, 602–604, 608, 609, 622, 745, 747
 - fixing, 84, 730
 - freedom, 95
 - function, 608
 - group, 161, 163, 178, 621
 - independence, 95
 - invariance, 103, 274, 490, 613, 731
 - light-cone, 157, 561
 - Lorentz, 602
 - potential, 161, 166
 - principle, 271, 580
 - spin connection, 743
 - symmetry, 100, 271, 274, 279, 501, 756
 - theory, ix, 35, 96, 174–176, 178, 179, 212, 216, 271, 274, 276, 580, 602, 607, 678, 698
 - lattice, 219
 - transformation, 101, 274, 602, 604, 607, 608
 - unification, 201, 203
 - GAUNT, D.S., 431
 - GAUNTLETT, J., 563
 - Gauss
 - approximation, 573, 583–585
 - fixed point, 417, 501
 - potential, 555, 556
 - Gauss-Bonnet theorem, 709
 - GAVA, E., 719
 - Gegenbauer polynomial, 88, 233, 389
 - GEISER, V., 547
 - GELL-MANN, M., viii, 197, 200, 203, 205, 207, 211, 224
 - Gell-Mann/Nishijima rule, 180
 - general relativity, xiii
 - generating functional, 317
 - GENG, Z.K., 364, 375, 385, 422
 - geodesics, 150
 - geometrical
 - moduli, 721, 722
 - phase, 92
 - representation, 174
 - geometry
 - Kerr, 773
 - Kerr-Newmann, 773, 774
 - Reissner-Nordström, 781, 782
 - GEORGI, H., 201, 205, 215, 217
 - GERHARDT, C., 482, 483
 - GERMAN, G., 623, 633, 634, 643, 644
 - GERRY, C.C., 81, 128
 - GERSTENMAYER, E., 719
 - GERVAIS, J.L., 55
 - GESHKENBEIN, V.B., 505
 - GHERGHETTA, T., 195
 - GHOSH, G., 81
 - ghost
 - anticommuting, 179
 - Faddeev-Popov, 730
 - field, 175, 176
 - mode, 603
 - GIACCONI, R., 778, 779, 796
 - GIACHETTI, R., x, 298, 313, 346
 - GIANETTA, R.W., 530
 - GIANSANTI, A., 41
 - GIAPINTZAKIS, J., 548
 - GIARMETTA, R.W., 547
 - Gibbs state, 142
 - GINSBERG, D.M., 548
 - GINSPIRG, P., 719
 - Ginzburg constant, 503
 - Ginzburg-Landau
 - energy functional, 353, 378
 - equation, 535, 569
 - field theory, 565, 574, 585
 - free energy, 526, 566
 - model, x, 485–487, 494, 497, 586
 - parameter, 487, 494, 497
 - theory, 520
 - GIULINI, D., 739
 - GLASHOW, S.L., 205
 - GLAUM, K., 245
 - gluon anomaly, 223
 - GÖCKELER, M., 432
 - GODFREY, S., 233
 - GODFRIN, H., 575, 611

- GODRÈCHE, C., 412
 GOLD, T., 775
 GOLDBERG, N., 695
 GOLDBERGER, M., 779
 GOLDBERGER, W.D., 195
 GOLDENFELD, N., 505, 769
 GOLDSCHMIDT, Y.Y., 466, 467
 Goldstone
 boson, 208–210, 215, 509, 744
 mode, 746
 propagator, 248, 254
 theorem, 609
 GOLDSTONE, J., 563
 GOLNER, G.R., 423
 GORDILLO, M.C., 41, 42
 GORISHNY, S.G., 385, 396, 411, 412, 443
 GORKOV, L., 531
 GORTER, C.J., 598
 Gottfried sum rule, 203
 GOTTFRIED, K., 203, 206
 GOTTLÖBER, S., 739
 GOULD, C.M., 534, 546
 GRABERT, H., 41
 GRACEY, J.A., 433, 443
 GRADSHTEYN, I.S., 633, 740
 GRAF, M.J., 549, 550
 graphical recursion relation, 235, 236, 238, 239, 247, 248, 250, 251, 256
 GRASSBERGER, P., 432
 Grassmann
 algebra, 176
 variables, 553, 556–558
 gravitation, xi, 678, 679, 685, 686
 bi-metric theories, 685, 688
 gravitational
 action, 699, 700, 710, 712, 745
 collapse, 759, 767, 773, 774, 788
 dynamics, 686, 687, 711
 exciton, 721, 723–727
 vacuum, 755
 graviton, 186, 187, 194, 747
 Green function, 36, 60–62, 66, 70, 71, 83–89, 91, 92, 95, 99, 102, 148, 259, 260, 318–320, 327, 416, 524, 731, 734, 737
 GREEN, M.S., 396, 411, 431
 GREENBLATT, J., 26
 GREENE, J.M., 562
 GREENE, R.L., 530
 GREINER, W., 633
 GREST, G.S., 660
 GRIBBIN, J., viii
 GRIBBIN, M., viii
 GRIBOV, V.N., 206
 GRINSTEIN, G., 467
 GROMES, D., 41, 432
 GROSBERG, A.Yu., 587
 GROSCHE, C., 128, 740
 GROSS, D., 200, 514
 GROSS, M., 708, 719
 Gross-Neveu model, 507–509, 511, 513, 514
 ground-state
 energy, 34, 43, 45, 47, 54, 107, 109, 284, 286, 291, 292, 294–296, 301, 302, 305, 307, 309–312, 315, 330, 347, 349, 352, 353, 363, 370–372
 polaron, 284
 wave function, 34, 35, 315, 316, 320–326, 469
 group
 dynamical, vii
 gauge, 161, 163, 178, 621
 Lie, 765
 GRUZALSKI, G.R., 466
 GUADAGNINI, E., 170
 GUIDA, R., 314, 374, 385, 397, 409, 411, 412, 422, 466
 GUITTER, E., 660, 661
 GULLY, W.J., 546
 GUNAYDIN, M., 204
 GUNDERSEN, S.A., 195
 GÜNTHER, U., 721, 727, 728
 GUPTA, A.K., 530
 GURALNIK, G.S., 432
 GURSEY, F., 204

-
- GURSKY, H., 779, 783, 796
 GURZADYAN, V., 797
 GUTFREUND, H., 279
 GUTH, A.H., 756
 GUTSFELD, C., 412
 GÜTTINGER, H., 410
 GUTTMANN, A.J., 314, 412, 431

 HAAS, S., 545
 HABA, Z., 129, 138
 HACKL, R., 548
 hadronization
 quark theory, ix
 hadrons, 208–210
 HAGA, Y., 549
 HAI, L.X., 335
 HÁJÍČEK, P., 731, 739
 Haldane conjecture, 470
 HALDANE, F.D.M., 482
 HALFKANN, F.J., 410
 Hall effect, 469, 482, 558
 HALL, L.J., 195
 HALLIDAY, I.G., 432
 HALLIWELL, J.J., 739
 HALPERIN, B.I., 487, 495, 505, 575
 HAMAMATSU, H., 346
 HAMBER, H.W., 708, 719, 720
 HAMEROFF, S., 280
 Hamilton-Jacobi equation, 98, 136, 773
 Hamiltonian, 54
 complex, 105
 dynamics, 446, 448, 452, 454–456
 Hubbard, 521, 528, 529
 non-Hermitian, 105
 separable, 95, 96, 101, 102
 vector fields, 150
 Weinberg-Salam, 181
 HAMMER, L., 346
 HAMMOND, R.T., 677, 682, 683
 HAMPRECHT, B., 347
 HAN, T., 195
 HANNI, R.S., 781, 783, 797
 HARA, T., 431
 HARDY, G.H., 374

 HARDY, V., 796
 HARDY, W.N., 505, 548
 HARKO, T., 756
 harmonic
 oscillator, 20, 23, 33, 35, 57, 58, 64,
 66, 67, 73, 74, 80, 107, 108,
 118, 120, 121, 124, 125, 127,
 212, 317, 322, 330, 339, 732,
 733, 761, 763
 potential, 19, 45
 HARRIS, A.B., 466, 467
 HARRIS, J.M., 530
 HARRIS, R., 466
 HARRISON, E.R., 757
 HARTLE, J.B., 72, 731, 739
 HASE, M., 483
 HASEGAWA, Y., 547, 549
 HASELWIMMER, R.K.W., 549
 HASENBUSCH, M., 422, 423
 HASHIDA, J., 195
 HASHIMOTO, M., 195
 HASSLACHER, B., 13
 HAUSSLING, R., 183
 HAWKING, S.W., 731, 739, 756, 780,
 797
 HAYASHI, K., 690, 694, 695
 HEFFNER, R.H., 547
 HEGGER, R., 432
 HEHL, F.W., 184, 678, 682, 695
 HEINEN, I., 547
 HEINZ, K., 346
 HEINZ, U., 495
 Heisenberg
 field transformation, 604
 magnet, 237
 operator, 606
 HEISENBERG, W., 797
 HELFRICH, W., 673
 helicity
 magnetic, 562
 helium, ix, x, 40, 353, 358, 374,
 377–379, 381, 383, 384, 433
 HELLER, E.J., 26, 27
 HELLWARTH, R.W., 298

- Helmholtz-Lie space problem, 689
 Hepp sector, 263, 264
 HERBUT, I.F., 495, 506
 HERMAN, M.F., 17, 26, 41
 HERZBERG, G., 346
 HIBBS, A.R., 25, 71, 92, 297
 HIGASHITANI, S., 550
 Higgs
 condensate, 608
 field, 173, 178, 179, 181, 182, 252, 253, 378, 753, 755
 core, 609
 mass, 428, 608
 mechanism, 742
 meson, 176
 model, 275, 276, 596, 597
 particle, 173, 746
 phase, 276
 potential, 178
 propagator, 248, 252, 253
 scalar, 182
 sector, 743
 HIGGS, P., 611
 high- T_c
 superconductivity, x, 517–520
 superconductor, 536, 538, 545
 high-temperature expansion, 57, 62, 64, 65, 67, 425
 HILL, C.T., 195
 HINDMARSH, M., 598, 599
 HIRSCHFELD, P.J., 505, 549
 HISANO, J., 195
 HO, R., 128
 HO, T.L., 546
 Hodge decomposition theorem, 618
 Hohenberg theorem, 519
 HOHENBERG, P.C., 530
 hole-doped cuprates, 517, 518, 520, 521, 525, 528, 529
 HOLM, C., 707, 719, 720
 HOLMAN, R., 575
 holomorphic
 coordinates, 73, 74, 76, 77, 80
 map, 130
 HOLOVATCH, YU., 457, 466, 467, 495, 506
 HOLSTEIN, B.R., 3, 13
 HONECKER, A., 478, 483
 HONKONEN, J.R., 443
 Hopfield model, 271–273, 276, 277
 HOPFIELD, J.J., 279
 HOPPE, J., 563
 HORGAN, R.R., 432
 horizon, 751, 753, 755, 775–778, 781, 782, 789, 790
 crossing, 754
 problem, 741, 742, 748, 751
 scale, 748, 760, 765
 HORNREICH, R.M., 506
 HORWITZ, L., viii
 HOUGHTON, A., 547
 HOWE, P., 673
 HOWSON, M.A., 505
 Hubbard
 Hamiltonian, 521, 528, 529
 model, 481, 517, 521, 538
 Hubble
 horizon scale, 760, 765
 parameter, 701
 radius, 753, 759
 HUBER, K.P., 346
 HULME, J.S., 26
 HUMAYUN, K., 456
 Hurwitz transformation, 73–75, 77, 80
 HURWITZ, A., 81
 HUSE, D.A., 456
 HWA, T., 660
 HWANG, D.S., 182, 183
 hydrogen atom, vii, viii, xi, 69, 73, 74, 80, 83, 92, 95, 96, 101–103, 117, 118, 307, 414, 733
 hyper-Borel transformation, 367, 368
 hypercharge, 180
 hypothesis
 limit curvature, 698
 scaling, 448, 449, 593
 identity

- exchange, 159
- Jacobi, 176, 199, 555
- Noether, 688, 691
- Ward, 503, 756
- IDO, N., 548
- IIZUKA, J., 224
- IKEDA, N., 138
- IMADA, M., 41
- importance sampling, 32, 37
- IMRY, Y., 467
- INAGAKI, T., 185
- incidence matrix, 717
- indefinite oscillator, 731
- inequality
 - Bell, xiii
 - Bogoliubov, 284, 285
 - Jensen-Peierls-Feynman, 134, 135, 275, 283–285, 287, 288, 291
- inflationary
 - model, 750, 751
 - scenario, 750, 752, 755
 - Universe, 71
- influence
 - functional, 20, 21
 - phase, 290, 292, 295
- infrared
 - behavior, 259, 260, 262, 264, 487, 489
 - divergence, 491, 647
 - rearrangement method, 389
- INOMATA, A., 41, 71, 117, 128
- instability, 572, 574
- instantons, 35, 220
- integration-by-parts algorithm, 389
- interaction
 - action, 44, 46, 47
 - Coulomb, 521, 528, 536, 543, 615
 - short-range, 595
 - strong, 219
- interquark potential, 625–631, 633, 636
- invariance
 - gauge, 103, 490, 613, 731
 - reparametrization, 52, 54, 57, 95, 96, 100, 102, 665, 669, 707, 730, 747
- topological, 87, 160, 161, 582–584
- ISAACSON, R.A., 769
- isentropic, 553
- ISGUR, N., 233
- ISHAM, C., 740
- ISHIDA, K., 549
- ISHIGURO, T., 545, 547, 550
- Ising
 - exponents, 708, 717
 - magnet, 237
 - model, 33, 276, 405, 415, 417, 421, 427, 445, 446, 452, 459, 464, 465, 498, 590, 591, 596, 707, 713, 716
 - phase transition, 717
 - spins, 272, 707, 710, 717, 718
 - system, 717, 718
 - transition, 707, 711, 717, 718
 - universality class, 421
 - update, 712
- isospin, 219
 - charge, 211, 215
 - symmetry, 219, 221–223
- ISRAELSSON, U.E., 364, 375, 385, 422
- ITZYKSON, C., 719
- IVANOV, A., 327
- IVANOV, M.A., 234
- IZAWA, K., 545, 550
- IZERGIN, A.G., 482
- JACKIW, R., 118, 128, 200, 412, 553, 562, 563
- Jackson q -derivative, 122
- Jacobi
 - action, 44, 46, 47
 - identity, 176, 199, 555
 - principle, 95, 96, 99–102
- JACUCCI, G., 41
- JAGANNATHAN, A., 466, 467
- JANKE, W., ix, xii, 34, 40, 41, 301, 313, 314, 326, 364, 374, 397, 719, 720
- JANSEN, K., 432

- JANSON, S., 155
 JANSSEN, H.K., 456
 JANTZEN, R.T., 796, 797
 JANYŠKA, J., 739
 JASCH, F., 365, 374, 422
 JAYAPRAKASH, C., 466
 JENSEN, C.A., 34, 40
 Jensen-Peierls-Feynman inequality,
 134, 135, 275, 283–285, 287, 288, 291
 JEROME, D., 536, 547
 JERSAK, J., 432
 JEVICKI, A., 55
 JEZABEK, M., 205
 JIANG, YU., 634, 643
 JOAS, C., 529
 JOHNSON, K., 198, 200
 Johnson-Treiman relation, 212
 JOLICOEUR, T., 482
 JONA LASINIO, G., 514
 JONES, H.F., 115, 313, 432
 JONSSON, T., 673
 JORDAN, H.F., 40
 JUKI, Y., 549
 JULIA, B., 175
 JUNKER, G., 128
 JUNOD, A., 505

 KAC, V.G., 178, 183
 KACIR, M., 482
 KADOWAKI, K., 548
 KAGEYAMA, H., 483
 KAHANA, D., 233
 KAJANTIE, K., 597, 599
 Kalb-Ramond field action, 618
 KALLONIATIS, A.C., 234
 KALOPEL, N., 728
 KALUZA, T., 175, 194
 Kaluza-Klein
 dimension, 175
 fermions, 185
 mode, 185–187
 model, 722
 KAMAL, S., 505
 KANTOR, Y., 661

 KANTZ, H., 456
 KAPITULNIK, A., 548
 KARBACH, M., 470, 482, 483
 KARDAR, M., 659–661
 KARRA, G., 575
 KASAI, M., 769
 KASPER, U., 694
 KASTENING, B., 245–247, 257
 KATAEV, A.L., 396, 411, 443
 KAUFFMAN, L.H., 157, 169, 170
 KAULKE, M., 483
 KAVOUSSANAKI, E., 575
 KAWASHIMA, N., 422
 KAYSER, R.F., 411
 Kazakov model, 708
 KAZAKOV, D.I., 396, 397, 443
 KAZAKOV, V.A., 707, 719
 KECK, B., 412
 KEE, H.Y., 545, 550
 KEIMER, B., 549
 KELLER, J.B., 13
 KEMURIYAMA, M., 279
 KENNA, R., 432
 Kepler problem
 five-dimensional, 74, 77, 78
 one-dimensional, 71
 three-dimensional, 73, 74
 two-dimensional, 74
 KERLICK, G.D., 682
 Kerr geometry, 773
 Kerr-Newmann geometry, 773, 774
 KHANDEKAR, D.C., 740
 KHAZAN, L.S., 298
 KHOKHLOV, A.R., 587
 KHOLODENKO, A.L., 587
 KHURANA, A., 466
 KIBBLE, T.W.B., 270, 593, 598, 611,
 678, 682
 KIEFER, C., 729, 739, 740
 KIGUCHI, M., 346
 Killing
 form, 130
 vector, 150, 152, 689
 KILMISTER, C.W., 184

- KIM, J.K., 432
 KIM, S., 200
 KIM, Y.B., 545, 550
 KIMBALL, J.C., 128
 KIMURA, N., 549
 KINCAID, J.M., 431
 KING, C., 157, 170
 KING, D.M., 531, 548
 KINI, A.M., 547
 KINNE, A., 346
 KIOMETZIS, M., 495, 506, 622
 KIRKPATRICK, S., 466
 KIRTLEY, J.R., 530, 538, 548
 KITAJIMA, Y., 346
 KITAOKA, Y., 549
 KIVELSON, S.A., 530
 KLEIN, M.L., 38, 41
 KLEIN, O., 175, 194
 KLEINBERG, R.L., 546
 Kleinert
 dual Lagrangian, 499
 duality map, 501
 scaling, 499, 503, 504
 square root trick, 322
 KLEINERT, A., xii, 173, 174, 208, 224
 KLEINERT, H., vii–xi, xiii, 13, 16, 26,
 29, 34, 40, 43, 54, 55, 57, 67–69, 71,
 73, 81, 83, 93, 95, 102, 103, 117, 128,
 129, 137, 138, 170, 173, 182, 194,
 195, 200, 201, 204, 207, 214, 216,
 224, 235, 245–247, 256, 257, 260,
 269, 279, 297, 298, 301, 302, 307,
 312–315, 325–327, 335, 338, 340,
 346–349, 363–365, 374, 375, 377,
 384, 385, 388, 390, 391, 396, 397,
 410–412, 414, 422, 423, 433, 434,
 437, 442, 443, 456, 458, 466, 485,
 487, 492, 494, 495, 498, 500, 503,
 505, 506, 514, 515, 529–531, 533,
 545–547, 565, 575, 577, 587, 590,
 599, 601, 602, 622, 623, 625, 626,
 633–636, 643, 660, 673, 677,
 693–695, 697, 705, 706, 719, 720,
 728, 729, 739, 740, 771, 781, 795
 KLÉMAN, M., 575
 KLIMIN, S.N., 297, 299
 KLUK, E., 17, 26
 KLÜMPER, A., 470, 482
 KNETTER, C., 483
 Knight shift, 540, 541
 KNIZHNIK, V.G., 707, 719
 knot
 diagram, 158, 438
 formation, 581, 586
 invariant, 158–162, 582
 theory, 157, 163, 433, 434, 438, 442,
 577, 581, 582
 KOBAYASHI, K., 346
 KODAMA, H., 768
 KOGUT, J.B., 200, 396, 410, 431, 495
 KOHLER, M., 689, 694
 KOHMOTO, K., 545
 KOHMOTO, M., 548–550
 KOKALES, J.D., 530
 KOKUFUTA, E., 660
 KOLB, E.W., 756
 KOLEZHUCK, A.K., 483
 KOLNBERGER, S., 495
 KOMAROV, L.L., 335
 KONCZYKOWSKI, M., 505
 KONDOH, H., 346
 KONDRATYUK, L., 233
 KONG, C.W., 195
 KONISHI, K., 314
 KONO, H., 41
 Kontsevich integrals, 157
 KOPCZYŃSKI, W., 690, 695
 KORZHENEVSKII, A.L., 467
 KOSHUT, T.M., 797
 KOSTERLITZ, J.M., 467, 531, 591,
 599
 Kosterlitz-Thouless
 recursion relation, 592
 theory, 528
 transition, 507, 510–514, 591, 594,
 596
 KOTTCKE, M., 346
 KPZ

- exponent, 707, 708, 716–718
 result, 708, 717, 718
 scaling, 717
 KRAMER, P., 183
 KRAMERS, H.A., 13, 505
 KRASTEV, P.I., 205
 KRAUSE, H.J., 410
 KREIMER, D., 438, 443
 KREISEL, E., 694
 KREMENTSOV, N., 562
 KREMER, M., 432
 KRISHNAN, B., 719, 720
 KRISKIV, S., 728
 KRISTJANSEN, C., 720
 KRUCHININ, S., 548
 KRUPKA, D., 739
 KRUSIUS, M., 546
 KÜBERT, C., 549
 KUBYSHIN, YU.A., 443
 KUHN, O., 26
 KULIKOV, V.D., 411
 KUMAR, P., 534, 546
 KUNIHIO, T., 315, 326, 769
 KURATSUJI, H., 128
 KUSTAAHEIMO, P., 81, 103, 117, 118, 128
 Kustaanheimo-Stiefel transformation, vii, 73, 74, 80, 92, 95, 102, 117, 118
 KÜSTER, J., 246, 412
 KVASNIKOV, I.A., 298
 KWON, H.J., 547

 LABASTIDA, J.M.F., 157, 170
 LAERMANN, E., 643
 LAGENDIJK, A., 81
 Lagrangian
 dual, 498, 499
 Einstein-Mayer, 690
 LAI, S.-N., 422
 LAM, Y.M.P., 270
 Lamb shift, 414
 LAMB, H., 562
 λ -transition, 374
 LAMBIASE, G., 625, 633, 635, 643, 697, 706
 Lanczos algorithm, 470
 Landau
 expansion, 500
 level, 92
 theory, 545
 LANDAU, L.D., 197, 200, 335, 431, 495, 562, 771, 795
 LANG, C.B., 432
 LANG, M., 547
 Langevin equation, 130, 137, 445, 447, 659
 Laplace-Beltrami operator, 133, 668, 738
 large
 N expansion, 187, 434, 439, 502, 514
 extra dimensions, 186, 189
 large-order
 behavior, 301, 305, 307, 365, 366, 368, 369, 371, 658
 coefficients, 301
 LARIN, S.A., 245, 364, 385, 396, 411, 412, 442, 443
 LARKIN, A.I., 505, 514
 LARSEN, D.M., 298
 LARSON, M., 364, 375, 385, 422
 lattice
 action, 737
 calculation, 197
 duality, 498
 field theory, 37, 425
 gauge theory, 219, 271, 274
 Ising model, 421
 measures, 718
 model, 426, 565, 708
 simulation, 427, 626, 636
 spacing, 37, 44, 449
 spin, 519, 540
 systems, 481
 theory, 426
 vortex, 497, 499, 503, 539
 LAWANDE, S.V., 34, 35, 40, 740
 LAWRIE, I.D., 505, 506

- LAZZIZZERA, I., 587
 LEACH, D.H., 548
 LEACH, R.W., 796
 learning, 271, 274, 276–278
 process, 271, 279
 least action principle, 95, 96, 99, 100, 102
 LEBOWITZ, J.L., 314, 431, 482, 506, 660
 LEDERER, T., 346
 LE DOUSSAL, P., 661
 LEDVINKA, T., 781, 797
 LEE, C.-Y., 182, 183
 LEE, C.H., 775, 796
 LEE, D.M., 533, 534, 545, 546
 LEE, T.D., 55, 740
 LEE, W.C., 548
 Legendre transformation, 98, 236, 243, 249, 251, 618
 LEGGETT, A.J., 546, 548
 LEGGETT, T., 534
 LE GUILLOU, J.C., 396, 407, 412, 422, 423, 431, 442, 443
 LEI, N., 660
 LEMBERGER, T.R., 506
 LEMMENS, L.F., 285, 297–299
 length
 coherence, 528, 539
 Compton wave, 84, 226, 232
 correlation, 37, 393
 critical, 579
 de Broglie wave, 4
 polymer, 580, 586
 string, 568, 571–573, 594–596, 636, 641, 642
 wave, 753, 765
 LEONCINI, X., 456
 LEUTWYLER, H., 224
 LEVI-CIVITÀ, T., 81
 LEVIN, K., 531
 LEVIN, Y., 410
 LÉVY, M., 412, 422
 LI, B., 422
 LI, X.P., 41
 LIANG, J.Q., 128
 LIANG, R., 505, 548
 LIANG, W.Y., 530
 Lie
 algebra, 160–165, 169, 176
 derivative, 149
 differential, 688
 group, 765
 LIEB, E.H., 472, 483
 Lieb-Schulz-Mattis theorem, 469, 471, 474
 LIEBSCHER, D.-E., 694
 LIEKE, W., 547
 Lifshitz point, 499, 500, 502, 505
 LIFSHITZ, E.M., 335, 495, 562, 644
 LIFSHITZ, Y., 773
 light-cone gauge, 157, 561
 limit
 BCS, 527
 Boltzmann, 204
 chiral, 221
 classical, 136
 continuum, 37, 38, 58, 133, 197, 425, 426, 498, 666
 London, 498
 low-energy, 175, 678, 679
 low-temperature, 315, 316, 321
 replica, 464
 scaling, 373, 498, 663–665, 667, 668, 671, 672
 semiclassical, 6
 strong-coupling, x, 302, 304, 307, 315, 322, 347–349, 352, 354, 355, 357, 358, 373, 379–383, 390–394, 505
 thermodynamic, 647, 707
 zero-mass, 259, 260, 264, 586
 zero-temperature, 284, 303
 LIN, C.C., 562
 LIN, D.H., 83, 93
 LIN, S.H., 41
 LIN, X.S., 170
 LIN, Z.H., 506
 LIN-LIU, Y.R., 534, 546

-
- LINDE, A.D., 756
 LINKE, V., 445
 linked-cluster method, 430
 linking number, 583
 LIPA, J.A., 364, 375, 377, 385, 422
 LIPATOV, L.N., 206
 liquid crystal, x, 535, 590
 LIU, A., 411
 LIU, F., 456, 575
 LIU, Y.J., 466
 LIVI, R., 456
 locality of divergences, 648
 LOESER, A.G., 548
 logarithmic
 behavior, 428, 430
 correction, 421, 428
 derivative, 352, 354, 355, 357, 358, 381, 382, 392, 393, 714
 divergence, 243, 407, 716
 LOLL, R., 720
 LOMBARDO, L.W., 548
 LOMBARDO, M.-P., 200
 London
 limit, 498
 model, 498
 long wavelength fluctuation, 38
 long-range
 correlation, 259, 260, 279, 413, 511
 disorder, 589, 591
 interaction, 198, 585
 order, 272, 508, 523, 589–591
 phase coherence, 517, 520, 529
 potential, 596
 long-term memory, 274
 LONZARICH, G.G., 549
 loop
 condensation, 568
 diagram, 66
 expansion, 488, 489, 494
 LORAM, J.W., 527, 530
 Lorentz
 covariance, 190, 692
 equation, 555
 gauge, 602
 gravity, 718
 invariance, 741, 744, 745, 755, 756
 transformation, 690–692, 742–744, 746
 LORENTZ, H.A., 688, 694
 LOTT, J., 183
 LOUKO, J., 739
 LOW, F., 197, 200
 low-energy
 sector, 219, 223
 theorem, 215
 LOWENSTEIN, J.H., 269
 LU, J.P., 531
 LUBAN, M., 506
 LUBENSKY, T.C., 467, 487, 495, 505, 660
 LUIJTEN, E., 411
 LUKE, G.M., 549
 LUNDVIQUIST, S., 128
 LÜSCHER, M., 431, 633, 634
 Lüscher term, 626–628
 LUSIGNOLI, M., 204
 LUSSIER, B., 549
 LUTHER, A., 467
 LUZHKOV, A.A., 467
 LUZIN, N., 562
 LYKKEN, J.D., 195
 LYNN, J.W., 549

 M-theory, xiii, 175, 721
 MA, B.Q., 206
 MA, S.-K., 467, 487, 495, 505
 MACDONALD, D.A., 797
 MACFARLANE, A.J., 124, 128
 Mach principle, 691, 692
 MACHIDA, K., 549
 MACHIDA, T., 548
 MACKENZI, A.P., 549
 MACKEOWN, P.K., 41
 MADRAS, N., 422
 MAENO, Y., 541, 545, 549, 550
 MAEZAWA, K., 549
 MAGGIO-APRILE, I., 549
 magnet

-
- Heisenberg, 237
 - Ising, 237
 - magnetic
 - charge, 615
 - coupling constant, 620
 - domain, 465
 - field, 71, 92, 284, 285, 472, 476–478, 482, 486, 487, 499, 503, 539, 542, 543, 555
 - flux, 89, 92, 540, 615
 - helicity, 562
 - induction, 498
 - interaction, 87
 - ion, 458
 - mass, 615
 - monopole, 71
 - penetration depth, 518, 528, 539
 - phase, 458
 - quantum number, 228
 - ringing, 534
 - vortex, 494
 - magnetization plateau, 469, 471, 474, 476–478, 481
 - magnetohydrodynamics, 562
 - MAGUIEJO, J., 742, 755, 756
 - MAHESHWARI, A., 156
 - MAIER, E., 505
 - Majorana neutrino, 201, 202
 - MAK, M.K., 756
 - MAKAROV, D.E., 27
 - MAKI, K., x, 533, 546–550
 - MAKRI, N., 15, 22, 23, 25–27, 41
 - MALBOUISSON, A.P.C., 259, 270, 495, 506
 - Malliavin formula, 142–146, 155
 - MANGANO, G., 206
 - manifold, 709
 - Banach, 149
 - Riemann, 139, 149, 150
 - symplectic, 139, 149, 150
 - MANKA, R., 611
 - MANOHAR, A., 215, 217
 - MANOLOPOULOS, D.E., 26
 - MANSKE, D., 517, 531
 - many-body
 - inequality, 288
 - problem, 285, 286
 - many-fermion
 - problem, 285
 - system, 286
 - MAO, Z.Q., 549, 550
 - map
 - exponential, 130
 - holomorphic, 130
 - MARCH-RUSSEL, J., 195, 728
 - MARCHESINI, G., 432
 - MARINO, A., 705
 - MARINOV, M.S., 71
 - Markov
 - chain, 32, 35
 - process, 137
 - property, 134
 - MARKOV, M., 705
 - MARKUM, H., 719, 720
 - MARQUARD, U., 41
 - MARSHALL, D.S., 530, 548
 - MARSHALL, J.T., 298
 - MARSTON, J.B., 547
 - MARTÍN-MAYOR, V., 422
 - MARTELLINI, M., 170
 - Maslov index, 16
 - mass
 - density, 586
 - matrix, 220
 - mathematical QED, 197
 - MATHEWS, G.J., 797
 - MATSUBA, K., 769
 - MATSUDA, Y., 545
 - MATSUI, T., 271
 - MATSUMOTO, H., 611
 - MATSUMURA, D., 346
 - matter
 - action, 745
 - density, 553, 692, 699, 705, 747
 - MATTIS, D., 472, 483
 - MAUNG MAUNG, K., 233
 - Maxwell equations, 3, 603, 606
 - MAYER, W., 695

- MAZAUD, A., 547
 MAZENKO, G.F., 456, 575
 MCCREA, J.D., 184, 695
 MCDUFF, D., 156
 MCKENZIE, S., 431
 MCCLAUGHLIN, D.W., 13, 138
 MCLEARY, J., 40
 mean-field
 approximation, 509, 510, 514, 573
 behavior, 501
 potential, 487
 theory, 271, 275, 277, 278, 458, 507, 527, 538, 545
 transition temperature, 517, 520
 measure
 computer, 711
 DeWitt, 712
 Feynman, 132
 Misner, 711
 theory, 162
 Wiener, 129, 139, 141, 144–146, 666
 MEEGAN, C.A., 797
 MEIRON, D.I., 257, 364, 410, 411, 442, 467
 MEISINGER, P.N., 105, 115
 Meissner effect, 520, 529, 782
 Mellin
 representation, 259, 262, 263
 transformation, 261–265
 Melosh transformation, 207, 212
 MELOSH, H.J., 212, 215–217
 melting, 174
 bulk, 344, 345
 crystal, x
 surface, 344, 345
 MELTZER, D.E., 530
 membrane, x, 535, 561, 645–647, 650, 655, 658, 659
 configuration, 646
 crystalline, 670, 672
 fluctuation, 664
 fluid, 670
 fractal dimension, 659
 polymerized, 645, 659, 660
 representation, 645
 self-avoiding, 645, 657, 673
 solid, 645
 tethered, 659
 MENOTTI, P., 718, 720
 MERMIN, N.D., 535, 546, 547
 Mermin-Wagner theorem, 590
 MERZBACHER, E., 13
 MESCHEDE, D., 547
 meson
 approximation by string, 629
 coupling to baryon, 212
 decay, 210, 212
 Higgs, 176
 properties, 208
 pseudoscalar, 210, 219, 220
 Regge trajectory, 225
 states, 201, 211
 vector, 219–222
 METHA, M.L., 719
 method
 infrared arrangement, 389
 linked-cluster, 430
 Monte Carlo, 15, 18, 19, 30, 31, 33, 35, 39, 40
 operator, 327, 332, 668
 Padé, 418, 429
 Padé-Borel, 434, 462, 463, 465
 renormalization, 434
 replica, 273, 459, 585
 metric-affine theories of gravitation, 685
 Metropolis
 algorithm, 32, 35, 37
 sampling, 34
 METROPOLIS, N., 30, 40
 METZLER, R., 346
 METZNER, W., 547
 MEYER, H., 313
 MEYERSON, B., 495
 MEZINCESCU, G.A., 109, 115
 MIELE, G., 205, 206
 MIELKE, E.W., 184
 MIGLIORE, G., 206

- MIGUCHI, H., 194, 195
 MIKKELSEN, A., 660
 MILAT, O., 547
 MILLER, W.H., 22, 23, 26, 27, 41
 MILLIS, A.J., 523, 530
 MILLS, R.L., 174
 MINEEV, V.P., 549
 minimal
 area, 619
 charge, 614
 coupling, 499, 585, 681, 687
 sensitivity principle, 323, 349
 subtraction scheme, 243, 378, 387,
 405, 406, 415, 435, 655, 657
 surface, 150
 MINKOWSKI, P., 205, 224
 MINTCHEV, M., 170
 MIRZA, K.A., 530
 Misner measure, 711
 MITROVIC, B., 531
 mixing, 219–221
 angle, 202, 222
 gluonic, 220–222
 operator, 208, 210, 211, 216
 QCD, 221
 singlet-octet, 220, 221
 triplet-singlet, 222
 MIYAHARA, S., 483
 MIYANAGA, T., 346
 mode
 collective, 543, 544
 excitation, 724
 ghost, 603
 Goldstone, 746
 Kaluza-Klein, 185–187
 zero, 60, 66
 model
 A dynamics, 446
 $t - J$, 481, 482
 big bang, 755
 Boulatov, 708
 boundary dimension, 436
 Chaplygin gas, 558, 561
 confinement, 226
 cosmological, 698, 704, 722, 730,
 733, 738
 curvature-saturated, 697, 699
 Debye, 343
 dual, 494, 499, 503, 504, 592, 597,
 614
 Edward, 646
 fluid membrane, 670
 free-fermion, 294
 Friedmann, 697, 699, 701, 704, 730
 gamma-ray burst, 786
 Ginzburg-Landau, x, 485–487, 494,
 497, 586
 Gross-Neveu, 507–509, 511, 513, 514
 Higgs, 275, 276, 596, 597
 Hopfield, 271–273, 276, 277
 Hubbard, 481, 517, 521, 538
 inflationary, 750, 751
 Ising, 33, 276, 405, 415, 417, 421,
 427, 445, 446, 452, 459, 464,
 465, 498, 590, 591, 596, 707,
 713, 716
 Kaluza-Klein, 722
 Kazakov, 707, 708
 lattice, 426, 565, 708
 London, 498
 Nambu-Goto, 561, 626, 636
 Nambu-Jona-Lasinio, 508
 neural network, 271
 polymer, 580, 584, 585
 quantum gravity, 102, 708, 717
 quark, 175, 207, 209, 212, 214, 215,
 219, 223
 Randall-Sundrum, 188, 189
 random
 anisotropy, 458
 walk, 672
 scalar field, 225
 self-avoiding walk, 427
 self-dual, 498
 sigma, 43, 426, 430, 439, 510
 spin, 707
 spring and bead, 645

-
- standard, 173, 174, 414, 428, 434, 721, 756
 - string, 630, 636, 637, 641, 643
 - superluminary, 749–753, 755
 - Veneziano, 207
 - model A dynamics, 446
 - moduli stabilization, 722
 - MOFFAT, J.W., 741, 756, 757
 - MOHAPATRA, R.N., 195, 205
 - molecular dynamics, 24
 - MOLER, K.A., 548
 - MØLLER, C., 690, 695
 - momentum density, 562
 - MOMONO, N., 548
 - MONHOUX, P., 531
 - MONIEN, H., 523, 530
 - MÖNNIGMANN, M., 411
 - monomer density, 580, 583, 586
 - monopole
 - action, 614
 - condensation, 613
 - Coulomb gas, 614
 - density, 613, 615, 616, 618, 621, 622, 749
 - gas, 613, 615, 617–619, 621, 622
 - magnetic, 71
 - potential, 617, 621, 622
 - Monte Carlo, 29, 30
 - calculation, 341, 502
 - data, 31, 32
 - dynamics, 445, 446, 452
 - estimate, 38
 - evaluation, viii, 30, 31
 - method, 15, 18, 19, 30, 31, 33, 35, 39, 40
 - path integral, 29
 - process, 37
 - sampling, 32–36
 - simulation, xi, 34, 40, 276, 345, 414, 427, 454, 494
 - update algorithm, 39
 - MONTHOUX, P., 548
 - MONTVAY, I., 432
 - MOORE, H., 796
 - MOORE, M.A., 467
 - MOPE, 648
 - MOREL, P., 546
 - MORETTE, C., 26
 - MORGAN, D.C., 548
 - MORI, Y., 549
 - MORITA, T., 34, 40
 - MORITA, Y., 545, 548
 - MORIYA, T., 548
 - MOROZUMI, T., 194
 - MORRISON, P.J., 562
 - Morse potential, 34, 71, 733, 737
 - MOSHE, M., 432
 - MOSTAPANENKO, V.M., 633
 - MOTOKAWA, M., 483
 - MOUNT, K.E., 13
 - MOZDOR, E.V., 298
 - MUÑOZ-SUDUPE, A., 422
 - MUKAMEL, D., 467
 - MUKHANOV, V., 706
 - MUKHANOV, V.F., 768
 - MUKHANOV, V.M., 769
 - MÜLLER, A., 224
 - MÜLLER, B., 633
 - MÜLLER, G., 470, 482
 - MÜLLER, K.A., 517, 529, 537, 547
 - MÜLLER, S., 346
 - MÜLLER-HARTMANN, E., 483
 - MÜLLER-HOISSEN, F., 128, 695
 - multicanonical sampling, 38
 - multigrid method, 38
 - multiloop calculations, 433
 - multivalued
 - action, 618
 - mappings, 129
 - MÜNSTER, G., 246, 412, 432
 - MURRAY, D.B., 364, 374
 - MUTA, T., 185, 195
 - MÜTTER, K.-H., 469, 482, 483
 - NAGAI, K., 550
 - NAGAI, S., 550
 - NAIR, V.P., 562
 - NAKANO, H., 194, 195

- NAKANO, T., 548
 Nambu notation, 524
 NAMBU, Y., 514, 759, 769
 Nambu-Goldstone
 bosons, 198
 mode, 534
 condensation, 601, 603
 Nambu-Goto
 action, 619, 622, 627
 model, 561, 626, 636
 string, 625–627, 635–639, 641–643
 theory, 561
 Nambu-Jona-Lasinio model, 508
 NANDI, S., 195
 NARANJO, E., 660
 NARAYANA SWAMY, P., 198, 200
 NAYFEH, A.H., 769
 NECHAEV, S., 587
 NEDELKO, S.N., 234
 NE'EMAN, Y., viii, 173, 182–184
 negative curvature spaces, 724
 NELSON, B., 156
 NELSON, D.R., 505, 660, 661
 NEMETSCHEK, R., 530
 NESTERENKO, V.V., 625, 633–635, 643, 644
 NESTOR, J.M., 682
 NEU, J., 245, 364, 385, 388, 396, 412, 442
 NEUMARK, D.M., 26
 NEUMEIER, J.J., 548
 neural network, 271, 274
 neuron, 271
 neutrino
 anomaly, 201
 Dirac, 202
 energy, 752
 flux, 202
 Majorana, 201, 202
 massive, 202
 oscillation, 201–203
 neutron
 scattering, 519, 539
 star, 772, 775, 779, 783, 786, 789, 794
 NEVEU, A., 13, 514
 NEWMAN, C.M., 431
 NEWMAN, D.J., 184
 NEWMAN, K.E., 423
 NEYZI, F., 563
 NG, K.W., 530
 NGUYEN, A.K., 506, 531, 597, 599
 NICHIPOR, I.V., 335
 NICKEL, B.G., 257, 364, 374, 396, 409–412, 422, 423, 442, 443, 467
 NICOLAI, H., 563
 NICOLE, D.A., 13
 NIEDERMAIER, M., 423
 NIEDERMAYER, F., 423
 NIEGAWA, A., 206
 NIELSEN, J.L., 720
 NIETO, M.M., 81
 NISHIZAKI, S., 549, 550
 NISSEN, J.A., 364, 375, 385, 422
 NITSCH, J., 695
 no-boundary condition, 730
 Noether
 identity, 688, 691
 theorem, 688
 NOGUEIRA, F.S., 495, 497, 506
 NOJIRI, H., 483
 non-Abelian gauge fields, 57
 non-analytic potential, 110
 non-commutative geometry, 179
 non-equilibrium, 565, 566
 duality, 568
 worldline duality, 565
 non-holonomic
 equivalence principle, 133
 mapping, ix, xi
 transformation, vii
 non-local
 action, 619
 field theory, 645
 non-minimal coupling, 681, 682, 698
 non-perturbative methods, 327
 nonlinear structure, 759

- NOORDAM, L.D., 81
 NORBURY, J.W., 233
 NORDSIECK, A., 260, 269
 NORMAN, M.R., 547–549
 NOUNESIS, G., 411
 NOZAKI, K., 769
 NUTKU, Y., 563
- octonions, 80, 201
 ODA, M., 548
 ODINTSOV, S.D., 633
 OHTA, T., 346
 OKADA, N., 195
 OKIJIMA, Y., 548
 OKUBO, S., 224
 OLESEN, P., 633
 OLVER, P., 563
 one-particle irreducible
 n -point function, 242
 diagram, 236, 239–245, 247, 248
 four-point function, 235, 236,
 240–243
 Onsager
 exponent, 707, 708, 714–717
 universality class, 717, 718
 ONSAGER, L., 590, 598
 OONO, Y., 769
 operator
 composite, 435
 creation and annihilation, 144, 145,
 155, 330, 523, 628
 current, 500
 density, 30
 derivative, 491, 742
 dipole, 654
 evolution, 15, 16, 18, 19, 22, 77, 97,
 99, 101
 field, 147
 Heisenberg, 606
 Laplace-Beltrami, 133, 668, 738
 lowering, 73, 76, 77, 80
 method, 327, 332, 668
 mixing, 208, 210, 211, 216
 normal order, 652
 number, 330, 332
 ordering, 95, 97, 101, 103, 285, 737
 product expansion, 198, 645, 648,
 651
 projection, 487
 spin, 211, 213, 470
 Taylor, 264–266
 tilt, 208
 time ordered, 147
 translation, 472
 vertex, 647
 OPPENHEIMER, J.R., 772, 795
 order
 antiferromagnetic, 471, 523
 ferromagnetic, 458, 459, 465
 long-range, 272, 508, 523, 589–591
 parameter, 276, 447, 471, 486–488,
 491, 492, 509, 520, 525–527,
 536, 538, 573, 605, 607, 609,
 610
 short-range, 523
 ordering dynamics, 447–449, 452, 453,
 455
 ORLOV, E.V., 423
 Ornstein-Uhlenbeck process, 668
 ORSZAG, S.A., 115
 oscillator
 anharmonic, 109, 301, 302, 305,
 307–310, 315, 316, 322, 333,
 335, 347, 349, 352, 353, 362,
 363, 370–372, 391
 harmonic, 20, 23, 33, 35, 57, 58, 64,
 66, 67, 73, 74, 80, 107, 108,
 118, 120, 121, 124, 125, 127,
 212, 317, 322, 330, 339, 732,
 733, 761, 763
 OSHEROFF, D.D., 533, 545, 546
 OSHIKAWA, M., 476, 483
 OSTERWALDER, K., 720
 OVEREND, N., 505
 OZAKI, M., 549
- Padé
 method, 418, 429

-
- resummation, 302
 - Padé-Borel method, 434, 462, 463, 465
 - PAGET, K.M., 506
 - pair
 - creations, 70
 - distribution function, 33
 - potential, 341
 - pairing potential, 518
 - PAO, C.-H., 531, 548
 - PAPADOPOULOS, N.A., 183
 - PAPAPETROU, A., 688, 694
 - PAPASTAMATIOU, N.J., 611
 - parabolic coordinates, 95, 96, 101
 - parallel transport, 179, 180, 679, 709
 - paramagnetic phase, 273, 426
 - parameter
 - Binder, 714
 - deformation, 122, 127
 - disorder, 499
 - Hubble, 701
 - order, 276, 447, 471, 486–488, 491, 492, 509, 525, 536, 538, 573, 605, 607, 609, 610
 - renormalization, 405
 - variational, 287, 293, 294, 296, 331, 332, 347–351, 353, 354, 357, 392, 393
 - PARIDA, M.K., 205
 - PARIKH, M.K., 797
 - Parisi approach, 492
 - PARISI, G., 206, 364, 410, 412, 422, 467, 495
 - parity reflection, 105, 106
 - PARK, C.H., 548
 - PARK, S.H., 515
 - PARKER, C.S., 313
 - PARRINELLO, M., 41
 - PARRY, J., 769
 - particle
 - free, 57, 58, 69, 78, 120–127, 682
 - Higgs, 173
 - in a box, 34
 - partition function, 30, 31, 33, 35, 36, 43, 46, 47, 58–61, 67, 227, 228, 230, 238, 272, 274, 284–287, 289, 290, 294, 316, 321, 327, 328, 330, 331, 333, 338, 339, 427, 459, 567, 568, 585, 594–596, 614–617, 620, 707, 710
 - PASTA, J., 456
 - path integral, vii
 - coordinate independence, 45, 46, 54
 - cosmological, 729, 730, 733
 - effective classical potential, 337, 338
 - gauge, 95, 96, 602
 - gravitation, 729
 - group, 602
 - holomorphic coordinates, 77, 80
 - hydrogen atom, 73, 74, 95, 96, 102, 103, 117
 - Jacobi's principle, 95, 102
 - measure, 707, 708, 711
 - Monte Carlo, 29, 345
 - nonlinear sigma model, 43
 - operator ordering, 103
 - path integral, 11
 - perturbative definition, 54, 57–59, 67
 - phase-space, 54
 - polymer, 580, 581, 584
 - quantum cosmology, 729
 - semiclassical, 3–7, 11, 13, 98
 - separable Hamiltonian, 101
 - PATI, J.C., 205
 - Pati-Salam intermediate symmetry, 201, 203
 - PATRASCIOIU, A., 423, 432
 - PATTERSON, J.D., 466
 - Pauli
 - blocking effect, 204
 - principle, 201, 203, 204
 - PAULI, W., 103
 - PCAC, 209, 210
 - PEAK, D., 71
 - PEETERS, B., 55, 68
 - PEETERS, F.M., 298
 - PEIRANO, P.P., 720
 - PELCOVITS, R.A., 467

- PELISSETTO, A., 411–413, 422, 423, 466, 467
 PELLEGRINI, C., 695
 PELSTER, A., xii, 68, 128, 245, 257, 298, 315, 326, 347, 505, 529, 547, 739, 740
 PENROSE, R., 280, 775, 796
 PÉREZ, E., 157, 170
 PEREZ-LORENZANA, A., 195
 PEREZ-RENDON, A., 183
 perfect fluid, 704, 726
 perturbation
 expansion, 43–47, 54, 301–303, 365, 370–373, 388, 391, 396
 series, 316
 theory, 316, 328, 334, 337, 345
 variational, ix, x, 301, 302, 308, 311, 312
 perturbatively defined path integral, 43–45, 53, 54
 PESKIN, M.E., 147, 156, 495
 PESTER, M., 547
 PETERS, C.J., 411
 PETERSON, C., 633, 644
 PETSCHKE, I.B., 660
 PETTINI, G., 456
 PETTINI, M., 456
 PHAM, F., 115
 phase
 antiferromagnetic, 538
 boundary, 276
 coherence, 508, 517, 520, 529
 condensed, 15
 confinement, 276
 decoherence, x
 diagram, 517, 519, 520, 525, 527–529
 ferromagnetic, 272, 273, 426, 457, 458
 fluctuation, 510, 511, 514, 527, 528
 fractal, 658
 geometrical, 92
 Higgs, 276
 influence, 290, 292, 295
 ordering dynamics, 445–447, 452, 454, 455
 paramagnetic, 273, 426
 polymer, 578
 spin glass, 273, 279, 458
 stiffness, 511, 512
 transformation, 92
 transition, ix, x, 236, 457–460, 465, 518, 520, 601, 610, 611, 741–743, 745, 748, 751–756
 cosmological, 593
 vacuum, 193
 ϕ^4 -theory, ix, x, 235, 236, 240, 243, 245, 247, 248, 256, 365, 373, 387, 388, 390, 391, 396, 425, 433
 PI, S.-Y., 562
 PIGORSCH, C., 128
 PINES, D., 523, 530, 548
 PINN, K., 422
 PINTERIC, M., 547
 pion dynamics, 223
 PIRNER, H.J., 660
 PIROZHENKO, I.G., 634, 643, 644
 PIS'MAK, YU.M., 443
 PISANTI, O., 206
 PISARSKI, R.D., 643
 PIS'MAK, YU.M., 443
 PITAEVSKII, L.P., 546, 644
 Planck scale, xiii, 186, 194, 698, 722, 723, 725–727, 729
 plasma-electromagnetic pulse, 788
 PLATZMAN, P.M., 298, 346
 PLEBAŃSKI, J., 695
 PLISCHKE, M., 466
 PLUNIAM, G., 633
 POILBLANC, D., 482
 point canonical transformation, 69, 734
 POIRIER, J.-P., 575
 Poisson formula, 89
 POKROVSKY, V.L., 546
 polaron
 action, 284
 Fröhlich, 283, 285
 ground-state energy, 284

- system, 286, 288
 POLLOCK, E.L., 37, 41
 POLTURAK, E., 599
 POLYAKOV, A.M., 622, 633, 673, 707, 719
 Polyakov-Kleinert action, xi, 619, 622
 polyatomic molecules, 15
 POLYCHRONAKOS, A.P., 562, 563
 polymer, 378, 460, 463, 578–586
 action, 585
 analog, 38
 artificial, 578
 branched, 672
 chain, 433
 entanglement, 577
 fluctuation, 664
 length, 580, 586
 linked, 577
 macromolecules, 578
 model, 580, 584, 585
 phase, 578
 physics, 577, 578
 property, 578
 solution, 237, 578
 synthetic, 578
 topological, 577, 580
 polymerized membrane, 645, 659, 660
 POMERANCHUK, I.YA., 197, 200, 431
 PONTECORVO, B., 202, 205
 POPOV, V.N., 175
 Pöschl-Teller potential, 71
 potential
 Alvarez, 629
 anharmonic oscillator, 316, 322, 349
 barrier, 4, 38, 583
 Casimir, 727
 chemical, 481, 482, 522, 566, 584
 complex, 111
 confinement, 294, 296
 Coulomb, 89, 511, 733
 double-well, 35, 38, 315
 effective, 23, 37, 187, 188, 191, 485–491, 494, 509–513, 613, 616, 618, 722, 724, 726, 727, 772, 773, 775, 776, 779, 781
 classical, 337, 342, 347
 electromagnetic, 71
 exponential, 71
 gauge, 161, 166
 Gauss, 555, 556
 harmonic, 19, 45, 69, 733
 Higgs, 178
 interatomic, 339
 interquark, 625–631, 633, 636
 inverse
 quadratic, 69
 square, 71
 Kepler, 71
 long-range, 596
 mean-field, 487
 monopole, 617, 621, 622
 Morse, 34, 71, 733, 737
 non-analytic, 110
 pair, 341
 pairing, 518
 Pöschl-Teller, 71
 quadratic, 5, 71
 quark, 630
 relativistic, 85
 slowly varying, 4
 square-well, 34, 111
 string, 626, 629, 631, 632, 636, 637, 641
 topological, 586
 two-body, 583, 584
 vector, 84, 87, 92, 119, 120, 498, 606
 velocity, 554, 561
 vibrational, 337
 POVODYREV, A.A., 411
 power-law
 behavior, 353, 392, 449, 452–454
 decay, 446, 449, 453
 inflation, 703
 regime, 405
 pre-formed Cooper-pairs, 517, 520, 527
 precocious temperature, 487
 precursor fluctuation, 507, 513

- PREECE, R.D., 797
 PREPARATA, G., 789, 797
 PRESTON, D., 432
 PRICE, R.H., 797
 primordial nucleosynthesis, 722
 principle
 action, 680
 equivalence, 133, 685–688, 693
 exclusion, 203
 gauge, 271, 580
 Jacobi, 95, 96
 least action, 95, 96, 99–102
 Mach, 691, 692
 minimal sensitivity, 323, 349
 non-holonomic mapping, xi
 non-local field theories, 645
 Pauli, 201, 203, 204
 relativity, 691, 692
 stationary action, 103
 symmetry, 103
 variational, 275, 283–286, 288,
 291–293, 295–297, 679, 681
 Schwinger, 146–148
 probability density, 34, 38
 problem
 back reaction, 768
 bipolaron stability, 284
 boundary value, 16
 cosmological, 590, 752
 constant, 750–752, 755
 Coulomb, 34, 117, 127
 eigenvalue, 113, 402
 energy, 685
 energy-momentum, 692
 flatness, 741, 742, 751
 horizon, 741, 742, 748, 751
 hydrogen atom, 117
 Kepler, 71, 73, 74, 77, 78
 many-body, 285, 286
 many-fermion, 285
 operator ordering, 103
 passive scalar, 659
 resummation, 434
 sign, 18
 space, 685–690, 693
 strong-coupling, 498, 505
 three-body, 227
 topological entanglement, 577, 581
 two-particle, 33
 unboundedness, 711
 process
 Chebyshev, 124, 125
 learning, 271, 279
 Markov, 137
 Monte Carlo, 37
 Ornstein-Uhlenbeck, 668
 stochastic, 129
 Wiener, 33
 product of distributions, 44–46, 48–51,
 53, 54
 propagator
 p -oscillator, 126
 q -free particle, 127
 q -objects, 126
 Coulomb, 117
 forward-backward, 15
 Goldstone, 248, 254
 Higgs, 248, 252, 253
 transmission, 11
 WKB, 6, 7
 property
 barrier, 9
 Markov, 134
 polymer, 578
 scaling, 61, 391, 427, 663, 672
 universal, 410, 457, 458
 PROZOROV, R., 530, 547
 pseudogap, 517, 519, 520, 525, 526,
 528, 529
 pseudoscalar, 220, 679, 682
 meson, 219, 220
 PUCHKARYOV, E., 547, 548
 PUGLIESE, A., 205
 PUJOL, P., 483
 pulsar, 772, 775
 pulsed harmonic oscillator, 125, 127
 PUNSLY, B., 781, 797
 PURWANTO, A., 195

- PYTTE, E., 467
- q -free particle, 122
- q -objects, 124
- QCD, 613, 619, 620, 622
 anomaly, 219
 vacuum, 225, 232
- QED
 triviality of, 197
- QIAO, S.N., 128
- quantum
 brain dynamics, 279
 chromodynamics, ix
 cosmology, 729–731, 737
 dot, 286
 electrodynamics, ix, 301
 field theory, 46
 fluctuation, 341, 514, 755, 768
 gravity, 707, 729
 model, 102, 708, 717
 interference, 15
 many-body systems, 285
 mechanics, 46
 spin systems, 469–471
 statistics, 31, 33, 173, 175, 176, 180,
 302, 327, 328, 347
 Universe, 181, 732
- quark
 constituent, viii
 current, viii
 model, 175, 207, 209, 212, 214, 215,
 219, 223
 parton distribution, 201
 potential, 630
 valence, 203
 wave function, 219
- quasi-particle, 535, 536, 539, 542, 591
- quaternions, 74, 80
- quenched disorder, x, 458, 459
- Quillen superconnection, 178
- QUILLEN, D., 179, 183
- RABUS, H., 346
- radial distribution function, 337
- radion, 725
- radius
 convergence, 347, 366–369, 391
 critical, 188, 194, 626, 630, 636
 Hubble, 753, 759
 Randall-Sundrum, 185, 186
- Radon-Nikodym derivative, 140
- RAHMAN, A., 41
- RAINER, M., 705
- RAJARAMAN, R., 13
- RAMAKRISHNAN, T.V., 530
- Raman scattering, 519, 539, 544
- RAMOND, P., 205
- RANDALL, L., 186, 189, 195
- Randall-Sundrum
 model, 188, 189
 radius, 185, 186
 warp factor, 186
- RANDEIRA, M., 548
- random
 anisotropy, 458, 464, 465
 model, 458
 chain, 580
 disorder, x
 paths, 663, 664
 surfaces, 663, 664, 669, 672
 walk, 585
 walk model, 672
- RANGFANI, A., 128
- RAVNDAL, F., 96, 102, 103, 195
- Rayleigh-Ritz principle, 284
- real-time path, 15
- recursion relation
 Bender-Wu, 322
 graphical, 235, 236, 238, 239, 247,
 248, 250, 251, 256
 Kosterlitz-Thouless, 592
- REDIVO-ZAGLIA, M., 385
- REDNER, S., 506
- reduced density matrix, 17, 20
- REES, M., 795
- REEVE, J., 412
- reflection
 coefficient, 12

- probability, 11
- Regge
 - action, 711
 - approach, 709
 - behavior, 225, 229, 231
 - calculus, 707, 708, 718
 - residue, viii
 - spectrum, 225, 232
 - trajectories, 225, 232
- REGGE, T., 183, 719, 775, 795
- regime
 - adiabatic, 566
 - asymptotic, 428
 - power-law, 405
 - scaling, 450
- REGNAULT, L.P., 530
- regularization, 53, 54, 59, 135
 - analytic, 353, 638
 - dimensional, 45, 46, 48, 54, 57, 387, 388, 406, 409, 435, 438, 487, 647, 650, 654
 - discretization, 664
 - zeta function, 629
- REHR, J.J., 346
- Reissner-Nordström geometry, 781, 782
- relation
 - Johnson-Treiman, 212
 - scaling, 668, 717
 - uncertainty, 98
- relativistic
 - astrophysics, 771, 772, 778
 - magnetohydrodynamics, 789
- relativity principle, 691, 692
- REMBLIENSKI, J., 170
- RENNER, C., 530, 548, 549
- renormalization, 197, 229, 232, 235, 236, 261, 263, 265, 269, 353, 389, 404, 409, 503, 628, 647, 649, 653, 654, 767
 - charge, 503
 - condition, 404–406
 - constant, 243, 353, 354, 373, 379, 380, 388, 390, 392, 437, 653, 656
 - effect, 176
 - field, 492
 - five-loop, 240
 - fluctuation, 768
 - group, 197, 365, 373, 387–389, 393, 395, 402, 414, 425, 457–460, 462–465, 467, 485, 486, 590, 593, 597, 598, 759–761, 764, 765, 768
 - density matrix, 470, 478
 - equation, 433, 434, 437, 465, 759, 760, 762, 763, 765–768
 - function, 354, 405, 406, 410, 413, 433–439, 442, 460, 656
 - transformation, 402, 460, 462, 465, 763, 765
- method, 434
- parameter, 405
- scale, 656
- scheme, 404, 406, 460, 463
- triangle anomaly, 198
- renormalized four-point coupling, 425, 426
- reparametrization invariance, 52, 54, 57, 95, 96, 100, 102, 665, 669, 707, 730, 747
- replica
 - limit, 464
 - method, 273, 459, 585
- REPPY, J.D., 599
- representation
 - Mellin, 259, 262, 263
 - membrane, 645
 - Schwinger proper-time, 567, 570
 - Schwinger-Bogoliubov, 261
- RESHETIKIN, N.Y., 482
- resummation, x, 384, 387, 388, 390, 395, 434, 459, 462, 464
- retarded interaction, 288
- REUTER, S., 706
- reweighting method, 713
- RHOADES, C., 796
- RIBAUT, M., 547
- Ricci

- rotation coefficients, 686, 690
- scalar, 690, 738
- tensor, 680, 700
- RICCIARDI, G., 206
- RICE, T.M., 482, 549
- RICHARDSON, B., 533
- RICHARDSON, R.C., 545, 546
- RIECK, C.T., 530
- RIECKERS, A., 183
- RIEDEL, E.K., 423
- RIEDLER, J., 719, 720
- Riemann manifold, 139, 149, 150
- rigid vertex graph, 158
- RITTENBERG, V., 183
- RIVASSEAU, V., 431
- RIVERS, R.J., 565, 575
- Robertson-Walker
 - cosmology, 70
 - theorem, 747
- ROCEK, M., 719
- ROLF, J., 720
- ROMEI, O., 611
- ROMEO, A., 633
- ROSA, L., 205, 206
- ROSEN, N., 688, 694
- ROSENBLUTH, A.W., 40
- ROSENBLUTH, M.N., 40
- ROSENSTEIN, B., 494, 515
- ROSKIES, R., 432
- ROSSI, P., 411, 422
- rotation Wigner, 212, 214
- ROULIN, M., 505
- ROWE, D.E., 40
- RUDNICK, J., 467
- RUEGG, H., 205
- RUELLE, D., 298
- RUFFINI, R.J., 769, 775–777, 780, 781, 784–786, 789, 790, 795–798
- RUFFO, S., 456
- RUIZ-LORENZO, J.J., 422
- Runge-Lenz vector, 74
- RUSS, H., 769
- RUUTU, V.M.H., 599
- RUVALDS, J., 530
- RYZHIK, I.M., 633, 740
- SA-YAKANIT, V., 313
- saddle-point approximation, 306, 369, 571, 618, 626
- SAFINYA, C.R., 660
- SAHLIN, H.L., 34, 40
- SAITO, G., 547
- SAKAKIBARA, K., 279
- SALAM, A., 205
- Salam-Strathdee superspace, 177
- SALAMON, M.B., 505
- SALANNON, M.H., 549
- SALMONSON, J.D., 797
- SALOMONSON, P., 55
- SALOPEK, D.S., 769
- sampling
 - importance, 32, 37
 - Metropolis, 34
 - multicanonical, 38
 - stochastic, 19, 33
 - umbrella, 38
- SANTORELLI, P., 205, 206
- SARACCO, P., 233
- SASA, S., 769
- SASAKI, K., 206
- SASAKI, M., 768
- satellite experiment, 365, 374
- SATO, M., 550
- SAUER, T., 29, 41, 314
- SAULS, A.J., 549
- SAVAGE, V.M., 115
- SAVOY, C.A., viii, 201, 204, 207, 216, 217
- SAWADA, O., 205
- SAYAKANIT, V., 128
- SCADRON, M.D., 224
- SCALAPINO, D.J., 531, 548
- scalar
 - curvature, 133
 - density, 690
 - field, 677–682, 710
 - model, 225
- scale horizon, 748, 760, 765

- scaling
 approach, 348, 355, 365, 366, 381, 384, 391, 393, 394
 argument, 581
 behavior, 415, 433, 448, 449, 455, 497, 499, 574, 663
 collapse, 451–453
 correction, 401–403, 405, 409, 410, 415, 429, 448, 449, 452
 dimension, 393, 435, 436, 449
 dynamic, 445, 446, 448
 factor, 122
 field, 416
 finite-size, 714, 715
 function, 414, 446, 451, 452, 455
 hypothesis, 448, 449, 593
 Kleinert, 499, 503, 504
 KPZ, 717
 law, 436
 limit, 373, 498, 663–665, 667, 668, 671, 672
 parameter, 264
 photon mass, 503
 property, 61, 391, 427, 663, 672
 regime, 450
 relation, 668, 717
 transformation, 62, 63
 variable, 714
SCARPETTA, G., 705
scattering
 amplitude, 260, 261
 deep inelastic, 201, 210
 electron-electron, 528
 multiple, 3
 neutrino, 202
 neutron, 519, 539
 pion-quark, 216
 Raman, 519, 539, 544
 rate, 542
scenario
 BCS, 508, 512, 513
 inflationary, 750, 752, 755
 large extra dimension, 186
 phase transition, 500, 502
 SCHÄFER, H., 547
 SCHÄFER, L., 432
 SCHAKEL, A., 574
 SCHAKEL, A.M.J., 485, 495, 506, 565, 575, 599, 622
 SCHALM, K., 68
 SCHAUB, B., 456
 SCHECK, F., 180, 183
 SCHERER, W., 183
 SCHERK, J., 678, 682
 SCHIFF, E.A., 599
 SCHIFF, L.I., 298
 SCHLOMS, R., 410
 SCHLUETER, J., 547
 SCHMIDT, C.F., 660
 SCHMIDT, H.-J., xii, 137, 694, 697, 705, 706, 739
 SCHMIDT, R., xii
 SCHMITT, A., 482, 483
 SCHMITTMANN, B., 456
 SCHOLLWÖCK, U., 482
 SCHOTTE, K.D., 483
 SCHRIEFFER, J.R., 508, 514, 523, 530, 531, 546
 SCHRÖDINGER, E., 73, 80, 314, 694
 SCHROEDER, D.V., 147, 156
 SCHROER, B., 200
 SCHULMAN, L.S., 16, 26, 41, 128, 138, 298
 SCHULTE, C., 183
 SCHULTE-FROHLINDE, V., x, 245, 246, 298, 347, 364, 374, 385, 396, 397, 412, 422, 442, 443, 456, 466, 795
 SCHULTZ, T., 472, 483
 SCHULZ, H.J., 482
 SCHULZ, M., 456
 SCHWARZ, J.H., 678, 682
 Schwarzschild black hole, 781
 SCHWEITZER, D., 547
 Schwinger
 proper-time representation, 567, 570
 variational principle, 146–148
 SCHWINGER, J., 797

- Schwinger-Bogoliubov representation, 261
 Schwinger-Dyson equation, 147
 SCIARRINO, A., 204, 205
 screening, 197, 592, 615
 SEGAL, G., 183
 SEGRÉ, E., 456
 SEILER, E., 423
 self-avoiding
 membrane, 645, 657, 673
 walk model, 427
 self-dual model, 498
 self-energy, 235, 236, 240–244
 SELKE, W., 506
 SELLMYER, J., 466
 semiclassical
 approximation, 4, 5, 11, 16, 18, 22, 25, 98, 136, 586, 732, 737
 expansion, 130, 136
 path integral, 3–7, 11, 13, 98
 SENGERS, J.V., 411
 SEPULVEDA, M.A., 26
 series
 asymptotic, 301, 302, 328, 371, 390, 457, 459, 460
 strong-coupling, 352
 weak-coupling, 302, 308, 311
 SESSLER, A.M., 546
 SEUFERLING, P., 432
 SHAFI, Q., 205
 SHAH, S., 587
 SHALAEV, B.N., 467
 SHANKAR, R., 538, 547
 Shanks transformation, 383
 SHAO, J., 22, 23, 27
 SHARP, D.H., 432
 SHARPE, B., 411
 SHASTRY, S., 483
 SHAUB, B., 467
 SHEKA, V.I., 298
 SHELLARD, E.P.S., 598
 SHEN, Z.X., 531, 539, 548
 SHENKER, S., 280
 SHI, J., 505
 SHIFMAN, M.A., 224
 SHIN, W.Y., 549
 SHIRAFUJI, T., 690, 694, 695
 SHIRAISHI, J., 545, 548, 549
 SHIRAMURA, W., 483
 SHIRKOV, D.V., 396, 397
 SHLOSMAN, S.B., 431
 Shockley-Hagedorn-Feynman transition, 565, 568, 574
 SHORE, G.M., 224
 short-range
 interaction, 595, 613
 order, 523
 SHPOT, M., 466, 661
 SHTRIKMAN, H., 505
 SHTRIKMAN, S., 506
 SHURYAK, E.V., 35, 41
 shuttle experiment, 377
 SI, Q., 531
 sigma
 fluctuation, 509
 model, 43, 426, 430, 439, 510
 sign problem, 18
 SIGRIST, M., 547, 549, 550
 ŠIJAČKI, D., 183
 SIMMONS JR., L.M., 81
 SIMON, B., 506
 simplex, 709
 SINGH, R.P., 483
 singular behavior, 571, 573
 singularity, 705
 Borel-transform, 418
 cosmological, 70, 698
 topological, 601, 602, 605–607, 610
 ultraviolet, 567
 van Hove, 526
 SKENDERIS, K., 55, 68
 SLANSKY, R., 205
 Slater sum, 33
 SLICHTER, C.P., 530
 SLOAN, T., 206
 slow motion dynamics, 760, 768
 SMADJA, G., 206
 SMIRNOV, A.J., 205

- SMIRNOV, V.A., 397
SMOLENSKI, P., 432
SMOLIN, L., 163, 170, 181, 183
SMOOT, G.F., 757
SNYDER, H., 772, 795
SODA, T., 546
SOFFEL, M.H., 769
SOFFER, J., 203, 206
soft mode, 469, 471–475, 478, 481
SOKAL, A.D., 417, 422, 431
SOKOLOV, A.I., 364, 423, 442, 467
solid membrane, 645
soliton, 544
SOLUNIN, S.A., 234
SOMMER, R., 643
Sommerfeld constant, 542
SOMPOLINSKY, H., 279
SONG, X., 26
SORACE, E., 204, 216
SORBA, P., 217
SORNBORGER, A., 706
SOURIAU, J.M., 183
space
 Fock, 124, 144, 146, 147, 604
 problem, 685–690, 693
space-time, 45, 46, 48
 de Sitter, 701–703
 Friedmann-Robertson-Walker, 759
 transformation, 83, 92, 729, 733
specific heat, 33, 347, 348, 353,
 357–359, 364, 374, 377, 402, 408,
 409, 414, 433, 470, 539, 541, 542,
 544, 714
SPECTOR, M.S., 660
speed of light, 741, 748, 751, 755
SPENCER, T., 506
spherical harmonics, 78, 80, 86, 88, 541
SPICER, W.E., 548
spin
 density wave, 523, 525, 526, 536, 545
 fluctuation, 517, 518, 520–523, 527,
 538
 glass, 273
 phase, 273, 279, 458
 ladder system, 469, 471, 477
 model, 707
 operator, 211, 213, 470
 Peierls transition, 478
 statistic theorem, 175
spontaneous
 broken symmetry, 106, 220, 602,
 605, 607
 magnetization, 409
SPRIK, M., 38, 41
spring and bead model, 645
SQUIRES, E.J., 233
SRIVASTAVA, A.M., 599
stabilization of internal spaces, 724
STACHEL, J., 739
STADLOBER, B., 530, 548
staging algorithm, 33, 38
standard model, 173, 174, 414, 428,
 434, 721, 756
STANLEY, H.E., 506
STANYUKOVICH, K., 563
star, 778
 binary, 778
 equilibrium, 772
 evolution, 772
 neutron, 772, 775, 779, 783, 786,
 789, 794
Stark effect, 301
STAROBINSKY, A., 728
state
 Gibbs, 142
 vacuum, 145, 197–199, 608, 727
state-vector collapse, 174
static structure factor, 469, 473, 474
stationary phase approximation, 6, 10,
 16–19, 23, 24
statistical mechanics, 445, 446
statistics
 Bose and Fermi, 31
STECH, B., 205
steepest-descent approximation, 735,
 737
STEGLICH, F., 547
STEINER, F., 128, 740

- STEINHARDT, P.J., 757
 STEPHENSON, G., 184
 STERNBERG, S., 183
 STEVENSON, P.M., 326, 364, 431
 STEWART, J.M., 769
 STIEFEL, E., 81, 103, 117, 118, 128
 stiffness, 512, 513, 626
 phase, 511, 512
 string, 619
 stochastic
 dynamics, 445, 448, 452–455
 process, 129
 sampling, 19, 33
 Stokes theorem, 616
 STOKKE, B.T., 660
 STONE, M., 495, 574
 STORA, R., 720
 STØRMER, C., 772, 795
 STORNAIOLO, C., 740
 Stratonovich equation, 131
 STRICKER, D.A., 364, 375, 385, 422
 string, x , 561, 677–679
 action, 619
 Brownian, 565, 568, 597
 density, 595
 dynamics, 625, 636
 effective action, 619, 621, 622
 free energy, 595, 637
 length, 568, 571–573, 594–596, 636, 641, 642
 model, 630, 636, 637, 641, 643
 potential, 626, 629, 631, 632, 636, 637, 641
 representation, 613, 614, 616, 617, 620, 621
 stiffness, 619
 theory, 677–679, 681
 vacuum, 639
 string-dilaton gravity, 698
 STRONG, I.B., 783, 784
 strong-coupling
 approach, 387, 388, 395, 431
 approximation, 394
 behavior, 355, 373
 coefficients, 348, 349, 352, 366, 369, 371, 372
 expansion, 301, 303, 305–312, 352, 355, 366–369, 371, 391, 425, 427, 430
 formula, 392
 limit, x , 302, 304, 307, 315, 322, 347–349, 352, 354, 355, 357, 358, 373, 379–384, 390–394, 505
 problem, 498, 505
 series, 352
 superconductor, 507, 508, 514, 518, 523
 theory, 377, 384, 387, 391, 395
 STRÖSSER, H., 411
 STRUNZ, W., 547
 STUART, C., 280
 SU, D.C., 128
 SUDAKOV, V.V., 431
 SUDBØ, A., 506, 531, 597, 599
 sum rule, 210, 216, 525
 Adler-Weisberger, 215
 Gottfried, 203
 SUMINO, Y., 205
 SUN, X., 23, 26
 SUNDRUM, R., 186, 189, 195
 SUNG, S.H., 466
 supercharge, 559, 561
 superconductor, ix, x , 497, 498, 501, 505
 d -wave, 533, 538–542
 BCS, 508, 509
 high- T_c , 517–520, 536, 538, 545
 strong-coupling, 507, 508, 514, 518, 523
 triplet, 544, 545
 unconventional, 533, 536, 538–541, 543, 545
 superconductor-normal phase
 transition, 485, 486
 superconnection, 179
 superdeterminant, 148
 superficial divergence, 266, 649

-
- superfluid, 237
 - density, 525, 528, 529, 541, 543
 - helium, ix, 374, 377–379, 381, 383, 384, 433
 - supergravity, 175
 - superluminary model, 749–753, 755
 - supermembrane, 561
 - superspace Salam-Strathdee, 177
 - superstring theory, xiii, 175, 721
 - supersymmetry, 175, 177, 553, 557, 560
 - charge, 560
 - surface, 337, 338, 342–345
 - area, 777, 778, 780
 - melting, 344, 345
 - minimal, 150
 - susceptibility, 402, 409, 426, 429, 433, 524, 525, 540, 666, 670, 714, 715, 717
 - SUTHERLAND, B., 483
 - SUURA, H., 270
 - SUZUKI, H., 314
 - SUZUKI, M., 41
 - SUZUKI, T., 346
 - SVAITER, N.F., 495, 506
 - SVIDZINSKII, A.A., 549
 - SVOBODA, K., 660
 - SWANSON, D.R., 364, 375, 385, 422
 - SWENDSEN, R.H., 720
 - SWIFT, A.R., 13
 - SYKES, M.F., 431
 - SYMANZIK, K., 200, 409, 412, 423, 431, 634
 - symmetry
 - breaking, 208, 210, 215
 - electroweak, 194
 - isospin, 223
 - spontaneous, 220, 741, 742, 744, 752, 755
 - isospin, 219, 221–223
 - principle, 103
 - symplectic manifold, 139, 149, 150
 - synapse, 271, 272, 274, 277
 - system
 - Coulomb, 89, 307, 308, 310–312
 - polaron, 288
 - ρ - ω , 219
 - SZEKERES, P., 767, 769
 - TABUNAGA, Y., 483
 - TACHI, M., 611
 - tachyons, 625, 626
 - TAILLEFER, L., 549
 - TAKAHASHI, M., 41
 - TAKAHASHI, T., 548
 - TAKAHASHI, Y., 280, 548, 611
 - TAKASAKA, A., 41
 - TAKATA, Y., 346
 - TAKEUCHI, T., 548
 - TAKIGAWA, M., 547
 - TALLON, J.L., 530
 - TANABASHI, M., 195
 - TANAKA, T., 660
 - TANATAR, M.A., 545, 550
 - TANCREDI, N., 206
 - TANG, A., 233
 - TANG, W.K., 234
 - tangent bundle, 153
 - TARAPHER, A., 530
 - TARASOV, O.V., 396, 443
 - target space, 45, 54
 - TASAKI, H., 431
 - TATE, J., 155
 - Taylor operator, 264–266
 - teleparallelism, 685, 690–692
 - TELLER, A.H., 40
 - TELLER, E., 40
 - temperature
 - effect, 601, 608, 610
 - zero, 24, 295, 447, 509, 510, 514, 637, 638, 641, 642, 743, 744
 - TEMPERE, J., 298
 - TEMPLETON, S., 412, 562
 - TEN WOLDE, A., 81
 - TER MARTIROSIAN, K.A., 431
 - TERASAKI, I., 483
 - TERENTYEV, M.V., 71
 - TEŠANOVIĆ, Z., 495, 506
 - TESSMANN, J., 495
 - tethered membrane, 645, 659

- TEUKOLSKY, S.A., 775, 796
 TeV scale, 723, 725, 726
 TEWORDT, L., 531, 550
 THEIS, W., ix
 theorem
 area, 780
 boson transformation, 604, 605
 Cauchy, 263
 central limit, 663, 664
 Coleman, 509
 Darboux, 154
 Federbush-Johnson, 199
 Gauss-Bonnet, 709
 Goldstone, 609
 Hodge decomposition, 618
 Hohenberg, 519
 Lieb-Schulz-Mattis, 469, 471, 474
 low-energy, 215
 Mermin-Wagner, 590
 Noether, 688
 Robertson-Walker, 747
 spin statistic, 175
 Stokes, 616
 virial, 35, 39
 Wick, 61, 63, 318
 Wigner-Eckart, 214
 theory
 BCS, 469, 518, 533, 538, 539, 545
 D-brane, xiii
 deterministic, xiii
 dual, 566, 568
 Eliashberg, 523
 gauge, ix, 35, 96, 174–176, 178, 179, 212, 216, 271, 274, 276, 580, 602, 607, 678, 698
 Ginzburg-Landau, 520
 Kosterlitz-Thouless, 528
 Landau, 545
 lattice, 426
 M-, xiii, 175, 721
 mean-field, 271, 275, 277, 278, 458, 507, 527, 538, 545
 measure, 162
 QCD, 219
 strong-coupling, 377, 384, 387, 391, 395
 superstring, xiii, 175, 721
 variational perturbation, ix, x, 315, 322, 347–349, 352, 353, 355, 358, 366, 370–372, 377
 WKB, 4, 6, 7, 111, 760
 thermal
 average, 338, 339
 conductivity, 540–543
 decomposition, 508
 excitation, 512, 513
 expansion, 342, 345
 fluctuation, 236, 497–499, 507, 510, 579, 583
 noise, 569
 phase fluctuation, 514
 photon, 753
 vibration, 337, 345
 THIERRY-MIEG, J., 175, 178, 182, 183
 THOMAS, P.R., 495, 574
 THOMPSON, K., 26
 THOMS, S., 313, 397
 'T HOOFT, G., vii, xiii, 45, 55, 68, 224, 396, 622
 THORLEIFSSON, G., 661
 THORNE, K.S., 781, 797
 THOSS, M., 27
 THOULESS, D.J., 531, 591, 599
 three-body problem, 227
 TIBULLO, V., 206
 time
 evolution amplitude, 316, 317, 320
 reversal, 105, 106
 transformation, 117–119, 122, 734, 738
 time-slicing, 57
 TKACHOV, F.V., 385, 396, 397, 411, 412, 443
 TOGNETTI, V., x, 298, 313, 346
 TOLMAN, E., 772
 TOLMAN, R.C., 769
 TOLSTIK, A.L., 335

- TOMIĆ, S., 545, 547
 TOMITA, K., 769
 TOMSOVIC, S., 26
 topological
 charge, 606, 607
 defects, x, 533, 534, 543, 589, 590,
 593, 598, 601, 602, 606, 607,
 610
 dual, 140
 effect, 83, 92, 581
 entanglement, 577, 581
 excitation, 589, 590, 592–595, 598
 field theory, 577, 581
 invariant, 87, 161, 582–584, 709
 polymer, 577, 580
 potential, 586
 singularity, 601, 602, 605–607, 610
 soliton, 544
 TÖRÖK, T., 422
 torsion, viii, xi, 129, 174, 175, 677–681,
 692, 693, 733
 TOU, H., 549
 TOYODA, T., 81, 96, 102, 103
 trace anomaly, 198, 199
 TRAMONTANO, F., 206
 transformation, 604
 Alfaro-Fubini-Furlen-Jackiw, 120
 Borel, 418
 boson, 604–607, 609
 canonical, 96, 332, 734
 chiral, 508
 conformal, 723
 coordinate, 43, 46, 54, 118, 123, 127,
 689, 690, 692, 724
 duality, 498, 614
 Duru-Kleinert, 301, 733
 Fierz, 187
 gauge, 101, 274, 602, 604, 607, 608
 Heisenberg field, 604
 holonomic, 118
 Hurwitz, 73–75, 77, 80
 hyper-Borel, 367, 368
 Kustaanheimo-Stiefel, vii, 73, 74,
 80, 92, 95, 102, 117, 118
 Legendre, 98, 236, 243, 249, 251, 618
 Lorentz, 690–692, 742–744, 746
 Mellin, 261–265
 Melosh, 207, 212
 non-holonomic, vii, 118
 phase, 92
 point canonical, 69
 renormalization group, 402, 460,
 462, 465, 763, 765
 scaling, 62, 63
 Shanks, 383
 space-time, 83, 92, 729, 733
 time, 117–119, 122, 734, 738
 Wigner, 22, 23
 transient attractor, 705
 transition
 Kosterlitz-Thouless, 507, 510–514,
 591, 594, 596
 lambda, 374
 Shockley-Hagedorn-Feynman, 565,
 568, 574
 spin Peierls, 478
 transmission
 and reflection formula, 4
 barrier, 11
 coefficient, 12
 probability, 11
 TRAVESSET, A., 661, 673
 TREDER, H.-J., 685, 694, 695
 trial frequency, 303
 triangle anomaly, 198
 triangulation, 669, 670, 709, 710, 712,
 713, 717
 tricritical point, x, 493, 494, 498–502,
 505
 TRIMPER, S., 445, 456
 TRINH, D.T., 115
 TRIOLO, L., 204
 triplet superconductor, 544, 545
 triviality
 of ϕ^4 -theory, 425
 of continuum limit, 425
 of QED, 197
 Trotter formula, 36, 135

- TROYAN, S.I., 206
 TRUGENBERGER, C.A., 623, 633
 TRULLINGER, S.E., 546
 TRUMAN, A., 137
 TRUNOV, N.N., 633
 TSEYTLIN, A.A., 706
 TSIEN, H.-S., 563
 TSOU, S., 183
 TSUEI, C.C., 530, 538, 548
 TSUNETO, T., 534, 546
 TUCKEY, P.A., 719
 TURBINER, A., 115
 turbulence, 659
 TURNER, M.S., 756, 757
 TUROK, N., 599, 756
 TUZI, T., 205
 two-point function, 223, 259, 264, 269
 TYLER, A.W., 549
 type-II superconductors, 487, 488, 494

 UCHINOKURA, K., 483
 UEDA, K., 483, 547, 548
 UEMURA, Y.J., 530
 UGAWA, A., 346
 UHRIG, G., 483
 ULAM, S., 30, 40, 456
 ultraviolet
 behavior, 573, 574
 convergence, 264
 divergence, 46, 52–54
 fixed-point, 197, 198
 singularities, 567
 ULYANENKOV, A.P., 335
 umbrella sampling, 38
 UMEZAWA, H., 280, 611
 ÑNAL, N., 73, 81
 uncertainty relation, 98
 unconventional superconductors, 533,
 536, 538–541, 543, 545
 universal
 amplitude, 410, 427, 428
 ratio, 402
 behavior, 579
 exponent, 391, 402
 function, 403
 property, 410, 457, 458
 ratio, 403, 410
 universality class, 236, 247, 365, 378,
 387, 394, 421, 446, 466, 494, 497,
 504, 591, 597, 718
 Universe
 closed, 764
 de Sitter, 752
 early, 589, 729, 741, 742, 744, 753,
 755, 756
 expansion, 705, 742, 751, 753, 759,
 764
 Friedmann-Robertson-Walker, 731,
 747, 751, 752, 759, 760, 762,
 764
 ground-state, 742
 inflationary, 71, 768
 inhomogeneous, 759, 760
 open, 750
 quantum, 181, 732
 radiation-dominated, 71, 748
 superluminary, 748, 750, 755
 wave function, 731
 URBACH, J.S., 548

 vacuum
 current, 607
 decay, 71
 diagram, 235, 238–240, 256, 721
 energy, 752
 density, 242, 751
 expectation value, 187, 188, 192,
 198, 741
 field equation, 700
 gravitational, 755
 phase, 193
 phase transition, 743
 polarization, 727, 784–786, 788, 789
 QCD, 225, 232
 state, 145, 197–199, 608, 727
 string, 639
 symmetry, 742–744
 VAFA, C., 706

- VAIA, R., 346
 VAKS, V.G., 514
 valence quark, 203
 VAN DEN BOSSCHE, B., 195, 245, 257, 377, 385, 411
 VAN HARLINGEN, D.J., 548, 549
 van Hove singularity, 526
 VAN LINDEN VAN DEN HEUWELL, H.B., 81
 VAN NIEUWENHUIZEN, P., 55, 68
 VANNIMENUS, J., 466
 VAN VLECK, J.H., 26
 Van Vleck determinant, 6, 16, 17
 variational
 approach, 283
 equations, 332
 functional, 283
 parameter, 276, 287, 293, 294, 296, 331, 332, 347–351, 353, 354, 357, 392, 393
 perturbation theory, ix, x, 301, 302, 308, 311, 312, 315, 322, 347–349, 352, 353, 355, 358, 366, 370–372, 377, 379, 384
 principle, 275, 283–286, 288, 291–293, 295–297, 679, 681
 resummation scheme, 304, 308
 Schwinger principle, 146–148
 VARMA, C.M., 530
 VAROQUAUX, E.J., 546
 VARTCHENKO, A.N., 270
 VASIL'EV, A.N., 443
 Vassiliev invariants, 157–161, 169
 vector
 fields, Hamiltonian, 150
 Killing, 150, 152, 689
 meson, 219–222
 potential, 84, 87, 92, 119, 120, 498, 606
 Runge-Lenz, 74
 velocity potential, 554, 561
 VELTMAN, M., 45, 55, 68, 396
 Veneziano model, 207
 VENEZIANO, G., 224
 VERBIST, G., 298
 VERGA, A.D., 456
 VERNON, J.F.L., 27
 VEROSKY, J., 563
 VERRUCCHI, P., 346
 vertex operator, 647
 VIBET, C., 546
 vibration
 thermal, 337, 345
 vibrational dynamics, 24
 VICARI, E., 411–413, 422, 423, 466, 467
 VILENKIN, A., 598, 756
 VILGIS, T.A., 587
 VINEN, W.F., 575
 VINETTE, F., 314, 364
 VINOKUR, V.M., 505
 VINTI, S., 422
 virial
 estimator, 39
 theorem, 35, 39
 VIROSZTEK, A., 530, 545, 550
 VITIELLO, G., 601, 611
 VLADIKAS, A., 432
 VLADIMIROV, A.A., 396, 397, 443
 VOGEL, P., 170
 VOIGT, L., xii
 VOLKOFF, G.M., 795
 VOLLHARDT, D., 546
 VOLOVIK, G.E., 539, 546, 549
 volume form, 139, 146, 149, 154
 VON BORZESZKOWSKI, H.-H., 685, 694, 695
 VON DER HEYDE, P., 682
 VON KARMAN, T., 563
 VON LAUE, M., 694
 vortex
 Abrikosov, 543
 charge, 591
 current, 498
 decay, 574
 density, 498, 574
 excitation, 596
 free energy, 592

- lattice, 497, 499, 503, 539
 lines, 498, 499
 loops, 494, 498, 499
 motion, 557
 pair, 511, 513, 591, 596
 separation, 574
 solution, 591, 608
 state, 539, 540, 542, 543
 string, 596
 vorticity, 553–557
 VOSS, R., 206
 VUORIO, M., 546

 WADIA, S.R., 623
 WAKAYAMA, S., 81
 WALKER, E., 505, 549
 WALSH, T.F., 643
 WALTON, A.R., 26
 WANG, G.F., 548
 WANG, G.J., 550
 WANG, H., 26, 27
 WANG, H.H., 547
 WANG, M.L., 128
 WANG, P., 128
 WANG, Q., 109, 115
 WANG, Z., 530
 WANNIER, G.H., 505
 Ward identity, 503, 756
 warp factor, 186, 189, 191, 194
 warped product topology, 722
 WARR, B.J., 515
 WATANABE, S., 138
 WATSON, P.J.S., 467
 wave
 charge density, 545
 function
 Bloch, 522
 coherent state, 17
 ground-state, 34, 35, 315, 316,
 320–326, 469
 no-boundary, 731
 quark, 219
 renormalization constant, 437
 transformation, 739
 Universe, 731
 WKB, 4
 length, 753, 765
 Compton, 84, 226, 232
 de Broglie, 4
 fluctuation, 753
 number, Fermi, 292
 spin density, 523, 525, 526, 536, 545
 weak-coupling
 coefficients, 348, 349, 351, 352, 355,
 356, 360
 expansion, 62, 301, 304, 307, 308,
 348, 355, 357, 358, 360, 363,
 426
 series, 302, 308, 311
 WEBB, R.A., 546
 Wegner expansion, 378
 WEGNER, F.J., 411, 660
 weight
 Boltzmann, 29, 59, 238, 647
 diagram, 17, 160, 235, 236, 238–240,
 243, 245, 409
 Lie algebra, 163
 Weinberg
 leading power, 266, 268
 renormalization scheme, 404, 406
 WEINBERG, S., 200, 215–217, 411
 Weinberg-Salam Hamiltonian, 181
 WEINGARTEN, D., 432
 WEISS, U., 41
 WEISSBACH, F., 315
 WEISZ, P., 423, 431, 634
 WELLS, B.O., 548
 WENIGER, E.J., 314
 WENTZEL, G., 13
 WENZEL, L., 346
 WETTERICH, C., 201, 205
 Weyl
 gravity, 699
 identity, 22
 nonmetricity, 686
 ordering, 734, 737, 739
 WEYL, H., 174, 694

- Weyl-Cartan space problem, 686, 689, 690, 693
- WHEATLEY, J.C., 534, 546
- WHEELER, J.A., 772–780, 795, 796
- Wheeler-DeWitt equation, 705, 730, 731, 733, 734, 736
- white dwarf, 772
- WHITE JR, H.J., 411
- WHITE, S.R., 482, 531, 548
- Wick theorem, 61, 63, 318
- Wiener
- chaos, 144, 146
 - integral, 33
 - measure, 129, 139, 141, 144–146, 666
 - process, 33
- WIESE, K.J., 645, 660, 661
- WIESE, U.J., 482
- WIESSNER, R., 482, 483
- Wigner
- approximation, 23, 24
 - function, 23
 - rotation, 212, 214
 - transformation, 22, 23
- WIGNER, E.J., 27
- Wigner-Eckart theorem, 214
- Wigner-Kirkwood expansion, 36
- WILCZEK, F., 781, 782, 797
- WILLIAMS, J.M., 547
- WILLIAMS, R.M., 719, 720
- WILLIAMSON, P.R., 364, 375, 385, 422
- Wilson
- line, 164–168
 - loop, 162–165, 167, 168, 582, 613, 614, 616–618, 620, 621
- WILSON, J.R., 781, 790, 797
- WILSON, K.G., 197, 200, 396, 410, 411, 431, 634
- winding number, 87, 608, 610
- WISE, M.B., 195
- Witten functional integral, 157, 160–162, 166, 169
- WITTEN, E., 161, 163, 170, 509, 515, 587
- WKB
- higher-order, 112
 - penetration factor, 4, 737
 - propagator, 6, 7
 - theory, 4, 6, 7, 111, 760
 - wave function, 4
- Wolff algorithm, 712
- WOLFF, U., 720
- WÖLFLE, P., 546
- WOLLMAN, D.A., 538, 548
- WOLTIER, L., 562
- WON, H., 533, 547–550
- WONG, C.C., 432
- WRIGHT, D.C., 535, 547
- WU, T.T., 175, 308, 313, 322, 326
- WUNDERLIN, A., 128, 740
- WURM, A., 139
- WUYTACK, L., 385
- XUE, S.-S., 789, 797, 798
- YAGI, S., 346
- YAMAGUCHI, M., 194, 195
- YAMAGUCHI, Y.Y., 769
- YAMAJI, K., 547
- YAMAMOTO, E., 549
- YAMANAKA, M., 476, 483
- YAMAWAKI, K., 195
- YAMAYA, K., 548
- YANAGIDA, T., 205
- YANG, C.N., 55, 174, 175, 184, 475, 483
- YANG, C.P., 475, 483
- YANG, G., 549, 550
- YANG, M.S., 483
- YANG, S.M., 482
- Yang-Baxter equations, 470
- Yang-Mills gauge theory, 175, 178, 179
- YANOFF, B.D., 549
- YASUOKA, H., 547
- YAVORS'KII, T., 466, 467
- YENNIE, D.R., 270
- YING, H.P., 482
- YIP, S.K., 549

- YOKOYA, J., 548
YOKOYAMA, T., 337, 346
YONAMOTO, Y., 346
YOSHIOKA, K., 194, 195
YOUNG, A.P., 466
YOUNG, M.J., 411
YOURGRAU, W., 694
YURKE, B., 599
- ZAKHAROV, V.E., 546
ZAMOLODCHIKOV, A.B., 708, 719
ZANNI, M.T., 26
ZASADZINSKI, J.A., 660
Zeeman effect, 301
ZEH, H.D., 740
ZEL'DOVICH, YA.B., 757
ZELDOV, E., 505
ZERBINI, S., 633
ZERILLI, F.J., 775, 795
- zero
 frequency fluctuation, 59
 mass, 586
 mode, 60, 66
 point energy, 626
 temperature, 24, 295, 447, 509, 510,
 514, 637, 638, 641, 642, 743,
 744
- ZERWAS, P.M., 643
zeta function, 66, 67, 109, 438, 627,
 629, 638
- ZHA, Y., 531
ZHANG, J., 530
ZHANG, K., 548
ZHANG, R.J., 195
ZHENG, B., 445, 456
ZHIROV, O.V., 41
ZHUK, A., 706, 721, 727, 728, 733
ZICHICHI, A., 515
ZIMAN, T.A.L., 482
ZIMMERMANN, W., 269
ZINN, S., 422
ZINN-JUSTIN, J., 374, 385, 396, 397,
 407, 409, 411, 412, 422, 423, 431,
 442, 443, 466, 719
- ZIPPELIUS, A., 467
ZOBIN, D., 466
ZUBER, J.B., 270, 719
ZUCKERMANN, M.J., 466
ZUREK, W.H., 575
Zweig rule, 219–221, 223
ZWEIG, G., 224