Karpacz 91, XXVII Winter School of Theoretical Physics

NONLINEAR FIELDS: CLASSICAL RANDOM SEMICLASSICAL

Karpacz, Poland 18 February - 1 March 1991

Editors

P. Garbaczewski and Z. Popowicz

Institute of Theoretical Physics
University of Wrocław

World Scientific
Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.
P O Box 128, Farrer Road, Singapore 9128
USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661
UK office: 73 Lynton Mead, Totteridge, London N20 8DH

NONLINEAR FIELDS: CLASSICAL, RANDOM, SEMICLASSICAL

ISBN 981-02-0647-X

Printed in Singapore by JBW Printers & Binders Pte. Ltd.

Foreword

Since 1964 Winter School of Theoretical Physics in Karpacz became a permanent institution in the scientific landscape of Poland. They are run voluntarily by members and graduates of the Institute of Theoretical Physics of the University of Wrocław, with subjects varying from year to year according to the current profiles of research activities. Karpacz 91 was devoted to Nonlinear Fields: Classical, Random, Semiclassical. Our main intention is to combine several nonlinear science trends in one activity, with the focus on benefits coming from diversity of ideas and methodologies. Principal themes represented in the invited lectures were: complete integrability impact on string theory, bi-Hamiltonian structures, hierarchies, links with quantum groups versus quantization of non-integrable systems, fundamental links between randomness and quantization, intrinsic randomness of deterministic motions.

We would like to thank our colleagues, members of the Organizing Committee: A. Borowiec, R. Gielerak, M. Mozrzymas and W. Hann for their help, and to acknowledge the sponsorship of the following institutions. Without their sponsored, the XXVII Winter School of Theoretical Physics in Karpacz would not have achieved its success.

Ecole Polytechnique Féderale de Lausanne, Lausanne, Switzerland
International Centre of Theoretical Physics, Trieste, Italy
Unviersity of Wrocław, Wrocław, Poland
Foundation for the University of Wrocław, Wrocław, Poland
Polish Academy of Sciences, Warsaw, Poland

Our thanks go to all persons and institutions involved.

Piotr Garbaczewski
Ziemowit Popowicz

Winter Schools of Theoretical Physics in Karpacz

I	1964 (R. S. Ingarden) Particles, Fields and Superconductivity, unpublished
II	1965 (W. Ziętek, ed.) Symmetries in Particle and Condensed Matters Physics, ITP mimeographed lecture notes
III	1966 (Z. Galasiewicz, ed.) Statistical Physics of Condensed Matter, Acta Univ. Wratislaviensis No. 80
IV	1967 (J. Rzewuski, ed.) Functional Methods in Quantum Field Theory and Statistical Mechanics, Acta Univ. Wratislaviensis Nos. 88, 89, 90
V	1968 (J. Łopuszański, ed.) Axiomatic Approach to Quantum Field Theory and Many Body Problem, Acta Univ. Wratislaviensis, Nos. 98, 99, 113
VI	1969 (J. Przystawa, ed.) Group Theory and Statistical Physics of Condensed Matter, ITP mimeographed lecture notes
VII	1970 (Z. Galasiewicz, ed.) Liquid Helium and Many Body Problems, Acta Univ. Wratislaviensis No. 141
VIII	1971 (J. Lukierski, ed.) New Developments in Relativistic Quantum Field Theory and Applications, Acta Univ. Wratislaviensis No. 164
IX	1972 (K. Wojciechowski, ed.) Theory of Metals and the Many Body Problem, Acta Univ. Wratislaviensis No. 181
X	1973 (J. Lukierski, ed.) New Developments in Relativistic Quantum Field Theory, Acta Univ. Wratislaviensis No. 207
XI	1974 (J. Łopuszański, A. Pękalski, J. Przystawa, eds.) Magnetism in Metals, Plenum Press, NY

XII 1975 (W. Garczyński, ed.)
Functional and Probabilistic Methods in Quantum Field Theory, Acta Univ. Wratislaviensis No. 368

XIII 1976 (J. Lukierski, ed.)
Recent Development in Relativistic Quantum Field Theory and Its Application, Acta Univ. Wratislaviensis No. 389

XIV 1977 (Z. Galasiewicz, ed.)
Collective Effects in Condensed Media, Acta Univ. Wratislaviensis No. 436

XV 1978 (W. Karwowski, ed.)
Mathematical Aspects of Quantum Field Theory, Acta Univ. Wratislaviensis No. 519

XVI 1979 (A. Pękalski, J. Przystawa, eds.)
Modern Trends in the Theory of Condensed Matter, Lecture Notes in Physics, vol. 115, Springer, Berlin

XVII 1980 (L. Turko, A. Pękalski, eds.)
Developments in the Theory of Fundamental Interactions, Harwood Acad. Publ., Chur

XVIII 1981 (W. Garczyński, ed.)
Gauge Field Theories: Theoretical Studies and Computer Simulations, Harwood Acad. Publ., Chur

1982 -

XIX 1983 (B. Milewski, ed.)
Supersymmetry and Supergravity, World Scientific, Singapore

XX 1984 (A. Pękalski, J. Sznajd, eds.)
Static Critical Phenomena in Inhomogeneous Systems Lecture Notes in Physics, vol. 206, Springer, Berlin

XXI 1985 (L. Michel, J. Mozrzymas, A. Pękalski, eds.)
Spontaneous Symmetry Breadown and Related Subjects, World Scientific, Singapore

XXII 1986 (A. Jadczyk, ed.)
Fields and Geometry, World Scientific, Singapore

XXIII 1987 (T. Paszkiewicz, ed.)
Physics of Phonons, Lecture Notes in Physics, vol. 285, Springer, Berlin

XXIV 1988 (R. Gielerak, W. Karwowski, eds.)
Stochastic Methods in Mathematics and Physics, World Scientific, Singapore

XXV 1989 (Z. Haba, J. Sobczyk, eds.)
Functional Integration, Geometry and Strings, Birkhäuser Verlag, Basel

XXVI 1990 (Z. Galasiewicz, A. Pękalski, eds.)
Ordering Phenomena in Condensed Matter Physics, World Scientific, Singapore

XXVII 1991 (P. Garbaczewski, Z. Popowicz, eds.)
Nonlinear Fields: Classical, Random, Semiclassical, World Scientific, Singapore

Lectures Delivered in the Course of the XXVII-th Session

1. H. Arodź (Cracow),	Classical Rigid String
2. I. Bakas (College Park)	Bi-Hamiltonian Structures in Integrable Systems and Conformal Field Theories
3. I. Barashenkov (Dubna)	Stability Diagram of Exact Phase-Locked Soliton in the Parametrically Driven Damped Nonlinear Schrödinger Equation
4. L. Bonora (Trieste)	Classical and Quantum Toda Lattices
5. M. Bożejko (Wrocław)	Q-Deformed Probability
6. J. Burzlaff (Dublin)	Vortex-Vortex Scattering
7. G. Casati (Milan)	Classical Chaos, Quantum Localization and Fluctuation
8. A. M. Cetto (Mexico)	Detailed Balance and Radiative Corrections in Stochastic Electrodynamics
9. I. H. Duru (Gebze)	Path Integral Quantization of Particles and Fields, and QFT with Classical Particle Trajectories
10. A. Fordy (Leeds),	The Hamiltonian Structure of Lax Equation and Its Finite–Dimensional Reduction
11. W. Fushchich (Kiev)	Construction of Solutions for Nonlinear d'Alembert, Maxwell and Yang-Mills Equations in Terms of Solutions of Nonlinear Spinor Equations

FOUNDATION

FOR THE KARPACZ WINTER SCHOOLS IN THEORETICAL PHYSICS

The Institute of Theoretical Physics of the University of Wroclaw has organized, since 1964, Winter Schools of Theoretical Physics at Karpacz. Starting from a small scale enterprise it gradually developed into a school well known all over the world. The idea of the School, lasting two weeks in the second half of February, is to invite some 15 leading physicists from the West and the East to provide a series of lectures centered around a common topic. Since the traditional sources for financing the School, like our University, are running dry, we are turning to our friends and we inform you that we have created.

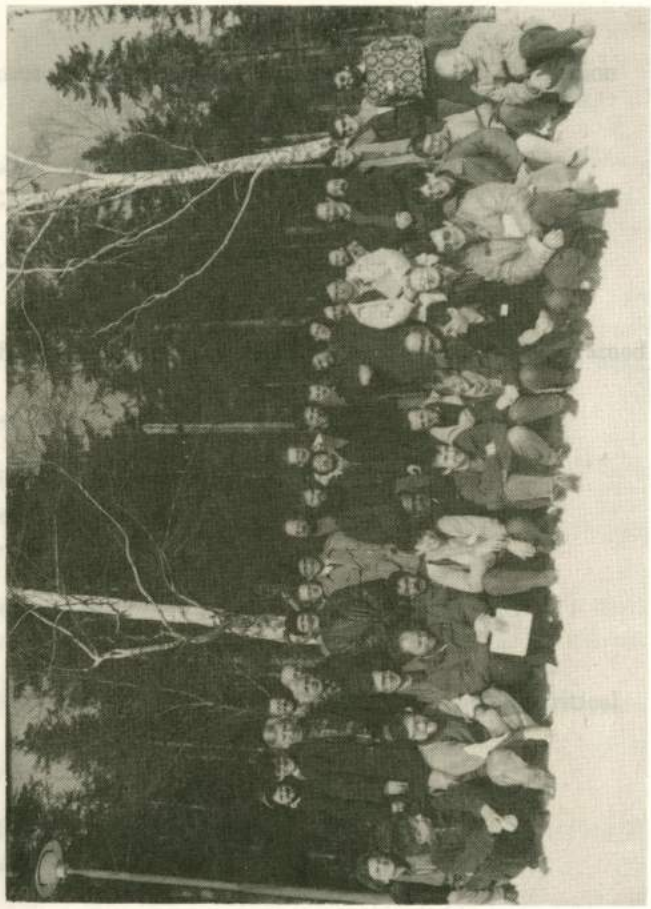
The Foundation for the Karpacz Winter Schools

with the obvious aim to help our School to survive the rainy days that are coming. Your generous help will be highly appreciated. The money we wish to use mainly for: covering the living, and if possible also travel, expenses for the invited lecturers, creating fellowships for young participants (also from countries poorer than ours), financial aid to our students who, also for years, organize in parallel with the School the so-called "Kindengarden in Theoretical Physics".

University of Wroclaw
Institute of Theoretical Physics
Cybulskiego 36, 50-205 Wroclaw, Poland
tel. (+48 71) 201 408 tx (+48 71) 5623 UWr
fax 402 800 E-mail IFTUWr at PLWRTU11

For the Foundation:
J. Rzewuski
P. Garbaczewski,
Z. Popowicz
R. Gielerak
T. Paszkiewicz,
A. Pekalski

Bank: PKO BP IV Oddzial Wroclaw nr: 93549-18193-132-3



Contents

PART I

CLASSICAL FIELDS: INTEGRABILITY, HIERARCHIES, VIRASORO ALGEBRAS

SELF–DUALITY, INTEGRABLE SYSTEMS, W–ALGEBRAS AND ALL THAT

Ioannis Bakas

Center for Theoretical Physics
Department of Physics and Astronomy
University of Maryland
College Park, MD 20742, USA

Table of Contents

To the memory of Stefanos Pneumatikos and Vassilis Xanthopoulos

0. Introduction and outline of the lectures

In last few years there have been many remarkable developments in the theory of integrable non–linear differential equations and 2–dim conformal field theories. Infinite dimensional algebras have provided the main framework for studing the general properties and exact solvability of these systems as well as their classification and other related physics problems (including critical phenomena, string theory and 2–dim quantum gravity). Although several questions remain unanswered up to this date, it is rather interesting that both integrable and conformal field theories in two dimensions share the same algebraic structures. It is the purpose of these lectures to present an introductory (yet comprehensive) account of this subject, while discussing the parallel connections between the symmetries and algebraic constructions traditionally used in the two areas. Special emphasis will be placed on the role of gauge symmetries which, as it turns out, enter into the theory in a fundamental way.

Operator algebras of 2–dim conformal field theory are generated by a collection of conserved (chiral) currents $\{W_s\}$, which also include the stress–energy tensor $T \equiv W_2$. Since the latter generates the Virasoro algebra

$$[T(x)\ ,T(x')] = (T(x)+T(x'))\delta'(x,x') + \frac{c}{12}\delta'''(x,x'),$$

all chiral operator algebras can be regarded as extended conformal symmetries whose structure constants obey the Jacobi identity constraints. The details of their structure depend crucially on the spin s (conformal dimension) of the additional generators they contain and in general they have quadratic determining relations. We shall refer to all such consistent extensions of the Virasoro algebra as W–algebras, which for the purposes of the present exposition will be assumed to be bosonic. The inclusion of fermionic generators becomes important in supersymmetric theories, which other lecturers of this school have considered in detail.

The classification of W–algebras and their unitary representations yield a large class (and possibly all) of rational conformal field theories depending on the value of the central charge c. This approach has been quite fruitful so far, both conceptually as well as practically, because it made the related 2–dim physics problems more tractable. A striking feature that emerged from this program is the realization that the structure and field theoretic representations of W–algebras can be

formulated entirely in terms of gauge symmetries and the classical theory of Lie algebras. These hidden relations between conformal and affine current algebras also arise in the Hamiltonian formulation of 2–dim integrable systems and in Polyakov's formulation of induced quantum gravity [1]. Their deep significance is by no means well understood, although much work has been done recently in this direction. The main core of my lectures centers around this theme, while discussing a selected list of topics related to my own work over the last few years.

In the first lecture I present the basic elements of the method of Hamiltonian reduction which yields W–algebras from affine current algebras by imposing constraints and using gauge fixing techniques. In this context I also discuss briefly the (twisted) Sugawara and free field realizations of W–algebras by adopting two different gauges. Classical W–algebras are also introduced from the point of view of integrable systems of KdV type and their free field realization is interpreted as the Miura transformation that maps KdV systems into the corresponding modified KdV equations. In the second lecture I provide an algorithm for embedding all KdV type flows into the self–dual Yang–Mills equations with two Killing symmetries and space–time signature $(2,2)$. This method explains naturally the origin of gauge symmetries in the theory of integrable systems and their formulation as zero curvature conditions. I also present a fractional class of integrable KdV type equations (together with an exotic generalization of bosonic W–algebras) using a non–standard set of constraints on the form of the 2–dim gauge connections for algebras with rank > 1.

In my opinion the intimate relations between conformal algebras and gauge symmetries should be regarded as one of the most important aspects of the 2–dim world. Apart from their numerous applications which are already known, the use of affine current algebras could be generalized appropriately to include a wide class of higher dimensional field theories. More precisely, it is conceivable that all gravitational theories (and even string theory) might exhibit hidden relations with gauge symmetries of Yang–Mills type for suitably chosen (possibly infinite dimensional) structure groups. Although there is circumstancial evidence to support this idea in certain cases, the exact formulation of the problem along these lines still remains unclear and out of reach in all generality. This point of view is partly put forward in my last lecture, where (area preserving) diffeomorphism algebras

are considered as *universal W–algebras* generated by an infinite tower of higher spin fields; in principle, any diffeomorphism group can be regarded as an infinite extension of the Virasoro algebra. I also present an alternative formulation of the self–dual Einstein equations in four dimensions, using a 2–dim theory with infinite dimensional gauge group. Hopefully, these "isolated" results could stimulate new directions for research in non–perturbative quantum field theory.

The search of hidden symmetries in higher dimensional theories and string theory remains undoubtedly a challenging question for future investigation. As it has been advocated by Witten in several occasions (see for instance [2]), better understanding of these issues is required before any real progress can be made in string theory and more generally in theories of quantum gravity.

These lectures are dedicated to the memory of Professors Stefanos Pneumatikos and Vassilis Xanthopoulos whose lives came to a tragic end in a shooting spree by a psychopath during a seminar at the University of Crete in Greece in November 1990.

1. Conformal algebras and gauge symmetries

1.1 *W–algebras and the method of Hamiltonian reduction*

Let us consider the space of gauge connections A which depend on a 1–dim (spatial) variable x and take values in the Lie algebra of a group G. It is convenient to represent the affine space of all such gauge connections by the covariant derivatives $D(A) = \partial_x + A$. Then, the infinite dimensional gauge group with values in G acts on $D(A)$ by conjugation

$$g^{-1}D(A)g = D(A^g) \tag{1.1}$$

and its orbits are simply described by the gauge transformations

$$A^g(x) = g^{-1}(x)A(x)g(x) + g^{-1}(x)g'(x). \tag{1.2}$$

The gauge connections A may also depend on additional (time) variables t, but at the moment this dependence will be suppressed. We will also restrict our discussion to the Lie groups $G \simeq SL(N)$ and use $N \times N$ traceless matrices to represent the values of the gauge connections $A(x)$.

Extended conformal symmetries exhibit hidden relations with affine current algebras, which can be easily described using the method of Hamiltonian reduction

[3, 4, 5]. For this purpose we introduce constraints on the form of the $SL(N)$ gauge connections, setting all strictly upper triangular matrix elements equal to zero apart from the elements of the first diagonal line whose values are conventionally normalized to -1. The Borel subgroup of gauge transformations with values in the strictly lower triangular matrices acts by conjugation on the gauge connections

$$D(A) = \partial_x + \begin{pmatrix} J_{1,1} & -1 & 0 & \dots & 0 \\ J_{2,1} & J_{2,2} & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ J_{N-1,1} & J_{N-1,2} & \dots & J_{N-1,N-1} & -1 \\ J_{N,1} & J_{N,2} & \dots & J_{N,N-1} & J_{N,N} \end{pmatrix}, \quad \sum_{i=1}^{N} J_{ii} = 0 \tag{1.3}$$

and clearly preserves the form of the upper triangular constraints. It is possible to show in all generality that the space of gauge equivalence classes that arises in this fashion can be parametrized by $\partial_x + Q$, where

$$Q = \begin{pmatrix} 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & -1 \\ u_N & u_{N-1} & \dots & u_2 & 0 \end{pmatrix} \tag{1.4}$$

depends on $N-1$ degrees of freedom. The residual symmetry on the orbit space (1.4) which is parametrized by $\{u_s\ ;\ s = 2,\ 3,\ \cdots,\ N\}$ is generated by the W_N algebra. In fact, the method of Hamiltonian reduction that we have just described can be regarded as the definition of classical W–symmetries of $SL(N)$ type. Their quantum version will be discussed briefly in the next section.

The simplest example of W_N algebras is provided by the Virasoro algebra which corresponds to $N = 2$. In this case we find that the diagonal degree of freedom in (1.3) can be gauged away, using the group element with $g_{1,1} = g_{2,2} = 1$, $g_{1,2} = 0$ and $g_{2,1} = J_{1,1}$. Explicit calculation also shows that the residual degree of freedom on the space of gauge equivalence classes is given by

$$u_2 = J_{2,1} - J_{1,1}^2 + J_{1,1}', \tag{1.5}$$

which is non–linear in the original J variables. Hence, it is reasonable to expect that the algebra generated by the field u_2 is not a subalgebra of the algebra generated by the J fields. Indeed, if we identify $J_{1,1} = J^{(0)}$ and $J_{2,1} = J^{(-)}$ and impose the $SL(2)$ Kac–Moody commutation relations

$$\{J^{(0)}(x)\ ,\ J^{(-)}(x')\} = J^{(-)}(x)\delta(x,x'), \tag{1.6a}$$

$$\{J^{(0)}(x)\ ,\ J^{(0)}(x')\} = -\frac{1}{2}\delta'(x,x'), \tag{1.6b}$$

we find immediately

$$\{u_2(x)\ ,\ u_2(x')\} = (u_2(x) + u_2(x'))\delta'(x,x') + \frac{1}{2}\delta'''(x,x'), \tag{1.7}$$

which is the Virasoro algebra with $c = 6$.

For the $SL(3)$ case similar calculation yields

$$u_2 = J_{3,2} + J_{2,1} - J_{1,1}^2 - J_{2,2}^2 - J_{1,1}J_{2,2} + (2J_{1,1} + J_{2,2})', \tag{1.8a}$$

$$u_3 = J_{3,1} + J_{1,1}J_{3,2} - (J_{1,1} + J_{2,2})(J_{2,1} + J_{1,1}J_{2,2}) + \\ + (J_{2,1} + J_{1,1}J_{2,2} + J_{1,1}')' - (J_{1,1} + J_{2,2})J_{1,1}', \tag{1.8b}$$

using the gauge transformation with $g_{1,1} = g_{2,2} = g_{3,3} = 1$, $g_{1,2} = g_{2,3} = g_{1,3} = 0$ and $g_{2,1} = J_{1,1}$, $g_{3,2} = J_{1,1} + J_{2,2}$, $g_{3,1} = J_{2,1} + J_{1,1}J_{2,2} + J_{1,1}'$. Imposing $SL(3)$ Kac–Moody commutation relations on the J fields and using the redefinition

$$w_3(x) = u_3(x) - \frac{1}{2}u_2'(x) \tag{1.9}$$

we arrive at the classical commutation relations of Zamolodchikov's W_3 algebra [6]

$$\{u_2(x)\ ,\ u_2(x')\} = (u_2(x) + u_2(x'))\delta'(x,x') + 2\delta'''(x,x'), \tag{1.10a}$$

$$\{u_2(x)\ ,\ w_3(x')\} = (w_3(x) + 2w_3(x'))\delta'(x,x'), \tag{1.10b}$$

$$\{w_3(x)\ ,\ w_3(x')\} = -\frac{1}{6}\delta''''(x,x') - \frac{5}{12}(u_2(x) + u_2(x'))\delta'''(x,x') + \\ + \frac{1}{4}(u_2''(x) + u_2''(x'))\delta'(x,x') - \frac{1}{3}(u_2^2(x) + u_2^2(x'))\delta'(x,x'). \tag{1.10c}$$

Eq.(1.10a) shows that the Virasoro algebra is contained in W_3 as a subalgebra (with $c = 24$) and eq.(1.10b) implies that w_3 is a primary conformal field with spin 3.

The operator content as well as the classical commutation relations of W_N algebras with arbitrary values of N can be described in a similar fashion using the $SL(N)$ Kac–Moody commutation relations

$$\{J^a(x)\ ,\ J^b(x')\} = f_c^{ab}J^c(x)\delta(x,x') - (T^a,T^b)\delta'(x,x'), \tag{1.11}$$

after making suitable identifications with the lower triangular matrix elements of the gauge connections (1.3). Here, $\{T^a\}$ are the Lie algebra generators of $SL(N)$ and $(T^a\ ,\ T^b) \equiv g^{ab}$ is the corresponding Killing metric. It turns out that u_2 generates the Virasoro subalgebra of W_N with central charge $c = N^3 - N$ (which as we will see later is related to the length of the Weyl vector of the $SL(N)$ algebra) and all other generators u_s with $s \geq 3$ have conformal dimension (spin) s. Although the presence of higher spin fields gives rise to non–linear terms in the commutation relations of W_N algebras with $N \geq 3$ (c.f. eq.(1.10c)), the details of their structure are always consistent with the Jacobi identity. We also point out that the higher spin fields u_s are not necessarily primary with respect to the u_2 transformations. Nevertheless, it is always possible to introduce suitable field redefinitions of the form

$$w_s(x) = \sum_{i,k} A^{i_1 \cdots i_p}_{N;k_1 \cdots k_p} u^{(k_1)}_{i_1}(x) \cdots u^{(k_p)}_{i_p}(x) \tag{1.12}$$

with $k_1 + \cdots + k_p + i_1 + \cdots + i_p = s \leq N$, so that w_s are primary with spin s. The field redefinition (1.9) is an example of this formula, but for arbitrary values of s and N the numerical coefficients $\{A^{i_1 \cdots i_p}_{N;k_1 \cdots k_p}\}$ are not universal and hence difficult to compute in closed form. The structure constants of the higher spin commutation relations of W_N algebras are also not available in closed form for arbitrary N. The method of Hamiltonian reduction, however, provides a systematic algorithm for their computation case by case.

From a mathematical point of view, the space of $SL(N)$ gauge connections A can be identified with the (smooth) dual of the Kac–Moody algebra (1.11). As such it is naturally equipped with a linear Poisson manifold structure given by the Kirillov bracket associated with the affine current algebra (1.11). The space of gauge equivalence classes (1.4) also inherits a Poisson manifold structure by Hamiltonian reduction, but it is non–linear since the corresponding W_N algebras are not Lie algebras. We could have implemented this procedure more generally, starting with the first order differential operators $k\partial_x + A$, with arbitrary values of k. Then, the central charge of the underlying Kac–Moody algebra would be k and certain structure constants of the resulting W_N commutation relations (including the central charge of the Virasoro subalgebra) would be modified accordingly. It is important to emphasize, however, that this generalization does not affect the operator content of the residual symmetry algebras.

1.2 *Field theoretic realizations and quantization*

The expressions for the generating fields u_s of the W_N algebras can be written in a more compact form using the constrained matrices J that appear in eq.(1.3). In particular we have

$$u_s = (-)^{s+1} \frac{1}{s} Tr J^s + \textit{derivative terms} \tag{1.13}$$

for all values of $s \geq 2$. Consequently the method of Hamiltonian reduction provides a realization of the W_N algebras in terms of the original $SL(N)$ currents *à la* Sugawara. The new feature however is that there are derivative terms present in the formalism. This modification is very important and accounts for the varying conformal dimension of the residual fields u_s, according to their location in the matrix (1.4). If u_2 was given by the original Sugawara construction, all J fields would clearly have spin 1. In the present case, the derivative terms introduce a twisting which changes the spin of the fields in the following way: the matrix elements of the first upper triangular line above the main diagonal acquire spin 0 and as we move down the spin increases by 1 across the diagonals in every step of the way. This assignment is consistent with the constraints imposed on the strictly upper triangular part of the gauge connections. Also, the spin of the fields with values in the Cartan subalgebra remains unchanged because u_2 contains only total derivative terms of the diagonal elements.

It is important to realize at this point that the parametrization (1.4) of the space of gauge equivalence classes with respect to the strictly lower triangular Borel subgroup of $SL(N)$ is by no means unique. Equally well we could have described the orbit space by gauge connections of the form [3, 4]

$$\tilde{Q} = \begin{pmatrix} q_1 & -1 & 0 & \dots & 0 \\ 0 & q_2 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & q_{N-1} & -1 \\ 0 & 0 & \dots & 0 & q_N \end{pmatrix} \quad , \quad \sum_{i=1}^{N} q_i = 0. \tag{1.14}$$

In this case the residual symmetry is generated by the spin 1 fields $\{q_i\}$ and consists of $N-1$ copies of the $U(1)$ Kac–Moody algebra.

The relation between the q and u fields in the two different slices (1.4) and (1.14) is quite interesting and has applications both in conformal field theory as well as in

the theory of integrable systems of KdV type. For $N = 2$ we find

$$u_2 = -q_1^2 + q_1', \tag{1.15}$$

while for more arbitrary values of N the transformation is described by the following differential operator equation

$$L_N \equiv \partial_x^N + u_2\partial_x^{N-2} + \cdots + u_N = (\partial_x + q_N)\cdots(\partial_x + q_1). \tag{1.16}$$

This formula follows immediately by equating the determinant of $\partial_x + Q$ with the determinant of $\partial_x + \tilde{Q}$. Since the q fields behave as $U(1)$ currents, the above relation provides a free field realization of W_N algebras. For $N = 2$, for example, if we identify $q_1 \equiv \partial\varphi/2$ we will obtain

$$u_2 = -\frac{1}{4}(\partial\varphi)^2 + \frac{1}{2}\partial^2\varphi, \tag{1.17}$$

which is nothing else but the Feigin–Fuchs free field realization of the Virasoro algebra.

More generally, let us choose to work with $k\partial_x + A$ with $k = i\alpha_0$ and let $\vec{\varphi} = (\varphi_1\ ,\ \varphi_2\ ,\ \cdots\ ,\ \varphi_{N-1})$ be an array of $N-1$ free massless scalar fields with Poisson bracket

$$\{\partial\varphi_i(x)\ ,\ \partial\varphi_j(x')\} = -2\delta_{ij}\delta'(x,x'). \tag{1.18}$$

We also introduce an overcomplete set of vectors $\{\vec{h}_k\ ;\ k = 1\ ,\ 2\ ,\ \cdots\ ,\ N\}$ in $(N-1)$–dim Euclidean space with

$$\sum_{k=1}^{N}\vec{h}_k = 0\quad ,\quad 2\vec{h}_k\cdot\vec{h}_l = \delta_{kl} - \frac{1}{N}. \tag{1.19}$$

Then, using the identification

$$\sum_{s=0}^{N} u_s(i\alpha_0\partial_x)^{N-s} = \prod_{k=1}^{N}(i\alpha_0\partial_x + \vec{h}_k\cdot\partial\vec{\varphi}), \tag{1.20}$$

with $u_0 = 1$ and $u_1 = 0$, we obtain the Feigin–Fuchs realization of the W_N algebra with arbitrary background charge α_0 [7]. The particular expressions are straightforward to derive from eq.(1.20) using Leibnitz's rule. Here we only give the result for the spin 2 field

$$u_2 = -\frac{1}{4}\sum_{i=1}^{N-1}(\partial\varphi_i)^2 + i\alpha_0\vec{\rho}\cdot\partial^2\vec{\varphi}, \tag{1.21}$$

where $\vec{\rho}$ is the Weyl vector of $SL(N)$

$$\vec{\rho} = \frac{1}{2}\sum_{k=1}^{N}(N+1-2k)\vec{h}_k \tag{1.22}$$

with $24\vec{\rho}^{\,2} = N^3 - N$.

Using the free field realization of W_N algebras, if we adopt normal orderings on the right hand side of eq.(1.20), we will obtain the quantum mechanical version of the Poisson bracket commutation relations of W_N. Of course, the non–linear terms that appear in the commutation relations of higher spin fields have to be regularized appropriately so that the resulting composite operators are well–defined. Although multiple contractions of the free fields lead to deformations of certain structure constants quantum mechanically, the main features of the algebras remain unchanged. In analogy with the classical case, there are no closed formulas for the quantum mechanical commutation relations of W_N algebras with arbitrary N. However, using eq.(1.21) one may easily verify that the value of the central charge of the Virasoro subalgebra becomes

$$c = (N-1)(1-\alpha_0^2 N(N+1)). \tag{1.23}$$

It is known that for $\alpha_0^2 = 1/(N+p)(N+p+1)$ with $p = 1,\ 2,\ 3, \cdots$ the unitary representations of W_N algebras are special in that the number of the corresponding building blocks is finite and hence the underlying series of 2–dim conformal field theories is "minimal." Detailed discussion of these results is beyond the scope of the present lectures and we refer the reader to the literature for further details [7, 8].

The free field realization of W algebras is also important in the theory of integrable systems of KdV type, where as we will see later the transformation law (1.16) maps KdV flows into their modified versions *à la* Miura. For completeness we also point out that W algebras can be constructed in a similar way for simple Lie algebras other than $SL(N)$. Their systematic presentation, however, is also beyond the scope of these lectures (for details see [9]).

1.3 *Associated systems of non–linear differential equations*

We turn now to the Hamiltonian formulation of integrable systems of KdV type in terms of W_N algebras. The r–th flow of the $SL(N)$ KdV hierarchy is traditionally defined using the differential operator $L_N = \partial_x^N + u_2\partial_x^{N-2} + \cdots + u_N$ (see for instance

[10] and references therein). The corresponding system of non–linear differential equations has the general form

$$\dot{u}_i = f_{N;r,i}(u, u', u'', \cdots), \tag{1.24}$$

where $f_{N;r,i}$ is a homogeneous expression of the u fields and their derivatives of degree (spin) $r+i$. The particular dependence of these evolutions on the fields $\{u_s\}$ is given by the scalar Lax equation

$$\frac{\partial L_N}{\partial t_r} = [(L_N{}^{r/N})_+ \ , \ L_N] \tag{1.25}$$

for all positive integer values of r. Here, $+$ denotes the purely differential part of $L_N{}^{r/N}$ which is uniquely defined within the calculus of formal pseudo–differential operators. The values $r = 0,\ N,\ 2N,\ 3N,\ \cdots$ are usually excluded because the evolutions are trivial in the sense that $\dot{u}_i = 0$. Also, the $r = 1$ flows are not interesting since they give rise to the linear differential equations $\dot{u}_i = u_i'$.

The space of differential operators L_N provides a useful parametrization of the space of gauge equivalence classes (1.4). Given the fact that the residual symmetry on this space is generated by the W_N algebra, it is reasonable to expect that all KdV type flows (1.25) are Hamiltonian with respect to the commutation relations of W_N. This is indeed the case, i.e.,

$$\dot{u}_i = \{u_i \ , \ \mathcal{H}_r\} \tag{1.26}$$

for appropriately chosen Hamiltonian functionals $\mathcal{H}_r$ (for details see [11] and references therein). We also point out for completeness that all evolutions of KdV type are bi–Hamiltonian in the sense that there is another bracket $\{\ ,\ \}_*$ which casts the non–linear equations (1.25) into Hamiltonian form. The latter is easily described in terms of W_N algebras as follows: if we shift the highest spin field u_N by a constant, say λ, the classical commutation relations of W_N acquire a linear modification which by itself can be used as a Lie–Poisson bracket for the KdV flows. In this case, however, the corresponding Hamiltonian functionals that one uses are different. It is important to emphasize that the existence of a bi–Hamiltonian structure can be used recursively to construct all conservation laws of the KdV hierarchies and prove the complete integrability of the systems. It turns out that for any value of N, all members of the hierarchy (1.25) share a common set of infinitely many

conserved quantities, which also coincide with the Hamiltonian functionals $\mathcal{H}_r$. The specific details will be omitted because they have become quite standard by now (see for instance [12, 13]) and have been discussed extensively by other lectures of this school.

There is an alternative formulation of the KdV flows (1.25) as zero curvature conditions of the form

$$[\partial_x + Q \ , \ \partial_{t_r} - H_r] = 0, \tag{1.27}$$

where Q is given by eq.(1.4) and H_r is an $SL(N)$ matrix whose elements depend on r [3]. This matrix description will be the starting point for embedding all KdV type flows into the self–dual Yang–Mills equations. In the next section we will also present a systematic way for constructing the matrix elements of the Hamiltonian matrices H_r. For the moment being (and for pedagogical reasons), we will summarize all the different formulations of the ordinary KdV flow, which is the simplest non–trivial example corresponding to $N = 2$ and $r = 3$. In this case, we have

$$L_2 = \partial_x^2 + u_2 \quad , \quad (L_2{}^{3/2})_+ = \partial_x^3 + \frac{3}{2}u_2\partial_x + \frac{3}{4}u_2' \tag{1.28}$$

and the Lax equation (1.25) yields

$$\dot{u}_2 = \frac{1}{4}(u_2''' + 6u_2u_2'). \tag{1.29}$$

The KdV equation (1.29) is Hamiltonian with respect to the commutation relations of the Virasoro algebra (1.7) provided that the Hamiltonian functional is chosen to be $\mathcal{H} = \frac{1}{4}\int u_2^2$. This equation also admits a zero curvature formulation in terms of $SL(2)$ gauge connections, provided that the Hamiltonian matrix in eq.(1.27) is given by

$$H = -\frac{1}{4}\begin{pmatrix} u_2' & -2u_2 \\ u_2'' + 2u_2^2 & -u_2' \end{pmatrix}. \tag{1.30}$$

The modified KdV type equations are related to the original flows (1.25) by the Miura transformation which maps the space of gauge equivalence classes (1.4) into (1.14). For example, for $N = 2$ we have the relation (1.15) which casts the KdV equation (1.29) into the modified form

$$\dot{q}_1 = \frac{1}{4}(q_1''' - 6q_1'q_1^2). \tag{1.31}$$

More generally, the explicit form of the Miura transformation follows from eq.(1.16) using Leibnitz's rule and the modified versions of all KdV type flows can be easily obtained by making the appropriate substitutions into the scalar Lax equation (1.25).

2. Self–duality and integrable systems of KdV type

2.1 *Systematic embedding of all SL(N) KdV flows into sdYM theory*

Next we will consider the embedding of KdV type flows into the self–dual Yang–Mills equations (sdYM)

$$F(A) = {}^*F(A). \tag{2.1}$$

The main motivation for this study is provided by the close connection of all KdV flows (1.25) with gauge symmetries. Hence, one hopes to understand the origin of gauge symmetries in the theory of W algebras and the associated systems of non–linear differential equations from a higher dimensional point of view in terms of Yang–Mills fields. In fact, the results that we will present here are part of a more general program that has the ambition to describe all integrable systems in two dimensions as dimensional reductions of the 4–dim self–dual Yang–Mills equations [14, 15, 16, 17]. In writing eq.(2.1) one has the freedom of choosing different gauge groups as well as space–times with signature $(4,0)$ or $(2,2)$. Usually, different ansatz for the gauge connections A lead to distinct classes of non–linear evolutions in two dimensions which can be integrated by the inverse scattering method. It turns out that for equations of KdV type, the appropriate starting point is the self–dual theory (2.1) defined on R^4 with signature $(2,2)$ and $SL(N)$ as gauge group.

Let us consider first KdV flows of $SL(N)$ type with $r < N$. In this case the Hamiltonian matrix H_r used in the zero curvature description of the equations is linear in the highest spin field u_N. Notice that the homogeneity of the KdV flows ensures that the matrix elements of H_r lying on the same diagonal line have the same degree (spin). In analogy with the matrix Q, the spin of the diagonal lines in H_r also increases by 1 every time we move across the diagonals towards the bottom left corner. For the r–th flow, the spin of the matrix element $H_{N,1}$ is $r+N-1 < 2N$ and since it is the maximum allowed we conclude that H_r can not be quadratic in u_N. Also, as far as the strictly upper triangular part of H_r is concerned, only its

first $r-2$ diagonal lines contain elements that depend on the u fields. Then, if we shift u_N by a constant, $u_N \to u_N + \lambda$, both matrices Q and H_r in eq.(1.27) will be modified linearly in λ : $Q \to Q - \lambda B$ and $H_r \to H_r - \lambda P_r$, where

$$B = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & 0 \\ 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 \end{pmatrix} \tag{2.2}$$

and P_r depends on the explicit form of H_r, which for the present argument is not very important. Requiring that the zero curvature condition (1.27) is consistent for all values of the parameter λ we obtain the system of equations

$$[\partial_x + Q \ , \ \partial_{t_r} - H_r] = 0, \tag{2.3a}$$

$$[\partial_x + Q \ , \ P_r] = [H_r \ , \ B], \tag{2.3b}$$

$$[P_r \ , \ B] = 0 \tag{2.3c}$$

to order λ^0, λ, and λ^2 respectively.

The main point of this description is that eqs.(2.3) are equivalent to the sdYM equations with two Killing symmetries [16, 17]. To make the connection more precise, let us introduce in R^4 the metric

$$ds^2 = 2dxdy + 2dzdt \tag{2.4}$$

with signature $(2,2)$. Using the convention $D = \partial - A$ for the covariant derivatives and $\epsilon_{xyzt} = -1$, we find that the self–duality conditions (2.1) take the form

$$F_{xt} = 0 \ , \ F_{xy} = F_{zt} \ , \ F_{yz} = 0. \tag{2.5}$$

Then, if we impose two null Killing symmetries with respect to the y and z coordinates and use the identification

$$A_x = -Q \ , \ A_y = P_r \ , \ A_z = B \ , \ A_t = H_r, \tag{2.6}$$

we can easily show that the self–dual system (2.5) is equivalent to (2.3). This result proves that all flows of the $SL(N)$ KdV hierarchies with $r < N$ can be embedded into the sdYM equations. The main point here is that with the help of the spectral parameter λ, the self–duality conditions (2.5) can be brought into a zero curvature

form which underlies the formulation (1.27) of the KdV flows after dimensional reduction.

The embedding of $SL(N)$ flows with $r < N$ into the sdYM equations can be generalized to all other cases, provided that the rank of the gauge group increases accordingly. We point out that for flows of the $SL(N)$ hierarchy with $r > N$, we can always choose a positive integer M which is sufficiently large so that $MN > r > N$. Recall that the r–th flow of the $SL(N)$ hierarchy can be viewed as a special case of the r–th flow for $SL(MN)$, provided that the corresponding scalar Lax operator L_{MN} is not arbitrary but depends only on $N-1$ functions in the following way:

$$L_{MN} = (\partial_x^N + u_2\partial_x^{N-2} + \cdots + u_N)^M. \tag{2.7}$$

The details of the proof will be omitted because they are quite standard and straightforward. Here, we only note that the self–dual embedding of $SL(N)$ flows with $r > N$ into the system (2.3) is identical to the $r < N$ case provided that we use $SL(MN)$ as gauge group and that the matrix elements of the $MN \times MN$ matrices Q and H_r incorporate the constraint (2.7) accordingly. Clearly, within the present framework, the standard KdV equation (1.29) can not be described as a special case of the sdYM equations using 2×2 matrices (c.f. eq.(1.30)). However, as we will see in the next section there is an alternative dimensional reduction of the self–dual $SL(N)$ Yang–Mills theory which is applicable to flows with $r = N+1$ (and hence (1.29)) and does not require the introduction of higher rank groups.

Having presented the main motivation for considering the dimensional reduction (2.3) of the sdYM equations, we are now in the position to describe in all generality the systematic derivation of all KdV type flows with $r < N$ from the system (2.3). It turns out that if we introduce the ansatz (1.4) and (2.2) for the matrices Q and B, the self–dual equations (2.3) are sufficient to determine all matrix elements of P and H, up to a choice of certain integration constants, as well as the time evolution of the fields $\{u_s\ ;\ s = 2,\ 3,\ \cdots,\ N\}$ [17]. These constants can be chosen to be zero or 1 by requiring homogeneity of the resulting differential equations. A careful analysis of the problem shows that there are N different ways of choosing the integration constants which then lead to the members of the $SL(N)$ KdV hierarchy with $r = 0,\ 1,\ 2, \cdots,\ N-1$. This way we obtain a systematic way for computing the element of the Hamiltonian matrix H_r in eq.(1.27) directly from the embedding

of the corresponding flows into the sdYM theory. We also point out that the choice of integration constants in this problem effectively fixes the u independent elements in the upper triangular part of H_r. More precisely, for any given value of $r < N$, the $(r-1)$–th strictly upper diagonal line has spin 1 and therefore its elements are all zero. The elements of the r–th line are constant (spin zero) and normalized to 1, while every other element above it, since it has negative spin, is set equal to zero for consistency.

For completeness we present the result of this method for the group $SL(3)$. In this case the only non–trivial flow with $r < N$ has $r = 2$ and the matrices P_r and H_r are given by the expessions

$$P = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad H = \frac{1}{3}\begin{pmatrix} 2u_2 & 0 & 3 \\ 2u_2' - 3u_3 & -u_2 & 0 \\ (2u_2' - 3u_3)' & u_2' - 3u_3 & -u_2 \end{pmatrix}. \tag{2.8}$$

Then, the corresponding KdV flow is described by the system

$$\dot{u}_2 = 2(u_3 - \frac{1}{2}u_2')', \tag{2.9a}$$

$$\dot{u}_3 = (u_3' - \frac{2}{3}u_2'' - \frac{1}{3}u_2^2)'. \tag{2.9b}$$

After elimination of the u_3 field we obtain the well–known Boussinesq equation

$$\ddot{u}_2 = -\frac{1}{3}(u_2'' + 2u_2^2)''. \tag{2.10}$$

Similarly, one may work out more complicated examples following the method we have outlined.

2.2 *Other dimensional reductions of sdYM equations*

Let us now consider an alternative dimensional reduction of the sdYM equations in a space–time with signature $(2,2)$ [16, 18]. We introduce the metric

$$ds^2 = dx^2 - dy^2 - 4dzdt \tag{2.11}$$

and use the convention $D = \partial - A$ and $\epsilon_{xyzt} = 1$. Then the system (2.1) becomes

$$F_{xt} = F_{yt}\ ,\ F_{xz} = F_{zy}\ ,\ 2F_{xy} = F_{tz}, \tag{2.12}$$

which in the presence of two Killing symmetries (one time–like with respect to y and one null with respect to z) yields

$$[\partial_x + Q\ ,\ \partial_t - H] = 0, \tag{2.13a}$$

$$[\partial_x + Q\ ,\ \partial_x - P] = [H\ ,\ B], \tag{2.13b}$$

$$[P\ ,\ B] = 0. \tag{2.13c}$$

Here we have assumed that B is a constant matrix (also given by eq.(2.2)) and that

$$Q = -A_x + A_y,\quad P = A_x + A_y,\quad A_z = B,\quad A_t = H. \tag{2.14}$$

Notice that the new system of equations is almost identical to (2.3) apart from P term in eqs.(2.3b) and (2.13b) respectively. This difference plays a significant role in the detailed structure of the resulting differential equations. In fact, if we introduce the ansatz (1.4) and (2.2) as before, we find that the form of H and P is uniquely determined by the requirement of homogeneity alone and the corresponding evolution of the u fields coincides with the $r = N + 1$ flow of the $SL(N)$ KdV hierarchy [18]. Although this formalism does not incorporate all possible flows of the $SL(N)$ KdV hierarchy, it has the advantage that the self–dual description of the $r = N + 1$ flow can be given entirely in terms of $N \times N$ traceless matrices. For the $N = 2$ case we find that H is given by eq.(1.30), while for the matrix P the only non–zero element is $P_{2,1} = -\frac{1}{2}u_2$. We point out that in this case P is not related to H in a simple way as in the dimensional reduction considered earlier.

To conclude our discussion of this topic we present the results of the next non–trivial example with $N = 3$ and $r = 4$. Explicit calculation shows that the non–zero matrix elements of P are

$$P_{2,1} = P_{3,2} = -\frac{1}{3}u_2,\quad P_{3,1} = -\frac{1}{3}(u_2' + u_3). \tag{2.15}$$

We also find that the Hamiltonian matrix is given by

$$H_{1,1} = \frac{2}{9}(u_2'' + u_2^2) - \frac{1}{3}u_3',$$

$$H_{2,1} = \frac{1}{9}(2u_2''' - 3u_3'' + 4u_2u_2' - 3u_2u_3),$$

$$H_{3,1} = -\frac{1}{3}u_3^2 + H_{2,1}',\quad H_{1,3} = \frac{1}{3}u_2,\quad H_{2,3} = \frac{1}{3}u_3,$$

$$H_{1,2} = \frac{1}{3}(u_3 - u_2'),\quad H_{2,2} = -\frac{1}{9}(u_2'' + u_2^2),$$

$$H_{3,2} = \frac{1}{9}(u_2''' - 3u_3'' + 2u_2u_2' - 6u_2u_3). \tag{2.16}$$

Then, the evolution equations (2.13a) yield the 4–th flow of the $SL(3)$ hierarchy (1.25) which explicitly reads as follows:

$$3\dot{u}_2 = -u_2'''' + 2u_3''' - (u_2^2)'' + 4(u_2u_3)', \tag{2.17a}$$

$$9\dot{u}_3 = -2u_2'''' + 3u_3'''' - 6u_2u_2''' - 12u_2'u_2''-$$

$$-4u_2^2u_2' + 6(u_2u_3')' + 6(u_3^2)'. \tag{2.17b}$$

2.3 *Fractional KdV hierarchies and $W_N^{(l)}$ algebras*

For algebras with rank > 1 there are more than one ways to constrain the upper triangular part of the gauge connections. For $SL(N)$, in particular, there are $l = 1, 2, \cdots, N-1$ different possibilities to choose which diagonal line of the strictly upper triangular part of the $SL(N)$ matrices is fixed to 1. So far we have explored the $l = 1$ case in connection with the theory of W algebras and the associated integrable systems of KdV type. In this section we consider the generalization $l > 1$ and in order to simplify the presentation we will restrict ourselves to the simplest example $N = 3$ and $l = 2$. Using the method of Hamiltonian reduction we find that the residual symmetry on the space of gauge equivalence classes Q is generated by four fields U, $G^{(\pm)}$ and T. A convenient parametrization of the orbit space is given by gauge connections of the form

$$Q = -\begin{pmatrix} -\frac{1}{2}U & 0 & 1 \\ G^{(+)} & U & 0 \\ T & G^{(-)} & -\frac{1}{2}U \end{pmatrix}, \tag{2.18}$$

which generalizes (1.4) to the present case.

The W algebra on the space of gauge equivalence classes (2.18) is denoted by $W_3^{(2)}$ and its spectrum is identical to the $N = 2$ super Virasoro algebra [19]. In particular, U is an abelian current with spin 1, $G^{(\pm)}$ have spin 3/2 and T has spin 2. Unlike the $N = 2$ super Virasoro algebra, however, all generators of $W_3^{(2)}$ are bosonic and hence its commutation relations have non–linear terms. It turns out

that the identification of T with the generator of the Virasoro subalgebra can be made correctly, provided that the field redefinition

$$T \to T - \frac{3}{4}U^2 \tag{2.19}$$

is taken into account. Straightforward calculation shows that the classical commutation relations of $W_3^{(2)}$ are given by

$$\{U(x)\ ,\ U(x')\} = \frac{2}{3}\delta'(x,x'), \quad \{T(x)\ ,\ U(x')\} = U(x)\delta'(x,x'),$$

$$\{U(x)\ ,\ G^{(\pm)}(x')\} = \pm G^{(\pm)}(x)\delta(x,x'), \quad \{G^{(\pm)}(x)\ ,\ G^{(\pm)}(x')\} = 0,$$

$$\{T(x)\ ,\ G^{(\pm)}(x')\} = (G^{(\pm)}(x) + \frac{1}{2}G^{(\pm)}(x'))\delta'(x,x'),$$

$$\{T(x)\ ,\ T(x')\} = (T(x)+T(x'))\delta'(x,x') - \frac{1}{2}\delta'''(x,x'),$$

$$\{G^{(+)}(x)\ ,\ G^{(-)}(x')\} = (3U^2(x) - T(x))\delta(x,x') + \frac{3}{2}(U(x)+U(x'))\delta'(x,x') + \delta''(x,x'). \tag{2.20}$$

Alternatively we could have considered another parametrization of the orbit space which is analogous to (1.14). For $l = 2$, however, we must introduce a two–component scalar field $\vec{\varphi}$ on the main diagonal as well as two bosonic scalar fields ϕ and ϕ^+ with spin $1/2$ which sit on the first line above the main diagonal. This way we obtain the free field realization of $W_3^{(2)}$ [19]

$$U = \frac{\sqrt{2}}{3}(\vec{e_2} - \vec{e_1})\cdot\partial\vec{\varphi} + \phi\phi^+, \tag{2.21}$$

$$G^{(+)} = \partial\phi^+ + \sqrt{2}(\vec{e_1}\cdot\partial\vec{\varphi})\phi^+ - (\phi^+)^2\phi, \tag{2.22}$$

$$G^{(-)} = \partial\phi + \sqrt{2}(\vec{e_2}\cdot\partial\vec{\varphi})\phi + \phi^+\phi^2, \tag{2.23}$$

$$T = \frac{2}{3}[(\vec{e_1}\cdot\partial\vec{\varphi})^2 + (\vec{e_2}\cdot\partial\vec{\varphi})^2 + (\vec{e_1}\cdot\partial\vec{\varphi})(\vec{e_2}\cdot\partial\vec{\varphi})] + \\ + \frac{1}{\sqrt{2}}(\vec{e_1}+\vec{e_2})\cdot\partial^2\vec{\varphi} + \frac{1}{2}(\partial\phi^+\phi - \partial\phi\phi^+), \tag{2.24}$$

where $\vec{e_1}$ and $\vec{e_2}$ are the simple (positive) roots of $SL(3)$ ($\vec{e_1} = \frac{1}{\sqrt{2}}(1,\sqrt{3})$, $\vec{e_2} = \frac{1}{\sqrt{2}}(1,-\sqrt{3})$), which are normalized as $\vec{e_1}^2 = \vec{e_2}^2 = 2$, $\vec{e_1}\cdot\vec{e_2} = -1$. Using the commutation relations of the free fields

$$\{\partial\varphi_a(x)\ ,\ \partial\varphi_b(x')\} = \frac{1}{2}\delta_{a,b}\delta'(x,x'), \quad \{\phi(x)\ ,\ \phi^+(x')\} = \delta(x,x'), \tag{2.25}$$

one may easily verify the commutation relations of the classical $W_3^{(2)}$–algebra (2.20). As in the $l = 1$ case, the values of certain structure constants are modified quantum mechanically, but this aspect of the problem is beyond the scope and applications of the present exposition.

Let us now consider in some detail the non–linear systems that follow from the sdYM equations (2.3) with Q given by the expression (2.18) [17, 20]. In this case we also choose B to be

$$B = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}. \tag{2.26}$$

Then, the self–duality conditions have two solutions: the first one is described by $H = -Q$, $P = -B$ and leads to the evolution

$$\dot{U} = U', \quad \dot{G}^{(\pm)} = G^{(\pm)\prime}, \quad \dot{T} = T', \tag{2.27}$$

while the second one is more interesting and corresponds to the choice

$$P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad H = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ 3U & 0 & 2 \\ G^{(+)} + G^{(-)} & 3U & 0 \end{pmatrix}. \tag{2.28}$$

Using the redefinition (2.19) and the sign change $\partial_x \to -\partial_x$, $\partial_t \to -\partial_t$ we obtain the following system of non–linear differential equations

$$\dot{U} = G^{(+)} - G^{(-)}, \quad \dot{T} = \frac{1}{2}(G^{(+)} + G^{(-)})',$$

$$\dot{G}^{(+)} = 3U^2 - T + \frac{3}{2}U', \quad \dot{G}^{(-)} = T - 3U^2 + \frac{3}{2}U'. \tag{2.29}$$

The flows (2.27) and (2.29) should be regarded as the $l = 2$ analogue of the $r = 1$ and $r = 2$ flows of the ordinary $SL(3)$ KdV hierarchy. These evolutions are Hamiltonian with respect to the commutation relations of the $W_3^{(2)}$ algebra (2.20), i.e. $\dot{X} = \{X\ ,\ \mathcal{H}\}$ with $X = U,\ G^{(\pm)},\ T$, provided that the Hamiltonian functional is

$$\mathcal{H} \equiv \mathcal{H}_4 = \int T \tag{2.30}$$

for the system (2.27) and

$$\mathcal{H} \equiv \mathcal{H}_3 = \int (G^{(+)} + G^{(-)}) \tag{2.31}$$

for the system (2.29). Here and in what follows, we use the notation $\mathcal{H}_s$ to indicate that the conformal dimension of the integrand is $s/2$. We also point out that the equations above admit an alternative Hamiltonian description using the bracket

$$\{U(x)\ ,\ G^{(\pm)}(x')\}_* = \mp\delta(x,x'),$$

$$\{T(x)\ ,\ G^{(\pm)}(x')\}_* = -\frac{3}{2}\delta'(x,x'), \tag{2.32}$$

provided that the Hamiltonian functional is chosen to be

$$\mathcal{H} \equiv \mathcal{H}_7 = -\frac{1}{3}\int \{T(G^{(+)} + G^{(-)}) + \frac{3}{2}U'(G^{(+)} - G^{(-)})\} \tag{2.33}$$

and

$$\mathcal{H} \equiv \mathcal{H}_6 = \frac{1}{3}\int \{G^{(+)}G^{(-)} - G^{(+)^2} - G^{(-)^2} - 3U(T - U^2)\} \tag{2.34}$$

for the evolutions (2.27) and (2.29) respectively.

The two Hamiltonian structures are related to each other in a very simple way. The commutation relations (2.32) arise as a modification to the commutation relations (2.20) under the variable redefinition $G^{(\pm)} \to G^{(\pm)} + \lambda$, with λ normalized for convenience to -1. This situation is analogous to the bi–Hamiltonian structure of the ordinary $SL(3)$ KdV equations and it is reflected in the choice of the constant matrix B. If we use $\mathcal{H}_6$ and $\mathcal{H}_7$ as Hamiltonian functionals for the algebra (2.20) we will obtain new systems of non–linear equations which also exhibit the same bi–Hamiltonian structure. Iterating this procedure we can obtain a whole hierarchy of non–linear flows of increasingly higher order, together with an infinite series of Hamiltonian functionals $\mathcal{H}_s$ [20]. These equations provide the $l = 2$ analogue of the conventional $SL(3)$ KdV hierarchy and they are integrable because all $\mathcal{H}_s$ are in involution

$$\{\mathcal{H}_s\ ,\ \mathcal{H}_{s'}\} = 0 = \{\mathcal{H}_s\ ,\ \mathcal{H}_{s'}\}_*. \tag{2.35}$$

The proof of the integrability is straightforward and relies on the presence of a bi–Hamiltonian structure, as in the ordinary KdV case [12, 13]. The values $s = 2,\ 5,\ 8,\ \cdots$ are excluded because they correspond to the trivial flows $\dot{X} = 0$, in analogy with the $r = 0,\ 3,\ 6,\ \cdots$ of the $l = 1$ $SL(3)$ hierarchy.

The explicit expressions for $\mathcal{H}_s$ and the associated systems of non–linear differential equations can be derived recursively for all values of s, starting from the simplest cases $s = 3$ and $s = 4$. However, there are no closed formulas available at the moment for arbitrary s, since the $l = 2$ analogue of the scalar Lax equation (1.25) is still lacking. For higher values of s the structure of the equations becomes considerably more involved and for this reason we will not present any other examples. We point out nevertheless that by eliminating the variables $G^{(\pm)}$, T from the system (2.29) we obtain the evolution [20]

$$U'' = -\frac{1}{3}(\ddot{U} - 6U^2)^{\cdot\cdot}, \tag{2.36}$$

which is identical to the Boussinesq equation (2.10) up to a suitable rescaling of U. A peculiar feature of eq.(2.36) is that the role of the x and t variables is interchanged and this difference can be attributed to the upper triangular form of the Hamiltonian matrix (2.28). More precisely by comparing the the expressions (2.8) with (2.28) we realize that the role of P and B is interchanged and the the same happens for the strictly upper triangular parts of Q and H. The $x \leftrightarrow t$ duality exhibited by the $r = 2$ and $s = 3$ flows of the $l = 1$ and $l = 2$ $SL(3)$ KdV hierarchies respectively is rather intriguing and it would be interesting to know whether it is a general feature of the theory. The two hierarchies are not equivalent, however, because their W algebras are quite different from each other.

In general it would be interesting to construct in a systematic way $W_N^{(l)}$ algebras with arbitrary values of N and l, together with the corresponding systems of non–linear KdV type equations. Since the conformal dimension of the generating fields is not necessarily integer for $l > 1$, we call the associated KdV flows *fractional*. The embedding of the ordinary KdV equations into sdYM theory that we have presented was quite inspiring in that regard, because it provided the starting point for introducing alternative ansatz for the gauge connections that led to new integrable flows with $l > 1$. Recent work has shown that the systematic classification of all $W_N^{(l)}$ algebras is closely related to the different embeddings of $SL(2)$ into higher rank groups [21, 22] (see also [23] for some related results in Toda field theory).

We also note that the operator content of W algebras with $l > 1$ could be changed by twisting the Virasoro generator,

$$T \to T + \frac{1}{2}U'. \tag{2.37}$$

For $W_3^{(2)}$ for example, the spin of the generators $G^{(+)}$ and $G^{(-)}$ changes to 2 and 1 respectively and alternatively, the commutation relations of $W_3^{(2)}$ describe the algebra of two spin 1 and two spin 2 fields put together in a rather convoluted way.

3. Large N limit considerations: W_∞, KP and sd Einstein equations

In the last lecture we will present some interesting results that arise in the large N limit of $SL(N)$ and W_N algebras. In this limit the structure of W algebras simplifies considerably and becomes linear [24, 25, 26, 27]. The commutation relations of W_∞ describe a central extension of the algebra of all differential operators in one dimension and therefore it is natural to expect that there is some interesting geometry associated with this limit. In fact, as we will demonstrate later, a suitably defined large N limit of the dimensional reduction of sdYM equations that led to integrable systems of KdV type can be also used to describe self–dual Einstein equations in space–times with signature $(2,2)$. Also, it turns out that W_∞ (more precisely $W_{1+\infty}$) provides a Hamiltonian structure for the KP hierarchy, which is a universal KdV system formally related to $SL(\infty)$.

Since W_N algebras are linear only for $N = 2$ and $N \to \infty$, we think that a more systematic study of the problems associated with the second case will be fruitful for future developments in field theory. We will only present a few well–established results in this area, which although are seemingly unrelated, they might be useful for demonstrating the relevance of such large N limit considerations in theoretical physics.

3.1 *The large N limit of W_N algebras*

The non–linearities that arise in the commutation relations of W_N algebras with $N \geq 3$ are characteristic of higher spin theories in any number of space–time dimensions. Inclusion of all higher spin generators, however, usually results into a linear structure, which for W_∞ is nothing else but a linear deformation of the algebra of area preserving diffeomorphisms. Within the standard framework of the Hamiltonian reduction we find that the classical commutation relations of W_N become ill–defined in the limit $N \to \infty$, because most of the structure constants depend implicitly on N and diverge rapidly for large values of N. For example, the central charge of the Virasoro subalgebra is $N^3 - N$ and so in order to make sense of W_∞

we have to introduce appropriate rescaling in the algebra. It turns out that if we define

$$[\ ,\ ]_N \equiv N^3\{\ ,\ \}, \quad U_s \equiv N^{-\frac{1s}{3}} u_s, \tag{3.1}$$

all divergent terms are not present in the large N limit and the classical commutation relations simplify considerably. We may also define new fields W^s using the generating function

$$W(t) = log(1 + U(t)), \tag{3.2}$$

where

$$W(t) = \sum_{s=2}^{\infty} W^s(x)t^s, \quad U(t) = \sum_{s=2}^{\infty} U_s(x)t^s \tag{3.3}$$

and $log(1+a) = a - \frac{a^2}{2} + \frac{a^3}{3} - \cdots$. W^s are primary fields and satisfy the commutation relations

$$[W^s(x)\ ,\ W^{s'}(x')]_\infty = ((s-1)W^{s+s'-2}(x) + (s'-1)W^{s+s'-2}(x'))\delta'(x,x'), \tag{3.4}$$

up to central terms in the Virasoro subalgebra only generated by $W^2(x)$. The proof of this statement follows by a straightforward but lengthy computation [25]. Introducing Fourier modes, the commutation relations of the algebra become

$$[W_m^s\ ,\ W_n^{s'}] = ((s'-1)m - (s-1)n)W_{m+n}^{s+s'-2} + \frac{c}{12}m(m^2-1)\delta_{s,2}\delta_{s',2}\delta_{m+n,0}, \tag{3.5}$$

which for $c = 0$ coincide with the area preserving diffeomorphism algebra of a cylinder.

There is another large N limit of W_N which is more natural for the purposes of quantum field theory [26, 27]. If we assume that the value of the background charge α_0 is not constant but depends on N as $\alpha_0 \sim 1/N$, then the quantum commutation relations of W_∞ become

$$[W_m^s\ ,\ W_n^{s'}] = ((s'-1)m - (s-1)n)W_{m+n}^{s+s'-2} + q^{2(s-2)}c_s(m)\delta_{s,s'}\delta_{m+n,0} +$$

$$+q^2 g_2^{ss'}(m,n)W_{m+n}^{s+s'-4} + q^4 g_4^{ss'}(m,n)W_{m+n}^{s+s'-6} + \cdots, \tag{3.6}$$

where q is a numerical constant which can be normalized to 1, and the coefficients of the central terms are

$$c_s(m) = \frac{c}{2}m(m^2-1)(m^2-4)\cdots(m^2-(s-1)^2)\frac{2^{2(s-3)}s!(s-2)!}{(2s-1)!!(2s-3)!!}. \tag{3.7}$$

Notice that the commutation relations (3.5) follow from (3.6) by the Lie algebra contraction $q \to 0$. Also, for $c = 0$ (but $q \neq 0$), the infinite dimensional algebra (3.6) is isomorphic to the Moyal algebra, i.e. the algebra of all differential operators in one variable (with $s \geq 2$).

The structure constants $g_l^{ss'}(m,n)$ are given by the formula [26]

$$g_l^{ss'}(m,n) = \frac{\varphi_l^{ss'}}{2(l+1)!} N_l^{ss'}(m,n), \tag{3.8}$$

where

$$N_l^{ss'}(m,n) = \sum_{k=0}^{l+1} (-1)^k \binom{l+1}{k} (2s-l-2)_k [2s'-k-2]_{l+1-k} \cdot \\ \cdot [s-1+m]_{l+1-k} [s'-1+n]_k \tag{3.9}$$

with $(a)_k \equiv a(a+1)(a+2)\cdots(a+k-1)$ and $[a]_k \equiv a(a-1)(a-2)\cdots(a-k+1)$. The numerical constants $\varphi_l^{ss'}$ are independent of m and n and their values are given by the following hypergeometric function

$$\varphi_l^{ss'} = \sum_{k\geq 0} \frac{(-\frac{1}{2}-2x)_k (\frac{3}{2}+2x)_k (-\frac{l+1}{2})_k (-\frac{l}{2})_k}{k!(-s+\frac{3}{2})_k (-s'+\frac{3}{2})_k (s+s'-l-\frac{3}{2})_k}, \tag{3.10}$$

with $x = 0$.

The commutation relations (3.6) of W_∞ can be extended consistently with the Jacobi identity to include the value $s = 1$, provided that the parameter x in the hypergeometric function (3.10) is chosen to be $x = -1/2$. Then, the resulting algebra describes an extension of W_∞ with a $U(1)$ field, which is usually denoted by $W_{1+\infty}$ [28]. The latter is a central extension of the algebra of all differential operators in one variable (including the functions) and in that regard, it is more natural to consider. Various field theoretic realizations of W_∞ and $W_{1+\infty}$ have already been constructed in terms of free bosonic and fermionic fields respectively and for further details we refer the reader to the literature [27, 29]. In this lecture we will be mostly concerned with the applications of these infinite dimensional structures in the theory of integrable non-linear differential equations, in particular the KP hierarchy.

3.2 $W_{1+\infty}$ *and the KP hierarchy*

The KP hierarchy is an integrable system of infinitely many non-linear differential equations which can be formulated in a compact way using the pseudo-differential operator

$$Q = \partial_x + q_0 \partial_x^{-1} + q_1 \partial_x^{-2} + q_2 \partial_x^{-3} + \cdots. \tag{3.11}$$

The r-th flow of the KP hierarchy is given by the evolution

$$\frac{\partial Q}{\partial_{t_r}} = [(Q^r)_+ \ , \ Q], \tag{3.12}$$

which generalizes the Lax equation (1.25) to infinitely many variables $\{q_i \ ; \ i = 0, \ 1, \ 2, \ \cdots\}$. Non-trivial flows arise for all values $r \geq 2$, but their explicit form is rather complicated. For pedagogical purposes we only write down the first few members of the $r = 2$ flow:

$$\dot{q}_0 = q_0''' + 2q_1', \tag{3.13a}$$

$$\dot{q}_1 = q_1''' + 2q_2' + 2q_0 q_0', \tag{3.13b}$$

$$\dot{q}_2 = q_2''' + 2q_3' + 4q_1 q_0' - 2q_0 q_0''', \tag{3.13c}$$

etc.

The KP equations (3.12) can be considered as a universal KdV system associated with the large N limit of the $SL(N)$ KdV hierarchies. Notice that the evolutions (1.25) are special cases of (3.12) provided that there exists a positive integer N so that $Q^N = L_N$ is differential. This condition introduces certain relations among the fields $\{q_i\}$ and only the first $N-1$ of them are independent. Then, the system of infinitely many equations (3.12) can be truncated consistently to the flows of the $SL(N)$ KdV hierarchy for all values of r.

The Hamiltonian structure of the KP hierarchy has already been discussed in the literature. Following Watanabe [30] we introduce the following bracket among the fields $\{q_i\}$,

$$\{q_i(x) \ , \ q_j(x')\} = K_{ij}(x)\delta(x,x'), \tag{3.14}$$

where K_{ij} is the linear differential operator

$$K_{ij} = \sum_{l=0}^{i} (-1)^l \binom{i}{l} q_{i+j-l} \partial_x{}^l - \sum_{l=0}^{j} \binom{j}{l} \partial_x{}^l q_{i+j-l}. \tag{3.15}$$

Here, the differential operator $\partial_x{}^l$ acts by Leibnitz's rule on the fields q_{i+j-l}. Then, the KP equations (3.12) can be written in Hamiltonian form $\dot{q}_i = \{q_i \ , \ \mathcal{H}_r\}$, provided that the functionals $\mathcal{H}_r$ are given by the expression

$$\mathcal{H}_r = \frac{1}{r+1} \int res Q^{r+1}, \tag{3.16}$$

where *res* denotes the residue of the pseudo-differential operator Q^{r+1}. Moreover, all $\mathcal{H}_r$ are in involution with respect to the bracket (3.14)

$$\{\mathcal{H}_r\ ,\ \mathcal{H}_{r'}\} = 0, \tag{3.17}$$

which proves that the KP hierarchy is integrable.

It has been demonstrated recently that the algebra (3.14) is isomorphic with $W_{1+\infty}$ with central charge $c = 0$ [31]. The appropriate change of variables that establishes this isomorphism is given by the formula

$$V_s(x) = \frac{(s-1)!}{2^{s-1}(2s-3)!!}\sum_{l=0}^{s-1}\binom{s-1}{l}\binom{2s-l-2}{s-1}q^{(l)}_{s-l-1}(x) \tag{3.18}$$

for all values $s \geq 1$. Then, after introducing Fourier modes, the commutation relations (3.14) transform into (3.6) with $q = -1/4$. Incidentally, we point out that W_∞ with $c = 0$ can also be obtained from the algebra (3.14) using the transformation

$$W^s(x) = \frac{s!}{2^{s-1}(s-1)(2s-3)!!}\sum_{l=0}^{s-2}\binom{s-1}{l}\binom{2s-l-2}{s}q^{(l)}_{s-l-1}(x) \tag{3.19}$$

for all $s \geq 2$. (Here we use the notation V_s and W^s to distinguish between the generators of $W_{1+\infty}$ and W_∞ respectively).

Using eqs.(3.18) and (3.19) we may eliminate the q variables and obtain a direct relation between W^s and V_s. Explicit calculation shows that [32]

$$W^s(x) = V_s(x) + \frac{B(s)}{s-1}\sum_{l=1}^{s-1}\frac{2s-2l-1}{B(s-l)}V^{(l)}_{s-l}(x), \quad s \geq 2 \tag{3.20}$$

where

$$B(s) = q^{s-2}\frac{2^{s-3}(s-1)!}{(2s-3)!!}. \tag{3.21}$$

Equation (3.20) is in fact valid for all values of q and c and provides the explicit relation between the two algebras W_∞ and $W_{1+\infty}$. For non-zero values of c, however, the central charge of W_∞ turns out to be -2 times the central charge of $W_{1+\infty}$. The implications of this twisting in the theory of topological W gravity have already been discussed in the literature [33].

The presence of W_∞ in the KP theory is not very surprising, because these equations provide a universal KdV hierarchy. The peculiar thing of the present formalism, however, is that the infinite dimensional algebra (3.14) does not correspond strictly speaking to the large N limit of W_N algebras. In the present case the

assignment of spin to the fields q_i is different (q_i has spin $i+1$), which explains the presence of the $U(1)$ field q_0 in the theory. Watanabe's commutation relations do not describe the large N limit of what is conventionally called the second Hamiltonian structure of the KdV type equations and therefore it is not surprising that the KP flows (c.f. eqs.(3.13)) appear to be inhomogeneous with respect to the scaling properties of the q variables.

It would be interesting to find another Hamiltonian structure for the KP flows (3.12) which is closer related to the standard description of $SL(N)$ KdV type systems. This might also provide a set of recursive relations for the conserved quantities (3.16) and clarify further the role of W_∞ (or $W_{1+\infty}$) in the KP theory. Another question that arises naturally in this context is to find whether $W_{1+\infty}$ with $c \neq 0$ allows for non-trivial deformations of the KP hierarchy, while preserving integrability.[1]

3.3 *Area preserving diffeomorphisms and sd Einstein equations*

In the remaining part of this lecture we will show that the dimensional reduction of sdYM equations used in the description of KdV flows of $SL(N)$ type can also incorporate the self-dual Einstein equations. It is known that the large N limit of $SL(N)$ can be identified with the algebra of area preserving diffeomorphisms of a 2-dim Riemann surface Σ (e.g., the torus $S^1 \times S^1$) [34]. To simplify our presentation and avoid questions related to the topology of the surface Σ that arises in the limit $N \to \infty$, we start directly with the self-dual Yang-Mills equations in R^4 with signature $(2,2)$ and gauge group $G \simeq SDiff\Sigma$. We will show that in this case, the dimensionally reduced self-dual system (2.3) is equivalent to the second Plebanski equation for self-dual gravity in four space-time dimensions with signature $(2,2)$ [16, 17].

The algebra of area preserving diffeomorphisms $SDiff\Sigma$ is generated by all divergenceless vector fields

$$X_f = f_{,p}\,\partial_q - f_{,q}\,\partial_p, \tag{3.22}$$

where f is an arbitrary function and (p,q) are local canonical coordinates on Σ. Equivalently, this algebra coincides with the Poisson bracket algebra of all functions on Σ. Then, the last equation of the self-dual system (2.3) implies that B and P

[1] I thank Y.-S. Wu for sharing his thoughts on this issue.

can be chosen to be the following (x,t) independent vector fields:

$$B = \partial_q \quad , \quad P = -\partial_p. \tag{3.23}$$

With this choice in mind, eq.(2.3b) states that there exists a function $\Theta(x,t,p,q)$ so that Q and H are represented by the vector fields associated to $-\Theta_{,q}$ and $\Theta_{,p}$ respectively, i.e.,

$$Q = -\Theta_{,pq}\,\partial_q + \Theta_{,qq}\,\partial_p, \tag{3.24a}$$

$$H = \Theta_{,pp}\,\partial_q - \Theta_{,pq}\,\partial_p. \tag{3.24b}$$

Substituting these expressions in eq.(2.3a) we obtain

$$\Theta_{,xp} - \Theta_{,tq} + \Theta_{,pp}\,\Theta_{,qq} - \Theta_{,pq}{}^2 = 0, \tag{3.25}$$

which is Plebanski's second heavenly equation for self-dual gravity on $T^*\Sigma$ with metric [35]

$$ds^2 = 2dxdp - 2dtdq - 2(\Theta_{,qq}\,dx^2 + \Theta_{,pp}\,dt^2 + 2\Theta_{,pq}\,dxdt). \tag{3.26}$$

In deriving Plebanski's equation from (2.3a) we have assumed that the right hand side of eq.(3.25), which turns out to be independent of p and q, is zero. Actually, we could have used any arbitrary function of x and t, say $F(x,t)$, instead of zero, but this dependence can be easily transformed away by redefining Θ as follows:

$$\Theta \to \Theta - p\int F(x,t)dx. \tag{3.27}$$

Since the flat metric $2dxdp - 2dtdq$ is trivially self-dual, eq.(3.25) provides the necessary condition for the Θ-dependent perturbations to preserve self-duality. Solutions to Plebanski equation will give rise to hyper-Kahler metrics on $T^*\Sigma$ which are mathematically allowed because the signature of the space-time is $(2,2)$. (If the signature was $(4,0)$, then only T^*S^2 would allow for such metrics). Although the global aspects of this problem are beyond the scope of the present lectures, we note that they are closely related to the Ooguri-Vafa formulation of $N = 2$ superstring theory in terms of self-dual Einstein equations [36]. It would be certainly interesting to study the quantization of the self-dual perturbations (3.26) within the context of integrable systems. More generally, the hidden relations between $SL(\infty)$ self-dual Yang-Mills and Einstein equations could provide new insights into 4-dim

theories of gravity, in analogy with Polyakov's treatment of induced 2-dim gravity formulated using $SL(2)$ gauge connections.

Apart from the theory of KdV type equations, the notion of W algebras also arises in another class of integrable systems, namely the Toda field equations. These theories can also be obtained by dimensional reduction from sdYM equations [15] and possess conserved currents which satisfy the W algebra commutation relations [37]. It is rather amusing to note that for $SL(N)$ groups in the large N limit, the Toda field equations also have a natural geometric interpretation in the context of self-dual Einstein equations [24, 38]. We think that better understanding of these issues and the limiting procedures which are involved in each case are necessary, before the concept of universal W algebras (like W_∞) can be employed successfully in the study of higher dimensional physics.

It would be interesting to know whether there are any direct relations between the KP hierarchy and self-dual gravity. It has been conjectured by Mason [16] that the self-dual Einstein equations may provide a universal integrable system that encompasses the flows of all $SL(N)$ KdV hierarchies. However, it is not clear at this point how to realize $SL(N)$ groups with $N > 2$ as subgroups of $SDiff\Sigma$. It also evident from the preceding discussion that the embedding of sd Einstein equations into the sdYM theory does not single out one embedding of $SL(2)$ into $SL(\infty) \equiv SDiff\Sigma$ from the others. Since there exist integrable hierarchies for non-standard embeddings too (with $l > 1$), it is unclear which particular KdV type flows (if any) can be obtained by further reduction of the sd Einstein equations. This problem should be studied more carefully in the future.

4. Outlook and future directions

We have presented a detailed account of the theory of W algebras and the associated integrable KdV systems of non-linear differential equations, from the point of view of gauge symmetries. The hidden relations between extended conformal algebras and gauge symmetries, as well as the zero curvature formulation of all KdV type systems *à la* Drinfeld and Sokolov, has been explained from a higher dimensional point of view, using dimensional reductions of sdYM equations in space-times with signature (2, 2). This description also enabled us to introduce generalized ansatz associated with non-standard embeddings of $SL(2)$ into higher rank groups and

construct new classes of integrable KdV hierarchies. Since we have sketched the fractional generalization of $SL(N)$ KdV flows only for $N = 3$ and $l = 2$, the general features of these equations should be investigated further in the future in the framework of $W_N^{(l)}$ algebras, using the method suggested in ref.[22].

It is rather tempting to speculate that there exist random matrix models of 2–dim quantum gravity, whose properties in the double scaling limit are governed by $W_N^{(l)}$ algebra constraints and fractional KdV hierarchies. The couplings of these models should not be arbitrary, but fall into an algebraic pattern that depends on the specific embedding of $SL(2)$ into higher rank groups. It would be also interesting to understand in this context the non–abelian generalization of the Toda field equations proposed by Leznov and Saveliev [23], as well as the different large N limits that emerge from these theories.

In conclusion, we note that the systematic embedding of various 2–dim integrable systems into the sdYM equations could be advantageous for addressing the problem of their quantization in terms of (hidden) gauge symmetries. Since the self–dual sector of Einstein's gravity also admits a similar 2–dim description, though with infinite dimensional structure group, it is very important to develop a quatization scheme for integrable systems based on self–dual Yang–Mills fields. In this "universal" framework one may also try to understand the notion of quantum groups and more generally the quantum inverse scattering method directly in terms of 4–dim self–dual Yang–Mills quantum fields.

Acknowledgements: I would like to thank Didier Depireux and Elias Kiritsis for fruitful collaboration in various stages of this research program.

5. References

1. A.M. Polyakov, Mod. Phys. Lett. A2 (1987) 893; V.G. Knizhnik, A.M. Polyakov and A.B. Zamolodchikov, Mod. Phys. Lett. A3 (1988) 819.

2. E. Witten, Phil. Trans. Roy. Soc. Lond. A329 (1989) 349.

3. V.G. Drinfeld and V.V. Sokolov, Sov. Math. Dokl. 23 (1981) 457; Jour. Sov. Math. 30 (1985) 1975.

4. I. Bakas, Phys. Lett. B213 (1988) 313; Phys. Lett. B219 (1989) 283; Commun. Math. Phys. 123 (1989) 627; M. Bershadsky and H. Ooguri, Commun. Math. Phys. 126 (1989) 49; J. Balog, L. Feher, P. Forgacs, L. O'Raifeartaigh and A. Wipf, Ann. Phys. (NY) 203 (1990) 76.

5. P. Mathieu, Phys. Lett. B208 (1988) 101; Y. Matsuo, Phys. Lett. B227 (1989) 209; D.J. Smit, Commun. Math. Phys. 128 (1990) 1.

6. A.B. Zamolodchikov, Theor. Math. Phys. 65 (1985) 1205; A.B. Zamolodchikov and V.A. Fateev, Nucl. Phys. B280[FS18] (1987) 644.

7. V.A. Fateev and S.L. Lykyanov, Int. Jour. Mod. Phys. A3 (1988) 507.

8. F.A. Bais, P. Bouwknegt, M. Surridge and K. Schoutens, Nucl. Phys. B304 (1988) 348; Nucl. Phys. B304 (1988) 371.

9. V.A. Fateev and S.L. Lykyanov, Kiev preprints, ITP–88–74, 75, 76 (1988).

10. G. Segal and G. Wilson, Publ. Math. IHES 61 (1985) 5.

11. I.M. Gelfand and L.A. Dickey, Funct. Anal. Appl. 10 (1976) 259; I.M. Gelfand and I. Dorfman, Funct. Anal. Appl. 15 (1981) 173; Yu.I. Manin, Jour. Sov. Math. 11 (1979) 1.

12. L.D. Faddeev and L.A. Takhtajan, *"Hamiltonian Methods in the Theory of Solitons,"* Springer Series in Soviet Mathematics, Springer–Verlag, (1987); S.P. Novikov, S.V. Manakov, L.P. Pitaevskii and V.E. Zakharov, *"Theory of Solitons,"* Consultants Bureau, (1984).

13. C. Rebbi and G. Soliani, *"Solitons and Particles,"* World Scientific, (1984); A. Das, *"Integrable Models,"* World Scientific, (1989).

14. R.S. Ward, Phil. Trans. Roy. Soc. Lond. A315 (1985) 451; M.F. Atiyah, unpublished; N.J. Hitchin, Proc. Lond. Math. Soc. 55 (1987) 59.

15. E. Witten, Phys. Rev. Lett. 38 (1977) 121; A.N. Leznov and M.V. Saveliev, Lett. Math. Phys. 3 (1979) 489; Commun. Math. Phys. 74 (1980) 111.

16. L.J. Mason and G.A.J. Sparling, Phys. Lett. A137 (1989) 29; L.J. Mason, Twist. Newslett. 30 (1990) 14.

17. I. Bakas and D.A. Depireux, *"The Origin of Gauge Symmetries in Integrable Systems of KdV Type,"* to appear in Int. Jour. Mod. Phys. A, (1991).

18. I. Bakas and D.A. Depireux, Mod. Phys. Lett. A6 (1991) 399.

19. A.M. Polyakov, Int. Jour. Mod. Phys. A5 (1990) 833; M. Bershadsky, *"Conformal Field Theories via Hamiltonian Reduction,"* Princeton preprint, IASSNS-HEP-90/44, (1990).

20. I. Bakas and D.A. Depireux, *"A Fractional KdV Hierarchy,"* to appear in Mod. Phys. Lett. A, (1991).

21. F.A. Bais, T. Tjin and P. van Driel, *"Covariantly Coupled Chiral Algebras,"* Amsterdam preprint, ITFA-90-31, (1990).

22. M.F. de Groot, T.J. Hollowood and L.J. Miramontes, *"Generalized Drinfeld-Sokolov Hierarchies,"* to appear as Princeton University preprint, (1991).

23. A.N. Leznov and M.V. Saveliev, Commun. Math. Phys. 89 (1983) 59; M.V. Saveliev, Mod. Phys. Lett. A5 (1990) 2223.

24. I. Bakas, Phys. Lett. B228 (1989) 57; *"Area Preserving Diffeomorphisms and Higher Spin Fields in Two Dimensions,"* in "Supermembranes and Physics in 2+1 Dimensions," edited by M. Duff et al., World Scientific (1990) 352.

25. I. Bakas, Commun. Math. Phys. 134 (1990) 487.

26. C. Pope, L. Romans and X. Shen, Phys. Lett. B236 (1990) 173; Nucl. Phys. B339 (1990) 191.

27. I. Bakas and E. Kiritsis, Nucl. Phys. B343 (1990) 185; Mod. Phys. Lett. A5 (1990) 2039.

28. C. Pope, L. Romans and X. Shen, Phys. Lett. B242 (1990) 401.

29. E. Bergshoeff, C. Pope, L. Romans, E. Sezgin and X. Shen, Phys. Lett. B245 (1990) 447; D.A. Depireux, Phys. Lett. B252 (1990) 586.

30. Y. Watanabe, Lett. Math. Phys. 7 (1983) 99; Ann. di Mat. Pura Appl. 137 (1984) 77.

31. F. Yu and Y.-S. Wu, *"Hamiltonian Structure, (Anti) Self-Adjoint Flows in KP Hierarchy and the $W_{1+\infty}$ and W_∞ Algebras,"* University of Utah preprint, UU-HEP-91/01, (1991); K. Yamagishi, *"A Hamiltonian Structure of KP Hierarchy, $W_{1+\infty}$ Algebra and Self-Dual Gravity,"* Livermore Laboratory preprint, LLNL-HEP-07/91, (1991).

32. I. Bakas and E. Kiritsis, unpublished.

33. C. Pope, L. Romans, E. Sezgin and X. Shen, *"W Topological Matter and Gravity,"* Texas AM University preprint, CTP TAMU-87/90, (1990).

34. J. Hoppe, Phys. Lett. B215 (1988) 706; E. Floratos, J. Iliopoulos and G. Tiktopoulos, Phys. Lett. B217 (1989) 285; D.B. Farlie and C. K. Zachos, Phys. Lett. B224 (1989) 101; M. Saveliev and A. Vershik, Commun. Math. Phys. 126 (1989) 387.

35. J. Plebanski, Jour. Math. Phys. 16 (1975) 2395; C. Boyer and J. Plebanski, Jour. Math. Phys. 26 (1985) 229.

36. H. Ooguri and C. Vafa, Mod. Phys. Lett. A5 (1990) 1389; *"Geometry of $N = 2$ Strings,"* Harvard University preprint, HUTP-91/A003, (1991).

37. A. Bilal and J.-L. Gervais, Phys. Lett. B206 (1988) 412; Nucl. Phys. B314 (1989) 646; Nucl. Phys. B318 (1989) 579.

38. Q-Han Park, Phys. Lett. B236 (1990) 429; Phys. Lett. B238 (1990) 287; R.S. Ward, Class. Quant. Grav. 7 (1990) L217.

Integrable Hierarchies, Virasoro Algebra, and Virasoro Constrained Hierarchies

A.M.Semikhatov

Theory Division, P.N.Lebedev Physics Institute

53 Leninsky prosp., Moscow SU-117924 USSR

ABSTRACT

The action of the Virasoro algebra on integrable hierarchies of non-linear equations and on related objects ('Schrödinger' differential operators) is investigated. The method consists in pushing forward the Virasoro action to the wave function of a hierarchy, and then reconstructing its action on the dressing and Lax operators. This formulation allows one to observe a number of suggestive similarities between the structures involved in the description of the Virasoro algebra on the hierarchies and the structure of conformal field theory on the world-sheet. This includes, in particular, an "off-shell" hierarchy version of operator products and of the Cauchy kernel. In relation to matrix models, which have been observed to be effectively described by Virasoro-constrained hierarchies, I propose to define general Virasoro-constrained hierarchies also in terms of dressing operators, by certain equations which carry the information of the hierarchy and the Virasoro algebra simultaneously and which suggest an interpretation as operator versions of the string/puncture equations in topological theories. The formulation in terms of dressing operators facilitates the analysis of the continuum limit of discrete (i.e., lattice) hierarchies with the Virasoro constraints. In particular, the KP hierarchy subjected to the Virasoro constraints is recovered as a scaling limit of the Virasoro-constrained Toda hierarchy.

1.Introduction.

These lectures are devoted to various aspects of the occurrence of the Virasoro algebra in the theory of integrable hierarchies [1][2][3], and to several related issues.

That the Virasoro algebra must have something to do with integrable systems, may be understood in different ways, but the most geometric one is to recall the Krichever construction [4] which says that integrable hierarchies are solved by Riemann surfaces (equipped with some extra data). On the other hand, the Virasoro algebra does act on Riemann surfaces. Consider, for instance, the KP hierarchy [1] or its various reductions (generalized (m)KdV's [3]) described in terms of pseudodifferential (ψDiff) operators. Then the Beilinson-Manin-Schechtman action [5] of the Virasoro algebra on the Riemann surface translates to an action on ψDiff operators [6]:

$$\begin{array}{ccc} \text{Riemann Surface} & \xrightarrow{\text{Krichever}} & \psi\textit{Diff operator} \\ \Big\downarrow \text{BMS [5]} & & \Big\downarrow [6] \\ \text{Riemann Surface} & \xrightarrow{\text{Krichever}} & \psi\textit{Diff operator} \end{array} \qquad (1)$$

This is an on-shell diagram, since the horizontal arrows determine *solutions* of the hierarchy, and therefore the right vertical arrow takes solutions into solutions. We will see, however, that it is possible to lift it off the solution locus, thus obtaining an "off-shell" Virasoro action on ψDiff operators.

Considering other hierarchies, the ψDiff operators get replaced with elements from other infinite-dimensional algebras (that of $\infty\times\infty$ matrices, for instance). Such Lie algebras $\mathfrak{g}$ play the role of a *phase space* of a hierarchy since it is in terms of these that integrable hierarchies usually admit a hamiltonian formulation (as, e.g., the Gelfand-Dikii hamiltonian structures on (pseudo)differential operators). Although we will not use the hamiltonian formulation, we will speak of the Lie algebra $\mathfrak{g}$ providing a chosen hierarchy with the hamiltonian structure, as of the phase space. Among elements of $\mathfrak{g}$ there are Lax operators for the hierarchy, since a peculiarity of the hamiltonian formulation of integrable hierarchies is that it renders just the Lax equations hamiltonian.

The description of a hierarchy in terms of flows on the phase space is

complementary (though equivalent[1]) to and in many respects simpler than the Hirota bilinear relations for the tau function. Simplifications occur for the phase space description because the way from the tau function to the phase space goes through the *wave function*, which is essentially a linear object. Knowing the wave function allows one to reconstruct the *dressing operators* which are in turn used in the construction of the Lax operators. That is, the Virasoro action can be pushed forward along the arrows of the diagram

$$\text{tau function} \longmapsto \text{wave function} \longmapsto \text{dressing operator} \longrightarrow \text{Lax operator} \qquad (2)$$

Recall that *solutions* to the hierarchies are given by Riemann surfaces. Then we get the "on-shell" chain of mappings

$$\begin{array}{ccccccc} \text{Riemann} & & \text{tau} & & \text{wave} & & \text{dressing} \\ \text{surface} & \longmapsto & \text{function} & \longmapsto & \text{function} & \longmapsto & \text{operator} \end{array} \longrightarrow \begin{array}{c} \text{Lax} \\ \text{operator} \end{array} \qquad (3)$$

This diagram embraces Riemann surfaces of different genera on equal footing, and it is this old argument that underlies attempts to reformulate conformal field theory entirely in terms of an integrable hierarchy. As an example of a thing that can be 'lifted off' Riemann surfaces, consider the coordinate present in the Krichever quintuple. While in the field–theoretic description this is just the Euclideanized world-sheet coordinate, on which operator insertions depend, from the integrable hierarchy standpoint it is nothing but an eigenvalue of a Lax operator:

$$Q\psi(t,z) = z\psi(t,z), \qquad Q \in \mathfrak{g} \qquad (4)$$

In that sense, the world-sheet is 'created' by non–linear differential equations. What about field operators and *dynamics* of a conformal theory? We would like to carry these over from the RHS of eq.(4) to its LHS and obtain a formulation based on replacing the coordinate z with the operator Q from the LHS of eq.(4), with all the two–dimensional fields depending on z, transformed into some geometrical

[1]Equivalence of the formulations based on the tau function and on the wave function, is a profound result in the KP theory, as well as Toda etc. However, for the super KP, for instance, the situation is more complicated.

objects associated to the hierarchy.

The fundamental dynamical world-sheet quantity is the energy-momentum tensor. Thus we come back to the Virasoro algebra and the problem of how to translate its action from Riemann surfaces to integrable hierarchies. Our strategy consists just in the use of eq.(3), and the result can be stated as follows:

There exist on the phase space a family of vector fields $\hat{\mathfrak{T}}(u)$, depending on a formal parameter u, which

(i) satisfy the centerless Virasoro commutation relations,

$$[[\hat{\mathfrak{T}}(u),\hat{\mathfrak{T}}(v)]] = \frac{1}{v}\frac{\partial}{\partial v}\,\delta(u,v)\cdot\hat{\mathfrak{T}}(u) - \frac{1}{u}\frac{\partial}{\partial u}\,\delta(u,v)\cdot\hat{\mathfrak{T}}(v) \qquad (5)$$

where $\delta(u,v) = \sum_{n\in\mathbb{Z}} \frac{u^n}{v^n}$ is a formal delta function (and $[[\ ,\ ]]$ is a commutator of vector fields on $\mathfrak{g}$, *not* just of *elements* of $\mathfrak{g}$);

(ii) are tangent to the space of *solutions* of the hierarchy;

(iii) by restriction to the Krichever locus, make the diagram commutative:

$$\begin{array}{ccc} \text{Riemann Surface} & \longrightarrow & \text{Lax operator} \\ \Big\downarrow T(u) & & \Big\downarrow \hat{\mathfrak{T}}(u) \\ \text{Riemann Surface} & \longrightarrow & \text{Lax operator} \end{array} \qquad (6)$$

where by "Riemann Surface" we actually mean a Krichever quintuple, and the left vertical arrow is given by the action, constructed in [5], of the vector field $\Sigma_{n\in Z} u^{-n-2} z^{n+1}\frac{\partial}{\partial z}$ (recall that u is a formal parameter).

As the conformal theory to start with, we will consider the bc theory of spin J [15]. Then, as we will see, the structure of $\hat{\mathfrak{T}}(u)$ repeats that of the energy–momentum tensor of the bc theory $(1-J)\partial b\cdot c - Jb\cdot\partial c$, although the operator 'analogues' of b and c do not, of course, involve any fermions (the trick which accounts for the "wrong statistics", is with the analogue of ∂). This result appears quite general and stems from general properties of integrable hierarchies or, more precisely, of their r-matrix [7] structure. Recall that in order to specify a hierarchy one fixes a Lie algebra $\mathfrak{g}$ (for instance that of $\infty\times\infty$ matrices (Toda) or ψDiff operators (KP etc.)) equipped with a decomposition into a sum of two sub-

algebras, $\mathfrak{g}=\mathfrak{a}+\mathfrak{b}$, which is essentially equivalent to fixing an r-matrix [7][2]. The form of the Virasoro constraints on the dressing operators follows by evaluating all the time derivatives acting on the wave function, in terms of generators of the flows on the phase space. Further, proceeding along the arrows of the diagram (3) we end up with an inner derivation $\mathrm{ad}\mathfrak{T}$ of the Lie algebra $\mathfrak{g}$, in which $\mathfrak{T}$ appears to have a universal form in terms of the r-matrix (and a character of $\mathfrak{a}$, see below).

Another context in which the Virasoro action on integrable hierarchies is important, is that of matrix models [8][9][10][11][12]. The partition function of a matrix model is the tau function of some integrable hierarchy (it seems that all possible reductions of the two-dimensional Toda lattice hierarchy comprise a wide enough class) [11][12]:

$$\tau(t) = \int (dM)\, \exp\left[- \mathrm{Tr} \sum_{r\geq 1} t_r M^r \right] \qquad (7)$$

with the integral going over (say, hermitian) matrices. Now, redefining M as $M \longmapsto M + \epsilon M^{n+1}$, $n\geq -1$, one readily discovers that the tau function is annihilated by the $(n\geq -1)$-Virasoro generators, of the type of those displayed in eqs.(2.16) below,

$$\mathscr{L}_n \tau(t) = 0, \qquad (8)$$

(These constraints, presumably, *characterize* those solutions to the hierarchy that come from matrix models.) Again, translating the constraints into the language of dressing operators, leads to constraints which are essentially universal for all integrable hierarchies admitting an r-matrix formulation [13]. For the Toda hierarchy with which, in particular, we will work below, the Virasoro constraints on the dressing operators read

[2]We are not going here into subtleties related to the fact that the most natural interpretation of the flows is in terms of the *co*adjoint action of one of the two subalgebras associated with the chosen r–matrix, see [7]. In the end, in any case, one has only commutators (i.e., the adjoint action), since the conventional Lax representations encountered in applications, require an identification of $\mathfrak{g}$ with its dual.

$$\mathfrak{L}_n \equiv \left[W\{[J(n+1) + \hat{\mathrm{p}}]\Lambda^n + \sum_{r\geq 1} r t_r \Lambda^{r+n}\} W^{-1} \right]_- = 0, \qquad (9)$$

where $n\geq 0$ or -1. The notations pertaining to the Toda case will be explained in the appropriate place, but before that we would like to point out that up to minor details and redefinitions, the Virasoro constraints are of essentially this form in any hierarchy. That is, the LHS of (9) is an element of the Lie algebra $\mathfrak{g}$, and $(\ldots)_-$ is the projection onto one of the two subalgebras entering the decomposition $\mathfrak{g} = \mathfrak{a}+\mathfrak{b}$ [7]. Further, $W \ldots W^{-1}$ is the adjoint action by the corresponding dressing operator, and Λ is the character (one–point coadjoint orbit) of one of these subalgebras. And finally, $\hat{\mathrm{p}}$ is chosen so that $\hat{\mathrm{p}}$ and Λ form a "Weyl" pair ($\Lambda \hat{\mathrm{p}} \Lambda^{-1} = \hat{\mathrm{p}}+1$). These general properties alone are sufficient to show that the corresponding vector fields $\hat{\mathfrak{L}}_n$ do form a Virasoro algebra.

The outline of the subsequent sections is the following. We start in Sec.2 with the Virasoro action on the KP hierarchy and after fixing our notations we review the results of ref.[6] on the Virasoro action. Analogies with the worldsheet formulation, suggested by the form of the Virasoro action, are further discussed in Sec.3. Although this section contains no rigid conclusions, it seems that structural similarities between the worldsheet theory and its hierarchy version deserve being pointed out.

It is stressed several times throughout the paper, and it applies in full to Sec.3, that inasmuch as the Virasoro action is concerned, the choice of a specific hierarchy is almost irrelevant. Another important example, besides the KP hierarchy, is provided by the (two-dimensional, to start with) Toda lattice hierarchy [2]. We introduce the necessary notation in Sec.4, and in Sec.5 proceed with the derivation of the Virasoro action on the Toda phase space (furnished by $\infty\times\infty$ matrices). The result reproduces the structures of Sec.2 in a different guise, and therefore Sec.3 can be copied with minor changes for the Toda case. (We spare this dull repetition, of course).

There is, however, another aspect which makes just the Toda hierarchy particularly important: associated to this hierarchy are second-order "two-dimensional" operators which are known [20] to "parametrize" what is a nice candidate for the universal moduli space. We recast in Sec.6 the Virasoro action into the language of objects intrinsically connected with these differential

operators.

In the second half of the paper we consider the application of the results on the Virasoro action to the problems that have appeared in the study of matrix models, where one encounters integrable hierarchies constrained by Virasoro highest-weight conditions. In Sec.7 we show that the 'bosonized' representation of the Virasoro constraints obtained in Sec.3 allows one to reinterpret them as a sort of recursion/puncture equation [24][25] for what presumably is a topological field theory. We actually work with the KP hierarchy, but, again, similar results hold for *any* hierarchy, thus suggesting a general correspondence of integrable hierarchies with topological theories in two dimensions. Further, in Sec.8 we establish the relation between the discrete (*lattice*) Virasoro-constrained hierarchies, such as Toda, and the continuum ones (such as KP). They are related by a scaling, which is easily performed at the level of dressing operators and which transforms the "discrete" Virasoro constraints to the continuum ones.

2. Virasoro action on the KP hierarchy.

We work with the KP hierarchy described in terms of ψDiff operators [1]. (Note in passing that there exists another description, which treats KP rather like an $n=\infty$–KdV hierarchy, i.e., in terms of $\infty\times\infty$ matrices [14].)

Recall that the KP hierarchy can be formulated as an infinite set of mutually commuting evolution equations

$$\frac{\partial K}{\partial t_r} = -(KD^rK^{-1})_-, \quad r\geq 1 \tag{2.1}$$

on the coefficients $w_n(x\equiv t_1,t_2,t_3,\ldots)$ of a ψDiff operator (more precisely, a ψDiff symbol) K of the form

$$K = 1 + \sum_{n\geq 1} w_n D^{-n} \tag{2.2}$$

Here and in the sequel,

$$D = \frac{\partial}{\partial x} \tag{2.3}$$

and we identify $x\equiv t_1$ according to eq.(2.1) with $r=1$:

$$\frac{\partial K}{\partial t_1} = -(KDK^{-1})K + (KDK^{-1})_+K = -KD + DK = \frac{\partial K}{\partial x} \tag{2.4}$$

That the flows defined by (2.1) commute, follows from the (modified) classical Yang–Baxter equation satisfied by the 'r–matrix'

$$r^-: \psi Diff \longrightarrow \psi Diff: A \longmapsto A_- \tag{2.5}$$

Introduce a 'matrix model potential'

$$\xi(t,z) = \sum_{r\geq 1} t_r z^r \tag{2.6}$$

(from now on t in the argument of a function denotes the set $(t_1,t_2,\ldots)$). The wave function and the adjoint wave function are defined by

$$\psi(t,z) = Ke^{\xi(t,z)}, \quad \psi^*(t,z) = K^{*-1}e^{-\xi(t,z)} \tag{2.7}$$

where K^* is a formal adjoint to K. Obviously, ψ satisfies eq.(4) with

$$Q = KDK^{-1} \tag{2.8}$$

We also define w and w^* by

$$\psi(t,z) = e^{\xi(t,z)}w(t,z),\quad \psi^*(t,z) = e^{-\xi(t,z)}w^*(t,z) \tag{2.9}$$

Along with ψ and ψ^*, these will be also referred to as wave functions. Clearly (cf. eq.(2.2)),

$$w(t,z) = 1 + \sum_{n\geq 1} w_n(t)z^{-n}.$$

A basic property of the wave function is its relation to the tau function:

$$\psi(t,z) = \frac{\Gamma(t,z)\tau(t)}{\tau(t)} = \frac{\tau(t-[z^{-1}])}{\tau(t)}e^{\xi(t,z)}, \tag{2.10}$$

$$\psi^*(t,z) = \frac{\Gamma^*(t,z)\tau(t)}{\tau(t)} = \frac{\tau(t+[z^{-1}])}{\tau(t)}e^{-\xi(t,z)} \tag{2.11}$$

where

$$\Gamma(t,z) = e^{\xi(t,z)}\, e^{-\sum_r \frac{1}{r} z^{-r}\frac{\partial}{\partial t_r}},\qquad \Gamma^*(t,z) = e^{-\xi(t,z)}\, e^{\sum_r \frac{1}{r} z^{-r}\frac{\partial}{\partial t_r}} \tag{2.12}$$

and, accordingly,

$$t \pm [z^{-1}] = (t_1 \pm z^{-1},\ t_2 \pm \tfrac{1}{2}z^{-2},\ t_3 \pm \tfrac{1}{3}z^{-3},\ \ldots) \tag{2.13}$$

We actually use the following combination of the mappings (3) and (4):

$$\begin{array}{ccccc} \text{Riemann Surface} & \longmapsto & \text{tau function} & \longmapsto & \psi Diff \text{ operator} \\ \Big\downarrow T(u) & & \Big\downarrow \mathcal{T}(u) & & \Big\downarrow \mathfrak{T}(u) \\ \text{Riemann Surface} & \longmapsto & \text{tau function} & \longmapsto & \psi Diff \text{ operator} \end{array} \tag{2.14}$$

(the diagram is projectively commutative). By "Riemann Surface" we mean a Krichever quintuple in which the line bundle is the bundle of J–differentials and the trivialization is determined by the parameter a_0 (see [6], where $a_0 = N + \frac{1}{2}$ and $J \leftrightarrow (1-J)$). Then

$$\mathcal{T}(u) = \sum_{p\in\mathbb{Z}} u^{-p-2}\,\mathcal{L}_p \tag{2.15}$$

with the usual expressions for the Virasoro generators:

$$\mathcal{L}_{p>0} = \frac{1}{2}\sum_{k=1}^{p-1}\frac{\partial}{\partial t_{p-k}}\frac{\partial}{\partial t_k} + \sum_{k\geq 1} kt_k\frac{\partial}{\partial t_{p+k}} + (a_0 + (J-\tfrac{1}{2})p)\frac{\partial}{\partial t_p}$$

$$\mathcal{L}_0 = \sum_{k\geq 1} kt_k\frac{\partial}{\partial t_k} + \tfrac{1}{2}a_0^2 - \tfrac{1}{24} \tag{2.16}$$

$$\mathcal{L}_{p<0} = \sum_{k\geq 1}(k-p)t_{k-p}\frac{\partial}{\partial t_k} + \frac{1}{2}\sum_{k=1}^{-p-1} k(-p-k)t_k t_{-p-k} + (a_0 + (J-\tfrac{1}{2})p)(-p)t_{-p}$$

These form the algebra

$$[\mathcal{L}_p, \mathcal{L}_q] = (p-q)\,\mathcal{L}_{p+q} + \delta_{p+q,0}(-p^3)(J^2 - J + \tfrac{1}{6}) \tag{2.17}$$

(Shifting $\mathcal{L}_0$ as $\mathcal{L}_0 \longmapsto \mathcal{L}_0 - \frac{1}{2}(J^2 - J + \frac{1}{6})$ we recover in (2.17) the "standard" central term $-\delta_{p+q,0}(p^3-p)(J^2 - J + \frac{1}{6})$).

Now we deform the tau function by means of the Virasoro generators as

$$\tau(t) \longmapsto \tau(t) + \delta\tau(t) = \tau(t) + \mathcal{T}(u)\tau(t) \tag{2.18}$$

and unravel the equations (2.7–10) with the new tau function. In this way we end up with a new dressing operator $K - \mathfrak{T}(u)K$, where [6] (see also [16] for J=1)

$$\mathfrak{T}(u) = (1-J)\frac{\partial\psi(t,u)}{\partial u}\circ D^{-1}\circ\psi^*(t,u) - J\psi(t,u)\circ D^{-1}\circ\frac{\partial\psi^*(t,u)}{\partial u} \tag{2.19}$$

(and we have set $a_0 = -\frac{1}{2}$ for simplicity). When moved to the right, D^{-1} gives rise to an infinite tail of negative powers of the differentiation via

$$D^{-1}\circ f(x) = f(x)D^{-1} + \sum_{n=1}^{\infty}(-1)^n\partial^n f\, D^{-n-1} \tag{2.20}$$

When acting with $\mathfrak{T}(u)$ on functions of K other than linear ones, one should view $\mathfrak{T}(u)K$ as components X^i of a vector field $X^i\frac{\partial}{\partial x_i}$. That is, we define the vector field as a derivation of the ring of functions of K,

$$\hat{\mathfrak{T}}(u).F(K) = \frac{d}{d\epsilon}F(K + \epsilon\mathfrak{T}(u)K)\Big|_{\epsilon=0} \qquad (2.21)$$

ψ and ψ^* in this formula are, of course, functions of K, see eqs.(2.7).

At this moment we can forget that we have been dealing with a *solution*, and view the above derivation of $\mathfrak{T}(u)$ merely as a way to guess the form of eq.(2.19) (and to prove the assertion (iii) of the Introduction). Eq.(5) follows by a direct calculation [6], and so does (ii). It is valid irrespectively of whether or not K corresponds to a smooth curve, or even is a solution to the KP hierarchy at all.

From the variation $\delta K = -\mathfrak{T}(u)K$, we readily get the variation of the wave function: eqs.(2.7) imply

$$\delta\psi(t,z) \equiv \hat{\mathfrak{T}}(u)\psi(t,z) = -(1-J)\omega(z,u,t)\frac{\partial\psi(t,u)}{\partial u} + J\frac{\partial\omega(z,u,t)}{\partial u}\psi(t,u) \qquad (2.22)$$

$$\omega(z,u;t) = \partial^{-1}(\psi(t,z)\psi^*(t,u)) = \int^z \psi(x',t_{\geq 2},z)\psi^*(x',t_{\geq 2},u)dx' \qquad (2.23)$$

This formula was re-derived, in a more geometrical context, in ref.[17]. (See also ref.[18] where the action of a half of the Virasoro algebra was found for the first time.)

3.Comments on "bosonization" on the hierarchy.

In this section we will discuss several issues related to the above translation of the *bc* theory to integrable equations.

Eq.(2.19) bears a striking similarity to the energy–momentum tensor of the spin-J *bc* system

$$-Jb\cdot\partial c + (1-J)\partial b\cdot c \qquad (3.1)$$

The fact that, unlike (3.1), the RHS of (2.19) is composed out of bosonic quantities, is "compensated for" by D^{-1} between the two wave functions.

Moreover, eq.(2.22) can be viewed as an abstract version of the usual local operator product $T(u)b(z) = \frac{Jb(u)}{(u-z)^2} + \frac{(1-J)\partial b(u)}{(u-z)}$ because $\omega(z,u;t)$ is an "off-shell" analogue of the Cauchy kernel: indeed, "*on shell*", i.e. for a given algebro–geometric *solution*[3], when z and u become genuine coordinates on a Riemann surface, one has [17]

$$\omega(z,u;t) = \frac{1}{u-z} + \textit{regular at } z=u \qquad (3.2)$$

Another important ingredient of the *bc* theory is the 'ghost' current $j = -bc$. It follows from the formulae of [6] that its KP version reads

$$j(u) = \psi(t,u)\circ D^{-1}\circ\psi^*(t,u) \qquad (3.3)$$

This can also be written as (with Q defined in eq.(2.8))

$$j(u) = (K\frac{1}{u}\delta(u,D)K^{-1})_- = \frac{1}{u}\delta(u,Q)_- \qquad (3.4)$$

Observe that the ψDiff analogue of the fusion $j(u)b(z) = \frac{b(u)}{u-z}$, that is,

$$\hat{j}(u)\psi(t,z) = \omega(z,u)\psi(t,u) \qquad (3.5)$$

is an immediate consequence of eq.(3.3).

---

[3]In the classical papers on algebro-geometric solutions, [26][30] and refernces therein, the algebro-geometric tau function is a *function*. Its analogues which are sections of the line bundle of J (or $1-J$ for ψ^*) differentials, as well as the corresponding tau functions, appeared in a number of papers more recently. See for example an operator *derivation* of these in [19].

These analogies raise the question of whether a ψDiff analogue of *bosonization* through a *scalar field* in ψDiff terms exists. Indeed, while eq.(2.19) calls for "fermionic" analogies, at the same time it can be *identically* rewritten as

$$\mathfrak{T}(u) = \tfrac{1}{2}(KP\tfrac{1}{u}\delta(u,D)K^{-1} + K\tfrac{1}{u}\delta(u,D)PK^{-1})_- - \tfrac{q}{2}\tfrac{\partial}{\partial u}\left[\tfrac{1}{u}\delta(u,Q)_-\right], \quad q = 2J-1 \quad (3.6)$$

with the operator

$$P = \sum_{r\geq 1} r t_r D^{r-1} \quad (3.7)$$

representing the derivative w.r.t. the spectral parameter (and hence being the main hero of Douglas' equation). We will call eq.(3.6) the 'bosonized' form of the energy-momentum tensor on the hierarchy.

The 'bosonized' form is very convenient for calculations of commutators of $\hat{\mathfrak{T}}(u)$ with other vector fields. Indeed, for any two vector fields (as in eq.(2.21)), $\hat{a}$ and $\hat{b}$, of the form

$$\hat{a} = aK\tfrac{\delta}{\delta K},\ \hat{b} = bK\tfrac{\delta}{\delta K} \text{ with } a = (KAK^{-1})_-,\ b=(KBK^{-1})_- \quad (3.8)$$

we evaluate

$$\hat{a}\hat{b}K = \hat{a}\left[(KBK^{-1})_-K\right]$$

$$= ((KAK^{-1})_-KBK^{-1})_-K - (KBK^{-1}(KAK^{-1})_-)_-K + (KBK^{-1})_-(KAK^{-1})_-K$$

$$= ((KAK^{-1})_-KBK^{-1})_-K - (KBAK^{-1})_-K + ((KBK^{-1})_-KAK^{-1})_-K$$

whence a useful expression follows,

$$[[\hat{a},\hat{b}]]K = (K[A,B]K^{-1})_-K \quad (3.9)$$

Using this to commute the energy–momentum tensor $\hat{\mathfrak{T}}(u)$ and the current $\hat{\jmath}(v)$ amounts to taking, respectively,

$$A = \left[-J\tfrac{1}{u}\tfrac{\partial}{\partial u} + \tfrac{1}{u}\tfrac{\partial\xi(t,u)}{\partial u}\right]\delta(u,D), \quad B = \tfrac{1}{v}\delta(v,D) \quad (3.10)$$

for which it follows

$$[A,B] = \tfrac{1}{uv}[x,\ \delta(v,D)]\ \delta(u,D) = \tfrac{1}{uD}\tfrac{\partial}{\partial v}\delta(v,D)\ \delta(u,D) = \tfrac{1}{u^2}\tfrac{\partial}{\partial v}\delta(v,D)\ \delta(u,D) \quad (3.11)$$

Therefore,

$$[[\hat{\mathfrak{T}}(u),\hat{\jmath}(v)]] = \tfrac{1}{u}\tfrac{\partial}{\partial v}\delta(u,v)\hat{\jmath}(u) \quad (3.12)$$

which, similarly to eqs.(5), is an "off-shell" equation reproducing the standard commutation relations.

We see, however, that the energy-momentum tensor written in the form (3.6), although resembling the well-known

$$T = \tfrac{1}{2}\,jj - \tfrac{q}{2}\,\partial j, \quad (3.13)$$

is not quite of this form in terms of the current $\hat{\jmath}$, eq.(3.4). In fact it *cannot* be just bilinear in $\hat{\jmath}$ because – again due to the absence of the central extension – the modes of $\hat{\jmath}$ all commute among themselves:

$\hat{\jmath}(u)\hat{\jmath}(v)$

$$= -\tfrac{1}{uv}\delta(u,v)(K\delta(u,D)K^{-1})_-K - ((K\tfrac{1}{u}\delta(u,D)K^{-1})_+K\tfrac{1}{v}\delta(v,D)K^{-1})_-K$$

$$-((K\tfrac{1}{v}\delta(v,D)K^{-1})_+K\tfrac{1}{u}\delta(u,D)K^{-1})_-K \quad (3.14)$$

which is symmetric in u,v (with whatever meaning given to the delta function as $u\to v$), and therefore would not lead to the commutation relations (5).

That $\jmath$ is 'deficient', can also be understood by going over to the KP hierarchy in the form of equations on the Lax operator Q defined in eq.(2.8):

$$\tfrac{\partial Q}{\partial t}_r = -[(Q^r)_-,Q] \ \text{ (and } = [(Q^r)_+,Q]\text{)} \quad (3.15)$$

The variation of K by $\hat{\jmath}(u)$ leads to the variation of Q of the form,

$$\delta Q = [\jmath(u),Q] = -\tfrac{1}{u}[\delta(u,Q)_+,Q] \quad (3.16)$$

Now, since $Q = KDK^{-1} = D + (\textit{integral operator})$,

$$\tfrac{1}{u}\delta(u,Q)_+ = \sum_{r\geq 0} u^{-r-1}\jmath_r \quad (3.17)$$

so that only half of the modes of $\jmath$ are relevant.

The presence of P is therefore necessary in order to 'compensate' for this deficiency of j. Introducing the vector field

$$\mathfrak{p}(u) = (KP\delta(u,D)K^{-1})_- \tag{3.18}$$

we find that

$$[[\hat{\mathfrak{p}}(u),\hat{\jmath}(v)]] = \frac{\partial\delta(u,v)}{\partial v}\,\hat{\jmath}(u) \tag{3.19}$$

It is this relation that underlies the commutation relation (3.12).

Analogously, the appearance of P is inevitable in what may be called a "bosonized" representation for ψ and ψ^*. The analogues of ψ and ψ^* on a complex plane, b and c, are expressed through the current j via

$$b = e^{-\phi}, \quad c = e^{\phi}, \quad \partial\phi = j \tag{3.20}$$

However, these very simple formulae have to be modified already on higher genus Riemann surfaces: there, ϕ is no longer uniquely determined by j as $\phi(x) \sim \int^x j$, because the integral jumps around the homologies. By the Baker-Akhiezer mechanism, the tentative expression

$$b(u)c(v) \sim \exp\int_u^v j \tag{3.21}$$

acquires an invariant meaning when inserted into the background of a certain operator $\mathcal{B}$ whose main ingredient is the theta function

$$\mathcal{B} \sim \theta(\ldots + \oint_b j) \tag{3.22}$$

depending on b-periods of the *operator* current j [19]. When the product of (3.21) and (3.22) is normal ordered, the argument of the theta function gets shifted as $\oint_b j \longmapsto \oint_b j + v - u$ (where v–u is the Jacobi map), thus maintaining a BA structure ($\exp\int_u^v j)\theta(v-u+\oint_b j)$.

The moral is that there are things of two kinds: the background $\mathcal{B}$, representing a Riemann surface with the *minimal possible* number of insertions, and extra insertions on it. The former corresponds to an operator such as Q (or K), while the extra insertions correspond to *vector fields* on the space of the Q's.

The elementary neutral insertion $b(u)c(v)$ is described by the vector field $\hat{\mathcal{O}}(u,v) = \mathcal{O}(u,v)K\frac{\delta}{\delta K}$ with

$$\mathcal{O}(u,v) = (Ke^{(u-v)P}\tfrac{1}{v}\,\delta(v,D)K^{-1})_- \tag{3.23}$$

This expression, which points to a "bosonic" interpretation, can indeed be rewritten in a "fermionized" form: first, by a straightforward application of the Campbell-Hausdorff formula,

$$e^{zP}\delta(v,D) = e^{\xi(t,v+z)\,-\,\xi(t,v)}\,\delta(v,D)$$

and further

$$Ke^{zP}\delta(v,D) = \sum_{n\geq 0} w_n(t)D^{-n}\circ e^{zz+\Sigma_{r>1}t_r((v+z)^r-v^r)}\,\delta(v,D)$$

$$= e^{\xi(t,v+z)\,-\,\xi(t,v)}\sum_{n\geq 0} w_n(t)(D+z)^{-n}\delta(v,D)$$

$$= \psi(t,v+z)e^{-\xi(t,v)}\delta(v,D)$$

Using this and similar arguments for $\delta(v,D)K^{-1}$ we arrive at

$$(Ke^{zP}\tfrac{1}{v}\,\delta(v,D)K^{-1})_- = \psi(t,v+z)e^{-\xi(t,v)}\frac{1}{D-v}e^{\xi(t,v)}\psi^*(t,v) = \psi(t,v+z)\circ D^{-1}\circ\psi^*(t,v) \tag{3.24}$$

which is a "fermionized" version of (3.23) and shows that $\hat{\mathcal{O}}(u,v)$ does describe an insertion of $b(u)c(v)$ into the background of $Q = KDK^{-1}$.

It is perhaps worthwhile to stress once more that we need *not* assume the KP evolution along the times t_r, $r\geq 2$; we can work just on the phase space of ψDiff operators in x. Accordingly, u and v in (3.23) and (3.24) are formal variables: the formal eigenvalues of Q, see eq.(4).

It is also instructive to calculate the behaviour of the "bc-insertion" under the Virasoro algebra; commuting $\hat{\mathfrak{T}}(z)$ with $\hat{\mathcal{O}}(u,v)$ we find terms involving $\delta(z,u)$ and those involving $\delta(z,v)$. The former, for instance, read

$$\left[\tfrac{1}{z}\,\frac{\partial\log\psi(t,z)}{\partial z}\,\delta(z,u) - J\frac{\partial}{\partial z}\left(\tfrac{1}{z}\,\delta(z,u)\right)\right]\hat{\mathcal{O}}(u,v) \tag{3.25}$$

which is readily seen to equal

$$\left[\delta(z,u)\frac{1}{u}\frac{\partial\psi(t,u)}{\partial u} - J\frac{\partial}{\partial z}\left(\frac{1}{z}\delta(z,u)\right)\psi(t,u)\right]\circ D^{-1}\circ\psi^{*}(t,v) \qquad (3.26)$$

i.e., it reduces to a natural variation of $\psi(t,u)$ which enters (3.24). Similarly with the $\delta(z,v)$-terms. Note that the square bracket in (3.26) is not quite the RHS of eq.(2.22). There is no contradiction, since what we have in (3.26) is the commutator $[[\hat{\mathfrak{T}}(z),\hat{\mathfrak{V}}(u,v)]]$ whereas in (2.22) we acted with $\hat{\mathfrak{T}}(z)$ on a *function* on the phase space.

Now that we have seen what the analogue of $e^{\pm\phi}$ is, a further question would be how this generalizes to $e^{\alpha\phi}$ in ψDiff terms. Answering this might allow one to carry over to integrable hierarchies the Dotsenko-Fateev construction, the conception of *screening*, etc.

4.Toda hierarchy – definitions [2].

Now we use the strategy similar to the one described in Secs.1,2 in order to find the action of the Virasoro algebra on the phase space of the two-dimensional Toda hierarchy[4]. First we recall the basic definitions. Those to whom the meaning of the ingredients of eq.(9) for the Toda case is clear, may skip to Sec.5. Note, however, that our W's are the $\hat{W}$'s of ref.[2].

Dressing operators are $\mathbb{I}\times\mathbb{I}$ matrices which act on a vector space whose vectors are written as $\sum_{s\in\mathbb{I}} v(s)|s\rangle$ (actually twice this space, one for $^{(\infty)}$ and one for $^{(0)}$). Define operators $\hat{\mathrm{p}}$ and Λ by

$$\hat{\mathrm{p}}|s\rangle = s|s\rangle, \quad \Lambda|s\rangle = |s{-}1\rangle \implies [\Lambda,\hat{\mathrm{p}}] = \Lambda \qquad (4.1)$$

Denote by $|u_\lambda\rangle$ the eigenvectors of Λ:

$$\Lambda|u_\lambda\rangle = \lambda|u_\lambda\rangle \qquad (4.2)$$

It is understood that $\hat{\mathrm{p}}$ is hermitian and Λ unitary w.r.t. the inner product $\langle s'|s\rangle = \delta_{s,s'}$. It readily follows then that $\langle s|u_\lambda\rangle = \lambda^s\langle 0|u_\lambda\rangle = \lambda^s$. Define the dressing matrix $W^{(\infty)}$, by

$$W^{(\infty)} = \sum_s |s\rangle\ \langle s|w^{(\infty)}(s;x,y;\Lambda), \qquad (4.3)$$

with $w^{(\infty)}(s,x,y,\lambda)$ expressed through the tau function as

$$w^{(\infty)}(s;x,y;\lambda) \equiv \sum_{j=0}^{\infty} w_j^{(\infty)}(s;x,y)\lambda^{-j} = \frac{\tau(s;x-[\lambda^{-1}],y)}{\tau(s;x,y)} \qquad (4.4)$$

We normalize $w_0 \equiv 1$. Directly from the definitions one gets useful identities

---

[4]From a geometric viewpoint, the main extra input in the two–dimensional Toda hierarchy over KP is a second puncture on the algebraic curve, hence the extra times (which are just the 'KP' times for this puncture), the doubling of fermions, of the Virasoro generators etc. "Physically" the two–punctured version is very appealing because the punctures are naturally interpreted as plus– and minus–time infinity [18].

$$\nabla(\mu,x)\log\tau(s,x) = (\mu\frac{\partial}{\partial\mu} - \nabla(\mu,x))\log w^{(\infty)}(s,\mu), \qquad w^{(\infty)}(s,\mu)=\mu^{-s}\langle s|\, W^{(\infty)}|\, u_{\mu}\rangle \quad (4.5)$$

$$\nabla(\mu,x) \equiv \sum_{r\geq 1}\mu^{-r}\frac{\partial}{\partial x_r}$$

Introduce further

$$W^{(\infty)^{-1}} = \sum_s w^{(\infty)*}(s;x,y;\Lambda)|s\rangle\,\langle s|, \quad (4.6)$$

where the adjoint wave functions $w^{(\infty)*}$ turn out to be [2]

$$w^{(\infty)*}(s;x,y;\lambda) = \frac{\tau(s;x+[\lambda^{-1}],y)}{\tau(s\,;\,x\,,\,y)} \quad (4.7)$$

The Lax operator L is defined by

$$L = W\Lambda W^{-1} \quad (4.8)$$

The associated linear problem, with the eigenvalue λ, can be written in terms of the vector

$$\hat{\Psi}(\lambda) = \sum_s |s\rangle\psi(s,\lambda), \quad \psi(s;x,y;\lambda) = \lambda^s w^{(\infty)}(s;x,y;\lambda) \quad (4.9)$$

We have

$$L\hat{\Psi}(\lambda) = \lambda\hat{\Psi}(\lambda) \quad (4.10)$$

This equation turns into an identity as soon as one notices that, knowing $W^{(\infty)}$, one can construct $\hat{\Psi}(\lambda)$ as

$$\hat{\Psi}(\lambda) = W^{(\infty)}|\, u_{\lambda}\rangle \quad (4.11)$$

The other dressing matrix of the two-dimensional Toda lattice hierarchy, $W^{(0)}$, is constructed as

$$W^{(0)} = \sum_s |s\rangle\,\langle s|w^{(0)}(s;x,y;\Lambda) \quad (4.12)$$

$$w^{(0)}(s;x,y;\lambda) = \frac{\tau(s+1;x,y-[\lambda])}{\tau\,(s;x,y)} \quad (4.13)$$

We will also have a chance to use the "full" wave functions [2]

$$\psi^{(\infty)}(s;x,y;\mu) = w^{(\infty)}(s;x,y;\mu)\mu^s e^{\xi(x,\mu)} \quad (4.14)$$

$$\psi^{(\infty)*}(s;x,y;\mu) = w^{(\infty)*}(s;x,y;\mu)\mu^{-s} e^{-\xi(x,\mu)} \quad (4.15)$$

The (x-half of the) Toda hierarchy equations read

$$\frac{\partial W^{(\infty)}}{\partial x_r} = -(L^r)_- W^{(\infty)} \quad (4.16)$$

$$\frac{\partial W^{(0)}}{\partial x_r} = (L^r)_+ W^{(0)} \quad (4.17)$$

where $(\ldots)_{\mp}$ refers to the strictly lower–triangular (non-strictly upper-triangular) part of a matrix.

We do not consider the dependence of our operators and functions on the y–times of the two–dimensional Toda hierarchy, nor will we introduce the corresponding second set of the Virasoro generators. Accordingly, we will sometimes omit indicating the dependence on y.

5. Toda hierarchy – the Virasoro action.

Now we apply the Virasoro generators to the tau function and use eqs.(4.1–17) to systematically express $\frac{\partial}{\partial x_r}$'s through the flows (4.16). The tau function is varied as

$$\tilde{\tau}(s) = \tau(s) + \sum_{k\in\mathbb{Z}} \mu^{-k-2} \mathscr{L}_k \tau(s) \equiv \tau(s) + \mathscr{A}(\mu)\tau(s) \tag{5.1}$$

where the Virasoro generators are given by eqs. (2.16) with $t_r \longmapsto x_r$ and a_0 related to $\hat{p}$ by

$$a_0 = \hat{p} + \tfrac{1}{2}q \tag{5.2}$$

with q arbitrary as yet.

5.1. THE $^{(\infty)}$-PIECE. For the time being, we drop the superscript $^{(\infty)}$. Under the variation (5.1), the wave function $w \equiv w^{(\infty)}$ changes as

$$\tilde{w}(s,\lambda) = \frac{\tilde{\tau}(s;x-[\lambda^{-1}])}{\tilde{\tau}(s;x)} \equiv \frac{\bar{\Gamma}(x;\lambda)\tilde{\tau}(s;x)}{\tilde{\tau}(s;x)} = w(s;x;\lambda) - w(s;x;\lambda)\frac{\mathscr{A}(\mu)\tau(s;x)}{\tau(s;x)} + \frac{\bar{\Gamma}(x;\lambda)\,\mathscr{A}(\mu)\tau(s;x)}{\tau(s;x)}$$

(we drop the dependence on y), where the energy-momentum tensor $\mathscr{A}(u)$ is given by

$$\mathscr{A}(u) = \mu^{-2}\Big[\tfrac{1}{2}\nabla(\mu,x)\nabla(\mu,x) + \mu\frac{\partial\xi(x,\mu)}{\partial\mu}\nabla(\mu,x) + a_0\nabla(\mu,x) - (J-\tfrac{1}{2})\mu\frac{\partial}{\partial\mu}\nabla(\mu,x)$$

$$+ \tfrac{1}{2}\Big[\mu\frac{\partial\xi(x,\mu)}{\partial\mu}\Big]^2 + \mu\frac{\partial\xi(x,\mu)}{\partial\mu}a_0 - (J-\tfrac{1}{2})\mu\frac{\partial}{\partial\mu}\Big[\mu\frac{\partial\xi(x,\mu)}{\partial\mu}\Big] + \tfrac{1}{2}a_0^2 - \frac{1}{24}\Big]$$

As before,

$$\xi(x,\mu) = \sum_{r\geq 1} x_r\mu^r, \quad \nabla(\mu,x) = \sum_{r\geq 1}\mu^{-r}\frac{\partial}{\partial x_r}$$

Evaluating now $\bar{\Gamma}(x,\lambda)\,\mathscr{A}(\mu)\bar{\Gamma}(x,\lambda)^{-1}$, we calculate δW as

$$\mu^2\delta W|u_\lambda\rangle = \mu^2\sum_s |s\rangle\,\delta w(s;\lambda)\lambda^s$$

$$= \sum_s (\nabla(\mu,x)\log\tau(s,x) + a_0)(\nabla(\mu,x) - F(\mu,\lambda))w(s;\lambda)|s\rangle\lambda^s$$

$$+ \Big[-(J-\tfrac{1}{2})\mu\frac{\partial}{\partial\mu} + \mu\frac{\partial\xi(x,\mu)}{\partial\mu} + \tfrac{1}{2}(\nabla(\mu,x) - F(\mu,\lambda))\Big](\nabla(\mu,x) - F(\mu,\lambda))\,W|u_\lambda\rangle$$

$$+ \sum(\tfrac{1}{2}a_0^2 - \tfrac{1}{24})w(s;\lambda)|s\rangle\lambda^s \tag{5.3}$$

with

$$F(\mu,\lambda) = \sum_{r\geq 1}(\mu/\lambda)^r \tag{5.4}$$

The use of the Toda hierarchy equations leads to

$$(\nabla(\mu,x) - F(\mu,\lambda))\,W|u_\lambda\rangle = (-\mathscr{A}(\mu)_- W - WF(\mu,\Lambda))|u_\lambda\rangle$$

$$= \delta(\mu,L)_- W|u_\lambda\rangle \tag{5.5}$$

where

$$\mathscr{A}(\mu) = \sum_{n\geq 0} L^n\mu^{-n} \tag{5.6}$$

Inserting (5.5) into (5.3), let us first concentrate on the second term on the RHS of (5.3): applying the square bracket to (5.5), we again make use of the Toda hierarchy equations. This involves a standard exercise with the $(\ldots)_+ - (\ldots)_-$-game, which, together with the identity

$$\mu\frac{\partial}{\partial\mu}\delta(\mu,L)_- = \delta(\mu,L)_- + F(\mu,L)^2 - (\mathscr{A}(\mu)^2)_-, \tag{5.7}$$

allows us to express the second term in (5.3) as

$$\Big[(\mathscr{A}(\mu)_-\delta(\mu,L))_- - \tfrac{1}{2}\delta(\mu,L)_- + (J\mu\frac{\partial}{\partial\mu} - \mu\frac{\partial\xi(x,\mu)}{\partial\mu})\delta(\mu,L)_-\Big]\,W|u_\lambda\rangle \tag{5.8}$$

Next consider the first term in (5.3). The summation over s is not immediate. First we use eqs.(4.5), thus representing the first term in (5.3) in the form,

$$-\sum_s\Big[(\mu\frac{\partial}{\partial\mu} - \nabla(\mu,x))\log w(s;\mu) + a_0\Big]|s\rangle\langle s|\delta(\mu,L)_- W|u_\lambda\rangle \tag{5.9}$$

Further, it is not hard to see that

$$\delta(\mu,L) = \sum_{s',s''} |s''\rangle w(s'',\mu)\langle s''|\delta(\mu,\Lambda)|s'\rangle w^*(s'+1;\mu)\langle s'| \qquad (5.10)$$

(A useful observation is that the delta function on the RHS is in fact a projector,

$$\delta(\mu,\Lambda) = |u_\mu\rangle\langle u_\mu|$$

which can be checked by the following formal manipulations,

$$|u_\lambda\rangle\langle u_\lambda| = \sum_{s,r} \lambda^s |s\rangle\langle r|\lambda^{-r} = \sum_{p,r} \lambda^p |r+p\rangle\langle r|$$

$$= \sum_{p,r} \lambda^p \Lambda^{-p} |r\rangle\langle r| = \sum_p \lambda^p \Lambda^{-p} .)$$

The operation of taking the lower diagonal part commutes with multiplication with diagonal matrices. Therefore eq.(5.9) becomes

$$-\sum_{s,s'} (\mu\frac{\partial}{\partial\mu} - \nabla(\mu,x) + a_0) w(s;\mu)|s\rangle\langle s|\delta(\mu,\Lambda)_-|s'\rangle w^*(s'+1;\mu)\langle s'|W|u_\lambda\rangle \qquad (5.11)$$

Finally, we use here eq.(4.5) and recall the hierarchy equations (4.16):

$$\sum_s |s\rangle\nabla(\mu,x)w(s,\mu)\langle s|\delta(\mu,\Lambda)_- = (\nabla(\mu,x)W\delta(\mu,\Lambda))_- = -(\mathscr{P}(\mu)_- W\delta(\mu,\Lambda))_- \qquad (5.12)$$

which gives for (5.11)

$$-\Big[(\mathscr{P}(\mu)_- + \tfrac{1}{2}q)\delta(\mu,L)\Big]_- W|u_\lambda\rangle$$

$$-\sum_{s,s'} |s\rangle(\mu\frac{\partial}{\partial\mu} + s)w(s;\mu)\langle s|\delta(\mu,\Lambda)_-|s'\rangle w^*(s'+1;\mu)\langle s'|W|u_\lambda\rangle \qquad (5.13)$$

Now, putting (5.8) and (5.13) together, we observe nice cancellations, with the result

$$\mu^2\delta W^{(\infty)} = -\sum_{s,s'} |s\rangle\Big[(\mu\frac{\partial}{\partial\mu} + s)w(s;\mu)\Big]\langle s|\delta(\mu,\Lambda)_-|s'\rangle w^*(s'+1;\mu)\langle s'|W^{(\infty)}$$

$$+\Big[J\mu\frac{\partial}{\partial\mu} - \mu\frac{\partial\xi(x,\mu)}{\partial\mu} - \frac{1}{2}q - \frac{1}{2}\Big]\delta(\mu,L)_- W^{(\infty)} + \Big[\frac{1}{2}\hat{p}^2 + \frac{1}{2}q\hat{p} + \frac{1}{8}q^2 - \frac{1}{24}\Big]W^{(\infty)}$$

$$\equiv -\mu^2\mathfrak{T}(\mu)W^{(\infty)} \qquad (5.14)$$

The final step consists in expressing $w^{(\infty)}$ and $w^{(\infty)*}$ entering (5.14), through the "full" wave functions (4.14,15). While doing that, we replace $\langle s|\mu^{-s}$ by $\langle s|\mu^{-\hat{p}}$ and similarly for $\mu^{s'+1}|s'\rangle$, and use the following consequence of eq.(4.1):

$$\mu^{-\hat{p}}\delta(\mu,\Lambda)_-\mu^{\hat{p}+1} = \mu\sum_{r\geq 1}\Lambda^{-r} \qquad (5.15)$$

Then

$$\mathfrak{T}(\mu) = \sum_s\sum_{r\geq 1}|s\rangle\Big[(1-J)\frac{\partial\psi(s;\mu)}{\partial\mu}\psi^*(s-r+1;\mu) - J\psi(s;\mu)\frac{\partial\psi^*(s-r+1;\mu)}{\partial\mu}\Big]\langle s-r|$$

$$-\mu^{-1}(J-\tfrac{1}{2}q-\tfrac{1}{2})\sum_s\sum_{r\geq 1}|s\rangle\psi(s;\mu)\psi^*(s-r+1;\mu)\langle s-r|$$

$$-\mu^{-2}(\tfrac{1}{2}\hat{p}^2 + \tfrac{1}{2}q\hat{p} + \tfrac{1}{8}q^2 - \tfrac{1}{24}) \qquad (5.16)$$

A natural choice of q now is

$$q = 2J-1 \qquad (5.17)$$

i.e., the value of the background charge. We have thus arrived at the energy–momentum tensor on the hierarchy as the following (non–strictly) lower–triangular matrix

$$\mathfrak{T}(\mu) = \sum_s\sum_{r\geq 1}|s\rangle\Big[(1-J)\frac{\partial\psi^{(\infty)}(s;\mu)}{\partial\mu}\psi^{(\infty)*}(s-r+1;\mu)$$

$$-J\psi^{(\infty)}(s;\mu)\frac{\partial\psi^{(\infty)*}(s-r+1;\mu)}{\partial\mu}\Big]\langle s-r|$$

$$-\tfrac{1}{2}\mu^{-2}(\hat{p}^2 + q\hat{p} + J^2 - J + \tfrac{1}{6}) \qquad (5.18)$$

The appearance of the extra piece in $\mathscr{L}_0$, proportional to $J^2 - J + \frac{1}{6}$, was already observed in Sec.2.

Eq.(5.18) is a direct Toda hierarchy analogue of our previous result (2.19). Indeed, suppressing the y–times of the Toda hierarchy, gives rise to the KP hierarchy in a matrix form (as in [14]). More precisely, what one gets in this way is

an infinite number of copies of the Schlesinger transformed KP hierarchy. (The Schlesinger transform is a shift along a "zero time", the discrete parameter s, [17].) The energy–momentum tensor (5.18) then corresponds to the KP one (2.19).[5] In Sec.8 we will obtain another, quite different, relation between the Virasoro algebras in the two cases, when the zero time s becomes a continuum variable identified as the first time $x \equiv t_1$ of the KP hierarchy.

5.2. THE $^{(0)}$-PIECE. It remains to see what happens to the other dressing matrix in the Toda theory, $W^{(0)}$. Again, we start from the tau function and vary it as before, eq.(5.1). Write eq.(4.13) as

$$w^{(0)}(s;x,y;\lambda) = \frac{\bar\Gamma(y,\lambda^{-1})e^{Q}\tau(s;x,y)}{\tau(s;x,y)}$$

where $e^{Q}\tau(s;x,y) = \tau(s+1;x,y)$, $e^{Q}a_0 e^{-Q} = a_0 - 1$ and

$$\bar\Gamma(x,\lambda^{-1}) = \exp -\sum_{r\geq 1}\frac{1}{r}\lambda^r\frac{\partial}{\partial x_r}$$

Now $\bar\Gamma(y,\lambda^{-1})$ commutes with $\mathscr{A}(\mu) \equiv \mathscr{A}(\mu;x)$; instead,

$$e^{Q}\mathscr{A}(\mu)e^{-Q} = \mathscr{A}(\mu) + a_0 + \tfrac{1}{2}\nabla(\mu,x) + \mu\frac{\partial\xi(x,\mu)}{\partial\mu} \qquad (5.19)$$

Hence

$$\tilde w^{(0)}(s;x,y;\lambda) = w^{(0)}(s;x,y;\lambda)$$

$$+\left[\tfrac{1}{2}\nabla(\mu,x) + \nabla(\mu,x)\log\tau(s) + \mu\frac{\partial\xi(x,\mu)}{\partial\mu} - (J-\tfrac{1}{2})\mu\frac{\partial}{\partial\mu} + a_0 + 1\right]\nabla(\mu,x)w^{(0)}(s;x,y;\lambda)$$

$$+(a_0 + \nabla(\mu,x)\log\tau(s) + \mu\frac{\partial\xi(x,\mu)}{\partial\mu})w^{(0)}(s;x,y;\lambda) \qquad (5.20)$$

Now the hierarchy equations (4.17) imply

$$\nabla(\mu,x)W^{(0)}(x,y) = (\mathscr{A}(\mu)_+ - 1)W^{(0)}(x,y) \qquad (5.21)$$

It is very convenient to replace here $\mathscr{A}(\mu)_+$ with $\delta(\mu,L)_+$. Then we get

[5] I wish to thank A.Yu.Orlov for a discussion of this point.

$$\mu^2\delta W^{(0)}(x,y)|u_\lambda\rangle$$

$$= \sum_s\left[\tfrac{1}{2}\nabla(\mu,x) + \nabla(\mu,x)\log\tau(s) + \mu\frac{\partial\xi(x,\mu)}{\partial\mu} - (J-\tfrac{1}{2})\mu\frac{\partial}{\partial\mu} + a_0 + 1\right]|s\rangle\times$$

$$\times\langle s|(\delta(\mu,L)_+ - 1)W^{(0)}|u_\lambda\rangle$$

$$+\sum_s(a_0 + \nabla(\mu,x)\log\tau(s) + \mu\frac{\partial\xi(x,\mu)}{\partial\mu})|s\rangle w^{(0)}(s;x,y;\lambda)\lambda^s \qquad (5.22)$$

The subtraction of 1 from the delta function proves very essential: after a number of cancellations, this becomes

$$\delta W^{(0)}|u_\lambda\rangle = -\tfrac{1}{2}\mu^{-2}W^{(0)}|u_\lambda\rangle$$

$$+\sum_s\sum_{r\geq 1}|s\rangle\left[(1-J)\frac{\partial\psi^{(\infty)}(s;\mu)}{\partial\mu}\psi^{(\infty)*}(s+r+1;\mu)\right.$$

$$\left.-J\psi^{(\infty)}(s;\mu)\frac{\partial\psi^{(\infty)*}(s+r+1;\mu)}{\partial\mu}\right]\langle s+r|W^{(0)}|u_\lambda\rangle$$

$$+\mu^{-1}(-J+\tfrac{1}{2}q+\tfrac{1}{2})\sum_s\sum_{r\geq 1}|s\rangle\psi(s;\mu)\psi^*(s+r+1;\mu)\langle s+r|W^{(0)}|u_\lambda\rangle \qquad (5.23)$$

whence the structure of the energy-momentum tensor acting on $W^{(0)}$ is read off. Again $q = 2J-1$ is a most natural choice, and again a generalized $(1-J)\partial b\cdot c - Jb\cdot\partial c$ structure emerges.

6. Virasoro action on Schrodinger operators.[6]

Let us digress for a while and recall attempts towards describing strings and, generally, conformal field theories, without Riemann surfaces, i.e., in terms of a universal moduli space. Such a formulation would imply, heuristically, a "non-trivial change of variables" from the world-sheet to some "non-perturbative" objects. (The Universal Grassmannian, as it is, appears too big.) A useful, and familiar to many, set of non-perturbative objects is provided [20] by second-order differential operators

$$H = \frac{\partial}{\partial z \partial \bar{z}} - b(z,\bar{z})\frac{\partial}{\partial \bar{z}} - c(z,\bar{z}) \tag{6.1}$$

with periodic (complex) coefficients:

$$f(z + l_1, \bar{z}) = f(z, \bar{z}),\ f(z, \bar{z} + l_2) = f(z, \bar{z}), \quad f = b, c \tag{6.2}$$

Every such operator encodes the information about a Riemann surface, as explained in [20]: That is, consider solutions of the eigenvalue problem for H and the formally adjoint H^*, for the eigenvalue $E = 0$,

$$H\psi = 0, \quad H^*\psi^* = 0 \tag{6.3}$$

with multiplicities w_1 and w_2 :

$$\begin{aligned} &\psi(z + l_1, \bar{z}) = w_1\psi(z, \bar{z}), \quad \psi(z, \bar{z} + l_2) = w_2\psi(z, \bar{z}) \\ &\psi^*(z + l_1, \bar{z}) = w_1^{-1}\psi^*(z, \bar{z}), \quad \psi^*(z, \bar{z} + l_2) = w_2^{-1}\psi^*(z, \bar{z}) \end{aligned} \tag{6.4}$$

The locus Γ of those (w_1, w_2) for which the solution exists, is a nice geometrical object: either a smooth finite genus Riemann surface, or a Riemann surface with singularities, or an infinite-genus Riemann surface. The crucial feature is that all the "infinities" that may occur, are controllable by the spectral theory of differential operators, so that a natural extension of the locus of finite genus Riemann surfaces can be *defined* in this way [20].

More precisely, the theorem says that if one starts with the spectral problem and its Bloch solution, then it is possible to construct a Riemann surface which is isomorphic to the spectral curve of H spanned by the multiplicators.

[6]The material of this section is based in part on a joint work with I.M.Krichever.

Moreover, the Bloch solutions become meromorphic functions on the Riemann surface, their poles being independent of x and $\bar{x}$ (the Baker-Akhiezer functions).

As in Sec.2, we are now interested in the description in terms of the H's not just of 'bare' Riemann surfaces, but also of other field theory ingredients. We again consider the energy-momentum tensor.

The insertion of $T(P)$ at a point $P \in \Gamma$ can be described as an infinitesimal deformation of the Riemann surface Γ [5]. Therefore the multiplicators w_1 and w_2 change, which can be translated into a deformation of ψ and ψ^* and then into a deformation of H. For us, however, it would be easier to find the transformation of ψ, ψ^* and H using an alternative procedure. We rely on the fact that there *is* an operator of the type of (6.1) associated to the two-dimensional Toda lattice hierarchy. (And vice versa, starting from such an operator one reconstructs the Toda hierarchy). Therefore we can use the results of Sec.5 on the Virasoro action if we identify H inside the Toda hierarchy and express the Virasoro variation of ψ, ψ^* and H in *intrinsic* terms of eq.(6.1). The latter essentially amounts to getting rid of Schlesinger-transformed components of $\psi(s)$ (i.e., those with a shifted s).

First let us extract H. Consider the evolutions of the wave function

$$\Psi(\lambda) = W^{(\infty)}e^{\xi(x,\Lambda)}|u_\lambda\rangle$$

(cf. eq.(4.9)) along the times x_n:

$$\frac{\partial\Psi(\lambda)}{\partial x_n} = (L^n)_+\Psi(\lambda) \tag{6.5}$$

or,

$$\frac{\partial\psi(s,\lambda)}{\partial x_n} = \sum_r \langle s|(L^n)_+|r\rangle\ \psi(r,\lambda) \tag{6.6}$$

In fact, this holds for both $\psi^{(\infty)}$ and $\psi^{(0)}$ defined in (4.15,16). Now define the matrix elements of L as

$$L = \sum_s \sum_{j\le 1} |s\rangle b_j(s)\langle s+j| \qquad (b_0 \equiv 1) \tag{6.7}$$

hence from (6.6) for n=1 it follows

$$\frac{\partial}{\partial x_1}\psi(s) = \psi(s+1) + b_0(s)\psi(s) \tag{6.8}$$

Similarly with the y-evolutions

$$\frac{\partial\psi(\lambda)}{\partial y_m} = -(\mathcal{M}^m)_-\psi(\lambda)$$

We write

$$M \equiv W^{(0)}\Lambda^{-1}W^{(0)-1} = \sum_s \sum_{j\geq -1} |s\rangle c_j(s)\langle s+j| \tag{6.9}$$

whence

$$\frac{\partial}{\partial y_1}\psi(s) = c_{-1}(s)\psi(s-1) \tag{6.10}$$

where, again, ψ is any of $\psi^{(\infty)}$, $\psi^{(0)}$. Combining eqs.(6.8) and (6.10) we find that $\psi(s)$ is annihilated by

$$H(s) \equiv \frac{\partial^2}{\partial x_1 \partial y_1} - b_0(s)\frac{\partial}{\partial y_1} - c_{-1}(s) \tag{6.11}$$

(that is, we identify $z = x_1$, $\bar{z} = y_1$). The same argument shows that the adjoint wave function $\psi^*(s+1)$ is a zero mode of the formally adjoint operator H^*:

$$H(s)\psi(s) = 0, \quad H^*(s)\psi^*(s+1) = 0 \tag{6.12}$$

At this point we can use results of Sec.5. First, however, we have to modify slightly the transformation obtained there so as to stay with operators L normalized as in eq.(6.7). The Virasoro transformation of the Lax operator,

$$\delta L = [\mathfrak{T}(\mu), L]$$

with $\mathfrak{T}(\mu)$ given by eq.(5.18), does *not* preserve the normalization $b_0 = 1$. The correct normalization is restored by an overall compensating gauge transformation, which amounts to dropping the second term on the RHS of (5.18). (It will, of course, reappear in the transformation formula of $W^{(0)}$). Then the variation of the wave function under the action of the "x–" Virasoro generators is read off from eq.(4.11):

$$\delta\psi^{(\infty)}(s;\lambda) = (1-J)\omega(s;\lambda,\mu)\frac{\partial\psi^{(\infty)}(s;\mu)}{\partial\mu} - J\frac{\partial\omega(s;\lambda,\mu)}{\partial\mu}\psi^{(\infty)}(s;\mu) \tag{6.13}$$

with

$$\omega(s;\lambda,\mu) \equiv \omega(s;\lambda,\mu;x,y) = \sum_{r\geq 1}\psi^{(\infty)}(s-r;\lambda)\psi^{(\infty)*}(s-r+1;\mu) \tag{6.14}$$

Recall that μ is the parameter of the transformation. However, the RHS of (6.14) involves an infinite tail of ψ's and ψ^*'s which are *not* eigenfunctions of H or H^*, while we would like to express $\delta\psi$ in intrinsically "Schrödinger" terms. A helpful observation is that by differentiating $\omega(s;\lambda,\mu)$ one gets, using eqs.(6.8,10),

$$\frac{\partial\omega(s;\lambda,\mu)}{\partial x_1} = \psi^{(\infty)}(s;x,y;\lambda)\psi^{(\infty)*}(s;x,y;\mu)$$
$$\frac{\partial\omega(s;\lambda,\mu)}{\partial y_1} = -c_{-1}(s)\psi^{(\infty)}(s-1;x,y;\lambda)\psi^{(\infty)*}(s+1;x,y;\mu) \tag{6.15}$$

Now let us use the above formulae in the situation described in the beginning of this section where we have no *spectral* problem and no coordinate system around punctures on the Riemann surface, and the solution ψ to H depends on a point on the surface Γ only through its Bloch momenta.

Therefore, let us be given an operator H of the form (6.1), and for $Q\in\Gamma$, let $\psi_i(x,y,Q)$ and $\psi_j^*(x,y,Q)$ $i,j = 1,2$, be solutions with Bloch momenta w_1 and w_2 determined by Q. Then the Virasoro action on these data is described as follows.

The insertion of the energy-momentum tensor at a point $P\in\Gamma$ leads to the following infinitesimal transformation of $\psi_1(x,y,Q)$,

$$\delta_P\psi_1(x,y;Q) = (1-J)\omega_{1i}(x,y;Q,P)\partial_P\psi_i(x,y;P) - J\partial_P\omega_{1i}(x,y;Q,P)\psi_i(x,y;P) \tag{6.16}$$

(with the sum going over $i=1,2$) where ω_{1i} is to be determined from

$$\frac{\partial\omega_{1i}(x,y;Q,P)}{\partial x} = -\psi_1(x,y;Q)\left[\frac{\partial\psi_i^*(x,y;P)}{\partial x} + b(x,y)\psi_i^*(x,y;P)\right] \tag{6.17a}$$

$$\frac{\partial\omega_{1i}(x,y;Q,P)}{\partial y} = -\frac{\partial\psi_1^*(x,y;Q)}{\partial y}\psi_i^*(x,y;P) \tag{6.17b}$$

(this system *is* consistent: the antisymmetrized mixed second derivative of ω is proportional to $H\psi_j$ or $H^*\psi_j^*$).

Further, the coefficients of H vary as

$$\delta_P b(x,y) = (1-J)\left[-\partial_P\psi_i(P)\frac{\partial\psi_i^*(P)}{\partial x} - \left[\frac{\partial}{\partial x}\partial_P\psi_i(P)\right]\psi_i^*(P)\right]$$
$$+ J\left[\psi_i(P)\frac{\partial}{\partial x}\partial_P\psi_i^*(P) - \left[\frac{\partial\psi_i(P)}{\partial x}\right]\partial_P\psi_i^*(P)\right], \tag{6.18}$$

$$\delta_P c(x,y)$$

$$= (1-J)\Big[-\Big[\frac{\partial}{\partial y}\partial_P\psi_i(P)\Big]\frac{\partial\psi_i^*(P)}{\partial x} - b(x,y)\Big[\frac{\partial}{\partial y}\partial_P\psi_i(P)\Big]\psi_i^*(P) - c(x,y)\partial_P\psi_i(P)\psi_i^*(P)\Big]$$

$$+ J\Big[\frac{\partial\psi_i(P)}{\partial y}\frac{\partial}{\partial x}\partial_P\psi_i^*(P) + b(x,y)\frac{\partial\psi_i(P)}{\partial y}\partial_P\psi_i^*(P) + c(x,y)\psi_i(P)\partial_P\psi_i^*(P)\Big] \qquad (6.19)$$

Note that, formally at least, the resulting transformation of H can be implemented *modulo* H by the commutation with the ψDiff operator

$$(1-J)\partial_P\psi_i(P)\circ\partial_y^{-1}\circ\psi_i^*(P)\circ\partial_y - J\psi_i(P)\circ\partial_y^{-1}\circ\partial_P\psi_i^*(P)\circ\partial_y \qquad (6.20)$$

This fact is helpful for the purposes of reversing the argument, i.e., in order to show that this action coincides with the one we had in Sec.5 provided the Toda hierarchy is reconstructed from the operator H and its solutions. The reconstruction of the Toda hierarchy from an H of the above type can, of course, be started by making the Toda lattice equations,

$$\frac{\partial b_0(s)}{\partial y} = c_{-1}(s) - c_{-1}(s+1)$$
$$\qquad (6.21)$$
$$\frac{\partial\log c_{-1}(s)}{\partial x} = b_0(s) - b_0(s-1)$$

to be tautologically satisfied: one fixes an integer, $s_\star$, and declares that

$$b = b_0(s_\star), \quad c = c_{-1}(s_\star) \quad \psi = \psi(s_\star) \quad \psi^* = \psi^*(s_\star+1) \qquad (6.22)$$

and then *defines*

$$b_0(s_\star-1) = b - \frac{\partial\log c}{\partial x}, \qquad \psi(s_\star+1) = (\frac{\partial}{\partial x} - b)\psi$$

$$c_{-1}(s_\star+1) = c - \frac{\partial b}{\partial y}, \qquad \psi(s_\star-1) = \frac{1}{c}\frac{\partial\psi}{\partial y} \qquad (6.23)$$

$$\psi^*(s_\star) = -\frac{\partial\psi^*}{\partial x} - b\psi^* \qquad \psi^*(s_\star+2) = -\frac{1}{c_{-1}(s_\star+1)}\frac{\partial\psi^*}{\partial y}$$

and further

$$c_{-1}(s_\star-1) = c + \frac{\partial b_0(s_\star-1)}{\partial y}$$

and so forth. *Higher* hierarchy equations are in principle reconstructed as symmetries of the lowest one.

However, one can stay with purely differential equations equivalent to the Toda hierarchy. These can be nicely formulated as follows. First, notice that eq.(6.8) leads to a formula for the Schlesinger transform of ψ [2]:

$$\psi(s+j;\lambda) = (\frac{\partial}{\partial x_1} - b_0(s+j-1))\ldots(\frac{\partial}{\partial x_1} - b_0(s))\psi(s;\lambda) \qquad (6.24)$$

This can be substituted back into eqs.(6.6) for $n\geq 2$. In this way one gets purely *differential* (*not* differential-difference) equations for $\psi(s)$. For instance,

$$\Big[\frac{\partial}{\partial x_2} - \Big\{\frac{\partial^2}{\partial x_1^2} - \frac{\partial b_0(s)}{\partial x_1} + b_{-1}(s) + b_{-1}(s+1)\Big\}\Big]\psi(s) = 0 \qquad (6.25)$$

and so on,

$$\Big[\frac{\partial}{\partial x_n} - \mathscr{D}_x^{(n)}(\partial/\partial x_1)\Big]\psi(s) = 0 \qquad (6.26)$$

for some n-th order differential operator $\mathscr{D}$ in $\partial/\partial x_1$. The y_2-analogues of (6.25,26) follow from (6.10) and read

$$\Big[\frac{\partial}{\partial y_2} - \Big\{\frac{\partial^2}{\partial y_1^2} + \Big[c_0(s-1) + c_0(s) - \frac{\partial\log c_{-1}(s)}{\partial y_1}\Big]\frac{\partial}{\partial y_1}\Big\}\Big]\psi(s) = 0 \qquad (6.27)$$

$$\Big[\frac{\partial}{\partial y_n} - \mathscr{D}_y^{(n)}(\partial/\partial y_1)\Big]\psi(s) = 0 \qquad (6.28)$$

Obviously, all these equations require consistency conditions, and of course the integrability is equivalent to the Toda hierarchy equations.

$$\frac{\partial L^m}{\partial y_n}_+ - \frac{\partial M^n}{\partial x_m}_+ + [L^m{}_+, M^n{}_+] = 0, \quad m,n \geq 1 \qquad (6.29)$$

and similar xx- and yy- conditions. The crucial fact, which relates the Toda hierarchy to the theory of rings of differential operators, is that by virtue of the hierarchy equations, all the commutators

$$\Big[\frac{\partial}{\partial x_n} - \mathscr{D}_x^{(n)}(\partial/\partial x_1)\,,\, \frac{\partial}{\partial y_n} - \mathscr{D}_y^{(n)}(\partial/\partial y_1)\Big]$$

and similar xx- and yy- commutators are proportional to the operator H, i.e., all the commutators lie in the left ideal generated by H in the ring of differential operators in all the times (x_n, y_m). For example,

$$\left[\frac{\partial}{\partial x_2} - \mathscr{D}_x^{(2)}(\partial/\partial x_1)\,,\ \frac{\partial}{\partial y_2} - \mathscr{D}_y^{(2)}(\partial/\partial y_1)\right] = 4(c_{-1}(s+1) - c_{-1}(s))H$$

Now one can replace the Toda hierarchy with the ring of differential operators

$$\frac{\partial}{\partial x_n} - \mathscr{D}_x^{(n)}(\partial/\partial x_1),\quad \frac{\partial}{\partial y_m} - \mathscr{D}_y^{(m)}(\partial/\partial y_1)$$

which is commutative modulo H [26]. The Virasoro action induced from the above action on H then agrees with the one we had in Sec.5. On the other hand, formulae (6.16–20) provide us with the Virasoro action on a realization of the universal moduli space.

7. Virasoro constraints and topological theories.

As noted in the Introduction, Virasoro constraints on partition functions of (1-)matrix models follow as Ward identities for the tau function represented as a matrix integral, and presumably characterize those solutions to the hierarchy that come from matrix models. In this section we consider the Virasoro constraints using the formalism developed in the previous sections. We choose the KP hierarchy, but from the remarks in the Introduction it follows that this should be looked at as a model example illustrating the general situation (it can be viewed also as an embryo for the generalized (m)KdV's, or, in anticipation of the results of Sec.8, as the continuum limit of the Virasoro-constrained Toda hierarchy). Thus, impose

$$\mathscr{L}_r\tau(t) = 0,\quad r\geq 0 \tag{7.1}$$

Then, in the variation of K under the action of $\hat{\mathfrak{T}}(u)$ only terms with u to the power ≥ -1 remain. These can be extracted from the "bosonized" form of $\mathfrak{T}(u)$, see Sec.3. As a result, the "reduced" conformal transformation

$$\bar{\delta}\tau = \sum_{r\leq -1} u^{-r-2}\mathscr{L}_r\tau \tag{7.2}$$

affects K as

$$\bar{\delta}K = (K(J\frac{1}{(D-u)^2}\frac{D}{u} - P\frac{1}{(D-u)}\frac{D}{u})K^{-1})_- K \tag{7.3}$$

Recall that the operator P, eq.(3.7), represents on ψDiff operators the operator of differentiation w.r.t. the spectral parameter. Then the RHS of (7.3) is "almost" the usual $T(u)b(z) = \frac{Jb(z)}{(u-z)^2} + \frac{\partial b(z)}{u-z}$ with the coordinate z "replaced" by a differentiation. We would like to stress that in (7.3) and similar formulae the meaning of $1/(D-u)$ is that of the series $\frac{1}{u}\Sigma_{n\geq 0}(u/D)^n$. (This can also be loosely expressed by saying that $|u| < |D|$, which is parallel to the *radial ordering* on the LHS of the operator product $T(u)b(z)$.)

As to the *constraints* themselves, they are written in a suggestive form

$$\mathfrak{L}_r \equiv (K(JrD^r + PD^{r+1})K^{-1})_- = 0,\quad r\geq 0 \tag{7.4}$$

and can be further summed up into the following generating expression,

$$(K\,(x+\tfrac{1}{u}J+\sum_{s\geq 2} st_s D^{s-1})\;D\;e^{D/u}\;K^{-1})_{-}=0 \qquad (7.5)$$

This is the full system of 'string-like' *operator* constraints which generalize similar string/puncture equations for correlation functions [24][25] (reduction to the generalized KdV's may be performed).

Based on our previous experience, we could expect that the parameter u, though introduced initially in a purely formal way, would reduce to the coordinate on algebro-geometric solutions. (Note, however, that *solutions* to the Virasoro constraints are "very far" from the algebro-geometric ones.)

We can view the constraint (7.5) similarly to the constraint that singles out the generalized KdV's [3] from the KP hierarchy: there, one imposes $(KD^NK^{-1})_- = 0$, that is, $A = KD^NK^{-1}$ is a *differential* operator. This A parametrizes independent degrees of freedom and therefore 'solves' the constraint. In the present case, we introduce a differential operator A and re-write eqs.(7.5) as

$$K(x)\circ(x+\tfrac{1}{u}J+\sum_{s\geq 2} st_s D^{s-1})D = A\circ K(x+\tfrac{1}{u}) \qquad (7.6)$$

It is understood that for a given model, the times t_r are set to zero for $r\geq N+1$. *Then* the order of A is just N, $A=\sum_{i=0}^{N} a_i D^i$, and eq.(7.6) turns into a very restrictive constraint saying that the argument, x, of K is *translated* by means of N-th order differential operators. More precisely, for K as in (2.2), we get

$$\sum_{j=m}^{N} jt_j w_{j-m}(x)=\sum_{i=m}^{N} a_i\sum_{j=m}^{i}\binom{i}{j}\partial^{i-j}w_{j-m}(x+\tfrac{1}{u}),\quad 2\leq m\leq N \qquad (7.7)$$

$$\sum_{j=1}^{N} jt_j w_{j-1}(x)+J\tfrac{1}{u}=\sum_{i=m}^{N} a_i\sum_{j=1}^{i}\binom{i}{j}\partial^{i-j}w_{j-1}(x+\tfrac{1}{u}), \qquad (7.8^1)$$

$$\sum_{j=1}^{N} jt_j w_j(x)+J\tfrac{1}{u}w_1(x)=\sum_{i=0}^{N} a_i\sum_{j=0}^{i}\binom{i}{j}\partial^{i-j}w_j(x+\tfrac{1}{u}), \qquad (7.8^2)$$

$$\sum_{j=1}^{N} jt_j w_{j-m}(x)+J\tfrac{1}{u}w_{1-m}(x)+mw_{-m}(x)=\sum_{i=0}^{N} a_i\sum_{j=0}^{i}\binom{i}{j}\partial^{i-j}w_{j-m}(x+\tfrac{1}{u}),\quad m\leq -1, \qquad (7.9)$$

The coefficients a_i are easily determined from eqs.(7.7,8) in terms of the first N coefficient functions of K, as

$$Nt_N = a_N, \qquad (7.10)$$

$$(N-1)t_{N-1}+Nt_N w_1(x)=a_{N-1}+a_N w_1(x+\tfrac{1}{u}), \qquad (7.11)$$

$$(N-2)t_{N-2}+(N-1)t_{N-1}w_1(x)+Nt_N w_2(x)$$

$$=a_{N-2}+a_{N-1}w_1(x+\tfrac{1}{u})+a_N\left\{\binom{N}{N-1}\partial w_1(x+\tfrac{1}{u})+w_2(x+\tfrac{1}{u})\right\} \qquad (7.12)$$

and so on. The rest of the constraints, *i.e.*, eqs.(7.9), with the a's thus determined, put severe restrictions on the dependence of K on the hierarchy times (and on x).

Does the equation (7.5) alone *imply* the KP hierarchy? I was not able to show this without *any* additional assumptions[7]. Of course, *assuming* the KP hierarchy and imposing eq.(7.5), one readily discovers that all the $\frac{\partial}{\partial t_r}$ derivatives of eq. (7.5) turn into identities (which is in fact guaranteed by eq.(3.12), but is also easily checked directly). Note that an interesting system of "non-local" evolution equations satisfied by $A=A(x,u;D)\in$Diff follows from (7.6):

$$\frac{\partial A}{\partial t_r}=Q(x)^r_+ A - AQ(x+\tfrac{1}{u})^r_+ \qquad (7.13)$$

Note also that when expanding A in μ, its modes A_r, which are *differential* operators of order $n(r)=1+r(n(1)-1)$, also satisfy the Virasoro algebra.

The form of eq.(7.5) suggests pushing forward the comparison with puncture/string equations [24][25]. Though the latter have been formulated for the KdV hierarchy, we go on with KP, as our main interest is in formal structures that arise on integrable hierarchies. We prefer, however, to add also $\mathscr{L}_{-1}\tau=0$ to the constraints (7.1). Then, instead of (7.5), we end up with

---

[7] I wish to thank G.Moore for his comments on the derivation of KdV equations from the string equation, see [27].

$$(K(P+\tfrac{1}{u}J)e^{D/u}K^{-1})_- = J(KD^{-1}e^{D/u}K^{-1})_- \qquad (7.15)$$

When pulling x out of $K...K^{-1}$, a K'-term appears:

$$\sum_{r\geq 1} rt_r(KD^{r-1}e^{D/u}K^{-1})_- + J(K(u^{-1}-D^{-1})e^{D/u}K^{-1})_- = -(K'e^{D/u}K^{-1})_- \qquad (7.16)$$

$$K' = \frac{\partial K}{\partial D} = \sum_{r\geq 1} w_r(-r)D^{-r-1} \qquad (7.17)$$

Hence the scaling operators are identified *via*

$$\begin{aligned} \mathfrak{P}(u) &\sim J(KD^{-1}e^{D/u}K^{-1})_- - J\tfrac{1}{\mu}(Ke^{D/u}K^{-1})_- - (K'e^{D/u}K^{-1})_-, \\ \mathfrak{S}_r(u) &\sim (KD^{r-1}e^{D/u}K^{-1})_- \end{aligned} \qquad (7.18)$$

The RHS's here should be viewed as 'components' of vector fields on the space of the K's (as explained in Sec.2). The LHS's might have also been written as $\mathfrak{P}(u)\mathfrak{P}(\infty)$ and $\mathfrak{S}_r(u)\mathfrak{P}(\infty)$, which anticipates the following geometrical meaning of the correspondences denoted by $\sim$: Picking up an algebro-geometric solution, we can use the vector fields defined by the RHS's of eqs.(7.18) to deform the Baker-Akhiezer function $\psi(t,z)$. This can be alternatively described as some operator insertions on the Riemann surface. These insertions, as always in the Krichever construction, would sit inside a patch that carries the coordinate system to which z and u now refer. The coordinate system is centered at $\{\infty\}$, which already gives one puncture $\mathfrak{P}(\infty)$. Now performing the contour integral along the boundary of the patch, one imitates an insertion, in addition to $\mathfrak{P}(\infty)$, of a scaling operator. Hence, in particular, the correlation functions

$$\langle \mathfrak{S}_r\mathfrak{P}\rangle = \frac{1}{2\pi i}\oint du\, \mathrm{res}(KD^{r-1}e^{D/u}K^{-1})_- = \mathrm{res}\,Q^r$$
$$\langle \mathfrak{P}\,\mathfrak{P}\rangle = -\partial\log\tau = \mathrm{res}\,K$$

(Note that $\mathrm{res}\,Q = \partial^2\log\tau$, $\partial = \partial/\partial x$).

8.Scaling limit of the Virasoro-constrained Toda hierarchy.

In this Section we demonstrate that when imposing the Virasoro constraints on the two basic examples which have been considered above, the KP and Toda hierarchies, not only do they demonstrate formal similarities, but are actually related by a scaling[8]. We start with adapting our definitions to the study of the scaling. The Toda hierarchy analogue of the Virasoro constraints in the form (7.4) reads

$$\mathfrak{L}_n \equiv \left[W^{(\infty)}\{[J(n+1)+\hat{\mathrm{p}}]\Lambda^n + \sum_{r\geq 1} rx_r\Lambda^{r+n}\}W^{(\infty)-1}\right]_- = 0, \quad n\geq 0 \text{ or } -1 \qquad (8.1)$$

These equations follow from the 'bosonized' form of the Virasoro generators on the phase space. Note that in a more abstract setting, the role of J is just to account for the arbitrariness in the relative ordering of $\hat{\mathrm{p}}$ and Λ^n. To obtain eq.(8.1), it suffices to take eq.(5.14) and use in it eq.(5.10) and the identity

$$W^{(\infty)}\hat{\mathrm{p}}\delta(\mu,\Lambda) = \sum_s |s\rangle\left[(s+\mu\tfrac{\partial}{\partial\mu})w^{(\infty)}(s,\mu)\right]\langle s|\delta(\mu,\Lambda) \qquad (8.2)$$

As we will consider in this section only the $^{(\infty)}$-part, we omit the $^{(\infty)}$ superscript.

8.1. REDEFINITIONS. The flows of the hierarchy being commutative, we use the possibility to linearly combine them, and introduce new times, $\tilde{x}_r$, $r\geq 1$, which will be more appropriate for the subsequent scaling. Note that a redefinition of hierarchy times has been used previously in [29] in order to relate two formulations of 1–dimensional Toda hierarchy; the obvious moral is that various choices of independent times may be appropriate depending on which properties of the hierarchy one wishes to simplify.

We set

[8]Different approaches to the study of the continuum limit of discrete hierarchies can be found in [28].

$$x_r = \frac{1}{r}\sum_{s=r}^{\infty}{}' \begin{bmatrix} s \\ r \end{bmatrix} (-1)^{s+r}(s+1)\tilde{x}_s, \qquad r \geq 1 \tag{8.3}$$

where $\begin{bmatrix} a \\ b \end{bmatrix}$ are the binomial coefficients.

Further, there also is a freedom to linearly combine the constraints (8.1) themselves. We define

$$\mathfrak{L}_{p-1} = \sum_{n=0}^{p}{}' \begin{bmatrix} p \\ n \end{bmatrix} \mathcal{L}_n (-1)^{n+p} \tag{8.4}$$

If there had been no $(\ldots)_-$-projection in eqs.(8.1), one could easily have done the sums which arise upon the substitution of (8.3) and (8.1) into (8.4). However, the main point of the whole construction is just the $(\ldots)_-$-projection; taking it into account is more involved, and that is what we are going to do.

To proceed, we wish to change slightly our interpretation of Λ. In anticipation of the continuum limit to be taken below, we view the vectors $\sum_s |s\rangle v(s)$ (see Sec.4), as functions

$$s \longmapsto v(s) \tag{8.5}$$

defined on $\mathbb{Z}$. Then Λ may be interpreted as the operator e^{∂}, where ∂ is formally $\partial/\partial s$, acting via

$$(e^{\partial} v)(s) = v(s+1) \tag{8.6}$$

Further, the operator $\hat{p}$ introduced in (4.1), becomes the operator of multiplication by the argument:

$$(\hat{p}v)(s) = sv(s); \tag{8.7}$$

abusing the notation, we may write s instead of $\hat{p}$. The dressing operator and its inverse now read,

$$W = \sum_{j\geq 0} w_j(\cdot) e^{-j\partial}, \qquad W^{-1} = \sum_{i\geq 0} e^{-i\partial} \circ w_i^*(\cdot) \tag{8.8}$$

where $w_j(\cdot)$ are the functions $s \mapsto w_j(s)$ which act by a *pointwise* multiplication.

8.2.SCALING. Now by scaling, we make s into a continuous variable: we put

$$s = \frac{t_1}{\epsilon}, \quad \partial = \epsilon D, \quad D \equiv \frac{\partial}{\partial t_1}, \quad \epsilon \longrightarrow 0 \tag{8.9}$$

One may consider ϵ dimensionful, of the dimension of length. Then so is t_1. Now Λ is identified with

$$\Lambda = e^{\epsilon D} \tag{8.10}$$

We also rescale the above $\mathfrak{L}$'s into

$$\tilde{\mathfrak{L}}_{p-1} = \sum_{n=0}^{p}{}' \begin{bmatrix} p \\ n \end{bmatrix} \mathcal{L}_n (-1)^{n+p} \epsilon^{-p+1} \tag{8.11}$$

The algebra of the $\tilde{\mathfrak{L}}$'s (more precisely, of their associated vector fields $\hat{\tilde{\mathfrak{L}}}$ on the space of dressing or Lax operators) reads

$$[[\hat{\tilde{\mathcal{L}}}_p, \hat{\tilde{\mathcal{L}}}_q]] = (p-q)\sum_{r\geq 0}{}' \hat{\tilde{\mathcal{L}}}_r \epsilon^{r-p-q} \begin{bmatrix} 1 \\ r-p-q \end{bmatrix} = (p-q)(\hat{\tilde{\mathcal{L}}}_{p+q} + \epsilon \hat{\tilde{\mathcal{L}}}_{p+q+1}) \tag{8.12}$$

However, the conclusion that the Virasoro commutation relations are recovered as the $\epsilon \to 0$ limit, would be premature, since, as we will see in a moment, the continuum limit implies several other infinitely large rescalings, and these affect the definition of the mapping $\mathcal{L}_n \longmapsto \hat{\mathcal{L}}_n$ and even of the space the $\hat{\mathcal{L}}_n$'s act on. That is, the bracket on the LHS of (8.12) does in a sense depend on ϵ in the course of the continuum limit. But of course, eq.(8.12) is a propitious sign.

Now we insert for the coefficients of $W \equiv W^{(\infty)}$, eqs.(4.3,4), the following scaling ansatz:

$$w_j = \sum_{l\geq 1} \epsilon^l \begin{bmatrix} j+l-1 \\ l-1 \end{bmatrix} k_l, \quad j \geq 1 \tag{8.13}$$

where k_i will become the coefficient functions of the KP operator K, denoted as w_i in (2.2). Then

$$W = k_0 + \sum_{l\geq 1} k_l \epsilon^l (1 - e^{-\epsilon D})^{-l} \tag{8.14}$$

where

$$k_0 = 1 - \sum_{l\geq 1} k_l \epsilon^l \tag{8.15}$$

Similarly, we set

$$w_i^* = \sum_{m\geq 1} \epsilon^m \binom{i+m-1}{m-1} k_m^*, \quad i\geq 1 \tag{8.16}$$

for some k_m^*. Finally, let

$$\tilde{x}_r = \frac{t_r}{\epsilon^{r+1}}, \quad r \geq 1 \tag{8.17}$$

and consider in the $\epsilon \longrightarrow 0$ limit all the k_l's and k_m^*'s finite as functions of (t_1, t_r). Then it follows that

$$W \longrightarrow K = 1 + \sum_{l\geq 1} k_l D^{-l}, \quad W^{-1} \longrightarrow \tilde{K} = 1 + \sum_{m\geq 1} D^{-m} \circ k_m^* \tag{8.18}$$

Further, the bilinear identity $WW^{-1} = 1$ yields as its $\epsilon \longrightarrow 0$ limit the bilinear identity for ψDiff operators, $K\tilde{K} = 1$ hence

$$\tilde{K} = K^{-1} \tag{8.19}$$

Next we wish to see what becomes of the constraints (8.11) after scaling. The result is given by eq.(8.36) below. As noted above, a subtlety lies in the $(\ldots)_-$-projection onto lower triangular matrices. We get in the end the Virasoro constraints for the KP hierarchy, which are written in terms of ψDiff operators K and K^{-1}, and involve a quite different operation (denoted by the same symbol $(\ldots)_-$, though) of projecting onto purely integral operators. The two $(\ldots)_-$'s being unrelated, one has to start from the definition of the matrix operation $(\ldots)_-$. That is, we write eqs.(8.1) as

$$\sum_{\substack{i,\ j\geq 0 \\ i+j\geq n+1}} w_j \left(J(n+1) + \frac{t_1}{\epsilon} - j\right) e^{-j\epsilon D} e^{n\epsilon D} e^{-i\epsilon D} \circ w_i^* + \sum_{r\geq 1} r x_r \sum_{\substack{i,\ j\geq 0 \\ i+j\geq n+r+1}} w_j e^{-(j+i-r-n)\epsilon D} \circ w_i^* = 0 \tag{8.20}$$

On the LHS of eq.(8.20) we perform the substitutions (8.13) and (8.16); we also insert (8.3) for the x's and go over to the $\tilde{\Omega}$'s as defined by eq.(8.11). The summation symbols thus proliferate, and one has to be careful to defer the summation of infinite power series in $e^{\epsilon D}$ till the very end, so as not to affect the convergence properties; first, one does all the *finite* sums.

Thus, consider the contribution to $\tilde{\Omega}_{p\text{-}l}$, proportional to $(r+1)t_{r+1}$. As a result of all the above substitutions, it becomes[9]

$$\sum_{n=0}^{p} (-1)^{n+p} \binom{p}{n} \sum_{s=1}^{r} (-1)^{s+r} \binom{r}{s} \epsilon^{-r-1} \times \times \sum_{\substack{i,\ j\geq 0 \\ i+j\geq n+s+1}} \sum_{l\geq 0} \sum_{m\geq 0} \epsilon^{l+m} k_l \binom{j+l-1}{l-1} e^{-j\epsilon D} e^{(n+s)\epsilon D} e^{-i\epsilon D} \circ \binom{i+m-1}{m-1} k_m^* \tag{8.21}$$

where the constraint $i+j \geq n+s+1$ singles out just the $(\ldots)_-$-part.

First we perform the summation over all i and j such that $i+j = b$ with b temporarily fixed. This is easily done with the help of the formula

$$\sum_{\substack{i,\ j\geq 0 \\ i+j=b}} \binom{j+l-1}{l-1} \binom{i+m-1}{m-1} = \binom{l+m+b-1}{l+m-1} \tag{8.22}$$

Similarly, we sum over the values of n and s satisfying $n+s=a$:

$$\sum_{s=1}^{r} \sum_{n=0}^{p} (-1)^{n+p} \binom{p}{n} (-1)^{s+r} \binom{r}{s} = \sum_{a=1}^{p+r} \binom{p+r}{a} (-1)^{r+p+a} \tag{8.23}$$

(this follows, in fact, as a result of several elementary combinatorial identities).

Thus eq.(8.21) becomes,

$$(-1)^{r+p} \epsilon^{-p-r} \sum_{l,\, m\geq 0} \epsilon^{l+m} k_l \sum_{c\leq 0} e^{c\epsilon D} \sum_{a=0}^{p+r} (-1)^a \binom{p+r}{a} \binom{l+m-1-c+a}{l+m-1} k_m^* \tag{8.24}$$

where we have performed yet another change of the summation indices $a - b = c$.

Now we are approaching the main point of the calculation: The sign-

[9] The subsequent manipulations are simplified by extending eqs.(8.13) and (8.16) to the values $j = 0$ and $i=0$, and the range of the summation indices, to $l=0$, $m=0$, with $\binom{l-1}{l-1}$ being always set to 1.

alternating sum over a of the two binomial coefficients equals

$$(-1)^{p+r}\begin{bmatrix} l+m-1-c \\ l+m-1-p-r \end{bmatrix} \tag{8.25}$$

which vanishes unless $l+m-1 \geq p+r$. (The vanishing condition can be stated as the identity

$$\sum_{k=0}^{n}{}'(-1)^k\begin{bmatrix} n \\ k \end{bmatrix}\begin{bmatrix} q+m+k \\ m \end{bmatrix} = 0 \quad \text{for} \quad m \leq n\ )$$

Therefore, the range of summation over l and m is constrained by

$$l+m \geq p+r+1 \tag{8.26}$$

Now it is straightforward to bring this short calculation to the end: using eq.(8.25), the sum over $c \equiv -k$ in eq.(8.24) becomes[10]

$$\frac{1}{(l+m-1-p-r)!}\sum_{k\geq 0}\frac{(l+m-1+k)!}{(p+r+k)!}e^{-k\epsilon D}$$

$$= \frac{1}{(l+m-1-p-r)!}\frac{(l+m-1)!}{(p+r-1)!}(1-e^{-\epsilon D})^{p+r-l-m}\sum_{k=0}^{l+m-p-r-1}{}'(-1)^k\begin{bmatrix} l+m-p-r-1 \\ k \end{bmatrix}\frac{e^{-k\epsilon D}}{p+r+k} \tag{8.27}$$

The finite sum over k here does not produce any zeroes or singularities as $\epsilon \longrightarrow 0$; we therefore replace it with its $\epsilon \longrightarrow 0$ limit, *i.e.*

$$\sum_{k=0}^{l+m-p-r-1}{}'(-1)^k\begin{bmatrix} l+m-p-r-1 \\ k \end{bmatrix}\frac{1}{p+r+k} = \frac{(l+m-p-r-1)!}{(p+r)(p+r+1)\ldots(l+m-1)} \tag{8.28}$$

---

[10] Note that had there been $l+m < p+r+1$ in (8.27), the sum over k would have behaved quite differently,

$$\sum_{k=0}^{\infty}{}'\frac{(l+m-1+k)!}{(p+r+k)!}e^{-k\epsilon D} = \frac{(l+m-1)!}{(p+r-1)!}\begin{bmatrix} p+r-1 \\ l+m-1 \end{bmatrix}\Bigg\{-e^{(p+r)\epsilon D}(e^{-\epsilon D}-1)^{p+r-l-m}\times$$

$$\times\ln(1-e^{-\epsilon D}) + e^{(l+m-1)\epsilon D}\sum_{k=1}^{p+r-l-m+1}{}'(-1)^k\begin{bmatrix} p+r-l-m \\ k-1 \end{bmatrix}\sum_{i=1}^{l+m+k-2}{}'\frac{1}{i}e^{(k-i)\epsilon D}\Bigg\}$$

This immediately brings eq.(8.27) to the form

$$(1-e^{-\epsilon D})^{p+r-l-m} \tag{8.29}$$

and therefore eq.(8.24) becomes

$$\sum_{\substack{l,\,m\geq 0 \\ l+m\geq p+r+1}}{}'\epsilon^{-p-r+l+m}k_l(1-e^{-\epsilon D})^{p+r-l-m}k_m^*$$

The $\epsilon \longrightarrow 0$ limit of this is just

$$\sum_{\substack{l,\,m\geq 0 \\ l+m\geq p+r+1}}{}'k_l D^{p+r-l-m}\circ k_m^* = (KD^rD^pK^{-1})_- \tag{8.30}$$

where $(\ldots)_-$ now refers to the algebra of ψDiff operators and means projecting onto the purely integral part. Recall that we have been considering the coefficient in front of $(r+1)t_{r+1}$. Thus we have reproduced the contribution

$$(K\sum_{r\geq 2}rt_rD^{r-1}D^pK^{-1})_- \tag{8.31}$$

to the $(p-1)^{th}$ Virasoro constraint in the *continuum* theory, see eq.(7.4). We have not obtained yet the first term in the sum over r which is proportional to t_1, the time which is singled out from the rest of the t's by the fact that $D = \partial/\partial t_1$, nor the other two terms on the LHS of (7.4) for $r=p-1$. All these come from the $\hat{p}$–term in eq.(8.1) for the discrete theory. To see this, consider what happens to the first sum over i,j in (8.20) under the substitutions (8.13), (8.16) and (8.11). As before, we first do the sum over i and j with a fixed value of $i+j = a$:

$$\sum_{\substack{i,\,j\geq 0 \\ i+j=a}}{}'\begin{bmatrix} l-1+j \\ l-1 \end{bmatrix}[Jn+J+\frac{t_1}{\epsilon}-j]\begin{bmatrix} m-1+i \\ m-1 \end{bmatrix}$$

$$= (Jn+J+\frac{t_1}{\epsilon})\begin{bmatrix} l+m-1+a \\ l+m-1 \end{bmatrix} - l\begin{bmatrix} l+m-1+a \\ l+m \end{bmatrix} \tag{8.32}$$

In accordance with the definition of the $\tilde{\Sigma}$'s, eq.(8.11), the next summation to be performed is over n. Note that the $(\ldots)_-$-projection meant $a \geq n+1$. Denoting $a=n+k$, we get from the last equation

$$\sum_{k=1}^{\infty}\epsilon^{-p+1}(-1)^p\sum_{n=0}^{p}(-1)^n\binom{p}{n}\left\{Jn\binom{l+m-1+k+n}{l+m-1}\right.$$
$$\left.+\left(J+\tfrac{t_1}{\epsilon}\right)\binom{l+m-1+k+n}{l+m-1}-l\binom{l+m-1+k+n}{l+m}\right\}e^{-k\epsilon D}$$

$$=\sum_{k=1}^{\infty}\epsilon^{-p+1}\left\{Jp\binom{l+m+k}{l+m-p}+\left(J+\tfrac{t_1}{\epsilon}\right)\binom{l+m-1+k}{l+m-1-p}-l\binom{l+m-1+k}{l+m-p}\right\}e^{-k\epsilon D} \tag{8.33}$$

Again, we see that $l+m$ gets restricted by $l+m \geq p$ or $p+1$. The series over k can be summed up as above, with the leading singularity as $\epsilon \to 0$ coming from

$$(Jp-l)\frac{\epsilon^{-p+1}}{(1-e^{-\epsilon D})^{l+m-p+1}}+\frac{t_1}{\epsilon}\frac{\epsilon^{-p+1}}{(1-e^{-\epsilon D})^{l+m-p}} \tag{8.34}$$

whence the contribution to $\tilde{\Omega}_{p-1}$ reads

$$\sum_{\substack{l,m\\ l+m\geq p}}(Jp-l+t_1D)k_lD^{-l-m+p-1}\circ k_m^* = \sum_{l,m}\left[k_lD^{-l}(Jp+t_1D)D^{p-1}D^{-m}\circ k_m^*\right]_-$$

$$=(K(Jp+t_1D)D^{p-1}\circ K^{-1})_- \tag{8.35}$$

Putting eqs.(8.31) and (8.35) together, we arrive at the Virasoro-constrained KP hierarchy[11]:

$$(K((J-1)p+N+\sum_{r\geq 1}rt_rD^r)D^{p-1}K^{-1})_- = 0,\ N=J \tag{8.36}$$

Finally, let us see what becomes of the ***linear problem*** (4.10) in the continuum limit. In accordance with eq.(8.5), $|u_\lambda\rangle$ is identified with the function $s\mapsto\lambda^s$, or after rescaling,

$$t_1\mapsto e^{zt_1}\ \text{with}\ \lambda=e^{\epsilon z} \tag{8.37}$$

[11]For the origin and geometrical meaning of the arbitrariness of adding the term proportional to $N\equiv a_0+\frac{1}{2}$, see ref.[6]. In Sec.2 N was set to zero.

Then the RHS of (4.11) assumes the form

$$\sum_{l\geq 0}\epsilon^{l}k_l\sum_{j\geq 0}e^{-j\epsilon D}\binom{j+l-1}{l-1}.e^{zt_1}=\sum_{l\geq 0}\epsilon^{l}k_l\frac{1}{(1-e^{-\epsilon D})^l}e^{zt_1}\longrightarrow\sum_{l\geq 0}k_lD^{-l}.e^{zt_1}$$
$$\equiv e^{-\sum_{r\geq 2}t_rz^r}\psi_{KP}(t,z) \tag{8.38}$$

As the exponential in (8.38) commutes with D, it follows that ψ_{KP} satisfies the KP linear problem [1]

$$Q\psi_{KP}(t,z)=z\psi_{KP}(t,z) \tag{8.39}$$

Therefore, ψ_{KP} obtained from the scaling limit coincides with the wave function of the KP hierarchy.

9.Discussion.

There are several directions in which the above considerations can be developed. First, the fact that the hierarchy times enter into the Virasoro hierarchy only *linearly*, makes it straightforward to study reductions of the Virasoro-constrained hierarchies. Many possibilities are open starting from the KP and Toda Virasoro-constrained hierarchies. It should be noted that when working in terms of the phase space constraints, the tau function of the reduced hierarchy becomes a secondary issue: one first performs the reduction and then derives (or, at worst, guesses) the hierarchy compatible with the reduced phase space Virasoro constraints.

The two-dimensional Toda hierarchy contains two infinite sets of the times, x and y. This allows one to view the tau function as a "vacuum vector" of any of the conformal theories expressible through two scalar fields. Leaving aside the question of the topological-theory counterpart of such constrained hierarchies, it is nevertheless interesting whether one can generalize the technique of Sec.5 to the case of constraints with mixed $\partial/\partial x$- and $\partial/\partial y$-derivatives acting on the tau function and translate them to the phase space. There might exist, for instance, a W_3- generalization of the Virasoro algebra representation (5).

ACKNOWLEDGEMENTS. I am grateful to I.M.Krichever T.Takebe and A.Zabrodin for valuable discussions. I thank A.Mironov for cooperation in typesetting. It is also a pleasure to aknowledge kind hospitality extended to me by the organizers of the Karpacz School where these lectures were presented.

References

[1] E.Date, M.Jimbo, M.Kashiwara, T.Miwa, *Transformation Groups for Soliton Equations*, in: Proc. RIMS Symp. on Non–Linear Integrable systems, M.Jimbo and T.Miwa (eds.), World Science, Singapore 1983, p.39–119.

[2] K.Ueno, K.Takasaki, *Toda Lattice Hierarchy*, in: Adv. Studies Pure Math. 4 (1984) 1–95.

[3] V.G.Drinfeld, V.V.Sokolov, *Lie Algebras and Equations of the Korteweg-De-Vries type*, in: Sovrem. Probl. Mat. 24, Moscow, Nauka 1984, p.81–178.

[4] I.M.Krichever, *Methods of Algebraic Geometry in the Theory of Non–Linear Equations*, Usp. Mat. Nauk (1977) 32:6, 183–208; *Integrating Non–Linear Equations by the Methods of Algebraic Geometry*. Funk. An. Prilozh. 11:2 (1977), 15–32.

[5] A.A.Beilinson, Yu.I.Manin, V.V.Schechtman, in: Lect. Notes Math. 1289 (Springer-Verlag 1987), 52

[6] A.M.Semikhatov, *Soliton Equations and the Virasoro Algebra*, Int. J. Mod. Phys. A 4 (1989) 467–479.

[7] M.A.Semenov-Tian-Shansky, *What is a Classical r-Matrix*, Funk. An. Prilozh., 17 (1983) 17–33.

[8] E.Brezin, V.A.Kazakov, *Exactly Solvable Field Theories of Closed Strings*, Phys. Lett. B 236 (1990) 144;

M.R.Douglas, S.H.Shenker, *Strings in Less Than One Dimension*, Nucl. Phys. B335 (1990) 635–654;

D.J.Gross, A.A.Migdal, *Nonperturbative Two Dimensional Quantum Gravity*, Phys. Rev. Lett. 64 (1990) 127.

[9] D.J.Gross, A.A.Migdal, *A Nonperturbative Treatment of Two Dimensional Quantum Gravity*, PUPT–1159;

T.Banks, M.R.Douglas, N.Seiberg, S.H.Shenker, *Microscopic and Macroscopic Loops in Non-Perturbative Two Dimensional Gravity*, RU–89–50;

M.R.Douglas, *Strings in Less Than One Dimension and Generalized KdV Hierarchies*, Phys. Lett. B238 (1990) 176–180.

[10]M.Fukuma, H.Kawai, R.Nakayama, *Continuum Schwinger-Dyson Equations and Universal Structures in 2D Quantum Gravity*, UT–562/KEK-TH-251;

R.Dijkgraaf, E.Verlinde, H.Verlinde, *Loop Equations and Virasoro Constraints in Non–Perturbative 2d Quantum Gravity*, PUPT-1184.

[11]A.Gerasimov, A.Marshakov, A.Mironov, A.Morozov, A.Orlov, *Matrix Models of 2D Gravity and Toda theory*, Lebedev Inst. prepr., July 1990.

[12] E.J.Martinec, *On the Origin of Integrability in Matrix Models*, EFI-90-67.

[13]A.M.Semikhatov, *Conformal Fields: from Riemann Surfaces to Integrable Hierarchies*, Lebedev Inst. preprint (November 1990).

[14] A.G.Reiman, M.A.Semenov–Tian–Shansky, *Hamiltonian Structure of Eqwuations of the Kadomtsev-Petviashvili Type*, Zap. Nauch. Sem. LOMI, 133 (1984) 212–227.

[15] D.H.Friedan, E.J.Martinec, S.H.Shenker, *Conformal Invariance, Supersymmetry and String Theory*, Nucl. Phys. B 271 (1986) 93–165.

[16] A.Yu.Orlov, *Hamiltonian Formalism and Gelfand-Dikii Identities for (2+1)-D Integrable Systems*, Moscow preprint (1988)

[17] P.G.Grinevich, A.Yu.Orlov, *Flag Spaces in KP Theory and Virasoro Action on DET$\bar{\partial}_j$ and Segal–Wilson Tau–Function*, Moscow preprint (1990).

[18]I.M.Krichever, S.P.Novikov, *Virasoro–Type Algebras, Riemann Surfaces and Structures of Soliton Theory*, Funk. An. Prilozh. 21:2 (1987) 46–63.

[19] A.M.Semikhatov, *Bosonization, τ–Functions and Operator Formalism on Riemann Surfaces*, Nucl. Phys. B 315 (1989) 222–248.

[20] I.M.Krichever, *Spectral Theory of Two–Dimensional Periodic Operators and its Generalizations*, Usp. Mat. Nauk, 44:2 (1989) 121–184.

[21] A.Mironov, A.Morozov, *On the Origin of Virasoro Constraints in Matrix Models: a Lagrangian Approach*, Lebedev Inst. preprint (July 1990).

[22] P.Di Francesco, D.Kutasov, *Unitary Minimal Models Coupled to 2D Quantum Gravity*, PUPT–1173.

[23] P.Ginsparg, M.Goulian, M.R.Plesser, J.Zinn-Justin, *(p,q) String Actions*, HUTP-90/A015 – PhT/90-049;

A.Jevicki, T.Yoneya, *Action Principle for Strings in Less Than One Dimension*, Mod. Phys. Lett. A5 (1990) 1615–1621.

[24] R.Dijkgraaf, E.Witten, *Mean Field Theory, Topological Field Theory and Multi-Matrix Models*, Nucl. Phys. B342 (1990) 486–522.

[25] E.Witten, *Two Dimensional Gravity and Intersection Theory on Moduli Space*, IASSNS-HEP-90/45

[26]I.M.Krichever, *Methods of Algebraic Geometry in the Theory of Non-Linear Equations*, Usp. Mat. Nauk, 32:6 (1977) 183 – 208.

[27] G.Moore, *Geometry of the String Equations*, YCTP–P4–90

[28] Yu.Makeenko, A.Marshakov, A.Mironov, A.Morozov, *Continuum Versus Discrete Virasoro in 1-Matrix Models*. Lebedev Inst. preprint (Nov. 1990).

L.Alvarez-Gaumé, C.Gomez, J.Lacki, *Integrability in Random Matrix Models*, CERN-TH.5875/90.

[29]T.Takebe, *Toda Lattice Hierarchy and Conservation Laws*, Commun. Math. Phys. 129 (1990) 281–318.

[30] B.A.Dubrovin, *Theta–Functions and Non–Linear Equations*, Usp. Mat. Nauk 36:2 (1981) 11–80.

T.Shiota, *Characterization of Jacobian Varieties in Terms of Solution Equations*. Invent. Math. 83 (1986) 333–382.

Integrable Hamiltonian Systems and their Fermionic and Supersymmetric Extensions

Allan P. Fordy
Department of Applied Mathematical Studies and
Centre for Nonlinear Studies,
University of Leeds, Leeds LS2 9JT, UK

Abstract

The purpose of this article is to give a brief introduction to and survey of the Hamiltonian theory of 'soliton equations'. Emphasis is on the relationship with the linear spectral problem and the systematic construction of Hamiltonian structure (Poisson bracket), constants of motion and corresponding commuting Hamiltonian flows. The role of gauge transformations and Miura maps as Poisson mappings will be emphasised. Fermionic and supersymmetric extensions are presented, using the sKdV equations and their 'modifications' as examples.

1. Introduction

Interest in completely integrable Hamiltonian systems has been rekindled in recent years, following the discovery of the soliton. Perhaps the most important feature of the soliton equation is its relationship with a linear spectral problem, from which all other important properties follow. In this paper I emphasise the close relationship between the linear spectral problem and the Hamiltonian properties.

First, I present a general synopsis of the relevant parts of Hamiltonian theory, starting with the classical (finite dimensional) theory [1] and emphasising those features which carry over to infinite dimensions [2,3], putting particular emphasis on the bi-Hamiltonian property.

Section 3 introduces the idea of an isospectral flow. It is shown how to systematically construct the Poisson bracket(s) corresponding to the isospectral flows and the infinite number of involutive constants of motion. Section 4 introduces the notion of a Miura map, giving rise to related integrable systems and co-ordinates in which

the Hamiltonian structures take simplified form. The main examples of sections 3 and 4 are coupled KdV systems associated with spectral dependent Schrödinger operators

In section 5, I apply the methods of this paper to some fermionic and supersymmetric extensions of the spectral problem of section 3. The fermionic extensions were first presented in [4], where more details can be found. The supersymmetric results are only preliminary, the work having been stimulated by discussions at this Winter School. Nevertheless, a systematic construction of the Hamiltonian structures of the "N=1" supersymmetric KdV equation is presented and points the way to several possible generalisations.

2. Hamiltonian Theory

When generalising from finite to infinite dimensional Hamiltonian systems we use the Poisson bracket rather than symplectic formulation of Hamiltonian mechanics.

The canonical Poisson bracket of two function f and g:

$$\{f,g\} = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right) \tag{2.1a}$$

is skew symmetric and bilinear and satisfies the Jacobi identity:

$$\{\{f,g\},h\} + \{\{g,h\},f\} + \{\{h,f\},g\} = 0 \tag{2.1b}$$

and the Leibnitz rule :

$$\{fg,h\} = \{f,h\}g + f\{g,h\}\ , \tag{2.1c}$$

where f,g,h are smooth functions. Hamilton's (canonical) equations, generated by Hamiltonian function $h(q_1,\ldots,p_n)$ are:

$$\dot{\mathbf{x}} = J\nabla h \tag{2.2}$$

where $\mathbf{x}=(q_1,\ldots,q_n,p_1,\ldots,p_n)^T$, $J=\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$, $\nabla h=\left(\frac{\partial h}{\partial q_1},\ldots,\frac{\partial h}{\partial p_n}\right)^T$. It is important to realise that any function h will generate a Hamiltonian flow (2.2). Using the chain rule, it is easy to see that:

$$\frac{df}{dt} = \{f,h\} \tag{2.3}$$

where f is a function of the 2n co-ordinates (q_i,p_i) but not explicitly of t. Thus any function which Poisson commutes with h is a constant of the motion (2.2). Two very important properties follow

from the Jacobi identity:

(i) the Poisson bracket $\{f,g\}$ of 2 constants of the motion f and g is again a constant of the motion.

(ii) Poisson commuting functions generate commuting Hamiltonian vector fields and Hamiltonian flows.

The natural (physically or geometrically defined) co-ordinates on phase space are not always canonical. However, we can abstractly define a Poisson bracket to be any skew symmetric, bilinear quadratic form on functions f and g satisfying (2.1b,c). With local co-ordinates $x_1,\ldots,x_N$ the Poisson bracket of two arbitrary smooth functions is given by:

$$\{f,h\} = \frac{\partial f}{\partial x_i} J_{ij} \frac{\partial h}{\partial x_j} = (\nabla f)^T J(\nabla h) \ , \qquad (2.4a)$$

where the structure functions J_{ik} are defined by:

$$J_{ik} = \{x_i, x_k\}. \qquad (2.4b)$$

In order that this Poisson bracket satsify the Jacobi identity (2.1b), the matrix elements J_{ik} must satisfy:

$$J_{im} \frac{\partial J}{\partial x_m}_{jk} + J_{jm} \frac{\partial J}{\partial x_m}_{ki} + J_{km} \frac{\partial J}{\partial x_m}_{ij} = 0. \qquad (2.4c)$$

Hamilton's equations now take the form :

$$\frac{dx_i}{dt_h} = \{x_i, h\} = J_{ik} \frac{\partial h}{\partial x_k} \ , \qquad (2.5)$$

where t_h denotes the time parameter along the Hamiltonian flow generated by h. Once again, we have (2.3) and the consequent importance of Poisson commuting functions. Any function $\mathcal{C}$ which Poisson commutes with every function on phase space M is called a Casimir function of the Poisson bracket. The gradient of a (non-constant) Casimir function is a null vector of matrix J.

Example: Lie-Poisson bracket

Associated with any Lie algebra we have the Lie-Poisson bracket, with structure functions :

$$J_{ij} = C^k_{ij} x_k \ , \qquad (2.6)$$

where C^k_{ij} are the structure constants of the Lie algebra *wrt* some basis. Furthermore, any Poisson bracket whose structure functions are linear in the phase space co-ordinates is the Lie-Poisson bracket corresponding to some Lie algebra.

Darboux Theorem

About each point in phase space M at which rank J = 2n, there exists a set of co-ordinates $q_1,\ldots,q_n,p_1,\ldots,p_n,z_1,\ldots,z_s$ (2n+s=N) with z_i being Casimirs and (q_i,p_i) being canonical. Thus, in finite dimensions, every Hamiltonian system (2.5) can be reduced to canonical form.

Poisson Maps

Suppose we have 2 Poisson manifolds M and N, with respective Poisson brackets $\{,\}_M$ and $\{,\}_N$ and Poisson matrices J_M and J_N.

A map $\varphi: M \to N$ is said to be a Poisson map if

$$\varphi^*\{f,g\}_N = \{\varphi^* f, \varphi^* g\}_M \ , \qquad (2.7a)$$

where $f,g: N \to \mathbb{R}$, $\varphi^* f = f \circ \varphi: M \to \mathbb{R}$. Let $d\varphi$ denote the Jacobian of this map. Then:

$$J_N = d\varphi J_M (d\varphi)^T \ . \qquad (2.7b)$$

In this paper, Poisson maps arise in the guise of Miura maps and are discussed in section 4.

We are mainly interested in completely integrable Hamiltonian systems. For a Hamiltonian system to be called completely integrable we require that the equations of motion can be solved (integrated) up to a finite number of quadratures. This leads us to Liouville's theorem [1]:

Theorem

Let $h = f_1, f_2, \ldots, f_n$ be n independent functions in involution (meaning that they mutually Poisson commute) on a 2n-dimensional symplectic manifold :

$$\{f_i, f_j\} = 0 \qquad i,j, = 1,\ldots,n.$$

Then Hamilton's equations (2.2) can be integrated by quadratures.

Infinite Dimensional Hamiltonian Systems

We now modify the above theory to apply to nonlinear evolution

equations (NLEEs) in (1+1)-dimensions. We wish our NLEE to be written in the form (2.5):

$$u_t \equiv \partial_{\mathcal{H}} u = B\delta\mathcal{H} \quad \text{or} \quad u_{it} = B_{ij}\delta_j\mathcal{H}\ ,\ i=0,\ldots,N-1\ , \tag{2.8a}$$

where the Poisson matrix J_{ij} has now been replaced (in general) by a skew symmetric matrix of differential operators, B_{ij}, and the gradient ∇h is replaced by the variational derivative $\delta\mathcal{H}$ of a density $\mathcal{H}$, defined by:

$$\delta\mathcal{H} \equiv \delta_u\mathcal{H} = (\delta_0\mathcal{H},\ldots,\delta_{N-1}\mathcal{H})^T\ ,\ \delta_i\mathcal{H} \equiv \frac{\delta\mathcal{H}}{\delta u_i} = \sum_{k=0}^{\infty}(-\partial)^k\frac{\partial\mathcal{H}}{\partial u_i^{(k)}}\ , \tag{2.8b}$$

where $u_i^{(k)} = (\partial)^k u_i$.

Remark

As usual, since $\delta f=0$ iff $f\in \mathrm{Im}\partial$ (i.e. an exact x-derivative), we have the equivalence relation:

$$f \sim g \text{ iff } f-g \in \mathrm{Im}\partial \text{ iff } \delta f = \delta g\ .$$

The equivalence class to which a given function H belongs is denoted by $\mathcal{H}$, which is called a density. We can avoid the discussion of boundary conditions if we write all our formulae in terms of densities, but will have to replace 'strict zero' by '$0(\mathrm{mod}\mathrm{Im}\partial)$' in several places.

We imitate formula (2.4) by the quadratic form:

$$\{\mathcal{K},\mathcal{H}\} = \frac{\delta\mathcal{K}}{\delta u_i}B_{ij}\frac{\delta\mathcal{H}}{\delta u_j}\ . \tag{2.9a}$$

The skew symmetry of matrix B_{ij} implies:

$$\{\mathcal{K},\mathcal{H}\} + \{\mathcal{H},\mathcal{K}\} = 0\ (\mathrm{mod}\mathrm{Im}\partial)\ . \tag{2.9b}$$

For (2.9a) to define a Poisson bracket, we also require the Jacobi identity to be satisfied:

$$\{\{\mathcal{H},\mathcal{J}\},\mathcal{K}\} + \{\{\mathcal{K},\mathcal{H}\},\mathcal{J}\} + \{\{\mathcal{J},\mathcal{K}\},\mathcal{H}\} = 0\ (\mathrm{mod}\mathrm{Im}\partial)\ . \tag{2.9c}$$

This imposes a strong restriction (analogous to (2.4c)) on the components of B (see [2]). When (2.9b,c) are satisfied, (2.9a) defines a Poisson bracket and B is called a Hamiltonian structure (or operator) and said to possess the Hamiltonian property. Note that constant coefficient, skew symmetric operators are always Hamiltonian.

Conservation Laws

Let $\mathcal{K}$ be any density. It follows from (2.8) and definition (2.9a) that:

$$\mathcal{K}_t = \{\mathcal{K},\mathcal{H}\} + F_x\ , \tag{2.10a}$$

with the exact derivative F_x arising through a series of integrations by parts. $\mathcal{K}$ and $\mathcal{H}$ will be said to Poisson commute if $\{\mathcal{K},\mathcal{H}\}=0$ $(\mathrm{mod}\mathrm{Im}\partial)$. In such a case we have the local conservation law:

$$\mathcal{K}_t = \mathcal{F}_x \tag{2.10b}$$

where $\mathcal{F}_x = \{\mathcal{K},\mathcal{H}\}+F_x$. With appropriate boundary conditions this gives:

$$\frac{d}{dt}\int\mathcal{K}dx = 0\ . \tag{2.10c}$$

It follows from the Jacobi identity (2.9c) that Poisson commuting Hamiltonian densities $\mathcal{H}$ and $\mathcal{K}$ generate commuting Hamiltonian flows.

Bi-Hamiltonian Systems

A system of evolution equations is said to be bi-Hamiltonian if there exist two Hamiltonian operators B_0 and B_1 and two Hamiltonians $\mathcal{G}$ and $\mathcal{H}$ such that :

$$u_t = B_0\delta\mathcal{G} = B_1\delta\mathcal{H}. \tag{2.11}$$

It is particularly interesting if the operator $B_0 + B_1$ is also Hamiltonian, in which case B_0 and B_1 are said to be compatible (in general the sum of the Poisson brackets would fail to satisfy the Jacobi identity). The importance of compatibility is that it enables us (under certain conditions) to construct an infinite hierarchy of (Poisson commuting) Hamiltonians [5].

Lemma 2.1

If B_0 and B_1 are compatible Hamiltonian operators, with B_0 non-degenerate, and :

$$B_1\delta\mathcal{G} = B_0\delta\mathcal{H}\ ,\quad B_1\delta\mathcal{H} = B_0K\ , \tag{2.12a}$$

then there exists a function $\mathcal{K}$ s.t. $K = \delta\mathcal{K}$.

To prove the existence of an infinite hierarchy of Hamiltonians, $\mathcal{H}_n$, related to compatible Hamiltonian operators B_0, B_1, we need to check that two conditions hold :

(i) $\exists$ an infinite sequence of vector functions $K_0, K_1, \ldots$ satisfying

$$B_1 K_n = B_0 K_{n+1}, \tag{2.12b}$$

(ii) $\exists$ two function(al)s $\mathcal{H}_0$ and $\mathcal{H}_1$ s.t.

$$K_0 = \delta\mathcal{H}_0, \quad K_1 = \delta\mathcal{H}_1.$$

It then follows from the Lemma that there exist function(al)s $\mathcal{H}_n$ s.t.

$$K_n = \delta\mathcal{H}_n \quad \forall\ n \geq 0\ . \tag{2.12c}$$

Remark

Condition (i) is not always easy to check, although it is for the systems discussed in this paper. Indeed, it may not even be satisfied, as shown by an example of Kupershmidt [6].

Theorem 2.1 [5]

Let $\{,\}_0$ and $\{,\}_1$ denote the Poisson brackets defined respectively by B_0 and B_1, which are assumed to be compatible. Let $\mathcal{H}_0, \mathcal{H}_1, \ldots$ be a sequence of functions defined by (2.12b,c). Then :

$$\{\mathcal{H}_n, \mathcal{H}_m\}_0 = \{\mathcal{H}_n, \mathcal{H}_m\}_1 = 0 \qquad \forall\ n,m \geq 0\ . \tag{2.13}$$

Proof

Let $m > n$. If $m=n+2k$, k a positive integer, then :

$$\{\mathcal{H}_n, \mathcal{H}_m\}_0 = \{\mathcal{H}_n, \mathcal{H}_{m-1}\}_1 = -\{\mathcal{H}_{m-1}, \mathcal{H}_n\}_1 = -\{H_{m-1}, \mathcal{H}_{n+1}\}_0$$

$$= \{\mathcal{H}_{n+1}, \mathcal{H}_{m-1}\}_0 = \ldots = \{\mathcal{H}_{n+k}, \mathcal{H}_{m-k}\}_0 = 0\ ,$$

since $n+k = m-k$. The cases when $(m-n)$ is an odd integer and/or starting with $\{\mathcal{H}_n, \mathcal{H}_m\}_1$ are similarly proved.

We can use (2.12b) to define an evolution parameter t_n by :

$$u_{t_n} = B_1 K_n = B_0 K_{n+1} \tag{2.14}$$

Members of this sequence of bi-Hamiltonian flows mutually commute.

Example

The KdV equation :

$$u_t = u_{xxx} + 6uu_x \tag{2.15a}$$

can be written in Hamiltonian form in 2 distinct ways :

$$u_t = B_0\delta(u^3 - \tfrac{1}{2}u_x^2) = B_1\delta(\tfrac{1}{2}u^2)\ , \tag{2.15b}$$

where $B_0 = \partial$ and $B_1 = \partial^3 + 4u\partial + 2u_x$. These Hamiltonian structures are

easily seen to be compatible [2]. Furthermore, starting with the Kernel of B_0 we can easily generate an infinite sequence $K_0, K_1, \ldots$, (see [2,3]):

$$K_0 = \tfrac{1}{2}, \quad K_1 = u, \quad K_2 = u_{xx} + 3u^2, \ldots \tag{2.15c}$$

satisfying (2.12b) with $B_0 K_0 = 0$.

Since $K_0 = \delta(\frac{1}{2}u)$ and $K_1 = \delta(\frac{1}{2}u^2)$ we conclude that $K_n = \delta\mathcal{H}_n$ for all subsequent n and that these generate an infinite sequence of commuting, bi-Hamiltonian flows.

Remark

Like the KdV equation above, many soliton equations possess an infinite number of conserved quantities in involution. However, we do not have Liouville's theorem in this context, so cannot automatically deduce complete integrability. Nevertheless, these equations are exactly soluble through the Inverse Spectral Transform (IST), so are 'completely integrable'. Associated linear spectral problems are discussed next.

3. Isospectral Flows and their Hamiltonian Structure

Many completely integrable Hamiltonian systems have arisen in recent years in the form of isospectral flows. This section contains a brief introduction to the main ideas, using the KdV hierarchy as the basic paradigm. I then present an interesting multi-component generalisation, associated with an energy-dependent Schrödinger operator [3,7,8].

KdV Paradigm

Consider the pair of linear equations:

$$\mathbb{L}\psi \equiv (\partial^2 + u - \lambda)\psi = 0\ , \quad \psi_t = \mathbb{P}\psi \equiv (\tfrac{1}{2}A\partial - \tfrac{1}{4}A_x)\psi \tag{3.1}$$

where $u(x,t)$ is the potential function and λ a spectral parameter. A is a function of λ, u and its x-derivatives. We seek the integrability conditions of (3.1), and find that:

$$\mathbb{L}_t - [\mathbb{P},\mathbb{L}] = A_x\mathbb{L}\ , \tag{3.2a}$$

giving the evolution of u by:

$$u_t = [\tfrac{1}{4}\partial^3 + u\partial + \tfrac{1}{2}u_x - \lambda\partial]A \equiv (B_1 - \lambda B_0)A\ . \tag{3.2b}$$

The important point here is that we started by seeking isospectral evolutions associated with spectral problem (3.1) and, after a simple manipulation, arrived at the two Hamiltonian structures B_0, B_1 of the KdV hierarchy. We next need to find A as a function of λ so that (3.2b) is identically satisfied in λ. The coefficient of λ^0 gives rise to the corresponding nonlinear evolution equation. To find the infinite sequence of polynomial evolutions we first consider the asymptotic series :

$$\mathcal{A} = \sum_{i=0}^{\infty} A_i \lambda^{-i} \qquad (3.3a)$$

as a solution of :

$$(B_1 - \lambda B_0)\mathcal{A} = 0 \ . \qquad (3.3b)$$

The polynomial forms of A are then given by :

$$A_{(m)} = (\lambda^m \mathcal{A})_+ , \qquad (3.4)$$

where $()_+$ means the truncation containing only non-negative powers of λ. Equation (3.3b) gives a recursion relation between the coefficients A_i :

$$B_1 A_i = B_0 A_{i+1} \ , \qquad (3.5a)$$

while the use of $A_{(m)}$ in (3.2b) leads to the isospectral flow :

$$u_{t_m} = B_1 A_m = B_0 A_{m+1} \ , \qquad (3.5b)$$

the second part being a consequence of (3.5a). We need to prove 2 results (given that B_0 and B_1 are compatible Hamiltonian operators) :

(i) that (3.5a) can be solved in terms of u and its x-derivatives for <u>each</u> coefficient A_{i+1}.

(ii) that A_m is the variational derivative of a functional $\mathcal{H}_m$, for each $m \geq 0$.

The first of these is proved by considering (3.3b) in more detail :

$$\tfrac{1}{4}\mathcal{A}_{xxx} + u\mathcal{A}_x + \tfrac{1}{2}u_x\mathcal{A} - \lambda\mathcal{A}_x = 0 \ . \qquad (3.6a)$$

Multiplying by $\mathcal{A}$ and integrating we get :

$$\tfrac{1}{2}\mathcal{A}\mathcal{A}_{xx} - \tfrac{1}{4}\mathcal{A}_x^2 + (u-\lambda)\mathcal{A}^2 = C(\lambda) \ , \qquad (3.6b)$$

where $C(\lambda)$ is a constant of integration, depending upon λ. Choosing

$C(\lambda) = -4\lambda$ we find $A_0 = 2$. Since the leading order term in the coefficient of $\lambda^{-(n-1)}$ is $2A_0A_n$ we can <u>always</u> solve for A_n in terms of previously calculated (differential) expressions in u. The first few of these are :

$$A_0 = 2, \quad A_1 = u, \quad A_2 = \tfrac{1}{4}(u_{xx} + 3u^2) \ . \qquad (3.6c)$$

To prove that each of these coefficients is a variational derivative we appeal to Lemma (2.1), by writing :

$$A_0 = \delta(2u) \ , \quad A_1 = \delta(\tfrac{1}{2}u^2) \ . \qquad (3.7a)$$

Thus, there exists a sequence of functions $\mathcal{H}_0, \mathcal{H}_1, \ldots$ s.t. :

$$A_m = \delta\mathcal{H}_m \ , \qquad (3.7b)$$

the first few being :

$$\mathcal{H}_0 = 2u, \quad \mathcal{H}_1 = \tfrac{1}{2}u^2, \quad \mathcal{H}_2 = \tfrac{1}{4}(u^3 - \tfrac{1}{2}u_x^2) \ . \qquad (3.8a)$$

These satisfy the bi-Hamiltonian ladder relation :

$$B_1\delta\mathcal{H}_n = B_0\delta\mathcal{H}_{n+1} \ , \qquad (3.8b)$$

with $\mathcal{H}_0$ being the Casimir function of B_0, and the equations of motion (3.5b) take bi-Hamiltonian form :

$$u_{t_m} = B_1\delta\mathcal{H}_m = B_0\delta\mathcal{H}_{m+1} \ . \qquad (3.9)$$

It follows from Theorem (2.1) that the functions $\mathcal{H}_m$ and that the corresponding flows (3.9) commute. The KdV equation (2.15a) corresponds to m=1.

This approach shows the clear connection between the isospectral and Hamiltonian properties of exactly soluble nonlinear evolution equations. A further connection will be made clear by the introduction of Miura maps in section 4.

Energy dependent Schrödinger

Consider the second order scalar spectral problem:

$$L\psi \equiv (\varepsilon\partial^2 + u)\psi \equiv \sum_0^N \lambda^i(\varepsilon_i\partial^2 + u_i)\psi = 0 \ , \qquad (3.10a)$$

with ε_i being constant and u_i functions of x. We look for time evolutions of the wave function ψ of the form :

$$\psi_t = \mathbb{P}\psi \equiv (\tfrac{1}{2}A\partial - \tfrac{1}{4}A_x)\psi \qquad (3.10b)$$

where A is a function of u_i and their x-derivatives, and of the spectral parameter λ. Proceeding as before:

$$\mathbb{L}_t - [\mathbb{P},\mathbb{L}] = A_x\mathbb{L}\ , \qquad (3.11a)$$

and:

$$u_t = \left(\tfrac{1}{4}\varepsilon\partial^3 + \tfrac{1}{2}(u\partial + \partial u)\right)A \equiv JA\ . \qquad (3.11b)$$

With ε and u defined by (3.10a), the operator J takes the form :

$$J = \sum_{k=0}^{N}\lambda^k J_k \equiv \sum_{k=0}^{N}\lambda^k\left(\tfrac{1}{4}\varepsilon_k\partial^3 + \tfrac{1}{2}(u_k\partial + \partial u_k)\right)\ , \qquad (3.12a)$$

with each J_k a copy of the KdV second structure. Equation (3.11b) then takes the form

$$\sum_{k=0}^{N}\lambda^k u_{kt} = \left(\sum_{k=0}^{N}\lambda^k J_k\right)A\ . \qquad (3.12b)$$

In [7] we continued the general development to include both KdV and Harry Dym type equations. Here we just consider the KdV reduction. This simplifies some of the formulae and statements :

KdV Case $u_N = -1$, $\varepsilon_N = 0$, so that $J_N = -\partial$.

To construct the 'polynomial' time evolutions we first seek a solution of $J\mathcal{A}=0$, where $\mathcal{A}$ is given by (3.3a). The co-efficients A_k must now satisfy:

$$J_0A_{k-N} + J_1A_{k-N+1} + \ldots + J_NA_k = 0, \quad \forall k \geq 0\ , \qquad (3.13)$$

with the polynomial expansion $A_{(m)}$ defined by (3.4). Upon substitution of $A_{(m)}$ into (3.12b) (with t_m parametrising the corresponding evolution) the coefficients of λ^k, $k \geq N$, are identically zero, whilst the remaining ones give the equations of motion for $u_0,\ldots,u_{N-1}$:

$$\begin{pmatrix} u_0 \\ \vdots \\ u_{N-1} \end{pmatrix}_{t_m} = \begin{pmatrix} 0 & \cdots & J_0 \\ \vdots & \iddots & \vdots \\ J_0 & \cdots & J_{N-1} \end{pmatrix}\begin{pmatrix} A_{m-N+1} \\ \vdots \\ A_m \end{pmatrix}\ . \qquad (3.14)$$

It is a remarkable fact that the scalar recursion relation (3.13) can be written as an N x N matrix equation in exactly (N+1) different ways :

$$B_nA^{(k-1)} = B_{n-1}A^{(k)}\ , \qquad n = 1,\ldots,N, \qquad (3.15a)$$

where $A^{(k)} = (A_{k-N+1},\ldots,A_k)^T$ and the matrix differential operators B_n are defined by:

$$B_n = \left(\begin{array}{ccc|ccc} & & J_0 & & & \\ & \iddots & \vdots & & 0 & \\ & & \vdots & & & \\ J_0 & \cdots & J_{n-1} & & & \\ \hline & & & -J_{n+1} & \cdots & -J_N \\ & 0 & & \vdots & \iddots & \\ & & & -J_N & & 0 \end{array}\right)\ . \qquad (3.15b)$$

In [7] we prove 3 basic facts :

(1) The operators B_n are each Hamiltonian and, furthermore, are mutually compatible.

(2) The recursion relation (3.13) can be solved for all k.

(3) The vectors $A^{(k)}$ given by (3.15a) are variational derivatives of a sequence of function(al)s $\mathcal{H}_k$ (the Hamiltonians).

Then it follows from (3.15a) that the equations of motion (3.14) can be written in Hamiltonian form in (N+1) distinct ways :

$$u_{t_m} = B_N\delta\mathcal{H}_m = \ldots = B_0\delta\mathcal{H}_{m+N}\ . \qquad (3.16)$$

I refer to [7] for the details. Here, I present one example.

Example : Dispersive water waves

To illustrate the above construction consider the simplest nontrivial example, N = 2. The resulting hierarchy is tri-Hamiltonian. Performing the invertible change of variables :

$$q = u_0 + \tfrac{1}{4}u_1^2 - \tfrac{1}{2}u_{1x}\ , \quad r = u_1\ , \qquad (3.17a)$$

changes the second order flow of (3.1a) into the standard DWW form:

$$q_{t_1} = \tfrac{1}{2}(-q_x + 2qr)_x\ , \quad r_{t_1} = \tfrac{1}{2}(r_x + 2q + r^2)_x\ . \qquad (3.17b)$$

The 3 Hamiltonian operators then take the form :

$$\bar{B}_0 = \begin{pmatrix} 0 & \partial \\ \partial & 0 \end{pmatrix}, \quad \bar{B}_1 = \frac{1}{2}\begin{pmatrix} q\partial + \partial q & -\partial^2 + r\partial \\ \partial^2 + \partial r & 2\partial \end{pmatrix},$$

$$\bar{B}_2 = \frac{1}{4}\begin{pmatrix} (r-\partial)(q\partial+\partial q)+(q\partial+\partial q)(r+\partial) & (r-\partial)^2\partial+2(q\partial+\partial q) \\ \partial(r+\partial)^2+2(q\partial+\partial q) & 2(r\partial+\partial r) \end{pmatrix}. \tag{3.18}$$

Numerous other examples can be found in [9], including Ito's equation (N=2, $\varepsilon=\lambda, u_0=\frac{1}{4}r^2$, $u_1=q$):

$$q_t = q_{xxx}+6qq_x+2rr_x, \quad r_t = 2(qr)_x. \tag{3.19}$$

Applying the above construction to other polynomial spectral problems results in analogous results. We have the same locally defined, compatible Hamiltonian operators (3.15b) so that the isospectral flows are multi-Hamiltonian of the form (3.16). The only difference is that the operators J_k take a different form. Fermionic extensions can be found in [4] and in section 5 of this paper, while generalised Zakharov-Shabat problems can be found [8]. Third order Lax operators are considered in [10].

Proof of Compatibility

The spectral problem (3.10a) depends polynomially upon the spectral parameter λ. A simple shift $\lambda\to\lambda+s$ induces a simple, invertible transformation of the potential functions which is generated by a particularly simple (master symmetry) vector field [8]:

$$\mathbb{L}(\lambda+s) = \mathbb{L}(\lambda)+s\frac{\partial\mathbb{L}}{\partial\lambda} + \cdots = \bar{\mathbb{L}}, \tag{3.20}$$

which is an operator of the same kind, depending upon new constants $\bar{\varepsilon}_i$ and new potential functions $\bar{u}_i$. Let us denote by $\mathbf{B}_i(s)$ ($\equiv \bar{B}_i$) the operator B_i when its coefficients depend upon the barred variables (with $B_i(0)=B_i$) (and similarly for $J_i(s)$). Let $\mathcal{J}(s)$ denote the Jacobian of the inverse of (3.20). If $\mathbf{v}$ denotes the infinitesimal generator of the transformation (3.20), then the Lie derivative $\mathcal{L}_v B_i$ is defined by:

$$\mathcal{L}_v B_i = \frac{d}{ds}\,\mathcal{J}(s)B_i(s)\mathcal{J}^T(s)\Big|_{s=0} \tag{3.21}$$

and is proportional to B_{i-1} as will be seen in the example below.

Furthermore, the quantity $\mathcal{J}(s)\,B_1(s)\,\mathcal{J}^T(s)$ is an s-dependent Hamiltonian operator which is a linear combination of $B_1(0)$, $B_{i-1}(0),\ldots,B_0(0)$. This is a direct generalisation of the well known example of the KdV hierarchy:

Example : KdV

$$\mathbb{L} = \partial^2+u-\lambda, \quad \lambda\to\lambda+s \;\Rightarrow\; \bar{u} = u-s,$$

$$B_1(s) = \partial^3+4\bar{u}\partial+2\bar{u}_x = \partial^3+4u\partial+2u_x-4s\partial = B_1-4sB_0. \tag{3.22}$$

This example is too simple since $\mathcal{J}(s)\equiv 1$, which is thus rendered superfluous.

Example : Two component case ($\varepsilon\equiv1$)

The two component example is nontrivial. With:

$$\mathbb{L}(\lambda) = \partial^2+u_0+u_1\lambda-\lambda^2, \tag{3.23}$$

the shift $\lambda\to\lambda+s$ induces the transformation:

$$\bar{u}_0 = u_0+u_1s-s^2, \quad \bar{u}_1 = u_1-2s, \tag{3.24a}$$

with infinitesimal generator:

$$\mathbf{v} = u_1\frac{\partial}{\partial u_0} - 2\frac{\partial}{\partial u_1} + prl, \tag{3.24b}$$

where *prl* denotes the usual prolongation of the vector field. It follows that:

$$J_0(s) = J_0+sJ_1+s^2J_2, \quad J_1(s) = J_1+2sJ_2, \quad J_2(s) = J_2, \tag{3.25a}$$

so that, with $\mathcal{J}(s)=\begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$, we have:

$$\mathcal{J}(s)B_2(s)\mathcal{J}^T(s) = B_2-2sB_1+s^2B_0,$$

$$\mathcal{J}(s)B_1(s)\mathcal{J}^T(s) = B_1-sB_0, \quad \mathcal{J}(s)B_0(s)\mathcal{J}^T(s) = B_0, \tag{3.25b}$$

which implies:

$$\mathcal{L}_{v_2}B_2 = -2B_1, \quad \mathcal{L}_{v_1}B_1 = -B_0, \quad \mathcal{L}_{v_0}B_0 = 0. \tag{3.25c}$$

These formulae show, very simply, that the Hamiltonian operators B_i are compatible.

4. Poisson Mappings

In this section we consider an interesting class of Poisson mappings which generalise the well known transformation of Miura [11], connecting solutions of the KdV and MKdV equations. We first consider this example:

The MKdV equation can be written in Hamiltonian form:

$$v_t = v_{xxx} - 6v^2 v_x = (-\partial)\frac{\delta \mathcal{H}^v}{\delta v}\ , \quad \mathcal{H}^v = \tfrac{1}{2}(v_x^2 + v^4)\ . \tag{4.1a}$$

Consider the variable u, defined by the mapping:

$$u = -v_x - v^2 \equiv M[v]\ . \tag{4.1b}$$

When v evolves by (4.1a), u evolves by:

$$u_t = -v_{xt} - 2vv_t = (-\partial - 2v)v_t \equiv M' v_t\ , \tag{4.1c}$$

where M' is the Fréchet derivative of M defined by $M'[v]w = \frac{d}{d\varepsilon} M[v+\varepsilon w]\Big|_{\varepsilon=0}$. Using (4.1a) and noticing that, up to equivalence, $\mathcal{H}^v$ can be obtained from $\mathcal{H}^u = \frac{1}{2}u^2$:

$$\mathcal{H}^v = \tfrac{1}{2}(-v_x - v^2)^2 - v^2 v_x \sim \mathcal{H}^u \circ M\ , \tag{4.1d}$$

and that under (4.1b) $\delta_v \mathcal{H}^v = (M')^\dagger \delta_u \mathcal{H}^u$, we have:

$$u_t = M' v_t = M'(-\partial)(M')^\dagger \delta_u \mathcal{H}^u\ . \tag{4.2a}$$

Noting further that

$$M'(-\partial)(M')^\dagger = (-\partial - 2v)(-\partial)(\partial - 2v) = \partial^3 + 4(-v_x - v^2)\partial + 2(-v_{xx} - 2vv_x), \tag{4.2b}$$

we obtain:

$$u_t = (\partial^3 + 4u\partial + 2u_x)\delta_u \mathcal{H}^u = u_{xxx} + 6uu_x\ . \tag{4.2c}$$

Thus, (4.1b) maps solutions of the MKdV equation (4.1a) onto solutions of the KdV equation (4.2c). Moreover, the second (local) Hamiltonian structure, $B_1 = \partial^3 + 4u\partial + 2u_x$, of the KdV equation is the image of the first (and only) local Hamiltonian structure, $(-\partial)$, of the MKdV equation.

Remark

It is important that the operator (4.2b), with coefficients given in terms of v is actually an operator with coefficients in terms of u. It is also important that $\mathcal{H}^v$ is (equivalent to) the pull back of some function $\mathcal{H}^u$.

The formula $B_1 = M'(-\partial)(M')^\dagger$ is just (2.7b) in the present setting. Thus, (4.1b) is a Poisson mapping. We may also regard (4.1b) as a change of co-ordinates which simplifies the operator B_1.

This simple example has been generalised in [12-15]. In the more general setting, with $\mathbf{u} = (u_0, \ldots, u_{N-1})^T$ and $\mathbf{v} = (v_0, \ldots, v_{N-1})^T$ the unmodified and modified variables respectively, we have:

Definition [14]

The mapping $\mathbf{u} = M[\mathbf{v}]$ is a Miura map for Hamiltonian operator $\tilde{B}$ (acting in the $\mathbf{v}$ space) if:

(i) M is not invertible

(ii) $B = M'\tilde{B}(M')^\dagger\big|_{u=M[v]}$ is locally defined in terms of $\mathbf{u}$ and its derivatives.

Remark

Whenever M is non-degenerate, the Hamiltonian nature of B follows from that of $\tilde{B}$.

Finding Miura Maps

We next discuss some techniques for constructing the Miura maps and modifications associated with a given spectral problem.

The Miura map (4.1b) can be obtained through the factorisation of the Schrödinger operator (3.1):

$$\partial^2 + u = (\partial + v)(\partial - v)\ . \tag{4.3a}$$

Defining $\varphi_2 = \varphi_{1x} - v\varphi_1$, we have [12]:

$$(\partial^2 + u)\varphi_1 = \lambda\varphi_1 \Rightarrow \begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix}_x = \begin{pmatrix} v & 1\\ \lambda & -v\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix}\ , \tag{4.3b}$$

which is a spectral problem for the MKdV hierarchy. Suppose instead that we considered the functions $\psi_1 \equiv \varphi_1$, $\psi_2 = \psi_{1x}$. Then:

$$(\partial^2+u)\varphi_1 = \lambda\varphi_1 \;\Rightarrow\; \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}_x = \begin{pmatrix}0 & 1\\ \lambda-u & 0\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} . \tag{4.4}$$

Comparing the definitions of (ψ_1,ψ_2) and (φ_1,φ_2) we find:

$$\Phi \equiv \begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} = \begin{pmatrix}1 & 0\\ -v & 1\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \equiv g\Psi . \tag{4.5a}$$

This is a gauge transformation connecting (4.3b) and (4.4) and:

$$\begin{pmatrix}v & 1\\ \lambda & -v\end{pmatrix} = g\begin{pmatrix}0 & 1\\ \lambda-u & 0\end{pmatrix}g^{-1}+g_x g^{-1} \;\Rightarrow\; u = -v_x-v^2 . \tag{4.5b}$$

In [7] we generalised the factorisation (4.3a) to energy dependent Schrödinger operators (3.10), obtaining a sequence of N Miura maps. We can also obtain these through gauge transformations.

Case N=2: Dispersive Water Waves

Written as a first order system, (3.10) is:

$$\Psi_x = Q\Psi \;,\; Q = \begin{pmatrix}0 & 1\\ \lambda^2-\lambda u_1-u_0 & 0\end{pmatrix} . \tag{4.6a}$$

The wave function Φ, defined by (4.5), satisfies:

$$\Phi_x = \hat{Q}\Phi \;,\; \hat{Q} = gQg^{-1}+g_x g^{-1} = \begin{pmatrix}v & 1\\ \lambda^2-\lambda u_1-u_0-v_x-v^2 & -v\end{pmatrix} . \tag{4.6b}$$

The 2 Miura maps are obtained by putting:

(i) $v = w_0$, $u_0 = -w_{0x}-w_0^2$, $u_1 = w_1$ (called $M^{(1)}$),

$$\hat{Q} = \begin{pmatrix}w_0 & 1\\ \lambda^2-\lambda w_1 & -w_0\end{pmatrix}; \tag{4.6c}$$

(ii) $v = v_0+\lambda v_1$, $u_0 = -v_{0x}-v_0^2$, $u_1 = -v_{1x}-2v_0v_1$ (called $M^{(2)}$),

$$\hat{Q} = \begin{pmatrix}v_0+\lambda v_1 & 1\\ \lambda^2(1-v_1^2) & -v_0-\lambda v_1\end{pmatrix} . \tag{4.6d}$$

Let the 3 Hamiltonian operators (3.15b) (with N=2) be labelled B_0^u, B_1^u and B_2^u. Then:

(i) The isospectral flows of (4.6b,c) are bi-Hamiltonian:

$$w_{t_m} = B_2^w\delta_w\mathcal{H}_m^w = B_1^w\delta_w\mathcal{H}_{m+1}^w \;,\; \mathcal{H}^w = \mathcal{H}^u\circ M^{(1)} , \tag{4.7a}$$

where

$$B_2^w = \begin{pmatrix}0 & \frac{1}{4}\partial(\partial-2w_0)\\ \frac{1}{4}(\partial+2w_0)\partial & w_1\partial+\frac{1}{2}w_{1x}\end{pmatrix} \;,\; B_1^w = \begin{pmatrix}-\frac{1}{4}\partial & 0\\ 0 & \partial\end{pmatrix} \tag{4.7b}$$

and

$$B_i^u = M^{(1)'}B_i^w(M^{(1)'})^\dagger\Big|_{u=M^{(1)}[w]} \;,\; i=1,2 ; \tag{4.7c}$$

(ii) The isospectral flows of (4.6b,d) are mono-Hamiltonian:

$$v_{t_m} = B_2^v\delta_v\mathcal{H}_m^v \;,\; \mathcal{H}_m^v = \mathcal{H}_m^v\circ M^{(2)} , \tag{4.8a}$$

where

$$B_2^v = -\frac{1}{4}\begin{pmatrix}0 & \partial\\ \partial & 0\end{pmatrix} \;,\; B_2^u = M^{(2)'}B_2^v(M^{(2)'})^\dagger\Big|_{u=M^{(2)}[v]} . \tag{4.8b}$$

General N

In [7] we present Miura maps M_n which reduce B_n to simpler form:

$$\tilde{B}_n = \frac{1}{4}\left(\begin{array}{cc|cc} 0 & -\partial & & \\ -\partial & 0 & & 0 \\ \hline & & 0 & \partial \\ & 0 & \partial & 0 \end{array}\right) . \tag{4.9}$$

Since $\tilde{B}_n$ has constant coefficients, it trivially satisfies the Jacobi identity, so that $B_n = M'_n\tilde{B}_n(M'_n)^\dagger$ also satisfies the Jacobi identity, thus giving a simple proof that B_n is Hamiltonian.

In [7] we also present a sequence of N Miura maps, relating the isospectral flows of (3.10) to a sequence of N modifications. The rth modification possesses (N-r+1) locally defined, mutually compatible Hamiltonian structures. If we denote the fields of the rth modification by $u^{(r)}=(u_0^{(r)},\dots u_{N-1}^{(r)})^T$, with $u^{(0)}\equiv u$ the original (unmodified) variable, then the rth modified hierarchy is written as:

$$u^{(r)}_{t_m} = B^r_{N-k}\,\delta\mathcal{H}^r_{m+k} \quad , \quad k=0,\ldots,N-r \quad , \quad m\geq 0 \quad , \tag{4.10}$$

where $B^r_r,\ldots,B^r_N$ are the (N-r+1) compatible Hamiltonian operators and $\mathcal{H}^r_{m+k}$ is the 'pull back' of $\mathcal{H}^0_{m+k}\equiv\mathcal{H}_{m+k}$ of the unmodified hierarchy.

We can represent these modifications and their Hamiltonian structures schematically as follows :

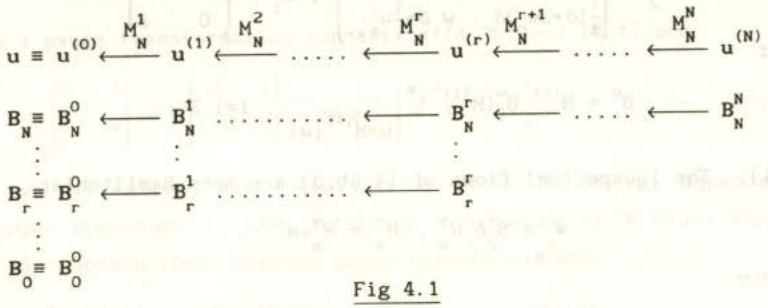


Fig 4.1

5. Fermionic and Supersymmetric Extensions

In this section I first describe fermionic extensions of the coupled KdV hierarchies of section 3 and the Miura maps of section 4. These are direct generalisations of Kupershmidt's [16] bi-Hamiltonian super-KdV equation and, as such, are <u>not</u> supersymmetric. These fermionic extensions were first presented in [4].

Secondly, I consider the "N=1" <u>supersymmetric</u> extension of the KdV equation [17] and, using the general approach described in this paper, <u>systematically</u> construct the associated Hamiltonian operators. As is known, the only local one corresponds to the <u>second</u> local structure of the KdV equation. Our construction gives the 'lower' structure only <u>implicitly</u>, although its inverse can be written down in <u>explicit</u> form. These results coincide with those of Oevel and Popowicz [18]. Extensions of these results will be presented elsewhere [19].

The Hamiltonian theory of section 2 has to be modified slightly to include fermionic and supersymmetric extensions. This material can be found elsewhere in these proceedings or in [20,21].

Fermionic Extensions

Here we just add some fermionic fields without enlarging our space-time to a 'superspace'. Thus the resulting structures and equations are <u>not</u> supersymmetric.

We start with Kupershmidt's spectral problem [16]:

$$(\partial^2+u)\psi+\xi\varphi = \lambda\psi \quad , \quad \partial\varphi+\xi\psi = 0 \tag{5.1a}$$

for the following bi-Hamiltonian fermionic extension of the KdV equation:

$$u_t = \tfrac{1}{4}(u_{xxx}+6uu_x+12\xi\xi_{xx}) \quad , \quad \xi_t = \tfrac{1}{4}(4\xi_{xxx}+6u\xi_x+3u_x\xi) \quad . \tag{5.1b}$$

The two Hamiltonian structures are:

$$B_1 = \begin{pmatrix} \partial^3+2(u\partial+\partial u) & 2\xi\partial+\partial\xi \\ 2\partial\xi+\xi\partial & \partial^2+u \end{pmatrix} \quad , \quad B_0 = \begin{pmatrix} 4\partial & 0 \\ 0 & 1 \end{pmatrix} \tag{5.1c}$$

and B_1 can be simplified by means of the Miura map

$$M:\ u = -v_x-v^2-\eta\eta_x \quad , \quad \xi = -\eta_x-v\eta \quad . \tag{5.1d}$$

The single local Hamiltonian structure $D=\begin{pmatrix}\partial & 0\\ 0 & 1\end{pmatrix}$ of the modified hierarchy is thus related to B_1 by:

$$B_1 = m(-D)m^{s\dagger} \quad , \tag{5.1e}$$

where the Fréchet derivative m of M and its super-adjoint are given by:

$$m = \begin{pmatrix} -\partial-2v & -\eta\partial+\eta_x \\ -\eta & -\partial-v \end{pmatrix} \quad , \quad m^{s\dagger} = \begin{pmatrix} \partial-2v & -\eta \\ -\partial\eta-\eta_x & \partial-v \end{pmatrix} \quad . \tag{5.1f}$$

The spectral problem (5.1a) is a super-generalisation of the Schrödinger equation. Here, we make a similar super-generalisation of the energy dependent Schrödinger equation (3.10a) (with ε=1):

$$(\partial^2+u)\psi+\xi\varphi = 0 \quad , \quad \partial\varphi+\xi\psi = 0 \quad , \tag{5.2a}$$

where u,ψ are even and ξ,φ are odd variables. We specify the λ-dependence of u and ξ to be:

$$u = \sum_0^N u_i\lambda^i \ , \ \xi = \sum_0^N \xi_i\lambda^i \qquad (5.2b)$$

with u_i,ξ_i being functions of x and with spectral parameter λ being even. In [4] we show that the time evolution:

$$\psi_t = 2p\psi_x - p_x\psi - \rho\varphi \ , \ \varphi_t = \rho\psi_x - (\rho_x+2\xi p)\psi \qquad (5.2c)$$

leads to the integrability conditions:

$$u_t = p_{xxx}+4up_x+2u_xp+3\xi\rho_x+\xi_x\rho \ , \ \xi_t = 3\xi p_x+2\xi_x p+\rho_{xx}+u\rho \ . \qquad (5.2d)$$

Introducing the operator $J=\sum_0^N J_k\lambda^k$ with:

$$J_0=\begin{pmatrix}\partial^3+2(u_0\partial+\partial u_0) & 2\xi_0\partial+\partial\xi_0\\ 2\partial\xi_0+\xi_0\partial & \partial^2+u_0\end{pmatrix} \ , \ J_k=\begin{pmatrix}2(u_k\partial+\partial u_k) & 2\xi_k\partial+\partial\xi_k\\ 2\partial\xi_k+\xi_k\partial & u_k\end{pmatrix} \ , \ k=1,\ldots,N. \qquad (5.3a)$$

we can write (5.2d) as:

$$U_t = JP \qquad (5.3b)$$

with $U=(u,\xi)^T$ and $P=(p,\rho)^T$. As before we have 2 cases, giving super-extensions of either coupled KdV systems ($U_N=(-1,0)^T$) or of Harry Dym systems ($U_0=(0,0)^T$). Here we consider only the coupled KdV reduction:

$$u_N=-1 \ , \ \xi_N=0 \ , \ J_N=\begin{pmatrix}-4\partial & 0\\ 0 & -1\end{pmatrix} . \qquad (5.4)$$

If we choose $P_{(m)}=\sum_{i=0}^m P_{m-1}\lambda^i$ to be polynomial in λ, with $P_k=(p_k,\rho_k)^T$ functions of u_i,ξ_i and their x-derivatives, then (5.3b) decouples into:

$$J_0P_{k-N}+J_1P_{k-N+1}+\ldots+J_NP_k = 0 \ , \ k=0,1,\ldots,m-1 \ , \qquad (5.5a)$$

and we obtain the evolution equation:

$$\begin{pmatrix}U_0\\ U_1\\ \vdots\\ U_{N-1}\end{pmatrix}_{t_m} = \begin{pmatrix} & & J_0\\ & 0 & \vdots\\ & \iddots & \vdots\\ J_0 & \cdots\cdots & J_{N-1}\end{pmatrix}\begin{pmatrix}P_{m-N+1}\\ \vdots\\ \vdots\\ P_m\end{pmatrix} . \qquad (5.5b)$$

These are exactly the same formulae as in (3.13) and (3.14), only now we have P_k a 2-vector and J_k a 2x2 matrix. In fact, we can proceed exactly as in section 3 and derive the same results, except our P_k and J_k are different. Thus we can consider the following asymptotic series and its truncations:

$$\mathcal{P} = \sum_{i=0}^{\infty} P_i\lambda^{-1} \ , \ P_{(m)} = (\lambda^m\mathcal{P})_+ \ , \qquad (5.6)$$

with P_k satisfying the recursion relation (5.5a), but for <u>all</u> k≥0. We can also construct the (N+1) operators B_n of (3.15b) and prove the same 3 basic facts (repeated here for emphasis):

(i) The operators B_n are each Hamiltonian and, furthermore, are mutually compatible.

(ii) The recursion relation (5.5a) can be solved for <u>all</u> k≥0.

(iii The vectors $P^{(k)}=(P_{m-N+1},\ldots,P_m)^T$ are variational derivatives of a sequence of functionals $\mathcal{H}_k$: $P^{(k)}=\delta\mathcal{H}_k$ and satisfy:

$$B_NP^{(k)} = B_{N-1}P^{(k+1)} = \ldots = B_0P^{(k+N)} \ .$$

It then follows from (5.5b) that the isospectral flows of (5.2a) can be written in Hamiltonian form in (N+1) distinct ways:

$$U_{t_m} = B_N\delta\mathcal{H}_m = \ldots = B_0\delta\mathcal{H}_{m+N} \ . \qquad (5.7)$$

The Miura maps of section 4 can also be extended to this case. Whereas before our basic building block was the KdV Miura map, now we start with Kupershmidt's map (5.1d). Thus all the formulae of section 4 still hold, but with m_k defined (with $v_N=-1$, $\eta_N=0$) by:

$$m_0=\begin{pmatrix}-\partial-2v_0 & -\eta_0\partial+\eta_{0x}\\ -\eta_0 & -\partial-v_0\end{pmatrix} \ , \ m_k=\begin{pmatrix}-2v_k & -\eta_k\partial+\eta_{kx}\\ -\eta_k & -v_k\end{pmatrix} \ , \ k=1,\ldots,N-1 \ , m_N=\begin{pmatrix}2 & 0\\ 0 & 1\end{pmatrix} . \qquad (5.8a)$$

We can thus reduce the Hamiltonian operator B_n to the simple form:

$$B_n = \left(\begin{array}{cc|cc} 0 & -D & & \\ & \ddots & & 0 \\ -D & 0 & & \\ \hline & & 0 & D \\ & 0 & & \ddots \\ & & D & 0 \end{array}\right) , \quad D = \begin{pmatrix} \partial & 0 \\ 0 & 1 \end{pmatrix} , \qquad (5.8b)$$

where the two diagonal blocks are respectively nxn and (N-n)x(N-n) (with each 'component' being 2x2). As before, there are N modifications, with the rth modification possessing (N+1-r) compatible Hamiltonian structures. These results can be depicted by a diagram analogous to figure 4.1.

The details of these fermionic extensions are presented in [4] where several examples can be found. Here, I just give one example.

Super dispersive water waves Here N=2 and the hierarchy is tri-Hamiltonian:

$$B_0 = \begin{pmatrix} -J_1 & -J_2 \\ -J_2 & 0 \end{pmatrix} , \quad B_1 = \begin{pmatrix} J_0 & 0 \\ 0 & -J_2 \end{pmatrix} , \quad B_2 = \begin{pmatrix} 0 & J_0 \\ J_0 & J_1 \end{pmatrix} , \qquad (5.9a)$$

with J_k given by (5.3a) and (5.4). The first few Hamiltonians are:

$$\mathcal{H}_0 = \tfrac{1}{2}u_1 , \quad \mathcal{H}_1 = \tfrac{1}{2}u_0 + \tfrac{1}{8}u_1^2 + \tfrac{1}{2}\xi_1\xi_{1x} , \quad \mathcal{H}_2 = \tfrac{1}{4}u_0u_1 + \tfrac{1}{16}u_1^3 + \tfrac{3}{4}u_1\xi_1\xi_{1x} + \xi_0\xi_{1x} . \qquad (5.9b)$$

There are 2 Miura maps, the first of which:

$$u_0 = -w_{0x} - w_0^2 - \kappa_0\kappa_{0x} , \quad u_1 = w_1 , \quad \xi_0 = -\kappa_{0x} - w_0\kappa_0 , \quad \xi_1 = \kappa_1 , \qquad (5.10a)$$

reduces B_1 to $\tilde{B}_1 = \begin{pmatrix} -D & 0 \\ 0 & D \end{pmatrix}$. The resulting hierarchy can be further modified by:

$$w_0 = v_0 , \quad w_1 = -v_{1x} - 2v_0v_1 - \eta_0\eta_{1x} - \eta_1\eta_{0x} , \quad \kappa_0 = \eta_0 , \quad \kappa_1 = -\eta_{1x} - v_0\eta_1 - v_1\eta_0 . \qquad (5.10b)$$

The composition of (5.10a,b) gives a map which reduces B_2 to $\tilde{B}_2 = \begin{pmatrix} 0 & -D \\ -D & 0 \end{pmatrix}$.

Supersymmetric Extensions

Here I show that the methods of this paper can be used to systematically construct the Hamiltonian structures associated with the supersymmetric KdV equation introduced by Manin and Radul [17]. In fact only the 'second' structure can be explicitly constructed, the 'first' being some highly non-local structure, but whose inverse can be explictly constructed. This reproduces the recent results of Oevel and Popowicz [18].

Here, as well as adding a fermionic field, we enlarge the space-time into a 'superspace' by adding the grassmanian variable θ, satisfying $\theta^2=0$. The supersymmetric co-variant derivative $\mathcal{D}$ is a 'square root' of ∂_x:

$$\mathcal{D} = \partial_\theta + \theta\partial_x , \quad \mathcal{D}^2 = \partial_x . \qquad (5.11)$$

We start with the spectral problem [17]:

$$\mathbb{L}\Psi = (\partial_x^2 + q\mathcal{D} - \lambda)\Psi = 0 , \quad q = \xi + \theta u , \qquad (5.12a)$$

where ξ is a fermionic field and u bosonic, so that q has odd parity. The operator $\mathbb{L}$ has even parity, so choosing $\Psi=\varphi+\theta\psi$, where φ,ψ are respectively odd and even, we find the component form:

$$\psi_{xx} + (u-\lambda)\psi = \xi\varphi_x , \quad \varphi_{xx} + \xi\psi = \lambda\varphi , \qquad (5.12b)$$

which reduces to the usual Schrödinger equation when all fermionic quantities are set to zero. The time evolution of Ψ can be written:

$$\Psi_t = \Big[(\alpha\mathcal{D}+\beta)\partial + \gamma\mathcal{D}\Big]\Psi \qquad (5.13a)$$

where α and γ are odd, whilst β is even, with

$$\alpha = \alpha^o + \theta\alpha^e , \quad \beta = \beta^e + \theta\beta^o , \quad \gamma = \gamma^o + \theta\gamma^e , \qquad (5.13b)$$

superfix "o" and "e" denoting odd and even parity functions. The integrability conditions imply:

$$\gamma = \tfrac{1}{2}(\mathcal{D}\beta - \alpha_x) , \quad q\alpha = \beta_x , \qquad (5.13c)$$

$$q_t = \tfrac{1}{2}\Big[\alpha_{xxx} + \mathcal{D}\beta_{xx} + 3q\beta_x + 2q_x\beta + (\mathcal{D}q)\alpha_x + q_x\mathcal{D}\alpha + 3q\mathcal{D}\alpha_x + (\mathcal{D}q)\mathcal{D}\beta\Big] . \qquad (5.13d)$$

In terms of components this gives:

$$\psi_t = \left[(\alpha^e+\beta^e)\partial-\frac{1}{2}(\alpha^e+\beta^e)_x\right]\psi+\left[\frac{1}{2}(\beta^0+\alpha^0_x)\partial-\lambda\alpha^0\right]\varphi \ , \quad \varphi_t = \left[\alpha^0\partial+\frac{1}{2}(\beta^0-\alpha^0_x)\right]\psi+\beta^e\varphi_x \ , \tag{5.14a}$$

and

$$\begin{pmatrix} u \\ \xi \end{pmatrix}_t = \begin{pmatrix} \frac{1}{2}\partial^3+2u\partial+u_x & \frac{3}{2}\xi\partial+\frac{1}{2}\xi_x \\ \frac{3}{2}\xi\partial+\xi_x & -\frac{1}{2}(\partial^2+u) \end{pmatrix}\begin{pmatrix} \alpha^e+\beta^e \\ -\beta^0-\alpha^0_x \end{pmatrix}-2\lambda\begin{pmatrix} \alpha^e_x \\ \alpha^0_x \end{pmatrix} , \tag{5.14b}$$

subject to the conditions:

$$\xi\alpha^0 = \beta^e_x \ , \quad u\alpha^0-\xi\alpha^e = \beta^0_x \ . \tag{5.14c}$$

The operator:

$$B_1 = \begin{pmatrix} \frac{1}{2}\partial^3+2u\partial+u_x & \frac{3}{2}\xi\partial+\frac{1}{2}\xi_x \\ \frac{3}{2}\xi\partial+\xi_x & -\frac{1}{2}(\partial^2+u) \end{pmatrix} \tag{5.14d}$$

is the known local Hamiltonian structure, which reduces to the second structure of the KdV equation when ξ=0. We can recursively construct polynomial (in λ) functions $\alpha^e, \beta^e, \alpha^0$ and β^0, from which we construct the time evolutions of the wave functions ψ and φ and the gradients of the Hamiltonians.

Remark

To execute this recursion we need both (5.14b) and the constraints (5.14c).

The mth flow:

$$\begin{pmatrix} u \\ \xi \end{pmatrix}_{t_m} = B_1\delta\mathcal{H}_m \tag{5.15a}$$

corresponds to:

$$\alpha = \sum_1^m \alpha_i\lambda^{m-i} \ , \quad \beta = \sum_1^m \beta_i\lambda^{m-i} \ , \quad \delta\mathcal{H}_m = (\alpha^e_m+\beta^e_m \ , \ -\beta^0_m-\alpha^0_{mx})^T \ . \tag{5.15b}$$

The first three Hamiltonians are:

$$\mathcal{H}_0 = u \ , \quad \mathcal{H}_1 = \frac{1}{2}(u^2-\xi\xi_x) \ , \quad \mathcal{H}_2 = u^3-\frac{1}{2}u^2_x+\frac{1}{2}\xi_x\xi_{xx}-2u\xi\xi_x \ , \tag{5.16a}$$

with $\mathcal{H}_1$ giving the supersymmetric KdV equations:

$$u_{t_1} = \frac{1}{2}(u_{xx}+3u^2-3\xi\xi_x)_x \ , \quad \xi_t = \frac{1}{2}(\xi_{xx}+3u\xi)_x \ . \tag{5.16b}$$

The Hamiltonian $\mathcal{H}_2$ gives the supersymmetric extension of Lax's fifth order KdV equation.

The lower Hamiltonian operator B_0 must satisfy;

$$\begin{pmatrix} \alpha^e_x \\ \alpha^0_x \end{pmatrix} = B_0\begin{pmatrix} \alpha^e+\beta^e \\ -\beta^0-\alpha^0_x \end{pmatrix} = B_0\begin{pmatrix} \alpha^e+\partial^{-1}\xi\alpha^0 \\ \partial^{-1}\xi\alpha^e-(\partial+\partial^{-1}u)\alpha^0 \end{pmatrix} , \tag{5.17a}$$

where we have used the constraints (5.14c) to eliminate β. It is not possible to calculate the components of B_0 in closed form. However:

$$B_0^{-1}\partial\begin{pmatrix} \alpha^e \\ \alpha^0 \end{pmatrix} = \begin{pmatrix} 1 & \partial^{-1}\xi \\ \partial^{-1}\xi & -\partial-\partial^{-1}u \end{pmatrix}\begin{pmatrix} \alpha^e \\ \alpha^0 \end{pmatrix} \tag{5.17b}$$

can be solved to give:

$$B_0^{-1} = \begin{pmatrix} \partial^{-1} & \partial^{-1}\xi\partial^{-1} \\ \partial^{-1}\xi\partial^{-1} & -1-\partial^{-1}u\partial^{-1} \end{pmatrix} , \tag{5.17c}$$

which is the formula found by Oevel and Popowicz [18].

Acknowledgements

I thank the organisers for creating the conditions for a stimulating atmosphere and providing some interesting people to talk to. The reported investigation of supersymmetric systems in section 5 is a direct consequence of conversations with Pierre Mathieu, Walter Oevel and Ziemek Popowicz.

References

[1] V.I.Arnol'd, "Mathematical Methods of Classical Mechanics", Springer-Verlag, Berlin, 1978.

[2] P.J.Olver, "Application of Lie Groups to Differential Equations". Springer-Verlag, Berlin, 1986.

[3] M.Antonowicz and A.P.Fordy, Hamiltonian structures of nonlinear evolution equations. Published in "Soliton Theory : A Survey of Results", pp.273-312, ed. A.P.Fordy, MUP, Manchester, 1990.

[4] M.Antonowicz and A.P.Fordy, Super-extensions of energy dependent Schrödinger operators. Commun.Math.Phys. 124, 487-500 (1989).

[5] F.Magri, A simple model of the integrable Hamiltonian equation. J.Math.Phys. 19, 1156-1162 (1978).

[6] B.A.Kupershmidt, Is a bi-Hamiltonian system necessarily integrable? Phys. Letts. A. 123, 55-9 (1987).

[7] M.Antonowicz and A.P.Fordy, Factorisation of energy dependent Schrödinger operators : Miura maps and modified systems. Commun.Math.Phys. 124, 465-86 (1989).

[8] A.P.Fordy, Isospectral flows: their Hamiltonian structures, Miura maps and master symmetries. Published in: Solitons in Physics, Mathematics and Nonlinear Optics, (eds) P.J.Olver and D.L.Sattinger, Springer, NY 1990.

[9] M.Antonowicz and A.P. Fordy, Coupled KdV equations with multi-Hamiltonian structures. Physica 28D, 345-57 (1987).

[10] M.Antonowicz, A.P.Fordy and Q.P.Liu, Energy dependent third order Lax operators, Nonlinearity (in press, 1991).

[11] R.M.Miura, Korteweg-de Vries equation and generalizations I. A remarkable explicit nonlinear transformation, J.Math.Phys. 9, 1202-1204 (1968).

[12] A.P.Fordy and J.Gibbons, Factorisation of operators I. Miura transformations. J.Math.Phys. 21, 2508-10 (1980).

[13] A.P.Fordy and J.Gibbons, Factorisation of operators II. J.Math.Phys. 22, 1170-5 (1981).

[14] B.A.Kupershmidt and G.Wilson, Modifying Lax equations and the second Hamiltonian structure, Invent.Math. 62, 403-36 (1981).

[15] V.G.Drinfel'd and V.V.Sokolov, Lie algebras and equations of Korteweg de Vries type, J.Sov.Math. 30, 1975-2036 (1985).

[16] B.A.Kupershmidt, A super KdV equation: an integrable system, Phys.Letts.A 102, 213-5 (1984).

[17] Yu.I.Manin and A.O.Radul, A supersymmetric extension of the Kadomtsev-Petviashvili Hierarchy, Commun.Math.Phys. 98, 65-77 (1985).

[18] W.Oevel and Z.Popowicz, The bi-Hamiltonian structure of fully supersymmetric Korteweg-de Vries systems, Commun. Math.Phys. (In press, 1991).

[19] A.P.Fordy and Q.P.Liu, Fermionic and supersymmetric KdV systems: Gauge transformations and Miura maps. Preprint 1991.

[20] D.A.Leites, Introduction to the theory of supermanifolds, Russ. Math.Surveys, 35, 1-64 (1980).

[21] B.A. Kupershmidt, Elements of superintegrable systems, Reidel, Dordrecht, 1987.

Toda theories, integrability and conformal invariance

L.Bonora
International School for Advanced Studies (SISSA/ISAS)
Via Beirut 2, 34014 Trieste, Italy
and INFN, Sezione di Trieste.

Abstract.
The conformal affine sl_2 Toda field theory is introduced and analyzed. Its reduction to the sinh-Gordon theory is studied. The vacuum corresponding to the sinh-Gordon theory is also determined. The classical affine sl_2 Toda theory is then defined on the lattice. Finally the Liouville theory is quantized on the lattice.

1 Introduction

It is well-known that Toda field theories based on finite dimensional Lie algebras together with the representation theory of the corresponding quantum group, provide a realization of W-algebra minimal models. It is also known that there are Toda field theories (based on loop algebras, e.g. the sine-Gordon theory) which are integrable but not conformal invariant, and are interesting from the physical point of view since they are used to describe off-critical situations. What is less known is the meaning and the role of the conformal invariant Toda theories based on affine Lie algebras. The comparison among these three types of theories offers us an excellent vantage point over the interplay between integrability and conformal invariance.

Let us try to clarify this issue. In the literature we meet integrable field theories – the typical example is the sine-Gordon model – which were solved essentially by using the Yang-Baxter equation, only the quantum group structure of the model being thus exploited. On the other hand, in conformal field theory, only the conformal algebra was first considered, and it took some time to understand how quantum

groups intervene. The complete structure is synthesized by the exchange algebra. What is not understood yet is the role the conformal algebra in integrable field theory, e.g. in the Sine-Gordon theory.

Let us consider the sl_2 case. The above hierarchy of three models (see ref.[1]) is formed by: the simple Liouville model (or simple sl_2 Toda field theory) defined by the equation

$$\partial_+\partial_-\varphi = m^2 e^{2\varphi} \tag{1}$$

the sinh-Gordon model (for simplicity we will be dealing with the latter instead of the sine-Gordon model)

$$\partial_+\partial_-\varphi = m^2\left(e^{2\varphi} - e^{-2\varphi}\right) \tag{2}$$

and the conformal affine Liouville model (conformal affine sl_2 Toda field theory) defined by three equations of motion

$$\partial_+\partial_-\varphi = m^2\left(e^{2\varphi} - e^{2(\eta-\varphi)}\right) \tag{3}$$

$$\partial_+\partial_-\eta = 0 \tag{4}$$

$$\partial_+\partial_-\xi = m^2 e^{2(\eta-\varphi)} \tag{5}$$

Here light-cone coordinates $z_\pm = x \pm t$, $x \in S^1, t \in R$, are being used.

The first and last models are conformal invariant, the second is not. To make contact with critical and off-critical descriptions, we could envisage the hierarchy of three equations as follows: eq.(1) is regarded as the initial conformal invariant (unperturbed) theory, which we perturb by adding the second term in the RHS of eq.(2) breaking in such a way conformal invariance; we then recover conformal invariance by adding suitable fields η and ξ. The third theory contains therefore, under suitable limiting conditions, both the initial unperturbed theory and its integrable perturbation: for example, by setting $\eta = \text{const}$ in eq.(3) we recover (2).

For simplicity in these lectures I will limit myself to the sl_2 case, but the above hierarchy can be generalized to Toda field theories based on any Lie algebra.

The program outlined here, and so far partially completed, consists in analyzing the classical phase space and, subsequently, in quantizing the affine Liouville theory (and of course, the Liouville one). The idea is to use not only the integrable structure but also the conformal one in the affine Liouville theory in order to derive results for the sinh-Gordon or sine-Gordon theory, which is to be considered the interesting physical theory.

The paper is organized as follows. In section 2 we introduce notation and models. In section 3 the affine Liouville model is analyzed from a lagrangian perturbative viewpoint: a subspace of the perturbative Hilbert space is recognized to describe the sinh-Gordon model, if conformal invariance is broken by suitably choosing the vacuum. Section 4 is devoted to analyzing the classical phase space of the affine Liouville model, and section 5 to its classical formulation on the lattice. In section 6

we study the reduction of the affine Liouville model to the sinh-Gordon model and the vacuum corresponding to the latter. Finally, in the last section the ordinary Liouville model is quantized on the lattice.

2 Classical Toda field theories. Generalities

2.1 Notation

Let E_+, E_-, H be the three generators of the Lie algebra sl_2

$$[H, E_\pm] = \pm 2E_\pm$$
$$[E_+, E_-] = H$$

The loop algebra $\tilde{sl}_2$ is the Lie algebra of traceless 2×2 matrices with entries which are Laurent polynomials in λ

$$\tilde{sl}_2 = C(\lambda, \lambda^{-1}) \otimes sl_2$$

The Lie algebra $\tilde{sl}_2'$ is the central extension of $\tilde{sl}_2$

$$\tilde{sl}_2' = \tilde{sl}_2 \oplus C\hat{c}$$

The affine Lie algebra $\hat{sl}_2$ is obtained from $\tilde{sl}_2'$ by adding the derivation $\hat{d} = \lambda\frac{d}{d\lambda}$.

More explicitly the $\hat{sl}_2$ algebra has generators $\hat{c}, \hat{d}$ and $X_k = \lambda^k \otimes X$ where $X = E_+, E_-, H$, which satisfy the algebra

$$[H_k, E_l^\pm] = \pm 2E_{k+l}^\pm \qquad [H_k, H_l] = 2k\delta_{k,-l}\hat{c}$$
$$[E_k^+, E_l^-] = H_{k+l} + k\delta_{k,-l}\hat{c} \qquad [d, X_k] = kX_k$$

The Cartan subalgebra $\hat{\mathcal{H}}$ is spanned by the elements H, $\hat{c}$ and $\hat{d}$. We have the decomposition

$$\hat{sl}_2 = \hat{\mathcal{N}}_- \oplus \hat{\mathcal{H}} \oplus \hat{\mathcal{N}}_+$$

$\hat{\mathcal{N}}_\pm$ admit the corresponding root decomposition. In particular, the simple roots eigenvectors are

$$E_{\pm\alpha_1} = E_\pm$$
$$E_{\pm\alpha_2} = \lambda^{\pm 1}E_\mp$$

and the positive root eigenvectors are

$$E_+, \qquad E_n^\pm, \qquad H_n, \qquad n > 0$$

We choose the inner product $(\cdot,\cdot)$ as follows

$$(\hat{c},\hat{d})=1, \qquad (H_n,H_{-n})=2, \qquad (E_n^+,E_{-n}^-)=1, \quad \forall n$$

Next we define

$$\mathcal{E}_+ = m(E_{\alpha_1}+E_{\alpha_2}) = m(E_+ + \lambda E_-) \tag{6}$$
$$\mathcal{E}_- = m(E_{-\alpha_1}+E_{-\alpha_2}) = m(E_- + \lambda^{-1}E_+) \tag{7}$$

We also introduce the element $t_0 \in \hat{\mathcal{H}}\otimes\hat{\mathcal{H}}$

$$t_0 = H\otimes H + 2\hat{c}\otimes\hat{d} + 2\hat{d}\otimes\hat{c} \tag{8}$$

We will need the two fundamental representations of $\widehat{sl}_2$. The corresponding highest weight vectors $|\Lambda_0>$ and $|\Lambda_1>$ are defined by:

$$H|\Lambda_0>=0, \qquad \hat{c}|\Lambda_0>=|\Lambda_0>, \qquad \hat{d}|\Lambda_0>=0$$
$$H|\Lambda_1>=|\Lambda_1>, \qquad \hat{c}|\Lambda_1>=|\Lambda_1>, \qquad \hat{d}|\Lambda_1>=0$$

At times we will consider also other finite dimensional Lie algebras beside sl_2. The corresponding Cartan subalgebra $\mathcal{H}$ will be spanned by the generators H_i. We will need

$$\mathcal{E}_+ = \sum_{\alpha\ simple} E_\alpha$$
$$\mathcal{E}_- = \sum_{\alpha\ simple} E_{-\alpha}$$

and

$$t_0 = \sum_i H_i\otimes H_i$$

2.2 Zero curvature representation

The Toda field theory models can be introduced by means of the zero curvature representation as follows. For any affine or finite dimensional Lie algebra $\hat{\mathcal{G}}$ or $\mathcal{G}$, respectively, let us introduce the field Φ valued in the Cartan subalgebra $\mathcal{H}$

$$\Phi = \tfrac{1}{2}\textstyle\sum_i H_i\otimes H_i$$

or in $\widehat{\mathcal{H}}$

$$\Phi = \tfrac{1}{2}\varphi H + \eta\hat{d} + \tfrac{1}{2}\xi\hat{c}$$

Then let us define the left and right connections

$$A_{z_+} = \partial_{z_+}\Phi + e^{ad\Phi}\mathcal{E}_+ \tag{9}$$
$$A_{z_-} = -\partial_{z_-}\Phi + e^{-ad\Phi}\mathcal{E}_- \tag{10}$$

The zero curvature condition

$$F_{z_+z_-} = \partial_{z_+}A_{z_-} - \partial_{z_-}A_{z_+} + [A_{z_+},A_{z_-}] = 0 \tag{11}$$

gives the equations of motion of the corresponding models. In particular for the $\widehat{sl}_2$ case, we obtain the equations of motion (3,4,5).

Remark 1. — The equations of motion are invariant under the transformation

$$z_+ = f(\tilde{z}_+)$$
$$z_- = g(\tilde{z}_-)$$

if the new fields are defined as

$$\tilde{\varphi}(\tilde{z}_+,\tilde{z}_-) = \varphi(f(\tilde{z}_+),g(\tilde{z}_-)) + \tfrac{1}{2}\log(f'g')$$
$$\tilde{\eta}(\tilde{z}_+,\tilde{z}_-) = \eta(f(\tilde{z}_+),g(\tilde{z}_-)) + \log(f'g')$$
$$\tilde{\xi}(\tilde{z}_+,\tilde{z}_-) = \xi(f(\tilde{z}_+),g(\tilde{z}_-)) + \tfrac{B}{4}\log(f'g')$$

The parameter B is arbitrary.

Remark 2. —Eqs.(1) and (2) can be obtained with the above construction by considering the sl_2 and $\widehat{sl}_2$ case respectively.

Remark 3. —The affine Liouville model can be obtained by starting from the Chern-Simons theory based on $\widehat{sl}_2$ in 2+1 dimensions : one derives the corresponding WZNW theory in 2 dimensions and applies a suitable Hamiltonian reduction [2].

2.3 The classical integrable structure

The integrable structure of the Toda model in an elementary form may be exposed as soon as we impose canonical Poisson brackets for the fields involved. Let us define the Cartan subalgebra valued conjugate momenta

$$\pi_\Phi = \partial_t\Phi$$

and impose, at equal times,

$$\{\pi_\Phi(x)\otimes\Phi(y)\} = t_0\delta(x,y) \tag{12}$$

Next, let us write

$$A \equiv A_x = A_{z_+} + A_{z_-}$$

Then, using eq.(12) we obtain

$$\{A(x)\otimes A(y)\} = \delta(x,y)\Big(e^\Phi\otimes e^\Phi[t_0,1+\mathcal{E}_+ - \mathcal{E}_+\otimes 1]e^{-\Phi}\otimes e^{-\Phi}$$
$$-e^{-\Phi}\otimes e^{-\Phi}[t_0,1\otimes\mathcal{E}_- - \mathcal{E}_-\otimes 1]e^\Phi\otimes e^\Phi\Big) \tag{13}$$

On the other hand let us introduce a matrix $r \in \mathcal{G} \otimes \mathcal{G}$ and write $A_1 = A \otimes 1$ and $A_2 = 1 \otimes A$. Then

$$[r, A_1 + A_2] = \delta(x,y)\Big(e^{\Phi} \otimes e^{\Phi}[r, 1 + \mathcal{E}_+ + \mathcal{E}_+ \otimes 1]e^{-\Phi} \otimes e^{-\Phi} + e^{-\Phi} \otimes e^{-\Phi}[r, 1 \otimes \mathcal{E}_- + \mathcal{E}_- \otimes 1]e^{\Phi} \otimes e^{\Phi}\Big) \quad (14)$$

Comparing eqs.(13) and (14) we see that if

$$[r, 1 \otimes \mathcal{E}_\pm + \mathcal{E}_\pm \otimes 1] = \pm[t_0, 1 \otimes \mathcal{E}_\pm - \mathcal{E}_\pm \otimes 1] \quad (15)$$

we have

$$\{A(x) \wp A(y)\} = \delta(x,y)[r, A_1(x) + A_2(y)]$$

The property (15) is a defining property for the r-matrix if we choose it to be antisymmetric. In particular for the $\widehat{sl}_2$ case we find

$$r = 2\Big(\sum_{\alpha \ positive} \frac{E_\alpha \otimes E_{-\alpha}}{(E_\alpha, E_{-\alpha})} - \frac{E_{-\alpha} \otimes E_\alpha}{(E_{-\alpha}, E_\alpha)}\Big) \quad (16)$$

If we call λ and μ the loop parameters corresponding to the first and second factor, respectively, in the tensor product, and set $x = \lambda/\mu$, we have

$$r = \frac{1+x}{1-x} H \otimes H + \frac{4}{1-x}(E_+ \otimes E_- + xE_- \otimes E_+)$$

3 A perturbative analysis.

After the above general presentation of the models, let us now concentrate on the affine sl_2 Toda field theory. The action corresponding to the equations of motion (3, 4, 5) is

$$S_0 = -\int dz_+ dz_- \left(\partial_+\varphi\partial_-\varphi + 2\partial_+\eta\partial_-\xi + m^2\left(e^{2\varphi} + e^{2\eta - 2\varphi}\right)\right) \quad (17)$$

Even though our intention is to solve the affine Toda theory exactly, we can use the perturbative Lagrangian viewpoint to have an insight into the quantum structure of the theory [3]. In particular we are interested in verifying how the classical reduction of the equation (3) to (2) by setting $\eta = \text{const}$, is recovered at the quantum level. This will give us information about the vacuum of the sinh-Gordon model.

To start perturbation theory one has to expand the action around a solution of the classical equations of motion (3,4) and (5). A very simple solution is

$$\varphi_0 = 0 = \eta_0, \qquad \xi_0 = -\tfrac{1}{2}m^2(z_+ - z_-)^2 \quad (18)$$

The last one is a periodic solution of

$$\partial_+\partial_-\xi_0 = m^2$$

Let us expand every field around this classical solution

$$\Phi = \Phi_0 + \Phi' \quad (19)$$

where Φ denotes collectively all the fields of the theory. Then, denoting for simplicity the fields Φ' with the same symbols as the original fields, the action (17) becomes

$$S = \int dxdt\Big[-\tfrac{1}{2}\varphi\Box\varphi - \eta\Box\xi - 2m^2(4\varphi^2 + 2\eta^2 - 4\eta\varphi + ...)\Big] \quad (20)$$

where dots denote cubic or higher interaction terms in which the ξ field never appears, and $\Box = \partial_t^2 - \partial_x^2$.

We now perform a Wick rotation and calculate the Fourier transformed propagators. They turn out to be

$$\begin{aligned} <\varphi,\varphi> &= \frac{1}{-\frac{k^2}{2} - 8m^2} \\ <\varphi,\xi> &= \frac{4m^2}{\frac{k^2}{2}(-\frac{k^2}{2} - 8m^2)} \\ <\eta,\xi> &= -\frac{1}{\frac{k^2}{2}} \\ <\xi,\xi> &= \frac{4m^2(-\frac{k^2}{2} - 4m^2)}{\frac{k^4}{4}(-\frac{k^2}{2} - 8m^2)} \end{aligned} \quad (21)$$

where $k^2 = k_0^2 + k_x^2$. The remaining propagators vanish. We see that the pole at $k^2 = -16m^2$ in the propagator $<\varphi,\varphi>$, corresponds to the massive particle of the sinh-Gordon model.

Let us consider now a scattering process whose external legs consist only of φ particles. From the structure of the propagators (21) and of the interaction terms, it is evident that in no connected graphs pertaining to this process will we have propagation of other modes than φ itself. The other modes decouple. Consequently for this kind of processes we can simply forget about η and ξ. This amounts to setting these fields to zero in the action. But then the action (20) becomes exactly the action of the sinh-Gordon model. *In other words as far as scattering processes involving massive φ modes are concerned, our theory coincides with the perturbative sinh-Gordon theory.*

Summarizing, the vacuum (18) is the correct vacuum to be chosen if we want to end up with the sinh-Gordon theory and the latter is recovered if we ignore some non-physical modes. Therefore we can regard the sinh-Gordon model as a spontaneously broken and reduced version of the conformal affine Liouville theory.

4 The Hamiltonian formalism

In the previous section we presented evidence that the sinh-Gordon theory can be regarded as a suitable reduction of a spontaneously broken conformal theory. With this in mind, we want now to proceed to find exact solutions of the affine Liouville model.
The first step is to single out the classical phase space as the space of local periodic solutions of the equations of motion (3, 4) and (5). To do this we will have to display the conformal structure of the affine Toda field theory. This amounts to splitting the chiralities, i.e. in particular to expressing the fields φ, η, ξ in terms of chiral free fields. As an intermediate step we will find the classical exchange algebra, which contains both the conformal structure and the integrable structure.
In this section we will follow closely ref.[4].

4.1 The energy-momentum tensor

We refer to the action (17), so that the conjugate momenta are

$$\begin{aligned} \pi_\varphi &= \partial_t \varphi \\ \pi_\eta &= \partial_t \xi \\ \pi_\xi &= \partial_t \eta \end{aligned}$$

The improved energy momentum tensor is defined as usual [1]:

$$\begin{aligned} T_{++} &= \tfrac{1}{4}(\mathcal{H} + \mathcal{P}) \\ T_{--} &= \tfrac{1}{4}(\mathcal{H} - \mathcal{P}) \end{aligned} \tag{22}$$

where

$$\begin{aligned} \mathcal{H} &= \pi_\varphi^2 + (\partial_x \varphi)^2 + 4n^2(e^{2\varphi} + e^{2\eta - 2\varphi}) - 2\partial_x^2 \varphi + \\ &\quad + 2\pi_\xi \pi_\eta + 2\partial_x \eta \partial_x \xi + B\partial_x^2 \eta - 4\partial_x^2 \xi \\ \mathcal{P} &= 2\pi_\varphi \partial_x \varphi + 2\pi_\eta \partial_x \eta + 2\pi_\xi \partial_x \xi - 2\partial_x \pi_\varphi + B\partial_x \pi_\xi - 4\partial_x \pi_\eta \end{aligned}$$

We easily check that

$$\partial_{z_-} T_{++} = 0 \qquad \partial_{z_+} T_{--} = 0$$

The Poisson bracket is defined by

$$\begin{aligned} \{\pi_\varphi(x), \varphi(y)\} &= \gamma \delta(x - y) \\ \{\pi_\eta(x), \eta(y)\} &= \gamma \delta(x - y) \\ \{\pi_\xi(x), \xi(y)\} &= \gamma \delta(x - y) \end{aligned}$$

Let

$$A(f) = \int_0^{2\pi} dx f(x) A(x)$$

Then

$$\{T_{++}(f), T_{++}(g)\} = -\gamma \int_0^{2\pi} dx f[\partial_x T_{++} + T_{++}\partial_x - \tfrac{1}{2}(1 - 2B)\partial_x^3]g$$

where we recognize the Virasoro algebra. Finally we notice that the current

$$\mathcal{J} = \partial_x \eta + \pi_\xi \tag{23}$$

is conserved.

4.2 Splitting the chiralities

The purpose of this section is to construct a periodic and local solution of the equations of motion (3,5), expressed in terms of chiral and antichiral free fields.
Let us define two periodic chiral fields taking values in the Cartan subalgebra $\widehat{\mathcal{H}}$ of $\widehat{sl}_2$,

$$\begin{aligned} P &= p_\theta H + 2p_\eta d + p_\zeta c \\ \bar{P} &= \bar{p}_\theta H + 2\bar{p}_\eta d + \bar{p}_\zeta c \end{aligned}$$

It was proved in [1] that there exists an object $Q_+(Q_-)$ which is an upper(lower)-triangular element of the group whose Lie algebra is $\widehat{sl}_2$ which can be represented by an infinite upper(lower) triangular matrix, is a function only of $z_+(z_-)$ and satisfies the *Drinfeld-Sokolov equations*

$$\partial_+ Q_+ = (P - \mathcal{E}_+) Q_+ \tag{24}$$

$$\partial_- Q_- = Q_-(-\bar{P} + \mathcal{E}_-) \tag{25}$$

These two equations are the basis of our construction. We consider the solution characterized by the boundary conditions

$$Q_+(0) = 1, \qquad Q_-(0) = 1$$

and for later use we introduce also the left and right monodromy matrices

$$S = Q_+(2\pi), \qquad \bar{S} = Q_-(2\pi)$$

We assume the following Poisson brackets

$$\begin{aligned} \{P(x) \overset{\otimes}{,} P(y)\} &= \tfrac{\gamma}{2} t_0 (\partial_x - \partial_y)\delta(x - y) \\ \{\bar{P}(x) \overset{\otimes}{,} \bar{P}(y)\} &= -\tfrac{\gamma}{2} t_0 (\partial_x - \partial_y)\delta(x - y) \end{aligned} \tag{26}$$

where we have used the definition

$$\delta(x) = \tfrac{1}{2\pi} \textstyle\sum_n e^{inx}$$

$P(x)$ and $\bar{P}(x)$ are assumed to be periodic, so we can Fourier expand them

$$\begin{aligned} P(x) &= \sum_n P_n e^{inx} \\ \bar{P}(x) &= \sum_n \bar{P}_n e^{inx} \end{aligned}$$

As a consequence we have the Poisson brackets

$$\begin{aligned} \{P_n \overset{\otimes}{,} P_m\} &= \tfrac{\gamma}{2\pi} in\delta_{n,-m} t_0 \\ \{\bar{P}_n \overset{\otimes}{,} \bar{P}_m\} &= -\tfrac{\gamma}{2\pi} in\delta_{n,-m} t_0 \end{aligned} \tag{27}$$

In the following an important role will be played by P_0 and $\bar{P}_0$. Eventually we will set

$$P_0 = \bar{P}_0 \tag{28}$$

but for the time being we keep them distinct. We will also need k_+ and k_-, defined as follows

$$\begin{aligned} k_+ &= \sum_{n\neq 0} \frac{iP_n}{n} \\ k_- &= \sum_{n\neq 0} \frac{i\bar{P}_n}{n} \end{aligned}$$

By definition we set

$$\begin{aligned} K_+(x) &= \int_0^x dy P(y) = k_+ + P_0 x + \sum_{n\neq 0} \frac{P_n}{in} e^{inx} \\ K_-(x) &= \int_0^x dy \bar{P}(y) = k_- + \bar{P}_0 x + \sum_{n\neq 0} \frac{\bar{P}_n}{in} e^{inx} \end{aligned}$$

The left and right degrees of freedom are independent, so

$$\{P_n \overset{\otimes}{,} \bar{P}_m\} = 0 \quad \forall n, m \tag{29}$$

4.3 Construction of a periodic local solution

In this subsection we explicitly construct a local periodic solution of (3,5). Here we summarize the construction as given in ref.[5], which follows closely ref.[4].
For any highest weight vector $|\Lambda_r>$ of $\widehat{sl}_2$ we define

$$\begin{aligned} \sigma^{(r)}(x) &= <\Lambda_r|Q_+(x) \\ \bar{\sigma}^{(r)}(x) &= Q_-(x)|\Lambda_r> \end{aligned} \tag{30}$$

These satisfy the exchange algebra

$$\begin{aligned} \{\sigma^{(r)}(x) \overset{\otimes}{,} \sigma^{(r')}(y)\} &= -\tfrac{\gamma}{2}\sigma^{(r)}(x)\otimes\sigma^{(r')}(y)\big[\theta(x-y)r^+ + \theta(y-x)r^-\big] \\ \{\bar{\sigma}^{(r)}(x) \overset{\otimes}{,} \bar{\sigma}^{(r')}(y)\} &= -\tfrac{\gamma}{2}\big[r^-\theta(x-y) + r^+\theta(y-x)\big]\bar{\sigma}^{(r)}(x)\otimes\bar{\sigma}^{(r')}(y) \end{aligned}$$

where

$$r^+ = t_0 + 4\sum_{\alpha\ positive} \frac{E_\alpha \otimes E_{-\alpha}}{(E_\alpha, E_{-\alpha})} \tag{31}$$

$$r^- = -t_0 - 4\sum_{\alpha\ positive} \frac{E_{-\alpha} \otimes E_\alpha}{(E_\alpha, E_{-\alpha})} \tag{32}$$

A solution to eqs.(3,5) is given (see [1]) by

$$e^{-\xi} = \sigma^{(0)} M \bar{\sigma}^{(0)}, \qquad e^{-\varphi-\xi} = \sigma^{(1)} M \bar{\sigma}^{(1)} \tag{33}$$

where M is a constant matrix to be determined. Since

$$\begin{aligned} \sigma^{(r)}(x+2\pi) &= \sigma^{(r)}(x) S \\ \bar{\sigma}^{(r)}(x+2\pi) &= \bar{S}\,\bar{\sigma}^{(r)}(x) \end{aligned}$$

to get a periodic solution, we must have

$$S\,M\,\bar{S} = M \tag{34}$$

We diagonalize the monodromy matrices (see [5])

$$\begin{aligned} S &= g\kappa g^{-1}, \qquad \kappa = e^{2\pi P_0} \\ \bar{S} &= \bar{g}^{-1}\bar{\kappa}\bar{g}, \qquad \bar{\kappa} = e^{-2\pi\bar{P}_0} \end{aligned} \tag{35}$$

Therefore we see that condition eq.(34) will be satisfied if

$$\begin{aligned} M &= gD\bar{g} \\ \kappa\bar{\kappa} &= 1 \end{aligned}$$

where $D \in \exp(\widehat{\mathcal{H}})$. The second condition simply means that $P_0 = \bar{P}_0$ and will be imposed at the end. The diagonal matrix D has to be chosen in such a way that the fields ξ and φ are local. The solution is

$$\begin{aligned} D &= \rho\bar{\rho} \\ \rho &= \theta e^{-k_+} \\ \bar{\rho} &= \bar{\theta} e^{k_-} \end{aligned}$$

and θ is related to the conjugate variable q of P_0 by

$$\theta = e^q$$

So

$$\begin{aligned} \{\theta \overset{\gamma}{\otimes} P_0\} &= \tfrac{\gamma}{2\pi}\theta \otimes 1 \cdot t_0 \\ \{\overline{\theta} \overset{\gamma}{\otimes} \overline{P}_0\} &= \tfrac{\gamma}{2\pi}\overline{\theta} \otimes 1 \cdot t_0 \end{aligned} \tag{36}$$

Therefore we can write

$$e^{-\xi} = \psi^0\,\overline{\psi}^0 \qquad e^{-\varphi-\xi} = \psi^1\,\overline{\psi}^1 \tag{37}$$

where we define the new objects (Block wave basis)

$$\psi^{(r)}(x) = \sigma^{(r)}(x) g\rho, \qquad \bar{\psi}^{(r)}(x) = \bar{\rho}\bar{g}\bar{\sigma}^{(r)}(x) \tag{38}$$

The ψ and $\bar{\psi}$ have diagonal monodromy

$$\psi^{(r)}(x+2\pi) = \psi^{(r)}(x)\,\kappa, \qquad \bar{\psi}^{(r)}(x+2\pi) = \bar{\kappa}\,\bar{\psi}^{(r)}(x) \tag{39}$$

and obey the exchange algebra

$$\begin{aligned} \{\psi^{(r)}(x) \overset{\gamma}{\otimes} \psi^{(r')}(y)\} &= -\tfrac{\gamma}{4}\psi^{(r)}(x) \otimes \psi^{(r')}(y)\Big\{\epsilon(x-y)(r^+ - r^-) - \\ &\quad -coth(\pi a d_1 P_0)(r^+ - t_0) - coth(\pi a d_2 P_0)(r^- + t_0)\Big\} \end{aligned} \tag{40}$$

$$\begin{aligned} \{\overline{\psi}^{(r)}(x) \overset{\gamma}{\otimes} \overline{\psi}^{(r')}(y)\} &= \tfrac{\gamma}{4}\Big\{\epsilon(x-y)(r^+ - r^-) + coth(\pi a d_1 \overline{P}_0)(r^- + t_0) \\ &\quad +coth(\pi a d_2 \overline{P}_0)(r^+ - t_0)\Big\}\overline{\psi}^{(r)}(x) \otimes \overline{\psi}^{(r')}(y) \end{aligned} \tag{41}$$

and since the two zero modes are considered as independent, we have

$$\{\psi^{(r)}(x) \overset{\gamma}{\otimes} \overline{\psi}^{(r')}(y)\} = 0 \tag{42}$$

Eq.(37) represents a general solution of eqs.(3, 5) which is both periodic and local and implements the chirality splitting.

We verify next that the solution (37) has all the required properties.
1 - *It is a solution*, as was proved in [4].
2 - *It is periodic.* To see this one should simply remember eqs.(39) and impose the condition (28).
3 - *It is local.* Using eqs.(40,41,42) one can verify that

$$\begin{aligned} &\{e^{-2\lambda^{(r)}_{max}(\Phi(x))}, e^{-2\lambda^{(r')}_{max}(\Phi(y))}\} = \tfrac{\gamma}{4}\psi^{(r)}(x) \otimes \psi^{(r')}(y)\cdot \\ &\Big\{[coth(\pi a d_1 P_0) + coth(\pi a d_2 \overline{P}_0)](r^+ - t_0) \\ &+[coth(\pi a d_2 P_0) + coth(\pi a d_1 \overline{P}_0)](r^- + t_0)\Big\}\overline{\psi}^{(r)}(x) \otimes \overline{\psi}^{(r')}(y) \end{aligned}$$

If $P_0 = \overline{P}_0$ this vanishes. It is now elementary to prove locality for the fields φ, η and ξ.

4.4 Conformal properties of ψ

In this subsection we briefly review the conformal properties of the objects introduced before. This is in order to single out those objects that behave tensorially under a conformal transformation. In this subsection we work in the + sector only and the corresponding label will be generally understood. Moreover we set, for simplicity, $\gamma = 1$. The energy momentum tensor of the + sector is

$$\mathcal{U} = -\tfrac{1}{2}\left(p_\theta^2 + p'_\theta + 2p_\eta p_\zeta - \tfrac{B}{2}p'_\eta + 2p'_\zeta\right)$$

Next we introduce the element Ξ, belonging to the Cartan subalgebra, defined by

$$\Xi = H + 4d - \tfrac{B}{2}c \tag{43}$$

It satisfies in particular

$$[\Xi, \mathcal{E}_+] = 2\mathcal{E}_+$$

Using this element we can write

$$\mathcal{U} = -\tfrac{1}{4}\left((P,P) + (\Xi,P')\right) \tag{44}$$

Let us set

$$\mathcal{U}(f) = \int_0^{2\pi} dx f(x)\mathcal{U}(x)$$

for any periodic function f. We find

$$\{\mathcal{U}(f), \mathcal{U}(g)\} = -\int_0^{2\pi} dx f\left(\partial_x \mathcal{U} + \mathcal{U}\partial_x - \tfrac{1}{2}(\Xi,\Xi)\partial_x^3\right)g \tag{45}$$

Proceeding as in ([4]), we obtain

$$\{\mathcal{U}(f), \psi^{(r)}(x)\} = f(x)\partial\psi^{(r)}(x) - \tfrac{1}{2}\lambda^{(r)}(\Xi)\partial f(x)\psi^{(r)}(x) \tag{46}$$

Therefore

$$h^{(r)} = -\tfrac{1}{2}\lambda^{(r)}(\Xi)$$

is the conformal weight of $\psi^{(r)}(x)$.
The basis with good tensorial properties is ψ basis.

5 The affine Liouville model on the lattice

Another useful way of formulating the affine Liouville theory, particularly in view of quantizing it, consists in putting the theory on the lattice.

We saw in section 5 that the main object in the theory is the Drinfeld-Sokolov linear system.

$$\partial Q_+ = (P - \mathcal{E}_+)Q_+$$

We have considered the solution of this equation normalized by

$$Q_+(0) = 1$$

It satisfies the following Poisson bracket: if $0 < x, y < 2\pi$ we have

$$\begin{aligned}\{Q_+(x) \overset{\otimes}{,} Q_+(y)\} &= \tfrac{1}{2} Q_+(x) \otimes Q_+(y) \cdot \\ &\cdot \Big\{\theta(x-y)[-r + Q_+^{-1}(y) \otimes Q_+^{-1}(y)(r - t_0)Q_+(y) \otimes Q_+(y)] \\ &+\theta(y-x)[-r + Q_+^{-1}(x) \otimes Q_+^{-1}(x)(r + t_0)Q_+(x) \otimes Q_+(x)]\Big\}\end{aligned} \tag{47}$$

Let us consider now a lattice with N sites of length Δ (later on we will impose periodic boundary conditions). It is natural to write the discrete version of (47) in the following way

$$\begin{aligned}\{Q_n \overset{\otimes}{,} Q_m\} &= \tfrac{1}{2} Q_n \otimes Q_m \cdot \\ &\cdot \Big\{\theta(n-m)[-r + Q_m^{-1} \otimes Q_m^{-1}(r - t_0)Q_m \otimes Q_m] \\ &+\theta(m-n)[-r + Q_n^{-1} \otimes Q_n^{-1}(r + t_0)Q_n \otimes Q_n]\Big\}\end{aligned} \tag{48}$$

We want to recover the above Poisson bracket in a constructive way. To this end we define on each lattice site the fields π_n, ϕ_n with values in $\widehat{\mathcal{H}}$ – one should not mistake ϕ_n with the discretization of the field φ of the previous sections – and Poisson bracket

$$\{\pi_n \overset{\otimes}{,} \phi_m\} = \tfrac{1}{2} t_0 \frac{\delta_{nm}}{\Delta}$$

Let us define next

$$L_n = e^{\Delta \pi_n} N^{-1} e^{\phi_{n-1} - \phi_n} \tag{49}$$

where

$$N = e^{\Gamma}, \qquad \Gamma = \sum_{n=0}^{\infty} \frac{\Delta^{2n+1}}{2n+1}(E_n^+ + E_{n+1}^-) \tag{50}$$

Our aim is to prove that

$$\begin{aligned}\{L_n \overset{\otimes}{,} L_m\} &= \tfrac{1}{2}\delta_{nm}[r, L_n \otimes L_m] \\ &- \tfrac{1}{2}\delta_{n,m+1} L_n \otimes 1 \cdot t_0 \cdot 1 \otimes L_m \\ &+ \tfrac{1}{2}\delta_{n,m-1} 1 \otimes L_m \cdot t_0 \cdot L_n \otimes 1\end{aligned} \tag{51}$$

where r is the matrix introduced in section 3, eq.(16).
A brief sketch of the proof. An explicit computation gives

$$\begin{aligned}\{L_n \overset{\otimes}{,} L_m\} &= -\tfrac{1}{2}\delta_{nm}[L_n \otimes 1 \cdot t_0 \cdot 1 \otimes L_n - 1 \otimes L_n \cdot t_0 \cdot L_n \otimes 1] \\ &+\tfrac{1}{2}\delta_{n,m-1} 1 \otimes L_m \cdot t_0 \cdot L_n \otimes 1 \\ &-\tfrac{1}{2}\delta_{n,m+1} L_n \otimes 1 \cdot t_0 \cdot 1 \otimes L_m\end{aligned}$$

The last two terms have the right form. Let us consider the first one. We have

$$\begin{aligned}\{L_n \overset{\otimes}{,} L_n\} &= e^{\Delta\pi_n} \otimes e^{\Delta\pi_n} \cdot N^{-1} \otimes N^{-1}\Big(1 \otimes N \cdot t_0 \cdot 1 \otimes N^{-1} - \\ &-N \otimes 1 \cdot t_0 \cdot N^{-1} \otimes 1\Big) e^{\phi_{n-1} - \phi_n} \otimes e^{\phi_{n-1} - \phi_n}\end{aligned}$$

On the other hand we have

$$[r, L_n \otimes L_m] = e^{\Delta\pi_n} \otimes e^{\Delta\pi_n}[r, N^{-1} \otimes N^{-1}] e^{\phi_{n-1} - \phi_n} \otimes e^{\phi_{n-1} - \phi_n}$$

So it is enough to prove that

$$[r, N^{-1} \otimes N^{-1}] = N^{-1} \otimes 1 \cdot t_0 \cdot 1 \otimes N^{-1} - 1 \otimes N^{-1} \cdot t_0 \cdot N^{-1} \otimes 1$$

This equation can be checked explicitly, bearing in mind in particular eq. (15).
Next we define

$$Q_n = L_n L_{n-1} \ldots L_1, \tag{52}$$

Then it is not hard to show that Q_n satisfies eq.(48). Conversely, if we define

$$L_n = Q_n Q_{n-1}$$

from (48) we find (51).
Next we consider periodic boundary conditions on the lattice

$$L_{N+n} = L_n$$

We define

$$S = L_N L_{N-1} \cdots L_1$$

We have in particular

$$Q_{N+n} = Q_n S$$

We find

$$\begin{aligned}\{Q_n \overset{\otimes}{,} S\} &= \tfrac{1}{2} Q_n \otimes S\Big[-r + Q_n^{-1} \otimes Q_n^{-1}(r + t_0)Q_n \otimes Q_n \\ &\qquad -1 \otimes S^{-1} \cdot t_0 \cdot 1 \otimes S\Big]\end{aligned} \tag{53}$$

$$\begin{aligned}\{S \overset{\otimes}{,} S\} &= \tfrac{1}{2} S \otimes S\Big[-r + S^{-1} \otimes S^{-1} \cdot r \cdot S \otimes S \\ &\qquad +S^{-1} \otimes 1 \cdot t_0 \cdot S \otimes 1 - 1 \otimes S^{-1} \cdot t_0 \cdot 1 \otimes S\Big]\end{aligned} \tag{54}$$

Finally, in the continuum case, we introduced the important matrix $\rho \in \exp(\widehat{\mathcal{H}})$. In the lattice we define

$$\rho = e^{-\phi_N} e^{\pi_0 \Delta} \tag{55}$$

and obtain immediately

$$\{\rho \,\hat{\otimes}\, \rho\} = 0 \tag{56}$$
$$\{Q_n \,\hat{\otimes}\, \rho\} = -\tfrac{1}{2} Q_n \otimes \rho \cdot t_0 \tag{57}$$
$$\{S \,\hat{\otimes}\, \rho\} = -\tfrac{1}{2} S \otimes \rho \cdot t_0 - \tfrac{1}{2} t_0 \cdot S \otimes \rho \tag{58}$$

The exchange algebra on the lattice is now easy to write down. From eq.(48), defining $\sigma_n = <\Lambda| Q_n$ for any highest weight vector $<\Lambda|$ of $\widehat{sl}_2$, we have

$$\begin{aligned} \{\sigma_n \,\hat{\otimes}\, \sigma'_m\} &= -\tfrac{1}{2} \sigma_n \otimes \sigma'_m \cdot r^+ \qquad n > m \\ \{\sigma_n \,\hat{\otimes}\, \sigma'_m\} &= -\tfrac{1}{2} \sigma_n \otimes \sigma'_m \cdot r^- \qquad n < m \\ \{\sigma_n \,\hat{\otimes}\, \sigma'_n\} &= -\tfrac{1}{2} \sigma_n \otimes \sigma'_n \cdot r^+ + \tfrac{t_0^{\Lambda\Lambda'}}{2} \sigma_n \otimes \sigma'_n \end{aligned} \tag{59}$$

where $t_0^{\Lambda\Lambda'}$ is the eigenvalue of t_0 on $|\Lambda> \otimes |\Lambda'>$. Then one can diagonalize S as in the continuum case

$$S = g e^{2\pi P_0} g^{-1}$$

where $P_0 \in \widehat{\mathcal{H}}$ is by definition the zero mode, and define the Bloch wave basis

$$\psi_n = \sigma_n g \rho$$

The corresponding exchange algebra can be straightforwardly calculated. Similarly in the other sector we can introduce $\bar{\psi}_n$ (related to the zero mode $\bar{P}_0$) and compute the relevant exchange algebra. Then one defines

$$e^{-\varphi_n} = \psi_n \bar{\psi}_n$$

This field is periodic and local in the reduced phase space $P_0 = \bar{P}_0$. For the sake of conciseness the details are not spelt out here.

6 Spontaneous symmetry breaking and reduction

So far we have seen how to single out the classical phase space of the theory, both in the continuum and on the lattice. However our main interest is in the sinh-Gordon theory. We already saw that, in order to end up with the latter theory, we have to reduce the phase space and choose a specific vacuum. At a classical level the vacuum is represented by the solution (18) and one can see that the reduction is implemented by imposing the the constraints

$$\mathcal{J}(x) = \partial_{z_+} \eta = 0 \tag{60}$$
$$\overline{\mathcal{J}}(x) = \partial_{z_-} \eta = 0 \tag{61}$$

Notice that we have

$$\{\mathcal{J}(x), \mathcal{J}(y)\} = 0$$

and a similar equation for $\overline{\mathcal{J}}$, so that we are dealing with first class constraints. It turns out that this reduction is well defined. For example, e^{φ} belongs to the reduced phase space, while e^{ξ} does not.
Let us begin now the discussion about the quantum theory. We have seen that the classical phase space splits into a chiral and an antichiral sector (up to the zero modes which must be common to both sector). The chiral (i.e. $+$) sector can be formulated in terms of the periodic field

$$P = p_\theta H + 2 p_\eta d + p_\zeta c$$

We recall that

$$p_\eta = -\partial_+ \eta_+$$

In conformity with eq.(27), we assume the following Dirac brackets

$$[P^n \,\hat{\otimes}\, P^m] = n \delta_{n,-m} t_0$$

where, for the sake of simplicity we have introduced the redefinition

$$P^n \longrightarrow \sqrt{\frac{\gamma}{2\pi}} P^n$$

The above corresponds to

$$[p_\theta^n, p_\theta^m] = n \delta_{n,-m} \tag{62}$$
$$[p_\eta^n, p_\zeta^m] = n \delta_{n,-m} \tag{63}$$

the remaining brackets being zero. We define the vacuum vector $|0>$ by

$$p_\theta^n |0> = 0, \qquad p_\zeta^n |0> = 0, \qquad p_\eta^n |0> = 0, \qquad n > 0 \tag{64}$$
$$p_\theta^0 |0> = \lambda_\theta |0>, \qquad p_\eta^0 |0> = \lambda_\eta |0>, \qquad p_\zeta^0 |0> = \lambda_\zeta |0>$$

where $\lambda_\theta, \lambda_\eta, \lambda_\zeta$ are numbers. The Fock space $\mathcal{F}_+$ is spanned by the vectors

$$|\psi> = p_\theta^{n_1} \ldots p_\theta^{n_r} p_\eta^{m_1} \ldots p_\eta^{m_s} p_\zeta^{l_1} \ldots p_\zeta^{l_t} |0> \tag{65}$$

with all $n_i, m_i, l_i < 0$.
The reduced Fock space $\mathcal{F}'_+$ will be determined by imposing the quantum version of (60)

$$p_\eta^n |phys> = 0, \qquad n > 0 \tag{66}$$

It is interesting to notice that, while $\mathcal{F}_+$ contains negative norm states, the states of $\mathcal{F}'_+$ have nonnegative norms, and once we mode out the zero norm states with a standard procedure, we obtain a Hilbert space.

6.1 The sinh-Gordon vacuum

Let us come now to the vacuum. The vacuum $|0>$, according to the previous definition, corresponds classically to the condition $P(x) = P_0 = \text{const}$. We can find the corresponding classical solution by using eqs.(37, 38). It is of course of crucial importance to see under what condition we can recover the sinh-Gordon vacuum (18).

To this end let us consider the case P =const, with the additional condition

$$p_\eta^0 = 2p_\theta^0, \qquad \bar{p}_\eta^0 = 2\bar{p}_\theta^0 \tag{67}$$

Then

$$P = p_\theta^0 H' + p_\zeta^0 \hat{c}, \qquad \bar{P} = \bar{p}_\theta^0 H' + \bar{p}_\zeta^0 \hat{c}$$

where we have introduced the notation $H' = H + 4\hat{d}$. Notice the commutator

$$[H', \mathcal{E}_\pm] = \pm 2\mathcal{E}_\pm \tag{68}$$

Then we have

$$\begin{aligned} Q_+(z_+) &= e^{(p_\theta^0 H_+ + p_\zeta^0 \hat{c} - m\mathcal{E}_+)z_+} = \\ &= e^{\frac{m}{2p_\theta^0}\mathcal{E}_+} e^{(p_\theta^0 H_+ + p_\zeta^0 \hat{c})z_+} e^{-\frac{m}{2p_\theta^0}\mathcal{E}_+} \end{aligned}$$

Similarly

$$\begin{aligned} Q_-(z_-) &= e^{-(\bar{p}_\theta^0 H_+ + \bar{p}_\zeta^0 \hat{c} - m\mathcal{E}_-)z_-} = \\ &= e^{-\frac{m}{2\bar{p}_\theta^0}\mathcal{E}_-} e^{-(\bar{p}_\theta^0 H_+ + \bar{p}_\zeta^0 \hat{c})z_-} e^{\frac{m}{2\bar{p}_\theta^0}\mathcal{E}_-} \end{aligned}$$

Let us define the diagonalizing elements

$$g = e^{\frac{m}{2p_\theta^0}\mathcal{E}_+}, \qquad \bar{g} = e^{\frac{m}{2\bar{p}_\theta^0}\mathcal{E}_-} \tag{69}$$

From (37) the solution of the equations of motion is given by

$$\begin{aligned} e^{-\xi(x,t)} &= <\Lambda_0|Q_+(z_+)g\rho\bar{\rho}\bar{g}Q_-(z_-)|\Lambda_0> = \\ e^{-\xi(x,t)-\varphi(x,t)} &= <\Lambda_1|Q_+(z_+)g\rho\bar{\rho}\bar{g}Q_-(z_-)|\Lambda_1> = \end{aligned}$$

A simple calculation (setting for simplicity all the q's to zero) gives

$$e^{-\xi(x,t)} = e^{p_\zeta^0 z_+ - \bar{p}_\zeta^0 z_-} e^{\frac{m^2}{4p_\theta^0 \bar{p}_\theta^0} e^{-2(p_\theta^0 z_+ + \bar{p}_\theta^0 z_-)}} \tag{70}$$

Setting now

$$p_\zeta^0 = \bar{p}_\zeta^0, \qquad p_\eta^0 = \bar{p}_\eta^0$$

we find

$$\xi(x,t) = -2p_\zeta^0 t - \frac{m^2}{4(p_\theta^0)^2} e^{-4p_\theta^0 t} \tag{71}$$

Similarly

$$e^{-\varphi(x,t)} = e^{p_\theta^0 z_+ - \bar{p}_\theta^0 z_-}$$

i.e.

$$\varphi(x,t) = -2p_\theta^0 t \tag{72}$$

We see that while (72) is regular in the limit $p_\theta^0 \to 0$, (71) is singular. However we may take advantage of the fact that the solution of eq.(5) is always determined up to a solution of $\partial_+\partial_-\xi = 0$. Therefore we may add to the solution (71) the term

$$\frac{m^2}{4(p_\theta^0)^2}(1 - 4p_\theta^0 t) + \alpha t + \beta$$

where α and β are arbitrary constants. This renders the solution smooth in the limit $p_\eta^0 = 2p_\theta^0 \to 0$. In this limit we finally obtain

$$\begin{aligned} \eta &= 0 \\ \varphi &= 0 \\ \xi &= \alpha t + \beta - 2m^2 t^2 \end{aligned} \tag{73}$$

In other words we have recovered the sinh-Gordon solution (18) by taking the limit $p_\eta^0 = 2p_\theta^0 \to 0$ of the general P =const solution and by suitably renormalizing it.

So far we have learnt how to reduce the affine Liouville theory to the sinh-Gordon theory, and how to recover the sinh-Gordon vacuum among the vacua of the conformal affine theory – through a limiting procedure. We should then find the quantum exchange algebra in the affine theory as an intermediate step in calculating correlation functions. As an introduction to this program let us see how the quantum exchange algebra can be calculated in the ordinary Liouville theory.

7 Quantum Liouville theory on the lattice

One way to introduce the necessary regularization of the quantum theory is to put the theory on the lattice. This was already done for the classical affine Liouville

theory in section 6. The corresponding formulae for the classical ordinary Liouville model are obtained from those of section 6 by simply removing the unwanted generators and fields.
The first thing to do is to find the quantum version of eq(51). A hint at the correct formulae is obtained by making the substitution

$$[\ ,\ ] = i\hbar\{\ ,\ \}$$

Suppose first $n = m$. Introducing the notation

$$\begin{aligned} L_{1n} &= L_n \otimes 1 \\ L_{2n} &= 1 \otimes L_n \end{aligned}$$

we find from eq(51)

$$L_{1n}L_{2n} - L_{2n}L_{1n} = \tfrac{i\hbar}{2}[r, L_{1n}L_{2n}]$$

or

$$(1 - \tfrac{i\hbar}{2}r)L_{1n}L_{2n} = L_{2n}L_{1n}(1 - \tfrac{i\hbar}{2}r)$$

We recognize the well-known formula

$$\begin{aligned} R_{12}L_{1n}L_{2n} &= L_{2n}L_{1n}R_{12} \qquad (74) \\ R_{12} &\longrightarrow 1 - \tfrac{i\hbar}{2}r + O(\hbar^2) \end{aligned}$$

Similarly we consider the case $m = n + 1$. A natural quantum formula is

$$L_{1n}L_{2,n+1} = L_{2,n+1}A_{12}L_{1n} \qquad (75)$$

where

$$A_{12} \longrightarrow 1 + \tfrac{i\hbar}{2}H \otimes H + O(\hbar^2)$$

Therefore, after introducing

$$A_{12} = A_{21} = e^{\frac{i\hbar}{2}H\otimes H}$$

we are lead to the following quantum version of eq.(51):

$$\begin{aligned} R_{12}L_{1n}L_{2n} &= L_{2n}L_{1n}R_{12} \qquad (76) \\ L_{1n}L_{2,n+1} &= L_{2,n+1}A_{12}L_{1n} \qquad (77) \\ L_{2n}L_{1,n+1} &= L_{1,n+1}A_{21}L_{2n} \end{aligned}$$

When we come to the quantum relations satisfied by the Q_n, something new happens. In fact the correct definition for Q_n is

$$Q_n = L_nBL_{n-1}B\cdots BL_1 \qquad (78)$$

where

$$B = e^{\frac{i\hbar}{4}H^2}$$

Then, as a consequence of eqs.(76,77), we find

$$\begin{aligned} R_{12}Q_{1n}Q_{2n} &= Q_{2n}Q_{1n}R_{12}, \qquad n < N \qquad (79) \\ Q_{1n}Q_{1m}^{-1}A_{21}R_{12}Q_{1m}Q_{2m} &= Q_{2m}Q_{1n}R_{12} \qquad n > m,\ \ n,m < N \qquad (80) \end{aligned}$$

In a periodic lattice with N sites we have, in particular,

$$\begin{aligned} L_{1N}L_{21} &= L_{21}A_{12}L_{1N} \\ L_{2N}L_{11} &= L_{11}A_{12}L_{2N} \end{aligned}$$

and we define

$$S = L_NBL_{N-1}\cdots BL_1$$

so that

$$Q_{n+N} = Q_nBS \qquad (81)$$

The quantum analogs of eqs.(54,53) are

$$\begin{aligned} R_{12}S_1A_{12}S_2 &= S_2A_{12}S_1R_{12} \qquad (82) \\ S_1Q_{1n}^{-1}A_{12}R_{12}Q_{1n}Q_{2n} &= Q_{2n}A_{12}S_1R_{12} \qquad (83) \end{aligned}$$

Finally, the quantum analog of the constant $\rho \in \exp(\mathcal{H})$ satisfies the following conditions

$$\begin{aligned} \rho_1\rho_2 &= \rho_2\rho_1 \qquad (84) \\ A_{12}S_1\rho_2A_{12} &= \rho_2S_1 \qquad (85) \\ Q_{1n}\rho_2A_{12} &= \rho_2Q_{1n} \qquad (86) \end{aligned}$$

7.1 Quantum exchange algebra.

With the above formalae we can immediately derive the exchange algebra. Let

$$\sigma_n^{(r)} = < \lambda_{max}^{(r)}|Q_n$$

then we have:
if $n \neq m$

$$\sigma_{1n}^{(r)}\ \sigma_{2m}^{(r')} = \sigma_{2m}^{(r')}\ \sigma_{1n}^{(r)}\ R_{12}^{\pm}(q) \qquad (87)$$

where +(-) refers to $n > m (n < m)$. If $n = m$

$$q^{\frac{1}{2}\lambda^{(r)}_{max}\cdot\lambda^{(r')}_{max}}\,\sigma^{(r)}_{1n}\,\sigma^{(r')}_{2n} = \sigma^{(r')}_{2n}\,\sigma^{(r)}_{1n}\,R^{+}_{12}(q) \tag{88}$$

Let us consider spin $\frac{1}{2}$ representations of sl_2. Then ($q = e^{-i\hbar}$)

$$R^{+}_{12} = \begin{pmatrix} q^{\frac{1}{2}} & 0 & 0 & 0 \\ 0 & q^{-\frac{1}{2}} & q^{\frac{1}{2}} - q^{-\frac{3}{2}} & 0 \\ 0 & 0 & q^{-\frac{1}{2}} & 0 \\ 0 & 0 & 0 & q^{\frac{1}{2}} \end{pmatrix} \qquad R^{-}_{12} = \begin{pmatrix} q^{-\frac{1}{2}} & 0 & 0 & 0 \\ 0 & q^{\frac{1}{2}} & 0 & 0 \\ 0 & q^{-\frac{1}{2}} - q^{\frac{3}{2}} & q^{\frac{1}{2}} & 0 \\ 0 & 0 & 0 & q^{-\frac{1}{2}} \end{pmatrix}$$

$$A_{12} = \begin{pmatrix} q^{-\frac{1}{2}} & 0 & 0 & 0 \\ 0 & q^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & q^{\frac{1}{2}} & 0 \\ 0 & 0 & 0 & q^{-\frac{1}{2}} \end{pmatrix}$$

Moreover $H^2 = 1$ and so $B = q^{-\frac{1}{4}}$.
Let us consider the exchange algebra (87, 88) in this case. Defining

$$(\sigma^1_n, \sigma^2_n) = (1, 0)Q_n$$

we find explicitly:
1) $n > m$

$$\begin{aligned} \sigma^1_n\sigma^1_m &= q^{\frac{1}{2}}\sigma^1_m\sigma^1_n \\ \sigma^1_n\sigma^2_m &= q^{-\frac{1}{2}}\sigma^2_m\sigma^1_n \\ \sigma^2_n\sigma^1_m &= q^{-\frac{1}{2}}\sigma^1_m\sigma^2_n + (q^{\frac{1}{2}} - q^{-\frac{3}{2}})\sigma^2_m\sigma^1_n \\ \sigma^2_n\sigma^2_m &= q^{\frac{1}{2}}\sigma^2_m\sigma^2_n \end{aligned} \tag{89}$$

2) $n < m$

$$\begin{aligned} \sigma^1_n\sigma^1_m &= q^{-\frac{1}{2}}\sigma^1_m\sigma^1_n \\ \sigma^1_n\sigma^2_m &= q^{\frac{1}{2}}\sigma^2_m\sigma^1_n + (q^{-\frac{1}{2}} - q^{\frac{3}{2}})\sigma^1_m\sigma^2_n \\ \sigma^2_n\sigma^1_m &= q^{\frac{1}{2}}\sigma^1_m\sigma^2_n \\ \sigma^2_n\sigma^2_m &= q^{-\frac{1}{2}}\sigma^2_m\sigma^2_n \end{aligned}$$

3) $n = m$

$$\sigma^1_n\sigma^2_n = q^{-1}\sigma^2_n\sigma^1_n$$

Just as in the classical case, we can give an explicit representation of all these objects and relations in terms of the scalar field ϕ_n and its conjugate momentum π_n, such that

$$[\pi_n, \varphi_m] = i\hbar\frac{\delta_{nm}}{\Delta}$$

(see [6]).
But let us concentrate now on the exchange algebra. Like in the classical case we want to arrive at the exchange algebra for the Bloch wave basis – the ψ_n basis (for the spin $\frac{1}{2}$ representation of sl_2). The monodromy matrix S is

$$S = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

By definition A and D are related to the zero mode P_0 by

$$A = q^{\frac{N-1}{4}}e^{2\pi P_0} \qquad D = q^{\frac{N-1}{4}}e^{-2\pi P_0}$$

For example, from eq.(82) we find that A, B, D commute.
Next let us define the matrix g by

$$S = g^{-1}\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} g$$

Therefore

$$g = \begin{pmatrix} 1 & \frac{B}{A-D} \\ 0 & 1 \end{pmatrix}, \quad g^{-1} = \begin{pmatrix} 1 & -\frac{B}{A-D} \\ 0 & 1 \end{pmatrix}$$

The matrix ρ has the form

$$\rho = \begin{pmatrix} \rho^1 & 0 \\ 0 & \rho^2 \end{pmatrix}$$

If we define the line vector $\psi_n = (\psi^1_n, \psi^2_n)$ by $\psi_n = \sigma_n g^{-1}\rho$, and use eqs.(83,85,86) and (89), we finally end up with
1) $n > m$

$$\begin{aligned} \psi^1_n\psi^1_m &= q^{\frac{1}{2}}\psi^1_m\psi^1_n \\ \psi^1_n\psi^2_m &= q^{-\frac{1}{2}}\psi^2_m\psi^1_n + (q^{\frac{1}{2}} - q^{-\frac{3}{2}})\psi^1_m\psi^2_n\frac{A}{A-D} \\ \psi^2_n\psi^1_m &= q^{-\frac{1}{2}}\psi^1_m\psi^2_n\frac{(qA - q^{-1}D)(q^{-1}A - qD)}{(A-D)^2} - (q^{\frac{1}{2}} - q^{-\frac{3}{2}})\psi^2_m\psi^1_n\frac{D}{A-D} \\ \psi^2_n\psi^2_m &= q^{\frac{1}{2}}\psi^2_m\psi^2_n \end{aligned}$$

2) $n = m$

$$\psi^1_n\psi^2_n - \frac{A-D}{q^{-1}A - qD}\psi^2_n\psi^1_n = 0$$

3) $n < m$

$$\begin{aligned} \psi^1_n\psi^1_m &= q^{-\frac{1}{2}}\psi^1_m\psi^1_n \\ \psi^1_n\psi^2_m &= q^{\frac{1}{2}}\psi^2_m\psi^1_n - (q^{-\frac{1}{2}} - q^{\frac{3}{2}})\psi^1_m\psi^2_n\frac{D}{A-D} \\ \psi^2_n\psi^1_m &= q^{\frac{1}{2}}\psi^1_m\psi^2_n\frac{(qA - q^{-1}D)(q^{-1}A - qD)}{(A-D)^2} + (q^{-\frac{1}{2}} - q^{\frac{3}{2}})\psi^2_m\psi^1_n\frac{A}{A-D} \\ \psi^2_n\psi^2_m &= q^{-\frac{1}{2}}\psi^2_m\psi^2_n \end{aligned}$$

We can find an analogous exchange algebra for the second chirality. As long as the zero modes of the two chiralities are independent dynamical variables, the two chiralities commute

$$\psi_{1n}\overline{\psi}_{2m} = \overline{\psi}_{2m}\psi_{1n}$$

7.2 Locality and periodicity

Let

$$\begin{aligned} \psi_n &= (\psi_n^1, \psi_n^2) \\ \overline{\psi}_n &= \begin{pmatrix} \overline{\psi}_n^1 \\ \overline{\psi}_n^2 \end{pmatrix} \end{aligned}$$

be the two Bloch waves as above. Define

$$e^{-\varphi_n} = \psi_n\overline{\psi}_n = \psi_n^1\overline{\psi}_n^1 + \psi_n^2\overline{\psi}_n^2 \qquad (90)$$

This field will be periodic (up to a trivial rescaling) if

$$A\overline{A} = D\overline{D}$$

Let us call $\mathcal{H}_{sub}$ the subspace of the Hilbert space such that

$$(A\overline{A} - D\overline{D})\mathcal{H}_{sub} = 0$$

This amounts to saying that $P_0 = \overline{P}_0$ on $\mathcal{H}_{sub}$. One can prove [6] that
1) The field $e^{-\varphi_n}$ admits a restriction to the subspace $\mathcal{H}_{sub}$.
2) The restriction of $e^{-\varphi_n}$ to $\mathcal{H}_{sub}$ is periodic and local, the latter property meaning that

$$[e^{-\varphi_n}, e^{-\varphi_m}]\mathcal{H}_{sub} = 0$$

We have seen how to obtain the Bloch wave quantum exchange algebra (in the simplest case, for a general treatment see [7]). The objects (90) are the building blocks to construct correlation functions. The program is to generalize this construction to the affine Liouville model and to calculate the reduction to the sinh-Gordon model.

More details and references to the literature can be found in the following:

References

[1] O.Babelon and L.Bonora, Phys.Lett. 244B(1990), 220.

[2] L.Bonora, M.Martellini and Y.Z.Zhang, Phys.Lett. 253B(1991), 373.

[3] O.Babelon and L.Bonora, in preparation.

[4] O.Babelon, L.Bonora and F.Toppan, "Exchange algebra and the Drinfeld-Sokolov theorem" I.S.A.S. preprint 65/90/EP, to appear in CMP.

[5] O.Babelon and L.Bonora, "Exchange algebra in the conformal affine sl_2 Toda filed theory" Proc. Rapallo Conference 1990.

[6] O.Babelon and L.Bonora, "Quantum toda theory", Phys.Lett. 253B (1991), 365.

[7] O.Babelon, "Universal exchange algebra for Bloch waves and Liouville theory" PAR LPTHE 91/11.

KdV TYPE EQUATIONS IN CONFORMAL FIELD THEORY AND SUBCRITICAL STRINGS

Pierre MATHIEU

Département de Physique
Université Laval, Québec, Canada, G1K 7P4

Lecture 1: **KdV type equations and conformal algebras** (summary)

In this lecture we review the relation between the second Hamiltonian structure of KdV type equations and extended conformal algebras. The Virasoro algebra realized in terms of Poisson brackets is the Fourier decomposition of the second Poisson structure for the KdV equation[1]. The Miura transformation, which is a canonical map from the natural mKdV Poisson structures to the KdV second Poisson structure, can be regarded as a free field representation of the classical Virasoro algebra[2]. The quantum Virasoro algebra can thus be constructed canonically from this Miura map, whose quantum part is the Feigin-Fuchs representation[3]. This result generalizes to W algebras which are just the second Gelfand-Dickii Poisson structures[4,5]. For instance the W_3 algebra of Zamolodchikov[6] corresponds to the second Hamiltonian structure of the Boussinesq equation[4,7]. From the Drinfeld-Sokolov theory[8] these W algebras can be classified in terms of Lie algebras. This connection also extends to the super case. The classical form of the N = 1 superconformal algebra is the second Poisson structure of two simple fermionic extensions of the KdV equation: the Kupershmidt system[9,10] which has no supersymmetry invariance and which is related

onto osp(2,1)[11] - and the Manin-Radul system[12,13] - which is supersymmetric and related to $osp(2,2)^{(2)}$[14]. (Both are bi-Hamiltonian[9,15]). Similarly the classical form of the N=2 superconformal algebra is the second Poisson structure of four 0(2) fermionic extensions of the KdV equation: the osp(2,2) KdV equation[16], which is the natural 0(2) extension of the Kupershmidt system and thus not supersymmetric, and three N = 2 super KdV equations[16,17]. However for one of these the integrability is only a conjecture[17]. The other two have been shown to be bi-Hamiltonian[15]. One of them is naturally related to sl(2,2)[18] and to the N=2 super sine-Gordon equation[19]. For N = 3, 4 superconformal algebras, there are no known related integrable super KdV equations. There are also two infinite sequences of "super" conformal algebras, the SO(N) and U(N) superconformal algebras,[20] which are neither supersymmetric nor linear (they contain quadratic terms, like the W algebras). The first is related to the osp(N,2) KdV equation[21,22] which can be regarded as the straightforward O(N) extension of the Kupershmidt system when rewritten in terms of appropriate (nonlocal) fields. Its Lax operator is known and its factorization yields a Miura map. Similarly the U(N) superconformal algebra can be related to an integrable (since it is bi-Hamiltonian) U(N) fermionic KdV equation[22] for which however the Lax operator is not known. Nevertheless by analogy with the SO(N) case, the Miura map can be found easily. When properly quantized these maps furnish a representation of these algebras in terms of one free boson, a decoupled (SO(N) or U(N)) Kac-Moody algebra and N free (real or complex) fermions. In contrast with the W algebras, these superconformal algebras where known before the importance of the KdV Virasoro relation was established. Thus in the supercase, the KdV Virasoro relation led to the discovery of new super KdV equations and, for the two infinite sequences, of a simple representation of these algebras.

Lecture 2: **Quantum KdV type equations in perturbed conformal field theories** (summary)

By considering the perturbation of the unitary minimal models in conformal field theory with relevant operators, namely with $\Phi_{1,3}$, $\Phi_{1,2}$ and $\Phi_{2,1}$ (in the BPZ notation[1]), Zamolodchikov noticed that the resulting massive theories have nontrivial commuting conservation laws, presumably in infinite number, which suggests that the resulting models are completely integrable[2]. It turns out that these sets of conservation laws correspond respectively to those of the quantum KdV equation, the quantum Gibbon-Sawada-Kotera equation and the quantum Kupershmidt equation[3]. In their classical versions these three equations are the only known integrable systems (modulo their higher order analogues in their respective hierarchy) which can be written as Hamiltonian systems in terms of the second KdV Poisson bracket. The quantum form of these equations is obtained canonically with the Poisson bracket replaced by the Virasoro algebra (or the operator product of the energy-momentum tensor with itself) and the classical Hamiltonian replaced by its quantum version, where field products are replaced by normal ordered products. These quantum equations are conjectured to be completely integrable. However a direct proof of integrability is still missing. (The standard classical tricks to prove integrability fail, i.e: the search for an additional Hamiltonian structure, the Lax representation, the search for an integrable deformation[3]). The equivalence between the conservation laws of the perturbed theory and those of the above quantum systems is established formally by first relating the latter to appropriate Toda systems[3,5] (sine-Gordon for KdV and Bullough-Dodd for the other two) at particuliar values of their coupling constants, and then to use the Feigin-Fuchs representation, where Virasoro primary fields are associated to specific vertex operators, to argue that the Hamiltonian of these Toda systems are, up to a trivial piece (a screening charge) just the integral of the appropriate perturbating field[3,5]. In the supercase there is only one relevant perturbation which preserves supersymmetry

and it is the superfield $\Phi_{1,3}$. Again the resulting massive theory appears to be integrable and its conservation laws can be shown to correspond to those of the quantum supersymmetric KdV equation[6]. Notice that this is the only known supersymmetric equation which can be written as a Hamiltonian system with a Poisson bracket corresponding to the classical form of the superconformal algebra. For N = 2 superconformal minimal models, three integrable perturbations[7,8] have been found and they can be put in a one-to-one correspondence with the three known integrable N = 2 supersymmetric KdV equations[7]. These results support the heuristic approach of classifying integrable perturbations of conformal field theory in terms of integrable KdV type equations which have a Hamiltonian structure whose Poisson bracket is the classical form of the corresponding conformal algebra.

Lecture 3: **KdV type equation in subcritical strings** (see the contribution of P. Mathieu and D. Sénéchal - these proceedings.)

Acknowledgments: This work was supported by NSERC (Canada) and FCAR (Québec)

References for lecture 1:

1. J.L. Gervais and A. Neveu. Nucl. Phys. B 209 (1982) 125; J.L. Gervais, Phys. Lett. B 160 (1985) 277, 279; B.A. Kupershmidt. Phys. Lett. A 109 (1985) 417

2. R.M. Miura, J. Math. Phys. 9 (1968) 1202; B.A. Kupershmidt and G. Wilson, Invent. Math. 62 (1981) 403

3. V.S. Dotsenko and V.A. Fateev, Nucl. Phys. B 1240 [F512] (1984) 312

4. P. Mathieu, Phys. Lett. B 208 (1988) 101; I. Bakas, Phys. Lett. B 213 (1988) 313; D.J. Smit, Utrech Preprint Thu 88/33

5. V.A. Fateev and S.L. Lykyanov, Inter. J. Mod. Phys. A3 (1988) 507; Additional symmetries in two-dimensional conformal field theory and exactly solvable models I, II, III, Kiev preprints

6. A.B. Zamolodchikov. Theor. Math. Phys. 65 (1986) 1205

7. S.V. Park, B.H. Cho and V.S.Myung, J. Phys. A 21 (1988) 1167; T.G. Khovanova. Theor. Math. Phys. 72 (1988) 899

8. V.G. Drinfeld and V.V. Sokolov, J. Sov. Math. 30 (1985) 1975

9. B.A. Kupersmidt, Phys. Lett. A 102 (1984) 213

10. M. Chainchian and P. Kulish, Phys. Lett. B. 183 (1987) 169

11. M. Gurses and O. Oguz, Phys. Lett. A 108 (1985) 437

12. Yu A. Manin and A.O. Radul, Comm. Math. Phys. 98 (1985) 65

13. P. Mathieu, J. Math. Phys. 29 (1988) 2499; Phys. Lett. B 203 (1988) 287; I. Yamanaka and R. Sasaki, Prog. Theor. Phys. 79 (1988) 1167

14. T. Inami and H. Kanno, Kyoto preprint RIFP-855

15. W. Oevel and Z. Popowicz, Padeborn-Wroclaw preprint 1990

16. C.A. Laberge and P. Mathieu, Phys. Lett. B 215 (1988) 718

17. P. Labelle and P. Mathieu, J. Math. Phys., to appear

18. T. Inami and H. Kanno, Kyoto preprint YITP/K-895

19. K.F. Kobayashi and T. Uematsu, KUCP-26/90

20. V.G. Knizhnik, Theor. Math. Phys. 66 (1986) 68; M. Bershadsky, Phys. Lett. B. 174 (1986) 295)

21. P. Mathieu and M. Thibeault, Lett. Math. Phys. 18 (1989) 9

22. P. Mathieu; Phys. Lett. B 218 (1989) 185

References for lecture 2:

1. A.A. Belavin, A.M. Polyakov and A.B. Zamolodchikov, Nucl. Phys. B 241 (1984) 333

2. A.B. Zamolodchikov, JETP Letters 46 (1987) 160; Int. Mod. Phys. A3 (1988) 4235; A4 (1988) 743; Adv. Study Pure Math. to appear

3. B.A. Kupershmidt and P. Mathieu, Phys. Lett. B 227 (1989) 245

4. R. Sasaki and T. Yamanaka, Adv. Stud. Pur. Math. 16 (1988) 271

5. T. Eguchi and S.K. Yang, Phys. lett. B 224 (1989) 373

6. P. Mathieu, Nucl. Phys. B 336 (1990) 338

7. P. Mathieu and M. Walton, Phys. lett. B 254 (1991) 106

8. P. Fendley, S.D. Mathur, C. Vafa and N. Warner, Phys. Lett. B243 (1990) 257
P. Fendley, W. Lerche, S. Mathur and N. Warner, Nucl. Phys. B348 (1991) 66.

KdV TYPE EQUATIONS IN SUBCRITICAL STRINGS

Pierre MATHIEU and David SÉNÉCHAL

Département de Physique
Université Laval, Québec, Canada, G1K 7P4

ABSTRACT

We review matrix models techniques applied to two-dimensional quantum gravity. We show how differential equations related to the KdV and mKdV hierarchies emerge of matrix models to describe the specific heat in the double-scaling limit, and we relate the order of these equations to the behavior of the eigenvalue density near its zeros. Finally we illustrate the integrability of matrix models outside the double scaling limit.

1. INTRODUCTION

In this work we review the matrix model approach to two-dimensional quantum gravity coupled to conformal matter. [1] Such a system may describe string theory embedded in $d-$dimensional space-time, where d is fixed by the matter content (i.e. the central charge) of the conformal model. For the solvable models discussed here, the dimension turns out to be fractional, and we do not even reach the treshold for field theory: $d = 1$. Hence we essentially deal with a finite number of degrees of freedom, and this accounts for the solvability of the models. The dimensionality 2 referred to above is the dimension of the world-sheet swept out by the moving string.

From the string theory point of view, these models are really toy models, whose study can improve our understanding of nonperturbative effects in more realistic models. However, from the point of view of statistical mechanics, they describe spin systems on random lattices, where lattice fluctuations mimick the coupling to gravity.

Within the context of statistical mechanics, our interest may be confined to models defined on surfaces having the topology of the sphere, or of the torus. All

[1] For recent reviews of matrix models, see [1-4].

calculations up to the last year were indeed restricted to these topologies, which are also the only ones for which the continuum theory has obtained precise results. However, in string theory one has to sum over surfaces of all topologies. The important breakthrough which occured recently was to realize that in a special limit (the double scaling limit) the specific heat of the matrix model, which is related to the partition function Z of two-dimensional gravity, satisfies a differential equation (the string equation) that incorporates the contribution of all topologies to Z.

It turns out that all the string equations found so far are integrable systems. Here we review in detail the derivation of these string equations for a variety of Hermitian one-matrix models, in which equations of the KdV and mKdV types appear.

2. DYNAMICAL TRIANGULATION

Quantizing two-dimensional gravity amounts to integrating over all possible metrics of two-dimensional surfaces. A simple and powerful method to accomplish this is dynamical triangulation [5-8]. It consists in replacing the surface by a set of tiles (triangles) where all the links have the same length, and in summing over all such triangulations. The first step amounts to fixing a metric, and the second to integrating over all metrics. One expects this discretization to reproduce the correct results in the continuum limit, i.e., the limit in which the area of the tiles goes to zero. This method is powerful because of its solvability, and it is solvable because of a one-to-one correspondence between the duals of the triangulations and the Feynman diagrams of some zero-dimensional field theory.

The dual of a triangulation is obtained by placing a vertex in the center of each tile, and by joining all such vertices by links, as ilustrated on fig.1. The dual triangulation of fig.1 may be regarded as a Feynman diagram in a ϕ^3 field theory. If, instead of triangles, we use squares (rhombi) to discretize the surface, then the corresponding field potential will be ϕ^4. Furthermore, a mixture of triangles and squares would correspond to a field potential of the type $\alpha\phi^3 + \beta\phi^4$. This generalizes in an obvious way. The sum over surfaces is then equal to the sum over connected diagrams, i.e., the free energy $\mathcal{F}$ of the zero-dimensional field theory.

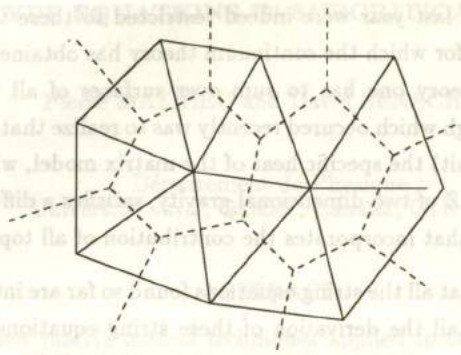


Fig.1 Dual triangulation

However, this simple approach does not allow us to control the contribution of each topology to the sum. Recall that the topology of closed two-dimensional surfaces is uniquely characterized by the number h of its handles, called *genus*. If L, F and V are respectively the numbers of links, faces and vertices of the dual triangulation, the genus of the corresponding surface is given by Euler's formula:

$$V - L + F = 2(1-h) \tag{2.1}$$

In the continuum approach to quantum gravity, the functional integration over metrics is to be done separately for each genus, followed by a discrete sum over genera:

$$\mathcal{F} = \sum_{h=0}^{\infty} \kappa^{2(h-1)} \mathcal{F}_h \tag{2.2}$$

$\mathcal{F}_h$ is the sum over surfaces of genus h, and κ is the string coupling constant, which weighs the various topologies differently. It is mainly the dependence of $\mathcal{F}$ on κ that is interesting. Furthermore, the sum over surfaces is well-defined only if each surface is weighed by a factor $e^{-\mu A}$, where A is the area of the surface, and μ is a parameter called the *cosmological constant*, like in four-dimensional gravity. In two-dimensional gravity, μ plays the role of a regulator, and it renormalizes the string coupling constant κ.

In order to gain control over the topology of triangulations, we replace the scalar field ϕ by an $N \times N$ Hermitian matrix Φ [9,10]. The propagator is then represented by two lines oriented oppositely, each one carrying an index running from 1 to N. Each face of the diagram then contains a closed loop with its index summed, contributing a factor $\delta_{ii} = N$ to the diagram.

Consider a polynomial matrix field theory, with potential $NU(\phi)$, where

$$U(\Phi) = g_2\Phi^2 + g_3\Phi^3 + \dots \tag{2.3}$$

In zero dimension there is no kinetic term, and the action is equal to the potential. Let us consider the contribution to $\mathcal{F}$ of a Feynman diagram with L propagators, F internal loops and V_n vertices of the type Φ^n (vertices where n propagators join) [11]. Each propagator contributes a factor $1/(g_2 N)$, each loop contributes N, and each vertex of type n contributes $-g_n N$, so that the diagram is equal to

$$(g_2 N)^{-L} N^F \prod_n (g_n N)^{V_n} \tag{2.4}$$

We can eliminate L and F by using the Euler relation (2.1), and the following self-evident relation: $2L = \sum_n n V_n$. The resulting expression for the diagram is

$$N^{2(1-h)} \prod_n (-g_n g_2^{-n/2})^{V_n} \tag{2.5}$$

Now, let us associate with each vertex (i.e. with each polygonal face of the original triangulation) an area $\frac{1}{2}(n-2)\delta^2$, where $n-2$ is the number of elementary triangles into which an n–sided polygon can be divided. The total area of the triangulated surface is then

$$A = \bar{A}\delta^2 \quad , \quad \bar{A} \equiv \tfrac{1}{2}\sum_n (n-2)V_n \tag{2.6}$$

The continuum limit of triangulations is obtained by sending $\delta \to 0$, and considering larger and larger diagrams, while keeping fixed the area A. In terms of $\bar{A}$, a diagram's contribution to $\mathcal{F}$ is $N^{2(1-h)}F_G(\bar{A})$, where F_G is independent of N. If we sum all the terms $F_G(\bar{A})$ for the diagrams with a given value of $\bar{A}$ and a given genus, we will generically obtain an exponential dependence, along with a power dependence:

$$F_h(\bar{A}) = \left(\sum_G F_G(\bar{A})\right)\Big|_{h,\bar{A}} \propto e^{-\bar{\mu}\bar{A}} \bar{A}^{\gamma_h - 3} \quad , \tag{2.7}$$

where we have introduced the constants $\bar{\mu}$ and γ_h. These constants are determined solely by the couplings g_i. Since $\bar{A}$ is porportional to the area A, $\bar{\mu}$ is related to the cosmological constant: $\mu = \bar{\mu}/\delta^2$. If all the couplings g_i are multiplied by an overall factor α^{-1}, then it follows from (2.5) that α only appears exponentiated by $\bar{A}$, which means that $\bar{\mu}$ depends only on an overall factor of the couplings. It is useful to normalize the g_i in such a way that the exponential dependence on $\bar{A}$

is eliminated, and to use an additional parameter α^{-1} multiplying the couplings, such that $\bar{\mu} = -\log \alpha$.

Let us now sum over all genera and all areas. After replacing the sum $\sum_{\bar{A}}$ by $\delta^{-2} \int \mathrm{d}A$, we obtain

$$\mathcal{F} = \sum_h N^{2(1-h)} \bar{\mu}^{2-\gamma_h} \mathcal{F}_h \quad , \tag{2.8}$$

where $\mathcal{F}_h$ is independent of δ. Let us first consider the contribution to $\mathcal{F}$ of surfaces having the topology of the sphere ($h = 0$). For a given value of μ, one sees that $\bar{\mu} \to 0$ in the continuum limit ($\delta \to 0$). A non-zero contribution then requires N to be infinite, with the product $N^2 \bar{\mu}^{2-\gamma_0}$ being kept finite (we assumed that $2 - \gamma_0 > 0$, which will be justified a posteriori). With $\bar{\mu} \sim N^{-2/(2-\gamma_0)}$, it is clear that, in order to get a finite contribution for each genus h, the exponent γ_h must be related to γ_0 by

$$\gamma_h - 2 = (\gamma_0 - 2)(1 - h) \tag{2.9}$$

Therefore one can write the genus expansion of the free energy as

$$\mathcal{F} = \sum_{h=0}^{\infty} \kappa^{2(h-1)} \mathcal{F}_h \tag{2.10}$$

where we defined the string coupling constant

$$\kappa \equiv N^2 \bar{\mu}^{2-\gamma_0} = N^2 \delta^{2(2-\gamma_0)} \mu^{2-\gamma_0} \tag{2.11}$$

The continuum limit (also called *double scaling limit* [12-14]) is obtained by sending $\delta^2 \to 0$ and $N \to \infty$, while keeping the combination $N^2 \delta^{2(2-\gamma_0)}$ equal to a constant, e.g. equal to 1. There is then a power relation between the string coupling constant κ and the cosmological constant μ. Having shown perturbatively that the partition function of two-dimensional gravity is equal to the free energy $\mathcal{F}$ of the matrix model in some large-N limit, we will now proceed to calculate the latter by direct integration.

3. MATRIX MODELS AND ORTHOGONAL POLYNOMIALS

The matrix model partition function is

$$Z = e^{-\mathcal{F}} = \int \prod_i \mathrm{d}\Phi_{ii} \prod_{i<j} \mathrm{dRe}\Phi_{ij}\ \mathrm{dIm}\Phi_{ij}\ \exp -N \mathrm{tr} U(\Phi) \tag{3.1}$$

The action is invariant with respect to unitary transformations of U: $U \mapsto M^\dagger U M$. Thus one can trade the variable Φ for the variables M and ϕ_i, where the ϕ_i denote the N eigenvalues of Φ. The Jacobian of this transformation is the square of the Vandermonde determinant,

$$\Delta(\phi) = \prod_{i<j} (\phi_i - \phi_j) \tag{3.2}$$

The integration with respect to M yields a factor which depends only on N, and which we drop since it only contributes an additive constant to $\mathcal{F}$, without altering its critical behavior. Therefore we have effectively reduced the number of degrees of freedom from N^2 to N:

$$Z = \int \left(\prod_i \mathrm{d}\phi_i \right) \Delta^2(\phi) \exp -N \sum_i U(\phi_i) \tag{3.3}$$

In order to deal with the Vandermonde determinant, we introduce a set of monic polynomials (i.e. whose highest power has unit coefficient) $P_i(\phi)$, orthogonal with respect to the measure [15]

$$\mathrm{d}\mu(\phi) \equiv \mathrm{d}\phi\ \exp -NU(\phi) \tag{3.4}$$

and with norm $\sqrt{h_i}$:

$$\langle P_i | P_j \rangle \equiv \int \mathrm{d}\mu(\phi) P_i(\phi) P_j(\phi) = h_i \delta_{ij} \tag{3.5}$$

From now on we restrict the discussion to the case of even potentials,

$$U(\phi) = \alpha^{-1} \sum_{i=1}^{d} g_{2i} \phi^{2i} \quad , \tag{3.6}$$

for which the polynomials P_i have a definite parity, and satisfy a three way recursion relation:

$$\phi P_i(\phi) = P_{i+1}(\phi) + R_i P_{i-1}(\phi) \tag{3.7}$$

To demonstrate (3.7), notice that $\phi P_i - P_{i+1}$ is necessarily a linear combination of polynomials of degree smaller than $i + 1$:

$$\phi P_i = P_{i+1} + \sum_{k \le i} \lambda_k P_k \tag{3.8}$$

By virtue of the orthogonality, we have

$$\langle P_l|\phi|P_i\rangle = \lambda_l \tag{3.9}$$

But since P_i is orthogonal to any polynomial of degree $< i$, and since ϕP_l is of degree $l+1$, we see that $\lambda_l = 0$ for $l < i-1$. We denote λ_{i-1} by R_i, and λ_i vanishes because of the definite parity of the polynomials. Notice also that

$$\begin{aligned}\langle P_i|\phi|P_{i-1}\rangle &= \langle P_i|P_i\rangle = h_i \\ &= \langle \phi P_i|P_{i-1}\rangle = R_i h_{i-1} \quad \Rightarrow R_i = h_i/h_{i-1}\end{aligned} \tag{3.10}$$

The Vandermonde determinant is nothing but

$$\Delta(\phi) = \det|\phi_i^{j-1}| = \det|P_{j-1}(\phi_i)| \tag{3.11}$$

Indeed, the latter form is obtained from the former by adding to the j^{th} row a linear combination of the preceding rows, which does not affect the determinant. The matrix model partition function thus becomes

$$Z \propto \int \left(\prod_i \mathrm{d}\mu(\phi_i)\right) \det|P_{j-1}(\phi_i)|^2 \tag{3.12}$$

Δ contains $N!$ terms; if one expands Δ^2, then the cross-terms do not contribute because of the orthogonality of the polynomials, and one obtains

$$Z \propto N! \prod_{i=0}^{N-1} h_i = N! h_0^N \prod_{i=0}^{N-1} R_i^{N-i} \tag{3.13}$$

For the free energy, this translates into

$$\mathcal{F} = const. \; - \sum_{i=0}^{N-1} (N-i) \log R_i \tag{3.14}$$

Therefore, up to an additive constant, $\mathcal{F}$ is determined by the recursion coefficients R_i.

We now assert that the coefficients R_i can be in principle determined recursively from the following relation:

$$ih_i = NR_i\langle P_i|U'|P_{i-1}\rangle \tag{3.15}$$

To demonstrate this, notice first that $\phi P_i' = iP_i + \lambda P_{i-2} + \ldots$, which allows us to write $ih_i = \langle P_i|\phi|P_i'\rangle$. Applying (3.7) to P_i, one has

$$\begin{aligned} ih_i &= \langle P_i'|P_{i+1}\rangle + R_i\langle P_i'|P_{i-1}\rangle \\ &= R_i\langle P_i'|P_{i-1}\rangle \\ &= NR_i\langle P_i|U'(\phi)|P_{i-1}\rangle - R_i\langle P_i|P_{i-1}'\rangle \end{aligned} \tag{3.16}$$

In the last step we have integrated by parts. The last inner product vanishes since the degree of P_{i-1}' is smaller than that of P_i, and we recover (3.15).

Since $U'(\phi)$ is a polynomial in ϕ, (3.15) can be evaluated by using (3.7) repeatedly. For instance, with $U(\phi) = g_2\phi^2 + g_4\phi^4$, (3.15) becomes

$$\frac{i\alpha}{N} = R_i(2g_2 + 4g_4(R_{i-1} + R_i + R_{i+1})) \tag{3.17}$$

If we add to $U(\phi)$ a sixth degree term $g_6\phi^6$, the following expression must be added to the r.h.s. of (3.17):

$$\begin{aligned} &6g_6R_i(R_{i-1}R_{i-2} + R_{i-1}^2 + 2R_iR_{i-1} + R_{i-1}R_{i+1} + \\ &R_i^2 + 2R_iR_{i+1} + R_{i+1}^2 + R_{i+1}R_{i+2}) \end{aligned} \tag{3.18}$$

4. EIGENVALUE DISTRIBUTION

The set of matrices Φ associated with the integration measure defined in (3.1), can be characterized by a statistical distribution of eigenvalues $\rho(\phi)$ [16]. This is natural, since the action depends only on these eigenvalues. To clarify this idea, we define the correlators

$$B_m \equiv N^{-1}\langle \mathrm{tr}\Phi^m\rangle \quad , \quad (m > 0) \tag{4.1}$$

We then adopt the following natural definition for a distribution of eigenvalues:

$$B_m \equiv \int \mathrm{d}\phi \; \rho(\phi)\phi^m \tag{4.2}$$

We will assume that $\rho(\phi)$ is normalized and, at least for $N \to \infty$, defined on a finite interval $[-c, c]$:

$$\int_{-c}^{c} \mathrm{d}\phi \; \rho(\phi) = 1 \tag{4.3}$$

If $U(\phi)$ is symmetric, then so is $\rho(\phi)$. The correlators B_m are the moments of ρ, and their knowledge suffices to reconstruct ρ. We also introduce the generating function F:

$$F(z) \equiv \frac{1}{N}\langle \mathrm{tr}\frac{1}{z-\Phi}\rangle = \frac{1}{z}\sum_{m=0}^{\infty}\frac{B_m}{z^m} \tag{4.4}$$

One sees that $F(z)$ is related to ρ as follows:

$$F(z) = \int \mathrm{d}\phi\, \frac{\rho(\phi)}{z-\phi} \quad \Rightarrow \quad \rho(\phi) = \frac{1}{\pi}\lim_{\epsilon\to 0^+} \mathrm{Im}F(\phi - i\epsilon) \tag{4.5}$$

4.1. The Schwinger-Dyson Equation.

The generating function can be determined in the large-N limit by use of the Schwinger-Dyson equation:

$$F(z)^2 = \frac{1}{N}\mathrm{tr}\langle U'(\Phi)\frac{1}{z-\Phi}\rangle \tag{4.6}$$

In order to demonstrate (4.6), we perform an infinitesimal change of variable in (3.1): $\Phi \mapsto \Phi + \epsilon f(\Phi)$, where f is for the time being an arbitrary function. Under such an infinitesimal transformation, the exponential of the action changes by $-N\mathrm{tr}(U'(\Phi)\epsilon f(\Phi))$, while the integration measure changes as follows:

$$\mathrm{d}\Phi \mapsto \mathrm{d}\Phi \left\{1 + \epsilon \oint \frac{\mathrm{d}z}{2\pi i} f(z)\left(\mathrm{tr}\frac{1}{z-\Phi}\right)^2\right\} \tag{4.7}$$

This last formula may appear nebulous, and so let us give more details. In general, under a transformation $\Phi \mapsto g(\Phi)$, the measure $\mathrm{d}\Phi$ changes as:

$$\mathrm{d}\Phi' = \mathrm{d}\Phi\ \det\left(\frac{\partial g_{ij}}{\partial \Phi_{rs}}\right) \tag{4.8}$$

Here the determinant is $N^2 \times N^2$. But, for ϵ infinitesimal, $\det(1+\epsilon A) = 1 + \epsilon \mathrm{tr}A$, and this allows us to write

$$\mathrm{d}\Phi \mapsto \mathrm{d}\Phi\left(1 + \epsilon\sum_{ij}\frac{\partial f_{ij}}{\partial \Phi_{ij}}\right) \tag{4.9}$$

Suppose now that $f(\Phi) = \Phi^m$. Then one easily calculates that

$$\sum_{ij}\frac{\partial f_{ij}}{\partial \Phi_{ij}} = \sum_{l=0}^{m-1}\mathrm{tr}\Phi^l\mathrm{tr}\Phi^{m-l-1} \tag{4.10}$$

But, on the other hand,

$$\left(\mathrm{tr}\frac{1}{z-\Phi}\right)^2 = \sum_{r,s=0}^{\infty}\frac{1}{z^{r+s+2}}\mathrm{tr}\Phi^r\mathrm{tr}\Phi^s \tag{4.11}$$

Therefore, by the residue theorem, the relation (4.7) is demonstrated for the case $f(\Phi) = \Phi^m$. For any function f that admits a power series in Φ, the demonstration is exactly the same, term by term.

Since a change of variable of integration should not affect the partition function, the change in the integration measure should exactly compensate the change in the exponential of the action. For the specific choice $f(\Phi) = 1/(z-\Phi)$, this yields

$$\frac{1}{N^2}\langle\left(\mathrm{tr}\frac{1}{z-\Phi}\right)^2\rangle = \frac{1}{N}\mathrm{tr}\langle U'(\Phi)\frac{1}{z-\Phi}\rangle \tag{4.12}$$

In the $N \to \infty$ limit, the average of the square of $\mathrm{tr}f(\Phi)$ is equal to the square of the average of $\mathrm{tr}f(\Phi)$, i.e. $\mathrm{tr}f(\Phi)$ does not fluctuate. Indeed, even if each eigenvalue ϕ_i fluctuates, these fluctuations tend to cancel out when the trace is taken, and they show up only at order $1/N$. This allows us, after using (4.4), to obtain the large-N Schwinger-Dyson equation (4.6).

4.2. Relation Between $U(\phi)$ and $\rho(\phi)$.

In (4.6), let us expand $U'(\Phi)$ in a Taylor series about $\Phi = z$. This series has a finite number of terms since U is a polynomial of degree $2d$. The Schwinger-Dyson equation then takes the following form:

$$F(z)^2 - F(z)U'(z) + \eta(z) = 0 \quad , \tag{4.13}$$

where $\eta(z)$ is a polynomial of degree $2d-2$, which can be expressed in terms of the lowest correlators B_m. The quadratic equation (4.13) is readily solved:

$$F(z) = \tfrac{1}{2}U'(z) \pm \sqrt{\tfrac{1}{4}U'(z)^2 - \eta(z)} \tag{4.14}$$

Let us now assume that $\rho(z)$ does not vanish on $[-c,c]$, except at isolated points. Then, (4.5) shows that $F(z)$ will be analytic in z everywhere except on the cut $[-c,c]$. Therefore, the polynomial $\eta(z)$ must be such that

$$F(z) = \tfrac{1}{2}U'(z) \pm G(z)\sqrt{z^2-c^2} \tag{4.15}$$

where $G(z)$ is a polynomial of degree $2d-2$, with $G^2(z^2-c^2) = \frac{1}{4}U'^2 - \eta$. Let us denote by P_+ the linear operator that keeps only the positive powers of z in an asymptotic expansion in $1/z$. By virtue of (4.5), $F(z) \sim 1/z$ when $z \to \infty$, and therefore $P_+(F(z)) = 0$. Since P_+ is the identity operator when applied to a polynomial in z, we have the relations

$$\begin{aligned} G(z) &= P_+\left(\frac{\frac{1}{2}U'(z)}{\sqrt{z^2-c^2}}\right) \\ U'(z) &= 2P_+\left(G(z)\sqrt{z^2-c^2}\right) \end{aligned} \tag{4.16}$$

These formulae allow us to relate the potential U to the eigenvalue distribution ρ, since, by virtue of (4.5),

$$\rho(\phi) = \frac{1}{\pi}G(\phi)\sqrt{c^2-z^2} \quad . \tag{4.17}$$

The relation (4.16) can be made more explicit with the help of the following asymptotic expansions:

$$\sqrt{z^2-c^2} = z\sum_{m=0}^{\infty} a_m (c/z)^{2m} \quad , \quad a_m = -\frac{|2m-3|!!}{m!2^m} \quad , \quad a_0 = 1 \tag{4.18}$$

$$\frac{1}{\sqrt{z^2-c^2}} = \frac{1}{z}\sum_{m=0}^{\infty} b_m (c/z)^{2m} \quad , \quad b_m = \frac{|2m-1|!!}{m!2^m} \tag{4.19}$$

The eigenvalue distribution corresponding to a generic potential $U(z)$ is then

$$\rho(\phi) = \frac{1}{\pi}\sqrt{c^2-\phi^2}\sum_{k=1}^{d}\sum_{m=0}^{k-1} k g_{2k} b_m c^{2m} \phi^{2(k-m-1)} \tag{4.20}$$

On the other hand, given a polynomial $G(z) = \sum_i G_{2i} z^{2i}$, the corresponding potential $U(z)$ is

$$U(z) = \sum_{i=1}^{d-1}\sum_{m=0}^{i} \frac{a_m c^{2m} G_{2i}}{i-m+1} z^{2(i-m+1)} \tag{4.21}$$

The above formulae provide an easy way to go from the potential to the eigenvalue distribution and vice-versa. However, they only make sense if the polynomial $G(z)$ is non-negative in the interval $[-c,c]$, otherwise the distribution ρ would be negative in some range. If, starting from some potential $U(z)$, the distribution ρ obtained from (4.20) is negative somewhere in $[-c,c]$, this simply means that the

formula (4.16) is not valid because $G(z)$ is not a polynomial, which happens if the distribution ρ has support on several disjoint cuts between $-c$ and c.

The constant c should be adjusted such that ρ is normalized in $[-c,c]$. To do so, one may use the following integral:

$$\int_{-c}^{c} \mathrm{d}z\ (c^2-z^2)^{k-\frac{1}{2}} z^{2q} = \pi(c/2)^{2q+2k}\frac{(2k)!(2q)!}{k!q!(q+k)!} \tag{4.22}$$

4.3. Relation Between $\rho(\phi)$ And Orthogonal Polynomials.

To establish the connection between the method of orthogonal polynomials and the eigenvalue distribution ρ, let us express the correlators B_m in terms of the recursion coefficients R_i. The correlators are defined by the following matrix integral:

$$\begin{aligned} B_m &= \frac{1}{NZ}\int \mathrm{d}\Phi\ \mathrm{tr}\Phi^m \exp -NU(\Phi) \\ &= \frac{1}{NZ}\int \prod_i \mathrm{d}\phi_i\ \Delta^2(\phi)\left(\sum_i \phi_i^m\right)\exp -N\sum_i U(\phi_i) \end{aligned} \tag{4.23}$$

After expanding Δ^2 in orthogonal polynomials, one obtains

$$B_m = \sum_{i=0}^{N-1}\frac{1}{h_i}\langle P_i|\phi^m|P_i\rangle \tag{4.24}$$

In order to calculate $\langle P_i|\phi^m|P_i\rangle$, one uses (3.7) repeatedly. The final expression is of course horrendously complicated, but can be simplified in the large-N limit.

Let us make the assumption that the recursion coefficients R_i vary slowly with i, and can be represented by a continuous fonction $R(x)$, with $x \equiv i/N$. This asumption is not always justified, as we will see below, but is valid in a large number of cases. It allows us to approximate $\langle P_i|\phi^m|P_i\rangle$ by a contour integral:

$$\frac{1}{h_i}\langle P_i|\phi^m|P_i\rangle \to \oint \frac{\mathrm{d}u}{u}\left(u+\frac{R(x)}{u}\right)^m \tag{4.25}$$

Indeed, the binomial expansion of the integrand mimicks the application of (3.7) on the r.h.s., and the residue integral selects the terms that would survive the orthogonality relation. The series (4.4) can then be formally summed like a geometric series, and one obtains

$$F(z) = \int_0^1 \mathrm{d}x \oint \frac{\mathrm{d}u}{u}\frac{1}{zu-u^2-R(x)} \tag{4.26}$$

If z is large enough, i.e., if the asymptotic expansion (4.4) is valid, the residue integral does not vanish, and

$$F(z) = \int_0^1 \mathrm{d}x \, \frac{1}{\sqrt{z^2 - 4R(x)}} \tag{4.27}$$

From this one can extract by analytic continuation an expression for $\rho(z)$, using (4.5):

$$\rho(z) = \frac{1}{\pi} \int_{D_z} \mathrm{d}x \, \frac{1}{\sqrt{4R(x) - z^2}} \tag{4.28}$$

The integration is performed over the range D_z of x on which the argument of the square-root is positive. This formula will allow us to relate the critical behavior of ρ to the critical behavior of the function $R(x)$.

5. KdV-TYPE STRING EQUATIONS

In this section we will show that the specific heat exponent γ_0 is restriced to the values $-1/k$, where k is a positive integer. This exponent characterizes the way the specific heat $\mathcal{F}'' \equiv \partial^2 \mathcal{F}/\partial \mu^2$ depends upon the cosmological constant μ when μ is large: $\mathcal{F}'' \sim \mu^{1/k}$. Under the hypothesis that the recursion coefficients R_i tend towards a single continuous fonction $R(x)$ when $N \to \infty$, we will see that the specific heat satisfies a non-linear differential equation in μ, belonging to the KdV hierarchy.

5.1. Critical behavior of $R(x)$.

When we replace the coefficients R_i by the continuous fonction $R(x)$, the expression (3.14) for the free energy becomes

$$\mathcal{F} = const. \; - \; N^2 \int_0^1 (1-x) \log R(x) \tag{5.1}$$

The dependence of $\mathcal{F}$ on μ is implicit in the function $R(x)$, and the exponent γ_0 is fixed by the way in which $R(x)$ approaches the critical value 1 as $x \to 1$. To extract this critical behavior, it is useful to introduce the following variables:

$$\alpha x \equiv 1 - a^{2k} t \quad , \quad R(x) \equiv 1 - a^2 f(t) \quad , \quad a \equiv N^{\gamma_0/(2-\gamma_0)} \tag{5.2}$$

Here we defined $k \equiv -1/\gamma_0$, which will shortly be proven to be a positive integer. Recall that we defined the continuum limit by fixing $\delta^2 = N^{-2/(2-\gamma_0)}$, and

therefore the value of the new variable t corresponding to $x = 1$ is nothing but the cosmological constant. When x is varied by $1/N$, t changes by a. Furthermore, in terms of t and f, the formula (5.1) for $\mathcal{F}$ becomes

$$\mathcal{F} = const. - \int_\mu^\infty \mathrm{d}t (t - \mu) f(t) \tag{5.3}$$

The specific heat $\mathcal{F}''$ is obtained by differentiating twice with respect to μ: $\mathcal{F}'' = f(\mu)$. Therefore, knowledge of the scaling function $f(t)$ tells us how the specific heat varies with the cosmological constant.

Now, let us go back to (3.17), and substitute the function $R(x)$ for R_i. At lowest order in $1/N$, we may ignore the difference between R_i and R_{i+1}, like we did in (4.25). The recursion relation (3.17) then becomes an algebraic equation for $R(x)$: $\alpha x = 2g_2 R + 12 g_4 R^2$. For a potential of degree $2d$, this equation is

$$\alpha x = \oint \mathrm{d}u U'(u + R/u) = \sum_{j=1}^{d} \frac{(2j)!}{j!(j-1)!} g_{2j} R^j \tag{5.4}$$

If we substitute $1 - a^2 f$ for R and expand, we realize that only integer powers of a^2 will occur, since the r.h.s. is a polynomial. Therefore, since the l.h.s. is nothing but $1 - a^{2k} t$, the value of k is restricted to positive integers. Furthermore, the condition that $R(1) = 1$ is equivalent to the following normalization constraint on the couplings:

$$\sum_{j=1}^{d} \frac{(2j)!}{j!(j-1)!} g_{2j} = 1 \tag{5.5}$$

This constraint was mentioned in sect.2, and is now made explicit.

For a given value of k, Eq.(5.4) is consistent only if $R(x) \sim 1 - (1-x)^{1/k}$ when $x \sim 1$, i.e. the critical behavior of $\mathcal{F}''$ as a function of μ is fixed by the behavior of $R(x)$ as x approaches 1: $f(\mu) \sim \mu^{1/k}$ [17,12-14].

5.2. Critical behavior of $\rho(x)$ and critical potentials.

The way the fonction $R(x)$ approaches 1 as x approaches 1 affects the way the distribution of eigenvalues $\rho(z)$ approaches 0 near the boundary c of its domain [18]. First, let us point out that the condition (5.5) on the couplings g_{2i} fixes the value of c to 2. This is easily verified with the help of (4.22) and (4.20). Next, let

us use (4.28) with $z = 2 - \delta$ and $R(x) = 1 - (1-x)^{1/k}$. Changing the integration variable to $\tau = 1 - x$, one has

$$\rho(2-\delta) = \frac{1}{\pi}\int_0^{\delta^k} \frac{\mathrm{d}\tau}{2\sqrt{\delta - \tau^{1/k}}} \tag{5.6}$$

This integral obviously scales like $\delta^{k-1/2}$.

Therefore the polynomial $G(z)$ has a zero of order $k-1$ at $z = 2$. The simplest symmetric polynomial for which this happens is $G_k(z) = (4-z^2)^{k-1}$, of degree $2k-2$. The corresponding potential is of degree $2k$, and can be obtained from (4.21):

$$U_k(\phi) = \sum_{j=1}^{k} (-1)^{j-1}\phi^{2j}\frac{k!(j-1)!}{(k-j)!(2j)!} \tag{5.7}$$

The simplest critical potentials are

$$U_1(\phi) = \tfrac{1}{2}\phi^2 \quad , \quad U_2(\phi) = \phi^2 - \tfrac{1}{12}\phi^4 \quad , \quad U_3(\phi) = \tfrac{3}{2}\phi^2 - \tfrac{1}{4}\phi^4 + \tfrac{1}{60}\phi^6 \tag{5.8}$$

Of course these critical potentials can also be obtained from (5.4), by assuming that the r.h.s. is equal to $1 - (1-R)^k$.

5.3. The string equations.

The function $R(x)$ obtained by solving Eq.(5.4) gives essentially the correct large t behavior of the scaling function $f(t)$. On the other hand, when t is small, i.e. when x is even closer to 1, the derivatives of $R(x)$ are less and less negligible, and Eq.(5.4) looses its validity. One needs to go back to (3.17), and to substitute directly $1 - a^2 f(t - aj)$ for R_{i+j}. If we Taylor expand the result around t, and keep the lowest non-trivial terms, we obtain a differential equation for the scaling function $f(t)$.

Consider for instance the potential αU_2 which corresponds to $k = 2$. Eq. (3.17) yields

$$1 - a^4 t = 1 - a^4(f^2 - \tfrac{1}{3}f'') + O(a^6) \tag{5.9}$$

that is, in the $a \to 0$ limit,

$$t = f^2 - \tfrac{1}{3}f'' \tag{5.10}$$

This is the Painlevé I equation.

In general, for the critical potential αU_k, this differential equation is $t = \mathcal{R}_k[f]$, where

$$\begin{aligned} \mathcal{R}_1[f] &= f \\ \mathcal{R}_2[f] &= f^2 - \tfrac{1}{3}f'' \\ \mathcal{R}_3[f] &= f^3 - ff'' - \tfrac{1}{2}(f')^2 + \tfrac{1}{10}f^{(4)} \\ &\text{etc.} \end{aligned} \tag{5.11}$$

These equations for the heat capacity $\mathcal{F}'' = f$ as a function of the cosmological constant are called *string equations* [12-14]. They belong to the KdV hierarchy, and can in general be written as follows:

$$t = \mathcal{R}_k[f] = A_k\left[-\tfrac{1}{2}\partial^2 + \partial^{-1}f\partial + f\right]^k \cdot 1 \quad , \quad A_k = (-\tfrac{1}{2})^{k+1}\frac{k!}{(2k-1)!!} \tag{5.12}$$

The KdV hierarchy proper is of course $\partial_\xi f = \partial_t \mathcal{R}_k[f]$, where ξ is the usual time and t is the spatial variable. The recursion operator appearing in (5.12) is the KdV recursion operator.

6. TWO-BRANCH ANSATZ AND THE mKdV HIERARCHY

6.1. The two-branch ansatz.

We have said that the scaling behavior of the recursion coefficients R_i near $i = N$ is described by non-linear differential equations belonging to the KdV hierarchy, provided we assume that these coefficients tend towards a continuous function $R(x)$ when $N \to \infty$. This fonction $R(x)$ should satisfy the boundary conditions $R(1) = 1$ and $R(0) = 0$. However, these conditions are not always satisfied. If, for instance, the 'mass' g_2 is negative, then Eq.(5.4) implies that $\partial R/\partial x\,|_{R=0} < 0$, and the fonction $R(x)$ must cross the $x = 0$ axis at some finite value of R before reaching $x = 1$, which is unacceptable (i.e. $R(x)$ is not single-valued). In this case, a description of the recursion coefficients R_i with the help of a single function is impossible as $N \to \infty$.

On the other hand, such a description is generally possible with the help of more than one function. Let us assume that the coefficients R_i tend towards two different functions, depending on the parity of i [19-24]:

$$R_{2i} \to R^{(1)}(x) \quad , \quad R_{2i+1} \to R^{(2)}(x) \tag{6.1}$$

We may use instead the two functions $T(x)$ and $Y(x)$, defined by $R^{(1,2)} \equiv T \mp Y$. Consider a general quartic potential $U(\phi) = g_2\phi^2 + g_4\phi^4$. Substitution of the two-branch ansatz (6.1) into the recursion formula (3.7) for i even and odd leads to the following solution:

$$T = -\frac{g_2}{4g_4} \quad , \quad Y = \frac{1}{\sqrt{4g_4}}\sqrt{x_c - x} \quad , \quad x_c \equiv \frac{g_2^2}{4g_4} \tag{6.2}$$

This solution is valid provided $g_2 < 0$ and $x < x_c$. Beyond x_c the function Y vanishes and a one-branch solution takes over. If $x_c = 1$, which happens in this case if $g_4 = \frac{1}{4}$, then this two-branch behavior reaches the critical value $x = 1$, and a new description of the critical behavior is needed.

Table 1 : critical couplings

degree	p	r	g_2	g_4	g_6	g_8
2	1	0	1/2			
4	2	0	1	-1/12		
4	1	1	-1	1/4		
6	3	0	3/2	-1/4	1/60	
6	2	1	-3/2	3/4	-1/12	
6	1	2	-1/4	-1/4	1/12	
8	4	0	2	-1/2	1/15	-1/280
8	3	1	-2	3/2	-1/3	1/40
8	2	2	-2/3	-1/2	1/3	-1/24
8	1	3	-2/5	-1/10	-1/15	1/40

The necessity of using two limiting functions $R^{(1,2)}$ is related to the existence of two degenerate minima of the quartic potential when $g_2 < 0$. In fact, the number of different functions $R^{(j)}$ necessary to describe the recursion coefficients R_i as $N \to \infty$ is generically equal to the number of degenerate minima of the matrix potential $U(\phi)$. Since the eigenvalues ϕ_i tend to accumulate around the minima of $U(\phi)$, the corresponding distribution $\rho(z)$ will have maxima roughly where the potential has minima. The critical potentials for which the n different

branches survive until $x = 1$ correspond to distributions whose various maxima are separated by zeros. For the especially simple case of two branches, these critical distributions have the following form (r and p are integers such that $r \geq 0$ and $p \geq 1$) [25]:

$$\rho(z) = c_{r,p} z^{2r}(4 - z^2)^{p-1/2} \quad , \quad c_{r,p} = \frac{1}{\pi}\frac{r!p!(r+p)!}{(2p)!(2r)!} \tag{6.3}$$

The corresponding critical potentials may be easily obtained from (4.21). Table 1 gives the critical couplings obtained for various values of p and r for potentials of degree less than or equal to 8:

Let us first relate the shape of the eigenvalue distribution to the critical behavior of $T(x)$ and $Y(x)$ as $x \to 1$. With the two-branch ansatz (6.1), Eq.(4.25) is replaced by

$$B_{2m} = \int_0^1 \mathrm{d}x \oint \frac{\mathrm{d}u}{u} \left(u + \frac{R^{(1)}(x)}{u}\right)^m \left(u + \frac{R^{(2)}(x)}{u}\right)^m \tag{6.4}$$

After summing the geometric series (4.4) and performing the contour integration, one obtains

$$F(z) = \int_0^1 \mathrm{d}x \, \frac{z}{\sqrt{z^4 - 4z^2T(x) + 4Y^2(x)}} \tag{6.5}$$

Again, the density of eigenvalues is obtained by performing the same integral over the range D_z of x such that the argument of the square-root is negative:

$$\rho(z) = \frac{1}{\pi}\int_{D_z} \frac{z}{\sqrt{4z^2T(x) - z^4 - 4Y^2(x)}} \tag{6.6}$$

To obtain the critical distributions (6.3) one needs the following critical behaviors for T and Y:

$$\begin{aligned} &i) \quad p \geq r: \quad Y^2 \sim \beta(1-x)^{1/r} \quad , \quad T \sim 1 - \alpha(1-x)^{1/p} \qquad (\beta \neq -4\alpha) \\ &ii) \quad p < r: \quad Y^2 \sim \beta(1-x)^{1/r} \quad , \quad 4T + Y^2 \sim 1 - \lambda(1-x)^{1/p} \end{aligned} \tag{6.7}$$

where α, β and λ are some constants. In all cases the multicritical index k is given by $k = \max(r,p)$. With $\alpha x \equiv 1 - a^{2k}t$, this ensures that $1 - T \sim a^2 t^{1/k}$.

6.2. Scaling functions, the specific heat and string equations.

The specific heat $\mathcal{F}''$, as obtained by double differentiation of (5.1), has the following form when expressed in terms of the n branches $R^{(i)}(x)$:

$$\mathcal{F}'' = \frac{1}{na^2}\sum_i \log R^{(i)}(1) \tag{6.8}$$

In the case of two branches, it therefore appears judicious to introduce the scaling functions $u(t) \equiv a^{-1}Y(x)$ and $\frac{1}{2}v(t) \equiv a^{-2}(1 - T(x))$. In terms of these two functions, the specific heat becomes $\mathcal{F}'' = \frac{1}{2}u^2 + \frac{1}{4}v^2$. However, we see from the second of eqs.(6.7) that the scaling function $w \equiv v + \frac{1}{2}u^2$ will be more useful than v. Note that each of these scaling functions may be considered as a series in powers of a, of which only the lowest term survives in the $a \to 0$ (i.e. $N \to \infty$) limit:

$$u = \sum_{i=0}^{\infty} u_i a^i \quad , \quad w = \sum_{i=0}^{\infty} w_i a^i \tag{6.9}$$

Let us now look at the consequences of (6.7) for three different cases:

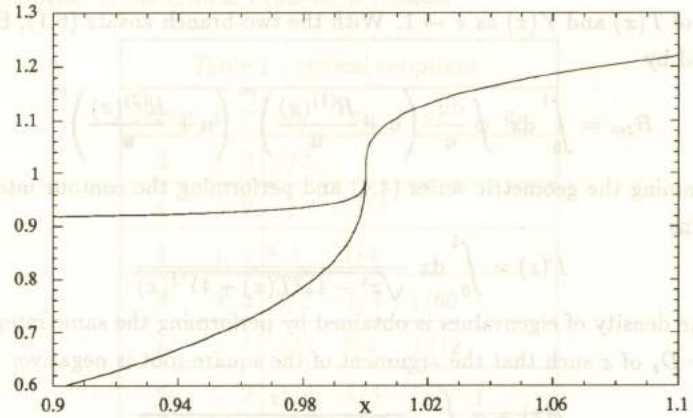


Fig.2

I) $p > r$ (hence $p = k$).

In this case the two branches merge like a cusp at $x = 1$, and Y scales like $a^{p/r}$. Fig.2 shows the two branches as a function of x for $p = 3$ and $r = 1$. Therefore u is zero at first order ($u_0 = 0$) and does not contribute to $\mathcal{F}''$. In this case, as far as the specific heat is concerned, the string equation may be derived with the single-branch ansatz (5.2), and one recovers the p^{th} KdV equation, namely eq.(5.12) for $f = \frac{1}{2}w$.

II) $p = r$.

From (6.7), one sees that Y scales like a, and $1 - T$ like a^2, i.e., neither w nor u vanishes at leading order, and both functions contribute to the specific heat:

$\mathcal{F}'' = \frac{1}{4}u^2 + \frac{1}{2}w$. The series expansions (6.9) are not necessary to unambiguously derive the correct string equations in this case: only the lowest terms matter. The result is a decoupled system of equations: w satisfies (5.12) with $k = p$, and u satisfies an equation of the mKdV type, namely

$$A'_r[-\partial^2 + u\partial^{-1}u\partial]^r \cdot u = ut \tag{6.10}$$

where the A'_r are some constants.

The simplest example of a model of this type comes from the potential

$$U(\phi) = -\phi^2 + \tfrac{1}{4}\phi^4 \quad , \tag{6.11}$$

for which $k = 1$. By substituting the ansatz

$$\begin{aligned} R_{2i} &\to R^{(1)}(x) = 1 - au(t) - \tfrac{1}{2}a^2 w(t) + \tfrac{1}{4}a^2 u^2(t) \\ R_{2i+1} &\to R^{(2)}(x) = 1 + au(t) - \tfrac{1}{2}a^2 w(t) + \tfrac{1}{4}a^2 u^2(t) \end{aligned} \tag{6.12}$$

into (3.17), for i even, and again for i odd, one finds, for the sum and differences of the two equations,

$$\begin{aligned} (S) \quad &: \quad t = 2w \\ (D) \quad &: \quad 0 = uw - \tfrac{1}{2}u^3 + u'' \end{aligned} \tag{6.13}$$

Replacing w in (D) by its expression obtained from (S), one finds

$$w = \tfrac{1}{2}t \quad , \quad ut = u^3 - 2u'' \tag{6.14}$$

The second of these equations is the Painlevé II equation.

III) $p < r$ (hence $k = r$).

In this case the two branches merge at $x = 1$ as in a bifurcation diagram. Fig.3 shows the two branches for $p = 1$ and $r = 2$. From (6.7) one sees that Y again scales like a to leading order. However, w is 0 to leading order, i.e. $T \sim 1 + \frac{1}{4}a^2 u$. To derive the string equations in this case, we substitute in the recursion relation (3.7) the series (6.9) for u and w. We thus obtain equations for each order in a, which allow us to express the w_i in terms of the u_i. After backsubstitution, one finds that the function u_0 (which becomes u when $a \to 0$) satisfies the r^{th} mKdV equation (6.10). Recall that the specific heat is $\mathcal{F}'' = \frac{1}{4}u^2$ in this case.

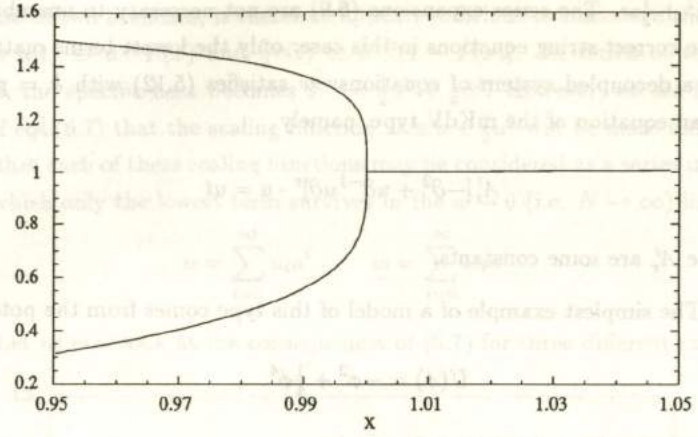


Fig.3

The simplest example of this class of models can be obtained from the following sixth degree potential (see (4.21) and (6.3)):

$$U(\phi) = -\tfrac{1}{2}\phi^2 - \tfrac{1}{4}\phi^4 + \tfrac{1}{12}\phi^6 \quad , \tag{6.15}$$

for which $k = 2$. After substitution of the critical variables t, u, and w in (3.17) and (3.18) we obtain:

$$\begin{aligned} a^4 t &= 4a^2 w + a^4(\tfrac{3}{2}u^2 w - 3w^2 - \tfrac{1}{2}(u')^2 + 2w'') + O(a^6) \qquad (S) \\ 0 &= 4a^3 uw + a^5(-\tfrac{3}{8}u^5 + \tfrac{3}{2}u^3 w - \tfrac{3}{2}uw^2 + 2u(u')^2 \\ &\quad - u'w' + \tfrac{5}{2}u^2 u'' - wu'' + 2uw'' - u^{(4)}) + O(a^7) \qquad (D) \end{aligned} \tag{6.16}$$

One then substitutes the series expansion (6.9). From (S) one finds that $w_0 = w_1 = 0$ and $w_2 = \frac{1}{4}t + \frac{1}{8}(u_0')^2$. Substitution of these results in (D) yields

$$ut = u^{(4)} - \tfrac{5}{2}u^2 u'' - \tfrac{5}{2}u(u')^2 + \tfrac{3}{8}u^5 \tag{6.17}$$

This is the second nontrivial member of the mKdV hierarchy. For a fixed value of p, the expression for w as a function of u seems universal, except that t appears only where appropriate for the corresponding value of k. For instance, analysis of the 8^{th} degree potential with $p = 1$ and $r = 3$ yields

$$\begin{aligned} w_0 = w_1 = 0 \quad &, \quad w_2 = \tfrac{1}{8}(u_0')^2 \\ w_3 = \tfrac{1}{4}u_0'u_1' \quad &, \quad w_4 = t + \tfrac{8}{5}u_0'u_2' + \tfrac{4}{5}(u_1')^2 + \tfrac{2}{3}u_0'u_0'' \end{aligned} \tag{6.18}$$

In that case u_0 satisfies (6.10) with $A_3' = 1$, while u_1 and u_2 are undetermined (i.e, they consistently cancel out). Therefore the same results would have been obtained without expanding u in a power series, but only w. The class $p = 1$ has also been considered in [26].

7. ORIGIN OF INTEGRABILITY IN MATRIX MODELS

In this section we want to point out that integrability is not limited to the string equations in the double scaling limit, but that it is rooted in the recursion relation (3.17), since this relation admits a Lax formulation of the form

$$[L, U'(L)_-] = 1 \tag{7.1}$$

The notation will be explained below; suffices to say at this point that the Lax operator L is completely specified by the basic recursion formula (3.7) for the orthogonal polynomials. Thus complete integrability turns out to be a direct consequence of the solvability of matrix models in terms of orthogonal polynomials [27-31]. (The results of this section are mainly borrowed from [31].)

Let us start by defining the orthonormal polynomials $Q_n = P_n/\sqrt{h_n}$, in terms of which (3.7) reads

$$\phi Q_n = \sqrt{R_{n+1}}Q_{n+1} + \sqrt{R_n}Q_{n-1} \tag{7.2}$$

This defines the multiplicative action of ϕ on the basis Q. Let us now work out the action of $\partial/\partial\phi$. Since differentiation is a linear operation, we must have

$$\frac{\partial}{\partial\phi}Q_n = \sum_{m<n} K_{mn}Q_m \tag{7.3}$$

The restriction in the summation expresses the fact that $\partial Q_n/\partial\phi$ is a polynomial of maximum degree $n-1$. Multiplying both sides by Q_l $(l < n)$ and integrating with respect to the measure $\mu(\phi)$, one finds

$$K_{nl} = \int \mathrm{d}\mu(\phi) Q_l \frac{\partial Q_n}{\partial\phi} \tag{7.4}$$

This is integrated by parts, and since $\partial Q_l/\partial\phi$ is necessarily orthogonal to Q_n for $n > l$, one finally obtains

$$K_{nl} = N \int \mathrm{d}\mu(\phi) Q_n Q_l U'(\phi) \tag{7.5}$$

We now claim that the recursion relation (3.15) is the compatibility equation of (7.2) and (7.5).

Let us work out in detail the simplest example, where U is a quartic potential:

$$K_{nl} = N \int d\mu(\phi) Q_n (2g_2\phi + 4g_4\phi^3) Q_l \tag{7.6}$$

Clearly $K_{nl} = 0$ unless $l = n-1$ or $n-3$, and the non-zero terms are readily calculated:

$$\begin{aligned} K_{n,n-1}/N &= 2g_2\sqrt{R_n} + 4g_4\sqrt{R_n}(R_{n-1} + R_n + R_{n+1}) \\ K_{n,n-3}/N &= 4g_4\sqrt{R_n R_{n-1} R_{n-3}} \end{aligned} \tag{7.7}$$

and we have

$$\frac{\partial Q_n}{\partial \phi} = K_{n,n-1}Q_{n-1} + K_{n,n-3}Q_{n-3} \tag{7.8}$$

Let us now differentiate (7.2) with respect to ϕ:

$$\phi\frac{\partial Q_n}{\partial \phi} + Q_n = \sqrt{R_{n+1}}\frac{\partial Q_{n+1}}{\partial \phi} + \sqrt{R_n}\frac{\partial Q_{n-1}}{\partial \phi} \tag{7.9}$$

We then use (7.8) and (7.2) to eliminate the derivatives and the factors of ϕ. Finally we equate to zero the coefficients of Q_{n-4}, Q_{n-2} and Q_n. With the expression (7.7), the first two vanish identically, while the coefficient of Q_n reads

$$1 + K_{n,n-1}\sqrt{R_n} - K_{n+1,n}\sqrt{R_{n+1}} = 0 \tag{7.10}$$

which implies that

$$\begin{aligned} n &= K_{n,n-1}\sqrt{R_n} \\ &= N(2g_2R_n + 4g_4R_n(R_{n-1} + R_n + R_{n+1})) \end{aligned} \tag{7.11}$$

This is the desired recursion relation.

The recursion relation can thus be written in the form of a Lax equation. Let us introduce the operator

$$L \equiv \Delta\sqrt{R} + \sqrt{R}\Delta \tag{7.12}$$

where the action of the operators Δ and $\sqrt{R}$ on the Q_n is given by

$$(\Delta Q)_n = Q_{n+1} \quad , \quad (\sqrt{R}Q)_n = \sqrt{R_n}Q_n \tag{7.13}$$

In terms of L, (7.2) becomes

$$LQ = \phi Q \quad . \tag{7.14}$$

This implies that $U'(\phi)Q_n = (U'(L)Q)_n$, so that (7.3) becomes

$$\frac{\partial Q}{\partial \phi} = U'(L)_- Q \tag{7.15}$$

Here we have introduced a decomposition of the operators in terms of the grading in Δ (i.e. $\mathrm{grad}(\Delta^l) = l$):

$$A = A_+ + A_0 + A_- \tag{7.16}$$

where A_+ contains all the terms in A with positive grading, and similarly for A_0 and A_-. For instance, $L_- = \sqrt{R}\Delta^{-1}$. The negative projection in (7.15) is equivalent to the restriction $m < n$ in the summation (7.3). The compatibility of (7.14) and (7.15) is just (7.1).

For the one-branch ansatz, the operator L has the following continuum limit:

$$L = 2 - a^2(\partial^2 - f) \tag{7.17}$$

so that $(2-L)/a^2$ is the usual KdV Lax operator: $\mathcal{L} = \partial^2 - f$. The continuous version of (7.1) is just [32]

$$1 = [\mathcal{L}, \mathcal{L}_+^{(2k+1)/2}] \tag{7.18}$$

where the + projects a pseudodifferential operator onto its differential part. These equations are just the t derivatives of eqs.(5.12).

8. CONCLUDING REMARKS.

8.1. Well-defined and ill-defined matrix models.

Matrix models are well-defined if the matrix potential $U(\phi)$ is bounded below, ensuring that the matrix integration measure $d\phi \exp -NU(\Phi)$ makes sense. For a potential of degree $2d$, this requires $g_{2d} > 0$, and a glance at table 1 shows that this condition is not always satisfied.

In general the highest coupling is proportional to the coefficient of the highest power of $G(z)$ (see (4.21)): $g_{2d} = G_{2d-2}/d$. For the potentials corresponding to the critical distributions (6.3), this means that $\mathrm{sign}(g_{2d}) = (-)^{p+1}$, and multicritical matrix models are well-defined only for p odd. When $p > r$, $k = p$

and so models with even k are ill-defined. This applies in particular to all the cases where the conditions for a one-branch ansatz are satisfied. Now, when $p < r$, this consequence can be avoided simply because p is no longer related to k (which is now equal to r). Therefore, if p is odd, the models are well-defined for all $k > p$. By choosing $p = 1$, one can construct well-defined models for all even values of k [25]. This seems to be the only way to obtain well-defined critical models with even k. In fact it is impossible to flow from a (well-defined) model with $k = 2l + 1$ to a model with $k = 2l$ [33]: attempts to *regularize* an ill-defined model by adding to U_{2l} a small piece proportional to U_{2l+1} induces chaotic behavior in the recursion coefficients, thus invalidating the assumptions of smoothness that led to the string equations [34-36].

8.2. Conformal field theory interpretation.

It has been argued that the models with $r = 0$ describe two-dimensional gravity coupled to minimal conformal matter of central charge

$$c = 1 - \frac{6(p-q)^2}{pq} \qquad (p, q \text{ coprime}) \tag{8.1}$$

with $q = 2$ and $p = 2k-1$. For $k = 2$, the central charge vanishes and this situation should correspond to pure gravity. Thus pure gravity should be described non-perturbatively by the Painlevé I equation [12-14], even though the corresponding matrix model is ill-defined. At this point there is no firm conformal field theory interpretation of the models with $r \neq 0$.

Other minimal conformal models can be described in terms of matrix models with more than one matrix. In fact, the required number of matrices coupled linearly in U is $q - 1$. The corresponding string equations are related to extended KdV equations, with a differential Lax operator of order q, the other operator of the Lax equation being of order p [32].

ACKNOWLEDGEMENTS

P.M. would like to thank the organizers of the School for their invitation to present this work, and for their excellent organization, which made the School so enjoyable and stimulating. This work was supported by NSERC (Canada) and F.C.A.R. (Québec).

REFERENCES

1. A. Bilal, CERN preprint TH-5867/90 (1990).
2. J. Zinn-Justin, Saclay preprint (1990).
3. F. David, Saclay preprint (1990).
4. D. Sénéchal, Laval preprint PHY-20/91 (1991).
5. V.A. Kazakov, Phys. Lett. **B150**, 282 (1985).
6. F. David, Nucl. Phys. **B257**, 45 (1985).
7. V.A. Kazakov, I.K. Kostov et A.A. Migdal, Phys. Lett. **B157**, 295 (1985).
8. J. Ambjørn, B. Durhuus et J. Fröhlich, Nucl. Phys. **B257**, 433 (1985).
9. G. 't Hooft, Nucl. Phys. **B72**, 461 (1974).
10. A.A. Migdal, Physics Reports **102**, 201 (1983).
11. D. Bessis, C. Itzykson et J.-B. Zuber, Adv. Appl. Math. **1**, 109 (1980).
12. E. Brézin et V.A. Kazakov, Phys. Lett. **B236**, 144 (1990).
13. M.R. Douglas et S.H. Shenker, Nucl. Phys. **B335**, 635 (1990).
14. D.J. Gross et A.A. Migdal, Phys. Rev. Lett. **64**, 127 (1990). Nucl. Phys. **B340**, 333 (1990).
15. D. Bessis, Commun. Math. Phys. **69**, 147 (1979).
16. E. Brézin, C. Itzykson, G. Parisi et J.-B. Zuber, Commun. Math. Phys. **59**, 35 (1978).
17. V.A. Kazakov, Phys. Lett. **A119**, 140 (1989).
18. H. Neuberger, *Regularized strings and flow equations*, Rutgers preprint RU-90-50 (1990).
19. L. Molinari, J. Phys. **A 21**, 1 (1988).
20. J. Jurkiewicz, Phys. Lett. **B245** (1990) 178.
21. M. Douglas, N. Seiberg and S. Shenker, Phys. Lett. **B244** (1990) 381.
22. K. Demeterfi, N. Deo, S. Jain and C.-I. Tan, Phys. Rev. **D42** (1990).
23. O. Lechtenfeld, R. Ray and A. Ray, *Phase diagram and orthogonal polynomials in multiple-well matrix models*, preprint IASSNS-HEP-90/69 (1990).
24. P. Mathieu et D. Sénéchal, *Tachyonic Hermitian one-matrix models*, Laval preprint PHY-22/90 (1990).
25. P. Mathieu and D. Sénéchal, *A well-defined multicritical series in Hermitian one-matrix models*, Laval preprint PHY-21/91 (1991).
26. C. Crnkovic and G. Moore, *Multicritical multi-cut matrix models*, Rutgers preprint RU-90-61 (YCTP-P16-90) (1990).
27. E. Witten, *Two-dimensional gravity and intersection theory on moduli space*, preprint IASSNS/HEP-90/45 (1990).
28. E. Martinec, *On the origin of integrability in matrix models*, Chicago U. preprint EFI-90-67 (1990).

29. A. Gerasimov et al., *Matrix models of 2-D quantum gravity and Toda theory*, Lebedev preprint (1990).
30. L. Alvarez-Gaumé, G. Gomez and J. Lacki, *Integrability in random matrix models*, CERN preprint TH/5875/90 (1990).
31. A.S. Fokas, A.R. Its and A.V. Kitaev, *Dicrete Painlevé equations and their appearance in quantum gravity*, Clarkson U. preprint INS/164.
32. M.R. Douglas. Phys. Lett. **B238**, 176 (1990).
33. M.R. Douglas, N. Seiberg et S.H. Shenker. Phys. Lett., **B244**, 381 (1990).
34. D. Sénéchal, *Chaos in the Hermitian one-matrix model*, Laval preprint PHY-22/91 (1991).
35. M. Sasaki and H. Suzuki, *On matrix realization of random surfaces*, preprint YITP/U-90-29 (1990).
36. O. Lechtenfeld, *On eigenvalue tunneling in matrix models*, preprint IASSNS-HEP-91/2.

(m)KdV-SOLITON SOLUTIONS ON QUASI-PERIODIC FINITE-GAP BACKGROUNDS

F. Gesztesy
Department of Mathematics, University of Missouri,
Columbia, MO 65211, USA

ABSTRACT. Commutation and τ-function methods together with algebro-geometric techniques are used to derive (m)KdV-soliton solutions relative to quasi-periodic finite-gap background solutions.

1. Introduction

In this contribution we report on a recent construction of (m)KdV-soliton solutions relative to quasi-periodic finite-gap background solutions in [18]. The methods employed are commutation and τ-function techniques combined with an algebro-geometric approach in terms of Baker-Akhiezer functions.

Historically, the first treatment of KdV-soliton solutions relative to a genus-one background KdV-solution appears to be due to [25] (see also [41]). The general finite-genus KdV-case has been studied in [24] and later on in [2], [8], [35], [39], [42]. Here we shall summarize the first complete treatment of such soliton solutions on nontrivial backgrounds in [18] simultaneously deriving analogous results for the mKdV-equation.

In Section 2 we briefly recall the construction of quasi-periodic finite-gap (m)KdV-solutions in terms of Riemann theta functions. In our main Section 3 we summarize the key results of [18] on (m)KdV-soliton solutions relative to the quasi-periodic finite-gap background solutions of Section 2. The single and double commutation methods (i.e., Darboux-type transformations) we use in Sections 2 and 3 are reviewed in Appendix A. Appendix B finally provides the τ-function techniques and Miura resp. Bäcklund-type transformations essential in our approach.

2. Quasi-Periodic Finite-Gap (m)KdV-Solutions

In this section we briefly recall the construction of quasi-periodic finite-gap (m)KdV-solutions in terms of Riemann theta functions. For reviews on this subject we refer e.g. to [7], [17], [28], [43].

We freely use the notation employed in Appendices A and B.

Let

$$-\infty < E_0 < E_1 < \ldots < E_{2g-1} < E_{2g} < \infty,\ g \in \mathbb{N} \tag{2.1}$$

and define

$$\rho_0 = (-\infty, E_0),\ \rho_n = (E_{2n-1},\ E_{2n}),\ 1 \leq n \leq g, \tag{2.2}$$

$$R_0(z)^{1/2} = \left[\prod_{n=0}^{2g}(E_n - z)\right]^{1/2},\ z \in \mathbb{C} \cup \{\infty\}. \tag{2.3}$$

Let $\prod_{\pm}$ be two copies of the cut plane $\prod_0 = \{\mathbb{C} \cup \{\infty\}\} \setminus \bigcup_{n=0}^{2g} \rho_n$. Then a topological model for the hyperelliptic Riemann surface S_g associated with (2.3) is realized in the usual way by joining the upper and lower rims of the cuts $\overline{\rho_n}$ of $\prod_{\pm}$ crosswise. Equivalently, we may consider the hyperelliptic curve K_g consisting of points

$$P = (z,\ R_0(z)^{1/2}),\ z \in \mathbb{C},\ P_\infty = (\infty, \infty) \tag{2.4}$$

with branch points

$$(E_n, 0),\ 0 \leq n \leq 2g,\ P_\infty = (\infty, \infty). \tag{2.5}$$

Here z denotes the projection of P onto $\mathbb{C} \cup \{\infty\}$ and the hyperelliptic involution $*$ (the sheet exchange map) is defined by

$$*: \begin{cases} K_g \longrightarrow K_g \\ P = (z, R_0(z)^{1/2}) \to P^* = (z, -R_0(z)^{1/2}). \end{cases} \tag{2.6}$$

The upper sheet $\prod_+$ is declared as follows: Define $\lim_{\epsilon \downarrow 0} R_0(\lambda + i\epsilon)^{1/2} = -|R_0(\lambda)|^{1/2}$ for $\lambda < E_0$ on $\prod_+$ and analytically continue in λ. Local parameters at $P_0 = (z_0, R_0(z_0)^{1/2}) \in K_g$ are given by

$$\zeta_\alpha = \begin{cases} z - z_0, & z_0 \in \mathbb{C} \backslash \{\{E_n\}_{n=0}^{2g} \cup \{\infty\}\} \\ (z - E_n)^{1/2}, & z_0 = E_n,\ 0 \leq n \leq 2g \\ z^{-1/2}, & z_0 = \infty \end{cases} \tag{2.7}$$

with (U_α, ζ_α), $\alpha \in J$ (an index set) denoting appropriate charts at P_0. A convenient homology basis $\{a_j, b_j\}_{j=1}^g$ on K_g is then chosen as follows: The cycle a_j surrounds the cut $\overline{\rho_j} = [E_{2j-1}, E_{2j}]$, $1 \leq j \leq g$ clockwise on $\prod_+$ while b_j starts at the lower rim of $\overline{\rho_j}$ on $\prod_+$, intersects a_j, then encircles E_0 clockwise thereby changing to the lower sheet $\prod_-$ and returns on $\prod_-$ to its initial point. Their intersection matrix is chosen to be

$$a_j \circ b_\ell = \delta_{j,\ell},\ 1 \leq j, \ell \leq g. \tag{2.8}$$

A basis for the Abelian differentials of the first kind (DFK) is given by

$$\eta_j = R_0(z)^{-1/2} z^{j-1} dz,\ 1 \leq j \leq g \tag{2.9}$$

and we use the standard normalization

$$\omega_j = \sum_{\ell=1}^{g} c_{j,\ell} \eta_\ell,\ \int_{a_j} \omega_\ell = \delta_{j,\ell},\ 1 \leq j, \ell \leq g. \tag{2.10}$$

The b-periods of ω_ℓ are then given by

$$\tau_{j,\ell} = \int_{b_j} \omega_\ell,\ 1 \leq j, \ell \leq g \tag{2.11}$$

and due to Riemann's period relation and (2.1), one infers

$$\tau_{j,\ell} = \tau_{\ell,j},\ 1 \leq j, \ell \leq g,\ \tau = iT,\ T > 0. \tag{2.12}$$

Abelian differentials of the second kind (DSK) $\omega^{(2)}$ are characterized by their vanishing residues and the normalization

$$\int_{a_j} \omega^{(2)} = 0,\ 1 \leq j \leq g. \tag{2.13}$$

A basis for DSKs at $P_\infty \in K_g$ holomorphic on $K_g \backslash \{P_\infty\}$ is provided by

$$\omega_{2k}^{(2)} = R_0(z)^{-1/2} z^{g+k} dz,\ k \in \mathbb{N}_0. \tag{2.14}$$

The Jacobi variety $J(K_g)$ of K_g is defined by

$$J(K_g) = \mathbb{C}^g / L_g, \tag{2.15}$$

where L_g denotes the period lattice

$$L_g = \{\underline{z} \in \mathbb{C}^g \,|\, \underline{z} = \underline{N} + (\tau \underline{M}),\ \underline{N}, \underline{M} \in Z^g\}. \tag{2.16}$$

The Abel map is defined by

$$\underline{A}_{P_0} : \begin{cases} K_g \to J(K_g) \\ P \ \to \underline{A}_{P_0}(P) = \left\{ A_{P_0,j}(P) = \int_{P_0}^{P} \omega_j \right\}_{j=1}^{g} \pmod{L_g},\ P_0 = (E_0, 0), \end{cases} \tag{2.17}$$

respectively by

$$\underline{\alpha}_{P_0} : \begin{cases} \mathrm{Div}(K_g) \to \ J(K_g) \\ \mathcal{D} \longrightarrow \ \underline{\alpha}_{P_0}(\mathcal{D}) = \sum_{P \in K_g} \mathcal{D}(P) \underline{A}_{P_0}(P), \end{cases} \tag{2.18}$$

where $\mathcal{D}$ denotes a divisor on K_g (i.e., a map $\mathcal{D} : K_g \to Z$, $\mathcal{D}(P) \neq 0$ for only finitely many $P \in K_g$) and $\mathrm{Div}(K_g)$ the set of divisors on K_g. The Riemann theta function associated with K_g is then defined by

$$\theta(\underline{z}) = \sum_{\underline{m} \in Z^g} \exp\left[2\pi i(\underline{m}, \underline{z}) + \pi i(\underline{m}, \tau \underline{m})\right],\ \underline{z} \in \mathbb{C}^g \tag{2.19}$$

(with $(.,.)$ the scalar product on $\mathbb{C}^g$). Finally, Riemann's vector $\underline{\zeta}_{P_0} = (\zeta_{P_0,1}, \dots, \zeta_{P_0,g})$ is defined by

$$\zeta_{P_0,j} = 2^{-1}(1 + \tau_{j,j}) - \sum_{\substack{\ell=1 \\ \ell \neq j}}^{g} \int_{a_\ell} \omega_\ell(P) \int_{P_0}^{P} \omega_j,\ 1 \leq j \leq g. \tag{2.20}$$

Given these preliminaries we introduce the normalized Baker-Akhiezer (B-A) function

$$\psi_0(P,t,x) = \exp\left[-i(x - x_0) \int_{P_0}^{P} \omega_0^{(2)} - 12i(t - t_0) \int_{P_0}^{P} \omega_2^{(2)}\right] \cdot$$

$$\cdot \left[\tau_0(P_1 t, x)\tau_0(t_0, x_0)\right] / \left[\tau_0(t,x)\tau_0(P, t_0, x_0)\right],\ P \in K_g \backslash \{P_\infty\},\ (t,x) \in \mathbb{R}^2, \tag{2.21}$$

where

$$\tau_0(P,t,x) = \theta\left(\underline{\zeta}_{P_0} - \underline{A}_{P_0}(P) + \underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}(t_0, x_0)\right) + \frac{(x - x_0)}{2\pi} \underline{U}_0 + \frac{6(t - t_0)}{\pi} \underline{U}_2\right), \tag{2.22}$$

$$\tau_0(t,x) = \tau_0(P_\infty, t, x) \tag{2.23}$$

and $\omega_{2k}^{(2)}$, $k = 0, 1$ are normalized DSKs with a single pole at P_∞ and principle part

$$\omega_{2k}^{(2)} = \left[\zeta_\alpha^{-2-2k} + 0(1)\right] d\zeta_\alpha \tag{2.24}$$

at P_∞ whose b-periods are $\underline{U}_{2k}$, $k = 0, 1$. In particular, $\omega_0^{(2)}$ is of the type

$$\omega_0^{(2)} = (-i/2) R_0(z)^{-1/2} \prod_{j=1}^{g} (\lambda_j - z) dz \tag{2.25}$$

with $\lambda_j \in (E_{2j-1}, E_{2j})$, $1 \leq j \leq g$. In addition we used in (2.22) the simplified notation

$$\underline{\alpha}_{P_0}(\underline{P}) = \underline{\alpha}_{P_0}(\mathcal{D}_{P_1 + \dots + P_g}) \tag{2.26}$$

for $\underline{P} = (P_1, \dots, P_g) \in K_g^g$, $P_j \neq P_\ell$ for $j \neq \ell$ and

$$\mathcal{D}_{P_1 + \dots + P_g} : \begin{cases} K_g \to \mathbb{N}_0 \\ P \ \to \mathcal{D}_{P_1 + \dots + P_g}(P) = \begin{cases} 1, & P \in \{P_1, \dots, P_g\} \\ 0, & P \in K_g \backslash \{P_1, \dots, P_g\}. \end{cases} \end{cases} \tag{2.27}$$

Finally, the (Dirichlet) divisor $\underline{\hat{\mu}}(t_0, x_0) = (\hat{\mu}_1(t_0, x_0), \dots, \hat{\mu}_g(t_0, x_0)) \in K_g^g$ is chosen such that the projection μ_j of each $\hat{\mu}_j = \left(\mu_j, R_0(\mu_j)^{1/2}\right) \in K_g$ lies in $\overline{\rho_j} = [E_{2j-1}, E_{2j}]$, $1 \leq j \leq g$.

The celebrated Its-Matveev result [22] (see also [28]) then can be summarized as follows: Define

$$V_0(t,x) = E_0 + \sum_{j=1}^{g} (E_{2j-1} + E_{2j} - 2\lambda_j) - 2\partial_x^2 \ln \tau_0(t,x) \tag{2.28}$$

and the Schrödinger operator in $L^2(\mathbb{R})$

$$H_0(t) = -\frac{d^2}{dx^2} + V_0(t),\ \mathcal{D}(H_0(t)) = H^2(\mathbb{R}). \tag{2.29}$$

Then

$$(H_0(t) - z)\, \psi_0(P^{(*)}, t) = 0, \tag{2.30}$$

$$(\partial_t - B_{V_0}(t))\, \psi_0(P^{(*)}, t) = 0 \tag{2.31}$$

(cf. (B.9)) and hence

$$\mathrm{KdV}(V_0) = 0 \tag{2.32}$$

(by (B.16)). The spectrum of $H_0(t)$ is purely absolutely continuous of multiplicity two and consists of finitely many gaps and bands

$$\sigma(H_0(t)) = \bigcup_{j=1}^{g} [E_{2(j-1)}, E_{2j-1}] \cup [E_{2g}, \infty). \tag{2.33}$$

In addition, V_0 is a quasi-periodic function of $t \in \mathbb{R}$ and $x \in \mathbb{R}$ and (2.28) describes all KdV-solutions that yield the finite-gap spectrum (2.33).

The above results are reviewed e.g. in [7], [8], [17], [28], [35]. In the remainder of this section we recall the corresponding construction of the quasi-periodic finite-gap mKdV-solutions following [17]. We start by introducing Floquet-type solutions $\psi_{0,\pm 1}$ of $H_0(t)$ as the two branches of the B-A function ψ_0 by

$$\psi_{0,+1}(z,t,x) = \psi_0(P,t,x), \ \psi_{0,-1}(z,t,x) = \psi_0(P^*,t,x), \tag{2.34}$$

$$P = \left(z, R_0(z)^{1/2}\right) \in \prod\nolimits_+ \backslash \overline{\rho_{H_0}}, \quad \overline{\rho_{H_0}} = [-\infty, E_0] \cup \bigcup_{j=1}^{g} [E_{2j-1}, E_{2j}]. \tag{2.35}$$

On $\overline{\rho_{H_0}} \backslash \{\infty\}$ we define

$$\psi_{0,\pm 1}(\lambda,t,x) = \lim_{\epsilon \downarrow 0} \psi_{0,\pm 1}(\lambda + i\epsilon, t, x), \ \lambda \in \overline{\rho_{H_0}} \backslash \{\infty\}. \tag{2.36}$$

We also assume that the path joining P_0 and P in $\psi_{0,\pm 1}$ (cf. (2.21), (2.22)) lies entirely on $\prod_\pm \cup \overline{\rho_{H_0}}$. Then

$$\begin{aligned} &0 < \psi_{0,\pm 1}(\lambda,t,.) \in L^2((R,\pm\infty)), \ \lambda < E_0, \ R \in \mathbb{R}, \\ &0 < \psi_{0,+1}(E_0,t,x) = \psi_{0,-1}(E_0,t,x). \end{aligned} \tag{2.37}$$

In addition, assuming

$$\lambda_0 < E_0, \quad \sigma_0 \in [-1,1] \tag{2.38}$$

we introduce

$$\psi_{0,\sigma_0}(t,x) = \frac{1}{2}(1-\sigma_0)\psi_{0,-1}(\lambda_0,t,x) + \frac{1}{2}(1+\sigma_0)\psi_{0,+1}(\lambda_0,t,x), \tag{2.39}$$

$$\phi_{\sigma_0}(t,x) = \partial_x \ln \psi_{0,\sigma_0}(t,x), \tag{2.40}$$

$$A_{\sigma_0}(t) = \frac{d}{dx} + \phi_{\sigma_0}(t), \ \mathcal{D}(A_{\sigma_0}(t)) = H^1(\mathbb{R}), \tag{2.41}$$

$$H_2(t) = A_\sigma(t)A_\sigma(t)^* = -\frac{d^2}{dx^2} + V_2(t), \ \mathcal{D}(H_2(t)) = H^2(\mathbb{R}), \tag{2.42}$$

$$V_2(t,x) = \phi_{\sigma_0}(t,x)^2 + \phi_{\sigma_0,x}(t,x). \tag{2.43}$$

Then

$$\partial_x^m \phi_{\sigma_0} \in C^\infty(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2), \ m \in \mathbb{N}_0, \ \sigma_0 \in [-1,1] \tag{2.44}$$

and we obtain from (B.25), (B.27)

$$\begin{aligned} \psi_{0,\pm 1}(\lambda,t,x) = \exp\left[\mp i(x-x_0)\int_{P_0}^{P} \omega_0^{(2)} \mp 12i(t-t_0)\int_{P_0}^{P} \omega_2^{(2)}\right] \bullet \\ \bullet \left[\tau_{0,\pm 1}(\lambda,t,x)\tau_0(t_0,x_0)\right] / \left[\tau_0(t,x)\tau_{0,\pm 1}(\lambda,t_0,x_0)\right], \end{aligned} \tag{2.45}$$

$$\begin{aligned} V_0(t,x) &= \phi_{\sigma_0}(t,x)^2 + \phi_{\sigma_0,x}(t,x) + \lambda_0 \\ &= E_0 + \sum_{j=1}^{g} (E_{2j-1} + E_{2j} - 2\lambda_j) - 2\partial_x^2 \ln \tau_0(t,x), \end{aligned} \tag{2.46}$$

$$\phi_{\pm 1}(t,x) = \mp i \int_{P_0}^{Q_0} \omega_0^{(2)} + \partial_x \ln\left[\tau_{0,\pm 1}(\lambda_0,t,x)/\tau_0(t,x)\right], \tag{2.47}$$

$$Q_0 = \left(\lambda_0, \lim_{\epsilon \downarrow 0} R_0(\lambda_0 + i\epsilon)^{1/2}\right),$$

where

$$\tau_{0,\pm 1}(\lambda,t,x) = \theta\left(\underline{\zeta}_{P_0} \mp \underline{A}_{P_0}(P) + \underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}(t_0,x_0)\right) + \frac{(x-x_0)}{2\pi}\underline{U}_0 + \frac{6(t-t_0)}{\pi}\underline{U}_2\right), \tag{2.48}$$

$$\tau_0(t,x) = \theta\left(\underline{\zeta}_{P_0} \mp \underline{A}_{P_0}(P_\infty) + \underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}(t_0,x_0)\right) + \frac{(x-x_0)}{2\pi}\underline{U}_0 + \frac{6(t-t_0)}{\pi}\underline{U}_2\right), \tag{2.49}$$

$$P = \left(\lambda, R_0(\lambda)^{1/2}\right) \in \prod\nolimits_+ \backslash \overline{\rho_{H_0}} \text{ or } P = \left(\lambda, \lim_{\epsilon \downarrow 0} R_0(\lambda + i\epsilon)^{1/2}\right) \in \prod\nolimits_+, \lambda \in \overline{\rho_{H_0}} \backslash \{\infty\}. \tag{2.50}$$

We have singled out $\phi_{\pm 1}$ in (2.47) since ϕ_{σ_0} is quasi-periodic in x iff $\sigma_0 = \pm 1$ [27]. In addition, (B.8), (B.26) imply

$$\begin{aligned} V_{0,\pm 1}(t,x) &= \phi_{\pm 1}(t,x)^2 - \phi_{\pm 1,x}(t,x) + \lambda_0 \\ &= E_0 + \sum_{j=1}^{g} (E_{2j-1} + E_{2j} - 2\lambda_j) - 2\partial_x^2 \ln \tau_{0,\pm 1}(\lambda_0,t,x). \end{aligned} \tag{2.51}$$

Applying (B.21), (B.24) and general commutation formulas [5], [19] the analogs of assertions (2.28)–(2.33) in the mKdV-context then read [17]:

$$\mathrm{KdV}(V_0) = 0, \quad \mathrm{KdV}(V_{0,\sigma_0}) = 0, \ \sigma_0 \in [-1,1], \tag{2.52}$$

$$\mathrm{mKdV}(\phi_{\sigma_0}; \lambda_0) = 0, \ \sigma_0 \in [-1,1], \tag{2.53}$$

where

$$V_{0,\sigma_0}(t,x) = \phi_{\sigma_0}(t,x)^2 - \phi_{\sigma_0,x}(t,x) + \lambda_0. \tag{2.54}$$

Define in $L^2(\mathbb{R})$

$$H_{0,\sigma_0}(t) = A_{\sigma_0}(t)^* A_{\sigma_0}(t) + \lambda_0 = -\frac{d^2}{dx^2} + V_{0,\sigma_0}(t), \ \mathcal{D}(H_{0,\sigma_0}(t)) = H^2(\mathbb{R}) \tag{2.55}$$

and

$$Q_{\sigma_0}(t) = \begin{pmatrix} 0 & A_{\sigma_0}(t)^* \\ A_{\sigma_0}(t) & 0 \end{pmatrix}, \ \mathcal{D}(Q_{\sigma_0}(t)) = \left[H^2(\mathbb{R})\right]^2 \tag{2.56}$$

in $\left[L^2(\mathbb{R})\right]^2$. Then

$$\sigma\left(H_{0,\sigma_0}(t)\right) = \begin{cases} \sigma(H_0(t)), \ \sigma_0 = \pm 1 \\ \sigma\left(H_0(t)\right) \cup \{\lambda_0\}, \ \sigma_0 \in (-1,1), \end{cases} \tag{2.57}$$

$$\sigma\left(Q_{\sigma_0}(t)\right) = \begin{cases} \bigcup\limits_{\substack{j=-g \\ j\neq 0}}^{g} \sum_j \cup \Big(-\infty, -|E_{2g} - \lambda_0|^{1/2}\Big] \cup \Big[|E_{2g} - \lambda_0|^{1/2}, \infty\Big), \ \sigma_0 = \pm 1 \\ \bigcup\limits_{\substack{j=-g \\ j\neq 0}}^{g} \sum_j \cup \{0\} \cup \Big(-\infty, -|E_{2g} - \lambda_0|^{1/2}\Big] \cup \Big[|E_{2g} - \lambda_0|^{1/2}, \infty\Big), \\ \qquad\qquad \sigma_0 \in (-1,1), \end{cases} \tag{2.58, 2.59}$$

$$\sum\nolimits_j = \Big[|E_{2(j-1)} - \lambda_0|^{1/2}, \ |E_{2j-1} - \lambda_0|^{1/2}\Big], \ \sum\nolimits_{-j} = -\sum\nolimits_j, \ 1 \leq j \leq g \tag{2.60}$$

and the essential spectra of $H_{0,\sigma_0}(t)$ and $Q_{\sigma_0}(t)$ are both purely absolutely continuous and of multiplicity two. In addition, V_{0,σ_0} and ϕ_{σ_0} are quasi-periodic functions of $t \in \mathbb{R}$ and $x \in \mathbb{R}$ iff $\sigma_0 = \pm 1$ and (2.47) describes all solutions of $\mathrm{mKdV}(.;\lambda_0) = 0$ that yield the finite-gap spectrum (2.58).

A simple Galilei transformation $(t,x) \to (t, y = x + 6\lambda_0 t)$ then transforms solutions ϕ of $\mathrm{mKdV}(.;\lambda_0) = 0$ into solutions χ of the mKdV-equation by noting the identity

$$\begin{aligned} &\phi_t - 6(\phi^2 + \lambda_0)\phi_x + \phi_{xxx} = \chi_t - 6\chi^2\chi_y + \chi_{yyy}, \\ &\phi(t,x) = \chi(t,y), \ y = x + 6\lambda_0 t. \end{aligned} \tag{2.61}$$

Similarly one has

$$\begin{aligned} &W_t - 6WW_x + W_{xxx} = U_t - 6UU_y + U_{yyy}, \\ &W(t,x) = U(t,y) + \lambda_0, \ y = x + 6\lambda_0 t. \end{aligned} \tag{2.62}$$

Relations (2.61) and (2.62) should also be compared with (B.23).

3. (m)KdV-Soliton Solutions on Quasi-Periodic Finite-Gap Backgrounds

In this section we report on the main results in [18] concerning (m)KdV-soliton solutions relative to quasi-periodic finite-gap backgrounds. According to the single and double commutation methods summarized in Appendix A we divide this section into two parts starting with the single commutation approach.

We suppose $V_0(t,x)$ to be the quasi-periodic finite-gap KdV-solution (2.28),

$$V_0(t,x) = \Lambda_0 - 2\partial_x^2 \ln \tau_0(t,x), \tag{3.1}$$

$$\mathrm{KdV}(V_0) = 0, \tag{3.2}$$

where

$$\Lambda_0 = E_0 + \sum_{j=1}^{g} (E_{2j-1} + E_{2j} - 2\lambda_j), \tag{3.3}$$

with the associated Floquet-type solutions (2.45),

$$\psi_{0,\pm 1}(\lambda,t,x) = F_{\pm}(\lambda,t,x)\tau_{0,\pm 1}(\lambda,t,x)/\tau_0(t,x), \tag{3.4}$$

$$F_{\pm}(\lambda,t,x) = \exp\left[\mp i(x-x_0)\int_{P_0}^{P} \omega_0^{(2)} \mp 12i(t-t_0)\int_{P_0}^{P} \omega_2^{(2)}\right] \tau_0(t_0,x_0)/\tau_{0,\pm 1}(\lambda,t_0,x_0) \tag{3.5}$$

and $\tau_{0,\pm 1}, \tau_0$ defined in (2.48)–(2.50). Given (3.4) (and (A.1)–(A.5)) we introduce the quantities $A_{\sigma_1,\dots,\sigma_N}(t)$, $\phi_{\sigma_1,\dots,\sigma_N}(t,x)$, $\psi_{0,\sigma_1,\dots,\sigma_N}(t,x)$, $H_{0,\sigma_1,\dots,\sigma_N}(t)$

and $V_{0,\sigma_1,\dots,\sigma_N}(t,x)$ according to (A.6)–(A.10), (A.12) (with σ_j t-independent) and hence have the results (A.13)–(A.25) at our disposal. E.g., (A.11) implies

$$(H_{0,\sigma_1,\dots,\sigma_N}(t)-\lambda)\psi_{0,\pm1,\sigma_1,\dots,\sigma_N}(\lambda,t)=0,\quad \lambda<\lambda_1 \tag{3.6}$$

and applying (B.18) (with $\beta_2=0$), (B.20) and (B.24) repeatedly, one infers

$$(\partial_t - B_{V_{0,\sigma_1,\dots,\sigma_N}}(t))\psi_{0,\pm1,\sigma_1,\dots,\sigma_N}(\lambda,t)=0,\quad \lambda<\lambda_1 \tag{3.7}$$

and hence

$$\begin{aligned}&\mathrm{KdV}(V_{0,\sigma_1,\dots,\sigma_N})=0,\ \sigma_j\in[-1,1],\ 1\le j\le N,\\ &\mathrm{mKdV}(\phi_{\sigma_1,\dots,\sigma_N};\lambda_1)=0,\ \sigma_j\in[-1,1],\ 1\le j\le N.\end{aligned} \tag{3.8}$$

A careful analysis [18] invoking the solution of Jacobi's inversion problem then yields for the asymptotic solutions $V_{0,\pm1,\dots,\pm1}$ (N times) and $\phi_{\pm1,\dots,\pm1}$ (N times) of $V_{0,\sigma_1,\dots,\sigma_N}$ and $\phi_{\sigma_1,\dots,\sigma_N}$ as $x\to\mp\infty$ (see (A.18), (A.19), (A.24), (A.25)) assuming for brevity that $\sigma_j\in(-1,1)$ for all $1\le j\le N$

$$\phi_{\pm1,\dots,\pm1}(t,x)=\mp i\int_{P_0}^{Q_1}\omega_0^{(2)}+\partial_x\ln\left[\tau_{0,\pm1,\dots,\pm1}(t,x)/\tau_{0,\pm1,\dots,\pm1}(t,x)\right], \tag{3.9}$$

(N times; N times; $(N-1)$ times)

$$\begin{aligned}V_{0,\pm1,\dots,\pm1}(t,x)&=\phi_{\pm1,\dots,\pm1}(t,x)^2-\phi_{\pm1,\dots,\pm1,x}(t,x)+\lambda_1\\ &=\Lambda_0-2\partial_x^2\ln\tau_{0,\pm1,\dots,\pm1}(t,x),\end{aligned} \tag{3.10}$$

(N times; N times; N times; N times)

where

$$\tau_{0,\pm1,\dots,\pm1}(t,x)=\theta\left(\underline{\zeta}_{P_0}\mp\underline{A}_{P_0}(P_\infty)+\underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}_{0,\pm1,\dots,\pm1}(t,x)\right)\right), \tag{3.11}$$

($(N+1-j)$ times; $(N+1-j)$ times)

$$\begin{aligned}\underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}_{0,\pm1,\dots,\pm1}(t,x)\right)&=\pm\sum_{\ell=j}^{N}\left[\underline{A}_{P_0}(P_\infty)-\underline{A}_{P_0}(Q_\ell)\right]+\underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}(t_0,x_0)\right)\\ &\quad+\frac{(x-x_0)}{2\pi}\underline{U}_0+\frac{6(t-t_0)}{\pi}\underline{U}_2,\end{aligned}$$

($(N+1-j)$ times)

$$Q_j=\left(\lambda_j,\lim_{\epsilon\downarrow0}R_0(\lambda_j+i\epsilon)^{1/2}\right),\ 1\le j\le N. \tag{3.12}$$

A comparison of the arguments of the two KdV-solutions $V_{0,\pm1,\dots,\pm1}$ yields

$$\underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}_{0,+1,\dots,+1}(t,x)\right)-\underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}_{0,-1,\dots,-1}(t,x)\right)=-2\sum_{j=1}^{N}\underline{A}_{P_0}(Q_j)(\mathrm{mod}\ L_g) \tag{3.13}$$

and hence

$$V_{0,+1,\dots,+1}=V_{0,-1,\dots,-1},\phi_{0,+1,\dots,+1}=\phi_{0,-1,\dots,-1}\ \text{iff}\ 2\sum_{j=1}^{N}\underline{A}_{P_0}(Q_j)=\underline{0}(\mathrm{mod}\ L_g). \tag{3.14}$$

One can show that the higher order terms in (A.18) and (A.19) are actually exponentially decaying in the present finite-gap context.

By definition $V_{0,\sigma_1,\dots,\sigma_N}$ is an M-soliton KdV-solution ($0\le M\le N$) relative to the background KdV-solution V_0 in (2.28), where M is the number of $\sigma_j\in(-1,1)$, $1\le j\le N$, i.e., the number of eigenvalues of $H_{0,\sigma_1,\dots,\sigma_N}(t)$ in the spectral gap $(-\infty,E_0)$ of $H_0(t)$ (see (A.17) and the comment following it). Next, assuming

$$H_{0,\sigma_1,\dots,\sigma_N}(t_0)\ge0 \tag{3.15}$$

for some (and hence for all) $t_0\in\mathbb{R}$ we define

$$\phi_{0,\pm}(t,x)=\partial_x\ln\psi_{0,\pm1}(0,t,x),\ \phi_{N,\pm}(t,x)=\partial_x\ln\psi_{0,\pm1,\sigma_1,\dots,\sigma_N}(0,t,x) \tag{3.16}$$

and in $\left[L^2(\mathbb{R})\right]^2$

$$\begin{aligned}Q_{0,\pm}(t)&=\begin{pmatrix}0 & -\frac{d}{dx}+\phi_{0,\pm}(t)\\ \frac{d}{dx}+\phi_{0,\pm}(t) & 0\end{pmatrix},\\ Q_{N,\pm}(t)&=\begin{pmatrix}0 & -\frac{d}{dx}+\phi_{N,\pm}(t)\\ \frac{d}{dx}+\phi_{N,\pm}(t) & 0\end{pmatrix},\end{aligned}$$

$$\mathcal{D}(Q_{0,\pm}(t))=\mathcal{D}(Q_{N,\pm}(t))=\left[H^1(\mathbb{R})\right]^2. \tag{3.17}$$

Then

$$V_0=\phi_{0,\pm}^2+\phi_{0,\pm,x},V_{0,\sigma_1,\dots,\sigma_N}=\phi_{N,\pm}^2+\phi_{N,\pm,x} \tag{3.18}$$

and

$$\mathrm{KdV}(V_0)=\mathrm{KdV}(V_{0,\sigma_1,\dots,\sigma_N})=0,\ \mathrm{mKdV}(\phi_{0,\pm})=\mathrm{mKdV}(\phi_{N,\pm})=0. \tag{3.19}$$

If $\lambda_1 > 0$, we define $\phi_{N,+}$ (resp. $\phi_{N,-}$) to be a 2M-soliton mKdV-solution relative to the background mKdV-solution $\phi_{0,+}$ (resp. $\phi_{0,-}$) since then $Q_{N,\pm}(t)$ has 2M eigenvalues in the spectral gap $(-|E_0|^{1/2}, |E_0|^{1/2})$ of $Q_{0,\pm}(t)$. If $\lambda_1 = 0$ and $\sigma_1 \in \{-1, +1\}$ then $\phi_{N,+}$ (resp. $\phi_{N,-}$) is also a 2M-soliton solution, whereas if $\lambda_1 = 0$ and $\sigma_1 \in (-1, 1)$ then $\phi_{N,+}$ (resp. $\phi_{N,-}$) is called a $(2M-1)$-soliton solution relative to $\phi_{0,+}$ (resp. $\phi_{0,-}$) since in the former case $0 \notin \sigma(Q_{N,\pm}(t))$, whereas $0 \in \sigma(Q_{N,\pm}(t))$ in the latter case. (Here we used the fact that the spectra of $Q_{0,\pm}, Q_{N,\pm}$ are symmetric with respect to the origin and easily determined from $Q_{0,\pm}^2, Q_{N,\pm}^2$ [19].) Since the Galilei transformation (2.61) leaves the spectra of $Q_{0,\pm}, Q_{N,\pm}$ invariant, such a transformation applied to the mKdV $(.; \lambda_1)$-solutions $\phi_{\sigma_1,\dots,\sigma_N}$ in (3.8) yields mKdV-soliton solutions on quasi-periodic finite-gap backgrounds in the above sense.

It remains to discuss the double commutation procedure. Given $\psi_{0,\pm}(\lambda, t, x) \equiv \psi_{0,\pm 1}(\lambda, t, x)$ in (3.4) (and (A.26)–(A.28)) we introduce the quantities $\psi_{0,j,\pm}(t,x)$, $c_{\gamma_1,\dots,\gamma_N,\pm,\ell,m}(t,x)$, $C_{\gamma_1,\dots,\gamma_N,\pm}(t,x)$, $\Psi^{(0)}_{\gamma_1,\dots,\gamma_N,\pm}(t,x)$, $V_{\gamma_1,\dots,\gamma_N,\pm}(t,x)$ and $H_{\gamma_1,\dots,\gamma_1,\dots,\gamma_N,\pm}(t)$ according to (A.29)–(A.34) (with γ_j t-independent) and hence conclude (A.35)–(A.41) for the case at hand. In addition, applying (B.16), (B.21), (B.22) repeatedly, one infers

$$\mathrm{KdV}(V_{\gamma_1,\dots,\gamma_N,\pm}) = 0, \tag{3.20}$$

$$\mathrm{mKdV}(\phi_{\gamma_1,\dots,\gamma_N,\pm}) = 0, \tag{3.21}$$

where

$$V_{\gamma_1,\dots,\gamma_N,\pm}(t,x) = \phi_{\gamma_1,\dots,\gamma_N,\pm}(t,x)^2 + \phi_{\gamma_1,\dots,\gamma_N,\pm,x}(t,x). \tag{3.22}$$

A similar analysis that led to (3.9)–(3.12) now yields for the asymptotic solutions $V_{0,\dots,0,\pm}$, $V_{\infty,\dots,\infty,\pm}$ and $\phi_{0,\dots,0,\pm}$, $\phi_{\infty,\dots,\infty,\pm}$ of $V_{\gamma_1,\dots,\gamma_N,\pm}$ and $\phi_{\gamma_1,\dots,\gamma_N,\pm}$ (see (A.38)–(A.41), (A.49)–(A.51))

$$\begin{aligned} V_{0,\dots,0,\pm}(t,x) &= V_0(t,x) = \Lambda_0 - 2\partial_x^2 \ln \tau_0(t,x), \\ V_{\infty,\dots,\infty,\pm}(t,x) &= \Lambda_0 - 2\partial_x^2 \ln \tau_{\pm,\underset{2N \text{ times}}{\dots\dots},\pm}(t,x), \end{aligned} \tag{3.23}$$

$$\phi_{0,\dots,0,\pm}(t,x) = \mp i \int_{P_0}^{Q_0} \omega_0^{(2)} + \partial_x \ln\left[\tau_\pm(t,x)/\tau_0(t,x)\right],$$

$$\phi_{\infty,\dots,\infty,\pm}(t,x) = \mp i \int_{P_0}^{Q_0} \omega_0^{(2)} + \partial_x \ln\left[\tau_{\pm,\underset{(2N+1) \text{ times}}{\dots\dots},\pm}(t,x)/\tau_{\pm,\underset{2N \text{ times}}{\dots\dots},\pm}(t,x)\right], \tag{3.24}$$

where τ_0 is defined in (2.49), $Q_0 = (0, \lim_{\epsilon \downarrow 0} R_0(i\epsilon)^{1/2})$ and

$$\tau_{\pm,\underset{2N \text{ times}}{\dots\dots},\pm}(t,x) = \theta\Big(\underline{\zeta}_{P_0} \mp \underline{A}_{P_0}(P_\infty) \pm 2\sum_{j=1}^{N}\left[\underline{A}_{P_0}(P_\infty) - \underline{A}_{P_0}(Q_j)\right] + \underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}_0(t,x)\right)\Big)$$

$$N \in \mathbb{N}_0, \tag{3.25}$$

$$\tau_{\pm,\underset{(2N+1) \text{ times}}{\dots\dots},\pm}(t,x) = \theta\left(\underline{\zeta}_{P_0} \mp \underline{A}_{P_0}(Q_0) \pm 2\sum_{j=1}^{N}\left[\underline{A}_{P_0}(P_\infty) - \underline{A}_{P_0}(Q_j)\right] + \underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}_0(t,x)\right)\right)$$

$$N \in \mathbb{N}_0, \tag{3.26}$$

$$\underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}_0(t,x)\right) = \underline{\alpha}_{P_0}\left(\underline{\hat{\mu}}(t_0,x_0)\right) + \frac{(x-x_0)}{2\pi}\underline{U}_0 + \frac{6(t-t_0)}{\pi}\underline{U}_2 \tag{3.27}$$

(the sum $\sum_{j=1}^{N} \dots$ in (3.25) and (3.26) is replaced by zero for $N = 0$) and the points Q_j, $1 \le j \le N$ are defined as in (3.12). In analogy to (3.14) one infers that $V_{\gamma_1,\dots,\gamma_N,\pm}(t,x)$ tends to $V_0(t,x)$ as $x \to \pm\infty$ iff $2\sum_{j=1}^{N} \underline{A}_{P_0}(Q_j) = \underline{0} (\mathrm{mod}\ L_g)$.

As in the single commutation case one can show that the higher order terms in (A.38) and (A.49) are actually exponentially decaying in this context.

Similarly to the discussion following (3.14), $V_{\gamma_1,\dots,\gamma_N,\pm}$ is an N-soliton KdV-solution relative to the background KdV-solution V_0 in (2.28) iff $\gamma_{j,\pm} \in (0,\infty)$, $1 \le j \le N$ since then $H_{\gamma_1,\dots,\gamma_N,\pm}(t)$ has precisely N eigenvalues $\{\lambda_j\}_{j=1}^N$ in spectral gaps of $H_0(t)$ (see (A.35)–(A.37)). A similar comment applies to $\phi_{\gamma_1,\dots,\gamma_N,\pm}$ in the context of the mKdV-equation.

While our approach to insert eigenvalues into spectral gaps of $H_0(t)$ is based on Darboux-type transformations (i.e., commutation methods), a different technique, deforming the hyperelliptic curve K_g into singular curves, has been studied in [29], [42], [43].

Appendix A. Commutation Methods

In this appendix we describe two methods of inserting additional eigenvalues into spectral gaps of a given "background" Schrödinger operator $H_0 = -\frac{d^2}{dx^2} + V_0$. For details we refer to [18].

First we summarize the single commutation method [4–6]. Assume

$$\frac{d^m}{dx^m} V_0 \in C^\infty(\mathbb{R}) \cap L^\infty(\mathbb{R}) \text{ to be real-valued, } m \in \mathbb{N}_0 \tag{A.1}$$

and define in $L^2(\mathbb{R})$ the self-adjoint operator

$$H_0 = -\frac{d^2}{dx^2} + V_0,\ \mathcal{D}(H_0) = H^2(\mathbb{R}) \tag{A.2}$$

($H^m(\mathbb{R})$, $m \in \mathbb{N}$ the standard Sobolev spaces). Let

$$E_0 = \inf \sigma(H_0),\ \lambda_1 < \lambda_2 < \ldots < \lambda_N < E_0,\ \sigma_j \in [-1,1],\ 1 \leq j \leq N, \tag{A.3}$$

(where $\sigma(.)$ denotes the spectrum) and pick a

$$\begin{aligned} &0 < \psi_{0,\pm 1}(\lambda,.) \in C^\infty(\mathbb{R}),\ \psi_{0,\pm 1}(\lambda, x_0) = 1 \text{ for some } x_0 \in \mathbb{R},\ \lambda < E_0, \\ &\psi_{0,\pm 1}(\lambda,.) \in L^2((R, \pm\infty)),\ R \in \mathbb{R},\ \lambda < E_0 \end{aligned} \tag{A.4}$$

satisfying

$$H_0 \psi_{0,\pm 1}(\lambda) = \lambda \psi_{0,\pm 1}(\lambda),\ \lambda < E_0 \tag{A.5}$$

in the distributional sense. Define

$$A_{\sigma_j,\ldots,\sigma_N} = \frac{d}{dx} + \phi_{\sigma_j,\ldots,\sigma_N},\ \mathcal{D}(A_{\sigma_j,\ldots,\sigma_N}) = H^1(\mathbb{R}), \tag{A.6}$$

$$\begin{aligned} \phi_{\sigma_j,\ldots,\sigma_N}(x) &= \frac{d}{dx} \ln \psi_{0,\sigma_j,\ldots,\sigma_N}(x), \\ \psi_{0,\sigma_N}(x) &= \frac{1}{2}(1-\sigma_N)\psi_{0,-1}(\lambda_N, x) + \frac{1}{2}(1+\sigma_N)\psi_{0,+1}(\lambda_N, x), \end{aligned} \tag{A.7}$$

$$\begin{aligned} \psi_{0,\sigma_j,\ldots,\sigma_N}(x) &= \frac{1}{2}(1-\sigma_j)\psi_{0,-1,\sigma_{j+1},\ldots,\sigma_N}(\lambda_j, x) \\ &\quad + \frac{1}{2}(1+\sigma_j)\psi_{0,+1,\sigma_{j+1},\ldots,\sigma_N}(\lambda_j, x),\ 1 \leq j \leq N-1, \end{aligned} \tag{A.8}$$

$$\psi_{0,\pm 1,\sigma_{j+1},\ldots,\sigma_N}(\lambda, x) = \frac{\left(A^*_{\sigma_{j+1},\ldots,\sigma_N} \cdots A^*_{\sigma_{N-1},\sigma_N} A^*_{\sigma_N} \psi_{0,\pm 1}(\lambda)\right)(x)}{\left(A^*_{\sigma_{j+1},\ldots,\sigma_N} \cdots A^*_{\sigma_{N-1},\sigma_N} A^*_{\sigma_N} \psi_{0,\pm 1}(\lambda)\right)(x_0)}, \quad \lambda < \lambda_{j+1},\ 0 \leq j \leq N-1, \tag{A.9}$$

$$\begin{aligned} H_{0,\sigma_j,\ldots,\sigma_N} &= A^*_{\sigma_j,\ldots,\sigma_N} A_{\sigma_j,\ldots,\sigma_N} + \lambda_j,\ 1 \leq j \leq N \\ &= A_{\sigma_{j-1},\ldots,\sigma_N} A^*_{\sigma_{j-1},\ldots,\sigma_N} + \lambda_{j-1},\ 2 \leq j \leq N, \\ \mathcal{D}(H_{0,\sigma_j,\ldots,\sigma_N}) &= H^2(\mathbb{R}),\ 1 \leq j \leq N. \end{aligned} \tag{A.10}$$

Then $0 < \psi_{0,\pm 1,\sigma_j,\ldots,\sigma_N}(\lambda,.) \in C^\infty(\mathbb{R})$ are linearly independent principal solutions of

$$H_{0,\sigma_j,\ldots,\sigma_N}\psi(\lambda) = \lambda\psi(\lambda),\ \lambda < \lambda_j,\ 1 \leq j \leq N \tag{A.11}$$

and writing

$$H_{0,\sigma_j,\ldots,\sigma_N} = -\frac{d^2}{dx^2} + V_{0,\sigma_j,\ldots,\sigma_N} \tag{A.12}$$

one obtains

$$\frac{d^m}{dx^m}\phi_{\sigma_j,\ldots,\sigma_N},\ \frac{d^m}{dx^m}V_{0,\sigma_j,\ldots,\sigma_N} \in C^\infty(\mathbb{R}) \cap L^\infty(\mathbb{R}),\ m \in \mathbb{N}_0 \tag{A.13}$$

and

$$\begin{aligned} V_{0,\sigma_j,\ldots,\sigma_N}(x) &= \phi_{\sigma_j,\ldots,\sigma_N}(x)^2 - \phi'_{\sigma_j,\ldots,\sigma_N}(x) + \lambda_j \\ &= V_0(x) - 2\frac{d^2}{dx^2} \ln\left[\psi_{0,\sigma_j,\ldots,\sigma_N}(x) \ldots \psi_{0,\sigma_{N-1}\sigma_N}(x)\psi_{0,\sigma_N}(x)\right],\ 1 \leq j \leq N. \end{aligned} \tag{A.14}$$

Moreover,

$$\begin{aligned} &\psi^{-1}_{0,\sigma_j,\ldots,\sigma_N} \in H^2(\mathbb{R}),\ \left(H_{0,\sigma_j,\ldots,\sigma_N} - \lambda_j\right)\psi^{-1}_{0,\sigma_j,\ldots,\sigma_N} = 0, \\ &\qquad \sigma_j \in (-1,1),\ \sigma_\ell \in [-1,1], j+1 \leq \ell \leq N, \\ &\psi^{-1}_{0,\pm 1,\sigma_{j+1},\ldots,\sigma_N} \notin L^2(\mathbb{R}),\ \sigma_\ell \in [-1,1],\ j+1 \leq \ell \leq N,\ 1 \leq j \leq N-1 \end{aligned} \tag{A.15}$$

implying

$$\sigma\left(H_{0,\sigma_j,\ldots,\sigma_N}\right) = \begin{cases} \sigma(H_{0,\sigma_{j+1},\ldots,\sigma_N}),\ \sigma_j = \pm 1 \\ \sigma(H_{0,\sigma_{j+1},\ldots,\sigma_N}) \cup \{\lambda_j\},\ \sigma_j \in (-1,1),\ 1 \leq j \leq N, \end{cases} \tag{A.16}$$

$$\sigma_{\mathrm{ess}}(H_{0,\sigma_j,\ldots,\sigma_N}) = \sigma_{\mathrm{ess}}(H_0),$$

$$\lambda_\ell \in \sigma_d(H_{0,\sigma_j,\ldots,\sigma_N}) \text{ iff } \sigma_\ell \in (-1,1),\ j \leq \ell \leq N,\ 1 \leq j \leq N. \tag{A.17}$$

(Here $\sigma_{\mathrm{ess}}(.)$, and $\sigma_d(.)$ denote the essential and discrete spectrum respectively.) The total number of eigenvalues of $H_{0,\sigma_j,\ldots,\sigma_N}$ inserted into the spectral gap $(-\infty, E_0)$ by the single commutation procedure thus equals the number of $\sigma_\ell \in (-1,1)$, $j \leq \ell \leq N$. In addition one verifies the asymptotic behavior

$$V_{0,\sigma_j,\ldots,\sigma_N}(x) \underset{x\to\pm\infty}{=} V_{0,\epsilon_{j,\pm},\ldots,\epsilon_{N,\pm}}(x) + o\left(V_{0,\epsilon_{j,\pm},\ldots,\epsilon_{N,\pm}}(x)\right), \tag{A.18}$$

$$\phi_{\sigma_j,\ldots,\sigma_N}(x) \underset{x\to\pm\infty}{=} \phi_{\epsilon_{j,\pm},\ldots,\epsilon_{N,\pm}}(x) + o\left(\phi_{\epsilon_{j,\pm},\ldots,\epsilon_{N,\pm}}(x)\right). \tag{A.19}$$

where

$$\epsilon_{\ell,\pm} = \begin{cases} \mp 1, & \sigma_\ell \in (-1,1) \\ \sigma_\ell, & \sigma_\ell \in \{-1,1\},\ 1 \leq \ell \leq N. \end{cases}$$

Following [4–6] it is possible to rewrite (A.9) and (A.14) in terms of Wronskians as follows:

$$\psi_{0,\pm,\sigma_j,\dots,\sigma_N}(\lambda,x)=\frac{W\left(\Psi_{\sigma_j,\dots,}\Psi_{\sigma_N},\psi_{0,\pm1}(\lambda)\right)(x)W\left(\Psi_{\sigma_j,\dots,}\Psi_{\sigma_N}\right)(x_0)}{W\left(\Psi_{\sigma_j,\dots,}\Psi_{\sigma_N}\right)(x)W\left(\Psi_{\sigma_j,\dots,}\Psi_{\sigma_N},\psi_{0,\pm1}(\lambda)\right)(x_0)}, \tag{A.20}$$

$$V_{0,\sigma_j,\dots,\sigma_N}(x)=V_0(x)-2\frac{d^2}{dx^2}\ln W\left(\Psi_{\sigma_j,\dots,}\Psi_{\sigma_N}\right)(x),\ 1\le j\le N, \tag{A.21}$$

where $W(f_1,\dots,f_M)(x)$ denotes the Wronskian of $f_1,\dots,f_M$ and

$$\Psi_{\sigma_j}(x)=\frac{1}{2}(1-\sigma_j)N_{-1,\sigma_j}\psi_{0,-1}(\lambda_j,x)+\frac{1}{2}(1+\sigma_j)N_{+1,\sigma_j}\psi_{0,+1}(\lambda_j,x), \tag{A.22}$$

$$\begin{aligned}&N_{\pm1,\sigma_N}=1,\ \sigma_N\in[-1,1],\\ &N_{\pm1,\sigma_j}=\left[\left(A^*_{\sigma_{j+1},\dots,\sigma_N}\cdots A^*_{\sigma_{N-1},\sigma_N}A^*_{\sigma_N}\psi_{0,\pm1}(\lambda_j)\right)(x_0)\right]^{-1},\\ &\qquad\sigma_j\in[-1,1],1\le j\le N-1.\end{aligned} \tag{A.23}$$

Inserting (A.20)–(A.23) into the leading asymptotic terms in (A.18), (A.19) yields

$$V_{0,\epsilon_j,\dots,\epsilon_N}(x)=V_0(x)-2\frac{d^2}{dx^2}\ln W\left(\psi_{0,\epsilon_j}(\lambda_j),\dots,\psi_{0,\epsilon_N}(\lambda_N)\right)(x), \tag{A.24}$$

$$\phi_{\epsilon_j,\dots,\epsilon_N}(x)=\frac{d}{dx}\ln\left[\frac{W\left(\psi_{0,\epsilon_j}(\lambda_j),\dots,\psi_{0,\epsilon_N}(\lambda_N)\right)(x)}{W\left(\psi_{0,\epsilon_{j+1}}(\lambda_{j+1}),\dots,\psi_{0,\epsilon_N}(\lambda_N)\right)(x)}\right], \tag{A.25}$$

$$\epsilon_\ell\in\{-1,1\},\ 1\le\ell\le N.$$

(Here the denominator in (A.25) is replaced by 1 for $j=N$.)

Next we briefly summarize the double commutation method which (in contrast to the single commutation method) enables one to insert eigenvalues in any spectral gap of H_0 (not just the lowest one). In addition to assumption (A.1) let $E_0=\inf\sigma(H_0)$ as before and choose

$$\lambda_j\in\rho(H_0)\cap\mathbb{R},\ \gamma_{j,\pm}>0,\ \lambda_j\neq\lambda_\ell \text{ for } j\neq\ell,\ 1\le j,\ell\le N \tag{A.26}$$

($\rho(\cdot)$ the resolvent set) and

$$\psi_{0,\pm}(\lambda,.)\in C^\infty(\mathbb{R}),\ \psi_{0,\pm}(\lambda,.)\in L^2\left((R,\pm\infty)\right),\ \lambda\in\rho(H_0)\cap\mathbb{R},\ R\in\mathbb{R} \tag{A.27}$$

satisfying

$$H_0\psi_{0,\pm}(\lambda)=\lambda\psi_{0,\pm}(\lambda),\ \lambda\in\rho(H_0)\cap\mathbb{R} \tag{A.28}$$

in the distributional sense. We introduce

$$\psi_{0,j,\pm}(x)=\psi_{0,\pm}(\lambda_j,x),\ 1\le j\le N, \tag{A.29}$$

$$c_{\gamma_1,\dots,\gamma_N,\pm,\ell,m}(x)=\mp\gamma_{\ell,\pm}\int_{\pm\infty}^{x}dx'\psi_{0,\ell,\pm}(x')\psi_{0,m,\pm}(x'),\ 1\le\ell,m\le N, \tag{A.30}$$

$$\begin{aligned}&C_{\gamma_1,\dots,\gamma_N,\pm}(x)=\left[c_{\gamma_1,\dots,\gamma_N,\pm,\ell,m}(x)\right]_{\ell,m=1}^N,\\ &\qquad\Psi^0_{\gamma_1,\dots,\gamma_N,\pm}=(\gamma_{1,\pm}\psi_{0,1,\pm},\dots,\gamma_{N,\pm}\psi_{0,N,\pm})^T,\\ &\qquad\Psi_{\gamma_1,\dots,\gamma_N,\pm}=\left[1+C_{\gamma_1,\dots,\gamma_N,\pm}\right]^{-1}\Psi^0_{\gamma_1,\dots,\gamma_N,\pm},\\ &\qquad\Psi_{\gamma_1,\dots,\gamma_N,\pm}=\left(\Psi_{\gamma_1,\dots,\gamma_N,1,\pm},\dots,\Psi_{\gamma_1,\dots,\gamma_N,N,\pm}\right)^T\end{aligned} \tag{A.31, A.32}$$

(note that $\det\left[1+C_{\gamma_1,\dots,\gamma_N,\pm}(x)\right]>1,\ x\in\mathbb{R}$),

$$V_{\gamma_1,\dots,\gamma_N,\pm}(x)=V_0(x)-2\frac{d^2}{dx^2}\ln\det\left[1+C_{\gamma_1,\dots,\gamma_N,\pm}(x)\right]. \tag{A.33}$$

Then $V_{\gamma_1,\dots,\gamma_N,\pm}\in C^\infty(\mathbb{R})$ and, in order to avoid technicalities, we assume in addition $\frac{d^m}{dx^m}V_{\gamma_1,\dots,\gamma_N,\pm}\in L^\infty(\mathbb{R})$, $m\in\mathbb{N}_0$. (This assumption is easily verified in our concrete applications of Section 3.) Defining

$$\begin{aligned}&H_{\gamma_1,\dots,\gamma_N,\pm}=-\frac{d^2}{dx^2}+V_{\gamma_1,\dots,\gamma_N,\pm},\ \mathcal{D}\left(H_{\gamma_1,\dots,\gamma_N,\pm}\right)=H^2(\mathbb{R}),\\ &\qquad\qquad\gamma_{j,\pm}>0,\ 1\le j\le N\end{aligned} \tag{A.34}$$

one can then prove

$$\Psi_{\gamma_1,\dots,\gamma_N,j,\pm}\in H^2(\mathbb{R}),\ ||\Psi_{\gamma_1,\dots,\gamma_N,j,\pm}||_2^2=\gamma_{j,\pm}, \tag{A.35}$$

$$H_{\gamma_1,\dots,\gamma_N,\pm}\Psi_{\gamma_1,\dots,\gamma_N,j,\pm}=\lambda_j\Psi_{\gamma_1,\dots,\gamma_N,j,\pm},\quad\gamma_{j,\pm}>0,\quad 1\le j\le N \tag{A.36}$$

and

$$\rho(H_{\gamma_1,\dots,\gamma_N,\pm})=\rho(H_0)\backslash\{\lambda_j\}_{j=1}^N,\ \gamma_{j,\pm}>0,\ 1\le j\le N. \tag{A.37}$$

In order to describe the asymptotic behavior of $V_{\gamma_1,\dots,\gamma_N,\pm}(x)$ as $x\to\pm\infty$, we now lift the restriction $\gamma_{j,\pm}>0$ and formally admit the values 0 and ∞ for $\gamma_{j,\pm}$. This yields

$$\begin{aligned}&V_{\gamma_1,\dots,\gamma_N,+}(x)\underset{x\to\pm\infty}{=}V_{\substack{0,\dots,0,+\\ \infty,\dots,\infty,+}}(x)+o\left(V_{\substack{0,\dots,0,+\\ \infty,\dots,\infty,+}}(x)\right),\\ &V_{\gamma_1,\dots,\gamma_N,-}(x)\underset{x\to\pm\infty}{=}V_{\substack{\infty,\dots,\infty,-\\ 0,\dots,0,-}}(x)+o\left(V_{\substack{\infty,\dots,\infty,-\\ 0,\dots,0,-}}(x)\right),\end{aligned} \tag{A.38}$$

where

$$V_{0,\dots,0,\pm}(x) = V_0(x), \tag{A.39}$$

$$V_{\infty,\dots,\infty,\pm}(x) = V_0(x) - 2\frac{d^2}{dx^2}\ln\det\left[C_{N,\infty,\pm}(x)\right], \tag{A.40}$$

$$C_{N,\infty,\pm}(x) = \left[\mp\int_{\pm\infty}^{x} dx'\psi_{0,\ell,\pm}(x')\psi_{0,m,\pm}(x')\right]_{\ell,m=1}^{N} \tag{A.41}$$

Next, defining

$$\Psi_{\gamma_1,\dots,\gamma_N,\pm}(\lambda,x) = \psi_{0,\pm}(\lambda,x) \pm \sum_{j=1}^{N}\Psi_{\gamma_1,\dots,\gamma_N,j,\pm}(x)\int_{\pm\infty}^{x} dx'\psi_{0,j,\pm}(x')\psi_{0,\pm}(\lambda,x'),$$
$$\gamma_j > 0,\ 1\le j\le N,\ \lambda\in\rho(H_0)\cap\mathbb{R} \tag{A.42}$$

one infers

$$H_{\gamma_1,\dots,\gamma_N,\pm}\Psi_{\gamma_1,\dots,\gamma_N,\pm}(\lambda) = \lambda\Psi_{\gamma_1,\dots,\gamma_N,\pm}(\lambda),\ \gamma_j>0,\ 1\le j\le N,\ \lambda\in\rho(H_0)\cap\mathbb{R} \tag{A.43}$$

in the distributional sense and

$$\Psi_{\gamma_1,\dots,\gamma_N,\pm}(\lambda_\ell,x) = \gamma_{\ell,\pm}^{-1}\Psi_{\gamma_1,\dots,\gamma_N,\ell,\pm}(x),\ 1\le\ell\le N. \tag{A.44}$$

Assuming in addition that

$$H_{\gamma_1,\dots,\gamma_N,\pm}\ge 0 \tag{A.45}$$

we may write

$$V_{\gamma_1,\dots,\gamma_N,\pm}(x) = \phi_{\gamma_1,\dots,\gamma_N,\pm}(x)^2 + \phi'_{\gamma_1,\dots,\gamma_N,\pm}(x), \tag{A.46}$$

where

$$\phi_{\gamma_1,\dots,\gamma_N,\pm}(x) = \frac{d}{dx}\ln\Psi_{\gamma_1,\dots,\gamma_N,\pm}(0,x),\ \gamma_j>0, 1\le j\le N \tag{A.47}$$

and hence

$$\frac{d^m}{dx^m}\phi_{\gamma_1,\dots,\gamma_N,\pm}\in C^\infty(\mathbb{R})\cap L^\infty(\mathbb{R}),\ m\in\mathbb{N}_0. \tag{A.48}$$

Thus one obtains

$$\phi_{\gamma_1,\dots,\gamma_N,+}(x) \underset{x\to\pm\infty}{=} \phi_{\substack{0,\dots,0,+\\ \infty,\dots,\infty,+}}(x) + o\left(\phi_{\substack{0,\dots,0,+\\ \infty,\dots,\infty,+}}(x)\right),$$
$$\phi_{\gamma_1,\dots,\gamma_N,-}(x) \underset{x\to\pm\infty}{=} \phi_{\substack{\infty,\dots,\infty,-\\ 0,\dots,0,-}}(x) + o\left(\phi_{\substack{\infty,\dots,\infty,-\\ 0,\dots,0,-}}(x)\right), \tag{A.49}$$

where

$$\phi_{0,\dots,0,\pm}(x) = \frac{d}{dx}\ln\psi_{0,\pm}(0,x), \tag{A.50}$$

$$\phi_{\infty,\dots,\infty,\pm}(x) = \frac{d}{dx}\ln\Psi_{\infty,\dots,\infty,\pm}(0,x). \tag{A.51}$$

Here

$$\Psi_{\infty,\dots,\infty,\pm}(\lambda,x) = \psi_{0,\pm}(\lambda,x) \pm \sum_{j=1}^{N}\Psi_{\infty,\dots,\infty,j,\pm}(x)\int_{\pm\infty}^{x} dx'\psi_{0,j,\pm}(x')\psi_{0,\pm}(\lambda,x'),$$
$$\lambda\in\rho(H_0)\cap\mathbb{R}, \tag{A.52}$$

$$\begin{aligned}&\Psi_{\infty,\dots,\infty,\pm}(x) = C_{N,\infty,\pm}^{-1}(x)\Psi_{\infty,\pm}^0(x),\\ &\Psi_{\infty,\pm}^0(x) = (\psi_{0,1,\pm}(x),\dots,\psi_{0,N,\pm}(x))^T,\\ &\Psi_{\infty,\dots,\infty,\pm}(x) = (\Psi_{\infty,\dots,\infty,1,\pm}(x),\dots,\Psi_{\infty,\dots,\infty,N,\pm}(x))^T\end{aligned} \tag{A.53}$$

and

$$H_{\infty,\dots,\infty,\pm}\Psi_{\infty,\dots,\infty,\pm}(\lambda) = \lambda\Psi_{\infty,\dots,\infty,\pm}(\lambda) \tag{A.54}$$

in the distributional sense, where

$$H_{\infty,\dots,\infty,\pm} = -\frac{d^2}{dx^2} + V_{\infty,\dots,\infty,\pm},\quad \mathcal{D}(H_{\infty,\dots,\infty,\pm}) = H^2(\mathbb{R}). \tag{A.55}$$

We emphasize that the methods described in this appendix not only apply to the quasi-periodic backgrounds V_0 discussed in Sections 2 and 3 but actually work under very general circumstances.

Single commutation methods such as (A.20)–(A.25) have a long history that dates back at least to Darboux. More recent accounts are [4–6], [36], [37]. Double commutation formulas (in the special case of constant or periodic backgrounds V_0) such as (A.30)–(A.37) originated with Kay and Moses [23] and were extensively used in the seminal paper of Gardner, Greene, Kruskal and Miura [16]. More recent applications can be found e.g. in [3], [9–12], [14], [26], [28–33] and in connections

with Bäcklund transformations and solitons for the KdV-equation in [1], [2], [8], [13], [15], [19], [25], [29], [33], [39], [40].

Appendix B. τ-Functions and Bäcklund Transformations

In this appendix we study auto-Bäcklund transformations of the KdV-equation and connections between KdV and mKdV-solutions related to each other by Miura's transformation [34]. Again we refer to [18] for details.
Throughout this appendix we assume V to be of the type

$$V(t,x) = C - 2\partial_x^2 \ln \tau(t,x),\ C \in \mathbb{C}, \tau :\ \mathbb{C}^2 \to \mathbb{C} \text{ holomorphic.} \tag{B.1}$$

Suppose $V_2(t,x)$ satisfies (B.1) (for some τ_2) and define the (possibly singular) differential expressions

$$A(t) = \partial_x + \phi(t,x),\ A(t)^+ = -\partial_x + \phi(t,x), \tag{B.2}$$

$$H_2(t) = A(t)A(t)^+ = -\partial_x^2 + V_2(t,x), \tag{B.3}$$

$$H_1(t) = A(t)^+ A(t) = -\partial_x^2 + V_1(t,x), \tag{B.4}$$

$$V_j(t,x) = \phi(t,x)^2 + (-1)^j \phi_x(t,x),\ j = 1,2, \tag{B.5}$$

$$\phi(t,x) = \partial_x \ln \psi_2(t,x), \tag{B.6}$$

where ψ_2 is a fixed solution of

$$(H_2(t)\psi(t))(x) = 0 \tag{B.7}$$

of the type

$$\psi_2(t,x) = Fe^{Dx+Et}\tau_1(t,x)/\tau_2(t,x), \tag{B.8}$$

where $\tau_j :\ \mathbb{C}^2 \to,\ j = 1,2$ are holomorphic and D, E, F are appropriate constants. We also introduce

$$B_V(t) = -4\partial_x^3 + 6V(t,x)\partial_x + 3V_x(t,x), \tag{B.9}$$

where V satisfies (B.1) and

$$H_0(t) = H_2(t) + \lambda_0 = -\partial_x^2 + V_0(t,x), \tag{B.10}$$

$$V_0(t,x) = V_2(t,x) + \lambda_0 = \phi(t,x)^2 + \phi_x(t,x) + \lambda_0, \tag{B.11}$$

$$\tilde{H}_0(t) = H_1(t) + \lambda_0 = -\partial_x^2 + \tilde{V}_0(t,x), \tag{B.12}$$

$$\tilde{V}_0(t,x) = V_1(t,x) + \lambda_0 = \phi(t,x)^2 - \phi_x(t,x) + \lambda_0 \tag{B.13}$$

for some $\lambda_0 \in \mathbb{C}$. Then

$$\frac{d}{dt}H_0 - [B_{V_0}, H_0] = V_{0,t} - 6V_0V_{0,x} + V_{0,xxx} = \mathrm{KdV}(V_0) \tag{B.14}$$

applied to ψ_2 yields

$$\psi_2\ \mathrm{KdV}(V_0) = -(H_0 - \lambda_0)(\partial_t - B_{V_0})\psi_2. \tag{B.15}$$

Thus

$$\mathrm{KdV}(V_0) = 0 \text{ iff } (H_0 - \lambda_0)(\partial_t - B_{V_0})\psi_2 = 0. \tag{B.16}$$

Next we assume

$$(H_0 - \lambda_0)(\partial_t - B_{V_0})\psi_2 = 0, \tag{B.17}$$

or equivalently,

$$(\partial_t - B_{V_0}(t))\,\psi_2(t,x) = \alpha_2(t)\psi_2(t,x) + \beta_2(t)\psi_2(t,x)\int_{y_0}^{x} dx'\psi_2(t,x')^{-2}, \tag{B.18}$$

$$\alpha_2, \beta_2 \in C^\infty(\mathbb{R}),\ y_0 \in \mathbb{R}.$$

Since

$$(H_0 - \lambda_0)\psi_2 = 0 \text{ implies } (\tilde{H}_0 - \lambda_0)\psi_2^{-1} = 0, \tag{B.19}$$

one infers from

$$(\tilde{H}_0 - \lambda_0)(\partial_t - B_{\tilde{V}_0})\psi_2^{-1} = -4\beta_2\psi_2^{-4}\psi_{2,x} \tag{B.20}$$

that

$$\mathrm{KdV}(\tilde{V}_0) = 0 \text{ iff } \beta_2 = 0 \text{ in (B.18) or } V_0 = \tilde{V}_0 = \lambda_0 \tag{B.21}$$

(in analogy to (B.16)). Introducing

$$\begin{aligned} \mathrm{mKdV}(\phi;\mu) &= \phi_t - 6(\phi^2 + \mu)\phi_x + \phi_{xxx}, \\ \mathrm{mKdV}(\phi) &= \mathrm{mKdV}(\phi;0) = \phi_t - 6\phi^2\phi_x + \phi_{xxx} \end{aligned} \tag{B.22}$$

and noting Miura's identities [34]

$$\begin{aligned} \mathrm{KdV}(V_0) &= (2\phi + \partial_x)\ \mathrm{mKdV}(\phi;\lambda_0), \\ \mathrm{KdV}(\tilde{V}_0) &= (2\phi - \partial_x)\ \mathrm{mKdV}(\phi;\lambda_0) \end{aligned} \tag{B.23}$$

one obtains

$$\mathrm{mKdV}(\phi;\lambda_0)=0 \text{ iff } \beta_2=0 \text{ in (B.18) or } V_0=\tilde{V}_0=\lambda_0=\phi^2. \tag{B.24}$$

Finally, invoking the τ-function formalism [21], [38] in (B.8) (and (B.1) for $V=V_2,\ \tau=\tau_2$) one verifies

$$V_0(t,x)=\lambda_0+C-2\partial_x^2\ln\tau_2(t,x), \tag{B.25}$$

$$\tilde{V}_0(t,x)=\lambda_0+C-2\partial_x^2\ln\tau_1(t,x), \tag{B.26}$$

$$\phi(t,x)=D+\partial_x\ln[\tau_1(t,x)/\tau_2(t,x)], \tag{B.27}$$

with

$$C-D^2=2D\tau_1^{-1}\tau_{1,x}-2D\tau_2^{-1}\tau_{2,x}-2\tau_1^{-1}\tau_2^{-1}\tau_{1,x}\tau_{2,x}+\tau_1^{-1}\tau_{1,xx}-\tau_2^{-1}\tau_{2,xx}. \tag{B.28}$$

Nonsingular solutions $V_0,\tilde{V}_0$ and ϕ are obtained if in addition $\psi_2>0$. However, the above formalism applies to all known classes of singular (m)KdV-solutions [18]. Moreover, the approach in this appendix is not limited to the (m)KdV-context but has been successfully applied to the Gelfand-Dikii and Drinfeld-Sokolov hierarchy in [20].

Acknowledgements

I would like to thank R. Svirsky for collaboration, J. Zagrodziński for discussions, and P. Garbaczewski and Z. Popowicz for their kind invitation to a most stimulating conference. I am indebted to W. Karwowski and J. Tafel for their great hospitality in Wroclaw and Warsaw.

Financial support from the Faculty International Travel Fund of the University of Missouri at Columbia is gratefully acknowledged.

References

1. M. Adler, J. Moser, Commun. Math. Phys. **61**, 1–30 (1978).
2. R. F. Bikbaev, R. A. Sharipov, Theoret. Math. Phys. **78**, 244–252 (1989).
3. M. Buys, A. Finkel, J. Diff. Eqs. **55**, 257–275 (1984).
4. M. M. Crum, Quart. J. Math. Oxford (2) **6**, 121–127 (1955).
5. P. A. Deift, Duke Math. J. **45**, 267–310 (1978).
6. P. Deift, E. Trubowitz, Commun. Pure Appl. Math. **32**, 121–251 (1979).

7. B. A. Dubrovin, Russ. Math. Surv. **36:2**, 11–92 (1981).
8. B. A. Dubrovin, V. B. Matveev, S. P. Novikov, Russ. Math. Surv. **31:1**, 59–146 (1976).
9. F. Ehlers, H. Knörrer, Comment. Math. Helvetici **57**, 1–10 (1982).
10. N.M. Ercolani, H. Flaschka, Phil. Trans. R. Soc. London **A315**, 405–422 (1985).
11. N. Ercolani, H. P. McKean, Invent. math. **99**, 483–544 (1990).
12. A. Finkel, E. Isaacson, E. Trubowitz, SIAM J. Math. Anal. **18**, 46–53 (1987).
13. N. E. Firsova, Math. USSR Sbornik **63**, 257–265 (1989).
14. H. Flaschka, D. W. McLaughlin, in Lecture Notes in Math. 515, R. M. Miura (ed.), Springer, Berlin, 1976, p. 253–295.
15. H. Flaschka, A.C. Newell, Physica **3D**, 203–221 (1981).
16. C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura, Commun. Pure Appl. Math. **27**, 97–133 (1974).
17. F. Gesztesy, in "Ideas and Methods in Mathematical Physics. In Memory of Raphael Høegh-Krohn", S. Albeverio, J. E. Fenstad, H. Holden, T. Lindstrøm (eds.), Cambridge Univ. Press.
18. F. Gesztesy, R. Svirsky, in preparation.
19. F. Gesztesy, W. Schweiger, B. Simon, Trans. Amer. Math. Soc. (1991).
20. F. Gesztesy, D. Race, R. Weikard, in preparation.
21. R. Hirota, in "Solitons", R. K. Bullough, P. J. Caudrey (eds.), Springer, Berlin, 1980, p. 157–176.
22. A. R. Its, V. B. Matveev, Theoret. Math. Phys. **23**, 343–355 (1975).
23. J. Kay, H.E. Moses, J. Appl. Phys. **27**, 1503–1508 (1956).
24. I.M. Krichever, Funct. Anal. Appl. **9**, 161–163 (1975).
25. E.A. Kuznetsov, A.V. Mikhailov, Sov. Phys. JETP **40**, 855–859 (1975).
26. B.M. Levitan, "Inverse Sturm-Liouville Problems", VNU Science Press, Utrecht, 1987.
27. L. Markus, R. A. Moore, Acta Math. **96**, 99–123 (1956).
28. H. P. McKean, In Lecture Notes in Math. **755**, M. Grmela, J. E. Marsden (eds.), Springer, Berlin, 1979, p. 83–200.
29. H. P. McKean, in "Partial Differential Equations and Geometry", C. I. Byrnes (ed.), Marcel Dekker, New York, 1979, p. 237–254.
30. H. P. McKean, Commun. Pure Appl. Math. **38**, 669–678 (1985).
31. H. P. McKean, Rev. Mat. Iberoamericana **2**, 235–261 (1986).

32. H. P. McKean, J. Stat. Phys. **46**, 1115–1143 (1987).
33. H. P. McKean, P. van Moerbeke, Invent. Math. **30**, 217–274 (1975).
34. R. M. Miura, J. Math. Phys. **9**, 1202–1204 (1968).
35. S. Novikov, S. V. Manakov, L. P. Pitaevskii, V. E. Zakharov, "Theory of Solitons", Consultants Bureau, New York, 1984.
36. A. V. Rybin, M. A. Sall', Theoret. Math. Phys. **63**, 545–550 (1985).
37. J. Satsuma, J. Phys. Soc. Japan **46**, 359–360 (1979).
38. G. Segal, G. Wilson, Publ. Math. IHES **61**, 5–65 (1985).
39. R. A. Sharipov, Russ. Math. Surv. **41:5**, 165–166 (1986).
40. R. A. Sharipov, Sov. Phys. Dokl. **32**, 121–123 (1987).
41. H. D. Wahlquist, in Lecture Notes in Math. **515**, R. M. Miura (ed.), Springer, Berlin, 1976, p. 162–183.
42. J. Zagrodziński, J. Phys. **A17**, 3315–3320 (1984).
43. J. Zagrodziński, Lecture Series at the Dept. of Physics of the Bharathidasan University, Tiruchirapalli, India, 1989.

DISPERSION EQUATION TECHNIQUE FOR PERIODIC SOLUTIONS OF NLPDE'S.

J. Zagrodziński

Institute of Physics, Polish Academy of Sciences 02-668 Warsaw, Poland.

ABSTRACT

The dispersion equation for the quasi-periodic solutions of soliton type nlpde's is considered. Introducing so-called T-functions the quasi-periodic, soliton and mixed solutions can be included in this formalism. There are discussed also the soliton limit of periodic solutions and the extremely effective Burnside - Baker technique for the determination of the high genus periodic solution parameters.

There are three reasons for this contribution. The first is to show that there exists an alternative and sometimes more convenient method to determine the parameters of the quasi-periodic solution of soliton type NLPDE's than the classical technique of integration along the canonical contours. This approach, by the arguments which will be obvious below, we call the dispersion equation (d-e) technique[7,9], although in the Russian terminology it is known as "effectivization of the solution" [3]. The second one is to demonstrate that this technique is quite the same in both cases of periodic and soliton - type solutions. Hence it has a common root with Hirota direct methods [5].On the other hand, also the multisoliton solutions on the background of multiphase periodic processes can be simply incorporated in this formalism . And finally, that due to the rediscovered recently the Burnside-Poincare series [1-4], the discussed dispersion equation technique becomes particularly convenient for the numerical evaluation, since the parameters of soliton solution make the first terms of the series which determine the parameters of the periodic solution. This is strongly connected with the uniformization of the Riemann surface determined by NLPDE.

Let us consider the functions g and f: $\mathbb{C}^g \to \mathbb{C}$ having the following factorization property [7-9]

$$g(x+y)\ f(x-y) = \sum_{\varepsilon} Y(y,\varepsilon)\ X(x,\varepsilon), \tag{1}$$

where the sum is over the finite subset $S^g \in \mathbb{Z}^g$ and the functions X and Y map $\mathbb{C}^g \otimes S^g \to \mathbb{C}$. This property will be called the addition property (ADDp) and if $f \neq g$ we will speak about mutual ADDp, otherwise about self ADDp, (sADDp). The point is that for the function f having sADDp one can calculate easily the derivatives with

$$\left[\ln \frac{f(x+y)}{f(x-y)}\right] = \sum_{\varepsilon} Y(y,\varepsilon)_{y_i} \frac{X(x,\varepsilon)}{f(x+y)\ f(x-y)}, \tag{2}$$

$$\left[\ln f(x)\right]_{x_i x_j} = \frac{1}{2} \sum_{\varepsilon} Y(0,\varepsilon)_{y_i y_j} \frac{X(x,\varepsilon)}{f(x)^2}, \tag{3}$$

where $[.]_{x_i}$ denotes the derivative with respect to x_i.

The higher order derivatives coincide with the second hierarchy of KdV equation. Indeed the fourth order operator leads directly to KdV equation

$$L_{ijkl} + 2(L_{ij}L_{kl} + L_{ik}L_{jl} + L_{il}L_{jk}) = \sum_{\varepsilon} Y(0,\varepsilon)_{ijkl} \frac{X(x,\varepsilon)}{f(x)^2}, \tag{4}$$

where by $(.)_i$ we denote the derivative with respect to the i-th component of argument and $L := \ln f(x)$.

The reader will recognize on the l.h.s. of (4) the operator which appears during the integration of KdV equation and hence a close relation to the Hirota technique. The approach we present here gives rise to some advantages. It is not confined to the soliton solutions, but it includes periodic solutions and also solution in a form of multisoliton processes on the background of multiphase quasi-periodic solution which will be called the mixed solutions.

The sixth order operator leads to the Kotera - Sawada equation

$$\underset{15\ \text{permutations}}{L_{ijklmn} + 2(L_{ij}L_{klmn} + \ldots)} + \underset{15\ \text{permutations}}{4(L_{ij}L_{kl}L_{mn} + \ldots)} = \sum_{\varepsilon} Y(0,\varepsilon)_{ijklmn} \frac{X(x,\varepsilon)}{f(x)^2}, \tag{5}$$

Thus we see that by invoking such elementary operation as differentiation one can deduce some hierarchy of NLPDE's, since all the equations (2-5) are the direct consequences of the sADDp (1). The question about the classes of the functions having this property is natural.

The fundamental class of the functions fulfilling sADDp are the Riemann theta functions defined as

$$\Theta(z|B) := \sum_{n \in \mathbb{Z}^g} \exp i\pi\ [2\langle z,n\rangle + \langle n,Bn\rangle], \tag{6}$$

where $B \in \mathbb{C}^{g\times g}$ represents the Riemannian matrix, (symmetric with positively defined imaginary part), $\langle z,n\rangle := \sum z_i n_i$ and argument $z \in \mathbb{C}^g$, [3,6,7,9].

Then sADDp can expressed either as

$$\Theta(x+y|B)\,\Theta(x-y|B) = 2^{-g} \sum_{2\varepsilon \in (\mathbb{Z}_2)^g} \Theta(2y+\varepsilon|2B)\ \Theta(2x+\varepsilon|2B), \tag{7}$$

or as

$$\Theta(x+y|B)\,\Theta(x-y|B) = \sum_{2\varepsilon \in (\mathbb{Z}_2)^g} Y(y,\varepsilon)\ \Theta(x+\varepsilon|B)^2, \tag{8}$$

where

$$Y(y,\varepsilon) := 2^{-g} \sum_{2\mu \in (\mathbb{Z}_2)^g} (-1)^{2\langle \varepsilon,\mu\rangle} \exp(i\pi\langle y,\mu\rangle)\ \Theta(2y+B\mu|2B)/\Theta(B\mu|2B), \tag{9}$$

The application of the reported relations for the most popular NLPDE's one can find in the literature, but below, in connection with the paper[5] we will illustrate the d-e technique by the example of the modified Korteweg de Vries (mKdV) equation.

We consider mKdV equation

$$v_t + 2(v^3)_x + v_{xxx} = 0, \tag{10}$$

which after the substitution of the ansatz $v = i\,[\ln(g/f)]_x$ reduces to

$$(g_t f - g f_t + g_{xxx} f - g f_{xxx} - 3g_{xx} f_x + 3g_x f_{xx}) - 3[\ln(g/f)]_x\,(g_{xx} f - 2g_x f_x + g f_{xx}) = C_1 gf, \tag{11}$$

where C_1 is the integration constant. We assume that both functions g and f depend on space and time in multilinear manner, i.e.

$$g(x,t) := g(z_1,z_2,\ldots,z_n), \quad \text{where } z_i = (\kappa_i x + \omega_i t + z_{ig}), \quad i=1,\ldots,g, \tag{12}$$

where z_{ig} are constant, and similarly for f(x,t). The mKdV equation will have the solution expressed in terms of the functions having mADDp if the system of algebraic equations

$$\sum_{i,j} \kappa_i \kappa_j Y(0,\varepsilon)_{ij} = C_2, \tag{13a}$$

$$\sum_i (\omega_i - 3C_2\kappa_i)Y(0,\varepsilon)_i \; - \sum_{i,j,k} \kappa_i\kappa_j\kappa_k Y(0,\varepsilon)_{ijk} = C_1 Y(0,\varepsilon) \; , \qquad (13b)$$

holds for any ε belonging to the set which is determined by validity of (1). Equations (13) are derived by the multiple differentiation of (1), in the assumption that $X(z,\varepsilon)$ are linearly independent.

The striking affinity with the Hirota formulas,[6]

$$(D_t + D_x^3)\, g\circ f = 0 \; , \quad D_x^2\, g\circ f = 0 \; , \qquad (14)$$

is obvious, although for the "pure" soliton solutions C_1 and C_2 were put to be zero; D represents the Hirota bilinear operator. In order to have the real solutions, g is chosen usually as $g = f^*$, (i.e. as c.c. of f).

For multi-phase periodic solutions $f = \Theta(z\text{-}s\,|\,B)$ and $g = \Theta(z+s\,|\,B) = \Theta(z\text{-}s\,|\,B)^*$. Equations (13) can be rewritten in more convenient form using instead of Y-coefficients so-called f-coefficient defined as

$$f_{\tau,i\ldots k}(s) := [\Theta(2s+B\tau\,|\,2B)\, \exp(i2\pi\langle s,\tau\rangle)]_{s_i\ldots s_k} \; , \quad \tau\in(\mathbb{Z}_2)^g \; . \qquad (15)$$

Then the system (13) becomes

$$\sum_{i,j} \kappa_i\kappa_j f(s)_{\tau,ij} = C_2 \; ,$$
$$\sum_i (\omega_i - 3C_2\kappa_i) f(s)_{\tau,i} \; - \sum_{i,j,k} \kappa_i\kappa_j\kappa_k f(s)_{\tau,ijk} = C_1 f(s)_\tau \; , \qquad (16)$$

for any $\tau\in(\mathbb{Z}_2)^g$.

In the simplest case of g=1, i.e. in the case of cnoidal waves, κ and ω are determined from equations

$$f_0''(s) = f_1''(s) \; , \qquad (17)$$
$$(\kappa\omega - 3f_0'')\,(f_0'f_1 - f_0f_1') = \kappa^3\,(f_0''f_1 - f_0f_1'') \; , \qquad (18)$$

provided that C_1 and C_2 do not vanish and all derivatives are taken at point s.

The two-phase quasi-periodic solution related to the matrix $B\in\mathbb{C}^{2x2}$ can be obtained in the similar manner.

For g>2 the system (16) is overdetermined what means that there exist some constrains on the matrix B, which here is a parameter. In the finite-gap integration formalism this fact expresses the restrictions imposed on B-matrix to be the period matrix of the suitably chosen Riemann surface,[3,7,10].

In order to link the periodic solution formalism with soliton type solutions it is natural to consider so called T-functions, being the objects which can be reduced either to the Riemann theta function or to the exponential functions appearing in the soliton solutions. The intermediate representations of T-function are suitable for the solutions in a form of multi-solitons on the background of multi-phase quasi periodic processes.

Let the Riemannian matrix $B\in\mathbb{C}^{gxg}$ be decomposed in blocks B^{ss}, B^{pp}, B^{sp}, $B^{ps}=(B^{sp})^t$, belonging to $\mathbb{C}^{sxs}$, $\mathbb{C}^{pxp}$, $\mathbb{C}^{pxs}$, $\mathbb{C}^{sxp}$, respectively; s+p=g. Similarly let argument z be decomposed into $z^s\in\mathbb{C}^s$ and $z^p\in\mathbb{C}^p$. Denoting $D^{ss} := \mathrm{Im\ Diag}\ B^{ss}$, we define T-function as the s-fold limit of the Riemann theta function

$$T(z^s,z^p) := \lim_{d\to\infty} \Theta\left[\begin{matrix} z^s - D^{ss}e^s \\ z^p \end{matrix}\,\middle|\, \begin{matrix} B^{ss} & B^{sp} \\ B^{ps} & B^{pp} \end{matrix}\right] \; , \qquad (19)$$

when all s elements of diagonal matrix D^{ss} tend to infinity, (e^s denotes the unit vector). It is easy to see that the limit exists and

$$T(z^s,z^p) = \sum_{n\in(\mathbb{Z}_2)^s} \exp\, i\pi[2\langle z^s,n\rangle + \langle n,\tilde{B}^{ss}n\rangle]\; \Theta(z + B^{ps}n\,|\,B^{pp}) \; , \qquad (20)$$

where $\tilde{B}^{ss} := B^{ss} - iD^{ss}$ and the sum contains 2^s terms.

It is obvious that if s=0, T-function coincides with Θ-function, but if p=0 then T-function reduces to the finite sum of exponential functions. The most important point is, however, that for T-functions defined according to (19) or (20), the relation (1) also holds. (The B-blocks of g and f must be of course the same). Thus T-functions have ADDp with all further consequences as relations (2-5), (10-13) and (16-18). The definition of f-coefficients looks now slightly different

$$f_\tau(w) := R(2w\,|\,\tau\,|\,2B)\, \exp(i2\pi\langle w,\tau\rangle) \; , \quad \tau\in(\mathbb{Z}_2)^g \; , \qquad (21)$$

where R is given by T-functions

$$R(2w\,|\,\tau\,|\,2B) := \sum_{(\mu^s)} \exp\{-i2\pi\langle\mu^s,[2w^s + \tilde{B}^{ss}(\tau^s-\mu^s) + B^{sp}\tau^p]\rangle\}$$
$$\times\; \Theta(2w^p + B^{pp}\tau^p + B^{ps}(\tau^s - 2\mu^s)\,|\,2B^{pp}) \; . \qquad (22)$$

Here the argument w, indices τ,μ and matrix B are also decomposed as above. The symbol (μ^s) denotes the sum over such $\mu^s\in(\mathbb{Z}_2)^s$ that also $(\tau^s\text{-}\mu^s)\in(\mathbb{Z}_2)^s$. In particular case p=0, i.e. in "pure" solitonic case R reduces to

$$R(2w^s\,|\,\tau^s\,|\,0) := \sum_{(\mu^s)} \exp\{-i2\pi\langle\mu^s,[2w^s + \tilde{B}^{ss}(\tau^s-\mu^s)]\rangle\} \; , \qquad (23)$$

and in opposite case, when s=0, (21) and (22) reproduce equation (15).

Putting $C_1 = C_2 = 0$, immediately we recover commonly known form of one-soliton solution [6]. Eqns.(13a) give $s=\pm 1/4$, eqns.(13b) prove the dispersion relation $\omega = \kappa^3$, and finally (20) determine g and f as $1 \pm i\exp[i2\pi(\kappa x+\omega t)]$, respectively, with κ imaginary to have the real solution v.

In order to write the analytical form for the simplest mixed solution (i.e. soliton on the background of periodic process) it is sufficient to put s=p=1 in (20). Then the solution is given once more by $v = i[\ln(g/f)]_x$, with g and f given by the functions $\Theta(z_2 \pm s_2 | b_{22}) + \exp[i2(z_1 \pm s_1)]\ \Theta(z_2 \pm s_2 + b_{12} | b_{22})$. The indices 1 and 2 refer to the soliton and periodic phases, respectively and b_{12} denotes off-diagonal element of the starting 2×2 B matrix which plays a role of the phase shift of the periodic wave. Indeed, since $z_1 \in i\mathbb{R}$ is linear in x, the asymptotic solutions when $x \to \pm\infty$ are given either by $i[\ln(\Theta(z_2+s_2|b_{22})/\Theta(z_2-s_2|b_{22})]_x$ or $i[\ln(\Theta(z_2+s_2+b_{12}|b_{22})/\Theta(z_2+s_2+b_{12}|b_{22})]_x$. Of course all parameters appearing in the above solutions follow from relevant version of the dispersion equations (16), or eventually can be determined according to the traditional finite gap integration, [10].

Hitherto we have discussed the mixed (and the soliton) solutions by the substitution the relevant ansatz with the suitable T-functions at once to the dispersion equation. One can however ask: whether and when the solution expressed in terms of the Riemann theta function has a limit expressed in terms of T-function, which we have called the mixed solution ?

The answer is not univocal. Two important points must be taken into account.

Firstly, the Riemannian B-matrix being a parameter of quasi-periodic solution is in general not arbitrary, but it must be the period matrix of a suitably chosen Riemann surface [10], what means that there are constraints imposed on its elements. Thus, the procedure suggested in (13) can lead to a matrix B which ceases to be a period-matrix and thus giving rise to the expression which is not a solution of initial equation.

The second point relates to interpretation of the soliton limit of the multi-phase periodic solution. Such limit is sometimes naively explained, without any restrictions, as a situation when the band structure of the spectrum determined by its end branch points reduces to the point spectrum - characteristic for the soliton solution.

Even in the case of cnoidal waves or two-phase solutions, (genus 1 or 2, respectively), where all the elements of B-matrix are arbitrary, there remains a problem of the correct choice of canonical contours and thus a choice of the cuts in the Riemann surface. In order to obtain the soliton solution as a limit (13), canonical contours must be chosen in a manner which enables two branch points coalesce into a single pole. We repeat here the example reported in [7]. In Fig.1. two choices of canonical contours are presented. The left configuration admits the

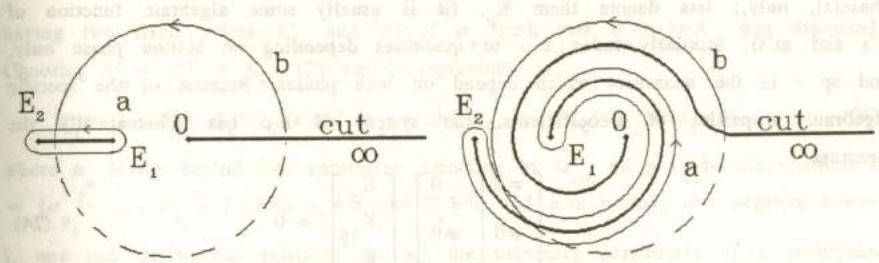


Fig.1.

soliton limit, in contrast to the right configuration, although the both represent the same solution. The transformation from one choice of canonical contours to another, for the fixed branch points, is governed by a modular transformation. This transformation is determined by the symplectic matrix $S \in \mathbb{Z}^{2g\times 2g}$, (detS=1), which transforms matrices of Abelian integrals (e.g. for the sG eqn.) [3,8,10]

$$H_{ij} = \oint_{a_i^+} dU_j\,, \qquad F_{ij} = \oint_{b_i} dU_j\,, \qquad dU_j = z^{j-1}\,[\prod_{k=0}^{2g}(z-z_k)]^{-1/2} \tag{24}$$

into new matrices F',H' of integrals over new contours a_i', b_i', the period matrix $B=FH^{-1}$ into new $B'=F'H'^{-1}$ and the argument z into z', as follows

$$\begin{bmatrix}F'\\H'\end{bmatrix} = \begin{bmatrix}S\end{bmatrix}\begin{bmatrix}F\\H\end{bmatrix}, \qquad S=\begin{bmatrix}a & b\\ c & d\end{bmatrix}, \qquad B'=(aB+b)(cB+d)^{-1}, \qquad z'=(cB+d)^{-1t}z. \tag{25}$$

The solution is expressed then by the theta functions of the "new" argument z' with B'-period matrix, which is related to the old one by

$$\Theta\begin{bmatrix}\alpha'\\ \beta'\end{bmatrix}(z'|B') = K\,\exp[i\pi\langle z,(cB+d)^{-1}cz\rangle]\ \Theta\,(z|B)\,, \tag{26}$$

where K is constant. Since z is linear in x and t, the Gaussian term in (25), at most gives rise to constant background in case of considered above mKdV. Thus, in the case of arbitrary choice of canonical contours, in order to obtain the correct

soliton limit, a suitable modular transformation should be performed as the indispensable step.

The last question we would like to elucidate in the context of d-e is the effect of the "dominance" of the periodic phases over the soliton ones in the mixed solutions. The effect follows the specific property of f-coefficients appearing in d-e. Let us consider the system of d-e as e.g (13). If the index p will be ascribed to the quantities which depend on the parameters of the periodic phase(s), only,; lets denote them K_p, (it is usually some algebraic function of κ's and ω's). Similarly index s - to quantities depending on soliton phase only, and sp - to the quantities which depend on both phases. Because of the specific algebraic properties of f-coefficients, the system of d-e has schematically the structure

$$\begin{bmatrix} \neq 0 & 0 \\ \neq 0 & \neq 0 \end{bmatrix} \begin{bmatrix} K_p \\ K_{sp} \\ K_s \end{bmatrix} = 0 \ . \qquad (24)$$

Thus the system splits out, and the quantities determining the quasi-periodic phase ($\kappa_p, \omega_p, B_{pp}$, etc) are completely independent of those related to the soliton one. The soliton and the coupling parameters, however, depend on the periodic phase. Hence the quasi-periodic phase turns out to be exactly the same as in the case when the soliton phase is absent, i.e. exactly the same as in the case when the quasi periodic process would be considered alone. This property, which we denote as the dominance of the periodic processes, coincides with the conclusion in [5], although it is in contradiction with the traditional approaches when usually the periodic part was considered as a small perturbation, ("phonon background").

From the applicative point of view, the formalism of d-e seems sometimes much more effective than the traditional Abelian integrals approach, particularly when the general form of solution is known and only the parameters of solution should be determined. Moreover, from the diagonal elements of the B-matrix one can easily evaluate how the periodic solution is close to the soliton one, (locally). These facts are well known to the people who apply numerical procedures. According to the author's experience, in any case when numerical results are essential, the d-e formalism gave (with genus not greater then 4) more legible results as a rule, than the approach starting from spectrum structure. This conclusion relates also to the situation when the system of d-es was overdetermined and approximate procedures had to be applied in order to determine its correct form. In case of mixed

solutions, the situation is much more favorable for d-e formalism.

Quite recently, the new and enormously fruitful for numerical treatment results have been achieved, or more precisely rediscovered and next generalized.

The fundamental results were obtained at the end of the last century by Burnside [4] next by Baker [1], and recently developed [3,4]. The three-parameter non-abelian group PSL(2,$\mathbb{C}$) generated by a homographic transformations

$$\sigma_n(z)=(\alpha_n z+\beta_n)(\gamma_n z+\delta_n)^{-1}, \qquad \alpha_n\beta_n-\gamma_n\delta_n=1. \qquad (25)$$

having two fixed points A'_n and A''_n of σ_n such that $\sigma_n(A_n)=A_n$ was discussed. Choosing $A'_n = -A''_n = A_n$, (25) can be represented as[9]

$$\sigma_n = (z\ \mathrm{ch}\xi_n + A_n \mathrm{sh}\xi_n)(z\ \mathrm{sh}\xi_n/A_n + \mathrm{ch}\xi_n), \qquad \mathrm{th}\xi_n=(\mu_n-1)(\mu_n+1), \qquad (26)$$

where μ_n is the second free parameter. Denoting by G_{mn} all possible combinations $G = \{\sigma_{n_1}^{j_1}, \ldots, \sigma_{n_k}^{j_k}\}$, with $n_1 \neq m$ and $n_k \neq n$, and with positive and negative powers j, one can derive the relations for all the necessary parameters of a multiphase periodic solution. Each parameter is expressed in the form of a series over that group, and what is the most important and surprising, the first terms in these series determine the soliton solutions.

This important for the modern soliton theory conclusion follows from the Schottky uniformization of the Riemann surface. There were analyzed the properties of the function $H(z,a):=\Sigma\ d\{\ln[\sigma_n(z)-a]\}/dz$, where the sum is over the discussed group. In general, H function transforms the interior of some circle into the exterior of another one. For the suitably chosen a, this function automatically determines the *normalized* Abelian integrals, leading directly to the period B-matrix of the Riemann surface. The details of the mathematical aspect can be found in the cited literature and here, in order to illustrate its role for the soliton theory and application, after [2,3] we report relations concerning KdV equation.

For the solution of KdV equation in form $u(x,t) = 2[\ln \Theta(z|B)]_{xx} + 2c$, where z is linear in x and t as in (12), and $\Theta(z|B)$ is given by (6), B, κ, ω and c can be determined from the series

$$i\pi B_{mn} = \sum_{\substack{\sigma\in G_{mn} \\ \sigma\neq I,\ \text{if } n=m}} \ln\{[A_m-\sigma(A_n)][A_m-\sigma(-A_n)]^{-1}\} = \ln\{[A_m-A_n][A_m+A_n)]^{-1}\} + \mathcal{O}(\mu), \qquad (27)$$

$$i2\pi\,\kappa_i = \sum_{\sigma\in G_{mn}} [\,\sigma(A_i)-\sigma(-A_i)\,] = 2A_i + \mathcal{O}(\mu), \tag{28}$$

$$i2\pi\,\omega_i = \sum_{\sigma\in G_{m}} [\sigma(A_i)^3-\sigma(-A_i)^3] = 2A_i^3 + \mathcal{O}(\mu), \tag{29}$$

$$c = \sum_{\substack{\sigma\in G\\ \sigma\neq I}} \gamma_i^{-2} = \mathcal{O}(\mu), \tag{30}$$

where $\mathcal{O}(\mu)$ denotes the terms of the order of μ (and higher), μ being the largest μ_n in (26), n = 1,2,...,g.

In the soliton limit ($\mu_n \to 0$) the leading terms in (27) - (30) represent just the multisoliton solution.

Thus the formalism developed one hundred years ago and quite forgotten later on, becomes now enormously fruitful in a computer era.

The author would like to thank Dr M. Jaworski for numerous enlightening discussions.

References.

1) Baker, H.F.,"Abel's theorem and the allied theory including the theory of theta functions", Cambridge University Press, (1897).

2) Bobenko, A.I. and Bordag, L.A., J. Phys. A 22, 1259 (1989).

3) Bobenko, A.I.,"Uniformization of Riemann surfaces and effectivization of theta functional formulae", Preprint 257 Tech. Univ. Berlin, (1990).

4) Burnside, W.,Proc. London Math. Soc. 23, 49 (1892).

5) Gesztesy, F., "Quasi-periodic solutions of mKdV eqn.", this proceedings.

6) Hirota, R., in "Solitons", eds R.Bullough, P.Caudrey, Springer,(1980); also in "Backlund Transformations", eds A.Dodd, B,Eckmann, Springer (1976).

7) Zagrodzinski, J., in "Syst. dynamique. non lineaires", ed. P. Winternitz (Univ. de Montreal, (1986); J. Phys. A 17, 3315 (1984).

8) Zagrodzinski, J., Phys. Rev. B 29, 1500 (1984).

9) Zagrodzinski, J., Jaworski, M. and Wyser, K., "From period. proces. to solitons and vice-versa", Proc.Conf. "Aspects of NL. Dynamics: Solitons and Chaos etc.", Brussel, Springer, (1990).

10) Zakharov, V.E., Manakov, S.V., Novikov, S.P. and Pitaevski, L.P., "Theory solitons", Plenum, (1984).

Gauge Transformations and Reciprocal Links for Integrable Equations

W. Oevel

Department of Mathematical Sciences
Loughborough University of Technology, LE11 3TU, U.K.

Abstract

Generalizing Sato's construction of the Kadomtsev-Petviashvili (KP) hierarchy we consider three classes of Lax equations on the algebra of (pseudo-) differential operators. The simplest realizations of these general classes are the hierarchies of the Korteweg-de Vries (KdV) equation, the modified KdV equation and the Harry Dym equation. Gauge transformations and reciprocal links of the Lax operators yield intimate relations between the classes. For second order Lax operators these transformations generate the Miura map between KdV and modified KdV as well as the known reciprocal link to the Dym equation. Similar results can be derived for integrable hierarchies associated with higher order Lax operators.

1 Introduction

The discovery of the celebrated Miura transformation linking the KdV and the modified KdV equations played a key role in the development of the Inverse Scattering Method [1] as a technique to solve integrable nonlinear partial differential equations. In fact, the linearization of the Riccati equation defining the Miura map immediately leads to the underlying linear problem for the KdV. On the other hand, invariances of the associated linear problems may be used to generate transformations leaving the corresponding nonlinear problems invariant. Typically, gauge transformations of the scattering operator may either change the explicit form of the linear problem –corresponding to different nonlinear equations associated to the different forms of the linear problem– or they will keep the linear problem invariant. The first case encodes a Miura type transform between the distinct sets of nonlinear equations, whereas the latter case represents auto-Bäcklund transformations. Various examples of such gauge transformations for scattering problems related to 1+1- and 2+1-dimensional integrable nonlinear hierarchies are given in [2]. A background account of the subject including recent results on coherent structures in 2+1 dimensions is to be found in a recent review article by Konopelchenko and Rogers [3].
Further, gauge transformations, in conjunction with reciprocal transformations, also play a key role in links between different scattering schemes. E.g., in [4] it was demonstrated that the AKNS and WKI scattering schemes are linked, when transformations of the independent variables are taken into account. Further, in the context of soliton theory, the Dym hierarchy as set down in [5] is known to be invariant under a reciprocal transformation. Reciprocal links between the Dym equation and KdV/modified KdV had been observed in the context of Painlevé analysis [6], similar results hold for their 2+1-dimensional counterparts, too [7].
We will demonstrate that gauge transformations and reciprocal links are natural tools to

describe Miura transformations, auto-Bäcklund transformations and reciprocal connections for the class of equations covered by Sato's famous construction of the KP hierarchy [8-10]. There the KP is constructed by generalized Lax equations of the type

$$L_t = [(L^q)_+, L] \ , \tag{1.1}$$

where L is a (pseudo-) differential operator and $(..)_+$ is the projection to the purely differential part of the operator. In [11] Adler gave a Lie algebraic interpretation for such equations, embedding integrable partial differential equations such as the KdV equation in a general Lie algebraic setting, now known as the celebrated Adler-Kostant-Symes (AKS) scheme. The AKS approach turned out to be a special case of an even more general algebraic structure: (classical) r-matrices. Following Drinfeld's ideas [12], Semenov showed how multi-Hamiltonian structures and dressing methods for integrable equations originate from general algebraic and group theoretic structures [13,14]. The AKS scheme is associated with the special r-matrices stemming from decompositions of the underlying Lie algebra. From these general ideas it is obvious that the Lax equations (1.1) will be integrable for any projection $(..)_+$ associated with a Lie algebra decomposition of the space of pseudo-differential symbols. It turns out that there are in fact three natural decompositions, leading to three classes of Lax equations (1.1), each with a slightly different meaning of the projection $(..)_+$. Such modifications of the original Lax equation were analyzed by Kupershmidt [15]. In the context of Sato's equations (1.1) these modified hierarchies have also been observed recently by Kiso [16].
A concise analysis of the general invariances and relating transformations for these three classes are set down in [17], a systematic analysis of 1+1-dimensional reductions, a survey of the related Hamiltonian concepts and a number of worked out examples is to be found in [18].
Here we want to give a simple introduction to these more general concepts by reviewing the meanwhile classical results for KdV, modified KdV and Dym equation in the light of generalized Sato equations (1.1).

2 KdV, modified KdV and Dym equation reviewed

A simple introduction to the more general results of [17] and [18] is given by the simplest realizations of the three classes of Sato's equations. In terms of Sato's construction these are the so-called 2-reductions associated with second order Lax operators, they represent the hierarchies of the KdV, the modified KdV and the Dym equation. We consider the operators

$$\begin{aligned}
&i)\ k=0: && L_{KdV} = \partial^2 + u && , && B_{KdV} = \partial^3 + \tfrac{3}{2}u\partial + \tfrac{3}{4}u_x \ , \\
&ii)\ k=1: && L_{mKdV} = \partial^2 + 2v\partial && , && B_{mKdV} = \partial^3 + 3v\partial^2 + \tfrac{3}{2}(v_x + v^2)\partial \ , \\
&iii)\ k=2: && L_{Dym} = w^2\partial^2 && , && B_{Dym} = w^3\partial^3 + \tfrac{1}{2}(w^3)_x\partial^2 \ .
\end{aligned} \tag{2.1}$$

The reason for the somewhat mysterious labeling $k = 0, 1, 2$ will become obvious soon. A simple calculation shows that these operators represent the Lax pairs for the equations

mentioned above, i.e. the Lax equation $L_t = [B, L]$ implies

$$\begin{aligned}
&i)\ k=0: && 4u_t = u_{xxx} + 6uu_x && \text{(Korteweg- de Vries)} \ , \\
&ii)\ k=1: && 4v_t = v_{xxx} - 6v^2v_x && \text{(modified Korteweg- de Vries)} \ , \\
&iii)\ k=2: && 4w_t = w^3w_{xxx} && \text{(Dym)} \ .
\end{aligned} \tag{2.2}$$

The crucial point for the following analysis is that in all cases the operators B can in fact be calculated directly from the given Lax operator L. For the KdV case this construction goes back to [19], the relevant procedure can be summarized as follows: Introducing a formal integration symbol ∂^{-1} we look at operators of the form

$$L^{1/2} = \partial + a_0 + a_1\partial^{-1} + a_2\partial^{-2} + a_3\partial^{-3} + \dots \ . \tag{2.3}$$

Here, the symbol ∂^{-1} is a formal inverse of the differential operator ∂, multiplication of this symbol with a function (=multiplication operator) is given by the generalized Leibnitz rule

$$\partial^{-1}a = a\partial^{-1} - a_x\partial^{-2} + a_{xx}\partial^{-3} - a_{xxx}\partial^{-4} + \dots \ . \tag{2.4}$$

Trying to turn the ansatz (2.3) into a formal square root of the Schrödinger operator $L = L_{KdV} = \partial^2 + u$, we identify the coefficients $a_0, a_1, a_2, ..$ by calculating

$$L^{1/2}L^{1/2} = \partial^2 + 2a_0\,\partial + (a_{0x} + a_0^2 + 2a_1) + (a_{1x} + 2a_0a_1 + 2a_2)\,\partial^{-1} + (..)\,\partial^{-2} + .. \tag{2.5}$$

using (2.4) and $\partial\partial^{-1} = \partial^{-1}\partial = 1$. Hence, requiring $(L^{1/2})^2 = \partial^2 + u$, we can calculate all the coefficients $a_0, a_1, a_2, ..$ recursively in terms of the field u. In particular, one finds the formal expansion

$$(\partial^2 + u)^{1/2} = \partial + \frac{1}{2}u\,\partial^{-1} - \frac{1}{4}u_x\,\partial^{-2} + \frac{1}{8}(u_{xx} - u^2)\,\partial^{-3} + (..)\,\partial^{-4} + \dots \ . \tag{2.6}$$

Calculating the 3rd power of (2.6) (or multiplying with $\partial^2 + u$) one finds the operator

$$(\partial^2 + u)^{3/2} = \partial^3 + \frac{3}{2}u\,\partial + \frac{3}{4}u_x + (..)\,\partial^{-1} + (..)\,\partial^{-2} + \dots \ , \tag{2.7}$$

which –by construction– commutes with $\partial^2 + u$. We observe that the operator B_{KdV} of (2.1) now can be read off the fractional power (2.7) as the purely differential part. In fact, the entire hierarchy of higher KdV equations can be obtained in the same fashion. Including further terms in the expansions (2.6) one can easily calculate $(\partial^2 + u)^{5/2}$ and consider its differential part. This operator will supplement $L = \partial^2 + u$ to form a Lax pair for the 5th order KdV flow $u_t = u_{xxxxx} + 10uu_{xxx} + 20u_xu_{xx} + 30u^2u_x$ commuting with the KdV.
A Lie algebraic interpretation of this construction was given by Adler [11], such that the KdV can be treated in the celebrated Adler-Kostant-Symes (AKS) scheme in the same fashion as many other integrable systems associated with Lie algebras (see e.g. [20] for a review and further references).
In a similar way one can extract the operators B_{mKdV} and B_{Dym} of (2.1) out of the given Lax operators L_{mKdV} and L_{Dym}. Using an ansatz (2.3) (modified to $L^{1/2} = w\partial + a_0 + a_1\partial^{-1} + ..$ for the Dym case) one calculates

$$\begin{aligned}
(\partial^2 + 2v\partial)^{1/2} &= \partial + v - \tfrac{1}{2}(v_x + v^2)\,\partial^{-1} + (..)\,\partial^{-2} + .. \ , \\
(w^2\partial^2)^{1/2} &= w\,\partial - \tfrac{1}{2}w_x + \tfrac{1}{8}(2w_x - \tfrac{w_x^2}{w})\,\partial^{-1} + (..)\,\partial^{-2} + .. \ ,
\end{aligned} \tag{2.8}$$

and

$$\begin{aligned}(\partial^2+2v\partial)^{3/2} &= \partial^3+3v\,\partial^2+\tfrac{3}{2}(v_x+v^2)\,\partial+(..)+(..)\,\partial^{-1}+..\,,\\ (w^2\partial^2)^{3/2} &= w^3\,\partial^3+\tfrac{1}{2}(w^3)_x\,\partial^2+(..)\,\partial+(..)+(..)\,\partial^{-1}+..\,.\end{aligned}\qquad(2.9)$$

Obviously, the operators B_{mKdV} and B_{Dym} of (2.1) are just the "leading order" terms of (2.9). More specifically: for the modified KdV we have to project the fractional power $L^{3/2}$ to terms of differential order ≥ 1, for the Dym case only orders ≥ 2 have to be taken into account. Just as in the KdV case, we can calculate higher powers such as $L^{5/2}$, project the resulting operators to their parts beyond the appropriate orders and thus find the Lax pairs for the next higher flows (5th oder modified KdV and 5th order Dym). In fact, we have established the following compact form describing the hierarchies of the KdV, the modified KdV and the Dym equation in the following unified way via their Lax representations:

$$L_t=[P_{\geq k}(L^q),L]\quad,\quad k=0,1,2\quad.\qquad(2.10)$$

Here q is a suitable fractional power (i.e. $q=n+\frac{1}{2}, n\in\mathbb{N}$, for the second order operators under consideration) and $P_{\geq k}$ are the projections to the differential orders $\geq k$ of the fractional power L^q as described above.

The Lax equations (2.10) originate from a general algebraic background, namely of Lie algebra decompositions of the space of pseudo-differential symbols introduced by (2.4). This explains [18] why k in (2.10) is restricted to the three distinguished cases $k=0,1$ and 2 discussed here. The abstract algebraic object granting the integrability of the equations (2.10) is a so-called r-matrix ([13]), which is closely related to the (multi-)Hamiltonian structures of the corresponding nonlinear equations (see [21,18] for a general discussion of the Hamiltonian aspects). The construction leading to (2.10) is purely algebraic, the fact that we had been dealing with second order Lax operators L in our motivation above does not have any significance whatsoever. In fact, starting with Lax operators L of any order, fixing $k=0,1$ or 2, we will find a hierarchy of commuting nonlinear equations from (2.10). Here q will label the different equations in the hierarchy, the commutativity of the equations for different q's is granted automatically by the general algebraic background. A detailed explanation as well as a systematic discussion of a large number of integrable nonlinear equations encoded in (2.10) is found in [18].

For $k=0$ the simplest equations associated with (2.10) are the KdV (as demonstrated) for second order Lax operators and the Boussinesq and Kaup-Kupershmidt equations for third order operators. For $k=1$ one encounters shallow water wave equations (discussed in detail in [15]), Burgers equation, the modified KdV (as demonstrated), a modified Boussinesq and the Sawada-Kotera equation. Many other integrable systems can be constructed by considering higher order Lax operators. For $k=2$ one finds a large class of coupled systems of "Dym-type".

In [17] is is shown that also three classes of 2+1-dimensional integrable equations are encoded in (2.10). For $k=0$ one has the famous construction of the Kadomtsev-Petviashvili (KP) equation of Sato's theory (see [10] for an elementary introduction and references), for $k=1$ one encounters the modified KP and $k=2$ leads fo an integrable 2+1-dimensional Dym equation first constructed in [22]. From the algebraic point of view there is no essential difference between the construction of the 1+1- and 2+1-dimensional equations encoded in (2.10).

At this stage we briefly review the reciprocal link of KdV/modified KdV to the Dym equation. One way of describing this link originates from the Painlevè analysis of integrable PDE's as introduced by Weiss et al [23]. For both the KdV and the modified KdV the singularity field Φ used in this analysis has to satisfy the nonlinear equation

$$4\,\frac{\Phi_t}{\Phi_x}=\{\Phi;x\}+\lambda\quad,\qquad(2.11)$$

where $\{\Phi;x\}=(\Phi_{xx}/\Phi_x)_x-(\Phi_{xx}/\Phi_x)^2/2$ is the Schwarzian derivative. In [6] it was noted that one may introduce the new independent variables $x'=\Phi(x,t)$ and $t'=t$. Then $w:=\Phi_x$, expressed in terms of x' and t', will satisfy the Dym equation $4w_{t'}=w^3w_{x'x'x'}$ in these new coordinates.

Having discovered KdV/modified KdV and Dym equation in their unified compact from (2.10), the natural question arise, how the well known Miura transformation and the reciprocal transformation $x\to x'=\Phi(x,t)$ between these three equations can be described in terms of the Lax representations (2.10). A general answer to this problem is given in [17], here we will motivate parts of these results by considering our second order prototypes (2.1).

A Lax equation $L_t=[B,L]$ implies the compatibility of the two linear problems

$$L\Phi=\lambda\Phi\quad,\quad\Phi_t=B\Phi\quad,\qquad(2.12)$$

which is used to implement an Inverse Scattering Transform for solving the nonlinear equations implied by the Lax pair L and B. Thus, starting with the KdV (2.1.i), we will consider an eigenfunction Φ satisfying

$$\Phi_{xx}+u\Phi=\lambda\Phi\quad,\quad\Phi_t=\Phi_{xxx}+\frac{3}{2}u\Phi_x+\frac{3}{4}u_x\Phi\quad.\qquad(2.13)$$

A simple gauge transformation $\tilde{L}:=\Phi^{-1}L\Phi$ leads to the gauge transformed operator

$$\tilde{L}=\partial^2+2\frac{\Phi_x}{\Phi}\,\partial+\frac{\Phi_{xx}+u\Phi}{\Phi}=\partial^2+2v\partial+\lambda\quad,\qquad(2.14)$$

where $v:=\Phi_x/\Phi$. Thus, for $\lambda=0$, the new operator $\tilde{L}$ is the Lax operator (2.1.ii) of the modified KdV. For $\lambda\neq 0$ we calculate

$$\tilde{L}_t=[P_{\geq 1}(\tilde{L}^{3/2}),\tilde{L}]\quad\leftrightarrow\quad 4v_t=v_{xxx}-6v^2v_x+6\lambda v_x\qquad(2.15)$$

following the above procedures. The question arises, whether $v=\Phi_x/\Phi$ will automatically satisfy the modified KdV (2.15), if the eigenfunction Φ evolves according to (2.13). By a direct computation it is easy to see that this is indeed the case. Hence, the gauge transformation $\tilde{L}=\Phi^{-1}L\Phi$ in terms of an eigenfunction Φ translates to

$$v=\frac{\Phi_x}{\Phi}\quad,\quad\Phi_{xx}+u\Phi=\lambda\Phi\quad.\qquad(2.16)$$

Eliminating Φ from (2.16) one rediscovers

$$u=\lambda-\frac{\Phi_{xx}}{\Phi}=\lambda-v_x+v^2\quad,\qquad(2.17)$$

i.e. the well known Miura transform between the KdV and (2.15). Thus, in terms of Lax equations (2.10), the Miura link is encoded in a simple gauge transformation $L \to \tilde{L} = \Phi^{-1} L \Phi$, where Φ solves the linear problem (2.13). Generally speaking, we have verified for the special case $L = \partial^2 + u$ and $q = 3/2$ that the $k = 0$ evolution (2.10), together with the time evolution (2.13), i.e. $\Phi_t = P_{\geq 0}(L^q)\Phi$, leads to the $k = 1$ evolution for $\tilde{L} = \Phi^{-1} L \Phi$:

$$\left.\begin{aligned} L_t &= [P_{\geq 0}(L^q), L] \\ \Phi_t &= P_{\geq 0}(L^q)\Phi \end{aligned}\right\} \quad \longrightarrow \quad \tilde{L}_t = [P_{\geq 1}(\tilde{L}^q), \tilde{L}] \ . \qquad \tilde{L} = \Phi^{-1} L \Phi \tag{2.18}$$

As the theorem of the next section will show this is not a special feature of the second order Lax operators considered here, but this holds for arbitrary Lax operators evolving according to (2.10).

The reciprocal link to the Dym equation can also be rephrased in terms of the Lax operators. Let us now start with an eigenfunction Ψ of the modified KdV (2.1.ii), i.e. we assume

$$\Psi_{xx} + 2v\Psi_x = \lambda\Psi \ , \quad \Psi_t = \Psi_{xxx} + 3\Psi_{xx} + \frac{3}{2}(v_x + v^2)\Psi_x \ . \tag{2.19}$$

Eliminating v by $2v = (\lambda\Psi - \Psi_{xx})/\Psi_x$, the time evolution (2.19) leads to

$$4\,\frac{\Psi_t}{\Psi_x} = \{\Psi; x\} + 3\lambda + \frac{3}{2}\lambda^2 \left(\frac{\Psi}{\Psi_x}\right)^2 \ . \tag{2.20}$$

For $\lambda = 0$ this coincides with the singularity manifold equation (2.11) shared by the KdV and the modified KdV. The transformation $x' = \Phi(x,t)$ studied in [6] suggests to introduce the new variables $x' = \Psi(x,t), t' = t$, where we assume Ψ to satisfy the linear equations (2.19). Thus, starting with the Lax operator $\tilde{L} = \partial^2 + 2v\partial$ of the modified KdV, we rewrite $\partial = \partial_x = \frac{\partial x'}{\partial x}\partial_{x'} = \Psi_x \partial'$ and obtain

$$\begin{aligned} L' = \tilde{L} = \partial^2 + 2v\partial = \ & \partial\Psi_x\partial' + 2v\Psi_x\partial' = \Psi_x^2\partial'^2 + (\Psi_{xx} + 2v\Psi_x)\partial' \\ = \ & \Psi_x^2\partial'^2 + \lambda\Psi\partial' = w^2\partial'^2 + \lambda x'\partial' \ , \end{aligned} \tag{2.21}$$

where $w := \Psi_x$. For $\lambda = 0$ this is the Lax operator (2.1.iii) associated with the Dym equation. For $\lambda \neq 0$ we may perform the same analysis as before, i.e. we calculate

$$(w^2\partial'^2 + \lambda x'\partial')^{3/2} = w^3\partial'^3 + \frac{1}{2}((w^3)_x + 3\lambda x' w)\partial'^2 + (..)\partial' + (..) + \ldots \ , \tag{2.22}$$

take the projection $P'_{\geq 2}$ to terms involving at least second powers of ∂' and calculate

$$L'_{t'} = [P'_{\geq 2}(L'^{3/2}), L'] \quad \leftrightarrow \quad 4w_{t'} = w^3 w_{x'x'x'} + 3\lambda^2 (x' - x'^2 \frac{w_{x'}}{w}) \ . \tag{2.23}$$

Again, the question arises, whether the dynamics of the (generalized) Dym equation (2.23) is the same as the dynamics (2.19) of the eigenfunction of the modified KdV, subject to the transformation $w = \Psi_x, x' = \Psi, t' = t$. It may be checked that this is indeed the case. Eliminating Ψ from the transformation

$$w = \Psi_x \ , \quad x' = \Psi \ , \quad t' = t \ , \quad \Psi_{xx} + 2v\Psi_x = \lambda\Psi \tag{2.24}$$

one finds

$$(w_x + 2vw)_x = \lambda w \ , \quad \frac{\partial x'}{\partial x} = w \ , \quad t' = t \ , \tag{2.25}$$

as a reciprocal auto-Bäcklund transformation between the modified KdV as in (2.2.ii) and the Dym equation (2.23).

Hence, we have verified for the special case $\tilde{L} = \partial^2 + 2v\partial$ and $q = 3/2$ that the $k = 2$ evolution of (2.10) is obtained from the the $k = 1$ evolution by a transformation $x' = \Psi$, where Ψ is an eigenfunction of the $k = 1$ Lax equation, i.e. $\Psi_t = P_{\geq 1}(\tilde{L}^q)\Psi$:

$$\left.\begin{aligned} \tilde{L}_t &= [P_{\geq 1}(\tilde{L}^q), \tilde{L}] \\ \Psi_t &= P_{\geq 1}(\tilde{L}^q)\Psi \end{aligned}\right\} \quad \begin{array}{c} x' = \Psi \ , \ t' = t \\ \longrightarrow \\ L' = \tilde{L} \end{array} \quad L'_{t'} = [P'_{\geq 2}(L'^q), L'] \ . \tag{2.26}$$

This result does not only hold for the second order Lax operators considered here. In fact, (2.26) holds in general for any given Lax operator satisfying the $k = 1$ hierarchy (2.10) and for any eigenfunctions associated with this.

As demonstrated with these simple examples, the well established Miura transform between KdV and modified KdV as well as the reciprocal link to the Dym equation have very simple counterparts on the level of Lax equations (2.10). The Miura transformations corresponds to a simple gauge transformation $L \to \Phi^{-1} L \Phi$ with an eigenfunction Φ, mapping the $k = 0$ case in (2.10) to the case $k = 1$. The reciprocal link is a transformation $x' = \Psi$ involving an eigenfunction Ψ, it maps the $k = 1$ case in (2.10) to the case $k = 2$.

3 General results

The calculations for second order operators presented in the last section suggest simple links between the tree cases $k = 0, 1$ and 2 of the Sato equations (2.10). Indeed, in [17] and [18] following general results are proven:

Theorem 1:

a) $(k = 0 \to k = 1)$ Let the operator L and the function Φ satisfy $L_t = [P_{\geq 0}(L^q), L]$ and $\Phi_t = P_{\geq 0}(L^q)\Phi$. Then $\tilde{L} = \Phi^{-1} L \Phi$ satisfies $\tilde{L}_t = [P_{\geq 1}(\tilde{L}^q), \tilde{L}]$.

b) $(k = 1 \to k = 2)$ Let the operator $\tilde{L} = \tilde{L}(x,t)$ and the function $\Psi = \Psi(x,t)$ satisfy $\tilde{L}_t = [P_{\geq 1}(\tilde{L}^q), \tilde{L}]$ and $\Psi_t = P_{\geq 1}(\tilde{L}^q)\Psi$. Then the transformation $x' = \Psi(x,t), t' = t$ leads to an operator $L'(x',t') = \tilde{L}(x,t)$ satisfying $L'_{t'} = [P'_{\geq 2}(L'^q), L']$, where $P'_{\geq 2}$ is the projection operator related to the differential symbol $\partial' = \Psi_x^{-1}\partial$.

The proof does not require any other assumption on the explicit form of L than those implied by the Lax equation $L_t = [P_{\geq k}(L^q), L]$. E.g., in order to prove a) one establishes the operator identity

$$\begin{aligned} & \left(\tfrac{1}{\Phi} L \Phi\right)_t - [\, P_{\geq 1}((\tfrac{1}{\Phi} L \Phi)^q) \, , \, \tfrac{1}{\Phi} L \Phi \,] \\ & = \tfrac{1}{\Phi} \left(L_t - [P_{\geq 0}(L^q), L] \right) \Phi \ - \ [\, \tfrac{1}{\Phi} \| \Phi_t - P_{\geq 0}(L^q)\Phi \|\, , \, \tfrac{1}{\Phi} L \Phi \,] \ , \end{aligned} \tag{3.1}$$

which holds for any arbitrary operator L and any arbitrary function Φ. Here $\|a\|$ is the multiplication operator with the function a. Obviously, from (3.1) the result a) of theorem

1 is obtained trivially.
In particular, starting with an N-th order Operator

$$L = \partial^N + u_{N-2}\partial^{N-2} + u_{N-3}\partial^{N-3} + ... + u_1\partial + u_0 \qquad (3.2)$$

we may define a hierarchy of integrable coupled nonlinear evolution equations for the fields $u_0, ..u_{N-2}$ by the Lax equations $L_t = [P_{\geq 0}(L^q), L]$. As demonstrated in section 2, for $N = 2$ this is the KdV hierarchy for the field u_0. For $N = 3$ the coupled system for u_0, u_1 encodes the Boussinesq hierarchy. An eigenfunction Φ satisfying $L\Phi = \lambda\Phi$ leads to a gauge transformed operator

$$\begin{aligned}\tilde{L} = \tfrac{1}{\Phi}L\Phi = {} & \partial^N + N\tfrac{\Phi_x}{\Phi}\partial^{N-1} + (u_{N-2} + \tfrac{N(N-1)}{2}\tfrac{\Phi_{xx}}{\Phi})\partial^{N-2} \\ & + ... + (u_1 + ..)\partial + \lambda \qquad (3.3) \\ = {} & \partial^N + v_{N-1}\partial^{N-1} + v_{N-2}\partial^{N-2} + ... + v_1\partial + \lambda \ ,\end{aligned}$$

where the new fields $v_{N-1} := \frac{N\Phi_x}{\Phi}$, $v_{N-2} := u_{N-2} + \frac{N(N-1)}{2}\frac{\Phi_{xx}}{\Phi}$, .. satisfy a "modified" hierarchy of integrable equations which can be determined easily from the Lax equation $\tilde{L}_t = [P_{\geq 1}(\tilde{L}^q), \tilde{L}])$. The Miura transformation is obtained by eliminating the eigenfunction Φ from the relations $v_{N-1} = N\Phi_x/\Phi$ etc. using $L\Phi = \lambda\Phi$.
Now we may start with the modified operator $\tilde{L}$ given by

$$\tilde{L} = \partial^N + v_{N-1}\partial^{N-1} + v_{N-2}\partial^{N-2} + ... + v_1\partial \qquad (3.4)$$

and an eigenfunction satisfying $\tilde{L}\Psi = \lambda\Psi$. The transformation $x' = \Psi(x,t), t' = t$ transforms $\tilde{L}$ into

$$\begin{aligned}L'(x',t') = \tilde{L}(x,t) = {} & (\Psi_x\partial')^N + v_{N-1}(\Psi_x\partial')^{N-1} + ... + v_1\Psi_x\partial' \\ = {} & \Psi_x^N\partial'^N + (v_{N-1}\Psi_x^{N-1} + \tfrac{N(N-1)}{2}\Psi_x^{N-2}\Psi_{xx})\partial'^{N-1} \\ & + ... + (v_2\Psi_x^2 + ..)\partial'^2 + \lambda\Psi\partial' \\ = {} & w_N^N\partial'^N + w_{N-1}\partial'^{N-1} + w_{N-2}\partial'^{N-2} + ... + w_2\partial'^2 + \lambda x'\partial' \ , \qquad (3.5)\end{aligned}$$

where the new fields $w_N = \Psi_x$, $w_{N-1} = v_{N-1}\Psi_x^{N-1} + \frac{N(N-1)}{2}\Psi_x^{N-2}\Psi_{xx}$, .., satisfy a "reciprocally linked" hierarchy of integrable equations which can be determined easily from the Lax equations $L'_{t'} = [P'_{\geq 2}(L'^q), L']$. We remark that the first order term A of $L' = w_N^N\partial'^N + ... + w_2\partial'^2 + A\partial'$ has to be $A = \lambda x'$, as $A = L'x' = \tilde{L}\Psi = \lambda\Psi$. The eigenfunction Ψ may then be eliminated from the transformation using $L'\Psi = \lambda\Psi$, leading to a reciprocal Bäcklund transformation between the equations for $v_1, .., v_{N-1}$ and $w_2, .., w_N$.
Hence, the links

$$\text{KdV} \xrightarrow{\text{(Miura)}} \text{modified KdV} \xrightarrow{\text{(reciprocal link)}} \text{Dym}$$

associated with second order Lax operators are just a special case of a more general picture:

$$\begin{array}{ccccc} \tilde{L} = \frac{1}{\Phi}L\Phi & & x' = \Psi, t' = t, L' = \tilde{L} & & \\ \downarrow & & \downarrow & & \\ L_t = [P_{\geq 0}(L^q), L] & \longrightarrow & \tilde{L}_t = [P_{\geq 1}(\tilde{L}^q), \tilde{L}] & \longrightarrow & L'_{t'} = [P'_{\geq 2}(L'^q), L'] \end{array}$$

To any integrable hierarchy associated with a N-th order Gelfand-Dikii spectral problem (3.2) there is a Miura transform to a modified hierarchy associated with the spectral problem (3.3), which admits a reciprocal link to a system of coupled equations of "Dym type" associated with the spectral problem (3.5). Examples and details are worked out in [18].

Further, the compact construction (2.10) for all these integrable hierarchies allow very general formulations of auto-Bäcklund transformations leaving these hierarchies invariant. We quote following results proved in [17]:

Theorem 2:

a) $(k = 0 \to k = 0)$ Let the operator L and the function Φ satisfy $L_t = [P_{\geq 0}(L^q), L]$ and $\Phi_t = P_{\geq 0}(L^q)\Phi$. Then $\tilde{L} = \Phi\partial\Phi^{-1}L\Phi\partial^{-1}\Phi^{-1}$ will again satisfy $\tilde{L}_t = [P_{\geq 0}(\tilde{L}^q), \tilde{L}]$.

b) $(k = 1 \to k = 1)$ Let the operator L and the function Φ satisfy $L_t = [P_{\geq 1}(L^q), L]$ and $\Phi_t = P_{\geq 1}(L^q)\Phi$. Then $\tilde{L} = \Phi^2\Phi_x^{-1}\partial\Phi^{-1}L\Phi\partial^{-1}\Phi_x\Phi^{-2}$ will again satisfy $\tilde{L}_t = [P_{\geq 1}(\tilde{L}^q), \tilde{L}]$.

c) $(k = 2 \to k = 2)$ Let the operator $L = L(x,t)$ and the function $\Phi = \Phi(x,t)$ satisfy $L_t = [P_{\geq 2}(L^q), L]$ and $\Phi_t = P_{\geq 2}(L^q)\Phi$. Introducing the transformation $x' = x - \Phi_x^{-1}\Phi$, $t' = t$, the operator $L'(x',t') = \Phi^2\Phi_x^{-1}\partial\Phi^{-1}L(x,t)\Phi\partial^{-1}\Phi_x\Phi^{-2}$ will again satisfy $L'_{t'} = [P'_{\geq 2}(L'^q), L']$.

The transformations of theorem 1 and theorems 2 may all be formulated as Darboux theorems as shown in [17]. This reveals an intimate relation between all these transformations which can be summarized in a collection of various commuting Bäcklund diagrams. In particular, sequences of exact solutions for the "Dym-type" equations can be extracted from these transformations in a straightforward way.

References

[1] M.J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*, SIAM Philadelphia (1982).

[2] B.G. Konopelchenko, *Nonlinear Integrable Equations*, Lecture Notes in Physics, Vol. 270, Springer Verlag, Heidelberg (1987).

[3] B.G. Konopelchenko and C. Rogers, in *Nonlinear Equations in the Applied Sciences*, (eds. W.F. Ames and C. Rogers), Academic Press, New York (in press).

[4] C. Rogers and P. Wong, Physica Scripta 30, 10 (1984).

[5] F. Calogero and A. Degasperis, *Spectral Transform and Solitons: Tools to Solve and Investigate Nonlinear Evolution Equations I*, North Holland, Amsterdam (1982).

[6] J. Weiss, J. Math. Phys. 24, 1405 (1983).

[7] C. Rogers , Phys. Lett. A 120, 15 (1987).

[8] M. Sato and Y. Sato, in *Nonlinear Partial Differential Equations in Applied Science*, (eds. H. Fujita, P.D. Lax and G. Strang), Kinokuniya/North Holland, Tokyo (1983).

[9] E. Date, M. Kashiwara, M. Jimbo and T. Miwa, in *Nonlinear Integrable Systems - Classical Theory and Quantum Theory*, (eds. M. Jimbo and T. Miwa), World Scientific, Singapore (1983).

[10] Y. Ohta, J. Satsuma, D. Takahashi and T. Tokihiro, Progr. Theor. Phys. Suppl. 94, 210 (1988).

[11] M. Adler, Invent. Math. 50, 219 (1978).

[12] V.G. Drinfeld, Soviet Math. Dokl. 27, 68 (1983).

[13] M.A. Semenov-Tian-Shansky, Funct. Anal. Appl. 17, 259 (1983).

[14] M.A. Semenov-Tian-Shansky, Publ. RIMS, Kyoto Univ., 21, 1237 (1985).

[15] B.A. Kupershmidt, Comm. Math. Phys. 99, 51 (1985).

[16] K. Kiso, Progr. Theor. Phys. 83, 1108 (1990).

[17] W. Oevel and C. Rogers, *Gauge Transformations and Reciprocal Links in 2+1 Dimensions*, preprint, Loughborough Univ. (1991).

[18] B.G. Konopelchenko and W. Oevel, *An r-matrix Approach To Nonstandard Classes of Integrable Equations*, preprint, Loughborough Univ. (1991).

[19] I.M. Gelfand and L.A. Dikii, Funct. Anal. Appl. 10, 259 (1976).

[20] T. Ratiu, in Lecture Notes in Mathematics, Vol. 775, Springer Verlag, Heidelberg (1980).

[21] W. Oevel and O. Ragnisco, Physica A 161, 181 (1989).

[22] B.G. Konopelchenko and V.G. Dubrovsky, Phys. Lett. A 102, 15 (1984).

[23] J. Weiss, M.Tabor and G. Carnevale, J. Math. Phys. 24, 522 (1983).

SYMMETRIES ON SOLITON MANIFOLD*

M Błaszak

Institute of Physics, A.Mickiewicz University, Poznań, Poland

Abstract

For a class of completely integrable, finite dimensional multi-soliton systems an algebra of hamiltonian symmetries which are polynomials in t variable is constructed. The approach is of purely algebraic nature.

1 Introduction

It is well known that soliton systems in (1+1) dimension can be divided into two classes as far as their time dependent symmetries are concerned. One class is made of systems for which we know only linear in time, non-hamiltonian symmetries [1]. A typical example is the famous Korteweg-de Vries (KdV) equation [2]. The other class is composed of these equations for which we know symmetries polinomial in t variable which all are hamiltonian vector fields [1]. A typical representative is the Benjamin-Ono (BO) equation [3,4]. As equations from both classes are integrable hamiltonian systems (we can solve the Cauchy problem for each of them), it seems reasonable to expect the existence of hamiltonian and non-hamiltonian symmetries, for each equation of both classes, which are polynomials of arbitrary degree in t variable.

In this lecture we present a systematic method of construction of polynomial in t variable hamiltonian symmetries for the systems of the KdV class, when we limit our considerations to the finite dimensional invariant submanifold, i.e. the so called multi-soliton submanifold. The method is based on the results obtained in the paper [5]. Section 2 presents the collection of the known facts concerning symmetries of soliton systems. In section 3, the polynomial in t variable, hamiltonian symmetries in linear representation are constructed. And finally, in section 4, we apply the results from the linear representation to the soliton representation.

*Work supported by MEN under Project P/04/193/90-2

2 Symmetries and integrals of motion of soliton systems

On a suitable manifold M we consider the hamiltonian evolution equation of the form

$$u_t = K(u) = \Theta \nabla H(u), \tag{2.1}$$

where $u = u(x,t) \in M$ denotes the field variable, $K(u)$ is a vector field on M, Θ is a hamiltonian (implectic or Poisson) operator [6,7] and $\nabla H(u)$ is a gradient of a scalar field H(u).

Definition 1. We admit the vector field $S = S(u,x,t)$ and the scalar field $G = G(u,x,t)$ to depend explicitely on the time variable t and the space variable x.

(i) The vector field $S(u,x,t)$ is said to be a symmetry for (2.1) if

$$\frac{\partial}{\partial t} S + [K,S] = 0. \tag{2.2}$$

(ii) The scalar field G(u,x,t) is said to be an integral of motion for (2.1) if

$$\frac{\partial}{\partial t} G + \langle \nabla G, K \rangle = 0. \tag{2.3}$$

[,] stands for Lie bracket of two vector fields and $\langle\ ,\ \rangle$ is the duality map (scalar product) between co-vectors and vectors.

The common feature of all soliton systems (2.1) is the existence of infinite hierarchy of hamiltonian symmetries

$$K_n(u) = \Theta \nabla H_n(u), \qquad n = 0,1,2,\ldots . \tag{2.4}$$

which do not depend explicitely on x and t variable and form an abelian Lie subalgebra : $[K_n, K_m]$=0. The respective scalar fields $H_n(u)$ are integrals of motion of (2.1) and are in involution with respect to the Poisson bracket defined by Θ:

$$\{H_n, H_m\}_\Theta := \langle \nabla H_m, \Theta \nabla H_n \rangle = 0. \tag{2.5}$$

Apart from the time-independent symmetries of (2.1) we can construct symmetries which depend explicitely on the time variable t. Such symmetries actually are polynomials in t variable. In the construction of these symmetries the notion of mastersymmetries [8] plays an important role. Let

$$ad\ K(S) := [K,S] \tag{2.6}$$

define an $ad\ K$ action on S vector field.

Definition 2. The vector field S(u,x) is called a mastersymmetry of degree r for an evolution equation (2.1) if

$$(ad\ K)^{r+1}\ S = 0. \tag{2.7}$$

It is not difficult to notice that there is one to one correspondence between a mastersymmetry of degree r and the symmetry $\Pi(u,x,t)$ which is a polynomial (of an order r) in time variable t, i.e.

$$\Pi(u,x,t) := \exp(-t\ ad\ K) S(u,x) = \sum_{n=0}^{r} \frac{(-t)^n}{n!} (ad\ K)^n S(u,x), \tag{2.8}$$

where Π fulfils condition (2.2).

As it was mentioned in Introduction, soliton systems in (1+1) dimension can be divided into two classes. One of them comprises the systems for which there exists the so called recursion operator $\phi(u)$ [9,10]. This operator $\phi(u)$ has an important property, i.e. when it acts on one symmetry of a given evolution equation it produces another one. For such systems we know only two hierarchies of mastersymmetries defined on the whole M. The first one contains mastersymmetries of degree 0:

$$K_n(u) = \phi^n(u) K_0(u) \tag{2.9}$$

which are just our time independent symmetries (2.4), and the second one contains mastersymmetries of degree 1:

$$\tau_n(u,x) = \phi^n(u) \tau_0(u), \tag{2.10}$$

related through (2.8) to the symmetries which are linear in t variable

$$\Pi_n(u,x,t) = \phi^n(u)(\tau_0(u,x) + f(\alpha) t K(u)), \qquad f(\alpha) = const. \tag{2.11}$$

$K_0(u)$ and $\tau_0(u,x)$ are generators of space translations and scaling, respectively. Moreover, contrary to $K_n(u)$, vector fields $\tau_n(u,x)$ are not hamiltonian. Actually, hamiltonian dynamical systems (2.1) with the recursion operator $\phi(u)$ are in fact multi-hamiltonian, i.e. all $\Theta_n := \phi^n \Theta$ are Poisson operators. Now, for an arbitrary Poisson operator Θ_m all vector fields K_n are hamiltonian and all but one vector fields τ_n are non-hamiltonian.

Example. The Korteweg-deVries equation.

The celebrated KdV equation [2] is given by

$$\begin{aligned} u_t &= u_{3x} + a u u_x = K_1(u) \\ &= \partial \nabla \int_{-\infty}^{+\infty} (-\frac{1}{2} u_x^2 + \frac{1}{6} a u^3)\, dx = \Theta \nabla H_1(u), \end{aligned} \tag{2.12}$$

where u is assumed to be an element of the Schwartz space of rapidly decreasing functions S(R). The hierarchies of commuting symmetries generated by the well known recursion operator $\phi(u)$

$$K_n(u) = \phi^n(u)K_0(u) = (\partial^2 + \frac{1}{3}a\partial u\partial^{-1} + \frac{1}{3}au)^n u_x \tag{2.13}$$

and mastersymmetries

$$\tau_n(u,x) = \phi^n(u)\tau_0(u,x) = (\partial^2 + \frac{1}{3}a\partial u\partial^{-1} + \frac{1}{3}au)^n(\frac{1}{2}xu_x + u), \tag{2.14}$$

related to the linear in time symmetries

$$\Pi_n(u,x,t) = \phi^n(u)(\tau_0 + \frac{3}{2}tK_1), \tag{2.15}$$

fulfill the hereditary algebra (2.12) with $\alpha = 1/2$. With respect to an arbitrary Poisson operator $\Theta_m = \phi^m\partial$, all vector fields K_n are hamiltonian: $K_n(u) = \Theta_m(u)\nabla H_{n-m}(u)$, and all vector fields $\tau_n(u,x)$ are non-hamiltonian exept one

$$\tau_{m-1}(u,x) = \Theta_m(u)\,\nabla F(u,x), \qquad F(u,x) = \frac{3}{2a}\int_{-\infty}^{+\infty} xu\,dx. \tag{2.16}$$

The second class of soliton systems contains the systems for which there is no recursion operator in the explicit form (in the non-generalized sense [11]). For such systems we can construct mastersymmetries of an arbitrary degree which constitutes the following algebra

$$[S_{r,n}, S_{p,m}] = (r(m+\alpha) - p(n+\alpha))S_{r+p-1,n+m}, \quad \alpha = const. \tag{2.17}$$

where $S_{k,l}$ stands for the l-th mastersymmetry of degree k. Of course to each mastersymmetry we can relate a polynomial in t variable symmetry

$$K_{p,m}(u,x,t) = \exp(-t\,ad\,K)S_{p,m}(u,x) = \sum_{n=0}^{p}\binom{p}{n}t^n S_{p-n,m+n}(u,x), \tag{2.18}$$

which fulfils the commutator relations (2.18). Moreover, all these symmetries are hamiltonian vector fields. One can notice that $K_{0,m}$-symmetries play the role of K_m -symmetries and do commute in pairs.

Example. The Benjamin-Ono equation.

The BO equation [3,4] is given by

$$u_t = 4auu_x + \mathcal{H}u_{xx} = \partial\,\nabla\int_{-\infty}^{+\infty}\left(\frac{1}{2}u\mathcal{H}u_x + \frac{2}{3}au^3\right)dx, \tag{2.19}$$

where $u \in S(R)$ and $\mathcal{H}$ stands for the Hilbert transform. A few first mastersymmetries are of the form

$$\begin{aligned} S_{0,0} &= u_x, \quad S_{0,1} = 8auu_x + 2\mathcal{H}u_{xx}, \dots \\ S_{1,0} &= xu_x + u, \quad S_{1,1} = xS_{0,1} + 4au^2 + 3\mathcal{H}u_x, \dots \\ S_{2,0} &= x^2u_x + 2xu, \dots \\ &\dots \end{aligned} \tag{2.20}$$

and fulfill the algebra (2.18) with $\alpha = 1$ [12], and all are hamiltonian vector fields with respect to $\Theta = \partial$.

Now, the question arises whether it is possible to construct hamiltonian symmetries which are polynomials in t variable (2.19) and fulfill the commutator relation (2.18) for the systems with a recursion operator and known hereditary algebra? The affirmative ansver in the case of the so called N-soliton submanifold $M_N \subset M$ is discussed in next section.

It is well known that evolution equations with recursion operator admit multisoliton solutions [13]. The N-soliton solutions can be described as elements of the following invariant submanifold M_N of M [14]

$$M_N = \{u \mid \text{there exist } \alpha_n \text{ such that } \sum_{n=0}^{N}\alpha_n K_n = 0\}. \tag{2.21}$$

We turn our attention to N-solitons which decompose asymptotically for $t \to \infty$ into one-solitons of the form

$$u_N \cong \sum_{i=1}^{N} s_i(x + c_i t + q_i). \tag{2.22}$$

If the velocities c_i and the phases q_i are considered as variables, then the set of these solutions forms a 2N-dimensional invariant submanifold M_N of M. For the submanifold M_N we may provide a new parametrization in the following way: we define a map Π which assigns to each $u_N(x,t)$ a set of asymptotic data $(q_1 \dots, q_N, c_1, \dots, c_N)$. It can be observed that although we refer to the asymptotic form of N-solitons, this new parametrization is defined for an arbitrary time.

Lemma 1. The quantities q_i, c_i are scalar fields on the submanifold M_N with the following time dependence

$$(u_N)_t = K(u_N) \xrightarrow{\quad\Pi\quad} \begin{cases} \frac{\partial}{\partial t}q_i = c_i \\ \\ \frac{\partial}{\partial t}c_i = 0. \end{cases} \tag{2.23}$$

Proof is given in [5].

Lemma 2. The projection onto N-soliton submanifold M_N preserves the hamiltonian structure of dynamics.

Lemma 1 shows that the flow (2.1) on M_N is linearized in our new coordinates. Lemma 2 says that algebraic objects introduced in this section, like recursion operator, Poisson operators, symmetries e.c.t., survive the reduction $M \to M_N$.

Now, the general procedure to recover the algebraic structure of the dynamical system (2.1) on M_N is very simple in principle. First we find the structure of the linear system (q_i, c_i) , then carry over the whole structure to the system represented by the field coordinate u_N. Although we don't know the explicit form of our variable transformation Π , we know how various algebraic objects behave under a change of coordinates.

3 Algebraic structure of a linear representation

In this section we restrict our considerations to the linear representation (q_i, c_i) (2.23) and examine some of its properties. This representation admits many different hamiltonian formulations

$$v_t = K(v) = \bar{\Theta}_0(\alpha)\,\nabla h = \begin{pmatrix} 0 & \Lambda_{1-\alpha} \\ -\Lambda_{1-\alpha} & 0 \end{pmatrix} \nabla \left(\frac{1}{1+\alpha} \sum_{i=1}^{N} c_i^{1+\alpha} \right), \quad 0 < \alpha \leq 1, \tag{3.1}$$

where $v = (q_1, \ldots, q_N, c_1, \ldots, c_N)^T, K(v) = (c_1, \ldots, c_N, 0, \ldots, 0)^T$ and Λ denotes the diagonal $N \times N$ - matrix

$$\Lambda_{1-\alpha} = \begin{pmatrix} c_1^{1-\alpha} & & 0 \\ & \ddots & \\ 0 & & c_N^{1-\alpha} \end{pmatrix}. \tag{3.2}$$

The recursion operator is of the form

$$\bar{\phi} = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_1 \end{pmatrix} \tag{3.3}$$

with the following eigenstates

$$\begin{aligned} \bar{A}_i = \frac{\partial v}{\partial q_i}, \qquad \bar{B}_i = \frac{\partial v}{\partial c_i}, \quad i = 1, \ldots, N \\ \bar{\phi}\,\bar{A}_i(\bar{B}_i) \;=\; c_i\,\bar{A}_i(\bar{B}_i), \end{aligned} \tag{3.4}$$

where c_i are eginevalues (each of them occurs twice). More details the reader can find in [5,16].

Now, for our further considerations, including the construction of the full structure in the Lie-algebra of vector fields for the nonlinear soliton systems, let us define

the fundamental algebra $\bar{\mathcal{A}}$ consisting of the following vector fields in the linear space

$$\begin{aligned} \bar{P}_{r,n} &:= \bar{\phi}^n \sum_{i=1}^{N} q_i^r \bar{A}_i, \\ \bar{M}_{s,n} &:= \bar{\phi}^{n+1} \sum_{i=1}^{N} q_i^s \bar{B}_i. \end{aligned} \tag{3.5}$$

These vector fields fulfill the commutator relations

$$\begin{aligned} [\bar{P}_{r,n}, \bar{P}_{s,m}] &= (s-r)\bar{P}_{r+s-1,n+m}, \\ [\bar{P}_{r,m}, \bar{M}_{s,m}] &= s\bar{M}_{s+r-1,n+m} - n\bar{P}_{s+r,n+m}, \\ [\bar{M}_{r,n}, \bar{M}_{s,m}] &= (m-n)\bar{M}_{r+s,n+m}. \end{aligned} \tag{3.6}$$

Generally $\bar{P}_{r,n}$ and $\bar{M}_{s,n}$ are non-hamiltonian vector fields, hence our first goal is to construct a subalgebra of vector fields which are hamiltonian with respect to the Poisson operator $\bar{\Theta}_p(\alpha) := \bar{\phi}^p \bar{\Theta}_0(\alpha)$. For a fixed p we define the following vector fields

$$\bar{S}^p_{r,n} := (-1)^r[(n+\alpha-p)\bar{P}_{r,n} - r\bar{M}_{r-1,n}], \tag{3.7}$$

fulfilling the commutator relations

$$[\bar{S}^p_{r,n}, \bar{S}^p_{s,m}] = (r(m+\alpha-p) - s(m+\alpha-p))\bar{S}^p_{s+r-1,n+m}. \tag{3.8}$$

Each vector field $\bar{S}^p_{r,n}$ is a hamiltonian vector field with respect to the Poisson operator $\bar{\Theta}_p(\alpha)$

$$\bar{S}^p_{r,n} = \bar{\Theta}_p\,\nabla \bar{T}_{r,n-p}, \tag{3.9}$$

where

$$\bar{T}_{r,n} = (-1)^r \sum_{i=1}^{N} q_i^r c_i^{n+\alpha} \tag{3.10}$$

are some scalar fields. The following commutator relations between these scalar fields hold

$$\{\bar{T}_{r,n}, \bar{T}_{s,m}\}_{\bar{\Theta}_p} = (r(m+\alpha) - s(n+\alpha))\bar{T}_{s+r-1,n+m+p}. \tag{3.11}$$

For our linear system (q_i, c_i) vector fields

$$\bar{K}^p_{r,n}(t) = \sum_{m=0}^{r} \begin{pmatrix} r \\ m \end{pmatrix} t^m\, \bar{S}^p_{r-m,n+m} \tag{3.12}$$

are hamiltonian w.r.t. $\bar{\Theta}_p$, polinomial in t variable symmetries and

$$\bar{H}^p_{r,n}(t) = \sum_{m=0}^{r} \begin{pmatrix} r \\ m \end{pmatrix} t^m\, \bar{T}_{r-m,n+m-p} \tag{3.13}$$

are suitable time-dependent conserved scalar fields.

4 Soliton manifold

In the previous section we studied the algebraic structure of linear representation (2.23). But of course we would like to carry over the structure on the physical N-soliton manifold and express all desired quantities in terms of the field variable $u_N(x,t)$. The details are presented in [5,16]. In this section we only develop the idea of the soliton fundamental algebra of vector fields. It allows us to reproduce in the common way the hierarchies of symmetries and mastersymmetries known from soliton systems.

Via the inverse of the pushforward Π' we define the soliton fundamental algebra $\mathcal{A}$ as the image of $\bar{\mathcal{A}}$ under $(\Pi')^{-1}$. Then the basic vector fields have the form

$$\begin{aligned} P_{r,n}(u_N) &:= \phi^n \sum_{i=1}^{N} q_i^r A_i, \qquad A_i = \frac{\partial u_N}{\partial q_i} \\ M_{s,m}(u_N) &:= \phi^{n+1} \sum_{i=1}^{N} q_i^s B_i, \qquad B_i = \frac{\partial u_N}{\partial c_i}. \end{aligned} \tag{4.1}$$

Since the pushforward is a Lie-algebra isomorphism the commutator relations of $\bar{\mathcal{A}}$ are also valid for $\mathcal{A}$. The same holds for all subalgebras, so we can apply in a direct way the results from the previous section.

KdV example. As was found in [5] for the KdV evolution equation, $\bar{\Theta}_0(\frac{1}{2})$ is related via Π' with the second Poisson operator $\Theta = \partial^3 + \frac{1}{3}au\partial + \frac{1}{3}a\partial u$. Hence, as $\alpha = \frac{1}{2}$, hamiltonian (w.r.t. Poisson operator Θ) mastersymmetries of degree n on M_N are of the form

$$\begin{aligned} S^0_{r,n} := S_{r,n}(u_N) &= (-1)^r \left[\left(n + \frac{1}{2} \right) P_{r,n} - r M_{r-1,n} \right] \\ &= (-1)^r \phi^n \sum_{i=1}^{N} q_i^{r-1} \left[\left(n + \frac{1}{2} \right) q_i (u_N)_{q_i} - r c_i (u_N)_{c_i} \right] \end{aligned} \tag{4.2}$$

and the related scalar fields

$$T_{r,n}(u_N) = -\frac{3}{a} \frac{1}{2n+1-r} \int_{-\infty}^{+\infty} x S_{r,n}(u_N)\, dx. \tag{4.3}$$

The last equality follows from the fact that for the KdV

$$F(u_N) = T_{1,0}(u_N) = \frac{2}{3a} \int_{-\infty}^{+\infty} x u_N \, dx, \tag{4.4}$$

hence from (3.11)

$$\{T_{r,n}(u_N), T_{1,0}(u_N)\}_\Theta = -\left(n + \frac{1}{2} - \frac{r}{2} \right) T_{r,n}(u_N) \tag{4.5}$$

and from the definition of Poisson bracket (2.5)

$$\{T_{r,n}(u_N), T_{1,0}(u_N)\}_\Theta = \langle \nabla T_{1,0}(u_N), S_{r,n}(u_N) \rangle = \frac{3}{2a} \int_{-\infty}^{+\infty} x S_{r,n}(u_N)\, dx. \tag{4.6}$$

Polinomial in t variable symmetries and conserved functionals on KdV N-soliton manifold are of the form (3.12) and (3.13), respectively, and of course fulfill conditions (2.2) and (2.3).

Notice that our vector fields and functional densities are differential functions of u_N variable, containing besides x-derivatives also q_i- and c_i-derivatives. We can eliminate these new derivatives. One can find that

$$\begin{aligned} K_n(u_N) &= P_{0,n}(u_N) = \phi^n \sum_{i=1}^{N} A_i = \sum_{i=1}^{N} c_i^n (u_N)_{q_i}, \\ \tau_n(u_N) &= M_{0,n}(u_N) - \frac{1}{2} P_{1,n}(u_N) = \sum_{i=1}^{N} \left(c_i^{n+1} (u_N)_{c_i} - \frac{1}{2} q_i c_i^n (u_N)_{q_i} \right). \end{aligned} \tag{4.7}$$

Hence $(u_N)_{q_i}$ and $(u_N)_{c_i}$ can be expressed as a linear combination of K_n and τ_n. But still the coefficients of our differential functions are c_i- and q_i-dependent.

References

[1] M.Błaszak," *Theory of Classical Soliton Particles*", UAM Press (1989).

[2] S.C.Gardner, J.M.Green, M.Kruskal and R.M.Miura, Comm.Pure Apply Math. **27** (1974) 97.

[3] T.B.Benjamin, J.Fluid Mech. **19** (1967) 559.

[4] H.Ono, J.Phys.Soc.Jpn. **39** (1975) 1082.

[5] G.Oevel, B.Fuchssteiner and M.Błaszak, Prog.Theor.Phys. **83** (1990) 395.

[6] I.M.Gel'fand, I.Y.Dorfman, Funct.Anal.Appl. **13** (1979) 248; **14** (1980) 223.

[7] A.Weinstein, J.Diff.Geom. **18** (1983) 523.

[8] B.Fuchssteiner, Prog.Theor.Phys. **70** (1983) 1508.

[9] F.Magri, J.Math.Phys. **19** (1978) 1156.

[10] B.Fuchssteiner, Noninear Analysis TAM **3** (1979) 849.

[11] M.Błaszak, J.Phys.A.:Math.Gen. **22** (1989) 451.

[12] A.S.Fokas, P.M.Santini, Commm.Math.Phys. **115** (1988) 375; **116** (1988) 449.

[13] M.J.Ablovitz, H.Segur, "*Solitons and the Inverse Scattering. Transform*, Siam Studies in Applied Mathematics (SIAM, Philadelphia,1981).

[14] B.Fuchssteiner, R.N.Aiyer, J.Phys.A.:Math.Gen. **20** (1987) 375.

[15] G.Oevel, B.Fuchssteiner, Rev.Math.Phys. (1990)

[16] M.Błaszak, "*Symmetries, conservation laws and multi-soliton perturbation theory*", to be published (1991)

Skyrmions and Their Scattering in (2+1) Dimensions

Bernard Piette
Wojciech J. Zakrzewski
Department of Mathematical Sciences,
University of Durham, Durham DH1 3LE, England

Abstract: We consider the question as to whether instanton and anti-instanton solutions of the S^2 σ-model in two Euclidean dimensions modified by the addition of appropriate potential and skyrme-like terms can be considered as static solitons (and anti-solitons) of the same model in (2+1) dimensions. We study the stability properties of these field configurations, which we call skyrmions, and also look at their scatterings. We find that in contradistinction to the pure S^2 σ model the addition of the potential and skyrme terms stabilises the skyrmions. Looking at their scattering properties we find that the force between them is repulsive. In the scattering process initiated at low relative velocities the skyrmions bounce back while at large velocities they scatter at at right angles to the original direction in the centre of mass. The scattering is quasi-elastic and the skyrmions preserve their shape after the collision. We discuss the dependence of the scattering properties of the skyrmions on the parameters of the model and on the parameters of the initial two-skyrmion configuration. We also look at the scattering processes involving systems consiting of skyrmions and antiskyrmions and of a skyrmion and a pure wave. We find that the skyrmions and antiskyrmions attract each other and annihilate into pure radiation. We also report our results on the scattering of skyrmions with pure radiation waves.

1. Introduction

Over the last few years sigma models in low dimensions have become an increasingly important area of research. In two-dimensions they arise as approximate models in a particle physics and a solid state physics context, they are low dimensional analogues of 4-dimensional Yang-Mills theories, their two-dimensional classical solutions can be treated as static solutions of a (2+1) version of the theory, and they are interesting from a purely mathematical point of view. Although many σ-models are integrable in two dimensions [1-3] it appears that only very special ones are integrable in (2+1) dimensions.[4]

From the point of view of particle physics and relativity, the most interesting models are those which are Lorentz invariant. But as all Lorentz invariant σ-models in (2+1) dimensions appear to be nonintegrable, it is natural to consider numerical evolutions in these cases.

The simplest Lorentz invariant σ-model in (2+1) dimensions is the S^2 model, which involves one real vector field, $\vec{\phi} \equiv (\phi^1, \phi^2, \phi^3)$. In (2+1) dimensions $\vec{\phi}$ is a function of the space-time coordinates (t, x, y), which we will also write as (x^0, x^1, x^2). The model is defined by the Lagrangian

density

$$\mathcal{L} = \tfrac{1}{4}(\partial^\mu \vec{\phi}).(\partial_\mu \vec{\phi}) , \tag{1.1}$$

together with the constraint $\vec{\phi}.\vec{\phi} = 1$, *i.e.* $\vec{\phi}$ lies on a unit sphere S^2_ϕ. In (1.1) the Greek indices take values $0, 1, 2$ and label space-time coordinates, and ∂_μ denotes partial differentiation with respect to x^μ. Note that we have set the velocity of light, c, equal to unity, so that in all our calculations we can use dimensionless quantities. The Euler-Lagrange equations derived from (1.1) are

$$\partial^\mu \partial_\mu \vec{\phi} + (\partial^\mu \vec{\phi}.\partial_\mu \vec{\phi})\vec{\phi} = \vec{0} . \tag{1.2}$$

To specify the problem completely we have to state the boundary conditions on $\vec{\phi}$. Using polar coordinates (r, θ) we take

$$\vec{\phi}(r, \theta, t) \to \vec{\phi}_0(t) \quad \text{as} \quad r \to \infty , \tag{1.3}$$

where $\vec{\phi}_0$ is independent of the polar angle θ. In two Euclidean dimensions (*i.e.* taking $\vec{\phi}$ to be independent of time) this condition ensures finiteness of the action, which is precisely the requirement for quantisation in terms of path integrals. In (2+1) dimensions it leads to a finite potential energy.

The boundary condition (1.3) introduces nontrivial topological aspects into the theory, since it is equivalent to imposing a one-point compactification of the xy-plane into a two-sphere S^2_{phys}. Hence there are distinct topological sectors, labelled by the winding number of the map $\vec{\phi}(x^\mu)$ from S^2_{phys} to S^2_ϕ. This winding number, N, can be interpreted as an integer-valued topological charge and may be expressed as the integral of a charge density over all space:

$$N = \frac{1}{8\pi} \int \epsilon_{ij} \vec{\phi}.(\partial_i \vec{\phi} \times \partial_j \vec{\phi})\, d^2x , \tag{1.4}$$

where ϵ_{ij} is the antisymmetric symbol on two indices such that $\epsilon_{12} = -\epsilon_{21} = 1$.

It is convenient to introduce another formulation of the model; instead of using the $\vec{\phi}$ fields we can express all the dependence on $\vec{\phi}$ in terms their stereographic projection onto the complex plane W. The $\vec{\phi}$ fields are then related to W by

$$\phi^1 = \frac{W + W^\star}{1 + |W|^2} , \quad \phi^2 = i\,\frac{W - W^\star}{1 + |W|^2} , \quad \phi^3 = \frac{1 - |W|^2}{1 + |W|^2} , \tag{1.5}$$

where * denotes the complex conjugation.

The W formulation is very useful, because it is in this formulation that the static solutions take their simplest form; namely, as originally shown by Belavin and Polyakov[5] and Woo,[6] they are given by W being any rational function of either $z = x + iy$ or of $z^\star = x - iy$. It is easy to see that the topological charge of these solutions is a positive or a negative integer respectively. By convention the first case corresponds to instantons and the other to anti-instantons.

Let us concentrate on the instanton solutions. Of these the simplest nontrivial solution corresponds to one instanton and is given by

$$W = \lambda \frac{z - a}{z - b} , \tag{1.6}$$

where $z = x + iy$ and a, b and λ are arbitrary complex numbers. The appearance of 6 real parameters in the solution is a reflection of the conformal invariance of the effective two-dimensional model (under such a transformation the values of these parameters change but the general form of (1.6) remains the same). It is easy to calculate the energy density, E, corresponding to the static solution (1.6). We find

$$E = \frac{8|\lambda|^2\,|a - b|^2}{(|z - b|^2 + |\lambda|^2\,|z - a|^2)^2} \tag{1.7}$$

and so we see that the instanton has a bell-like shape, with its position and size determined by

$$\frac{a|\lambda|^2 + b}{|\lambda|^2 + 1} \quad \text{and} \quad \frac{|\lambda|^2\,|a - b|^2}{|\lambda|^2 + 1}$$

respectively.

Turning now our attention to the problem of the time dependence it is clear that we can consider all instanton solutions as static solutions of the same model in (2+1) dimensions. Can we have any nonstatic solutions? Of course, the static solutions can be made to move with arbitrary velocity, simply by Lorentz boosting. Being extended structures with localised energy, they resemble the familiar examples of solitons in (1+1) dimensions. But are they solitons in the strict sense? In particular, are they stable under small perturbations and do they preserve their shape and velocity in scattering processes?

These are some of the problems we will discuss in this talk. A part of our talk will be based on the work done in collaboration with R.A. Leese and M. Peyrard and which has been the subject of a series of papers.[7] [8] [9] [10]

Returning to the question of stability we observe that in many simple models there is an intrinsic scale, which essentially fixes the size of all solitons, but this is not the case in the S^2 model. In other words, the model admits the existence of solitons of arbitrary size. Hence it is possible that under small perturbations they could either expand indefinitely or they could shrink to become infinitely tall spikes of zero width. Our simulations have shown that this is exactly what happens in this model. In fact, as soon as the solitons are perturbed, *e.g.* start moving, they start shrinking. Such shrinking can even be induced by considering a static soliton on a lattice in the $\vec{\phi}$ formulation. (In this case the approximation of the real laplacian by its lattice version introduces a perturbation which destabilises the soliton).

We have analysed this problem in some detail and have found[7] that the solitons of the S^2 σ model are unstable. This is true not only in the full simulation of the model but also[11] in the approximation to the full simulation provided by the so-called "collective coordinate" approach. In this approach the real time evolution is approximated by the geodesic motion on the manifold of static solutions. Such an approximation is clearly very reliable at small velocities; however, all our studies have shown[11][7] that it is also reliable even at unexpectedly high velocities (~ 0.5 of the velocity of light).

Let us finish this section with a few words about our numerical procedures. Some of our simulations were performed in Los Alamos using a 4th order Runge-Kutta method of simulating the time evolution. We used the Los Alamos Connection Machine working in double precision and also some Los Alamos Crays. We also performed some calculations, using double precision, on the Floating Point System Machine and on Multiflow Trace. Almost all our simulations were performed on fixed lattices which varied from 201×201 to 512×512, with lattice spacing $\delta x = \delta y = 0.02$. The time step was 0.01. Some simulations were also performed, using double precision, on Durham Suns and the Durham HEP Vax.

We performed most of our simulations using the $\vec{\phi}$ formalism. In this case we have the problem of the constraint $\phi^i\phi^i = \vec{\phi}.\vec{\phi} = 1$. As we evolve all three components of ϕ^i separately, due the unavoidable numerical truncation errors introduced at various stages of the calculations, $\vec{\phi}$ gradually moves away from the unit sphere and the constraint is no longer satisfied. To overcome this problem, we rescaled all fields

$$\vec{\phi} \rightarrow \vec{\phi}\,' = \vec{\phi} \Big/ \sqrt{\vec{\phi}.\vec{\phi}} \tag{1.8}$$

every few iterations. In fact (just before the rescaling of $\vec{\phi}$) we evaluated $\mu_e = \vec{\phi}.\vec{\phi} - 1$ at each lattice point. If we treat $\mu_{\max} \equiv \max|\mu_e|$ as a measure of the numerical errors we have found that, in general, $\mu_{\max} \approx 10^{-4} \sim 10^{-9}$, depending on the process in question. On the other hand, we checked that when we used an unsound numerical procedure (like taking the time step in the Runge-Kutta procedure larger than the lattice spacing) this has always reflected itself in the rapid growth of $\mu_{\max}$. Hence we believe $\mu_{\max}$ to be a good estimate of the overall numerical accuracies and the error associated with this procedure is of the order of the accuracy of our calculations.

In addition, we know that the total energy and the topological charge are constants of motion and so our numerical procedure should exhibit this property. Of course, due to the numerical truncation errors we would expect both conservations to be slightly violated and, as is easy to check, without the rescaling of $\vec{\phi}$ the total energy increases (albeit very slowly at first). It is interesting to note that with the rescaling of $\vec{\phi}$ both the total energy and the topological charge exhibit small oscillations, but never alter much from their initial values (the errors are less than

1.00%). Hence we believe that our procedure of treating all three ϕ's as independent and then, every few iterations, rescaling them, avoids the build-up of numerical errors.

So far as the boundary conditions are concerned most of our simulations were performed with fixed boundary conditions as all the effects associated with the variation of the fields at the boundaries are very small. However, even though small, they are nonzero and so we tested their effects by introducing some absorption or by extrapolating the fields at the boundaries. We have found that the waves coming from the boundaries or the waves reflected from the boundaries can effect our results quite significantly. In particular, some preliminary results obtained on smaller size lattices, were not confirmed in our bigger lattice simulations. Having tested our results by changing the lattice size and varying the boundary conditions we are reasonably confident of our results; although we believe some more work would be required to be absolutely certain.

2. Skyrme Model

So how can we stabilise our solitons? Clearly, we have to introduce a scale into our model and to introduce it in such a way that this scale prevents the instantons from both shrinking or expanding. Guided by the ideas of Skyrme[12][13] we chose to add to our Lagrangian density the following extra terms

$$\mathcal{L}_I = -\tfrac{1}{4}\Big(\theta_1\Big((\partial^\mu\vec{\phi}.\partial_\mu\vec{\phi})^2 + (\partial^\mu\phi.\partial^\nu\phi)(\partial_\mu\phi.\partial_\nu\phi)\Big) + \theta_2(1+\phi^3)^4\Big), \tag{2.1}$$

where θ_1 and θ_2 are two new (real) parameters of the model. It is clear that the model based on the Lagrangian with these terms is still Lorentz invariant and for positive values of θ's its Hamiltonian is positive definite. Moreover, despite the appearance to the contrary, the Lagrangian does not contain time derivatives higher than two and so its equation of motion takes the conventional form. This equation is given by

$$\begin{aligned}\partial_\mu\partial^\mu\phi^i - (\vec{\phi}.\partial_\mu\partial^\mu\vec{\phi})\phi^i - 2\theta_1\Big[\partial_\mu\partial^\mu\phi^i(\partial_\nu\vec{\phi}.\partial^\nu\vec{\phi}) + \partial_\nu\phi^i(\partial_\mu\partial^\nu\vec{\phi}.\partial^\mu\vec{\phi}) \\ -\partial_\nu\partial_\mu\phi^i(\partial^\nu\vec{\phi}.\partial^\mu\vec{\phi}) - \partial_\mu\phi^i(\partial^\nu\partial_\nu\vec{\phi}.\partial^\mu\vec{\phi}) + (\partial_\mu\vec{\phi}.\partial^\mu\vec{\phi})(\partial_\nu\vec{\phi}.\partial^\nu\vec{\phi})\phi^i \\ -(\partial_\nu\vec{\phi}.\partial_\mu\vec{\phi})(\partial^\nu\vec{\phi}.\partial^\mu\vec{\phi})\phi^i\Big] + 2\theta_2(1+\phi^3)^3(\delta_{i3} - \phi^i\phi^3) = 0.\end{aligned} \tag{2.2}$$

The equation (2.2) is rather difficult to solve, but if we restrict ourselves to looking for static solutions and then consider $\vec{\phi}$, which corresponds to W being analytical (*i.e.* $W = W(z)$, $z =$

$x + iy$) then it is easy to check that

$$W = \lambda(z - a) \tag{2.3}$$

is a static solution if

$$\lambda = \lambda_0 = \sqrt[4]{\frac{\theta_2}{2\theta_1}}. \tag{2.4}$$

Observe that this is a particular case of the one instanton solution of the S^2 σ model (1.6), but with the fixed "size" (determined by λ_0). The total energy of the field configuration (2.3) is given by

$$E(\lambda) = 2\pi\Big(1 + \frac{4}{3}\theta_1\lambda^2 + \frac{2}{3}\frac{\theta_2}{\lambda^2}\Big). \tag{2.5}$$

This shows very clearly the stabilising nature of the new terms in the Lagrangian; the Skyrme term contains a factor λ^2 in the numerator and so prevents the skyrmions from being too spiky while the potential term contains the factor λ^2 in the denominator and so prevents the skyrmions from being too spread-out. For $\lambda = \lambda_0$ the energy (mass of the skyrmion) is

$$M = E(\lambda_0) = 2\pi\Big(1 + \frac{8}{3}\sqrt{\frac{\theta_1\theta_2}{2}}\Big). \tag{2.6}$$

It is easy to show that this configuration corresponds to a minimum of the potential in (2.1). Moreover, it is stable with respect to any small perturbation. In fact, if we try to evolve it with (2.4) different from λ_0 we find that the system has an excess of energy which it uses to bring its size up or down to the correct value and at the same time it reduces its energy by sending out a wave of radiation.

So what are the scattering properties of our skyrmions? First we looked at the behaviour of two static skyrmions. Thus we considered the field configuration described by

$$W = \frac{(z - a)(z + a)}{2\mu a}. \tag{2.7}$$

This configuration describes two skyrmions (located at $\pm a$); their widths are the same and are given by $\lambda_0 = \frac{1}{\mu}$. Moreover, as close to the position of each skyrmion *i.e.* for $z \sim a$ or $z \sim -a$ the W field $W \sim \frac{(z-a)}{\mu}$ and $W \sim -\frac{(z+a)}{\mu}$, respectively, we see that the phases of both skyrmions are opposite thus demonstrating that the angle of the relative orientation of the two skyrmions in the S^2 space is 180° . As (2.7) is not a solution of (2.2) it evolves and as it evolves the system develops some kinetic energy. Looking at this evolution we have found that the forces acting on the skyrmions are quite complicated; they have both repulsive and attractive components. At first the attractive forces win and the skyrmions approach each other. However, this does not last long; very soon the process is reversed, the skyrmions repel and move away from each other.

During this reversal the system performs some internal oscillations. In fact, in our simulations we observed two internal oscillations - which involved the kinetic energy flowing twice in the original direction of motion followed by its flow at 90° before the system stabilised and the skyrmions started moving away from each other. As they moved away they accellerated. Moreover, our studies have shown that the qualitative behaviour of the interaction does not depend on the values of θ's; as we increase their values all effects are the same but become more pronounced.

Next we looked at the scattering properties of two skyrmions sent towards each other at some velocity v. We implemented this idea by starting with the field configuration

$$W(x, y, t) = \frac{(z - a + vt)(z + a - vt)}{2\mu(a - vt)} \tag{2.8}$$

and then calculating from it $W(x, y, 0)$ and $\partial_t W(x, y, 0)$. We chose $a = 1.0$ and considered the dependence of the evolution on the values of v. Again, we found that for all values of θ's the qualitative properties of the scattering were the same. At small values of velocity the skyrmions scattered back to back. When we increased the initial velocity the skyrmions came closer and closer together before scattering back to back, then they spent longer and longer in a quasi-trapped state before bouncing back, and finally above a certain critical value of the velocity v_{cr} they scattered at 90° to the original direction of motion in their centre of mass. In fig 1 and 2 we present sets of pictures of the energy density for two such typical scattering cases; one at the speed below, the other at the speed above the critical velocity.

When the initial velocity is critical or higher the skyrmions manage to come very close together before scattering; at their closest they form a ring and it is from this ring that the outgoing skyrmions emerge. When the ring is formed the skyrmions loose their identity - hence it does not make sense to enquire which skyrmion goes where.

We also looked at the dependence of the critical velocity on the values of θ's. We found a strong dependence and in fig 3 we present our results. We observe an increase of v_{cr} with the increase of θ's and a sort of levelling off (or even a small decrease) at larger values of θ's.

One interesting question is the existence or not of the critical velocity for the pure S^2 model. Due to the shrinking of skyrmions in the pure S^2 model we cannot check this directly; hence we can only consider the limit of our curve for the vanishing values of θ's. The results suggest either a very low or no critical velocity but it is clear that we would have to go to much larger lattice sizes to be more definitive.

We have also looked at the "quasi-stable" state formed as the skyrmions move towards each other with their velocities approaching their critical value. We observed the lengthening of the time during which the skyrmions stayed together implying the trapped nature of the quasi bound-state formed by the skyrmions.

Finally, we looked at the time evolution of the initial configuration corresponding to two skyrmions put initially on top of each other. Such a configuration, for which the energy density describes a ring, is very reminiscent of the ring formed during the scattering process. The observed evolution showed small oscillations in the value of the size of the ring, followed, quite suddenly, by the separation of skyrmions and their motion away from each other with some finite velocity. We looked at this case in some detail and found that the velocities of the skyrmions moving away from each other were very close to the effective critical velocity (for the corresponding value of θ's). Hence we see that the mechanism of the ring formation and the possibility of having two skyrmions on top of each other have the same origin.

3. More General Two-Skyrmion Configurations

We also looked at the evolution of a more general initial condition for W describing two skyrmions, namely:

$$W(x,y,t) = \frac{\sqrt{1+R^2}}{2\mu}\frac{(z-a+vt)(z+a-vt)}{(a-vt+iRz)} \tag{3.1}$$

where R is a real parameter and μ is related to the θ's as in the previous section.

Like the configuration of the previous section this configuration corresponds to two skyrmions located initially at $x = \pm a$, moving towards each other. However, as close to the positions of the skyrmions the W field has the phase factor $exp(\pm i\varphi)$, where $tan(\varphi) = -R$ we see that the relative orientation of the two skyrmions in the S^2 space depends on R. Of course, when $R = 0$ (3.1) reduces to (2.8). Notice also that for $R \neq 0$, (3.1) is not symmetric with respect to the reflection along the x axis.

Moreover, the shape of the energy density of the system also depends on R; as R increases, the energy of the solitons increases slightly, but at the same time the energy is more spread out. It is easy to see that as $r = \sqrt{x^2+y^2}$ goes to infinity, the energy density goes to zero as a^2r^{-6} when $R = 0$ but only like r^{-4} for other values of R. As the local excess of energy decreases when R increases, the critical velocity increases and the scattering is much slower. For example, for $\theta_1 = 1/512$, $\theta_2 = 5$,'s and $a = 1$, $v_{cr} = 0.40$ for $R = 0$ and $v_{cr} = 0.64$ for $R = 1$. For larger values of v, as they collide, the skyrmions merge to form a ring from which they emerge as two skyrmions of different sizes. The ring, however, is not symmetric; it is still circular in shape but it is higher on one side. In fig 4 we present a few pictures of the energy density seen in a simulation for $R \neq 0$ for a velocity above its critical value. As the skyrmions separate they oscillate in size until they both reach the correct size for a single skyrmion.

One can wonder whether the 90° scattering we have observed is not a lattice artifact. After all, we use a regular lattice and most of our simulations exhibited some symmetries which could have forced the scattering to be back to back or only at 90°. However, the more general field configuration discussed above did break the symmetry and the scattering was still qualitatively the same. To check it further we decided to add some random noice to the original configuration (2.7) and to perform it for a velocity just above its critical value. Then, if the scattering proceeded through the formation of a fireball-like structure followed by its decay the randomnes could lead to the arbitrary direction of the decaying structure. Our simulation produced, however, a scattering at 90° suggesting once again the general nature of this phenomenon.

4. Ring Formation and Its Effect

Some understanding of our results can be gained by thinking more seriously about the mechanism of the scattering; *i.e.* that the scattering at 90° is only possible if it is preceded by the formation of a ring. This suggests a discussion based on the energy balance; namely the initial skyrmions must have enough energy to form the ring in order to be able to scatter at 90°. Such a discussion provides a lower bound on the value of the critical velocity and at the same time, if successful, provides a further support for the mechanism of the scattering. The discussion is quite complicated for the most general field configuration; it becomes quite simple if one neglects all effects associated with the finite separation between initial skyrmions (which are expected to be rather small). Such a discussion was performed by Sutcliffe[14] . In his estimate Sutcliffe took the ring field configuration in the form

$$W = \frac{1}{\lambda_1}z^2 \tag{4.1}$$

and then minimised the energy by changing λ_1. He found that the minimum was obtained for

$$\lambda_1^2 = \lambda_r^2 = \frac{(2\lambda_0)^4}{5}, \tag{4.2}$$

where λ_0 is the canonical size of the skyrmion and so is given by (2.4). For that value of λ_1 the energy of this ring structure is given by

$$E_r = 4\pi\Big(1 + 2\pi\sqrt{\frac{5\theta_1\theta_2}{32}}\Big). \tag{4.3}$$

On the other hand E_r, the energy of two skyrmions at infinity, each of the canonical size and moving with velocity v, is $E_v = 2M\gamma$, where M is the static energy of each skyrmion (2.6), and γ is the usual Lorentz factor $\gamma = (1+v^2)^{\frac{1}{2}}$. Then the equality of E_r and E_v gives a lower bound on the critical velocity

$$v_{cr}^2 \leq \frac{\Theta(\alpha-\beta)(2+(\alpha+\beta)\Theta)}{(1+\beta\Theta)^2}, \tag{4.4}$$

where $\alpha = \frac{\pi\sqrt{5}}{2}$, $\beta = \frac{8}{3}$ and $\Theta = \sqrt{\frac{\theta_1\theta_2}{2}}$. Looking at the expression for (4.4) we see that v_{cr} increases monotonically with Θ approaching $v_{cr} \sim 0.86$. This is in qualitative agreement with

our results on the dependence of v_{cr} as a function of θ_2 (shown in fig 3), in which we altered both θ_1 and θ_2 to keep λ fixed to compare the scattering of skyrmions of the same size.

We have also looked at scattering in which we kept θ_2 fixed and altered θ_1 and so also λ_0. In fig 5 we present our results. They show a steady increase followed by a sudden jump. If we look in detail at the behaviour of our skyrmions as they come together we find that as θ_1 increases the behaviour of our skyrmions becomes less and less regular. For small values of θ_1 (corresponding to $\lambda = .2$ to 0.5) and for velocities close to the critical velocity, the skyrmions, after the original slowing down, approached each other with a reasonably constant deceleration; for larger values of θ_1 they changed their speed more often. This is partly responsible for the sudden increase of v_{cr} for $\lambda_0 = 0.8$. The sudden change of speed is partly due to the wave emitted by one skyrmion hitting the other one. In fig 6 we present the plots of the time dependence of the distance of the skyrmions from $x = y = 0$ obtained for two values of λ for the velocity just above its critical value.

The calculation of Sutcliffe reproduced the main features of our results quite well. However, his results refer to skyrmions initially at infinity. Our results refer to skyrmions initially located at $x = \pm 1$, $y = 0$ Clearly there is some initial interaction between the skyrmions, and moreover, our configurations are not optimal resulting in some initial loss of energy required for the rearrangement. In addition, the overlap between the skyrmions increases with λ_0. Thus a more realistic estimate of v_{cr} would involve the calculation starting with skyrmions located at $x = \pm 1$, $y = 0$. This has to be done numerically (to calculate their energies); then assuming that the total difference between the energy of the ring and the energy of the skyrmions corresponds to the kinetic energy of motion of the skyrmions gives us, again, a bound on the critical velocity. Our results are given in fig 7. They correspond to the lower bound but they are not that different from the velocities observed in the simulations. This supports our interpretation of the mechanism. The results disagree more for larger values of λ_0. However, as we said before, due to the overlap of the skyrmions and the irregularity in their motion the simple picture begins to break down for those values of λ_0. Recall that as λ_0 increases so does increase the total energy and also the contribution of the skyrme and potential terms. Hence we conclude that our understanding of the mechanism of scattering is more reliable for the configurations for which the corrections to the original S^2 model are not too large.

Can we explain qualitatively our results on the dependence of the critical velocity of the skyrmions on their relative orientation in the S^2 space? The discussion of Sutcliffe is of no help here as the energy of well separated skyrmions does not depend on their orientation. Moreover, our discussion must include the observation that when $R \neq 0$ the ring the skyrmions form is not symmetrical. To do this we chose to repeat our previous calculation but this time comparing our results with the energy of the ring given by

$$W_{mr} = \kappa \frac{z^2}{a + iRz}, \tag{4.5}$$

i.e. corresponding to a ring whose shape is determined by the initial configuration (a in the denominator in (4.5)). This choice of the ring may be incorrect but it should produce some information of more qualitative nature. Our results showed the increase of the critical velocity with R; we found (for $\lambda = 0.25$) that as $v_{cr} = 0.2095$ for $R = 0$ its value increases to $v_{cr} = 0.2339$ for $R = 1$ and $v_{cr} = 1.043$ for $R = 10$, in a qualitative agreement with the results of our numerical simulation.

We also looked at the R dependence of the initial motion of two skyrmions placed at $x = \pm 1$, $y = 0$ and released from rest. We know they repel but the detailed investigation of the numerical results showed that for $R = 0$ they originally attracted each other and only repelled after some rearrangement, while for large values of R we had also initial repulsion. To look at this initial motion we assumed that our skyrmions have some accelleration; hence we parametrised their positions by $\pm(a - b\frac{t^2}{2})$ and tried to determine b from the equation of motion. Of course, we could not do this at all values of the field; we looked only at the positions of our skyrmions and so calculated b only at $z = \pm a$. We found that b depends on R, being a positive real number for $R = 0$, becoming complex for small R and ending up a negative real number for large R, in complete agreement with our numerical simulations. This supports once again the notion that some understanding can be gained from treating skyrmions as extended objects whose parameters are determined by the geodesic approximation.

5. More General Systems

We also looked at systems consisting of one skyrmion and one antiskyrmion. In this case we considered as our initial configuration

$$W = \frac{(z-a)(z^\star + a)}{2\mu a}. \tag{5.1}$$

It is easy to check that (5.1) is not a solution of the equation of motion. Moreover, the system is clearly unstable and when started at rest the two extended structures approached each other and annihilated into pure radiation. During their approach the skyrmion and antiskyrmion attracted each other and so accellerated while moving towards each other. In fig 8 we present some pictures of the total energy density for a few representative values of time. We observe that before their annihilation the skyrmion and antiskyrmion preserved their identities very well. After the interaction the system represented just pure waves; what is interesting is that their maxima flowed at 90° to the original direction of motion. It is easy to check that the outgoing structures really represented pure radiation waves and not skyrmions and antiskyrmions; it is enough to observe that they moved with the velocity of light and that their topological charge gradually decreased. In fig 9 we present a plot of the time dependence of the position of the maximum of the energy density obtained in our simulation. Observe that this maximum originally refers to

the skyrmions and later to the radiation. It is easy to check that the radiation moved with the velocity of light.

Later we introduced a nonzero velocity v and performed several simulations varying this initial value of v and changing the values of θ_2. We found no significant dependence on θ's showing that for a system of skyrmions and antiskyrmions, which is characterised by strong attractive forces already at the S^2 level, the additional forces generated by the potential and skyrme terms in the Lagrangian have little effect on the main features of the scattering. We performed several others simulations including the interesting case of a skyrmion and an antiskyrmion rotating around each other. In this last case we found that the skyrmion and the antiskyrmion had almost got trapped in an orbit around each other; however, due to their interaction, at a certain time they slowed down their circular motion, moved towards each other and then annihilated into pure radiation. The maximum of the radiation was again sent out at 90° to the direction motion just before the annihilation.

This last case suggested a serious look at the interaction of skyrmions with waves of radiation. We have very little time left to discuss this case but let us point out that we created our wave by perturbing the initial condition for the time derivative of the field. Our simulations have given us very interesting results; in particular we found that as the initial wave hit the skyrmion it modified skyrmion's shape and that the perturbed skyrmion emited a circular wave. Looking at the scattering in more detail we observed that the skyrmion had absorbed a part of the incident wave and had left a gap in the front wave which, within the limits of our integration, propagated without being "filled out". We also observed that during the interaction, the skyrmion was at first pushed forward, then pulled back very quickly and as the wave left the skyrmion, the skyrmion begun to move backwards (*i.e.* towards the direction where the wave came from). Moreover, it moved in this direction at a constant speed. The observed motion of the skyrmion is a little unexpected, but as the skyrmion emitted a wave, this wave carried the momentum which counterbalanced the momentum of the moving skyrmion. Also, we observe that the skyrmion absorbed some energy of the wave thus explaining the appearance of the gap left behind in the wave front.

6. Some General Comments

We have seen that the skyrmions behave very much like real solitons. In the scattering involving only skyrmions they preserve their shape and although during the scattering some radiation effects are present, these effects are always very small. The situation is different, however, for systems involving skyrmions and anti-skyrmions; they interact with each other very strongly and annihilate into pure radiation. The outgoing radiation follows the scattering of skyrmions above their critical velocity; namely the peaks of radiation are sent out at 90° to the original direction of motion in the centre of mass. This phenomenon has also been observed in many scatterings of other extended structures, such as monopoles or vortices[15] and so we believe that it is probably very typical of all scattering of extended structures in (2+1) dimensions.

Of course our results depend quite crucially on the specific form of the skyrme and potential term in (2.1). Recently[16] we have started looking also at some other forms of the potential term. In particular, we have found that a simple modification of this term can give a model which has static configurations corresponding to two separated skyrmions. The preliminary results have exhibited many interesting phenomena; however, when both skyrmions were sent towards each other with sufficiently high velocity they still scattered at 90°. Thus we believe that this type of scattering is generic to (2+1) dimensional solitons.

In conclusion we see that the modified S^2 model, although non-integrable, is almost integrable in that it has many features in common with many integrable models. Most differences or deviations are rather small. As most physically relevant models are not integrable our results suggest that the results found in some integrable models should not be dismissed as not relevant; it is quite likely that some of these results may also hold in models which, strictly speaking, are not integrable but whose deviations from integrability are rather small.

Most of our results agree with the results obtained in the collective coordinate approximation. This suggests that this approximation is much better than could be first thought of on purely general grounds. Finally, our results suggest that the modified S^2 σ model is a good candidate for being a toy model of solitons in (2+1) dimensions.

ACKNOWLEDGEMENTS

A large part of the work reported in this paper was carried out using various numerical simulation programs developed jointly with Michel Peyrard when we were all visiting the Center for Nonlinear Studies at the Los Alamos National Laboratory. We wish to thank Michel Peyrard for his collaboration and David Campbell for the hospitality at the Center.

Some numerical simulations reported here were carried out on the Durham HEP Vax and on Durham Suns. We wish to thank Mike Whalley and Peter Craig for their help and advice in the implementations of our programs.

We also wish to acknowledge the support of NATO - grant SA.9.15. (922/88).

REFERENCES

1. Y.Y. Goldschmidt and E. Witten, *Phys. Lett*, **91 B**, 392 (1980)
2. V.E. Zakharov and A.V. Mikhailov, *Sov. Phys. JETP* **47**, 1017 (1979)
3. J. Harnad, Y. Saint-Aubin and S. Shnider, *Comm. Math. Phys.*, **92**, 329 (1984)
4. R.S. Ward, *Nonlinearity*, **1**, 671 (1988)
5. A.A. Belavin and A.M. Polyakov, *JETP Lett*, **22**, 245 (1975)
6. G. Woo *J. Math. Phys.*, **18**, 1264 (1977)
7. R.A. Leese, M. Peyrard and W.J. Zakrzewski, *Nonlinearity* **3**, 387 (1990)
8. R.A. Leese, M. Peyrard and W.J. Zakrzewski, *Nonlinearity* **3**, 773 (1990)
9. W.J. Zakrzewski - Soliton-like Scattering in the $O(3)$ σ model in (2+1) Dimensions - *Nonlinearity*, to appear, (1990)
10. M. Peyrard, B. Piette and W.J. Zakrzewski - Soliton Scattering in the Skyrme Model in (2+1) Dimensions 1. and 2., Durham University preprints DTP-90/37 and 39 (1990)
11. R.A. Leese, Low-energy Scattering of solitons in the CP^1 model, *Nucl. Phys.* **B**, to appear (1990)
12. T.H.R. Skyrme - *Proc. Roy. Soc.***A 260**, 127 (1961)
13. E. Witten - *Nucl. Phys.* **B 223**, 433 (1983), G. Adkins, C. Nappi and E. Witten - *Nucl. Phys.* **B228**, 552 (1983)
14. P.M. Sutcliffe - The Interaction of Skyrme-like Lumps in (2+1) Dimensions; Durham preprint DTP-90/25 (1990)
15. for an extended discussion see ref. 11
16. B. Piette and W.J. Zakrzewski - work in preparation

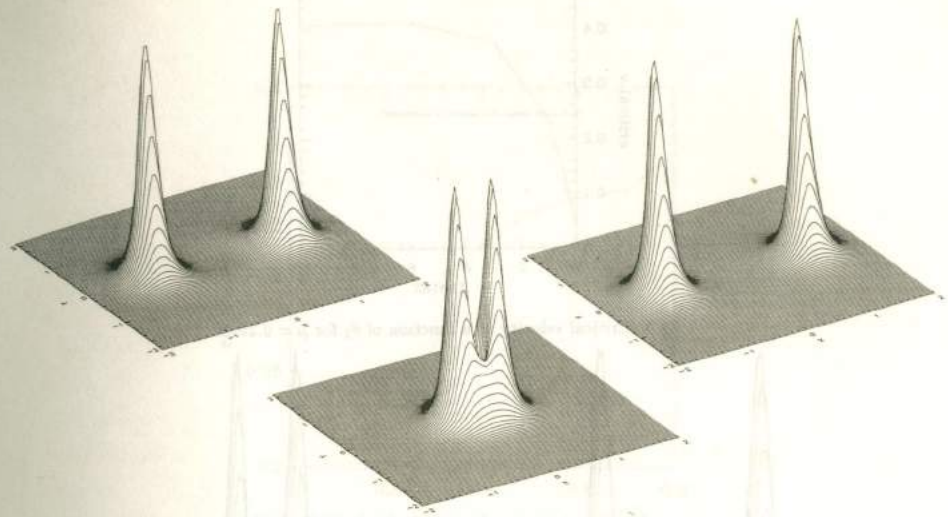


fig 1: Back to back scattering of two skyrmions below v_{cr}

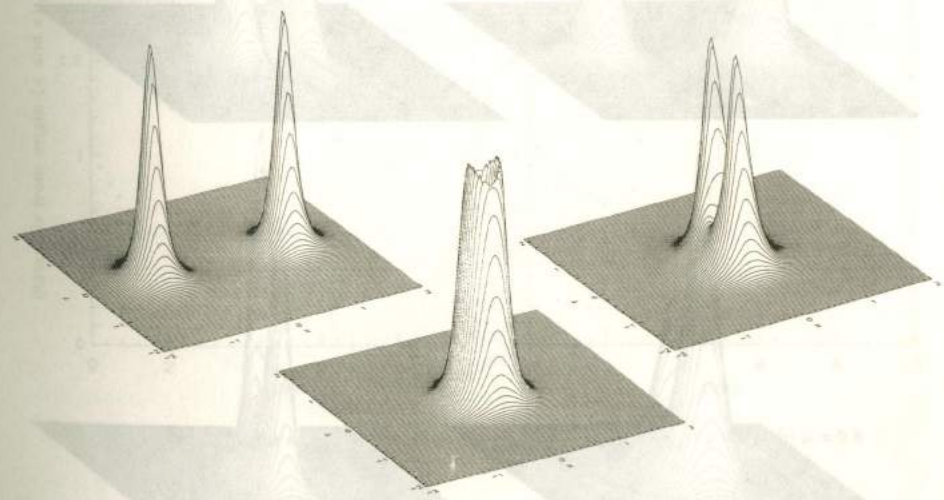


fig 2. 90 degrees scattering of two skyrmions above v_{cr}

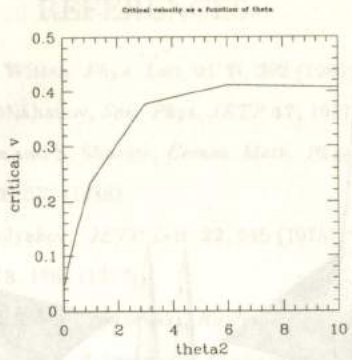


fig 3. Critical velocity as a function of θ_2 for $\mu = 0.25$.

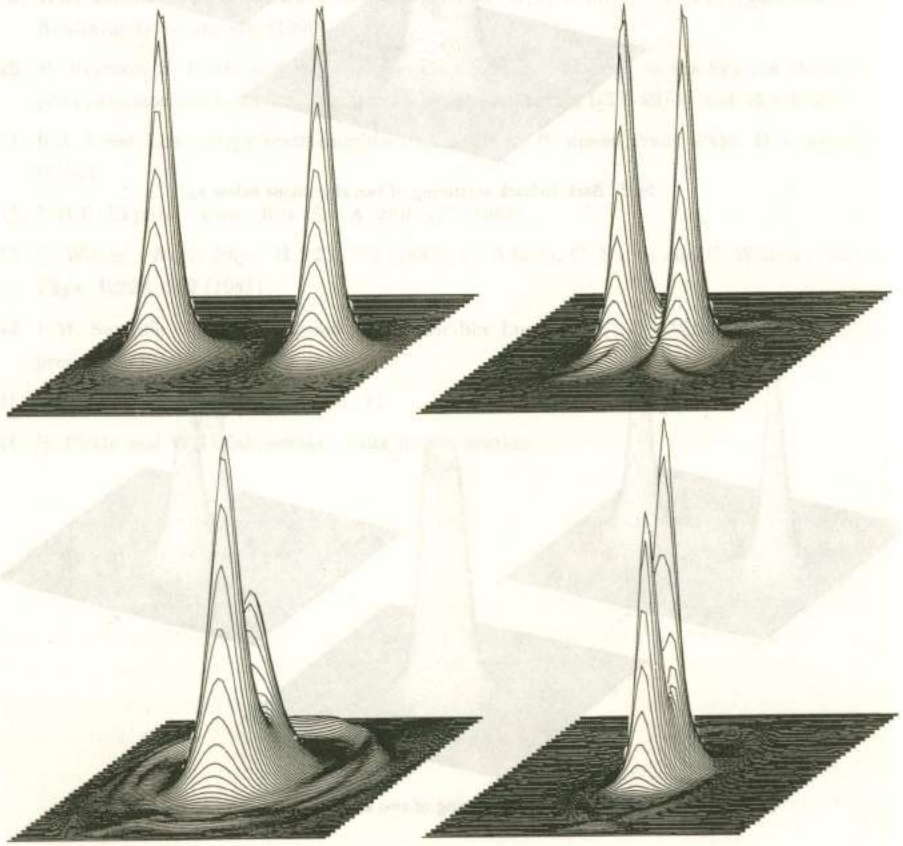


fig 4. Energy density for the modified initial condition $R = 1$, $\mu = .25$, $v = .6$

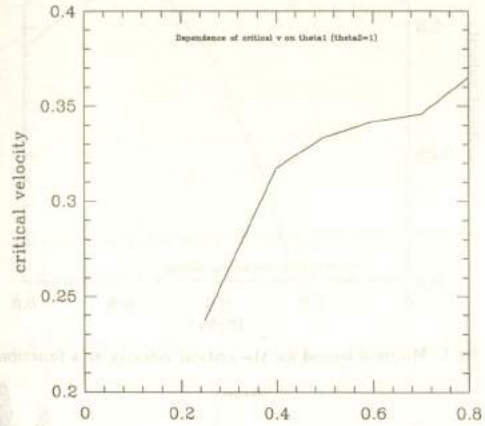


fig 5. Critical velocity as a function of μ, $\theta_2 = 1$

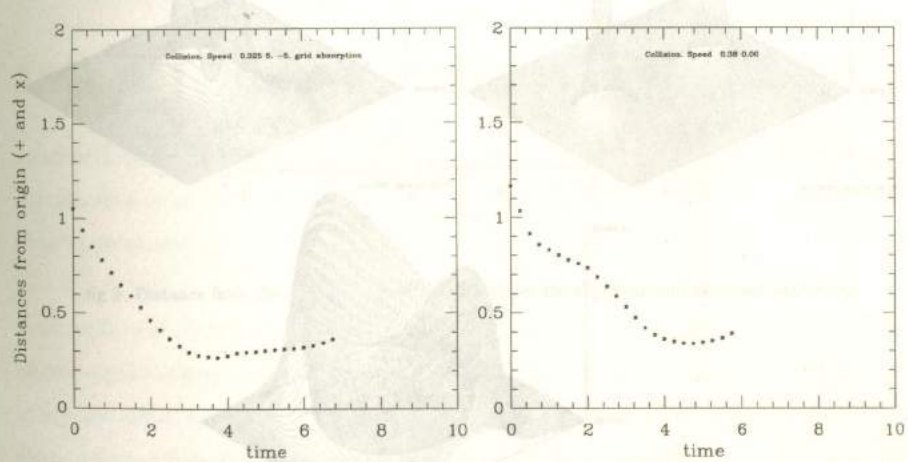


fig 6. Distance of the skyrmion from the origin as a function of time. a) $\mu = 0.7$, b) $\mu = 0.8$

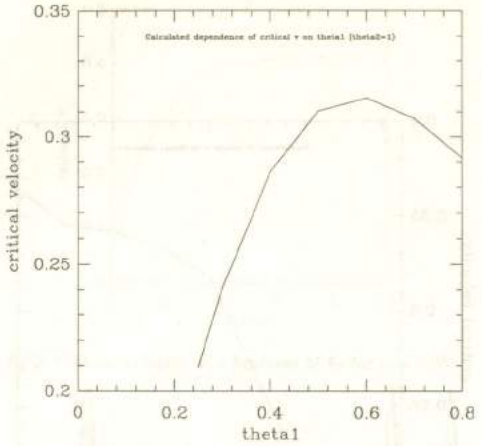


fig 7. Minimal bound for the critical velocity as a function of μ.

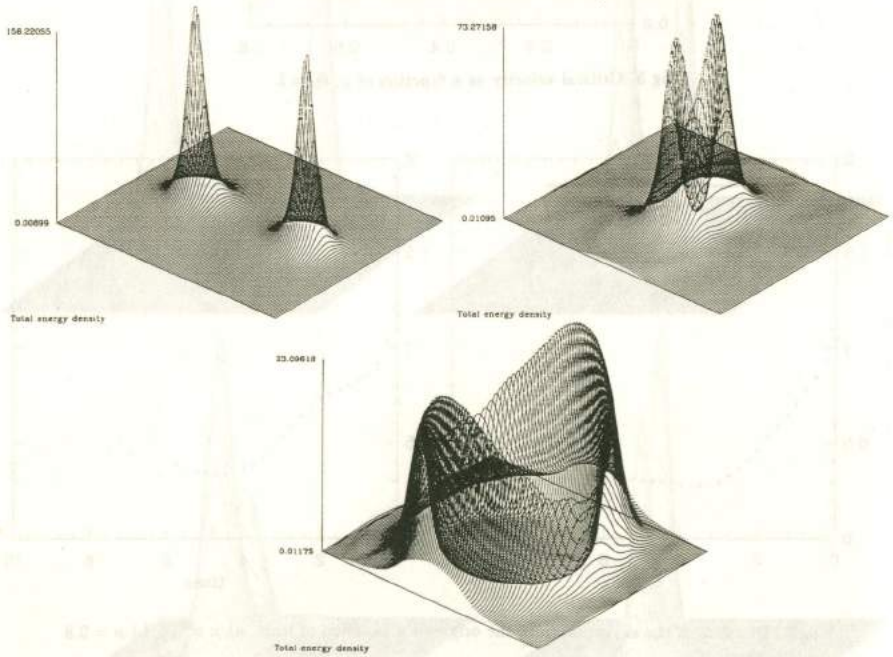


fig 8. Total energy density for a skyrmion anti-skyrmion scattering.

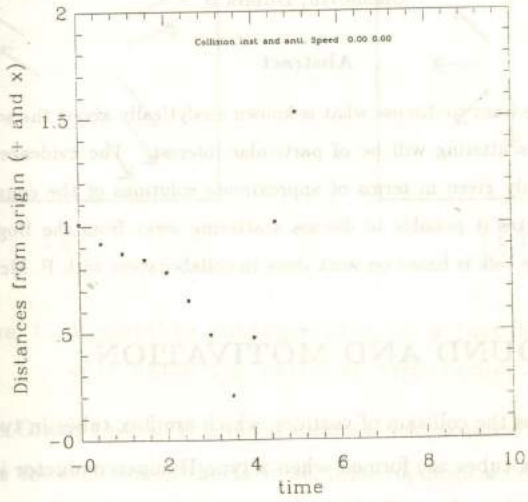


fig 9. Distance from the origin as a function of time for the skyrmion anti-skyrmion scattering.

VORTEX-VORTEX SCATTERING

J. Burzlaff
School of Mathematical Sciences
Dublin City University
Glasnevin, Dublin 9

Abstract

In this talk we want to discuss what is known analytically about the scattering of vortices. 90° scattering will be of particular interest. The evidence for 90° scattering is mainly given in terms of approximate solutions of the equations of motion. This makes it possible to discuss scattering away from the Bogomolnyi limit. Most of this talk is based on work done in collaboration with P. McCarthy.

1 BACKGROUND AND MOTIVATION

We are going to discuss the collision of vortices, which are flux tubes in type - II superconductors. These flux tubes are formed when a type-II superconductor in a magnetic field is cooled down below the critical temperature. A type-I superconductor would then expel the magnetic field whereas the magnetic field can penetrate the type-II superconductor and does so in the form of small flux tubes of quantized flux. The flux tube sustains (and is itself sustained by) a supercurrent of Copper pairs flowing around the flux tube. This microscopic picture shows that vortex is an appropriate name for the phenomenon. It also shows that in order to produce a smooth configuration, the superconductor should be in the normal state at the center of the vortex.

In the framework of the Ginzburg-Landau theory the vortex can easily be understood as a topological defect[1]. In this theory a complex field ϕ is the density amplitude for

Cooper pairs. In the normal state it is zero and below the critical temperature it takes a nonzero value, which can be represented as an arrow whose direction represents the argument of the complex number. Just below the critical temperature, due to fluctuations, these directions may be chosen independent of each other in different regions of space. These regions only grow together and smooth out well below the critical temperature. This can lead to topological defects which are the vortices (see fig. 1). Since

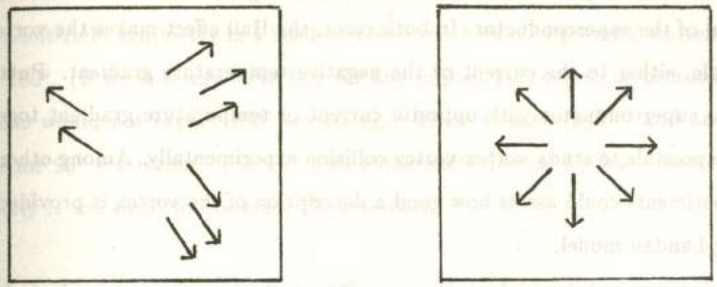


Figure 1. A possible configuration of ϕ just below and well below the critical temperature.

also in certain Grand Unified Theories a complex field acquires a nonzero value at a certain time in the evolution of the universe, these topological defects can again occur in cosmological models and are called cosmic strings[2].

Due to its topological nature, the vortex has a stability which makes it a soliton-like object. Since solitons show very interesting behaviour during collisions, it is natural to study the collision of soliton-like objects. This investigation has been started. Most of the analytical work is based on an idea by Manton[3] for the collision of slowly moving monopoles. Using Manton's idea, in particular Atiyah and Hitchin[4] have made important contributions to our understanding of monopole-monopole scattering. Manton's technique was also used on other soliton-like objects, e.g. by Ward[5] on moving CP^1

instantons and by Ruback[6] on vortices (strings). Furthermore, a great amount of numerical work has been done, in particular on skyrmions[7] and cosmic strings[8].

What distinguishes vortices from the other soliton-like objects is that they can be produced in the laboratory. They also can be set in motion[9], either by a current orthogonal to the vortex or a temperature gradient. In the first case, the vortex is moved by the Lorentz force, in the second case, the entropy is lowered if the vortex moves to the cooler end of the superconductor. In both cases, the Hall effect makes the vortex move at an angle, either to the current or the negative temperature gradient. Putting two parts of a superconductor with opposite current or temperature gradient together, it should be possible to study vortex-vortex collision experimentally. Among other things, these experiments could assess how good a description of the vortex is provided by the Ginzburg-Landau model.

From a theoretical point of view, one expects the scattering process for vortices to be simpler than the scattering of monopoles. In the case of vortices, there is a $2n$-parameter family of static n-vortex solutions, which means that the position of the vortices are the only degrees of freedom, and the behaviour during a collision does truly reflect the soliton-like nature of vortices. Monopoles, on the other hand, have further degrees of freedom, internal phases, and different types of scattering processes are produced by

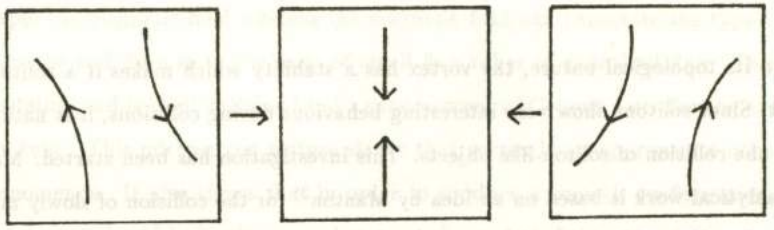


Figure 2. Head-on collision as the limit of collisions with nonzero impact parameter.

changing these inner degrees of freedom. This implies we have more confidence in general arguments which predict the outcome of a vortex-vortex collisions.

A very simple argument of this type rules out scattering at angles other than 0°, 90° and 180°. As fig. 2 illustrates, a head-on collision can be considered as the limit of a sequence - and of its mirror image - of collisions with nonzero impact parameter. This leads to a left-right symmetry in a head-on collision which only permits scattering at 0°, 90° and 180°. (If there is any deflection for any impact parameter, as presumed in fig. 2, one also would not expect 180° scattering.) There is numerical[10] and analytical[6] evidence for 90° scattering. There is also numerical evidence for back scattering[10] if the velocity is too small to overcome the repulsion.

The analytical evidence given by Ruback[6] is based on Manton's idea[3] for the investigation of slowly moving monopoles, which can be applied to other soliton-like objects. Manton's technique starts with the $(4n-1)$-parameter solution $F(s_1,\cdots,s_N)$ for n static monopoles in the Bogomolnyi-Prasad-Sommerfield limit. Time dependence is introduced by making the parameters time dependent. The time dependent configurations $F(s_1(t),\cdots s_N(t))$ are then used in the Lagrangian to find the term of the form $\frac{1}{2}g_{ij}\dot{s}_i\dot{s}_j$ which defines the metric g_{ij} in the parameter space. (In most applicatons this procedure is not literally followed. Instead, symmetry arguments are used to determine the metric.) It is argued that the geodesics in the parameter space describe the motion of n slowly moving monopoles.

Ruback studies the symmetries of the metric of the $2n$-parameter space of static vortex solutions in the Bogomolnyi limit and finds evidence of 90° scattering. We will formulate his arguments in such a way that we can extend them to the case of small repulsion between the vortices. We find that 90° scattering persists. This is not an obvious result, since, although the repulsion is small, so is the velocity which has to overcome

the repulsion. Possible extensions we do not discuss are the introduction of a relative angle between the colliding cosmic strings (vortices) and an increase in velocity. It can be argued that 90° scattering for nonparallel strings leads to intercommutation. To our knowledge, no progress has been made to find higher order corrections in the velocity to the approximation for slowly moving soliton-like objects.

2 EVIDENCE FOR 90° SCATTERING

To describe the dynamics we use the Ginzburg-Landau model

$$\mathcal{L} = \frac{1}{2}(D_\mu\phi)(D^\mu\phi)^* - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{8}\lambda(\phi\phi^* - 1)^2, \tag{1}$$

where the covariant derivative $D_\mu\phi$ and the electromagnetic field $F_{\mu\nu}$ are given by

$$D_\mu\phi = \partial_\mu\phi - iA_\mu\phi, \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{2}$$

for $\mu,\nu = 0,1,2$ and metric is diag $(+,-,-)$. The corresponding Euler-Lagrange equations are

$$D_\mu D^\mu\phi - \frac{1}{2}\lambda\phi(\phi\phi^* - 1) = 0, \tag{3}$$

$$\partial_\mu F^{\mu\nu} + \frac{i}{2}[\phi^* D^\nu\phi - \phi(D^\nu\phi)^*] = 0. \tag{4}$$

Our aim is to find an approximate solution for $t \in (-\varepsilon,\varepsilon), \varepsilon << 1$, i.e. for the time from shortly before to shortly after the collision, which shows evidence of 90° scattering. Since there is only a choice between 0°, 90° and 180° scattering, we hope to be able to decide which case is realised, by studying the two vortices when they are close together.

We start with the static, finite-energy 2-vortex solution[11], which describes two vortices sitting on top of each other:

$$\phi(r,\theta) = e^{i2\theta}f(r), \qquad A_0(r,\theta) = 0,$$

$$A_j(r,\theta) = -\frac{2}{r}\varepsilon_{jk}x_k a(r), \quad j,k = 1,2, \tag{5}$$

where f and a satisfy the equations

$$r(rf')' - 4f(a-1)^2 - \frac{1}{2}r^2\lambda f(f^2-1) = 0,$$

$$\left(\frac{a'}{r}\right)' - \frac{f^2}{r}(a-1) = 0, \tag{6}$$

and the conditions

$$f(0) = a(0) = 0, \quad \lim_{r\to\infty} f(r) = \lim_{r\to\infty} a(r) = 1. \tag{7}$$

We now introduce two types of approximations. First, we seek solutions for $\lambda = 1 + \tilde{\lambda}$ where $0 < \tilde{\lambda} << 1$ which implies that we can linearize in $\tilde{\lambda}$. (The solutions (5) to equations (3) and (4) are analytic in $\tilde{\lambda}$.[12]) In addition, we introduce the time dependence through the ansatz

$$\phi(t,\vec{x}) = \hat{\phi}(\vec{x}) + \tilde{\phi}(t,\vec{x}) = \hat{\phi}(\vec{x}) + \tilde{\lambda}\varphi(\vec{x}) + t\xi(\vec{x}),$$

$$A_i(t,\vec{x}) = \hat{A}_i(\vec{x}) + \tilde{A}_i(t,\vec{x}) = \hat{A}_i(\vec{x}) + \tilde{\lambda}a_i(\vec{x}) + tB_i(\vec{x}),$$

$$A_0(t,\vec{x}) = 0. \tag{8}$$

Here, $(\hat{\phi},\hat{A}_i)$ are the solutions (5) for $\lambda = 1$. Since we consider solutions only for $t \in (-\varepsilon,\varepsilon)$ where $0 < \varepsilon << 1$, we can linearize in ξ and B_i. Furthermore, because all finite energy solutions are solutions of the Bogomolnyi equations[13),14)], ξ and B_i can be found by solving the Bogomolnyi equations and the equation (2.4) for $\nu = 0$, where all equations are linearized in ξ and B_i.

Weinberg[15] and Ruback[6] have solved these equations and found the solutions

$$\xi = 2(\alpha + i\beta)e^{i\theta}f'(r),$$

$$B_1 + iB_2 = \frac{2}{rf^2}(\beta - i\alpha)(rff'' - rf'^2 + ff'). \tag{9}$$

These solutions are of the form

$$\hat{\phi}(\vec{x}) + t\hat{\xi}(\vec{x}) = e^{i\chi}\hat{\phi}(x + 2t\alpha, y + 2t\beta),$$

$$\hat{A}_i(\vec{x}) + tB_i(\vec{x}) = \hat{A}_i(x + 2t\alpha, y + 2t\beta) + \partial_i\chi, \tag{10}$$

where

$$\chi = t\left(\frac{2f'}{f} - \frac{4}{r}\right)(\beta\cos\theta - \alpha\sin\theta), \tag{11}$$

to first order in t. This means the solutions (9) are gauge transformed and spatially translated versions of the solution (5).

More interesting are the other solutions Weinberg and Ruback found:

$$\xi = 2f(r)k(r), \quad B_1 + iB_2 = \frac{-2i}{r}e^{-i\theta}[rk'(r) + 2k(r)], \tag{12}$$

and

$$\xi = 2if(r)k(r), \quad B_1 + iB_2 = \frac{2}{r}e^{-i\theta}[rk'(r) + 2k(r)]. \tag{13}$$

Here

$$-\frac{1}{r}(rk')' + (f^2 + \frac{4}{r^2})k = 0, \tag{14}$$

$$rf' - 2(1-a)f = 0, \tag{15}$$

$$\frac{4}{r}a' + f^2 - 1 = 0. \tag{16}$$

The ordinary differential equations (14)-(16) cannot be solved in closed form. Series solutions to these equations are discussed in the appendix and in detail in ref. 16.

Finite-energy solutions to eq. (14) go like e^{-r} (after normalisation) at infinity and like $c_1 r^{-2} + c_2 r^2$ at the origin. In fact, the r^{-2} term must be present, because its absence would imply the existence of a point r_0 with $k(r_0) > 0, k'(r_0) = 0$ and $k''(r_0) \leq 0$. This is not consistent with (14). Therefore, k is strictly monotonic decreasing from infinity to zero, as r increases, and there exists a point $r = \rho > 0$ such that $k(\rho) = 1/(2|t|)$. For the solution (12) we have

$$|\phi|^2 = f^2(1 + 4 - k\cos 2\theta + 4t^2k^2)$$

$$\geq f^2(1 - 2|t|k)^2. \tag{17}$$

The zeros of the Higgs field, which give the location where the magnetic field penetrates the superconductor, are therefore at $r = \rho, \theta = 0$ and $\theta = \pi$ for $t < 0$, and at $r = \rho, \theta = \frac{\pi}{2}$ and $\theta = \frac{3\pi}{2}$ for $t > 0$. This solution describes 90° scattering. A similar argument shows that the solution (13) describes 90° scattering as well. Further evidence of 90° scattering is given by the form of the energy density plotted for $t = -1, t = 0$ and $t = +1$ in fig. 3. Of course, only for $t \in (-\varepsilon, \varepsilon), \varepsilon << 1$, this is the energy density of an approximate solution. Ruback[6] has described this 90° scattering in terms of geodesic motion in the parameter space of solutions.

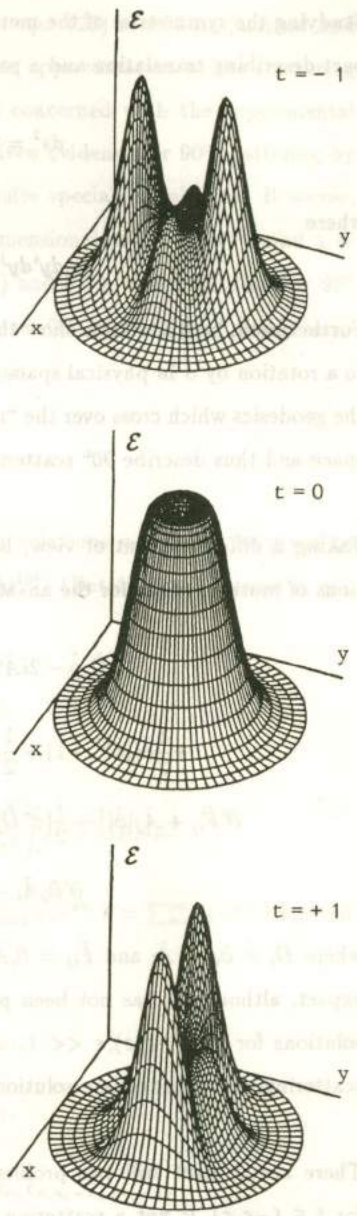


Figure 3. The energy density for the configuration (12).

Studying the symmetries of the metric he has shown that the metric can be split into a part describing translation and a part describing relative motion:

$$ds^2 = 2d\vec{X} \cdot d\vec{X} + g_{ij}dy^i dy^i, \tag{18}$$

where

$$g_{ij}dy^i dy^j = h^2(\rho)d\rho^2 + g^2(\rho)d\psi^2. \tag{19}$$

Furthermore, he was able to show that a rotation by 2α in parameter space corresponds to a rotation by α in physical space. Since the metric is that of a "snub-nosed" cone[17], the geodesics which cross over the "nose" lead to a change of angle by 180° in parameter space and thus describe 90° scattering in physical space.

Taking a different point of view, here we have argued entirely in terms of the equations of motion, which for the ansatz (8) we have linearized in $\tilde{\phi}$ and $\tilde{A}_i$:

$$\begin{aligned}
&\hat{D}_i\hat{D}_i\tilde{\phi} - 2i\tilde{A}^i\hat{D}_i\hat{\phi} - i\hat{\phi}\partial^i\tilde{A}_i + \frac{1}{2}\tilde{\phi}(|\hat{\phi}|^2 - 1) \\
&+\frac{1}{2}\tilde{\phi}(|\hat{\phi}|^2 - 1) + \frac{1}{2}\hat{\phi}[\hat{\phi}\tilde{\phi}^* - \hat{\phi}^*\tilde{\phi}] + \frac{1}{2}\tilde{\lambda}\hat{\phi}(|\hat{\phi}|^2 - 1) = 0, \\
&\partial^i\tilde{F}_{ij} + \tilde{A}_j|\hat{\phi}|^2 + \frac{i}{2}[\tilde{\phi}^*\hat{D}_j\hat{\phi} - \tilde{\phi}(\hat{D}_j\hat{\phi})^* + \hat{\phi}^*\hat{D}_j\tilde{\phi} - \hat{\phi}(\hat{D}_j\tilde{\phi})^*] = 0, \\
&\partial^i\partial_0\tilde{A}_i - \frac{i}{2}[\hat{\phi}^*\partial_0\tilde{\phi} - \hat{\phi}\partial_0\tilde{\phi}^*] = 0,
\end{aligned} \tag{20}$$

where $\hat{D}_i = \partial_i - i\hat{A}_i$ and $\tilde{F}_{ij} = \partial_i\tilde{A}_j - \partial_j\tilde{A}_i$. We have solved these equations and we expect, although it has not been proven rigorously, that we have found approximate solutions for $t \in (-\varepsilon,\varepsilon), \varepsilon << 1$. Among the three possibilities of 0° , 90° and 180° scattering, the approximate solutions decide in favour of 90° scattering.

There are at least still two problems which have to be addressed. First, a solution for $t \in (-\varepsilon,\varepsilon)$ is not a scattering solution. However, we can take the configuration for $t = 0$ as initial data of a solution for $t \in (-\infty,\infty)$ which we know exists[18]. For

$t \in (-\varepsilon,\varepsilon), \varepsilon << 1$, the linearization which leads to eqs. (20) should be justified and the solution we discussed for $t \in (-\varepsilon,\varepsilon)$ should be an approximation to the scattering solution for $t \in (-\infty,\infty)$. The second problem is concerned with the experimental realisation of the 90° scattering process. We have given evidence for 90° scattering by presenting special approximate solutions, which require special initial data. However, since the parameter space for static vortices is 4-dimensional and we have found a 4-parameter family of approximate solutions (9), (12) and (13), which all describe 90° scattering possibly with a spatial translation, we expect 90° scattering for slowly moving vortices for all initial data which lead to a collision.

APPENDIX: SERIES SOLUTIONS

We want to draw briefly the attention to the series solutions which can be given to eqs. (14)-(16). We use techniques developed by Rühl[19]. (Details can be found in ref. 16.) At the origin, we rewrite the equations as

$$\begin{aligned}
f(r) &= \gamma_0 r^2 - 2r^2\int_0^r \frac{a(s)f(s)}{s^3}ds, \\
a(r) &= \frac{1}{8}r^2 - \frac{1}{4}\int_0^r sf^2(s)ds, \\
k(r) &= \frac{c_1}{r^2} + c_2r^2 + \int_0^r \left(\frac{r^2}{4\rho} - \frac{\rho^3}{4r^2}\right) f^2(\rho)k(\rho)d\rho.
\end{aligned}$$

The coefficients in the series $f = \sum_{n=1}^{\infty} f_n r^{2n}, a = \sum_{n=1}^{\infty} a_n r^{2n}, k = \sum_{n=-1}^{\infty} r^{2n}$ are given by the recursion relation

$$\begin{aligned}
f_n &= \frac{1}{1-n}\sum_{n_1,n_2=1}^{\infty} a_{n_1}f_{n_2}\delta_{n,n_1+n_2}, \\
a_n &= \frac{-1}{8n}\sum_{n_1,n_2=1}^{\infty} f_{n_1}f_{n_2}\delta_{n,n_1+n_2+1}, \\
k_n &= \frac{1}{4(n^2-1)}\sum_{n_1,n_2,n_3=1}^{\infty} f_{n_1}f_{n_2}k_{n_3}\delta_{n,n_1+n_2+n_3+1}.
\end{aligned}$$

By induction, one can prove

$$|f_n| \leq \frac{M^n}{(n+1)^2}, \quad |a_n| \leq \frac{M^n}{(n+1)^2}, \quad |k_n| \leq \frac{M^n}{(n+1)^2},$$

and hence convergence for $r < 1/\sqrt{M}$. This provides series solutions near the origin, which can be shown to include the solutions to (14)-(16) discussed in section 2.

At infinity, we can use eqs. (14)-(16) to find asymptotic series of the form $f = \sum_{n=0}^{\infty} f_n(r)e^{-nr}, a = \sum_{n=0}^{\infty} a_n(r)e^{-nr}, k = \sum_{n=0}^{\infty} k_n(r)e^{-nr}$ with polynomially bounded coefficient functions. f_n satisfies the recursion relation

$$f_n(r) = \frac{i\pi}{4}e^{nr}\int_r^{\infty} \rho[H_0^{(2)}(i\rho)H_0^{(1)}(ir) - H_0^{(1)}(i\rho)H_0^{(2)}(ir)]$$

$$\times \left\{ \frac{3}{2}\sum_{n_1,n_2=1}^{\infty} f_{n_1}f_{n_2}\delta_{n,n_1+n_2} + \frac{4}{r^2}\sum_{n_1,n_2=1}^{\infty} a_{n_1}a_{n_2}\delta_{n,n_1+n_2} \right.$$

$$\left. + \frac{1}{2}\sum_{n_1,n_2,n_3=1}^{\infty} f_{n_1}f_{n_2}f_{n_3}\delta_{n,n_1+n_2+n_3} + \frac{4}{r^2}\sum_{n_1,n_2,n_3=1}^{\infty} a_{n_1}a_{n_2}f_{n_3}\delta_{n,n_1+n_2+n_3} \right\} e^{-n\rho}d\rho.$$

a_n and k_n satisfy similar relations. One can show that

$$\sup_{r>R} |r f_n(r) e^{-nr/2}| < \frac{M^n}{(n+1)^2}$$

and establish similar bounds for a_n and k_n. This implies that the series converge for $r > R$ and $r > 2\log M$.

REFERENCES

1. Kibble, T.W.B., J. Phys. A9, 1397 (1976).
2. Vilenkin, A., Phys. Rep. 121, 263 (1985).
3. Manton, N.S., Phys. Lett. 110B, 54 (1982).
4. Atiyah, M.F. and Hitchin, N.J., "The geometry and Dynamics of Magnetic Monopoles", Princeton University Press 1988.
5. Ward, R.S., Phys. Lett 158B, 424 (1985).
6. Ruback, P.J., Nucl. Phys. B296, 669 (1988).
7. Zakrzewski, W.J., "Scattering of Skrymions in (2+1) Dimensions", in these Proceedings, and references therein.
8. Laguna, P. and Matzner, R.A., Phys. Rev. D 41, 1751 (1989), and references therein.
9. Huebener, R.P., "Magnetic Flux Structures in Superconductors", Springer 1979.
10. Matzner, R.A., Comp. Phys. 2, 51 (1988).
11. Plohr, B., Princeton Unversity Doctoral Dissertation, 1980; J. Math. Phys. 22, 2184 (1981).
12. Chow, S.N. and Hale, J.K., "Methods of Bifurcation Theory", Springer 1982.
13. Bogomolnyi, E.B., Yad. Fiz. 24, 861 (1976) [Sov. J. Nucl. Phys. 24, 449 (1976)].
14. Jaffe, A. and Taubes, C., "Vortices and Monopoles", Birkhäuser 1980.
15. Weinberg, E.J., Phys. Rev. D 19, 3008, (1979).
16. Burzlaff, J. and McCarthy, P., in preparation.
17. Shellard, E.P.S. and Ruback, P.J., Phys. Lett. 209B, 262 (1988).
18. Burzlaff, J. and Moncrief, V., J. Math. Phys. 26, 1368 (1985).
19. Rühl, W., University of Kaiserslautern, unpublished report, 1977.

PART II

Q-DEFORMATIONS AND QUANTUM GROUPS

Solvable Lattice Models and Representation Theory of Quantum Groups

Tetsuji MIWA

RIMS, Kyoto Univ., Kyoto Japan

1. The 6 vertex model

About ten years ago, Rodney Baxter wrote a book "Exactly Solved Models in Statistical Mechanics" [1], in which he gave beautiful expositions on solvable lattice models such as the Ising model, the 6 vertex model, the 8 vertex model, the hard hexagon model, and so on. Since then a flood of new ideas is coming out of this area and flowing into mathematics and mathematical physics. In this lecture, I will explain some of the developments in the ten years after Baxter's book.

I will mainly discuss on the 6 vertex model. This is a statistical mechanical model on 2 dimensional square lattice.

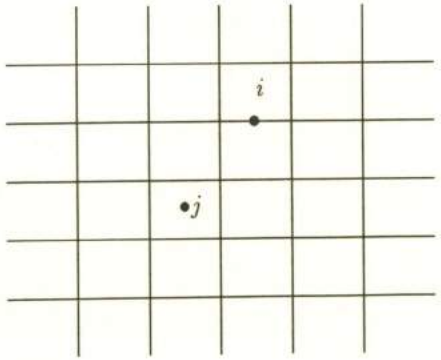


*Fig.*1

Let i denote a bond of the lattice. We associate it with a fluctuation variable σ_i taking value 1 or 2. A *configuration* $\sigma = (\sigma_i)$ is a collection of $\sigma_i = 1$ or 2 assigned to each i. A configuration must satisfy the following *restriction*: Around each vertex, only one of the following 6 configurations is allowed.

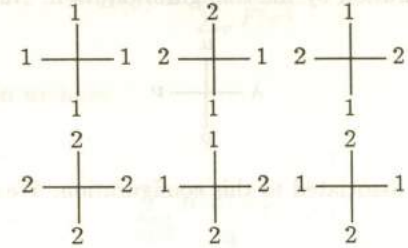


*Fig.*2

We introduce a dual configuration $\tau = (\tau_j)$ in the following way. This time j indicates a face, not a bond, and τ_j is a fluctuation variable associated to the face j taking value in $\mathbf{Z}$ (Fig. 1) Restriction for τ is

$$(1) \qquad \tau_{j'} = \tau_j \pm 1$$

if j and j' are adjacent, i.e.

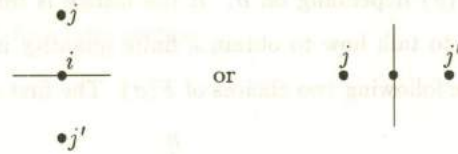


*Fig.*3

In fact, we can make σ out of τ by defining

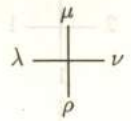
$$(2) \qquad \sigma_i = 1 \quad \text{if} \quad \tau_{j'} = \tau_j + 1$$
$$= 2 \quad \text{if} \quad \tau_{j'} = \tau_j - 1$$

where i is the bond between j and j' as indicated in Fig.3. Note that if you define σ from τ in this way, σ satisfies the 6 vertex condition, namely, around each vertex only one of 6 configurations in Fig.2 occurs. Conversely, if σ is a configuration

satisfying the 6 vertex condition, then one can find τ satisfying (2). For a given σ, the choice of τ is not unique, but later we will fix this arbitrariness.

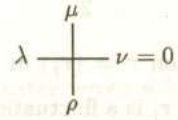
Let us come back to Fig.2. To each of 6 configurations, we associated a complex number, called the Boltzmann weight. Let us denote the Boltzmann weight associated to a configuration by the configuration itself. Namely, we represent by

(3)



the Boltzmann weight associated to this configuration. We set

(4)



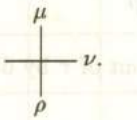
if (2) is none of those in Fig.2.

A configuration sum is a sum over configurations

$$\sum_{\sigma} F(\sigma) \tag{5}$$

of certain function $F(\sigma)$ depending on σ. If the lattice is infinite, such a sum diverges. I am going to talk how to obtain a finite quantity in such a situation soon. We consider the following two choices of $F(\sigma)$. The first one is

(6) $F(\sigma) = \prod_{vertex}$



Here we take the product of the Boltzmann weights for all the vertices on the lattice determined by a configuration σ. The second one is as follows. Fix a face j, say $j = 0$. Choose an integer a. Then, the second choice, which I denote by $G(a;\sigma)$, is

$$G(a;\sigma) = \delta_{\tau_0,a} F(\sigma), \tag{7}$$

where δ_{ab} is the Kronecker delta and $F(\sigma)$ is as in (σ). Here I suppose that we fixed the choice of τ for a given σ in some way. Now consider the following ratio

$$H(a) = \frac{\sum_{\sigma} G(a;\sigma)}{\sum_{\sigma} F(\sigma)}. \tag{8}$$

Then, by the definition we have

$$\sum_{a \in \mathbf{Z}} H(a) = 1. \tag{9}$$

In fact, $F(\sigma)/\sum_{\sigma} F(\sigma)$is the probability of the occurrence of the configuration σ, and therefore $H(a)$ is nothing but the probability of the variable τ_0 taking the value a . If one starts from a finite size lattice, computes (8) and takes the infinite lattice limit, then one gets a finite answer. We call it *1 point function.*

Remember that we have not fixed the one to one correspondence between σ and τ, therefore the definition of the 1 point function is not yet completed. I postpone this task to a later section.

2. The R matrix as intertwiner

I will not reproduce Baxter's arguments on the solvability of lattice models, the commuting transfer matrices and all that, but start with the following formulation of the Yang-Baxter equation. Suppose that we have certain operators

$$R_i(x,y) \tag{10}$$

where $i = 1,2,3,\cdots,m-1$ and x, y are points on some manifold, say $\mathcal{S}$. We call $\mathcal{S}$ a spectral manifold, and x, y spectral parameters.

The Yang-Baxter equation is

$$R_i(x,y)R_{i+1}(x,z)R_i(y,z) = R_{i+1}(y,z)R_i(x,z)R_{i+1}(x,y). \tag{11}$$

together with

$$[R_i(x,y), R_j(x',y')] = 0 \quad \text{if} \quad |i-j| > 1. \tag{11}'$$

A solution $R_i(x,y)$ to the Yang-Baxter equation is called an R matrix. Let us give an example of R matrix, which is related to the 6 vertex model. Consider a two dimensional vector space $V \cong \mathbf{C}v_1 \oplus \mathbf{C}v_2$ with the distinguished basis v_1, v_2. We identify the Boltzmann weights of the 6 vertex model with the matrix elements of a matrix R acting on $V \otimes V$:

$$Rv_\mu \otimes v_\nu = \sum_{\lambda,\rho} \lambda \begin{array}{c} \mu \\ \downarrow \\ \nu \rightarrow \\ \rho \end{array} \nu\, v_\lambda \otimes v_\rho \tag{12}$$

We introduce two parameters x and q and give the Boltzmann weights as follows.

$$\begin{array}{c} \mu \\ \downarrow \\ \mu \rightarrow \\ \mu \end{array} \mu = xq - x^{-1}q^{-1} \tag{13}$$

$$\begin{array}{c} \mu \\ \downarrow \\ \mu \rightarrow \\ \nu \end{array} \nu = \begin{cases} x(q-q^{-1}) & \mu < \nu \\ x^{-1}(q-q^{-1}) & \mu > \nu \end{cases}$$

$$\begin{array}{c} \mu \\ \downarrow \\ \nu \rightarrow \\ \mu \end{array} \nu = x - x^{-1}$$

Let us denote by $R(x)$ the matrix in (12) with the special choice of weights (13). Consider tensor product

$$\underbrace{V \otimes V \otimes V \otimes \cdots \otimes V}_{m},$$

and define $R_i(x,y)$ to be the operator on it such that it acts trivially except for the ith and the $(i+1)$th components on which it acts as $R(x/y)$. This is the R matrix of the 6 vertex model.

Now I will explain how one can find this R matrix by using the representation theory of the quantum group $U_q\widehat{\mathfrak{sl}}(2)$ [2],[3]. The idea is that the R matrix is an intertwiner of two isomorphic tensor product representation.

Let me start from $\mathfrak{sl}(2)$ and its two dimensional representation V. The Lie algebra $\mathfrak{sl}(2)$ has generators e, f and h satisfying $[h,e] = 2e, [h,f] = -2f$ and $[e,f] = h$. The representation is given by

$$ev_1 = 0, \quad ev_2 = v_1, \quad fv_1 = v_2, \quad fv_2 = 0,$$
$$hv_1 = v_1 \quad \text{and} \quad hv_2 = -v_2.$$

We can define the action of $\mathfrak{sl}(2)$ on the tensor product $V \otimes V$. The tensor product splits into the symmetric tensor and the anti-symmetric tensor.

$$\square \otimes \square = \square\square \oplus \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} \tag{14}$$
$$2 \times 2 = 3 + 1$$

A matrix

$$R : V \otimes V \longrightarrow V \otimes V \tag{15}$$

is called an intertwiner if it commutes with the action of $\mathfrak{sl}(2)$. $R = id$ is an intertwiner, and $R = P$ (the permutation, i.e. $P(x \otimes y) = y \otimes x$) is also an intertwiner. In general, $a \cdot id + bP$ $(a, b \in \mathbf{C})$ are the intertwiners. This is because $V \otimes V$ splits into two irreducible representations.

Now consider $U_q\mathfrak{sl}(2)$ This is an algebra generated by e, f and t satisfying the following relation

$$
\begin{aligned}
tet^{-1} &= q^2 e, \\
tft^{-1} &= q^{-2} f, \\
[e, f] &= \frac{t - t^{-1}}{q - q^{-1}}.
\end{aligned} \tag{16}
$$

We can make a representation of $U_q\mathfrak{sl}(2)$ on V by the same actions for e and f as $\mathfrak{sl}(2)$ and by putting $t = q^h$. Here we consider q as a generic complex number. Or more precisely, we consider q as an indeterminate. Then we must change V as

$$
V = \mathbf{C}(q)v_1 \oplus \mathbf{C}(q)v_2, \tag{17}
$$

namely we take rational functions in q as coefficients.

If we consider the tensor product $V \otimes V$, the effect of q deformation really appears. For $X \in \mathfrak{sl}(2)$, $X \otimes 1 + 1 \otimes X$ gives the action, but for $U_q\mathfrak{sl}(2)$ there is no such right-left symmetric action. We take the following actions of the generators on $V \otimes V$.

$$
\begin{aligned}
e &\longmapsto e \otimes 1 + t \otimes e \\
f &\longmapsto f \otimes t^{-1} + 1 \otimes f \\
t &\longmapsto t \otimes t
\end{aligned} \tag{18}
$$

Now, let us consider the intertwiner (15) in this situation. The identity is still an intertwiner, but the permutation P is not because of the asymmetry of the action

on the tensor product. But, still there are two dimensional intertwiners. This is because the splitting similar to (14) takes place. In fact,

$$
\begin{aligned}
&v_1 \otimes v_1 \\
&v_2 \otimes v_1 + qv_1 \otimes v_2 \\
&v_2 \otimes v_2
\end{aligned} \tag{19}
$$

are the symmetric tensors and

$$
-qv_2 \otimes v_1 + v_1 \otimes v_2 \tag{19}'
$$

is the anti-symmetric tensor.

Now we consider $U_q\widehat{\mathfrak{sl}}(2)$. This is an algebra generated by $e_i, f_i, t_i (i = 0, 1)$. In fact $U_q\widehat{\mathfrak{sl}}(2)$ contains two $U_q\mathfrak{sl}(2)$, the one generated by e_1, f_1, t_1 and the other generated by e_0, f_0, t_0. Let us call these two $U_q\mathfrak{sl}(2)$ of color 1 or 0, respectively. These two are combined by the so called Serre relation.

$$
\begin{aligned}
e_0^2 e_1 - (q + q^{-1}) e_0 e_1 e_0 + e_1 e_0^2 &= 0 \\
e_1^2 e_0 - (q + q^{-1}) e_1 e_0 e_1 + e_0 e_1^2 &= 0
\end{aligned} \tag{20}
$$

We are interested in finite dimensional irreducible representations of $U_q\widehat{\mathfrak{sl}}(2)$. A simple way of obtaining such representation is to start from a finite dimensional irreducible representation of the $U_q\mathfrak{sl}(2)$ of color 1. Then a question is if it is possible to extend this action to the whole algebra $U_q\widehat{\mathfrak{sl}}(2)$ consistently or not. The answer is yes, but the solution is not unique. Take any non zero complex parameter x. Then the action of e_0, f_0, t_0 defined by the substitution $e_0 \mapsto xf_1$, $f_0 \mapsto x^{-1}e_1$, $t_0 \mapsto t_1^{-1}$, is consistent. This appearance of an arbitrary parameter x is a new phenomena which we observe when we consider finite dimensional irreducible representations of affine Lie algebras (with or even without q).

Let us denote by V_x the irreducible 2 dimensional representation of $U_q\widehat{\mathfrak{sl}}(2)$ obtained from the color 1 $U_q\mathfrak{sl}(2)$ action on V of (17). Now let us consider an intertwiner

$$R(x,y) : V_x \otimes V_y \longmapsto V_y \otimes V_x. \tag{21}$$

Now, the choice $R(x,y) = id$ does not work. It is not at all clear whether $V_x \otimes V_y$ and $V_y \otimes V_x$ are isomorphic. The fact is that for generic x, y

(i) $V_x \otimes V_y$ and $V_y \otimes V_x$ are still irreducible

and

(ii) $V_x \otimes V_y$ and $V_y \otimes V_x$ are isomorphic.

Therefore, the intertwiner exists and it is unique up to scalar multiple. With a suitable normalization of this scalar multiple, the intertwiner is given by (12) and (13).

We conclude this section by noting that the Yang-Baxter equation follows from the fact that R is an intertwiner. In fact, both sides of (11) (Take $m = 3$ and $i \doteq 1$ for simplicity.) intertwines two irreducible representations $V_x \otimes V_y \otimes V_z$ and $V_z \otimes V_y \otimes V_x$, and therefore they must coincide (up to a scalar which can be fixed to 1 by a simple argument).

3. Link invariants

The braid group B_m of m strings is a group generated by invertible elements $\sigma_i \quad (i = 1, \cdots, m-1)$ satisfying

$$\sigma_i\sigma_{i+1}\sigma = \sigma_{i+1}\sigma_i\sigma_{i+1} \tag{22}$$

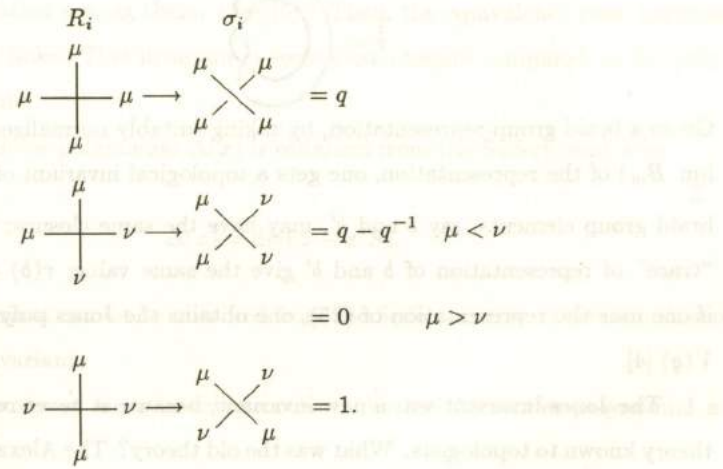
$$[\sigma_i, \sigma_j] = 0 \quad \text{if} \quad |i-j| > 1. \tag{23}$$

(Do not confuse this σ_i with fluctuation variable σ_i in the 6 vertex model.) These relations are the same with the Yang-Baxter equation if we forget the spectral parameters. Suppose that the spectral manifold $\mathcal{S}$ is just $\mathbf{C}^\times = \mathbf{C} - \{0\}$, and the dependence of R_i on x and y is of the form $R_i(x/y)$. Then we might expect that $\sigma_i = R_i(1)$ give a braid group representation. But, we find in most cases that $R_i(1) = const.$, the trivial representation. The correct choice is $x^{\pm 1} \to \infty$. With suitable normalization of x power we can, in most cases, obtain a non trivial braid group representation:

$$\lim_{x^{\pm 1}\to\infty} x^{\pm(power)} R_i(x) = \sigma_i^{\pm 1}. \tag{24}$$

For example, the braid limit of the 6 vertex R matrix reads

(25)



Now we are going to talk about topological invariants of links. For each braid group element b one can define its closure $\hat{b}$, a link obtained as follows. To each elementary braid $\sigma_i^{\pm 1}$ we associate the following braid of strings

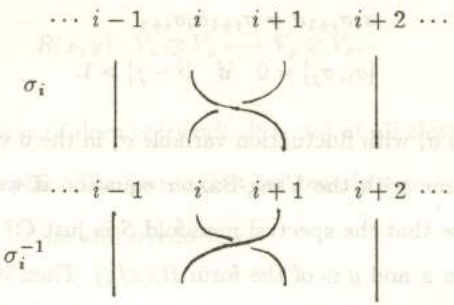


*Fig.*4

The product of braid group elements corresponds to the composition of braids of strings. For example, consider the case $m = 2$. Take $b = \sigma_1^2$. Then the corresponding topological braid is

$b =$ 

*Fig.*5

Closing this braid, we get the following link which is called the Hopf link.

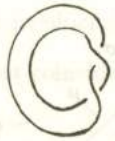
$\hat{b} =$ 

*Fig.*6

Given a braid group representation, by taking suitably normalized trace $\tau(b)$ ($b \in \varinjlim_m B_m$) of the representation, one gets a topological invariant of links. Different braid group elements, say b and b', may have the same closure; $\hat{b} = \hat{b'}$. But, the "trace" of representation of b and b' give the same value; $\tau(b) = \tau(b')$. In fact, if one uses the representation of (25), one obtains the Jones polynomial invariant $V(q)$ [4].

The Jones invariant was a new invariant, because it never reduces to the old theory known to topologists. What was the old theory? The Alexander polynomial

is one of previously known invariants. Seifert matrix invariant is known as more general construction of invariants. Let me briefly explain the Seifert theory.

We put a surface on a link. For example, starting from two discs, then connecting them by two thin strips, and finally twisting both of them, we get the Hopf link with surface on it.

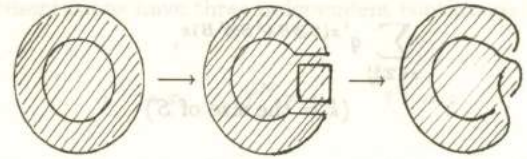


*Fig.*7

Now, take a basis of homology 1 cycles on the surface and define a matrix S to be the linking matrix of these cycles. Namely, the size of S is equal to the dimension of the 1-st homology group of the surface, and the (i, j) entry of S is the linking number between the i-th and the j-th cycles. This is called the Seifert matrix.

The way of putting a surface on a link is not unique. The choice of cycles is also not unique. Therefore, the Seifert matrix is not uniquely determined for a given link. Even the size may vary.

To get "invariant" of links from the Seifert matrices, one introduces a certain equivalence relation among them. (See [5].) Then, the equivalence class becomes an invariant of links. This invariant is somewhat obscure compared to the polynomial invariants.

The Alexander polynomial $\Delta(x)$ is obtained from the Seifert matrix by

$$\Delta(x) = \det(S - x^t S). \tag{26}$$

The Jones invariant cannot be obtained from the Seifert matrix. Therefore, it was called a new invariant.

We can construct more invariants other than the Alexander polynomial as follows.

Fix integers l and N, and take an arbitrary bilinear form B such that

$$B : \mathbf{Z}_N^l \times \mathbf{Z}_N^l \longrightarrow \mathbf{Z}_N. \tag{27}$$

Then, the following Gauss sum gives an invariant ($q^N = 1$).

$$\sum_{x \in \mathbf{Z}_N^{kl}} q^{{}^t x(S \otimes B + {}^t S \otimes {}^t B)x}, \tag{28}$$

(k = the size of S) .

The questions are

(i) Can one find a representation of the braid group such that the invariant (28) is its consequence?

(ii) Can one find an R matrix with spectral parameters such that that braid group representation is its consequence?

I will answer to these questions in the next section.

4. Roots of unity

Now we are interested in representations of U_q at root of unity, say $q = e^{2\pi i/N}$ where N is odd, and the corresponding R matrix.

The big difference between the cases where q is generic and $q^N = 1$, is that for the latter case U_q contains a big center. For example, consider $U_q\mathfrak{sl}(2)$. One can show that e^N, f^N, t^N commute with any element in $U_q\mathfrak{sl}(2)$. Therefore, the irreducible representations have three independent continuous parameters

$$e^N = x, \qquad f^N = y, \qquad t^N = z. \tag{29}$$

If you start from a finite dimensional representation of $U_q\mathfrak{sl}(2)$ with generic q, and take a limit where q approaches to a root of unity, you get a restricted representation such that $x = 0, y = 0$ and $z = 1$. We will consider non-restricted representations.

Weights of the representation take value at one of the N-th roots of z. Therefore, there are N different weight spaces. Since e and f (we assume $x \neq 0, y \neq 0$) give isomorphisms between weight spaces, the multiplicities of weight spaces are the same. In fact, if we consider an irreducible representation, each weight space is 1 dimensional for $U_q\mathfrak{sl}(2)$, and the total representation space is N dimensional.

Now, take two such irreducible representations, say V and V'. Suppose (29) for V and x', y', z' for V'. Consider $V \otimes V'$ and $V' \otimes V$.

Are they isomorphic? This is a crucial question. The answer is NO, for generic choice of V and V'. The reason is simple. Let us compute e^N on $V \otimes V'$.

From (18) we have

$$e^N \to e^N \otimes t^N + 1 \otimes e^N = xz' + x' \tag{30}$$

for $V \otimes V'$. Therefore, if $V \otimes V'$ and $V' \otimes V$ are isomorphic, we need

$$xz' + x' = x'z + x. \tag{31}$$

Therefore, the intertwiners exist only for special choices of parameters in the representations.

In order to get R matrix with spectral parameters we need to consider $U_q\hat{\mathfrak{sl}}(2)$. In fact, we need to add two more central elements z_0, z_1 to $U_q\hat{\mathfrak{sl}}(2)$. They are trivial centers for multiplicative structure, but we deform comultiplicative structure. This means that we change (18) as follows.

$$
\begin{aligned}
e_i &\longmapsto e_i \otimes 1 + z_i t_i \otimes e_i \\
f_i &\longmapsto f_i \otimes t_i^{-1} + z_i^{-1} \otimes f_i \\
t_i &\longmapsto t_i \otimes t_i \\
z_i &\longmapsto z_i \otimes z_i .
\end{aligned} \tag{32}
$$

Recall that representations of $U_q\hat{\mathfrak{sl}}(2)$ acquire one more continuous parameter when it is extended from representations of $U_q\mathfrak{sl}(2)$. Thus, in the total, we have 6 independent continuous parameters in the representations of $U_q\hat{\mathfrak{sl}}(2) \oplus \mathbf{C}z_0 \oplus \mathbf{C}z_1$.

Let us denote these 6 parameters by ξ. Take two representations V_ξ and V_η. Now our R matrix should be

$$
R(\xi, \eta) : V_\xi \otimes V_\eta \longrightarrow V_\eta \otimes V_\xi . \tag{33}
$$

As I explained above, for generic ξ, η we have no such intertwiner. The fact is the 6 dimensional parameter space has fibrations such that each fiber $\mathcal{S}$ is $\mathcal{C} \times \mathcal{C}$ with some curve $\mathcal{C}$ ($\mathcal{C}$ varies from fiber to fiber) and the intertwiner $R(\xi, \eta)$ exists if and only if ξ and η lies on the same fiber $\mathcal{S}$. The corresponding solvable lattice model is found to be the chiral Potts model.(See [6],[7].) We can proceed in the same way for $U_q\mathfrak{sl}(n)$ in the case of the minimal representations, i.e., the representations whose weight spaces have dimension 1. Since the number of weight spaces is N^l ($l = n-1$), these are N^l dimensional representation. Starting from this, after a lengthy calculation we get new R matrices as the intertwiners (33). (See [8],[9].) Again, the spectral manifold has the same structure $\mathcal{S} = \mathcal{C} \times \mathcal{C}$ (the curves $\mathcal{C}$ are

different for different n). I am not going into the details about this calculation, but will discuss its braid limit [10].

As I have explained, the spectral manifold splits into two curves

$$
\begin{array}{ccc}
\mathcal{S} & = & \mathcal{C} \times \mathcal{C} \\
\cup\!\!| & & \cup\!\!| \\
\xi & = & (u, v)
\end{array} \tag{34}
$$

Accordingly, the R matrix splits into four pieces:

$$
R(\xi, \eta) = \bar{S}(u, v')T(u, u')\bar{T}(v, v')S(v, u') \tag{35}
$$

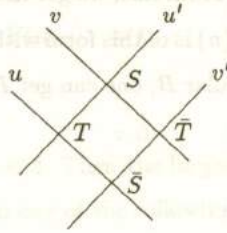


Fig.8

In fact, we can show that $T(u, u')$ $(u, u' \in \mathcal{C})$ satisfies the Yang-Baxter equation among itself.

Among the spectral manifolds, there is a particular one such that $\mathcal{C} \cong \mathbf{C}^\times$. In this case, we can take the braid limit of T. In the following, I will describe this braid representation.

Recall the bilinear form B of (27). Define

$$
A = \frac{1}{2}(B - {}^tB).
$$

Let us denote by $a, b, \cdots$, elements of $\mathbf{Z}_N^l$. We take $l = n-1$. Consider an algebra $\mathcal{A}_m$ which is generated by elements x_i^a $(i = 1, \cdots, m-1; a \in \mathbf{Z}_N^l)$. We impose the following relations.

$$
\begin{aligned}
&[x_i^a, x_j^b] = 0 \quad \text{if} \quad |i-j| > 1 \\
&x_i^a x_i^b = q^{A(b,a)} x_i^{a+b} \\
&x_i^a x_{i+1}^b = q^{B(a,b)} x_{i+1}^b x_i^a ,
\end{aligned} \tag{36}
$$

This is a C-algebra with dimension

$$D = N^l. \tag{37}$$

Define σ_i $(i = 1, \cdots, m-1)$ by

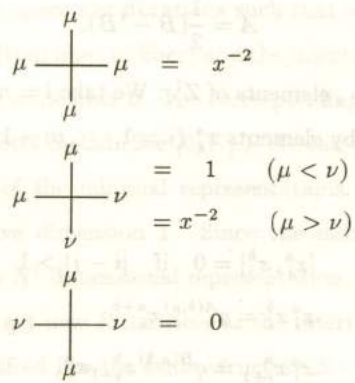
$$\sigma_i = \frac{1}{\sqrt{D}} \sum_{a \in \mathbf{Z}_N^l} q^{\frac{-1}{2}B(a,a)} x_i^a. \tag{38}$$

They satisfy the braid relation. From this, we get the invariant (28) by the routine procedure. The limit of T for $\mathfrak{sl}(n)$ is of this form with a particular choice of B, and surprisingly, only for this particular B, one can get R matrix with (trigonometric) spectral parameters.

5. Low temperature limit

Now we change the subject, and come back to the 6 vertex model. I explained the role of the spectral parameter x. From the physical point of view the parameter q is much more important. Let us consider the case $0 < q < 1$. Here q is temperature; $q = 0$ is the low temperature limit and $q = 1$ is the critical temperature. Let us see what happens to the R matrix in the low temperature limit. Up to some scaling factor, we have

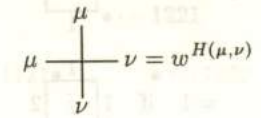
(39)



First of all, in this low temperature limit, because of the last equation, configurations $2 \begin{smallmatrix} 1 \\ + \\ 1 \end{smallmatrix} 2$ or $1 \begin{smallmatrix} 2 \\ + \\ 2 \end{smallmatrix} 1$ are not allowed. Therefore the fluctuation variable σ_i (Do not confuse it with a braid generator!) must take the same value along the NE to SW diagonal direction. In other words, the system freezes to 1 dimensional system.

Now, let us consider the other weights. Set $w = x^{-2}$. One can rewrite (39) as

(40)



$$H(\mu,\nu) = 1 \quad \begin{matrix} \mu & \nu \\ 1 & 1 \\ 2 & 1 \\ 2 & 2 \end{matrix}$$

$$= 0 \quad \begin{matrix} 1 & 2 \end{matrix}$$

We consider the case $0 < w < 1$. Then the largest contributions to the configuration sum (5), (6), come from one of the following.

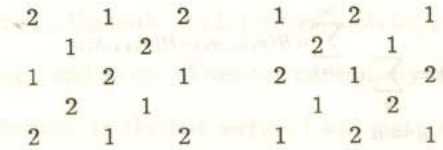


Fig.9

They are called the ground state configurations. We choose and fix one of them. From now on, we consider such configurations that are different from the ground state configuration at finitely many places. This gives the proper definition of the 1 point function. (In fact, we need to care about how to take the large lattice limit, but I am not going into details on this routine matter.)

Chapter 14 of Baxter's book deals with the 1 point function of the hard hexagon model by means of the corner transfer matrix method. I am not going to explain this method here, since I have nothing to add to his book for the moment, as far as the CTM method itself is in concern.

The reader must know that this method provides us with a reduction of 2 dimensional configuration sum to 1 dimensional configuration sum. Before explaining this reduction in the 6 vertex case, we must settle the question about the

one to one correspondence between σ configurations and τ configurations. Let us denote the ground state configuration for σ by $\bar{\sigma}$. We define the corresponding ground state configuration for τ, which we denote by $\bar{\tau}$, as follows.

(41)
$$\bar{\tau}_j = 0 \quad \text{if} \quad \begin{array}{ccc} & 2 & \\ 2 & \boxed{j} & 1 \\ & 1 & \end{array}$$
$$= 1 \quad \text{if} \quad \begin{array}{ccc} & 1 & \\ 1 & \boxed{j} & 2 \\ & 2 & \end{array}$$

*Fig.*10

Note that, given σ which is equal to $\bar{\sigma}$ almost everywhere there is a unique τ which satisfies (2) and is equal to $\bar{\tau}$ almost everywhere.

The computation of the 1 point function reduces to the computation of the following 1 dimensional configuration sum.

(42)
$$S(a) = \sum_{\sigma} q^{\sum_{i=1}^{\infty} i(H(\sigma_{i+1},\sigma_i) - H(\bar{\sigma}_{i+1},\bar{\sigma}_i))}$$
$$\tau_0 = a$$

Here σ is a configuration on the half infinite 1 dimensional chain and τ is the dual configuration.

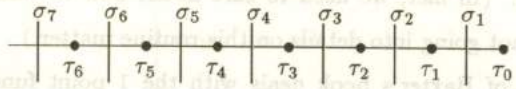


*Fig.*11

In the sum (42) we take only such σ that differs from the following ground state $\bar{\sigma}$ at finitely many places:

(43)
$$\bar{\sigma} = (\cdots 1212).$$

Let us denote the exponent of q in (42) by $\omega(\sigma)$.

The following is a plot of several paths (We call a 1 dimensional configuration a *path*.) in the (ω, τ_0)-plane.

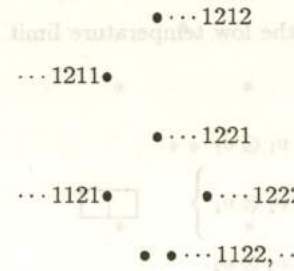


*Fig.*12

Note that the 1 dimensional configuration sum (42) is the generating function of the multiplicities of paths lying along a vertical string. The fact observed is that Fig. 12 is nothing but the diagram of the weight multiplicities for the basic representation $L(\Lambda_0)$ of $U_q\hat{\mathfrak{sl}}(2)$ [11]. The path $\cdots 1212$ corresponds to the highest weight vector v_{Λ_0}, the path $\cdots 1211$ corresponds to $f_0 v_{\Lambda_0}$, the path $\cdots 1221$ corresponds to $f_1 f_0 v_{\Lambda_0}$, and so on. When the multiplicity is greater than 1, such correspondence is obscure. In the last section I will make a precise assertion by using the *crystal base*.

6. Crystal base

Let us go back to (19) and (20). They give the decomposition of the tensor product $V \otimes V (V = \mathbf{C}(q)v_1 \oplus \mathbf{C}(q)v_2)$ into the symmetric tensor $\square\square$ and the anti-symmetric tensor $\begin{smallmatrix}\square\\\square\end{smallmatrix}$. In the low temperature limit $q \to 0$, they become as follows.

(44) $v_1 \otimes v_1$, $v_2 \otimes v_1$, $v_2 \otimes v_2$ $\Big\}$ $\square\square$

(45) $v_1 \otimes v_2$ $\begin{smallmatrix}\square\\\square\end{smallmatrix}$.

A remarkable thing here is that the decomposition in the limit $q \to 0$ is given by pure tensors. Linear combinations have disappeared.

In general, we have the following. Let

$$M = \underbrace{V \otimes V \otimes \cdots \otimes V}_{m} \tag{46}$$

Denote by A the subring of $\mathbf{C}(q)$ consisting of those which have no pole at $q = 0$. Let

$$B = \{v_{i_1} \otimes \cdots \otimes v_{i_m}\}. \tag{47}$$

This is a set of 2^m elements. Let

$$L = \bigoplus_{b \in B} Ab. \tag{48}$$

Then we have $M = \mathbf{C}(q) \underset{A}{\otimes} L$. The statement is that we can decompose M in such a way that each irreducible component is spanned by such vectors that are equal to elements in B modulo qL [12].

I will come back to the consideration of tensor product later. Now let us consider $U_q\mathfrak{sl}(3)$. Let us consider the adjoint representation $\begin{smallmatrix}\square\square\\\square\end{smallmatrix}$. This is 8 dimensional and the diagram of the weight multiplicities is as follows

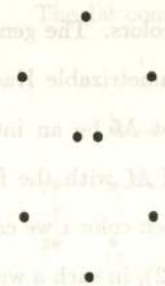


*Fig.*13

Note that the multiplicity 2 occurs in the weight space of weight 0. $U_q\mathfrak{sl}(3)$ contains two $U_q\mathfrak{sl}(2)$, one generated by e_1, f_1, t_1 and the other generated by e_2, f_2, t_2. Let us call them $U_q\mathfrak{sl}(2)$ of color $i = 1, 2$, respectively. We denote them by $U_q\mathfrak{sl}(2)_i \quad (i = 1, 2)$,

The representation $\begin{smallmatrix}\square\square\\\square\end{smallmatrix}$ splits into the irreducible pieces with respect to $U_q\mathfrak{sl}(2)_1$ as follows.

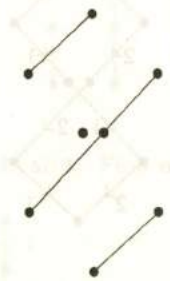


*Fig.*14

Also, it decomposes with respect to $U_q\mathfrak{sl}(2)_2$.

*Fig.*15

But these two decompositions are not compatible in the sense that the choices of two vectors in the weight space of weight 0 are different in Fig.14 and Fig.15. You must change linear combinations. But in the limit $q \to 0$, we have compatible decompositions for both colors. The general statement is as follows [13]. Let us consider an arbitrary symmetrizable Kac–Moody Lie algebra containing $U_q\mathfrak{sl}(2)_i \quad (i = 1, \cdots l)$ in it. Let M be an integrable highest weight module. Then, there exists a subset B of M with the following property. Define L by (48). The statement is that for each color i we can decompose M into irreducible components with respect to $U_q\mathfrak{sl}(2)_i$ in such a way that each component has basis consisting of those vectors which are equal to vectors in B modulo qL.

By this theorem we can unify Figures 14 and 15 to make B into a colored oriented graph.

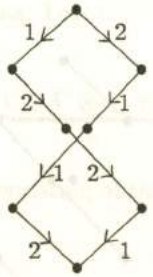


*Fig.*16

The set B is called the crystal base, and the graph is called the crystal graph.

Now we return to tensor product. I will give an answer to the question how to construct the crystal graph of tensor product.

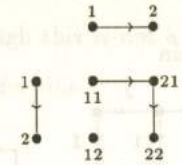
Since we can argue in each $U_q\mathfrak{sl}(2)_i$ separately, it is enough to answer the question for $U_q\mathfrak{sl}(2)$. First we consider $\square \otimes \square$. The crystal graph of each $\square$ is

*Fig.*17

The crystal graph for $\square \otimes \square$ is $B \times B$ as a set (before putting arrows on it). We know how to draw arrows by (44) and (45).

The 1st component

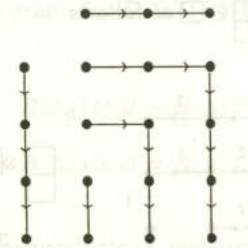
The 2nd component



$2 \times 2 = 3 + 1$

*Fig.*18

The general rule is similar. For example,



$3 \times 4 = 6 + 4 + 2$

*Fig.*19

Now let us consider $\square \otimes \square$ for $U_q\mathfrak{sl}(3)$. First of all, the crystal graph for $\square$ is

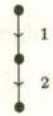


*Fig.*20

If we consider $U_q\mathfrak{sl}(2)_1$, we have

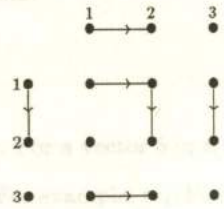


*Fig.*21

Similarly, $U_q\mathfrak{sl}(2)_2$ gives

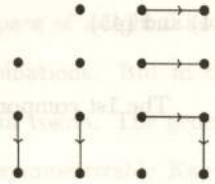


*Fig.*22

By unifying these two, we obtain

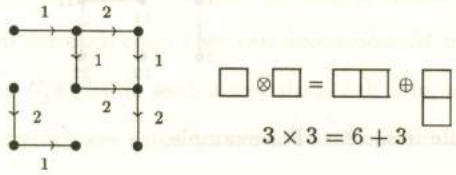


*Fig.*23

Similarly, we can decompose $\square \otimes \square$ as follows

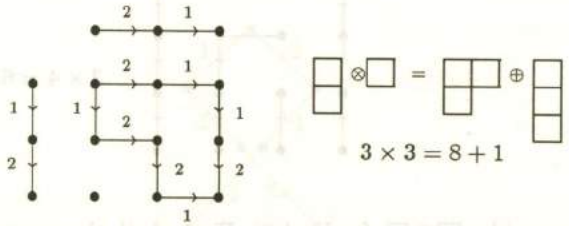


*Fig.*24

7. Paths and affine crystal graphs

In this section we combine the ideas in Sections 5 and 6. We will see that the paths for the 6 vertex model constitute the crystal graph for the basic representation $L(\Lambda_0)$ of $U_q\widehat{\mathfrak{sl}}(2)$ [14],[15]. Here I introduce Kashiwara's argument. In the construction of the R matrix for the 6 vertex model we started from the 2 dimensional representation of $U_q\widehat{\mathfrak{sl}}(2)$. Although this is not a highest weight representation, the following is the crystal graph for this.

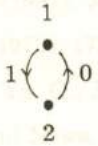


*Fig.*25

Let $B(\Lambda_i)$ denote the crystal graph for $L(\Lambda_i)$ $(i = 0, 1)$. We claim that

$$B(\Lambda_0) \otimes B = B(\Lambda_1) \tag{49}$$
$$B(\Lambda_1) \otimes B = B(\Lambda_0).$$

The left hand sides are tensor products of two crystal graphs. They are defined purely combinatorially by using the tensor product rule given in Fig.19.

Let us consider how one can get highest weight vectors in the product graph. Here I mean by the highest weight vectors those which have no incoming arrows. The condition is as follows. The tensor $b_1 \otimes b_2$ is a highest weight vector if and only if

(i) b_1 is a highest weight vector

and

(ii) $\epsilon_i(b_2) \leq \langle h_i, wt(b_1) \rangle$.

Let me explain the notations. For a vector b in a crystal graph, $\epsilon_i(b)$ is the length of the color i string above b. For example, $(\epsilon_1(b), \epsilon_2(b))$ for b in Fig.16 is as follows.

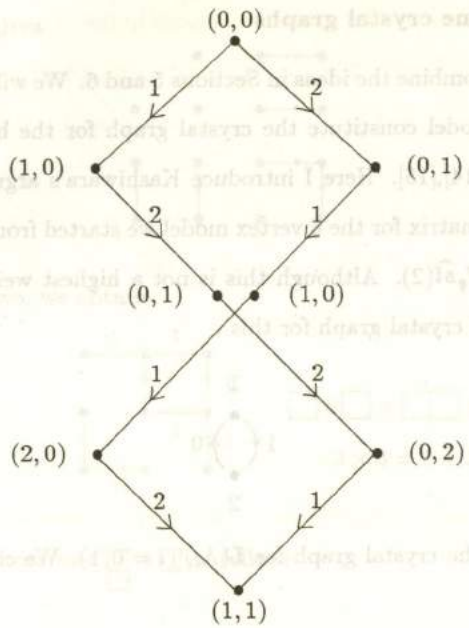


*Fig.*26

$wt(b)$ is the weight of b and h_i is a simple co-root. Therefore, $\langle h_i, wt(b)\rangle$ for a highest weight vector b is the length of the color i string below b.

Let us apply the above consideration to get the highest weight vectors $b_1 \otimes b_2$ in the tensor product $B(\Lambda_0) \otimes B$. From (i) we have $b_1 = v_{\Lambda_0}$, where v_{Λ_0} is the highest weight vector in the crystal graph of $L(\Lambda_0)$. Then from (ii) the unique choice of b_2 is $b_2 = 1$. (Here 1 in the right hand side means the vector 1 in Fig.25). The weight of the highest weight vector $v_{\Lambda_0} \otimes 1$ is Λ_1. Therefore, we can deduce that $B(\Lambda_0) \otimes B = B(\Lambda_1)$ (although this is not a proof). Applying (49) several times, we get

$$B(\Lambda_0) = B(\Lambda_0) \otimes B^{\otimes 2k} \tag{50}$$

for any k. Moreover, the highest weight vector on the right hand side is

$$v_{\Lambda_0} \otimes \underbrace{1 \otimes 2 \otimes \cdots \otimes 1 \otimes 2}_{2k}.$$

In the limit $k \to \infty$, the $B(\Lambda_0)\otimes$ and $v_{\Lambda_0}\otimes$ disappears. In such a way, we get the identification between the paths and the affine crystal graph.

References

[1] Baxter, R. J., Exactly solved models in statistical mechanics, Academic, London (1982).

[2] Drinfel'd, V. G., in Proc. of ICM 86, Berkeley (1987), 798.

[3] Jimbo, M., L.M.P. 10 (1985), 63.

[4] Jones, V. F. R., Ann. Math. 126 (1987), 335.

[5] Trotter, H. F., Inven. Math. 20 (1973), 173.

[6] Bazhanov, V. V. and Stroganov, Yu. G., J. Stat. Phys. 51 (1990), 799.

[7] Date, E., Jimbo, M., Miki, K. and Miwa, T., RIMS preprint 706, to appear in Publ. RIMS.

[8] Bazhanov, V. V., Kashaev, R. M., Mangazeev, V. V. and Stroganov, Yu. G., preprint (1990).

[9] Date, E., Jimbo, M., Miki, K. and Miwa, T., RIMS preprint 715 (1990), to appear in C.M.P.

[10] Date, E., Jimbo, M., Miki, K. and Miwa, T., RIMS preprint 729 (1990), to appear in Pacific J. of Math.

[11] Date, E., Jimbo, M., Kuniba, A., Miwa, T. and Okado, M., L.M.P. 17 (1989) 69.

[12] Date, E., Jimbo, M. and Miwa, T., in Physics and Mathematics of Strings, World Scientific, Singapore, (1990).

[13] Kashiwara, M., C.M.P. 133 (1990), 249.

[14] Misra, K. and Miwa, T., C.M.P. 134 (1990), 79.

[15] Jimbo, M., Misra, K., Miwa, T. and Okado, M., RIMS preprint 709 (1990), to appear in C.M.P.

INTEGRABILITY OF STRINGS

Satoru SAITO
Department of Physics, Tokyo Metropolitan University
Setagaya-ku, Tokyo 158 Japan†

ABSTRACT

The correspondence between the string theory and soliton theory is explored and a q − deformation of string theory which preserves integrability is discussed.

1. INTRODUCTION

In the last ten years much progress in theoretical physics has been made in the study of nonlinear problems. This was achieved by the development of systematic methods to solve certain class of nonlinear equations with no reference to perturbative methods. The soliton theory provided a number of approaches to investigate a broad class of nonlinear equations in various fields of physics[1]. These methods, such as the inverse scattering method, Bäcklund transformation method, Hirota's method, etc., are correlated each other and based on the geometry of two dimensional surface. Through the investigation of them the notion of completely integrable nonlinear systems became clear and the space of solutions were fully understood[2]. The inverse method was shown being a useful tool also to solve various two dimensional lattice models in statistical mechanics. Intensive studies during the last decade have uncovered rich structure of solvable lattice models and the connection with integrable nonlinear problems[3].

† *address after April 1991: Department of Physics, Tokyo Metropolitan University, Minamiohsawa 1-1, Hachiohji City, Tokyo, 192 Japan*

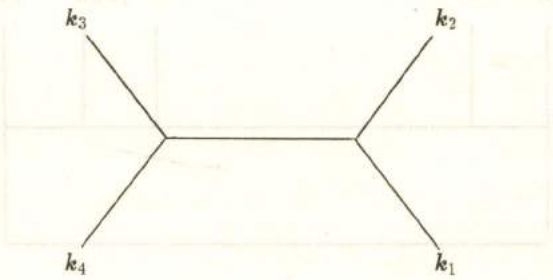
The importance of the integrable nonlinear systems was further explored when they were shown to describe the theory of particle physics. The intimate correspondence became aware[4,5] between the soliton theory and the string theory, which were completely uncorrelated fields before that time. The string theory is a theory supposed to describe everything in particle physics including the theory of gravity. It seems now quite obvious that the integrable nonlinear systems, which appear in various fields of physics, are universal. They share the same mathematical origin and play the fundamental role irrespective to the difference of subjects. The solutions to the integrable nonlinear equations are stable and behave regularly against perturbations. This means that we can deform equations somehow holding stability of solutions until we encounter chaos. The most of the parts of the space of equations are filled by chaos even if the equations are those of deterministic. Therefore it is a fundamental question to ask how far one can deform the integrable systems without loosing stability of solutions.

I would like to show in what sense the string theory is related with the soliton theory. This correspondence enables us to consider the string theory as an integrable system. It is then natural to ask how far one can extend the theory without violation of integrability. In the theory of solvable lattice models the Yang-Baxter relation plays an essential role and quantum groups appear quite naturally. I will discuss, in the latter part of this talk, a q-deformation of strings as one of possible extensions of string theory which preserve integrability.

2. STRING THEORY

The string theory was developed as a theory which describes all fundamental forces, including gravity, in particle physics in a unified way[6]. Since some of you may not be familier with this theory in detail I start with a brief review of the theory.

To begin with we consider a scattering of point particles associated with the diagram



The corresponding scattering amplitude is given by

$$\int_{t_1}^{t_4} dt_3 \int_{t_1}^{t_3} dt_2 \langle k_4 | e^{-i\int_{t_3}^{t_4} dt H_0} V(k_3) e^{-i\int_{t_2}^{t_3} dt H_0} V(k_2) e^{-i\int_{t_1}^{t_2} dt H_0} | k_1 \rangle \equiv \int_{t_1}^{t_4} dt_3 \int_{t_1}^{t_3} dt_2 F_4(k_1,t_1;k_2,t_2;k_3,t_3;k_4,t_4). \tag{1}$$

In the first quantized level the momentum eigenstate is defined by

$$|k\rangle \equiv e^{ikx}|0\rangle,$$

so that $p|k\rangle = k|k\rangle$ follows to $[x,p] = i$. Since our discussion is independent on the real space-time dimension we ignore their indices. The variable t_j should be interpreted as propertimes. In the simplest case the free Hamiltonian and the vertex operator are given by

$$H_0 = \frac{1}{2}p^2, \qquad V(k) = g e^{ikx}, \qquad g : const. \tag{2}$$

The latter corresponds to a point interaction without form factor. After a simple calculation we obtain

$$\frac{1}{T}\int_{t_1}^{t_4} dt_3 \int_{t_1}^{t_3} dt_2 F_4 = \frac{ig^2}{(k_1+k_2)^2 - m^2} e^{\frac{i}{2}Tm^2} \delta(k_1+k_2+k_3+k_4), \tag{3}$$

in the large $T = (t_1 - t_4)$ limit. Here m is the mass of the particles. The result shows that there exists a bound state of mass m in the $s = (k_1+k_2)^2$ channel and that's all.

If we replace the point particles by strings a very rich structure appears both in physics and mathematics. The free Hamiltonian is now

$$H_o = \frac{1}{2}p^2 + \sum_{n=1}^{\infty} n a_n^\dagger a_n, \qquad [a_m, a_n^\dagger] = \delta_{mn}, \tag{4}$$

and the vertex operator is given by (we put g = 1)

$$V(k,z) =: e^{ikX(z)} :\equiv e^{ikX^+(z)} e^{ikX^-(z)}, \tag{5}$$

where the string coordinate $X(z)$ replaces $x + pt$ of the particle:

$$x + pt \longrightarrow X(z) = X^+(z) + X^-(z), \qquad z = e^{-it},$$
$$X^-(z) = x + \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} a_n z^n, \qquad X^+(z) = ip\ln z + \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} a_n^\dagger z^{-n}. \tag{6}$$

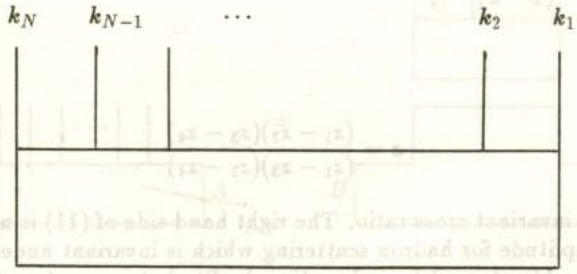
If we define the vacuum state $|0\rangle$ by $a_n|0\rangle = p|0\rangle = 0, \quad \forall n$, and use

$$V(k,z)V(k',z') = (z-z')^{kk'} : V(k,z)V(k',z') : \tag{7}$$

we find

$$F_N \equiv \langle 0 | \prod_j^{N\to 1} V(k_j, z_j) | 0 \rangle = \prod_{i<j} (z_i - z_j)^{k_i k_j}. \tag{8}$$

Here $\prod_j^{N\to 1}$ means the ordered product $\prod_j^{N\to 1} A_j \equiv A_N A_{N-1} \cdots A_2 A_1$, which we often write as $\prod_j$ for simplicity.



The first nice thing is that F_N behaves covariantly under the Möbius map of the variables (Koba-Nielsen variables[7])

$$z_j \longrightarrow z_j' = \frac{\alpha z_j + \beta}{\gamma z_j + \delta}, \qquad \alpha, \beta, \gamma, \delta \in \mathbf{C}. \tag{9}$$

Then we find

$$F_N(z_1', \cdots, z_N') = \prod_j \Big(\frac{\gamma z_j + \delta}{\sqrt{\alpha\delta - \beta\gamma}} \Big)^{m^2} F_N(z_1, \cdots, z_N), \tag{10}$$

which implies $\prod_j dz_j (z_j - z_{j+r})^{-a} F_N(z_1, \cdots, z_N)$ is invariant under this transformation if $a = 1 - \frac{1}{2}m^2$ for an integer $r (1 \leq r \leq N-1)$. Therefore we are tempted to define the scattering amplitude by

$$\int C \prod_{j=1}^{N} \frac{dz_j}{(z_{j+1} - z_{j-1})^a} F_N(z_1, \cdots, z_N),$$

where

$$C^{-1} = \frac{dz_a dz_b dz_c}{(z_b - z_c)(z_c - z_a)(z_a - z_b)}, \qquad z_a, z_b, z_c \in \{z_j\},$$

is the measure which is overcounted due to the symmetry. In the case of $N = 4$, choosing $z_a = z_1, z_b = z_3$ and $z_c = z_4$, we obtain

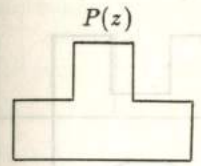
$$\begin{aligned} &\int_{z_1}^{z_3} dz_2 \frac{(z_3 - z_4)(z_4 - z_1)}{(z_2 - z_4)^2 (z_1 - z_3)} \Big(\frac{(z_1 - z_2)(z_3 - z_4)}{(z_1 - z_3)(z_2 - z_4)} \Big)^{k_1 k_2} \Big(\frac{(z_2 - z_3)(z_1 - z_4)}{(z_1 - z_3)(z_2 - z_4)} \Big)^{k_2 k_3} \\ &= \int_0^1 dx\, x^{k_1 k_2} (1 - x)^{k_2 k_3}, \end{aligned} \tag{11}$$

where

$$x = \frac{(z_1 - z_2)(z_3 - z_4)}{(z_1 - z_3)(z_2 - z_4)}$$

is the Möbius invariant cross ratio. The right hand side of (11) is nothing but the Veneziano amplitude for hadron scattering which is invariant under the exchange of k_1 and k_3 and/or k_3 and k_4, exhausting duality between s, t, u channels.

So far we have considered interaction of strings with particles in the ground state. In general we must investigate interactions of strings. The off-shell three string vertex operator is given by[8]



$$\begin{aligned} &\equiv W(P, g) \\ &=: \exp\Big[\frac{1}{2\pi} \oint \frac{dz}{z} P(z) X(g(z)) \Big] : \\ &=: \exp\Big[\frac{1}{2\pi i} \oint X(g(z)) dY(z) \Big] : \end{aligned} \tag{12}$$

where g is a function of z analytic at $z = 0$ and $P(z) = -iz\frac{d}{dz}Y(z)$ is an operator associated with an external string momentum whose coordinate $Y(z)$ consists of independent set of oscillators from those of X. If $P(z)$ is a constant k we notice that $W(P, g)$ becomes the vertex $V(k, g(0))$ for a ground state particle of momentum k.

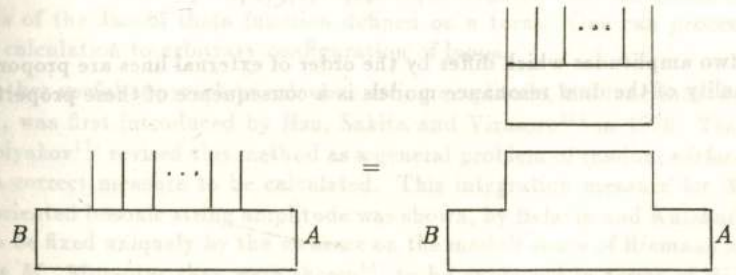
The use of this operator W should be justified in the following way[9]. We denote the states of Y-string by $|\cdot\rangle_Y$ and $:\exp[ikY(z)]:$ by $V_Y(k, z)$. Then we can show

$$V_Y(k, z) W(P, g)|0\rangle_Y = W(P, g)|0\rangle_Y V(k, g(z)), \tag{13}$$

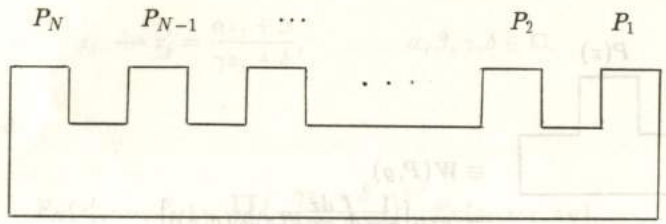
from which follows

$$\langle B| \prod_l V(k_l, g(z_l))|A\rangle \quad = \quad {}_Y\langle 0| \prod_l V_Y(k_l, z_l) \langle B|W(P, g)|A\rangle |0\rangle_Y. \tag{14}$$

This formula represents the situation depicted in the Figure;



Using W's we can write down N-string tree amplitudes as



$$\equiv \langle 0| \prod_j^{N\to 1} \otimes W(P_j, g_j)|0\rangle.$$

The three string vertex operator turns out to be nothing but $\langle 0|W\otimes W\otimes W|0\rangle$ and the propagator $\langle 0|W\otimes W|0\rangle$.

Important properties of these amplitudes are *factorizability* and the *crossing symmetry*. The factorizability is obvious in our operator formalism since the amplitudes are given by products of vertex operators. The crossing symmetry is a symmetry of the amplitude under the exchange of adjacent external particles (or strings). This owes to the fact that

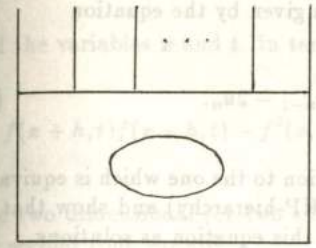
$$\begin{aligned} V(k_i,z_i)V(k_j,z_j) &= R_{ij}V(k_j,z_j)V(k_i,z_i), \\ R_{ij} &= e^{i\pi\epsilon(z_i-z_j)k_ik_j}, \end{aligned} \tag{15}$$

(similar relation for W's) which follows to (7), where

$$\epsilon(z-z') = \pm 1, \quad |z| \gtrless |z'|.$$

Hence two amplitudes which differ by the order of external lines are proportional. The duality of the dual resonance models is a consequence of these properties.

The calculation of general amplitudes is simplified by the use of our vertex operator W's. For instance the one loop amplitude with N-external ground state particles is given by calculating



$$= \sum_K \langle K| \prod_j^N V(k_j,z_j)|K(w)\rangle$$

where $|K(g)\rangle$ is a coherent state defined by

$$|K(g)\rangle = W(K,g^{-1})|0\rangle, \qquad \langle K(g)| = \langle 0|W(-K,g^{-1}).$$

The trace over K is the trace over the coherent state, *i.e.*, over the components of $K(z)$. After a little calculation we obtain

$$\begin{aligned} &\sum_K \langle K| \prod_j^N V(k_j,z_j)|K(w)\rangle \\ &= \Big(\frac{-\pi}{\ln w}\prod_{n=1}^{\infty}\frac{\pi}{1-w^n}\Big)^{1/2} \prod_{i<j}^N \Big(\exp\Big[-\frac{\pi^2\phi_{ij}}{\ln w}\Big]\frac{\theta_1(\phi_{ij},w^{1/2})}{\theta_1'(0,w^{1/2})}\Big)^{k_ik_j} \prod_j (2\pi i z_j)^{-\frac{1}{2}k_j^2}, \end{aligned} \tag{16}$$

where $\phi_{ij} = \frac{1}{2\pi i}\ln(z_i/z_j)$, and θ_1 is the Jacobi theta function. This expression agrees with the well-known result[10], up to the last factor which comes from the definition of the vertex operator. It is important to notice that the result is given in terms of the Jacobi theta function defined on a torus. One can proceed this type of calculation to arbitrary configuration of loops.

Another useful approach to calculate string amplitude, the functional integral method, was first introduced by Hsu, Sakita and Virasoro[11] in 1970. Ten years later Polyakov[12] revised this method as a general problem of random surfaces and found a correct measure to be calculated. This integration measure for M-loop closed oriented bosonic string amplitude was shown, by Belavin and Knizhnik[13] in 1986, to be fixed uniquely by the measure on the moduli space of Riemann surface of genus M. Moreover they were shown[14] to be expressed in terms of Riemann theta-functions. So far is the story about string models commonly understood before 1986 and necessary to our following discussions.

3. INTEGRABLE NONLINEAR SYSTEMS

We now change the subject to the soliton physics which was completely uncorrelated field with the string theory before 1986. The Toda lattice[15] is one of the well-studied systems which is quite simple, completely integrable and contains all characteristic features of soliton physics. It is given by the equation

$$\frac{d^2}{dt^2}\ln(1+u_n) = u_{n+1} + u_{n-1} - 2u_n. \tag{17}$$

The aim of this section is to generalize this equation to the one which is equivalent to infinitely many soliton equations (called the KP-hierarchy) and show that the string amplitudes are completely determined by this equation as solutions.

To begin with let us consider in (17) the following transformation of variables[16]

$$\tau = h^3 t, \qquad x = hn - (h - h^5)t, \qquad u_n(t) = h^2 u(x,\tau). \tag{18}$$

The equation becomes

$$\Big(\frac{\partial}{\partial\tau} - (h^{-2}-h^2)\frac{\partial}{\partial x}\Big)^2 \ln(1+h^2u(x,\tau)) = \frac{1}{h^4}\Big(u(x+h,\tau)+u(x-h,\tau)-2u(x,\tau)\Big). \tag{19}$$

In the limit of $h \longrightarrow 0$, this equation reduces to the KdV equation

$$\frac{\partial}{\partial\tau}u(x,\tau) + \frac{1}{4}\frac{\partial}{\partial x}u^2(x,\tau) + \frac{1}{24}\frac{\partial^3}{\partial x^3}u(x,\tau) = 0, \tag{20}$$

whereas it recovers the Toda equation when $h = 1$. Eq.(18) relates the discrete variable to a continuous variable. An important observation is that the equation (19) itself is completely integrable for an arbitrary value of h, as it was proved by means of the inverse scattering method[16]. The Toda equation and the KdV equation are special cases of integrable systems which are connected each other analytically through (19). Since equation (19) is derived from the Toda equation (17) via the variable transformations they are equivalent. We shall get nothing new by studying (19) other than known results in the Toda lattice. This example, however, shows the possibility that soliton equations are not independent but related through change of variables or derived from a single equation by taking various limits of parameters.

In order to see this, it is more convenient to use a new function f defined by

$$u(x,\tau) \equiv h^4\frac{\partial^2}{\partial t^2}\ln f(x,t),$$

and the variables x and t. In terms of these notation we have

$$f(x+h,t)f(x-h,t) - f^2(x,t) = h^2\Big(f(x,t)\frac{\partial^2}{\partial t^2}f(x,t) - \Big(\frac{\partial f(x,t)}{\partial t}\Big)^2\Big). \tag{21}$$

The two dimensional (or two time) Toda lattice is then obtained by introducing the second time variable s,

$$f(x+h,t,s)f(x-h,t,s) - f^2(x,t,s) = h^2\Big(f\frac{\partial^2}{\partial t^2}f - \Big(\frac{\partial f}{\partial t}\Big)^2 + f\frac{\partial^2}{\partial s^2}f - \Big(\frac{\partial f}{\partial s}\Big)^2\Big). \tag{22}$$

This equation is also known integrable. We are now tempted to generalize this equation to the following difference equation of three symmetric variables

$$\begin{aligned}\alpha f(x+h,t,s)f(x-h,t,s) + \beta f(x,t+\epsilon,s)f(x,t-\epsilon,s)\\ + \gamma f(x,t,s+\delta)f(x,t,s-\delta) = 0,\end{aligned} \tag{23}$$

where $\alpha+\beta+\gamma = 0$. This reduces to (22) in the $\epsilon, \delta \longrightarrow 0$ limit. Hirota derived[17] this equation in 1981 and showed that many soliton equations can be obtained by taking various limits of parameters in various combinations. Among them, besides the KdV and Toda equations, there are modified KdV equation, the sine-Gordon equation, Kadomtsev-Petviashivili (KP-) equation, Benjamin-Ono equation etc.. Furthermore it was proved by Miwa[18] that every solution to the KP-hierarchy satisfies Hirota's bilinear difference equation (HBDE) Eq.(23).

The KP-hierarchy is an infinite set of soliton equations including those above mentioned. General solutions (called τ-functions) to these equations have been known[2] and identified with points on an infinite dimensional Grassmann manifold. Explicit form of the τ-function is given by

$$\tau(t) = e^{Q(t)}\theta(\phi(t)). \tag{24}$$

Some explanations of this expression are in order. First t represents a set of infinite number of variables $(t_1, t_2, \cdots)$ of soliton equations. θ is the Riemann

theta-function whereas $\phi(t)$ and $Q(t)$ are given by the first and second kinds of Abel differentials which can be explicitly written, for an arbitrary vector ζ, as

$$\phi(t) = \zeta + \sum_{n=1}^{\infty} \frac{t_n}{2\pi i} \oint \omega z^{-n},$$
$$Q(t) = \frac{1}{2} \sum_{m,n=1}^{\infty} \frac{t_m t_n}{(m-1)!(n-1)!} \frac{\partial^m}{\partial x^m} \frac{\partial^n}{\partial y^n} \ln \frac{E(x,y)}{x-y}\Big|_{x=y=0}. \tag{25}$$

Here $\omega_k \quad (k = 1, 2, \cdots, M)$ and $E(x, y)$ are the first Abel differentials and the prime form associated to the Riemann surface of genus M. From the expression it is apparent that τ is a modular function defined on some Riemann surface. This form of solution was first shown by Krichever[19] to satisfy the KP-hierarchy. It was then proved by Mulase[20] and also Shiota[21] that every solution to the KP-hierarchy is given by (24). This means, at the same time, that the KP-hierarchy characterizes the Jacobi variety among Abel varieties thus solving the Schottky problem.

To see this it is most convenient to use the transformation of variables which owes to Miwa[18]:

$$t_n = \frac{1}{n} \sum_j k_j z_j^n, \qquad n = 1, 2, 3, \cdots. \tag{26}$$

Here $\{k_j, \quad j = 1, 2, \cdots\}$ are new variables and $\{z_j\}$ are parameters. We notice that this transformation is a generalization of the change of variables (18) used to derive the KdV equation from the Toda lattice. In terms of the new variables, (24) becomes

$$\tau(k) = \prod_{i,j=1}^{\infty} \left(\frac{E(z_i, z_j)}{z_i - z_j}\right)^{\frac{1}{2}k_i k_j} \theta\Big(\zeta + \sum_{j=1}^{\infty} k_j w(z_j)\Big). \tag{27}$$

If we choose any three of k's, say k_1, k_2, k_3, out of $\{k_j\}$ and define

$$\frac{x}{h} = k_2 + k_3 - 1, \quad \frac{t}{\epsilon} = k_3 + k_1 - 1, \quad \frac{s}{\delta} = k_1 + k_2 - 1,$$

the HBDE becomes (we write $f(x, t, s)$ as $f(k_1, k_2, k_3)$)

$$\alpha f(k_1+1, k_2, k_3) f(k_1, k_2+1, k_3+1) + \beta f(k_1, k_2+1, k_3) f(k_1+1, k_2, k_3+1) + \gamma f(k_1, k_2, k_3+1) f(k_1+1, k_2+1, k_3) = 0. \tag{28}$$

The substitution of $\tau(k)$ into HBDE leads to

$$\begin{aligned} &E(z_0, z_1)E(z_2, z_3)\theta(\zeta' + w(z_0) + w(z_1))\theta(\zeta' + w(z_2) + w(z_3)) \\ &+E(z_0, z_2)E(z_3, z_1)\theta(\zeta' + w(z_0) + w(z_2))\theta(\zeta' + w(z_3) + w(z_1)) \\ &+E(z_0, z_3)E(z_1, z_2)\theta(\zeta' + w(z_0) + w(z_3))\theta(\zeta' + w(z_1) + w(z_2)) = 0, \end{aligned} \tag{29}$$

where

$$w(z_j) = \int_{z_0}^{z_j} \omega,$$

and

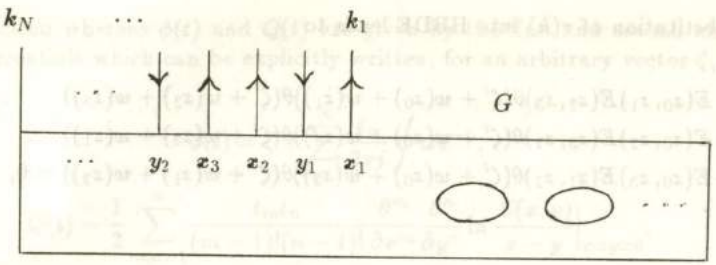
$$\zeta' = \zeta + \sum_j k_j w(z_j) - w(z_0).$$

What we have just obtained is called Fay's trisecant formula[22] for algebraic curves. It is remarkable that this formula holds irrespective to the variables k_j's. The formula is an identity as far as $w(z_j)$'s are elements of a Jacobi variety. Therefore $\tau(k)$ solves HBDE.

4. STRING-SOLITON CORRESPONDENCE

From above arguments we know that in one hand the string amplitudes are modular functions defined on Jacobi variety, while, on the other hand, HBDE characterizes the Jacobi variety and its solution is given by modular functions on the Jacobi variety. Therefore we are tempted to conjecture[4,5] that the string amplitudes satisfy the HBDE and HBDE characterizes all of the string amplitudes. Only the problem which remains is what are the variables in string amplitudes corresponding to the variables of soliton equations. The answer is that the momenta k_j of external particles in the string amplitudes are related to the variables $\{t_n\}$ through (26) and the parameters $\{z_j\}$ in (26) are nothing but the Koba-Nielsen variables in the string amplitudes which specify the positions of external lines on the world sheet.

To see this first we recall that the string amplitudes with N-external particles are given, using the factorizability and crossing symmetry, in the form $F_{N,G} = \langle 0| \prod_j^N V(k_j, z_j) G |0\rangle$. Here G is an operator which represents internal configuration of loop diagrams.



We can determine G from the information of solution to HBDE as follows. Consider $\tau(k)$ of (27) in the case such that each k_j takes values ± 1 under the condition $\sum_j^N k_j = 0$. Let $x_1, x_2, \cdots, x_{N/2}$ $(y_1, y_2, \cdots, y_{N/2})$ be the local coordinates associated with positive (negative) k_j's. Then $F_N\tau(k)$ becomes (F_N being the N-point tree amplitude)

$$\frac{\prod_{i<j}^{N/2} E(x_i,x_j)E(y_i,y_j)}{\theta(\zeta)\prod_{i,j}E(x_i,y_j)}\theta\Big(\zeta-\sum_{j=1}^{N/2}\int_{x_j}^{y_j}\omega\Big)=\det\Big[\frac{\theta\Big(\zeta-\int_{x_i}^{y_j}\omega\Big)}{\theta(\zeta)E(x_i,y_j)}\Big]. \qquad (30)$$

The second expression owes to Fay's addition theorem[22] for algebraic curves. This determinant can be further expressed by using fermionic fields $\psi(z)$ and $\bar\psi(z)$ as

$$\langle 0|\prod_j^{N/2}\psi(x_j)\bar\psi(y_j)\exp\Big[\frac{1}{4\pi^2}\oint dx\oint dy\bar\psi(y)\Big(\frac{\theta(\zeta-\int_x^y\omega)}{\theta(\zeta)E(x,y)}-\frac{1}{y-x}\Big)\psi(x)\Big]|0\rangle.$$

This is the τ-function discussed in Refs.23) in connection with the correspondence between fermionic string amplitudes and the KP-hierarchy. But now remember that the τ-function satisfies HBDE for arbitrary values of k_j's. The restriction of the values of k_j to ± 1 are too strong. I like to show[4,5,24] that the τ-function with any values of k_j really correspond to bosonic string amplitudes with k_j being momenta of external particles.

For this purpose we notice the fact that the bosonic vertex operators $V(\pm 1, z)$ behave like fermionic fields, as can be seen from (7). Therefore we identify $\psi(z)$ and $\bar\psi(z)$ with V's:

$$\psi(z) = V(1,z), \qquad \bar\psi(z) = V(-1,z),$$

and define

$$G=\exp\Big[\frac{1}{4\pi^2}\oint dx\oint dy V(-1,y)\Big(\frac{\theta(\zeta-\int_x^y\omega)}{\theta(\zeta)E(x,y)}-\frac{1}{y-x}\Big)V(1,x)\Big]. \qquad (31)$$

Owing to the factorizability of the string amplitudes the operator which describes

the loop configuration of the amplitudes does not depend on external particle states. Hence it must be given by G of (31).

We show, in the following, that

$$\tau_G(k) = \langle 0| : \prod_j V(k_j, z_j) : G|0\rangle$$

solves HBDE. The proof goes parallel to the argument in Ref.2). Because of the fermionic nature of $V(\pm 1, z)$ we can prove

$$\begin{aligned} I &= \oint dz\langle 0| : \prod_j^N V(k_j,z_j) : V(1,z)G|0\rangle\langle 0| : \prod_j V(k'_j,z_j) : V(-1,z)G|0\rangle \\ &= \oint dz\langle 0| : \prod_j V(k_j,z_j) : GV(1,z)|0\rangle\langle 0| : \prod_j V(k'_j,z_j) : GV(-1,z)|0\rangle. \end{aligned}$$

The right hand side, however, vanishes upon integration over z along a closed contour around $z = 0$. By pushing $V(\pm 1, z)$ to the left, on the other hand, we find

$$I=\oint dz\prod_j(z-z_j)^{k_j-k'_j}\langle 0|e^{i\sum_j k_jX^-(z_j)+iX^-(z)}G|0\rangle\langle 0|e^{i\sum_j k'_jX^-(z_j)-iX^-(z)}G|0\rangle.$$

Let us choose the variables k_j's and k'_j's such that $k'_0 = k_0 - 1$, $k'_j = k_j + 1$ for $j = 1,2,3$ and $k'_j = k_j$ otherwise. We obtain

$$\begin{aligned} 0 &= \frac{1}{2\pi i}\oint dz\frac{z-z_0}{(z-z_1)(z-z_2)(z-z_3)} \\ &\quad\times\langle 0|e^{i\sum_j k_jX^-(z_j)+iX^-(z)}G|0\rangle\langle 0|e^{i\sum_j k'_jX^-(z_j)-iX^-(z)}G|0\rangle \\ &= \frac{z_0-z_1}{(z_1-z_2)(z_1-z_3)}\tau_G(k_0,k_1+1,k_2,k_3)\tau_G(k_0-1,k_1,k_2+1,k_3+1) \\ &+\frac{z_0-z_2}{(z_2-z_3)(z_2-z_1)}\tau_G(k_0,k_1,k_2+1,k_3)\tau_G(k_0-1,k_1+1,k_2,k_3+1) \\ &+\frac{z_0-z_3}{(z_3-z_1)(z_3-z_2)}\tau_G(k_0,k_1,k_2,k_3+1)\tau_G(k_0-1,k_1+1,k_2+1,k_3) \end{aligned} \qquad (32)$$

We already know that (32) is the HBDE, thus the proof is completed.

In terms of τ_G the string amplitude for N-external ground state particles with loop configuration associated with G is given by

$$F_{N,G} = \langle 0| \prod_j^N V(k_j, z_j) G |0\rangle = F_N(k) \tau_G(k), \tag{33}$$

which is a modular function defined on a Jacobi variety.

Another expression of G, hence of amplitude, is possible[24] by using string coordinate $X(z)$. To see this we notice that, since $\langle 0| : \prod_j V(k_j, z_j) :$ is a coherent state,

$$\langle 0| : \prod_j V(k_j, z_j) : dX^+(z) = \langle 0| : \prod_j V(k_j, z_j) : \sum_j k_j \frac{idz}{z_j - z},$$

holds. Thus we can write

$$G = \theta\Big(\zeta - \frac{1}{2\pi} \oint dX(z) w(z)\Big) \exp\Big[\frac{1}{8\pi^2} \oint dX(x) \oint dX(y) \ln \frac{E(x,y)}{x-y}\Big]. \tag{34}$$

Here $\frac{E(x,y)}{x-y}$ and $w(z)$ in G are assumed being expanded in positive powers of arguments. The expression(34) of G is given by the off-shell vertex operator W alone.

Finally I like to mention that explicit calculation of bosonic N-point amplitudes with arbitrary loops was performed[25] by using the off-shell vertex oprator W and gave the same expression of ours.

5. INTEGRABLE DEFORMATION OF STRINGS

As we have seen in the previous sections the integrability of string amplitudes is a consequence of duality. The latter determines the amplitudes uniquely when external lines and the number of internal loops are given. In our formulation the amplitudes are given by correlation functions of the vertex operators and the duality is implied by the crossing symmetry (15) between vertices. In this section we are interested in studying the symmetries which underly this property of the amplitudes. As we have seen before the amplitudes thus obtained remain invariant under the change of variables (9) which are generated by the three generators

$$L_m = z^m (z \frac{\partial}{\partial z} + a_m), \qquad m = 0, \pm 1, \tag{35}$$

of $SL(2, C)$ algebra

$$[L_+, L_-] = -2L_0, \qquad [L_0, L_\pm] = \pm L_\pm, \tag{36}$$

if $a_1 + a_{-1} = 2a_0$. This symmetry enabled people to write down the scattering amplitudes at tree level. It was found by Virasoro[26] that the amplitudes possess

much higher symmetry generated by the operators (35) with arbitrary m. They form a closed algebra, *i.e.*, the Virasoro algebra

$$[L_m, L_{m'}] = (m' - m) L_{m+m'}, \tag{37}$$

if

$$(m' - m) a_{m+m'} = m' a_{m'} - m a_m.$$

This symmetry corresponds to the general coordinate transformation $z \longrightarrow g(z)$, g; analytic at $z = 0$.

Since we can write the vertex operator as

$$V(k, z) = \lim_{\bar{z} \to z} \Big(\frac{1}{z} - \frac{1}{\bar{z}}\Big)^{k^2/2} e^{ikX(z)}, \tag{38}$$

we see that it transforms covariantly as

$$z^{k^2} V(k, z) \longrightarrow \Big(\frac{dg(z)}{dz}\Big)^{k^2/2} g(z)^{k^2} V(k, g(z)), \tag{39}$$

under the transformation. Therefore the string theory is a conformal field theory of fields with conformal dimension $k^2/2$.

Moreover, since the string amplitudes are solution to soliton equations the string theory has a symmetry possessed by the soliton theory. Such symmetry is quite large and has been known[1,2] to be $GL(\infty)$.

We are interested in a deformation of the string models from conformal field theory, which preserves integrability. The deformation must mix holomorphic and antiholomorphic parts of the amplitudes in non-trivial way[27,28]. Then the string coordinate $X(z)$ becomes a function of $\zeta = \rho z$, $\bar{\zeta} = \rho z^{-1}$, *i.e.*,

$$X(\zeta, \bar{\zeta}) = x + \frac{i}{2} \ln(\zeta/\bar{\zeta}) + \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} (a_n \zeta^n + a_n^\dagger \bar{\zeta}^n). \tag{40}$$

When $\rho = 1$ we recover $X(z)$.

Let us pause here and recall what is meant by integrability of the string theories. They are integrable in the sense that they are characterized by completely integrable equations. The integrability of the latter was established by means of the inverse scattering method. The Koba-Nielsen variables are nothing but eigenvalues of the inverse problems. The inverse method, on the other hand, was used

to solve statistical models on two dimensional lattice. The solvability condition of the lattice models is given by the Yang-Baxter relation for the transfer matrices. The eigenvalues of the inverse problem appear as parameters through the Boltzmann weight. From these connections, together with the fact that the integrability of the strings is implied by the crossing relation of vertex operators, a Yang-Baxter relation of strings must result from braid relations of three vertices[28].

If B_{ij} denotes the operator which represents the braid of two adjacent vertices $V(k_i, z_i)$ and $V(k_j, z_j)$ when z_j moves from the region $|z_i| < |z_j|$ to $|z_j| < |z_i|$ anticlockwise it is clear that

$$\begin{aligned} B_{i,i+1}B_{j,j+1} &= B_{j,j+1}B_{i,i+1} \qquad if \quad |i-j| \geq 2 \\ B_{ij}B_{ik}B_{jk} &= B_{jk}B_{ik}B_{ij} \qquad if \quad j = i+1, k = j+1. \end{aligned} \tag{41}$$

They are the relations of the braid group. The Yang-Baxter relation is induced by the braid group to include arbitrary parameters.

The introduction of ρ deforms R_{ij} in the crossing (or braid) relation (15) into

$$R_{ij} = \left(\frac{z_i - \rho^2 z_j}{z_j - \rho^2 z_i}\right)^{k_i k_j}. \tag{42}$$

Let u be $\ln(z_j/z_i)$ and write $R_{ij} = R_{ij}(u)$ in (42) and apply the braid operations to $V(k_i, z_i)V(k_j, z_j)V(k_k, z_k)$ $(\equiv V_iV_jV_k)$ into the reversed order in two different ways. We find

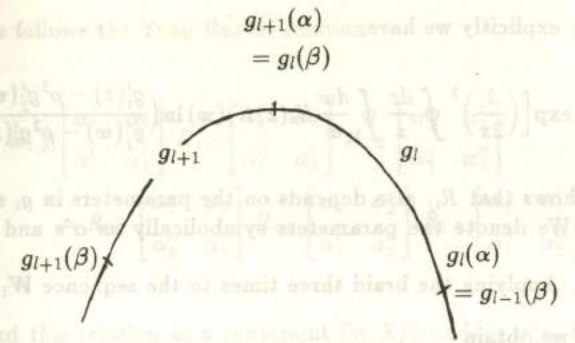
$$\begin{aligned} V_iV_jV_k &= R_{ij}(u)R_{ik}(v)R_{jk}(v-u)V_kV_jV_i \\ &= R_{jk}(v-u)R_{ik}(v)R_{ij}(u)V_kV_jV_i, \end{aligned} \tag{43}$$

where $v = \ln(z_k/z_i)$. The relation we have just obtained depends on z's, the Koba-Nielsen variables. It is, however, not the Yang-Baxter relation, since R's are not matrices but c-numbers. Nontrivial Yang-Baxter relations are obtained if we go beyond the scattering of ground state particles.

Let us consider a sequence of three vertices W_i, W_j, W_k in this order. The order of W's is determined by the order of $g_l(0) = z_l$, which we assume $|z_i| < |z_j| < |z_k|$. Further we assume that the l-th external momenta K_l lies along the line between $z = \alpha$ and $z = \beta$ and g_l satisfies the connection conditions

$$g_{l+1}(\alpha) = g_l(\beta), \quad \forall l.$$

This condition suffices that momenta of two adjacent vertices are joined properly.



Now the braid $B_{l,l+1}$ of W_{l+1} and W_l should bring g_l and g_{l+1} into g'_l and g'_{l+1};

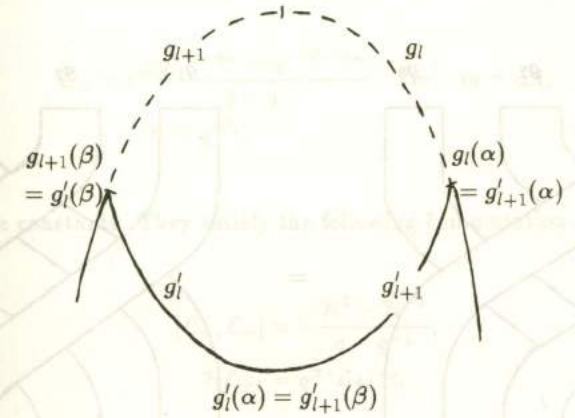
$$\begin{aligned} W(K_{l+1}, g_{l+1})W(K_l, g_l) &\longrightarrow W(K_{l+1}, g'_{l+1})W(K_l, g'_l) \\ &= R_{l+1,l}W(K_l, g'_l)W(K_{l+1}, g'_{l+1}), \end{aligned} \tag{44}$$

such that

$$g'_l(0) = g_{l+1}(0) = z_{l+1}, \quad g'_{l+1}(0) = g_l(0) = z_l,$$

together with the connection conditions

$$g'_l(\beta) = g_{l+1}(\beta), \quad g'_{l+1}(\alpha) = g_l(\alpha), \quad g'_l(\alpha) = g'_{l+1}(\beta).$$



Therefore g'_l is determined depending on parameters in g_l and g_{l+1}.

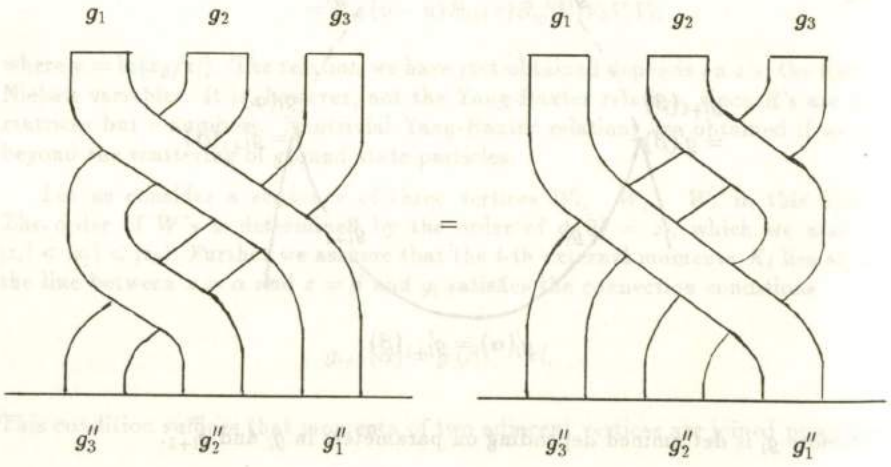
Writing R_{ij} explicitly we have

$$R_{i,j} = \exp\Big[\Big(\frac{1}{2\pi}\Big)^2 \oint \frac{dz}{z} \oint \frac{dw}{w} K_i(z) K_j(w) \ln\Big(\frac{g_i'(z) - \rho^2 g_j'(w)}{g_j'(w) - \rho^2 g_i'(z)}\Big)\Big], \tag{45}$$

which clearly shows that R_{ij} also depends on the parameters in g_i and g_j as well as g_i' and g_j'. We denote the parameters symbolically as α's and write $R_{i,j} = R_{i,j}\begin{bmatrix} \alpha_i & \alpha_j \\ \alpha_i' & \alpha_j' \end{bmatrix}$. Applying the braid three times to the sequence $W_1 W_2 W_3$ in two possible orders we obtain

$$\begin{aligned} &W(K_1, g_1) W(K_2, g_2) W(K_3, g_3) \\ &= R_{1,2}\begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_1' & \alpha_2' \end{bmatrix} R_{1,3}\begin{bmatrix} \alpha_1' & \alpha_3 \\ \alpha_1'' & \alpha_3' \end{bmatrix} R_{2,3}\begin{bmatrix} \alpha_2' & \alpha_3' \\ \alpha_2'' & \alpha_3'' \end{bmatrix} W(K_3, g_3'') W(K_2, g_2'') W(K_1, g_1'') \\ &W(K_1, g_1) W(K_2, g_2) W(K_3, g_3) \\ &= R_{2,3}\begin{bmatrix} \alpha_2 & \alpha_3 \\ \alpha_2' & \alpha_3' \end{bmatrix} R_{1,3}\begin{bmatrix} \alpha_1 & \alpha_3' \\ \alpha_1' & \alpha_3'' \end{bmatrix} R_{1,2}\begin{bmatrix} \alpha_1' & \alpha_2' \\ \alpha_1'' & \alpha_2'' \end{bmatrix} W(K_3, g_3'') W(K_2, g_2'') W(K_1, g_1''). \end{aligned} \tag{46}$$

The situation is shown in the Figure;



From this follows the Yang-Baxter relation

$$\begin{aligned} &R_{1,2}\begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_1' & \alpha_2' \end{bmatrix} R_{1,3}\begin{bmatrix} \alpha_1' & \alpha_3 \\ \alpha_1'' & \alpha_3' \end{bmatrix} R_{2,3}\begin{bmatrix} \alpha_2' & \alpha_3' \\ \alpha_2'' & \alpha_3'' \end{bmatrix} \\ &= R_{2,3}\begin{bmatrix} \alpha_2 & \alpha_3 \\ \alpha_2' & \alpha_3' \end{bmatrix} R_{1,3}\begin{bmatrix} \alpha_1 & \alpha_3' \\ \alpha_1' & \alpha_3'' \end{bmatrix} R_{1,2}\begin{bmatrix} \alpha_1' & \alpha_2' \\ \alpha_1'' & \alpha_2'' \end{bmatrix}. \end{aligned} \tag{47}$$

We can regard this relation as a constraint for K_l's and/or g_l's. For the configuration satisfying the constraint the correlation functions of W's can be calculated explicitly.

6. q-DEFORMATION OF VIRASORO ALGEBRA

We now consider the symmetry of the deformed strings associated with the Virasoro algebra. It must contain the operator $q^{\rho\partial_\rho}$ which generates the deformation and must reduce into the Virasoro algebra (37) in the conformal ($q \to 1$) limit. To find such symmetry we recall that the study of the Möbius map, or $SL(2,C)$ symmetry, led people to find the Virasoro algebra in the old string models. In the present problem the corresponding object is the q-deformed $SL(2,C)$ algebra, which has been extensively studied[29]. The generators of $SL_q(2,C)$ are given by

$$\begin{aligned} \mathcal{L}_m &= z^m \frac{q^{z\partial_z + a_m} - q^{-z\partial_z - a_m}}{q - q^{-1}} \quad for \quad m = \pm 1, \\ \mathcal{H} &= q^{z\partial_z}, \end{aligned} \tag{48}$$

where a_m are constants. They satisfy the following commutation relations

$$\begin{aligned} [\mathcal{L}_+, \mathcal{L}_-] &= -\frac{\mathcal{H}^2 - \mathcal{H}^{-2}}{q - q^{-1}}, \\ \mathcal{H}\mathcal{L}_{\pm 1} &= q^{\pm 1}\mathcal{L}_{\pm 1}\mathcal{H}, \end{aligned} \tag{49}$$

when $a_+ + a_- = 0$. The importance of this algebra lies on the fact that it is a Hopf algebra, hence defines a coalgebra. This additional structure enables us to operate the symmetry transformations consistently to states of tensor products.

An algebra A is called a Hopf algebra if there exist operations, a coproduct Δ, a counit ϵ and an antipode γ, which operate as

$$\begin{aligned}\Delta(ab) &= \Delta(a)\Delta(b),\\ \epsilon(ab) &= \epsilon(a)\epsilon(b),\\ \gamma(ab) &= \gamma(b)\gamma(a),\end{aligned} \tag{50}$$

and satisfy the following criteria of a Hopf algebra;

$$\begin{aligned}(id\otimes\Delta)\Delta(a) &= (\Delta\otimes id)\Delta(a),\\ (id\otimes\epsilon)\Delta(a) &= (\epsilon\otimes id)\Delta(a) = a,\\ m(id\otimes\gamma)\Delta(a) &= m(\gamma\otimes id)\Delta(a) = \epsilon(a),\end{aligned} \tag{51}$$

for any elements a, b of A. In the last equation m means $m(a\otimes b) = ab$. Since the states of strings in our interest are of the form of tensor products the Hopf structure will play a central role in the theory of the q-deformed string theories.

Possible extensions of $SL_q(2,C)$ to the Virasoro algebra were studied by some groups[30,31,32,33]. For instance Curtright and Zachos[31] found that the generators

$$\mathcal{L}^{CZ}_m = z^m\frac{q^{2z\partial_z}-1}{q-q^{-1}}, \qquad m\in\mathbf{Z}. \tag{52}$$

satisfy a q-deformed Virasoro algebra;

$$q^{m-m'}\mathcal{L}^{CZ}_m\mathcal{L}^{CZ}_{m'} - q^{m'-m}\mathcal{L}^{CZ}_{m'}\mathcal{L}^{CZ}_m = [m-m']\mathcal{L}^{CZ}_{m+m'},$$

where $[x]$ means the q-valued x defined by

$$[x] = \frac{q^x - q^{-x}}{q-q^{-1}}.$$

Some of these algebras can be extended to have a central charge. Such possibilities were studied by Chaichian, Ellinas and Popowicz[34] and also by Aizawa and Sato[35]. However, they do not have a total Hopf algebra structure, or it has not been given yet explicitly.

We want to construct a q-analogue of the Virasoro algebra which is a Hopf algebra[32]. A natural extension of $SL_q(2,C)$ into infinite dimensional case is to accept $\mathcal{L}_m$ in (48) for all $m\in\mathbf{Z}$. Then we immediately find that a commutator, say $[\mathcal{L}_1,\mathcal{L}_2]$, requires a new generator in which q is replaced by other value in some power of q. Let us include all of them into our algebra and define

$$\mathcal{L}_{mn} = z^m\frac{[nz\partial_z + a_{mn}]}{[n]}. \tag{53}$$

Notice that the suffix n appears always together with q. In other words we are considering various deformations with various powers of q simultaneously.

We can readily convince ourselves that they form a closed algebra

$$[\mathcal{L}_{mn},\mathcal{L}_{m'n'}] = C^{mn}_{m'n'}\mathcal{L}_{m+m',n+n'} + C^{mn}_{m',-n'}\mathcal{L}_{m+m',n-n'}, \tag{54}$$

where

$$C^{mn}_{m'n'} = \frac{[(nm'-n'm)/2][n+n']}{[n][n']} \tag{55}$$

and become the Virasoro algebra in the $q\to 1$ limit, irrespective to n and n'.

The generators of $SL_q(2,C)$ correspond to $\mathcal{L}_{\pm1,1}$ and the odd looking generator $\mathcal{H} = q^{z\partial_z}$ is now included naturally in $\mathcal{L}_{0,n}$ and does not play any special role. We should emphasize that this is a usual (although infinite dimensional) algebra depending on a parameter q.

The equation (54) is correct if $a_{mn} = n(\frac{1}{2}m+a)$, $\forall a$. We can, however, relax this condition to

$$q^{a_{mn}+a_{m'n'}-a_{m+m',n+n'}} = \sqrt{\frac{q^{s-r-mn'}-q^{-m'n}}{q^{m'n}-q^{s-r+mn'}}},$$

if we define the commutator in (54) by $[A,B]\equiv q^rAB - q^sBA$.

The central charge will be obtained by solving†

$$C^{mn}_{m'n'}\phi(m+m',n+n',m'',n'') + C^{mn}_{m,-n'}\phi(m+m',n-n',m'',n'') + \textit{cyclic perm.}\ (m,n;m',n';m'',n'') = 0$$

for ϕ, although an explicit form has not been given yet.

† This formula owes to H.Hiro-oka.

So far the algebra has nothing to do with the deformation of strings, since $\mathcal{L}_{mn}$ do not mix the holomorphic and anti-holomorphic sections. We need an operator which mix $\zeta = \rho z$ and $\bar{\zeta} = \rho z^{-1}$. For this purpose we define

$$\mathcal{K}_l = q^{l\rho\partial_\rho}. \tag{56}$$

This takes account deformation in the direction of ρ in various steps. Accordingly we also deform $\mathcal{L}_{mn}$ into

$$\mathcal{L}_{mn} = \zeta^m \frac{[nz\partial_z + a_{mn}]}{[n]}. \tag{57}$$

They satisfy

$$[\mathcal{K}_l, \mathcal{K}_{l'}] = 0, \quad \forall l, l', \qquad \mathcal{K}_l \mathcal{L}_{mn} = q^{lm} \mathcal{L}_{mn} \mathcal{K}_l, \quad \forall l, m, n. \tag{58}$$

For (54) and (58) to form a q-deformation of the Virasoro algebra we have to demonstrate that it is a Hopf algebra. In fact we can show

$$\begin{aligned} \Delta(\mathcal{L}_{mn}) &= \mathcal{L}_{mn} \otimes \mathcal{K}_m + \mathcal{K}_m \otimes \mathcal{L}_{mn}, & \Delta(\mathcal{K}_l) &= \mathcal{K}_l \otimes \mathcal{K}_l, \\ \epsilon(\mathcal{L}_{mn}) &= 0, & \epsilon(\mathcal{K}_l) &= 1, \\ \gamma(\mathcal{L}_{mn}) &= -\mathcal{K}_m^{-1} \mathcal{L}_{mn} \mathcal{K}_m^{-1}, & \gamma(\mathcal{K}_l) &= \mathcal{K}_l^{-1}, \end{aligned} \tag{59}$$

satisfy the criteria and that (54) and (58) are invariant under these operations. Hence our q-Virasoro algebra is really a q-deformation of the Virasoro algebra.

The symmetry which we have just found guarantees that the crossing matrices R_{ij} satisfy the Yang-Baxter relation and hence the integrability of the q-deformed string amplitudes.

In conclusion we have discussed the correspondence between the string theory and the soliton theory. They were developed independently in two uncorrelated fields. It was, however, shown that they are characterized by the same equation. This equation is completely integrable and describes a broad class of theories in various different fields. We have then discussed a possible extension of this system which preserves integrability, although there remain many problems to be further studied.

Acknowledgements

I would like to present my sincere thanks to Professors Z.Popowicz and P. Garbaczewski and other members of Institute of Theoretical Physics, University of Wrocław, for invitation and their warm hospitality during my stay in Karpacz. I would also like to thank Yoshida Foundation for Science and Technology for the financial support of my travel.

REFERENCES

1. See for example "Nonlinear Integrable Systems" *ed.* M.Jimbo and T.Miwa (World Scientific, 1983);
   Newell, A., "Solitons in Mathematics and Physics" (Society for Industrial and Applied Mathematics, Philadelphia, 1985).
2. Sato,M. and Sato,Y., Publ. Res. Inst. Math. Sci. (Kyoto Univ.) 388, 183 (1980); *ibid.* 414, 181 (1981); Sato, M., *ibid.* 433, 30 (1981),
   Date,E., Jimbo,M., Kashiwara,M. and Miwa,T., J. Phys. Soc. Jpn 50, 3806, 3813 (1981).
3. See, for example, "Yang-Baxter Equation in Integrable Systems", *ed.* M.Jimbo (World Scientific, Singapore, 1990).
4. Saito,S., Tokyo Metropolitan Univ. preprint TMUP-HEL-8613, 8615 (1986); Phys. Rev. *D36*, 1819 (1987); Phys. Rev. *D37*, 990 (1988) ;
   in "Strings '88" *ed.* S.J.Gates,Jr., C.R.Preitschopf and W.Siegel (World Scientific, 1989) p.436 ;
   Saito,S. and Kato,H., in "Nonlinear Physics", *ed.* Chaohao Gu, Yishen Li and Gui Zhang Tu (Springer-Verlag, Berlin, 1990) p.258.
5. Saito, S., Phys. Rev. Lett. 59, 1798 (1987).
6. See, for example, Green, M.B., Schwarz,J.H. and Witten,E., "Superstring Theory" (Cambridge University Press, Cambridge, 1987)
7. Koba, Z. and Nielsen, H.B., Nucl. Phys. *B10*, 633 (1969); *ibid.* *B12*, 517 (1969).
8. Della Selva, A. and Saito, S., Lett. Nuovo Cim. 4, 689 (1970).
9. Saito, S., Phys. Lett. *34B*, 72 (1971); Kosterlitz, J.M. and Saito, S., Nucl. Phys. *B34*, 557 (1971); Saito, S. and Kosterlitz, J.M., Nucl. Phys. *B38*,

269 (1972).
10. Frye, G. and Susskind, L., Phys. Lett. 31*B*, 537 (1970);
Gross, D.J., Neveu, A., Scherk, J. and Schwarz, J.H., Phys. Rev. *D*2, 697 (1970); Shapiro, J., Phys. Rev. *D*5, 1945 (1972).
11. Hsu, C.S, Sakita, B. and Virasoro, M.A., Phys. Rev. *D*2, 2857 (1970).
12. Polyakov, A., Phys. Lett. 103*B*, 207, 211 (1981).
13. Belavin, A.A. and Knizhnik, V., Phys. Lett. 168*B*, 201 (1986).
14. Manin, Yu.I., Phys. Lett. 177*B*, 184 (1986).
15. Toda, M., J. Phys. Soc. Jpn. 22, 431 (1967); "Theory of Nonlinear Lattices" (Springer-Verlag, 1981).
16. Saitoh, N., J. Phys. Soc. Jpn. 49, 409 (1980).
17. Hirota, R., J. Phys. Soc. Jpn. 50, 3787 (1981); "Nonlinear Integrable Systems" *ed.* M.Jimbo and T.Miwa (World Scientific, 1983) p.17.
18. Miwa, T., Proc. Jpn. Acad. 58*A*, 9 (1982);
Date, E., Jimbo, M. and Miwa, T., J. Phys. Soc. Jpn. 51, 4116, 4125 (1982).
19. Krichever, I.M., Russian Math. Surveys 32, 185 (1977).
20. Mulase, M., J. Diff. Geom. 19, 403 (1984).
21. Shiota, T., Invent. Math. 83, 333 (1986).
22. Fay, J.D., "Theta Functions on Riemann Surfaces", Lect. Notes in Math. 352, (Springer-Verlag, 1973).
23. Ishibashi, N., Matsuo, Y. and Ooguri, H., Mod. Phys. Lett. *A*2, 119 (1987), Alvarez-Gaumé, L., Gomez, C. and Reina, C., Phys. Lett. 190*B*, 55 (1987).
24. Kato, H. and Saito, S., Lett. Math. Phys. 18, 117 (1989).
25. Carow-Watamura, U., Ezawa, Z.F. and Watamura, S., Nucl. Phys. *B*315, 166 (1989); Carow-Watamura, U., Ezawa, Z.F., Tezuka, A. and Watamura, S., Phys. Lett. 221*B*, 299 (1989); Nucl. Phys. *B*319, 187 (1989).
26. Virasoro, M.A., Phys. Rev. *D*1, 2933 (1970).
27. Bernard, D. and LeClair, A., Phys. Lett. 227*B*, 417 (1989).
28. Câteau, H. and Saito, S., Phys. Rev. Lett. 65, 2487 (1990).
29. Kulish, P.P. and Reshetikhin, N.Yu., Zapiski Nauch. LOMI 101, 101 (1980) ; J. Sov. Math. 23, 2435 (1983) ;
Sklyanin, E.K., Funct. Anal. Appl.16, 27 (1982); 17, 34 (1983);

Jimbo, M., Lett. Math. Phys. 10, 63 (1985); 11, 247 (1986);
Int. Jour. Mod. Phys. *A*4, 3759 (1989) ;
Drin'feld, V.G., Sov. Math. Dokl. 32, 254 (1985);
"Quantum Groups", ICM Proceedings, (Berkley, 1986) P.898;
Woronowicz, S., Publ. RIMS-Kyoto 23, 117 (1987);
Roche, P. and Adnaudon, D., Lett. Math. Phys. 17, 295 (1989);
Witten, E., Nucl. Phys. *B*330, 285 (1990);
Fairlie, D.B., J. Phys. A: *Math. Gen.* 23, L183 (1990).
30. Chaichian, M., Kulish, P. and Lukierski, J., Phys. Lett. *B*237, 401 (1990).
31. Curtright,T. and Zachos, C., Phys. Lett. *B*237, 237 (1990).
32. Hiro-oka, H, Matsui, O., Naito, T. and Saito, S., Tokyo Metropolitan Univ. preprint, TMUP-HEL-9004 (1990), Saito, S., TMUP-HEL-9020 (1990).
33. Devchand, Ch. and Saveliev, M.V., Moscow preprint IHEP-90-144 (1990), Polychronakos, A.P., University of Florida preprint UFIFT-HEP-90-14.
34. Chaichian, M., Ellinas, D. and Popowicz, Z., Phys. Lett. *B*248, 95 (1990).
35. Aizawa, N. and Sato, H., Hiroshima Univ. preprint HUPD-9007 (1990).

A q-DEFORMED PROBABILITY,NELSON'S INEQUALITY AND CENTRAL LIMIT THEOREMS.

MAREK BOŻEJKO

Instytut Matematyczny Uniwersytetu Wrocławskiego

Wrocław,Plac Grunwaldzki 2/4,Poland.

ABSTRACT:We present some new ideas in the probability theory coming from the q-deformed oscillators and twisted Fock spaces,which were considered by D.Voiculescu in the case $q=0$ and by R.Speicher and the author in the case $-1 \leq q \leq 1$. In the case $q=1$ we have the classical probability; $q=0$ corresponds to "free" probability,which was founded by D.Voiculescu and $q=-1$ is the fermionic calculus.

O.INTRODUCTION. The paper is divided on four parts. We start with a measure system and free probability,which was introduced and developed by D.Voiculescu [V 1] .We present the "free" central limit theorem,for which the limit measure is Wigner semicircle low $2/\pi \sqrt{1-x^2}\, dx$. Nextly we shall prove the analogue of Nelson's hypercontractive inequality for free independent Gaussian variables G_i :

$$\| \sum_{|i|=k} b_i \, G_i \| \leq c\sqrt{k} \, (\sum_{|i|=k} b_i^2)^{1/2} .$$

In the second part we consider the probability system (A, φ , Ψ) with two states φ, Ψ on a C^*-algebra A and we present the central limit theorem.This was done by R.Speicher and the author [B-S 1]. In the section 3 we are looking on the geometry of the permutation group S_n .The important fact which we use in the construction of the twisted Fock space $F_q(H)$ is the theorem that the matrix $(q^{|\sigma\tau|})_{\sigma,\tau \in S_n}$ is strictly positive definite for $q \in (-1,1)$. This was proved by T.Januszkiewicz,R.Spatzier and the author [B-J-S] for all Weyl (Coxeter) groups, but here we use that theorem in the case of the finite permutation group S_n .

Applying that fact about the length function on the permutation group S_n we construct the twisted Fock space $F_q(H)$,(H is a fixed Hilbert space) ,which is the completion of the full Fock space $F(H) = \bigoplus_{n=0}^{\infty} H^{\otimes n}$,with respect the following new scalar product:
For a tensor $x_1 \otimes x_2 \otimes \ldots \otimes x_n \in H^{\otimes n}$

we put

$$\| x_1 \otimes x_2 \ldots \otimes x_n \|_q^2 = \sum_{\sigma \in S_n} q^{|\sigma|} \langle x_{\sigma(1)}, x_1 \rangle \ldots \langle x_{\sigma(n)}, x_n \rangle .$$

This was done by R.Speicher and the author [B-S 2].One can see that in the case $q=\pm 1$ we get the classical bosonic (fermionic) Fock space;the case $q=0$ was considered by D.Voiculescu [V 1].

Introducing the q-deformed ,independent creation and annihilation operators we state the central limit theorem for q-probability.The q-Gaussian measure μ_q has compact support for which the corresponding orthogonal polynomials are q-Hermite polynomials satisfying the following recurence relations:

$$x H_n^{(q)} = H_{n+1}^{(q)} + (1-q^n)(1-q)^{-1} H_{n-1}^{(q)} .$$

For more details see the paper [B-S 3] .

1.MEASURE SYSTEM; FREE PROBABILITY.

Following classical aproach we consider as unital C^*-algebra A,i.e. a closed,$*$-subalgebra of all bounded operators on Hilbert space,such that $1 \in A$.

A measure system is a pair (A, φ),where φ is a state on A i.e. a normalized,positive linear functional on A , $\varphi(1) = 1$. By a random variable we mean a hermitian element of A. A distribution of a random variable $a=a^* \in A$ in the measure system (A, φ) is the probability measure on the real line- μ_a such that for every continuous function with a compact support f ,we have

$$\int_{\mathbb{R}} f \, d\mu_a = \varphi(f(a)) .$$

By Riesz representation theorem such a measure μ_a is unique probability measure on $\mathbb{R}$ with support included in the spectrum of an element a .

Now we consider two important examples of measure systems.

Example 1. If $(\Omega, \mathcal{B}, P)$ is a standard probability system and let f be a bounded random variable on Ω . Let us take the algebra $A(\Omega) = \{ M_f : f \in L^\infty(\Omega) \}$, where $M_f : L^2(\Omega) \to L^2(\Omega)$ is the multiplication operator defined as follows: For $h \in L^2(\Omega)$, $M_f(h) = f \cdot h$. If we define the state ψ as

$\psi(M_f) = \int_\Omega f \, dP$, then in the measure system $(A(\Omega), \psi)$ the distribution of a random variable f is the same as in the classical case : $F(x) = P(f > x)$,or more precisely $dF(x) = d\mu_{M_f}(x)$.

Example 2. Let $A = L(l^2(N))$ be the algebra of all bounded operators on $l^2(N)$, and for $T \in A$ we define the state as follows : $\varepsilon(T) = \langle T \delta_o, \delta_o \rangle$,where

$\delta_o(k) = \begin{cases} 1 \text{ if } k=0 \\ 0 \text{ if } k \neq 0 \end{cases}$. Let S be the one-sided shift operator defined on $l^2(N)$: $S(x_0, x_1, \ldots,) = (0, x_0, x_1, \ldots)$, then the distribution of the operator $W = 1/2(S + S^*)$ in the measure system (A, ε) is the Wigner semicircle low

$$du_W(x) = 2/\pi \sqrt{1-x^2} \, dx .$$ For more examples see the papers of D.Voiculescu [V 1], [V 2], [V 3] .

By a probability system we shall mean a measure system

together with some form of "independence" .

In the commutative probability system (A, φ) ,an algebra A is commutative and elements $a,b \in A$ are independent if for all $n,m = 0,1,2,\ldots$ we have

$$\varphi(a^n b^m) = \varphi(a^n)\ \varphi(b^m) \ .$$

The idea here is following: the "supports" of independent random variable a and b live on separable domains of variables i.e. $a = a(x)$,$b = b(y)$.

Now we consider a different concept of independence: free independence .Before giving a formal definition we present a natural example:

Example 3. Let our algebra A is the algebra of all polynomials in non-commuting variables x,y; for a polynomial $p= p(x,y)$ we put

$$\varphi(p) = p(0,0) = \text{constant term in the polynomial } p.$$

If we take for example polynomials $a = p_1(x)$ and $b = p_2(y)$,where p_i are polynomials,then one can see that in our measure system they are free independent.In general elements $a,b \in (A, \varphi)$ are free independent if we have no relation between a and b modulo φ ; The definition we shall give later.

We follow now D.Voiculescu [V 1] and we present another very important class of free independent random variables.

Example 4. Let H be a Hilbert space with an orthonormal basis $e_1,e_2,\ldots,e_N$,$(N = \infty,1,2,\ldots)$ and $\tilde{\Omega}$ -a fixed normalized vector out of H . We form the full Fock space

$$F(H) = \mathbb{C}\tilde{\Omega} \oplus \bigoplus_{n=1}^{\infty} \underbrace{H \otimes H \otimes \ldots \otimes H}_{n} \ .$$

Let for a bounded,linear operator T on $F(H)$ we define the natural vacuumstate $\varepsilon_N(T) = \langle T\tilde{\Omega}, \tilde{\Omega} \rangle$.

Now for $f \in H$ we form the creation operator as follows:

$$a^+(f)(x_1 \otimes x_2 \ldots \otimes x_n) = f \otimes x_1 \otimes \ldots \otimes x_n \ ;$$

and $a(f)(x_1 \otimes x_2 \ldots \otimes x_n) = \langle f, x_1 \rangle x_2 \otimes \ldots \otimes x_n$ -

the annihilation operator .

If we consider the operators $l_i = a(e_i)$ and $l_i^+ = a^+(e_i)$, $(i=1,2,\ldots,N)$,then we get the family of random variables $\{l_i, l_i^+\}$, which is free independent in the measure system $\langle L(F(H)) , \varepsilon_N \rangle$.

Remark. One can compute (see also [V 1]) that the distribution of the random variable $G_i = 1/2(l_i + l_i^+)$ is again Wigner semicircle low. The variable G_i are called free Gaussian distributions.

We state now the central limit theorem for free independent random variable,which in particular case of free group was proved by the author [B] ,but in full generality by D.Voiculescu [V 1] ,[V 2]. .

THEOREM (central limit theorem).

If $b_i = b_i^* \in (A, \varphi)$ are free independent and

1^o $\quad \varphi(b_i) = 0$

2^o $\quad \varphi(b_i^2) = 1/2$

3^o $\quad \| b_i \| \leq C$, then the random variable $\tilde{S}_n = 1/\sqrt{n}\,(b_1 + b_2 + \ldots + b_n)$ tends weakly to Wigner

distribution $2/\pi \sqrt{1-x^2}\, dx$ or in another words:

for all natural k, $\lim_n \varphi(\tilde{S}_n^k) = \begin{cases} \frac{1}{m+1}\binom{2m}{m} & \text{if } k = 2m \\ 0 & \text{if } k \text{ is odd} \end{cases}$.

Now we prove an analogue of Nelson's hypercontractive inequality for the free Gaussian variables.

For this we need some notations:

Let $\underline{i} = (i_1, i_2, \ldots, i_k)$ be a multiindex, $i_j \in N$;

we put $|\underline{i}| = k$ and $l_{\underline{i}} = l_{i_1} l_{i_2} \ldots l_{i_k}$, $l_{\underline{i}}^+ = (l_{\underline{i}})^+$

and $l_{\underline{i},\underline{\varepsilon}} = l_{i_1}^{\varepsilon_1} l_{i_2}^{\varepsilon_2} \ldots l_{i_k}^{\varepsilon_k}$, where $\varepsilon_j \in \{1, +\}$

and l_i^+ (l_i) are creation (resp. annihilation) operators on the full Fock space defined in Example 4.

Before the Theorem we state a Lemma.

LEMMA 1. 1° For multiindices $\underline{i}$, $\underline{j}$ such that $|\underline{i}| = |\underline{j}|$, we have $l_{\underline{i}}\, l_{\underline{j}}^+ = \delta_{\underline{i},\underline{j}}$.

2°. $$\Big\| \sum_{\substack{|\underline{i}|=k \\ |\underline{j}|=l}} a_{\underline{i},\underline{j}}\, l_{\underline{i}}^+\, l_{\underline{j}} \Big\| \le \Big(\sum_{\substack{|\underline{i}|=k \\ |\underline{j}|=l}} |a_{\underline{i},\underline{j}}|^2 \Big)^{1/2}$$

Proof. 1°. Since $l_i\, l_j^+ = \delta_{i,j}$, therefore by a simple induction we obtain 1° .

2°. Let us first prove our inequality in the case $|\underline{i}| = |\underline{j}| = 1$. For this we define a map

$\alpha : M_N(\mathbb{C}) \longrightarrow L(F(H))$ as follows

$$\alpha(a_{i,j}) = \sum_{i,j=1}^{N} a_{i,j}\, l_i^+\, l_j$$

and this map is a $*$-isomorphism , i.e. $\alpha(a\,b) = \alpha(a)\alpha(b)$, $\alpha(a^*) = \alpha(a)^*$ and $\alpha(a) = 0$ if and only if $a = 0$,

This follows immediately from 1° .
Therefore by the standard theorem on C^*-algebras we get that α is an isometry i.e.

$$\Big\| \sum a_{i,j}\, l_i^+\, l_j \Big\| = \| a_{i,j} \| \le \Big(\sum |a_{i,j}|^2 \Big)^{1/2}.$$

The same proof can be repeated in the case of multiindices.

THEOREM(an analogue of Nelson's inequality).

Let l_i (l_i^+) be the annihilation (creation) operator on the full Fock space and $G_i = 1/2(l_i + l_i^+)$,

for $\underline{i} = (i_1, \ldots, i_k)$ let $G_{\underline{i}} = G_{i_1} \ldots G_{i_k}$, then

(I) $\| \sum_{|\underline{i}|=k} b_{\underline{i}} \, l_{\underline{i}} \| = \left(\sum_{|\underline{i}|=k} |b_{\underline{i}}|^2 \right)^{1/2}$

(II) $\| \sum_{\substack{|\underline{s}|=m \\ \underline{\varepsilon}}} c_{\underline{s},\underline{\varepsilon}} \, l_{\underline{s},\underline{\varepsilon}} \| \leq \sqrt{m+1} \left(\sum_{\substack{|\underline{s}|=m \\ \underline{\varepsilon}}} |c_{\underline{s},\underline{\varepsilon}}|^2 \right)^{1/2}$

(III) $\| \sum_{|\underline{i}|=k} d_{\underline{i}} \, G_{\underline{i}} \| \leq C \, k^{-1/2} \left(\sum_{|\underline{i}|=k} |d_{\underline{i}}|^2 \right)^{1/2}$.

Proof. (I). If $B = \sum_{|\underline{i}|=k} b_{\underline{i}} \, l_{\underline{i}}$, then $\| B \|^2 =$

$= \| B \, B^* \| = \| \sum_{|\underline{i}|=|\underline{j}|=k} b_{\underline{i}} \, \overline{b}_{\underline{j}} \, l_{\underline{i}} \, l_{\underline{j}}^+ \|$.

Since by Lemma 1 , $l_{\underline{i}} \, l_{\underline{j}}^+ = \delta_{\underline{i},\underline{j}}$ we at once get (I).

(II) . Let us note that every element $l_{\underline{s},\underline{\varepsilon}}$, $|\underline{s}| = m$ is of the following form: $l_{\underline{s},\underline{\varepsilon}} = l_{\underline{i}}^+ \, l_{\underline{j}}$,where

$|\underline{s}| = |\underline{i}| + |\underline{j}| = m$ and $\underline{\varepsilon} = (\underbrace{+,+,\ldots,+}_{|\underline{i}|},\underbrace{1,1,\ldots,1}_{|\underline{j}|})$.

Therefore

$$\sum_{\substack{|\underline{s}|=m, \\ \underline{\varepsilon}}} c_{\underline{s},\underline{\varepsilon}} \, l_{\underline{s},\underline{\varepsilon}} = \sum_{k=0}^{m} \sum_{\substack{|\underline{i}|=k \\ |\underline{j}|=m-k}} c_{\underline{i},\underline{j}}^{(k)} \, l_{\underline{i}}^+ \, l_{\underline{j}}$$

and by Lemma 1 and Schwarz inequality we have

$$\| \sum_{\substack{|\underline{s}|=m \\ \underline{\varepsilon}}} c_{\underline{s},\underline{\varepsilon}} \, l_{\underline{s},\underline{\varepsilon}} \| \leq \sum_{k=0}^{m} \left(\sum_{\substack{|\underline{i}|=k \\ |\underline{j}|=m-k}} |c_{\underline{i},\underline{j}}^{(k)}|^2 \right)^{1/2}$$

$$\leq \sqrt{m+1} \left(\sum_{\substack{|\underline{s}|=m \\ \underline{\varepsilon}}} |c_{\underline{s},\underline{\varepsilon}}|^2 \right)^{1/2} .$$

(III) This inequality follows directly from (I)and (II) .

Remark.1. If we take $c_{\underline{s},\underline{\varepsilon}} = 1$ for $|\underline{s}| = m$,then for $m \longrightarrow \infty$ we obtain $\| \sum_{\substack{|\underline{s}|=m \\ \underline{\varepsilon}}} l_{\underline{s},\underline{\varepsilon}} \| / \sqrt{m+1} \longrightarrow 1$.

This implies that the constant $\sqrt{m+1}$ is the best possible.

Remark 2.In the case of the free group the similar inequality was obtained by U.Haagerup [H],but the constant is $(m+1)$.

2.TWO STATES PROBABILITY SYSTEMS.

Now we are interested in the probability systems (A, φ, ψ) ,where A is some C^*-algebra with unit and φ , ψ are normalized states on A .

We introduce the following definition of the independence.

Definition. Subalgebras $(A_i)_{i \in I} \subset A$ are bi-independent if for each $b_j \in A_{i_j}$, $i_1 \neq i_2 \neq \ldots \neq i_k$ such that $\psi(b_j) = 0$ we have

$$\varphi(b_1 b_2 \ldots b_k) = \varphi(b_1)\,\varphi(b_2)\ldots\varphi(b_k)$$

and for each $c_j \in A$, $i_1 \neq i_2 \neq \ldots \neq i_k$

such that $\varphi(c_j) = 0$ we have

$$\Psi(c_1 c_2 \ldots c_k) = \Psi(c_1)\,\Psi(c_2)\ldots\Psi(c_k) \quad .$$

Elements $\{a_j\}_{i\in I}$ are <u>bi-independent</u> ,if the subalgebras A_j generated by $\{a_j\}$ are bi-independent in the measure system (A, φ, Ψ) .

If $\varphi = \Psi$,then the bi-independence coincides with <u>free independence</u>. For others examples see the paper of R.Speicher and the author [B- S 1].

To get a large family of bi-independent random variables we proved the analoque of Kolmogoroff's product theorem

<u>THEOREM</u>. Let $(A_i, \varphi_i, \Psi_i)_{i\in I}$ be a two states probability systems,then there exists a two system (A, φ, Ψ) such that

(i) Algebras A_i are embedded in A for each $i \in I$.

(ii) $\varphi|_{A_i} = \varphi_i$; $\Psi|_{A_i} = \Psi_i$,for $i \in I$.

(iii) Subalgebras $(A_i)_{i\in I}$ are bi-independent in the system (A, φ, Ψ) .

The proof of that theorem is rather complicated and we only make remark that the main problem is in the construction of a big Hilbert space and a special $*$-representations. For more details see the paper of R.Speicher and the author [B-S 1] .

In two states systems we also have central limit theorem.

<u>THEOREM</u> . Let $\{b_j\}$ are bi-independent ,hermitian elements in two states probability system (A, φ, Ψ) such that

(i) $\varphi(b_j) = \Psi(b_j) = 0$,for $j \in I$.

(ii) $\varphi(b_j^2) = \alpha^2$, $\Psi(b_j^2) = \beta^2$,then

the distribution of the operator $S_n = 1/\sqrt{n}\,(b_1 + \ldots + b_n)$

tends weakly in the system (A, φ) to the measure $\mu_{\alpha,\beta}$, and symetrically in the system (A, Ψ) to the measure $\mu_{\beta,\alpha}$. The measure is of the following form:

$$\mu_{\alpha,\beta}(t) = c_{\alpha,\beta}\left[\delta\left(\frac{\alpha^2}{\sqrt{\alpha^2-\beta^2}}\right) + \delta\left(-\frac{\alpha^2}{\sqrt{\alpha^2-\beta^2}}\right)\right] +$$

$$+ \chi_{[-2\beta,2\beta]}\, 1/2\pi \frac{\alpha^2\sqrt{4\beta^2 - t^2}}{\alpha^4 - (\alpha^2-\beta^2)t^2}\, dt \quad ,\text{where}$$

$$c_{\alpha,\beta} = 1/4\;\alpha^2/(\alpha^2-\beta^2)\,\max(1-2\beta^2/\alpha^2\,,\,0) \quad .$$

<u>Remarks</u>. 1°. For a special values of a parameters α, β the limit measure has a discrete part with two atoms.

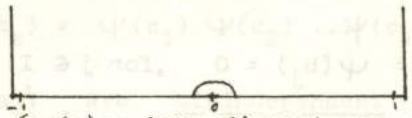
2°. The proof is combinatorial and was proved by M.Leinert and the author [B-L], see also general aproach [B-S 1]. In the case $\alpha = \beta$ we get again Wigner low .

3°. Assume that $\alpha = 1$ and denote $u_{1,\beta} = u_\beta$,

then we have the following measures:

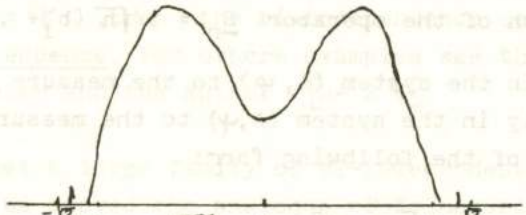
1°. $\beta = 0$, $u = 1/2(\delta_{-1} + \delta_{1})$ - this corresponds to fermionic case.

2°. For β near 0, the picture of μ_{β} is following:



At points $(\pm 1\pm\varepsilon)$ we have discrete masses.

3°. For β^2 near $1/2$ we have the following measure:

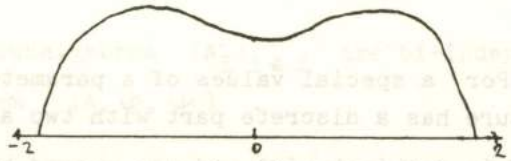


Yet at the points $\pm\sqrt{2\pm\varepsilon}$ we have a small discrete measure.

4°. For $\beta^2 = 1/2$, the measure μ_{β} is without discrete part, supportet on the interval $[-\sqrt{2},\sqrt{2}]$ and is of the form

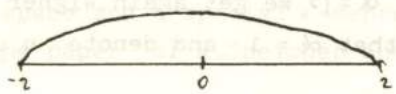
$d\mu_{\beta}(t) = 1/\pi \ (2-t^2)^{-1/2} dt$. In that case the orthogonal polynomials with respect to the measure μ_{β} are Tchebychef polynomials of the first kind i.e. $T_n(x) = \cos(n\theta)$, $x = \sqrt{2}\cos\theta$.

5°. If $1/2 < \beta^2 < 1$, the graph of the measure is as follows,



6°. For $\beta^2 = 1$ we get again Wigner semicircle low

$1/2\pi \sqrt{4- t^2}$ dt



The corresponding orthogonal polynomials in that case are Tchebychef polynomials of the second kind i.e.

$U_n(x) = \sin(n+1)\theta / \sin\theta$, $x = 2\cos\theta$.

7°. If $\beta \longrightarrow \infty$, then μ_{β} tends weakly to the Dirac measure at 0 .

3. GEOMETRY OF PERMUTATION GROUP S_n; TWISTED FOCK SPACES.

Let S_n be the permutation group on n letters $1,2,.,n$. The group S_n is generated by the following transpositions:

$\pi_1 = (1,2)$, $\pi_2 = (2,3), \ldots, \pi_{n-1} = (n-1,n)$.

One can verify that they satisfy the following relations:

1°. $\pi_i^2 = 1$ for $i = 1,2,\ldots,n-1$.

2°. $\pi_i \pi_j = \pi_j \pi_i$ if $|i-j| \geq 2$

3°. $\pi_i \pi_j \pi_i = \pi_j \pi_i \pi_j$ if $|i-j| = 1$.

It is not difficult to show that every permutation $\sigma \in S_n$ is of the following form

$$\sigma = \pi_{i_1} \pi_{i_2} \ldots \pi_{i_k} .$$

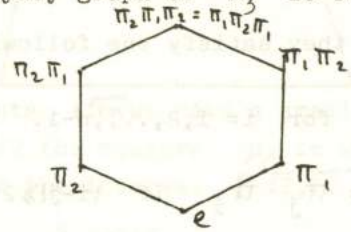
That representation is not unique but the number k is the same for other representations of σ; k is equal to the number of inversions of the permutation σ . We shall denote $k = |\sigma| = \#\{(i,j): i < j \leq n, \sigma(i) > \sigma(j)\}$. For example if

$\sigma = \begin{pmatrix}1234\\4312\end{pmatrix}$ then $|\sigma| = 5$.

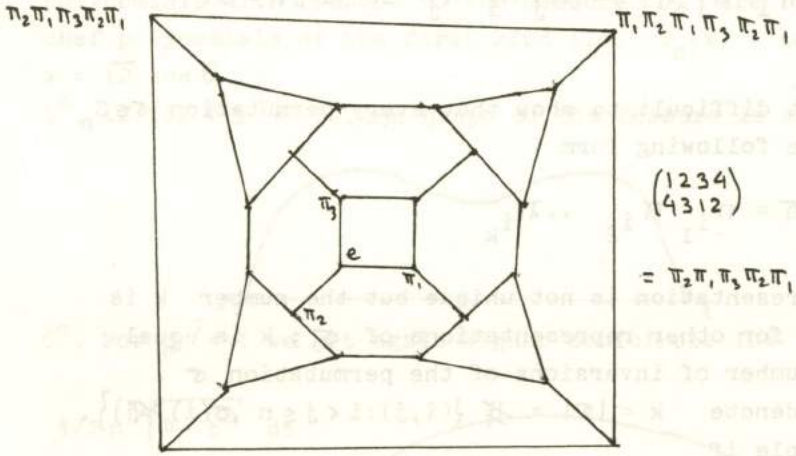
If we form so called Cayley graph of the group S_n, i.e. the graph,which vertices are elements of S_n and the edges are of the following form: $(\sigma, \sigma\pi_i)$, $\sigma \in S_n$ and $i=1,2,\ldots,n-1$, then we have the following well known fact.

PROPOSITION. For $\sigma, \tau \in S_n$, $|\sigma^{-1}\tau|$ is equal of the distance of the points σ, τ on Cayley graph of the permutation group.

For example Cayley graph od S_3 is following:



The group S_4 has the following Cayley graph:



$\begin{pmatrix}1234\\4312\end{pmatrix}$

$= \pi_2\pi_1\pi_3\pi_2\pi_1$.

For our futher considerations we need the following theorem.

THEOREM. (i) There exists an embedding $\gamma : S_n \longrightarrow \mathbb{R}^m$, $m = \binom{n}{2}$,such that for $x,y \in S_n$ we have

$$\|\gamma(x) - \gamma(y)\|^2 = |y^{-1}x| \quad .$$

(ii) For $-1 < q < 1$,the operator (matrix)

$$P_q^{(n)}(y^{-1}x) = q^{|y^{-1}x|}$$ is strictly positive

definite on $l^2(S_n)$.

(iii) $$\det\left(P_q^{(n)}\right) = \prod_{k=1}^{n-1} \left(1 - q^{k(k+1)}\right)^{a_n(k)} \quad .$$

The part (i) of Theorem was proved by T.Januszkiewicz, R.Spatzier and the author [B-J-S] ; parts (ii) and (iii) was founded by R.Speicher and the author [B-S2] .
It was communicated in the paper of O.Greenberg [G] ,that D.I.Favel proved that $a_n(k) = n!(n-k)/k(k+1)$.
Let us note that the part (iii) holds in all finite, irreducible Weyl groups.

THEOREM. If (W,S) is a finite,irreducible Weyl group, and for $J \subset S$ let $w_o(J)$ be the longest element in the parabolic subgroup W_J , then if $P_q^{(W)}(x) = q^{|x|}$,then

$$\det\left(P_q^{(W)}\right) = \prod_{\substack{J \subset S \\ J\text{-connected}}} \left(1 - q^{2|w_o(J)|}\right)^{a(J)}$$

Now we construct the twisted Fock space $F_q(H)$. Let H be a Hilbert space with an orthonormal basis $e_1, e_2, \ldots, e_N$, $N=\infty,1,2,\ldots$; Ω -a vacuum vector, $\|\Omega\| = 1$. Let $q \in [-1,1]$ and we define the following operator (q-symmetrizator) on a tensor power of H : $H^{\otimes n}$.

$$P_q^{(n)}(x_1 \otimes \ldots \otimes x_n) = \sum_{\sigma \in S_n} q^{|\sigma|} x_{\sigma(1)} \otimes \ldots \otimes x_{\sigma(n)} .$$

By the previous Theorem we have that for $\xi \in H^{\otimes n}$

$$\langle P_q^{(n)} \xi , \xi \rangle \geq C \|\xi\|^2 ,$$ (the scalar product is taken from the full Fock space).

Therefore the operator $P_q = \bigoplus_{n=0}^{\infty} P_q^{(n)}$ is positive on the full Fock space; It is without kernel for $-1 < q < 1$.

Now we introduce the new scalar product on the full Fock space $F(H)$ putting for $x,y \in F(H)$

$$\langle x , y \rangle_q = \langle P_q x, y \rangle .$$

Let $F_q(H)$ be the completion of $F(H)$ with respect that new scalar product. For $f \in H$ we define q-annihilation and q-creation operators in the following way:

$$a_q(f)(x_1 \otimes \ldots \otimes x_n) = \sum_{k=1}^{n} q^{k-1} \langle f, x_k \rangle x_1 \otimes \ldots \otimes \check{x}_k \otimes \ldots \otimes x_n .$$

$$a_q^+(f)(x_1 \otimes \ldots \otimes x_n) = f \otimes x_1 \otimes \ldots \otimes x_n .$$

Let us denote $\partial_i = a_q(e_i)$ and $x_i = a_q^+(e_i)$, where $\{e_i\}$ is a fixed orthonormal basis in Hilbert space H.

The following theorem was proved by R.Speicher and the author [B-S1].

THEOREM. (i) In the twisted Fock space $F_q(H)$ for $f \in H$, we have $a_q^*(f) = a_q^+(f)$;

In particular, we have $\partial_i^* = x_i$.

(ii) For $0 \leq q < 1$, the operators $a_q(f)$ and $a_q^+(f)$ are bounded and $\|a_q(f)\| = 1/\sqrt{1-q}\ \|f\| = \|a_q^+(f)\|$.

(iii) For $-1 \leq q \leq 0$, $\|a_q(f)\| = \|f\| = \|a_q^+(f)\|$.

(iv) $$\partial_i x_j - q\, x_j \partial_i = \delta_{ij} .$$

(v) For $f,g \in H$ we have

$$a_q(f)\, a_q^+(g) - q\, a_q^+(g)\, a_q(f) = \langle f,g \rangle .$$

Remark 1. The q-commutator relation as in (v) was considered by O.W.Greenberg, [G], without a construction.

Remark 2. If a Hilbert space $H = \mathbb{C}$, is one dimensional with a fixed unit vector $e_1 = x$, then

$$F_q(H) = \left\{ \sum_{n=0}^{\infty} \alpha_n x^n : \sum |\alpha_n|^2 [n]_q! < \infty \right\}, \quad [n]_q = (1-q^n)(1-q)^{-1}$$

and

$$[n]_q! = [1]_q [2]_q \cdots [n]_q .$$

One also verify directly from the definition that

$$\partial_1(1) = 0 \ , \ \partial_1(x^n) = [n]_q x^{n-1} \ ,\text{and for } f \in F_q(\mathbb{C})$$

we have

$$\partial_1(f)(x) = \ (f(x) - f(qx))/ (x - qx) \ ;$$

This is the classical Euler derivation operator. The scalar product on $F_q(\mathbb{C})$ is such that

$$\langle x^n, x^n \rangle_q = \sum_{\sigma \in S_n} q^{|\sigma|} = [1]_q [2]_q \cdots [n]_q = [n]_q! \ .$$

Let us note that the creation operator $a_q(x^n) = x^{n+1}$.

If we consider the measure system

$\langle L(F_q(C)) \ , \ \varepsilon \rangle$,where $\varepsilon(T) = \langle T\,1, 1 \rangle$,(here $\Omega = 1$),

then the distribution of the operator $T = \partial_1 + x_1$ is the probability measure μ_q for which the orthogonal polynomials $H_n^{(q)}$ satisfy the following relation:

$$x\, H_n^{(q)}(x) = H_{n+1}^{(q)}(x) + n\ H_{n-1}^{(q)}(x) \ .$$

In the case q=1 , we obtain Hermite polynomials and the measure $d\mu_1(x) = \exp(-x^2)dx$ is the well know Gauss distribution.

If $q = 0$,then $[n]_q = 1$ and we get that $H_n^{(o)} = U_n$, where U_n are Tchebyschef polynomials of second kind and our measure

$$d\mu_o(x) = 1/2\pi \sqrt{4-x^2}\ dx \ ,$$

so again Wigner semicircle low .

4. A q-DEFORMED PROBABILITY.

Let us consider the operators $G_q^{(i)} = x_i + \partial_i$ in the twisted Fock space $F_q(H)$ as models for q-independent q-Gaussian random variables.

As we saw , if q= 1,then the distribution of G_1 in Fock space is Gauss measure - $\exp(-x^2)\,dx$ on the real line. In the case $-1 \leq q < 1$, the distribution μ_q of the operator G_q was computed by S.Szegö and it is of the following form:(See R.Askey , M.Ismalil [A-I]).

$$d\mu_q(x) = (1-q)/2\pi \prod_{n=1}^{\infty} (1-q^n) \,|\, 1 - q^n e^{2i\theta} \,|\, \sqrt{4/(1-q) - x^2}\ dx,$$

where $x = 2/\sqrt{1-q}\ \cos\theta$, $0 < \theta \leq \pi$; the support of the measure μ_q is the interval $\left[-2/\sqrt{1-q} \ , \ 2/\sqrt{1-q} \right]$.

Now we can state the central limit theorem for q-probability and we shall see that the measure μ_q is the analogue of Gauss distribution. We introduce the notion of q-independence for random variables in the measure system (A, φ) . We shall say that the operators $T_i \in (A, \varphi)$ are q-independent,if there exist operators $\tilde{T}_i$ of the form

$$\tilde{T}_i = x_i + \sum_{k=0}^{\infty} d_{k,i}\, \partial_i^k \quad , \ i=1,2,\ldots, \ \text{inthe}$$

twisted q-Fock space $F_q(H)$ such that the distribution of T_i is equal to the distribution of the operator $\tilde{T}_i$, for all i=1,2,... , . This means that for all m=0,1,2,...

$$\varphi(T_i^m) = \varepsilon(\tilde{T}_i^{\,m}) \ .$$

THEOREM.(central limit for q-probability).

Let (A, φ) be a measure system and $T_i^* = T_i \in A$ be q-independent random variable such that:

(i) $\varphi(T_i) = 0$, $\varphi(T_i^2) = 1$, $\| T_i \| \leq C$,

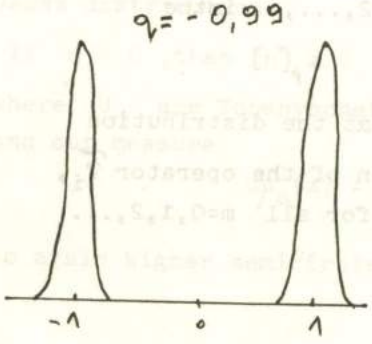
then the distribution of $\tilde{S}_n = 1/\sqrt{n}\,(T_1 + T_2 + \ldots + T_n)$ tends weakly to q-Gaussian measure μ_q.

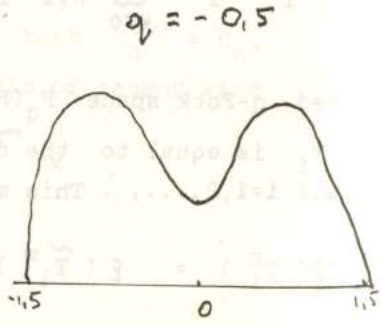
Remark 1. We can also consider the analogues of Poisson measure and Wiener processes. In the case q=0 this was done by R.Speicher [S] .See also the paper [B-S 1].

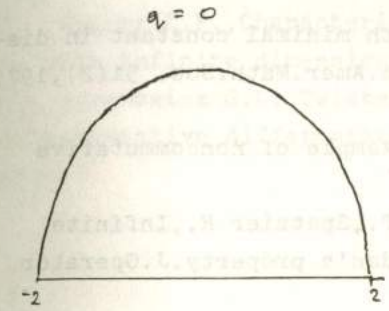
Remark 2. There are natural relations with Woronowicz quantum group $SU_q(2)$; see L.C.Biedenharn [BIE] and [B-S 2] .

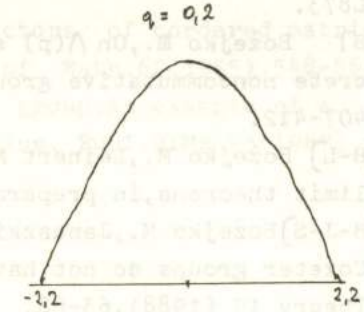
Remark 3. D.Voiculescu (see [V 1,2,3]) founded nice application of the free probability to a construction of special II_1- factors and random matrices .
We hope that q-independence gives the similar applications .

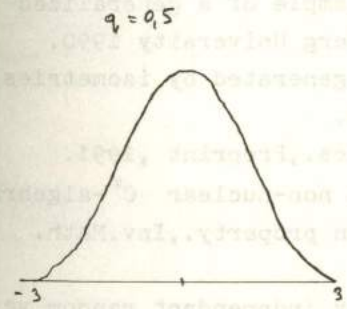
Now we give the pictures of the measure μ_q for different values of q in order to see the interpolation between the fermionic,the free and the bosonic case.

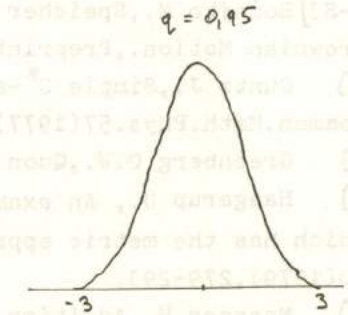












References.

[A-I] Askey R.,Ismail M., Recurrence relations,continued fractions and orthogonal polynomials.Memoires of the AMS, vol.49, (300),1984.

[BIE] Biedenharn L.C., The quantum group $SU_q(2)$ and a q-analoque of the boson operators.Journ.of Physic A 22(1989), L873.

[B] Bożejko M.,On $\Lambda(p)$ sets with minimal constant in discrete noncommutative groups,Proc.Amer.Math.Soc. 51(2),1975, 407-412.

[B-L] Bożejko M.,Leinert M.,An example of noncommutative limit theorems,in preparation.

[B-J-S] Bożejko M.,Januszkiewicz T.,Spatzier R.,Infinite Coxeter groups do not have Kazhdan's property.J.Operator Theory 19 (1988),63-68.

[B-S1] Bożejko M.,Speicher R., Ψ -Independent and symmetrized White Noises.,To appear in Quantum Probability V.

[B-S2] Bożejko M.,Speicher R.,An example of a generalized Brownian Motion, To appear in Commun.Math.Phys.

[B-S3] Bożejko M.,Speicher R.,An example of a generalized Brownian Motion.,Preprint,Heidelberg University 1990.

[C] Cuntz J.,Simple C^*-algebras generated by isometries. Commun.Math.Phys.57(1977),173-185.

[G] Greenberg O.W.,Quon Statistics.,Preprint ,1991.

[H] Haagerup U., An example of a non-nuclear C^*-algebra which has the metric approximation property.,Inv.Math. 50(1979),279-293.

[M] Maassen H.,Addition of freely independent random variables. Preprint,Nijmegen 1990.

[MAC] Macfarlane A.J.,On q-analoques of the quantum harmonic oscillator and the quantum group $SU(2)_q$. Journ.of Physics A22(1989),4581 .

[S] Speicher R.,A new example of "independence" and 'White Noise' .Probab.Th.Rel.Fields 84(1990),141-159.

[V 1] Voiculescu D.,Symmetries of some reduced free product C^*-algebras. Lectures Notes inMathematics 1132,Springer Verlag,Heidelberg 1985.

[V 2] Voiculescu D., Addition of certain non-commuting variables. J.Funct.Anal.66(1986),323-346.

[V 3] Voiculescu D.,Limit lows for random matrices and free products.Preprint IHES/M/89/21.

[W] Wigner E.P.,Characteristic vectors of bordered matrices with infinite dimensions.,Ann.of Math.62(1955),548-564.

[WO] Woronowicz S.L.,Twisted SU(2) group.An example of a non-commutative differential calculus. Publ.RIMS 23(1987), 117-181.

SUPERCOVARIANT q-OSCILLATORS*

M. CHAICHIAN†
Centre de Physique Theorique (CPT)
CNRS - Luminy, 70, route Leon Lachamp
13288 Marseille Cedex 9, France

P. KULISH
Steklov Mathematical Institute of USSR Acad. Sciences
Leningrad Branch, 191011 Leningrad, USSR

J. LUKIERSKI
Institute for Theoretical Physics, University of Wroclaw
ul. Cybulskiego 36, P-50205 Wroclaw, Poland

1 Introduction

Quantum deformations (the so-called q-deformations) of Lie groups and Lie algebras (cf. [1-3]) have found many applications in diverse domains of both theoretical physics and mathematics (see e.g.[4-6]). These mathematical objects which appeared as a byproduct of quantum inverse scattering method [7-8] have found applications in integrable models [9,6], in the formulation of the examples of noncommutative geometry [10-13] and in the description of rational conformal field theories [14]. It should be added that the q-deformations have been subsequently extended to supergroups and superalgebras [12,15-17].

Independently different deformations of the Heisenberg algebra (bosonic creation and annihilation operators) had already been studied in seventies (see e.g. [18]). Recently the system of independent q-oscillators (bosonic as well fermionic) was used for the q-oscillator realizations of different quantum Lie algebras [19-21] and quantum superalgebras [16,12]. It is known that the set of n creation or annihilation operators forms a basis for $U(n)$ fundamental representation. Using the

*Lecture given by J.L.

†On leave of absence from the Department of High Energy Physics, University of Helsinki, Siltavuorenpenger 20C, SF-00170 Helsinki, Finland

equivalence between the first order differential calculus and second quantization procedure the algebra of n q-oscillators covariant under the coaction of $SU_q(n)$ was given firstly in [23,24]. Interestingly enough, it appeared that $SU_q(n)$-covariance requires a particular coupling of different modes, which determines also the noncommutativity factors in $SU_q(n)$-covariant differential calculus [10-13].

The aim of this note is to construct the algebra of bosonic and fermionic q-oscillators covariant under the quantum supergroup $SU_q(n|m)$ (q-superoscillators). The formulae obtained here provide the construction of the differential geometry on non-commutative superspaces. After recalling the covariant formulation for $n = 2, m = 0$ (bosonic, nonsupersymmetric case) we shall discuss in detail the case $n = m = 1$. We shall show that the scheme can be made covariant under the discrete group of transformations, representing the permutations of the q-superoscillator labels.

Further, the $SU_q(n|m)$-invariant form defining the supersymmetric Hamiltonian for the q-superoscillators will be given. It appears that such a Hamiltonian also commutes with the quantum algebra $\mathcal{U}_q(U(n|m))$ with generators described by mixed bilinear products (creation, annihilation) of q-superoscillators. We shall see that the quantum algebra $\mathcal{U}_q(SU(1|1))(n = m = 1)$ can be treated as q-deformation of the $N = 2$ supersymmetry in quantum mechanics. Subsequently, the realizations of the supercovariant system of q-oscillators in terms of free q-oscillators (commuting different modes) as well as in terms of usual ones will be given.

It was pointed out recently by one of the authors that the q-deformed oscillators can be described by the R-matrix formalism (see e.g. [3]), following the description of scattering amplitudes of integrable models by Zamolodchikov-Fateev algebra (see e.g. [25]).

2 $SU_q(1/1)$-covariant pair of q-oscillators

The quantum supergroup $GL_q(1|1)$ is a graded associative Hopf algebra with two even (a,d) and two odd (β, γ) generators satisfying the following algebraic relations (see e.g. [13])

$$\Gamma^{1,1} \in GL_q(1|1): \quad \begin{array}{ll} \alpha\beta = q\beta a & d\beta = q\beta d \\ a\gamma = q\gamma a & d\gamma = q\gamma d \\ \beta\gamma = -\gamma\beta & ad - da = -r\beta\gamma \\ \beta^2 = 0 & \gamma^2 = 0 \quad (g \neq -1) \end{array} \tag{2.1}$$

Introducing q-deformed superdeterminant one defines

$$SL_q(1|1): \qquad s\,det_q T^{1,1} \equiv ad^{-1} - \beta d^{-1}\gamma d^{-1} = 1 \tag{2.2}$$

Further the *-involution restricts consistently the algebra (2.1) by imposing the relations

$$U_q(1|1): \quad T^{1,1} \in GL_q(1|1): \quad c = -\frac{1}{aa^*}\, b^* \,\frac{1}{a^*}, \quad d = (a^*)^{-1} \tag{2.3}$$

and

$$SU_q(1|1): \quad T^{1,1} \in U_q(1|1): \quad sdet_q T^{1,1} = a^*a + \beta\beta^* = 1 \tag{2.4}$$

We introduce the supersymmetric pair of q-oscillators (A, B) by postulating the following relations

$$\begin{aligned} AB &= qBA \\ A^+B^+ &= \frac{1}{q}B^+A^+ \end{aligned} \qquad B^2 = B^{+2} = 0 \tag{2.5a}$$

$$AA^+ - q^2A^+A = 1 \tag{2.5b}$$

$$BB^+ + B^+B = 1 + \left(q^2 - 1\right) A^+A \tag{2.5c}$$

$$AB^+ = qB^+A \qquad BA^+ = qA^+B \tag{2.5c}$$

The coaction of $GL_q(1|1)$

$$\begin{aligned} A' &= aA + \beta B & B' &= \gamma A + dB \\ A^{+\prime} &= a^*A^+ - \beta^*B^+ & B^{+\prime} &= \gamma^*A^+ + d^*B^+ \end{aligned}$$

with the generators of quantum group $GL_q(1|1)$ graded-commuting with the q-oscillators (A, B, A^+, B^+) leave invariant (2.5a), the relations (2.5d) (if $T^{1,1} \in U_q(1|1)$), and the relations (2.5b-c) (if $T^{1,1} \in SU_q(1|1)$). The relations (2.5a-d) distinguish the deformation of bosonic and fermionic sector (no operator tail on the rhs of (2.5b)). One can show however that the following "permuted" set of relations

$$AB = qBA \qquad B^2 = 0 \tag{2.6a}$$

$$AA^+ - \frac{1}{q^2}A^+A = 1 + \left(\frac{1}{q^2} - 1\right) B^+B \tag{2.6b}$$

$$BB^+ + B^+B = 1 \tag{2.6c}$$

$$AB^+ = \frac{1}{q}B^+A \tag{2.6d}$$

is covariant with respect to the same choices of quantum supergroups ((2.6a) for $GL(1|1)$, (2.6d) for $U(1|1)$, (2.6b-c) for $SU_q(1|1)$). The $U_q(1|1)$-invariant Hamiltonian is described by the following bilinear form

$$H^{1,1} = A^+A + B^+B \tag{2.7}$$

Introducing $d = dz\nabla_z + d\theta\nabla_\theta$ as an exterior differential on $R_q^{1|1}$ quantum superplane (z, θ), and δ as dual operator intertvinning k-th order and (k-1)-th order differential forms on $R_q^{1|1}$, after symmetrization one can interprete the operator (2.7) as the Laplace-Beltrami operator on $R_q^{1|1}$.

Now we shall calculate the quantum superalgebra $U_q(U(1|1))$ with the generators

$$\begin{aligned} Q &= A^+B & Q^+ &= B^+A \\ X &= A^+A - B^+B & Z &\equiv H^{1,1} \end{aligned} \tag{2.8}$$

One obtains :

$$Q^2 = Q^{+2} = 0 \tag{2.9}$$

$$\begin{aligned} q^2QQ^+ \;& +q^{-2}Q^+Q = q^{-2}Z + \\ & + \tfrac{1}{2}\left(1 - \tfrac{1}{q^2}\right)(X + Z) + \tfrac{1}{4}(q^2 - 1)(Z + X)^2 \end{aligned} \tag{2.10a}$$

$$Q^+Y - q^2ZQ^+ = -2Q^+ + (1 - q^2)ZQ^+ \tag{2.10b}$$

$$QY - q^2YQ = 2Q - \left(1 - q^2\right) ZQ \tag{2.10c}$$

$$[Q, Z] = [Q^+, Z] = [X, Z] = 0 \tag{2.10d}$$

Due to the relation (2.10d) we see that the Hamiltonian (2.7) describes the central extension of quantum algebra with basic generators Q, Q^+, Y, obtained by putting $Z = O$ in the relations (2.10a-c). The formulae (2.10) describe the q-extension of the N=2 supersymmetry in quantum mechanics ($\{Q, Q^+\} = H^{1,1}$, $Q^2 = Q^{+2} = 0$), with the supersymmetric doublets ($|E>, Q|E>$) having the same energy values (see (2.10d)).

It should be mentioned that for quantum $GL(1|1)$ one can introduce also a two-parameter R-matrix

$$R = \begin{pmatrix} q & 0 & 0 & 0 \\ 0 & s^{-1} & 0 & 0 \\ 0 & r & s & 0 \\ 0 & 0 & 0 & q^{-1} \end{pmatrix} \qquad \begin{cases} r = q - q^{-1}, \\ s - complex \end{cases} \tag{2.11}$$

defining the quantum group $GL_{q,s}(1|1)$ by means of the relation $R(T \otimes T) = (I \otimes T)(T \otimes I)R$ as also follows :

$$GL_{q,s}(1|1): \quad \begin{array}{ll} a\beta = qs^{-1}\beta a & d\beta = qs^{-1}\beta d \\ a\gamma = qs\gamma a & d\gamma = qs\gamma d \\ \beta\gamma = -s^2\gamma\beta & ad - da = rs\gamma\beta \\ \beta^2 = 0 & \gamma^2 = 0 \end{array} \tag{2.12}$$

It can be shown that the quantum superdeterminant of $GL_{q,s}(1|1)$ is central, i.e. one can define consistently $SL_{q,s}(1|1)$.

It would be interesting to generalize the q-superoscillator algebra (2.5) or (2.6) which is covariant under the coaction of the quantum supergroup (2.12) restricted by suitable unitarity conditions.

3 General case

For arbitrary values of n, m we propose the following set of relations for q-deformed superoscillators variables (q real)

$$\begin{array}{ll} A_iA_j = qA_jA_i, & i < j \\ A_i^+A_j^+ = \frac{1}{q}A_j^+A_i^+, & i < j \\ A_iA_j^+ = qA_j^+A_i, & i \neq j, \end{array} \tag{3.1}$$

$$A_iA_i^+ - q^2A_i^+A_i = 1 + \left(q^2 - 1\right)\sum_{k<i} A_k^+A_k \tag{3.2}$$

$$\begin{array}{l} A_iB_r = qB_rA_i, \\ A_i^+B_r^+ = \frac{1}{q}B_r^+A_i^+, \\ A_iB_r^+ = qB_r^+A_i, \end{array} \qquad \begin{cases} 1 \leq l \leq n \\ 1 \leq r \leq m \end{cases} \tag{3.3}$$

$$\begin{array}{ll} B_rB_s = -qB_sB_r, & r < s \\ B_r^+B_s^+ = -\frac{1}{q}B_s^+B_r^+ & r < s \\ B_rB_s^+ = qB_s^+B_r^+, & r \neq s, \\ B_rB_r^+ + B_r^+B_r = 1 + (q^2-1)\sum_{k=1}^{n} A_k^+A_k & \\ + (q^2-1)\sum_{s<r} B_r^+B_r & \end{array} \tag{3.4}$$

The supercovariant Hamiltonian is given by the formula

$$H^{n,m} = \sum_{i=1}^{n} A_i^+A_i + \sum_{r=1}^{m} B_r^+B_r \tag{3.5}$$

which is commuting with the $U_q(U(n|m))$ quantum algebra with the q-deformed Cartan-Weyl basis obtained by the bilinear Hermitian products of the superoscillators (3.1-4).

In order to obtain the formulae (3.1-4) let us introduce n bosonic and m fermionic independent q-oscillators [19-21]

$$\begin{array}{c} a_ia_j^+ - q^{\delta_{ij}}a_j^+a_i = q^{-N_i}\delta_{ij} \\ [a_i, a_j] = [a_i^+, a_j^+] = 0 \end{array} \tag{3.6}$$

$$\begin{array}{c} b_rb_s^+ + q^{-\delta_{ij}}\, b_s^+b_r = q^{-M_r}\delta_{rs} \\ \{b_r, b_s\} = \{b_r^+, b_s^+\} = 0 \end{array} \tag{3.7}$$

and (no summation)

$$a_i^+a_i = [N_i]_q \qquad b_r^+b_r = [M_r]_{q^{-1}} \tag{3.8}$$

One can also introduce another type of q-independent oscillators

$$\begin{array}{ll} \alpha_i = q^{\frac{N_i}{2}}\, a_i & \beta_r = q^{\frac{M_r}{2}}\, b_r \\ \alpha_i^+ = a_i^+ q^{\frac{N_i}{2}} & \beta_r^+ = b_r^+ q^{\frac{M_r}{2}} \end{array} \tag{3.9}$$

satisfying the algebra

$$\begin{array}{c} \alpha_i\alpha_j^+ - q^{2\delta_{ij}}\alpha_j^+\alpha_i = \delta_{ij}, \\ [\alpha_i, \alpha_j] = [\alpha_i^+, \alpha_j^+] = 0, \end{array} \tag{3.10a}$$

$$\begin{array}{c} \beta_r\beta_s^+ + \beta_s^+\beta_r = \delta_{rs}, \\ \{\beta_s, \beta_r\} = \{\beta_s^+, \beta_r^+\} = 0. \end{array} \tag{3.10b}$$

The relations (3.8) are replaced by the following ones

$$\alpha_i^+\alpha_i = [N_i]_{q^2}^c, \qquad \beta_r^+\beta_r = [M_i]_{q^2}^c \tag{3.11}$$

where $[a]_q^c = \dfrac{q^a - 1}{q - 1}$, or

$$q^{2N_i} = 1 + \left(q^2 - 1\right) \alpha_i^+ \alpha_i, \tag{3.12a}$$

$$q^{2M_i} = 1 + \left(q^2 - 1\right) \beta_r^+ \beta_r. \tag{3.12b}$$

The formulae (3.12a-b) can be generalized as follows :

$$q^{2\sum\limits_{i<k} N_k} = 1 + \left(q^2 - 1\right) \sum_{i<k} A_k^+ A_k, \tag{3.13a}$$

$$q^{2\sum\limits_{r<s} M_s} = 1 + \left(q^2 - 1\right) \sum_{r<s} \tilde{B}_s^+ \tilde{B}_{s'} \tag{3.13b}$$

where A_k satisfy the relations (3.1) and B_r the relation (3.4) with $n = 0$. In order to obtain the set of all formulae (3.1-4) we should introduce

$$A_i = q^{\sum\limits_{i<k} N_k} \alpha_i = q^{\left(\sum\limits_{i<k} N_k + \frac{N_i}{2}\right)} a_i, \tag{3.14a}$$

$$B_r = q^{\left(\sum\limits_{i=1}^{n} N_k + \sum\limits_{r<s} M_s\right)} \beta_i = q^{\left(\sum\limits_{i=1}^{n} N_k + \sum\limits_{r<s} M_s + \frac{M_r}{2}\right)} b_r. \tag{3.14b}$$

In order to obtain the representations of the algebra (3.1-4) the relations (3.14a-b) are very useful. One can construct the representations of the relations (3.6-7) or (3.9-10) in the Fock spanned by the normalized eigenstate $|n_{\alpha_i}, m_{\beta_r}>$ of the number operators $N_i, M_r (i = 1..n; r = 1..m)$. In particular the spectrum of the Hamiltonian (3.5) is given by the following formula.

$$\begin{aligned} &H^{(n|m)} |n_{\alpha_1} \dots n_{\alpha_n}; m_{\beta_1} \dots m_{\beta_m}> \\ &= [N + M]_q^c |n_{\alpha_1} \dots n_{\alpha_n}; m_{\beta_1} \dots m_{\beta_m}>, \end{aligned} \tag{3.15}$$

where $N + M = \sum\limits_{l=1}^{n} n_{\alpha_i} + \sum\limits_{r=1}^{m} m\beta_r$.

We have chosen in the formulae (3.1-4) the one-parameter family of the q-super-oscillators corresponding to the standard $GL_q(n|m)$ R-matrix given in [16].

4 Final remarks

We would like to mention that

a) The question arises whether q-oscillators describe quantum algebra defined as a quasitriangular Hopf algebra [1].

It appears that recent proposals for the comultiplication of q-oscillators obtained as the property of a particular class of realizations. It should be mentioned that adding the relations (3.8) to (3.6-7) means looking for the realizations of (3.6-7) invariant under the mapping $q \to q^{-1}$ (see e.g. [26]).

b) In order to describe the supercovariance of real q deformed graded phase space one shall consider also the Bogolubov type of transformations. For just one bosonic q-oscillator the transformations

$$\begin{aligned} A' &= aA + bA^* \\ A'^* &= b^* A + a^* A^* \end{aligned} \tag{4.1}$$

describe the q-deformed symplectic transformations $Sp_q(2) \simeq SL_q(2 : R)$ (see [27]).

We conjecture that for the case $n = 1$, $m = 1$ (Sect.2) the quantum covariance supergroups $SU_q(1|1)$ is extended to by allowing the Bogolubov type of transformations $OSp_q(2|2)$, and for arbitrary values of n and m (Sect.3) we obtain in such a way the extension of $SU_q(n|m)$ to $OSP_q(2n|2m)$.

c) Contrary to the case of conventional quantum mechanics ($q = 1$) the algebra of q-oscillators (3.6) has nonequivalent representations. Let us consider the system of n bosonic q-oscillators. The following n linearly independent operators

$$Z_i = q^{-N_i} \left([N_i]_q - a_i^+ a_i\right) \qquad i = 1, 2, \dots, n \tag{4.2}$$

describe the n-dimensional nontrivial center. We see that the conditions (3.8) imply that $Z_i = 0$. Analogously one can define nontrivial centers for the fermionic degree of freedom (3.7) (see [26])

d) It should be mentioned that the algebra (3.7) has a 2×2 realization the same for all values of q. Indeed, putting

$$b = \begin{pmatrix} q & 0 \\ 1 & 0 \end{pmatrix}, \qquad b^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad and \qquad M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{4.3}$$

one obtains that

$$q^{-M} = \begin{pmatrix} q^{-1} & 0 \\ 0 & 1 \end{pmatrix} = bb^+ + q^{-1} b^+ b. \tag{4.4}$$

More details about the realizations of the bosonic and fermionic q-oscillator algebras, including the description of coherent states for q-oscillators, is given in [27].

References

[1] V.G. Drinfeld, in Proceedings of Intern. Congress of Math. (MSRI, Berkeley, 1986), p.798

[2] M. Jimbo, Lett. Math.Phys. **10**(1985),63 ibid, **11** (1986),247

[3] L. Faddeev, N. Reshetikhin and L. Takhtajan, Algebra i Analiz **1** (1989), 178

[4] C.N. Yang and M.L. Ge (Eds.) "Braid Group, Knot Theory and Statistical Mechanics", World Scientific, Singapore 1989

[5] L. Brink, D. Friedan and A.M. Polyakov (Eds.) "Physics and Mathematics of String" (Memorial volume for V. Knizhnik), World Scientific, Singapore 1990

[6] P. Kulish (Ed.), Proc. First EIMI Workshop on Quantum Groups (October - December 1990), Springer Verlag, to appear in 1991

[7] L. Faddeev, in "Le Houches XXXIX, ed. J.-B. Zuber and R. Stora, Elsevier, Amsterdam 1984, course 8

[8] P. Kulish and E. Sklyanin, Lecture Notes in Physics, **151** (1984), p.61

[9] V. Pasquier and H. Saleur, Nucl. Phys. **B330** (1990), 523

[10] S.L. Woronowicz, Publ. RIMS (Kyoto) **23**, No 1(1987),117

[11] J. Wess and B. Zumino, CERN preprint TH-5697/90

[12] Yu. I. Manin, Comm.Math.Phys. **123** (1989), 163

[13] W.B. Schmidke, S.P. Volios and B. Zumino, Zeitschr. Phys. **C48**, 249 (1990)

[14] see e.g. L. Alvarez-Gaum'e, G. Sierra and Gomez in [5], p.16

[15] P. Kulish, Kyoto preprint, RIMS-615(1989); P. Kulish and N. Reshetikhin, Lett. Math. Phys. **18** (1989),143

[16] M. Chaichian and P. Kulish, Phys. Lett. **B234** (1990), 72

[17] V.N. Tolstoy, in Proc. of the Quantum Group Workshop, Clausthal, July 1989; publ. World Scientific, Singapore, 1990

[18] D. Coon, S. Yu and M. Baker, Phys. Rev. **D5** (1972), 1429

[19] A. MacFarlane, Journ.Phys. **A22** (1989), 4581

[20] L. Biedenharn, Journ.Phys. **A22** (1989), L873

[21] T. Hayashi, Comm. Math. Phys. **127** (1990) 129

[22] M. Chaichian, P. Kulish and J. Lukierski Phys. Lett. **B237** (1990), 401

[23] W. Pusz and S.L. Woronowicz, Rep. Math. Phys. **27** (1989), 231

[24] W. Pusz, Rep. Math.Phys. **27** (1989), 349

[25] P. Kulish, Zap. Naucz. Semin. LOMI, **109** (1981), 83; the lecture in [6], to be published

[26] P. Kulish and E.V. Damaskinsky, J.Phys. **A23**, 415(1990)

[27] A.T. Filipov, D. Gangopadhyay and A. Isaev, J.Phys. **A24**, L63(1991)

[28] P. Kulish and E.V. Damaskensky , Zap. Naucz. Sem. LOMI, **Vol. 189**, 1991 (in press).

Q-DEFORMATIONS OF CONFORMAL ALGEBRA AND ITS CENTRAL EXTENSION

A. Isaev
Lab. Theor. Physics, JINR, Dubna, USSR

Z. Popowicz
Institute of Theoretical Physics, University of Wrocław,
Cybulskiego 36, 50-205 Wrocław, Poland

Abstract

We describe q-deformations of the Witt (conformal) algebra realizations depending on the conformal dimension parameter Δ. Restricting the consideration to the q-deformed subalgebras with three generators $L_{\pm 1}, L_0$, we obtain the two-parameter generalization of $U_q(sl(2,R))$.

The special numbers $\Delta = 0, 1/2, 1, 2$ give us a list of known alternative q-deformations of $sl(2,R)$ which have been presented earlier by other authors.

The nontrivial central extension term describing q-deformation of the Virasoro algebra is derived.

This report is largely based on the results presented in [1].

1 Introduction

Recently, a great deal of attention has been paid to the study of quantum groups and algebras [2]. These new mathematical objects were applied to some physical models, e.g., in the conformal field theories [3] as well as in the vertex and spin models (see e.g. [4]) and quantum optics [5]. The deformed creation and annihilation operators (see e.g. [6]) were also applied to constructing different q-(super)algebras and new statistics with violation of Pauli principle [7].

The Virasoro algebra as a $d = 2$ conformal algebra has played a prominent role in physics in the last twenty years while the study of its q-deformation started only recently. In particular, in [8] it has been shown how it is possible to construct the q-deformed oscillators. In such a way, one obtains the q-deformation of the conformal algebra with conformal dimension $\Delta = 0$. Next, in [10] (see also [12]) a central term for the deformed Virasoro algebra has been obtained. However, it appeared in [10] that it reproduced only a trivial part of the usual central term in the limit $q \to 1$.

Our starting point in this report is to consider the q-deformations of the differential realizations of the conformal algebra depending on arbitrary conformal dimension Δ, what is represented in the sect. 2. It appears that the cases $\Delta = 0, 1$ are much easily considered separately. We do this in Sect. 3. where we also consider the interesting cases $\Delta = 1/2, 2$ and present the subalgebras $U_q\ sl(2,R)$ embedded in the q-deformed conformal algebra for $\Delta = 0, 1/2, 1, 2$. These different q-deformations of $sl(2,R)$ are related to the deformations obtained by Witten, Drinfeld, Jimbo and Sklyanin (see [11] and references therein). It seems that the case $\Delta = 3/2$ does not give us the known q-deformation of $sl(2,R)$.

In Sect. 4., using two different approaches we obtain the nontrivial central extension term of the q-deformed conformal algebra. In the one of these we utilize the infinitesimal q-deformed conformal transformations of the stress- energy tensor. These transformations are described by the $\Delta = 2$ realization of the q-deformed Virasoro algebra.

The final remarks and some outlook are presented in the last section.

2 Deformation of Conformal Algebra Realizations with arbitrary Conformal Dimension.

Let $A_\Delta(z)$ be an arbitrary primary field with the conformal dimensions Δ. This means that $A_\Delta(z)$ has the following conformal transformation property.

$$A_\Delta(z) \to (\phi'(z))^\Delta A_\Delta(\phi(z)), \tag{2.1}$$

where $\phi(z)$ is an arbitrary function, prime denotes the derivative and Δ is an arbitrary number called the conformal dimension. Formula (2.1) in the infinitesimal

form reads as

$$\delta_\epsilon(A_\Delta(z)) = \epsilon^{1-\Delta}\partial(\epsilon^\Delta A_\Delta(z)), \tag{2.2}$$

where $\phi(z) = z + \epsilon(z)$. If we take $\epsilon(z) = z^{n+1}$, we obtain

$$\delta_n A_\Delta(z) = l_n A_\Delta(z) = z^{(1-\Delta)(n+1)}\partial(z^{\Delta(n+1)}A_\Delta) = \\ z^n(z\partial + \Delta(n+1))A_\Delta \tag{2.3}$$

and the generators l_n fulfill the usual commutation relations

$$[l_n, l_m] = (m-n)l_{n+m}. \tag{2.4}$$

It seems that the most general way to deform the algebra (2.4) with generators represented in (2.3) is to write

$$L_n = z^n f_n(z\partial), \tag{2.5}$$

where $f_n(z\partial)$ are some arbitrary functions to be defined below. It is easy to obtain the basic deformed commutators for L_n

$$[L_n, L_m]_{(1,y^{nm})} \equiv L_nL_m - y^{nm}L_mL_n = \\ = \{\phi_{nm}(z\partial) - y^{nm}\phi_{mn}(z\partial)\}\, L_{n+m}, \tag{2.6}$$

where y^{nm} are arbitrary numbers and

$$\phi_{nm}(z\partial) = \frac{f_n(z\partial - m)f_m(z\partial - n - m)}{f_{n+m}(z\partial - n - m)}.$$

We see that the structure constants in the commutators (2.6) depend on the operator $z\partial$ related to $L_0 = f_0(z\partial)$ (see eq. (2.5)).

The main problem now is to find the functions $f_n(z\partial)$ in (2.5) and to obtain the deformed algebra (2.6) which possesses the following requirements:
1) the algebra (2.6) has a nontrivial central extension term
2) this algebra has the structure of the Hopf algebra.

These two requirements must give us some restrictions on the form of the functions $f_n(z\partial)$. Resolving these restrictions we could define the algebra (2.6) explicitly. However, it seems that such a general way is very difficult to realize. Therefore, we shall limit ourselves to the consideration of some special cases.

Let us construct the q-analogy of formula (2.2). For this purpose, let us define it by

$$\delta(A_\Delta(z)) = (\epsilon(z))^{1-\Delta}D_q(\epsilon(z)^\Delta A_\Delta(z)), \tag{2.7}$$

where the q-deformed derivative D_q is defined as usual

$$D_q = \frac{1}{z}[z\partial]. \tag{2.8}$$

Here $[x] = \dfrac{q^x - q^{-x}}{q - q^{-1}}$ and q is the deformation parameter.

Introducing $\epsilon = z^{n+1}$ into (2.7) we obtain

$$\delta_n^q A_\Delta(z) \;= L_n^{(\Delta)}A_\Delta(z) = z^{n-\Delta(n+1)}[z\partial]z^{\Delta(n+1)}A_\Delta(z) \\ = z^n[z\partial + \Delta(n+1)]A_\Delta(z). \tag{2.9}$$

Notice that the operators $L_n^{(\Delta)}$ have been also introduced in [12]. We can rewrite L_n^Δ in the form of L_n represented in eq. (2.5) if we choose the functions $f_n(z\partial)$ as

$$f_n^{(\Delta)}(z\partial) = [z\partial + \Delta(n+1)]. \tag{2.10}$$

In this way, the deformed commutators (2.6) take the form

$$\left[L_n^{(\Delta)}, L_m^{(\Delta)}\right]_{(1,y_\Delta^{nm})} \\ = \frac{1}{q - q^{-1}}\left\{q^{N_\Delta}\left(q^{-n} - y_\Delta^{nm}q^{-m}\right) - q^{-N_\Delta}\left(q^n - y_\Delta^{nm}q^m\right)\right\}L_{n+m}^{(\Delta)}, \tag{2.11}$$

where $N_\Delta = z\partial + \Delta$, i.e. $L_0^{(\Delta)} = [N_\Delta]$ and

$$y_\Delta^{nm} = \frac{[m(\Delta-1)][n\Delta]}{[n(\Delta-1)][m\Delta]} = (y_\Delta^{mn})^{-1}. \tag{2.12}$$

Let us notice that the numbers y_Δ^{nm} are chosen in such a way that the structure constants of the algebra (2.11) take the simple form like $y_1q^{N_\Delta} + y_2q^{-N_\Delta}$. We see that these operator-valued structure constants depend on the conformal dimension Δ. On the other hand, from the definition (2.9) we can deduce that

$$L_n^{(\Delta')} = (U_{n,\Delta'-\Delta})^{-1}L_n^{(\Delta)}U_{n,\Delta'-\Delta}, \quad U_{n,\Delta} = z^{\Delta(n+1)}. \tag{2.13}$$

Thus, we have the gauge equivalence between $L_n^{(\Delta)}$ for the same n and different Δ. Unfortunately, only for the conventional commutator the algebra is invariant under the mapping (2.13) and for arbitrary two dimensions this transformation relates two different q-deformed algebras.

In view of the explicit form of the numbers y_Δ^{nm} (see (2.12)), the cases $\Delta = 0, 1$ require special consideration. In the next Section, we will consider these cases and also three other $\Delta = 1/2, 3/2, 2$.

3 Q-Deformation of Conformal algebra Realizations and its Subalgebras $U_q(sl(2,R))$ with $\Delta = 0, 1/2, 1, 3/2, 2$.

Let us write the realizations (2.9) for $\Delta = 0, 1$

$$L_n^{(0)} = [z\partial - n]z^n, \quad L_n^{(1)} = [z\partial + 1]z^n \tag{3.1}$$

from which we have ($\Delta = 0, 1$)

$$L_n^{(\Delta)} L_m^{(\Delta)} = [z\partial + \Delta - n]\, L_{n+m}^{(\Delta)} \tag{3.2}$$

From the last formulas follows

$$\left[L_n^{(\Delta)}, L_m^{(\Delta)}\right]_{(1, y^{nm})} = \frac{1}{q - q^{-1}} \left\{ q^{N_\Delta} \left(q^{-n} - y^{nm} q^{-m}\right) - q^{-N_\Delta} \left(q^n - y^{nm} q^m\right) \right\} L_{n+m}^{\Delta}, \tag{3.3}$$

where $\Delta = (0, 1)$, $N_\Delta = z\partial + \Delta$, $L_0^{(\Delta)} = [N_\Delta]$ and y^{nm} are arbitrary numbers. In particular, choosing $q^{-n} - y^{nm} q^{-m} = 0$ and redefinig $\hat{L}_n^{(\Delta)} = q^{N_\Delta} L_n^{(\Delta)}$ one obtains [8,9]

$$\left[\hat{L}_n^{(\Delta)}, \hat{L}_m^{(\Delta)}\right]_{(q^{n-m}, q^{m-n})} = [m - n]\hat{L}_{n+m}^{(\Delta)}. \tag{3.4}$$

If we put $y^{nm} = 1$ in eq. (3.3), then we obtain surprising result: the commutation relations (3.3) reduce to ($\Delta = 0, 1$)

$$\left[L_n^{(\Delta)}, L_m^{(\Delta)}\right] = \left[\frac{m-n}{2}\right] \left\{ q^{N_\Delta - \frac{n+m}{2}} + q^{-N_\Delta + \frac{n+m}{2}} \right\} L_{n+m}^{(\Delta)}. \tag{3.5}$$

It is worth noting that the same commutation relations can be obtained from eqs. (2.11) if $\Delta = 1/2$ ("spinorial case"). One should choose here $y_{1/2} = 1$ and we have

$$\left[L_n^{(1/2)}, L_m^{(1/2)}\right] = \left[\frac{m-n}{2}\right] \left\{ q^{N_{1/2} - \frac{m+n}{2}} + q^{-N_{1/2} + \frac{n+m}{2}} \right\} L_{n+m}^{(1/2)}. \tag{3.6}$$

The relations (3.5) and (3.6) have been recently considered in [13].

Let us also consider the algebra (2.11) for $\Delta = 2$, with the generators

$$L_n^{(2)} = z^{-(n+1)} D_q z^{2(n+1)} \tag{3.7}$$

which reads as follows:

$$\left[L_n^{(2)}, L_m^{(2)}\right]_{(q^m + q^{-m}, q^n + q^{-n})} = [m - n] \left(q^{N_2} + q^{-N_2}\right) L_{n+m}^{(2)}, \tag{3.8}$$

where again we have $L_0^{(2)} = [N_2]$. This algebra will be important for our further consideration. Introducing a new generator

$$\bar{L}_0^{(2)} = \left(q^{N_2} + q^{-N_2}\right) \tag{3.9}$$

one can rewrite eq. (3.8) in the form

$$\left[L_n^{(2)}, L_m^{(2)}\right]_{(q^m + q^{-m}, q^n + q^{-n})} = [m - n]\, \bar{L}_0^{(2)} L_{m+n}^{(2)}, \tag{3.10}$$

$$\left[\bar{L}_0^{(2)}, L_m^{(2)}\right]_{(q^m + q^{-m}, 2)} = \left(q^m - q^{-m}\right) \left(q - q^{-1}\right) L_0^{(2)} L_m^{(2)}, \tag{3.11}$$

$$\bar{L}_0^{(2)} \bar{L}_-^{(2)} - (q - q^{-1})^2 L_0^{(2)} L_0^{(2)} = 4. \tag{3.12}$$

As it will be seen below the subalgebra $L_0, \bar{L}_0, L_{\pm 1}$ reproduce the Sklyanin's trigonometric deformation of $sl(2,R)$ [11,14]. Thus, it is natural to call the algebra (3.10)-(3.12) the trigonometric deformation of the conformal algebra.

At the end of this section we consider the q-deformations of $sl(2,R)$ algebras emerged as subalgebras $\left\{L_0^{(\Delta)}, L_{\pm 1}^{(\Delta)}\right\}$ of the algebras (2.11) and (3.3). From eq. (2.11) we have the two-parameter deformation of $sl(2,R)$

$$\left[L_0^{(\Delta)}, L_\pm^{(\Delta)}\right]_{(1, \frac{[\Delta-1]\Delta}{[\Delta][\Delta-1]})} = \left\{ [N_\Delta] - \frac{[\Delta - 1]\Delta}{[\Delta](\Delta - 1)} \, [N_\Delta \mp 1] \right\} L_\pm^{(\Delta)}, \tag{3.13}$$

$$\left[L_+^{(\Delta)}, L_-^{(\Delta)}\right] = -\left\{ q^{N_\Delta} + q^{-N_\Delta} \right\} L_0^{(\Delta)}, \quad L_0^{(\Delta)} = [N_\Delta].$$

For all these algebras we can introduce the comultiplication

$$\Delta(N_\Delta) = N_\Delta \otimes 1 + 1 \otimes N_\Delta, \quad \Delta\left(L_\pm^{(\Delta)}\right) = L_\pm^{(\Delta)} \otimes q^{N_\Delta} + q^{-N_\Delta} \otimes L_\pm^{(\Delta)}$$

which specify Drinfeld - Jimbo type $U_q(sl(2,R))$ algebra.

Now we would like to consider some special cases. Let us put $\Delta = 1/2$ in (3.13), then we obtain

$$\left[L_0^{(1/2)}, L_\pm^{(1/2)}\right] = \left\{ \left[N_{1/2}\right] - \left[N_{1/2} \mp 1\right] \right\} L_\pm^{(1/2)}, \tag{3.14a}$$

$$\left[L_+^{(1/2)}, L_-^{(1/2)}\right] = -\left\{ q^{N_{1/2}} + q^{-N_{1/2}} \right\} L_0^{(1/2)}. \tag{3.14b}$$

Using $L_0^{(1/2)} = [N_{1/2}]$ one can rewrite eqs. (3.14) in the form

$$\left[N_{1/2}, L_\pm^{(1/2)}\right] = L_\pm^{(1/2)}, \quad \left[L_+^{(1/2)}, L_-^{(1/2)}\right] = -\left[2N_{1/2}\right] \tag{3.15}$$

and (3.15) give us the well known Drinfeld - Jimbo deformation of $sl(2,R)$ [2,11].

Let us consider the subalgebra $\left\{L_0^{(\Delta)}, \bar{L}_0^{(\Delta)}, L_{\pm n}^{(\Delta)}\right\}$ in the case $\Delta = 2$ (see eqs. (3.9) - (3.12)). Then we have the following algebra ($q = exp(2u)$):

$$[L_0^{(2)}, L_n^{(2)}]_- = \frac{tanh(nu)}{2sinh(2u)}\,[\bar{L}_0^{(2)}, L_n^{(2)}]_+,$$

$$\left[\bar{L}_0^{(2)}, L_n^{(2)}\right]_- = 2tanh(nu)sinh(2u)\left[L_0^{(2)}, L_n^{(2)}\right]_+ \tag{3.16}$$

$$\frac{1}{4}\left(\bar{L}_0^{(2)}\right)^2 - sinh^2(2u)\left(L_0^{(2)}\right)^2 = 1.$$

Putting $n = \pm 1$ and redefining the generators as

$$L_0 \to \frac{1}{2\,sinh(2u)cosh(u)}\,S_3, \quad \bar{L}_0 \to \frac{1}{sinh(u)}\,S_0, \quad L_\pm \to \pm\frac{1}{2\,sinh(2u)}\,S_\pm,$$

we obtain Sklyanin's trigonometric deformation of $sl(2,R)$ [11,14]

$$[S_0, S_3]_- = 0, \quad [S_+, S_-]_- = 4S_0S_3, \quad [S_3, S_\pm]_- = \pm[S_0, S_\pm]_+,$$

$$[S_0, S_\pm]_- = \pm tanh^2u[S_\pm, S_3]_+, \quad S_0^2 - S_3^2tanh^2u = 4sinh^2u.$$

The cases $\Delta = 0, 1$ rewritten in the form (3.4) yield Witten's second deformation [11,15]

$$[\hat{L}_+, \hat{L}_-]_{(q^2,q^{-2})} = \hat{L}_0$$

$$[\hat{L}_0, \hat{L}_\pm]_{(q^{\pm 1},q^{\mp 1})} = \pm\hat{L}_\pm. \tag{3.17}$$

So one can see that the cases $\Delta = 0, 1$ give us Witten's deformation of $sl(2,R)$, the case $\Delta = 1/2$ yields the Drinfeld - Jimbo - type of $U_q(sl(2,R))$ while $\Delta = 2$ reproduces Sklyanin's trigonometric deformation.

At the end of this Section, for completeness we bring the deformation of $sl(2,R)$ for $\Delta = 3/2$

$$\left[L_+^{(3/2)}, L_-^{(3/2)}\right] = -\left\{q^{N_{3/2}} + q^{-N_{3/2}}\right\}L_0^{(3/2)},$$

$$\left[L_0^{(3/2)}, L_\pm^{(3/2)}\right]_{(q^{-1}+1+q,\,3)} = \left\{\pm\left(q^{N_{3/2}} + q^{N_{-3/2}}\right) + [N_{3/2}] - [N_{3/2} \mp 1]\right\}L_\pm^{(3/2)}.$$

4 Deformed conformal algebra with central extension and q-deformed stress-energy tensor.

Let us notice that formula (2.11) can be cast also into the following form:

$$\left[L_n^{(\Delta)}, L_m^{(\Delta)}\right]_* = [m-n]L_{n+m}^{(\Delta)}, \tag{4.1}$$

where

$$\left[L_n^{(\Delta)}, L_m^{(\Delta)}\right]_* = S_{n,m}^{\Delta}L_n^{(\Delta)}L_m^{(\Delta)} - S_{m,n}^{\Delta}L_m^{(\Delta)}L_n^{(\Delta)} \tag{4.2}$$

with

$$S_{n,m,}^{\Delta} = [m-n]\left\{[N_\Delta - n] - y_\Delta^{nm}[N_\Delta - m]\right\}^{-1} = y_\Delta^{mn}S_{m,n}^{\Delta} \tag{4.3}$$

And y_Δ^{nm} have been defined in (2.12). Let us notice that $S_{m,n}^{\Delta}$ are in fact operators. The simplest version of the algebra (4.1) - (4.3) has been presented in (3.4) for $\Delta = 0, 1$. Taking double deformed commutators, we easily see that our algebra (4.1) - (4.3) implies the following relation:

$$\left[L_k^{(\Delta)}, \left[L_n^{(\Delta)}, L_m^{(\Delta)}\right]_*\right]_* = [m-n][m+n-k]L_{m+n+k}. \tag{4.4}$$

From this relation it is easy to obtain that the double commutators (4.4) have to satisfy the following q-deformed Jacobi relation:

$$(q^k + q^{-k})\left[L_k^{(\Delta)}, \left[L_n^{(\Delta)}, L_m^{(\Delta)}\right]_*\right]_* + \text{cycl.perm.} = 0. \tag{4.5}$$

We can obtain the three-linear relations on generators $L^{(\Delta)}$ as a consequence of the relation (4.5) and definition (4.2). These relations in the simplest case of the algebra (3.4) have the form

$$\left(q^{m+k} - q^{-m-k}\right)\left(q^{m-k}L_mL_nL_k - q^{k-m}L_kL_nL_m\right) + \text{cycl.perm.} = 0. \tag{4.6}$$

It should be stressed that only in the limit $q \to 1$ this formula takes the form of an algebraic identity.

The main merit of relation (4.5) is in searching the deformation of the central term of the Virasoro algebra. Indeed, let us write

$$\left[L_n^{(\Delta)}, L_m^{(\Delta)}\right]_* = [m-n]L_{n+m}^{(\Delta)} + \hat{c}_{n,}(q) \tag{4.7}$$

and assume that $\hat{c}_{n,m}(q) = \delta_{n+m,0}\hat{c}(n,q)$, i.e. $\hat{c}$ behaves in the deformed commutators as $L_0^{(\Delta)}$. Further, we assume that

$$\left[L_n^{(\Delta)}, \hat{c}(k,q)\right]_* = S_{n,0}L_n^{(\Delta)}\hat{c}(k,q) - S_{0,n}\hat{c}(k,q)L_n^{(\Delta)} = 0. \tag{4.8}$$

In particular, if $S_{n,0} = q^n$ and $S_{0,n} = q^{-n}$, corresponding to the algebra (3.4), the operator part of the solution $\hat{c}(k,q)$ of eq. (4.8) does not depend on k, i.e. one can write

$$\hat{c}(k,q) = c(k,q)q^{2N_\Delta}. \tag{4.9}$$

For such a case the substitution of (4.7) - (4.9) into (4.5) gives

$$(q^k + q^{-k})[m-n]c(k,q)\delta_{k+m+n,0} + \text{cycl. perm} = 0. \tag{4.10}$$

The general solution of (4.10) can be written down in the following compact form:

$$c(n,q) = c(q)\,\frac{[n-1][n][n+1]}{q^n+q^{-n}} + c_o(q)[2n], \tag{4.11}$$

where $c(q)$ and $c_0(q)$ are arbitrary functions of q.

Now let us note that the explicit form of the central term (4.9), (4.11) follows from our commutation relation (4.1) and Jacobi relation (4.5). To justify deeply this we obtain the same result using quite different approach. For this we introduce the notion of q-deformation $T_q(z)$ of the stress-energy tensor $T(z)$ playing important role in the 2-dimensional conformal field theories.

It is well known [16] that the holomorphic field $T(z)$ with conformal dimension $\Delta = 2$ is not transformed as the primary field (see (2.1)). Indeed, its conformal transformations look as follows:

$$T(z) \to (\phi'(z))^2 T(\phi(z)) + c\left(\frac{\phi'''}{\phi'} - \frac{3}{2}\left(\frac{\phi''}{\phi'}\right)^2\right), \tag{4.12}$$

where the last term proportional to the central charge c describes the central extension of the conformal transformation algebra (see [16]). In the infinitesimal form we have

$$\delta_\epsilon T(z) = (\epsilon(z)\partial + 2\epsilon'(z))T(z) + c\epsilon'''(z). \tag{4.13}$$

After the substitution $\epsilon(z) = z^{n+1}$, formula (4.13) reduces to

$$\delta_n T(z) = l_n T(z) + g(n) z^{n-2}, \tag{4.14}$$

where l_n is given by (2.3) with $\Delta = 2$ and

$$g(n) = c(n-1)n(n+1). \tag{4.15}$$

If the central term in eq. (4.14) is unknown, its form can be obtained if we assume that the variations (4.14) commute as the conformal algebra (2.4), i.e.

$$[\delta_m, \delta_n]T(z) = (m-n)\delta_{n+n}T(z). \tag{4.16}$$

Substituting here the definition (4.14) we obtain the equation

$$q(m)(2n+m) - g(n)(2m+n) = (m-n)g(m+n)$$

with the general solution $g(m) = cm^3 + c_0 m$ where c and c_0 are arbitrary constants. Using the redefinition $T(z) \to T(z) + \frac{1}{2}(c-c_0)z^{-2}$ we can fix $c_0 = c$ and arrive to the known result (4.15).

Now we would like to apply this approach to the q-deformed algebra (3.7), (3.8).

The q-analogue of the formula (4.14) is

$$\delta_n^q T_q(t) = L_n^{(2)} Tg(z) + c(n,q) z^{n-2}. \tag{4.17}$$

This formula can be considered as the definition of the deformed stress - energy tensor $T_q(z)$ which is the holomorphic field with the conformal q-dimension $\Delta = 2$. The q-deformed conformal transformations (4.17) of this field contains the nonhomogeneous part $c(n,q)z^{n-2}$ which is identified with the central terms. To find the explicit expression for the function $c(n,q)$ we rewrite the consistency condition (4.16) in the q-deformed form (compare with eq. (3.8))

$$[\delta_m^q, \delta_n^q]_{(q^m+q^{-m}, q^n+q^{-n})} T_q(z) = [m-n]\left\{q^{N_2} + q^{-N_2}\right\}\delta_{m+n}^q T_q(t). \tag{4.18}$$

Substituting (4.17) into (4.18) and using (3.8) we obtain

$$\begin{aligned}&\left(q^m + q^{-m}\right)[2m+m]c(m,q) - \left(q^n + q^{-m}\right)[2m+n]c(n,q)\\ &= [m-n]\left(q^{n+m} + q^{-n-m}\right)c(m+n,q).\end{aligned} \tag{4.19}$$

This equation coincides with equation (4.10). Thus, the general solution of eq. (4.19) is given by formula (4.19). Let us note that we can redefine the deformed field $T_q(z) \to T_q(z) + c_0(q)z^{-2}$ in eq. (4.17) and thus to put $c_0(q) = 0$ in the solution (4.11).

Now we introduce the Fourier coefficients L_n^q for the field $T_q(z)$

$$L_n^q = \frac{1}{2\pi i}\oint dz\, z^{n+1} T_q(z), \quad T_q(z) = \sum_k z^{-k-2} L_k^q. \tag{4.20}$$

Using (4.20) we can rewrite the variation (4.17) in terms of L_n^q in the form

$$\begin{aligned}\delta_n^q L_m^q &= [n-m]L_{n+m}^q + c(n,q)\delta_{m+n,0},\\ c(n,q) &= c(q)\frac{[n-1][n][n+1]}{q^n+q^{-n}}.\end{aligned} \tag{4.21}$$

The consistency condition (4.18) in terms of L_n^q looks like

$$[\delta_m^q, \delta_n^q]_{(q^m+q^{-m}, q^n+q^{-n})} L_k^q = [m-n]\left(q^k + q^{-k}\right)\delta_{m+n}^q L_k^q \tag{4.22}$$

According to the nondeformed case, as in [16] we conjecture that the conformal variations δ_n^q (4.21) are generated by the Fourier coefficients L_n^q of the deformed stress-energy tensor. It means that

$$\delta_n^q L_m^q = [[L_m^q, L_n^q]] = -[[L_n^q, L_m^q]], \tag{4.23}$$

$$[[L_m^q, L_n^q]] = [n-m]L_{n+m}^q + c(n,q)\delta_{n+m,0}, \tag{4.24}$$

where $[[\cdot,\cdot]]$ is some antisymmetric bilinear mapping (bracket product) such that $\{L_m^q\} \otimes \{L_n^q\} \to \{L_k^q\}$. We would like to stress here that this mapping was introduced also in [17] but its motivation was somewhat obscure.

From the fact that $\delta_n^q c(m,q) = 0$ we obtain

$$[[L_n^q, c(m,q)]] = 0. \tag{4.25}$$

Then, using eqs. (4.21), (4.23) and the formula $\delta_m^q \delta_n^q L_k^q = [[\,[[L_k^q, L_n^q]], L_m^q]]$ we can rewrite the condition (4.22) in the form of the Jacobi "identity" for the brackets $[[\cdot,\cdot]]$ as (compare with eq. (4.5))

$$\left(q^k + q^{-k}\right)[[\,[[L_n^q, L_m^q]],\ L_k^q]] + \ cycl.perm. = 0. \tag{4.26}$$

In view of eq. (4.26) the following question naturally emerges whether it is possible to realize the bracket mapping $[[\cdot,\cdot]]$, satisfying the axioms (4.23), (4.24) and (4.26), in the form of the q deformed commutators like (4.2).

$$[[L_n^q, L_m^q]] = S_{n,m} L_n^q L_m^q - S_{m,n} L_m^q L_n^q. \tag{4.27}$$

In general, $S_{m,n}$ can be the functions of the parameter q and the number operator related to L_0^q. Assuming that the operator products $L_m L_n L_k$, $L_n L_k L_m, \ldots$ (for diverse m,n,k) are independent objects and substituting (4.27) into (4.26) we obtain the braiding relations for $S_{m,n}$ which, as one can see, are inconsistent for $q \neq 1$. Hence, we conclude that in general the realization (4.27) is not possible for the mapping $[[\cdot,\cdot]]$. The algebraic meaning of this mapping is an open problem up to now.

At the end of this section we note that the relations (4.26) coincide with eqs. (4.5) which have been used to find the central extension for the q-deformed Virasoro algebra (4.7). The central terms in (4.11) and (4.21) are the same up to some inessential redefinitions.

Thus, we conclude that our second approach to introduce central extension for the q-deformed conformal algebra justifies our approach used in the beginning of this section.

5 Conclusion

In this report we have presented two different ways to derive the nontrivial central extension for the q-deformed conformal algebra (to obtain the q-deformed Virasoro algebra). In both approaches we have obtained the identical results. The starting point of our consideration is the construction of the q-deformations of the conformal algebra $\left\{L_n^{(\Delta)}\right\}$ depending on the conformal dimension parameter Δ.

It is interesting that considerating the subalgebras generated by the three operators $L_\pm^{(\Delta)}$ and $L_0^{(\Delta)}$, we reproduce for $\Delta = 0, 1/2, 1, 2$ the well-known $U_q(sl(2,R))$ algebras.

We have not discussed here the problem of deducing the comultiplication for the conformal algebra (2.11). These were attempts [13] to do this for some special cases $\Delta = 0, 1/2, 1$. But these attempts seem to be not successful because the reduced comultiplication for the subalgebras $U_q(sl(2,R))$ does not reproduce the known comultiplications [2,3,11,15].

As an interesting idea how to apply our results we would like to note the possibility to derive the q-deformation of the KdV equation and its hierarchy using the relations (4.7) and (4.24) as canonical commutators.

At last, we would like to stress that our class of Δ-dependent (representation - dependent) q-deformations (2.11) of the Witt algebra may be only an intermediate step to introduce a Δ-independent q-deformed conformal algebra. It is our aim now to find such a "genuine" algebra.

Acknowledgments
We are indebted to S. Saito and V. Tolstoy and especially to M. Chaichian, J. Lukierski and P. Prešnajder for valuable discussions and help.

References

[1] M. Chaichian, A. Isaev, J. Lukierski, Z. Popowicz, P. Prešnajder, preprint CERN-TH. 5933/90, 1990, to appear in Phys. Lett. B.

[2] V.G. Drinfeld, Proc. ICM Berkeley, Vol. 1, 798 (Academic Press, 1987); M. Jimbo, Lett. Math. Phys. **11**, (1986) 247;
L. Faddeev, N. Reshetikhin and L. Takhtajan, Algebra and Analysis, 1, N1 (1989) 178 (in Russian)
Yu. I. Manin, Quantum Groups, Univ. de Montreal, 1988;
S.L. Woronowicz, Publ. RIMS, Kyoto Univ. **23** (1987) 117

[3] L. Alvarez - Gaume, C. Gomez and G. Sierra, Nucl. Phys. **B330** (1990) 347;
G. Moore and N. Reshetikhin, Nucl. Phys. **B220** (1989) 557

[4] H. de Vega, Intern. Journ. Mod. Phys. **A4** (1989) 2371;

[5] M. Chaichian, D. Ellinas and P. Kulish, Phys. Rev. Lett. **65** (1990) 980

[6] A.J. Macfarlane, J. Phys. **A22** (1989) 4581;
L.C. Biedenharn, J. Phys. **A22** (1989) L873;
M. Chaichian and P. Kulish, Phys. Lett. **B234** (1990) 72;

P. Kulish and E.V. Damaskinsky, J. Phys. **A23** (1990) L415;
A.T. Filippov, D. Gangopadhyay and A. Isaev, J. Phys. **A24** (1991) L63

[7] O.W. Greenberg, Phys. Rev. Lett. **64** (1990) 705;
R.N. Mohapatra, Phys. Lett. **B242** (1990)407

[8] M. Chaichian, P. Kulish and J. Lukierski, Phys. Lett. **B237** (1990) 401

[9] T. Curtright and C. Zachos, Phys. Lett. **B243** (1990) 237

[10] M. Chaichian, D. Ellinas and Z. Popowicz, Phys. Lett. **B248** (1990) 95

[11] C. Zachos, "Paradigms of Quantum Algebras", preprint ANL-HEP-PR-90-61 (1990)

[12] H. Hiro-Oka, O. Matsui, T. Naito and S. Saito, preprint TMUP-HEL 9004 (1990);
S. Saito - Talk on this school

[13] Ch. Devchand and M.V. Saveliev, preprint IHEP, Protvino (1990)

[14] E. Sklyanin, Funct. Anal. Appl. **16** (1982) 285

[15] E. Witten, Nucl. Phys. **B330** (1990) 285

[16] A.A. Belavin, A.M. Polyakov and A.B. Zamolodchikov Nucl. Phys. **B241** (1984) 33

[17] N. Aizawa and H. Sato, preprint Horoshima University, HUPD - 9012 (1990)

PART III

RANDOMNESS VERSUS QUANTIZATION

Topics in the quantum theory of non-integrable systems

R. Vilela Mendes[2]
TH Division, CERN - CH-1211 Geneva 23

1. Introduction

According to what is taught in any quantum mechanics textbook, to solve a quantum problem one should :
(i) Diagonalize (or identify the spectral subspaces) of a self-adjoint operator (the Hamiltonian) acting in a Hilbert space.
(ii) Find a complete set of operators commuting with the Hamiltonian, their eigenvalues defining the one-dimensional subspaces in Hilbert space (or eventually in the dual of some nuclear subspace).

Comparing with classical mechanics one sees that a similar situation is only expected to occur for a very rare class of systems, called the integrable systems, where the existence of N constants of motion in involution allows the foliation of the 2N-dimensional phase space into N-dimensional invariant tori. How are we then to characterize the Hilbert space of non-integrable systems if already at the classical level no complete set of analytical invariants is known to exist ? What new quantum

[2]On leave from Centro de Física da Matéria Condensada, Av. Gama Pinto 2, 1699 Lisboa Codex, Portugal

phenomena should we expect for non-integrable systems ? Is there anything similar to the complex and chaotic behaviour of classical systems ? These are questions that need to be clarified to complete the framework of quantum theory.

These are important questions not only because the class of integrable systems is relatively small but also because they begin to have a direct bearing on the interpretation of experimental results. I refer for example to the excitation and ionization of hydrogen in a microwave cavity, the behaviour of atom beams in strong magnetic fields, the dynamical interpretation of nuclear energy levels and possibly the practical applications of complex dynamical behaviour inside condensed matter.

A complete, satisfactory, formulation of the quantum theory of non-integrable systems does not yet exist. A few problems have however been identified and intensely pursued. One is the establishment of a connection between the energy level separation statistics and the complexity of the corresponding classical system. Another problem which is to quantum mechanics as KAM theory is to classical mechanics, is the *stability problem.* It may be formulated as follows :

Consider a system described by a Hamiltonian

$$H = H_0 + V$$

where H_0 is integrable and V a non-integrable perturbation which may or may not depend on time. Given a system initially in a state which is localized in phase space, does the evolution under the perturbed Hamiltonian lead to delocalization ? In classical mechanics, by *localized state* we mean a state living in one of the N-dimensional tori of the phase space foliation associated to H_0. In quantum mechanics, states in the pure point spectrum stay confined in time in some region of space and states in the continuous spectrum escape to infinity. Quantum stability is therefore equivalent to the question : "If one perturbs a quantum system whose unitary evolution operator has a pure point spectrum, will the resulting evolution operator still have a pure point spectrum ?" Of interest is not only the discussion of resonant versus non-resonant perturbations, but also the question of whether situations that are stable for small perturbations will reach a threshold of instability when the intensity of the perturbation is increased.

Because of its general interest as a background for most of the current research in the field, I will present in Sections 2-4 a short review of the results on recurrence and

stability in quantum mechanics. Finally, in Sections 5 and 6 I will report on some recent attempts to characterize the notions of sensitive dependence and entropy as a measure of dynamical complexity in quantum mechanics.

I want to emphasize that this paper does not attempt to be a review of the whole field of quantization of non-integrable systems. The particular topics chosen reflect merely my personal interests. Important topics which are not touched upon are, for example, the discussion of the statistics of energy level separation, functional integral techniques and the semi-classical limit.

2. Quantum recurrence

In classical mechanics Poincaré's recurrence theorem states that, for a system confined to the support of a finite volume measure in phase space, almost any initial configuration will be repeated as accurately as one wishes infinitely many times in the future. For quantum systems discreteness of the energy spectrum plays for recurrence the same role as the finiteness of the invariant measure in classical mechanics. In fact quantum recurrence does not hold in general if the system has a continuous spectrum.

Theorem : (Quantum recurrence)[1-3]
Let a quantum system have a discrete energy spectrum and $\psi(t)$ and $\rho(t)$ be a wave function and a density matrix. Then, for an arbitrarily small positive ϵ, there are infinitely many τ and τ' such that

$$\|\psi(t+\tau)-\psi(t)\|<\epsilon \qquad \forall t$$

$$\|\rho(t+\tau')-\rho(t)\|<\epsilon \qquad \forall t$$

Furthermore the sets of all τ's and τ''s are relatively dense, i. e. there is a number L such that any interval of the real line of length L contains at least one element of the set.

Proof :

Expanding the wave function in normalized energy eigenstates

$$\psi(t)=\sum_{n=0}^{\infty} a_n \phi_n e^{-iE_n t}$$

one obtains

$$\|\psi(t+\tau)-\psi(t)\|^2=2\sum_{n=0}^{\infty}|a_n|^2(1-\cos E_n\tau)$$

Normalizability of the wave function implies that for sufficiently large N

$$\sum_{n=N}^{\infty}|a_n|^2(1-\cos E_n\tau) < \epsilon^2/2$$

The result now follows from the fact that there are infinitely many, relatively dense, τ's such that the remaining trigonometrical polynomial satisfies

$$\sum_{n=0}^{N-1}|a_n|^2(1-\cos E_n\tau) < \epsilon^2/2$$

a standard result in the theory of almost-periodic functions.

The proof for the density matrix is similar. The lenghts are defined from the scalar product $(\rho,\rho')=\mathrm{Tr}(\rho^{\dagger}\rho')$. In the energy representation

$$\rho(t)=\sum_{n,n'=0}^{\infty}\rho^{nn'}(0)\, e^{i(E_n-E_{n'})t}$$

with $\rho^{nn'}(t)=P_n\ \rho(t)\ P'_n$, P_n being the projection operator on the E_n energy eigenstate.
The finite sum approximations

$$\sigma^{NN'}(t)=\sum_{n=0}^{N}\sum_{n'=0}^{N'}\rho^{nn'}(t)$$

converge uniformly to $\rho(t)$ because

$$\left\|\rho(t)-\sigma^{NN'}(t)\right\|^2=\sum_{n=N+1}^{\infty}\sum_{n'=N+1}^{\infty}\left|\rho^{nn'}(0)\right|^2$$

and the length of the density matrix is not larger than one. The result then follows, as for the wave function, because $\sigma^{NN'}(t)$ is almost-periodic. □

The same argument applies to prove recurrence for the expectation values of bounded operators.

Hogg and Huberman[4] have generalized the proof of quantum recurrence for systems with Hamiltonian

$$H=H_0+V(t)$$

where H_0 has a discrete spectrum and $V(t)$ is a time-periodic non-resonant potential such that $\|\dot{V}\|$ is bounded.

These results have, for quantum systems, the same consequences as classical Poincaré recurrence in invalidating the unrestricted formulation of the Boltzmann H-theorem (i.e. the formulation without additional postulates of a statistical nature). However their significance concerning the absence of quantum chaos in quantum systems with point spectrum has been exaggerated. To show the absence of chaos (in the sense of sensitive dependence on initial conditions) in a particular quantum system, it is not enough to invoke the almost-periodicity of the wave function. What one would have to prove would be, for example, the near equality of recurrence times of the expectation values of bounded and unbounded operators for two wave functions that evolve from two nearby initial conditions.

3. Spectral properties and large time behaviour

As stated in the introduction, the stability problem in quantum mechanics is related to the question of the nature of the spectrum of the perturbed non-integrable Hamiltonian. Two classes of results are of interest, the first pertaining to the behaviour of the wave function and the second to the rate of growth of expectation values of operators.

A wave function $\psi_t=\exp(-itH)\psi_0$ is said to be localized for all time if for any $\epsilon>0$, there is $R>0$ such that

$$\sup_{t\in\mathbb{R}} \int_{|x|\geq R} d^N x\, |(e^{-iHt}\psi_0)(x)|^2<\epsilon \tag{3.1}$$

Conversely it is said to delocalize or decay locally if for any $R>0$

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T dt \int_{|x|\leq R} d^N x\, |(e^{-iHt}\psi_0)(x)|^2=0 \tag{3.2}$$

Ruelle[5], Amrein and Georgescu[6] have shown that for a large class of Hamiltonians the wave function is localized for all time if and only if ψ_0 belongs to the point spectral subspace of H and decays locally if and only if ψ_0 belongs to the continuous spectrum subspace. Furthermore, if the Hamiltonian has no singular continuous spectrum the time average may be omitted and the local decay condition reads

$$\lim_{t\to\infty} \int_{|x|\leq R} d^N x\, |(e^{-iHt}\psi_0)(x)|^2=0 \tag{3.3}$$

The Ruelle-Amrein-Georgescu theorem was generalized by Yajima[7], Enss and Veselic[8] also for a large class of time periodic Hamiltonians, $H(t+T)=H(t)$. For time periodic Hamiltonians the spectrum subspaces to be considered are those of the time evolution operator over one period $U(t+T,t)$, called the Floquet operator (see Sect. 4).

These results make rigorous the intuitive notion that the point and continuous subspaces correspond to bound and scattering states, respectively.

I will sketch the main ideas of the proof. Denote by F_r the projector onto the subspace of states localized in the sphere $|x|<r$, by U_t the time evolution operator and by M_0 and M_∞ the following two subsets of the Hilbert space

$$M_0=\{\phi\in\mathcal{H} : \lim_{r\to\infty} \sup_{t\in\mathbb{R}} \left\|(I-F_r)U_t\phi\right\|^2=0\} \tag{3.4}$$

$$M_\infty=\{\phi\in\mathcal{H} : \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\, \left\|F_r U_t\phi\right\|^2=0 \quad \forall r\} \tag{3.5}$$

Let $\mathcal{H}=\mathcal{H}_p\oplus\mathcal{H}_c$ be the decomposition into point and continuum spectral subspaces. A simple computation shows that M_0 and M_∞ are closed linear subspaces of $\mathcal{H}$ and M_0 is orthogonal to M_∞. If ϕ belongs to the point spectrum, $U_t\phi=e^{-iEt}\phi$. Therefore

$$\left\|(I-F_r)U_t\phi\right\|^2=\left\|(I-F_r)\phi\right\|^2$$

This obviously converges to zero when $r\to\infty$. Therefore $\mathcal{H}_p\subset M_0$. Using this inclusion and the orthogonality of M_0 and M_∞ one obtains

$$\mathcal{H}_c=\mathcal{H}_p^\perp\supset M_0^\perp\supset M_\infty$$

The inclusions $\mathcal{H}_p\subset M_0$ and $\mathcal{H}_c\supset M_\infty$ are thus seen to hold in general. The "if and only if" identification of localized and decaying states with elements of the point and continuum spectra requires $\mathcal{H}_p=M_0$ and $\mathcal{H}_c=M_\infty$. For this to hold, some restrictions have to be put on the evolution operator. A general characterization of these restrictions is provided by the following theorem (proved in Ref.6) :

Theorem: Let H be a self-adjoint operator and $\{F_r\}$ r=1,2,... a family of orthogonal projections such that s-$\lim_{r\to\infty} F_r=I$. Suppose there exists a family $\{S_n\}$ n=1,2,... of linear operators acting on $\mathcal{H}$ such that

(i) $\forall n$ S_n is a bounded operator commuting with H.

(ii) The sequence S_n converges strongly to an operator S of dense range in $\mathcal{H}$.

(iii) E_c being the projector on $\mathcal{H}_c$, $F_r S_n E_c$ is compact $\forall r,n$.

Then $\mathcal{H}_p=M_0$ and $\mathcal{H}_c=M_\infty$.

To prove the theorem one has to prove $\mathcal{H}_c\subset M_\infty$. Because M_∞ is closed it suffices to verify the inclusion for a set D dense in $\mathcal{H}_c$ and according to the hypothesis

on S, this may be the set $S\mathfrak{H}_c$. Then to prove

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\,\left|F_r U_t S\psi\right|^2=0$$

with $\psi\in\mathfrak{H}_c$, one approximates S by the sequence S_n, and because a compact operator is a limit of finite rank operators, the essential property to be used is the following characterization of the continuous spectrum subspace

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\,|(\phi,U_t\psi)|^2=0 \qquad (3.6)$$

for $\psi\in\mathfrak{H}_c$ $\phi\in\mathfrak{H}$.

To prove $\mathfrak{H}_p=M_0$ and $\mathfrak{H}_c=M_\infty$ for a particular self-adjoint Hamiltonian it suffices to find the appropriate sequence S_n and to prove compactness of $F_r S_n E_c$. For example with $S_n=S=(H+\mathrm{i})^{-1}$ a large class of Hamiltonians of the form $H=\phi(\mathrm{p})+V$, where $\phi(\mathrm{p})$ is a self-adjoint operator of the momentum operators, is proved to fulfil the conditions.

For time periodic systems the corresponding requirement is compactness of $F_r(K-z)^{-1}$, K being the quasi-energy operator (see Sect.4).

Another important question which is also related to the spectral nature of the states is the time asymptotic behaviour of matrix elements of unbounded operators. Let H be a Hamiltonian for a system defined on the n-dimensional torus T^n

$$H=H_0+V$$

where $H_0=H_0(-\mathrm{i}\frac{\partial}{\partial x_i})$ is a function of the momentum operators only. H_0 has a point spectrum, namely $\phi_m=\mathrm{e}^{\mathrm{i}m.x}$ $m\in\mathbb{Z}^n$ are the eigenstates of energy E_m. Under the time evolution induced by H_0 the kinetic energy $T=\sum_k(-\mathrm{i}\frac{\partial}{\partial x_k})^2$ remains bounded for all times. If the set $\{E_m\}$ is a discrete set and V a bounded perturbation, the spectrum of H is still pure point. If however H_0 has a dense point spectrum, the V perturbation may change the nature of the spectrum. This is important for the asymptotic behaviour of the expectation values because of the following result:

Theorem[9]: If ψ belongs to the continuous spectral subspace of H then $(\psi(t),T\psi(t))$ is not bounded in time.

Let $c_m(t)=(\phi_m,U_t\psi)$. Using the characterization (3.6) of the continuous spectral subspace,

$$\lim_{T\to\infty} a_m(T)=\lim_{T\to\infty}\frac{1}{T}\int_0^T|c_m(t)|^2\,\mathrm{d}t=0$$

Then, given $\epsilon>0$ and $M\in\mathbb{N}$, there is $T_{\epsilon,M}>0$ such that $T>T_{\epsilon,M}$ implies

$$\sum_{|m|\le M} a_m(T)\le\epsilon$$

Hence

$$\frac{1}{T}\int_0^T\mathrm{d}t\,(\psi_t,T\psi_t)=\sum_{m\in\mathbb{Z}^n} m^2\,a_m(T)\ge M^2\sum_{|m|>M} a_m(T)\ge M^2(1-\epsilon)$$

Therefore

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\mathrm{d}t\,(\psi_t,T\psi_t)=\infty$$

The conclusion is that for the kinetic energy to stay bounded in time it is necessary that the spectrum remain pure point. This is only a necessary condition, not sufficient in general, because T being an unbounded operator it is also necessary that ψ_t be at all times in the domain of T.

A similar result holds, when $V(t)$ is a time periodic function, for the continuous spectral subspace of the Floquet operator.

4. Time-periodic systems. Spectral properties of the Floquet operator.

For a time-dependent Hamiltonian $H(t)$ acting on a Hilbert space $\mathfrak{H}$, the solutions of the Schrödinger equation

$$H(t)\,\psi(t)=\mathrm{i}\hbar\frac{\partial\psi}{\partial t} \qquad (4.1)$$

may be expressed as

$$\psi(t)=U(t,s)\,\psi(s) \qquad (4.2)$$

the evolution operator $U(t,s)$ being the solution of the integral equation

$$U(t,s)=1-\frac{\mathrm{i}}{\hbar}\int_s^t\mathrm{d}\tau\,H(\tau)\,U(\tau,s) \qquad (4.3)$$

A classical time-dependent problem may be converted into a formally time-independent one by considering physical time as a coordinate conjugate to the energy. The new Hamiltonian is

$$K=H(t)+E \qquad (4.4)$$

From Hamilton's equation $\dot{t}=\frac{\partial t}{\partial \eta}=\frac{\partial K}{\partial E}=1$ it follows that the new, fictitious, time η differs from t only by the addition of a constant. This formalism has been applied by Howland[10] to quantum mechanics. With $E=-i\hbar\frac{\partial}{\partial t}$

$$K = H(t) - i\hbar\frac{\partial}{\partial t} \tag{4.5}$$

and

$$K\phi = i\hbar\frac{\partial\phi}{\partial\eta} \tag{4.6}$$

If the solutions of (4.1) belong to the Hilbert space $\mathfrak{H}$, the solutions of (4.6) then lie in the extended Hilbert space $\mathcal{L}^2(\mathbb{R},\mathfrak{H})$ of $\mathcal{L}^2$ functions in $\mathbb{R}$ with values in $\mathfrak{H}$. The operator $\exp(-\frac{i}{\hbar}\ \eta\ K)$ is related to the evolution operator of the time-dependent Schrödinger equation (4.1) by

$$(e^{-\frac{i}{\hbar}\eta K}\ \phi)(t) = U(t,t-\eta)\ \phi(t-\eta) \tag{4.7}$$

Let now $H(t)$ be periodic of period T, $H(t+T)=H(t)$. Then $U(t+T,s+T)=U(t,s)$ implies $U(t+NT,s)=U^N(t+T,t)\ U(t,s)$ and the problem is reduced to that of finding the time evolution within one period. The one-period evolution operator $U(t+T,t)=F_t$ is called the Floquet operator.

The spectral properties of the Floquet operator F_t do not depend on t, because of the unitary equivalence

$$F_s = U(s,t)\ F_t\ U^\dagger(s,t) \tag{4.8}$$

K commutes with the one-period translation operator and in the extended (Howland) formalism one may restrict oneself to the subspace $\overline{\mathfrak{H}} = \mathcal{L}^2([0,T],\mathfrak{H})$ of periodic functions with the boundary conditions $\phi(T)=\phi(0)$.
The restriction K_T of the K operator to this subspace is called the quasi-energy operator.

$U(t,0)$ and $e^{i\frac{t}{\hbar}K_T}$ are unitarily equivalent (Yajima[11]).
The spectrum of the Floquet operator (or equivalently the spectrum of the quasi-energy operator) is called the quasi-energy spectrum.

From the spectral decomposition of the evolution operator

$$U(t,s) = \int e^{-i\lambda(t-s)}\ dP_\lambda \qquad \lambda\in[0,2\pi] \tag{4.9}$$

one obtains the wave function $\psi(t)$ given any initial condition. Hence the quasi-energy spectrum contains all the information needed to solve the quantum problem.

The nature of the quasi-energy spectrum and in particular the question of the stability of the point spectrum have been studied in a number of cases. Below I list a selection of some of the results that have appeared in the literature.

a) Hamiltonians quadratic in p and q with time periodic coefficients

Because the Hamiltonian is quadratic the quantum evolution is simply a unitary implementation of the classical phase space flow. Hagedorn, Loss and Slawny[12] have proved that, if the coefficients are T-periodic and piecewise continuous, the Floquet operator $U(T,0)$ has either a pure point spectrum or a transient absolutely continuous spectrum. The pure point spectrum corresponds to the elliptic case in the classical motion and the transient absolutely continuous spectrum to the hyperbolic or parabolic cases.

A state belongs to the transient absolutely continuous spectrum if it is the limit of a sequence of vectors ϕ_n such that $|(\phi_n,U_t\phi_n)|$ decays faster than any inverse power of t[13].

b) Time periodic perturbations of the harmonic oscillator

Let $K=D+P$ where $D=-i\frac{\partial}{\partial t}+\frac{1}{2}(-\frac{\partial^2}{\partial x^2}+x^2)$ and P is a bounded operator in $\overline{\mathfrak{H}}$. To define a norm for P one uses the matrix elements in the basis of eigenvectors $\psi_{m,n}$ of D

$$\psi_{m,n}(x,t)=e^{im\Omega t}\otimes\phi_n(x),\quad \phi_n \text{ being the Hermite polynomials.}$$

Consider any $\gamma>0$ and a perturbation P such that $\|P\|<C\gamma^2$ with the norm and the constant C as defined in the conditions ($\mathcal{A}$) and ($\mathcal{C}$) of Ref.[14]. Then $D+P$ has a pure point spectrum for frequencies Ω away from a resonant set of Lebesgue measure less than γ.

c) The quantized Fermi accelerator

The Hamiltonian is $H=-\frac{\partial^2}{\partial x^2}$ defined on the time-dependent Hilbert space $\mathfrak{H}(t)=\mathcal{L}^2(0,l(t))$, $l(t)$ denoting the motion of the wall. It has been proved[15] that the spectrum of the Floquet operator is pure point for almost every wall oscillation, provided the functions $l(t)$ are five times differentiable.

d) Time periodic perturbations of Hamiltonians with increasing gaps between successive eigenvalues

Let H_0 have a discrete spectrum with eigenvalues of simple multiplicity satisfying

$$\Delta\lambda_n = \lambda_n - \lambda_{n-1} \geq cn^\alpha$$

with $\alpha > 0$. Let $V(t)$ be strongly C^r and

$$\int_0^{2\pi} V(t)\, dt = 0$$

Then if $r = 1 + \frac{1}{\alpha}$ the quasi-energy operator $K = -i\frac{\partial}{\partial t} + H + V(t)$ has no absolutely continuous spectrum[16].

e) The pulsed rotator[9]

$$H = -\alpha \frac{d^2}{dx^2} + \mu V(x,t) \qquad x \in \mathbb{R}/2\pi\mathbb{Z}$$

and $V(x,t)$ is a 2π-periodic regular function, admitting an analytical continuation to the strip $\{|\mathrm{Im}x| < r, |\mathrm{Im}t| < r\}$. Then $\forall \epsilon > 0$ there is a closed set $\Omega \subset [1,\infty)$ with the measure of $\bar{\Omega}$ less than ϵ and a constant $\mu_C > 0$ such that if $\alpha \in \Omega$ and $|\mu| < \mu_C$ the Floquet operator has a pure point spectrum.

f) The kicked rotator

$$H = -\alpha \frac{d^2}{dx^2} + V(x) \sum_{n \in \mathbb{Z}} \delta(t - 2\pi n) \qquad x \in \mathbb{R}/2\pi\mathbb{Z}$$

For the resonant case, i.e. if α is rational, it was proved[17] that for generic V in the space of continuous functions in the torus, the spectrum of the Floquet operator is continuous.

For the weakly resonant case, i.e when there is a sequence $\frac{p_n}{q_n}$ of rationals such that $|\alpha - \frac{p_n}{q_n}| < c_1 e^{-c_2 q_n}$, the result[18] is that for generic V the Floquet operator spectrum is also continuous.

For α an irrational diophantine number there is numerical evidence, from the localization of the wave functions, that the spectrum of the Floquet operator is pure point. This however, to my knowledge, has not yet been proven rigorously.

g) Rotators kicked by rank-one potentials

$$H = H_0 + \lambda |\phi\rangle\langle\phi| \sum_{n=-\infty}^{\infty} \delta(t - nT)$$

The main interest in using finite rank kicks as opposed to multiplicative potential kicks lies in that it seems easier to obtain non-perturbative results. Combescure[19] has proven that if the spectrum of H_0 is pure point and if ϕ is cyclic for H_0 with an absolutely convergent expansion in the basis of eigenstates of H_0

$$\sum_{-\infty}^{\infty} |a_n| < \infty \qquad \phi = \sum_{-\infty}^{\infty} a_n \phi_n$$

then the Floquet operator has a pure point spectrum for almost every λ.

If the absolute convergence condition is not verified, the Floquet operator may have a continuous spectrum. For example : let H_0 be the Hamiltonian of the harmonic oscillator, with frequency ω, and let $a_n = 2\pi n^{-\gamma}$ $\frac{1}{2} < \gamma < 1$. Then if $\omega T/2\pi$ is diophantine, ϕ belongs to the continuous spectrum of the Floquet operator for any λ.

h) Results also exist for time-dependent but non-periodic perturbations, either deterministic or random[12,20,21].

5. Sensitive dependence in quantum systems

Chaos in classical mechanics becomes a precise notion if one identifies it with the existence of at least one positive Lyapunov exponent. It is related to the physical notion of sensitive dependence on initial conditions. The positive Lyapunov exponent measures the local rate of separation of orbits which, at time zero, differ by a small vector along the unstable manifold. The positive Lyapunov exponent does not tell us that two orbits with initial conditions differing by a small vector along the unstable manifold will not eventually come close to each other at some later time. In fact they will, with probability one, for a bounded system. Therefore the Lyapunov exponent (and sensitive dependence) are not statements about a global property like recurrence. What the Lyapunov exponent is, in fact, is the average of a local property, i. e. the average local separation of orbits in the support of some measure.

The almost-periodicity of the wave functions (and density matrices) of bounded quantum systems is very suggestive of regular motion. There is however nothing mysterious about it. It is simply a consequence of the linearity of the Schrödinger equation. The non-linearity in the classical evolution is, in the quantum Schrödinger picture, traded off for the spectral problem of the Hamiltonian operator in the infinite-dimensional Hilbert space. Already if one uses the Heisenberg picture, the non-linearity becomes manifest in the time evolution of the operators.

It therefore clear that, to decide whether and how chaoticity (in the sense of

sensitive dependence) is modified or destroyed when one passes from classical to quantum mechanics, it is not enough to invoke almost-periodicity of the wave functions. Instead one should see whether (and when) quantum mechanical systems are sensitively dependent on initial conditions, as an averaged local property of separation of the dynamics, starting from two close initial wave functions $\psi(t_0,x)$ and $\psi(t_0,x)+\delta\psi(t_0,x)$.

By the Ehrenfest theorem one knows that sufficiently localized wave packets have to follow the classical trajectories, at least during some time. Therefore for a hyperbolic system one may find wave packets representing locally separating orbits. However if, at time t_0, one has

$$\|\psi+\delta\psi - \psi\|_{t_0} = \delta$$

then the unitarity of the evolution operator implies that this distance remains the same for all times. The norm distance of the evolving wave functions remains the same at all times no matter which deformation direction one takes. Still the wave packets must, at least during some time, be separating like the classical orbits. This only shows that statements based purely on the behaviour of the wave function risk to be trivial and to reflect merely the linear nature of the Schrödinger equation. Non-trivial statements have necessarily to involve expectation values of operators and to make full use of their spectral properties.

Let $\psi(t)$ be a pure state in the Hilbert space $\mathcal{H}$ and $\mathcal{A}$ a symmetric operator $\{(\mathcal{A}\phi,\psi)=(\phi,\mathcal{A}\psi)\}$. The tangent space to $\mathcal{H}$ at $\psi(t)$ being $\mathcal{H}$ itself, one considers norm-preserving infinitesimal deformations $\psi_\delta(0)=\psi(0)+\delta\psi(0)$

$$\delta(\psi,\psi)=0 \quad \Rightarrow \quad (\delta\psi,\psi)=0$$

i. e. the deformations are $\delta\psi(0)=\delta\cdot\phi(0)$ with $\phi(0)$ in the orthogonal complement of $\psi(0)$ in $\mathcal{H}$ (hence also $(\phi(t),\psi(t))=0$). The expectation values of the operator $\mathcal{A}$ are used to measure the separation of the orbits with nearby initial conditions. Let

$$\Delta(t)=(\psi_\delta(t),\mathcal{A}\psi_\delta(t))-(\psi(t),\mathcal{A}\psi(t))\simeq 2\delta.\ \mathrm{Re}(\phi(t),\mathcal{A}\psi(t))$$

The time derivative of $\ln|\Delta(t)|$ measures the local rate of separation of the orbits and the average value of this separation is obtained from a time average along the orbit. Hence, whenever the following limit exists

$$\lim_{T\to\infty}\ \frac{1}{T}\int_0^T \frac{\partial}{\partial t}\ln|\mathrm{Re}(\phi(t),\mathcal{A}\psi(t))|\ \mathrm{d}t = \lim_{T\to\infty}\ \frac{1}{T}\ \ln|\mathrm{Re}(\phi(T),\mathcal{A}\psi(T))| = \lambda_{\mathcal{A},\phi} \tag{5.1}$$

($\phi(t)=e^{-iHt}\phi(0)\perp\psi(t)$), it defines a **quantum characteristic exponent at ψ along the ϕ-direction**. A positive $\lambda_{\mathcal{A},\phi}$ would require exponential growth of $|\mathrm{Re}(\phi(t),\mathcal{A}\psi(t))|$. Anticipating eventual rates of growth different from exponential in quantum mechanics we define

Def.: There is **$\mathcal{A}$–sensitive dependence at the state ψ** if there is at least a $\phi\perp\psi$ such that for any T and $M>0$ there is a $t>T$ implying $\left\|\frac{\Delta(t)}{\Delta(0)}\right\|>M$.

The above definition allows for oscillations of the ratio; it only requires that $\left\{\left\|\frac{\Delta(t)}{\Delta(0)}\right\|;\ t>T\right\}$ be unbounded. Also, for the moment, I will not be concerned with measure considerations in Hilbert space and consider only the behaviour on the orbit $\psi(t)$.

 # If $\mathcal{A}$ is defined everywhere in $\mathcal{H}$, $\mathcal{A}$–sensitive dependence cannot take place.

Proof : By the Hellinger-Toeplitz theorem a symmetric, everywhere-defined operator is necessarily bounded. If $\mathcal{A}$ is bounded, $\exists\ c$ such that

$$\|\mathcal{A}\psi\|<c\|\psi\| \qquad \forall\psi\in\mathcal{H}$$

and by Schwartz's inequality $\exists M$ such that

$$|(\phi(t),\mathcal{A}\psi(t))|<M$$

□

Henceforth I will assume that $\mathcal{A}$ is unbounded with dense domain $\mathcal{D}(\mathcal{A})$ in $\mathcal{H}$.

Sensitive dependence is a notion distinct from unboundedness of expectation values. Let $\psi(t)$ and $\phi(t)$ be states in the continuum spectrum component of a Hamiltonian H. Let $\{x_n\}$ be a basis in the continuum spectral subspace and $\mathcal{A}$ an operator such that

$$(x_{n'},\mathcal{A}x_n) = \mathrm{f(n)}\ \delta_{n,n'} \quad \text{with } \mathrm{f'(n)}>0 \text{ and } \mathrm{f(n)} \text{ unbounded}$$

From $\phi(t)=\sum_n a_n(t)\ x_n$ and $\psi(t)=\sum_n b_n(t)\ x_n$, using the RAGE-theorem[22], it follows that $\forall\epsilon>0$ and N $\exists T$ such that

$$\frac{1}{T}\int_0^T \sum_{n\le N} |\ a_n(T)|^2\ \mathrm{d}t \le \epsilon$$

Then, as in Sect.3, it is easy to show that $\forall\epsilon>0$ and $M>0$ there is $T_{\epsilon,M}$ such that for $T>T_{\epsilon,M}$

$$\frac{1}{T}\int_0^T (\phi(t),\mathcal{A}\phi(t))\,dt \geq M(1-\epsilon)$$

$$\frac{1}{T}\int_0^T (\psi(t),\mathcal{A}\psi(t))\,dt \geq M(1-\epsilon)$$

implying that the matrix elements $(\phi(t),\mathcal{A}\phi(t))$ and $(\psi(t),\mathcal{A}\psi(t))$ are unbounded in time.

For the matrix elements involved in sensitive dependence

$$(\phi(t),\mathcal{A}\psi(t)) = \sum_n a_n^*(t)\, b_n(t)\, f(n) \qquad (5.2)$$

because $\phi\perp\psi \quad \sum_n a_n^*(t)\, b_n(t) = 0$ and any bound on the sum of a finite number of terms implies the same bound on the remaining terms. Whether the matrix element (5.2) is bounded or unbounded depends on the detailed nature of the time evolution. The argument leading to unboundedness of expectation values does not imply unboundedness of the matrix elements of sensitive dependence. This suggests that these two phenomena might not always occur simultaneously.

The point spectrum of bounded quantum systems implies recurrence of the wave function. For sensitive dependence one has the result

\# **If the energy has a point spectrum without accumulation points nor infinite degeneracies, there is no sensitive dependence for any operator $\mathcal{A}$ when restricted to a bounded energy subspace.**

Proof : By bounded energy subspace of $\mathcal{H}$ we mean a subspace where vectors have components of energy less than K only. Let

$$\phi(t)=\sum_n a_n\phi_n e^{-iE_nt} \text{ and } \quad \psi(t)=\sum_n b_n\phi_n e^{-iE_nt}$$

with $E_n<K$. Then in

$$\mathrm{Re}(\phi(t),\mathcal{A}\psi(t))=\mathrm{Re}\sum_{n,n'} a_{n'}^* b_n\,(\phi_{n'},\mathcal{A}\phi_n)\; e^{-i(E_n-E_{n'})\,t} \qquad (5.3)$$

if there are no accumulation points nor infinite degeneracies of the spectrum in the range $[0,K)$, the sum on the r. h. s. has only finitely many terms and, being a trigonometrical polynomial, is bounded for all t. □

That a point spectrum does not, by itself, exclude sensitive dependence is seen from an example. Let $\phi(0)=\sum_n a_n\,\chi_n$, $\psi(0)=\sum_n b_n\,\chi_n$

$$(\chi_k,\mathcal{A}\,\chi_n)=(-1)^n\,\delta_{k,n+1}+(-1)^{n-1}(n-1)\,\delta_{k,n-1}$$

and

$$a_n=\frac{2\sqrt{6}}{\pi n} \;(n \text{ even}) \qquad a_n=0 \;(n \text{ odd})$$

$$b_n=\frac{2\sqrt{2}}{\pi n} \;(n \text{ odd}) \qquad b_n=0 \;(n \text{ even})$$

Then

$$(\phi(t),\mathcal{A}\psi(t))=\frac{8\sqrt{3}}{\pi^2}\sum_{n\atop \text{even}}\frac{1}{n}\left\{-e^{i(E_n-E_{n-1})\,t}+\frac{n}{n+1}\,e^{-i(E_{n+1}-E_n)\,t}\right\} \qquad (5.4)$$

If, for example

$$E_n=\frac{1}{2(n+1)} \;(n \text{ even}) \qquad E_n=\frac{1}{2n} \;(n \text{ odd})$$

the term in curly brackets on the r. h. s. of (5.4) becomes $-\exp\{i(E_n-E_{n-1})t\}+\frac{n}{n+1}$ (n even). For $t=\pi N!!$ all (E_n-E_{n-1}) up to order $N-1$ are odd multiples of π. Hence the sum in (5.4) reaches arbitrarily high values for sufficiently large t. Notice that the sum of the remaining terms cannot cancel this growth because for large n it has a positive real part. Similar examples may easily be constructed involving, instead of an accumulation point, a sequence of arbitrarily large energy components.

Bellissard[9] had already pointed out that a pure point spectrum is not sufficient for boundedness in time of expectation values if some of the eigenstates of H are not in the domain of the operator $\mathcal{A}$. In this case however, one even notices that all eigenstates of H as well as $\psi(0)$ and $\phi(0)$ are in the domain of $\mathcal{A}$. Subtler domain questions must however be involved because under time evolution the norm $\|\mathcal{A}\psi(t)\|$ is also reaching successively higher values, approaching in some sense the boundary of the domain of $\mathcal{A}$. Notice also that $(\phi(t),\mathcal{A}\phi(t))=(\psi(t),\mathcal{A}\psi(t))=0 \;\; \forall t$, showing that in this case sensitive dependence and unbounded growth of expectation values are different phenomena.

Notice the difference in the behaviour of $\mathrm{Re}(\phi(t),\mathcal{A}\psi(t))$ and, for example, the quantity $\|\psi(t)-\psi(0)\|$ that is used in the proof of recurrence of the wave function

$$\|\psi(t)-\psi(0)\|=2\sum_n |b_n|^2(1-\cos E_nt)$$

Because of the normalizability of the wave function, this norm difference is approximated to arbitrary accuracy by a trigonometric polynomial, which is bounded and almost-periodic. In $\mathrm{Re}(\phi(0),\mathcal{A}\psi(0))$, however, the series need not be absolutely convergent. In fact, absolute convergence of the series would be a sufficient condition

for boundedness.

If $S=\sum_{n,n'=1}^{\infty}\left|a^*_{n'}b_n\ (\phi_{n'},\mathcal{A}\phi_n)\right|$ **is finite then** $(\phi(t),\mathcal{A}\psi(t))$ is bounded in time.

S finite $\Rightarrow \forall\epsilon\ \exists N$ such that $\sum_{n,n'>N}^{\infty}\left|a^*_{n'}b_n\ (\phi_{n'},\mathcal{A}\phi_n)\right|<\epsilon$. Then

$\sum_{n,n'=1}^{N} a^*_{n'}b_n\ (\phi_{n'},\mathcal{A}\phi_n)$ is a (bounded) trigonometric polynomial that approximates $(\phi(t),\mathcal{A}\psi(t))$. □

The point nature of the spectrum, by itself, does not preclude the occurrence of quantum sensitive dependence. Notice however that the non-trivial cases necessarily involve states with arbitrarily many energy components near an accumulation point or components of arbitrarily large energy. Numerical simulation of quantum systems tend to take into account only finitely many energy components and non-trivial quantum behaviour risks to be excluded, not because of the quantum dynamics but because of the choice of initial conditions[23]. The critical nature of the initial conditions (i. e. the state $\psi(0)$ defining the orbit where sensitive dependence is studied) is apparent in the example. Indeed, if $\psi(0)$ is an arbitrary state with a finite number of energy components, then no matter how these components are chosen, $\mathrm{Re}(\phi(t),\mathcal{A}\psi(t))$ is always a trigonometric polynomial and there is no $\mathcal{A}$−sensitive dependence. The sequence $\psi_p(t)=\sum_{n=1}^{p} b_n\phi_n e^{-iE_nt}$ with $b_n\propto\frac{1}{n}$ is a sequence of vectors without $\mathcal{A}$−sensitive dependence, which approaches an $\mathcal{A}$−sensitive limit point.

The phenomenon of quantum localization is related to situations where the perturbed time evolution operator has a point spectrum. However, in the literature related to the numerical simulation of quantum systems, it is pointed out as a further symptom of the stability of quantum mechanics that, even in the case of a continuum spectrum, the reversibility of time evolution is not affected by the computer round-off errors. This in contrast to the situation in the numerical simulation of classical chaotic systems. This difference in behaviour might have a simple interpretation in terms of the nature of the quantum sensitive dependence. In the classical orbits the round-off errors essentially explore the whole of the (finite dimensional) tangent space to the orbit. In quantum mechanics however the tangent space is infinite-dimensional and the computer errors, referring to errors in a finite number of components, only explore a

finite dimensional subspace of the infinite dimensional tangent space. Furthermore it seems to be usually the case that the unstable directions are associated to vectors with an infinite number of components in the unperturbed basis. Then, any error on finitely many components is simply a deformation along a direction that is not sensitively dependent and reversibility will not be affected.

I will end up with a simple example of a quantum rotator kicked by a sequence of rank-one potentials. Rank-one kicks, as compared to multiplicative potential kicks, are easier to handle analytically, allowing in particular stronger non-perturbative results[19].

Consider the following time-dependent Hamiltonian

$$H(t)=-\frac{\alpha}{2\pi}\frac{\partial^2}{\partial x^2}+\lambda\sum_n|\phi_n><\phi_n|\,\delta(t-n) \tag{5.5}$$

defined on the circle $x\in[0,1)$, where $|\phi_n>$ is a sequence of Hilbert space vectors.

If all ϕ_n's are the same this would be a time-periodic Hamiltonian. Hamiltonians of the form (5.5) seem to be good testing grounds to study quantum sensitive behaviour phenomena. Here I will only illustrate its usefulness in an extremely simple case. I will be particularly concerned with the sequence

$$|\phi_n>=\gamma_n\sum_{k=0}^{n}\frac{1}{2k+1}\,|\,2\ k> \tag{5.6}$$

where γ_n is a normalization factor and $|k>=e^{i2\pi kx}$ an eigenvector of the unperturbed Hamiltonian $H_0(t)=-\frac{\alpha}{2\pi}\frac{\partial^2}{\partial x^2}$.

The operator of time evolution from "just after the $(n-1)$ th kick" to "just after the n th kick" is

$$U_n(\lambda)=e^{-i\lambda|\phi_n><\phi_n|}\ e^{i\frac{\alpha}{2\pi}\frac{\partial^2}{\partial x^2}}=(1+\mu\,|\phi_n><\phi_n|)\ e^{i\frac{\alpha}{2\pi}\frac{\partial^2}{\partial x^2}} \tag{5.7}$$

$\mu=e^{-i\lambda}-1$

Consider the (resonant) case $\alpha=1$. The free evolution operator $U_0(0)=\exp\{i\frac{\alpha}{2\pi}\frac{\partial^2}{\partial x^2}\}$ has an infinitely degenerate point spectrum, all vectors in the Hilbert space belonging to the eigenvalue one. The eigenvalue problem for each of the $U_n(\lambda)$ operators is extremely simple. The spectrum remains pointlike, with the Hilbert space splitting into two components. The vector $|\phi_n>$ belongs to the eigenvalue $e^{-i\lambda}$, whereas all vectors

in the orthogonal complement to $|\phi_n>$ have eigenvalue one. This simple structure allows a trivial computation of the time evolution. Let for example

$$\psi(0) = c \sum_{k=1}^{\infty} \frac{1}{k^2} \; |\, 2\,k+1> \tag{5.8}$$

Because $\psi(0)$ is orthogonal to all $|\phi_n>$'s , $\psi(t)=\psi(0)$ for $t\in\mathbb{Z}$. Let $\mathcal{A}=H_0$. $\psi(0)$ is not in the domain of $\mathcal{A}$ but in a rigged Hilbert space sense the matrix elements $<\psi(t)\ \mathcal{A}\ \psi(t)>$ are well defined, finite and constant for all $t\in\mathbb{Z}$.

For the matrix elements $<\phi(t)\ \mathcal{A}\ \psi(t)>$ with $\phi(t)\perp\psi(t)$, if $\phi(0)$ is chosen also in the subspace orthogonal to all $|\phi_n>$'s , the matrix element is fixed for all times. If however $\phi(0)$ has a non-zero overlap with one $|\phi_n>$, the matrix element $<\phi(t)\ \mathcal{A}\ \psi(t)>$ grows unboundedly in time. This example may be readily extended to the study of other more complex (non-resonant) situations.

6. Entropy

In classical dynamical systems the Kolmogorov[25]-Sinai[26] entropy is an important indicator of orbit complexity. Through Pesin's theory it is closely related to the distribution of positive Lyapunov exponents, at least when the invariant measure is absolutely continuous with respect to Lebesgue measure[27].

Several attempts have been made to carry this notion to quantum systems[28-30]. In the definition of the Kolmogorov-Sinai entropy a finite partition of a measure space plays an essential role. This is generalized to the quantum case by considering a finite dimensional abelian subalgebra of a von Neumann algebra or an abelian sublattice in the lattice of projections. I.e. an abelian model of the system is constructed, the entropy being then defined as a supremum over all abelian models. First I will sketch a Hilbert space version of Srinivas[28] construction, which is directly inspired by the Kolmogorov definition. Then one notices that the resulting quantity is not purely an indicator of the dynamical diversity in the neighbourhood of an orbit, but it involves in an essential way the reduction features of the quantum measurement process. Using a cylinder measure centred on a pure state, a new quantity is defined in the spirit of the topological Brin-Katok[31] definition.

In the Kolmogorov construction one considers a partition $\mathcal{P}$ and its refinement $\bigvee_n \mathcal{P}$, the elements of $\bigvee_n \mathcal{P}$ being the non-empty intersections

$$\left\{a_0 \cap T^{-1} a_1 \cap \cdots \cap T^{-n} a_n\right\} \qquad a_i \in \mathcal{P}$$

If μ is the invariant measure, the entropy of the partition is

$$h(T,\mathcal{P}) = -\lim_{n\to\infty} \frac{1}{n} \sum_{b\in \bigvee_n \mathcal{P}} \mu(b) \log \mu(b) \tag{6.1}$$

The entropy of the dynamics T is the supremum of $h(T,\mathcal{P})$ over all partitions. If at time zero a particular point in phase space belongs to the element $b=a_0\cap T^{-1}a_1\cap\cdots\cap T^{-n}a_n$ of the partition $\bigvee_n \mathcal{P}$, it means that a positive answer is assigned to the proposition *"the orbit at time k is in the element a_k of $\mathcal{P}$"*.

In Hilbert space quantum mechanics the set of propositions is the orthocomplemented lattice $\mathcal{L}$ of projection operators. The interpretation of the elements of $\bigvee_n \mathcal{P}$ as propositions suggests that partitions in the space of quantum events be identified with Boolean sublattices of $\mathcal{L}$. The physical meaning of the partition is provided by the spectral theorem, which associates with each self-adjoint operator $\mathcal{A}$ (observable) a spectral measure, i. e. a function from the Borel sets Δ_i on $\mathbb{R}$ to the lattice of projections $\Delta_i \to E(\Delta_i)\doteq E_i\in\mathcal{L}$. $E(\Delta_i)$ represents the measurement that selects systems with values of the observable $\mathcal{A}$ in the range Δ_i.

Because the sublattice so defined is abelian and $E(\cup_i \Delta_i)=\sum_i E_i$ for disjoint Δ_i's, each state yields a measure on the partition, $E_i\to\mu(E_i)$, which by Gleason's theorem is defined by a density operator ρ_μ

$$\mu(E_i)=\mathrm{Tr}(\rho_\mu\, E_i)$$

If U_t is the unitary operator of time evolution, the projections evolve as

$$E_i^{(t)} = U_t\, E_i\, U_t^{-1}$$

The elements of the refinement are

$$\epsilon_{\{i_n\cdots i_0\}} = E^{(n)}(\Delta_{i_n}) \cdots E^{(1)}(\Delta_{i_1})\ E(\Delta_{i_0}) \tag{6.2}$$

$\epsilon_{\{i_n\cdots i_0\}}$ corresponds to the experiment $E(\Delta_{i_0})$ at time zero followed by $E^{(1)}(\Delta_{i_1})$ at time one, etc. However $\epsilon_{\{i_n\cdots i_0\}}$ is a projection if and only if the projections $E^{(n)}(\Delta_{i_n})$ commute.

The measure in the refinement is

$$\mu(\epsilon_{\{i_n\cdots i_0\}}) = \mathrm{Tr}\left(\rho_\mu\ \epsilon^*_{\{i_n\cdots i_0\}}\ \epsilon_{\{i_n\cdots i_0\}}\right) \tag{6.3}$$

and the straightforward quantum generalization of Kolmogorov's partition entropy

becomes

$$h(\mu,\mathcal{A})=-\lim_{n\to\infty}\frac{1}{n}\sum_{\epsilon_{\{i_n\cdots i_0\}}}\mu(\epsilon_{\{i_n\cdots i_0\}})\ \log\mu(\epsilon_{\{i_n\cdots i_0\}}) \tag{6.4}$$

For the entropy $h(\mu)$ one would take the supremum over the partitions induced by all self-adjoint operators in $\mathcal{H}$.

At this point one notices however that the partition refinement as defined in (6.2) plays a role very different from $\bigvee_n \mathcal{P}$ in the classical mechanics case. In the classical case the assignment of the trajectory at a certain time, to an element of the phase-space partition, does not change the orbit. Therefore the rate of refinement of the partition does not depend on the measurement process, but only on the diversity of orbit behaviour. In contrast, the quantum mechanical projections in (6.2) do change the dynamical state and in the end the entropy will depend as much on the non-commutativity of $\mathcal{A}$ with the time evolution U_t as on the state that defines the measure μ.

To define a quantity with a meaning closer to the classical one, we now use a construction in the spirit of the Brin-Katok[31] definition. Let ψ be a pure state and suppose we want to characterize the diversity of dynamical behaviour in a neighbourhood of ψ. Let $\mathcal{A}$ be a self-adjoint operator that as in Sect.5 one uses to characterize orbit separation. For the measure one chooses a Gaussian cylindrical measure centred on ψ, with covariance defined by a strictly positive operator $\mathcal{B}$ commuting with time evolution.

Let $e_i(t)=U_t e_i(0)$ be an orthonormal basis in an n-dimensional subspace contained in the orthogonal complement of $\psi(t)$ in $\mathcal{H}$. The Gaussian measure is characterized by the following finite dimensional (cylinder) densities.

$$d\mu(\xi_1\cdots\xi_n)=\frac{1}{(2\pi)^{n/2}}\exp\Big\{-\tfrac{1}{2}\sum_{ij}x_i^*\,(\Lambda^{-1})_{ij}\,x_j\Big\}\ dx_1\cdots dx_n \tag{6.5}$$

with $\xi_i=\psi+x_i\,e_i$ and $(\Lambda)_{ij}=(e_i\,,\,\mathcal{B}\,e_j)$. The μ so defined is finitely additive, but σ-additive if and only if $\mathcal{B}$ is trace-class.

As in Sect.5 $\mathcal{A}$-separation between nearby normalized states ψ and $\psi+\delta\phi$ is $|2\ \mathrm{Re}(\delta\phi,\mathcal{A}\psi)|$ with $\delta\phi\perp\psi$. For each finite dimensional subspace V_n in the orthogonal complement of ψ in $\mathcal{H}$, we define the (ϵ,T) n-ball around ψ as the set

$$B_\epsilon^{(n)}(T,\psi)=\Big\{\delta\phi\,:\,|2\ \mathrm{Re}(\delta\phi(t),\mathcal{A}\psi(t))|\le\epsilon\ \ 0\le t\le T\ ,\ \delta\phi\in V_n\subset\mathcal{H}\ ,\ \delta\phi\perp\psi\Big\} \tag{6.6}$$

with $\delta\phi(t)=U_t\delta\phi$ and $\psi(t)=U_t\psi$.

The measure associated to the (ϵ,T) n-ball is

$$\mu(B_\epsilon^{(n)}(T,\psi))=(2\pi)^{-\frac{n}{2}}\int_{\substack{|2\ \mathrm{Re}(\delta\phi(t),\mathcal{A}\psi(t))|\le\epsilon\\ 0\le t\le T}} dx_1\cdots dx_n\ e^{-\frac{1}{2}\sum_{ij}x_i(\Lambda^{-1})_{ij}\,x_j} \tag{6.7}$$

Provided the limits exist, the quantities

$$h^{(n)}(\psi,\epsilon)=-\lim_{T\to\infty}\frac{1}{T}\ \log\mu(B_\epsilon^{(n)}(T,\psi)) \tag{6.8}$$

$$h^{(n)}(\psi)=\lim_{\epsilon\to 0}h^{(n)}(\psi,\epsilon) \tag{6.9}$$

are cylinder set quantum generalizations of the Brin-Katok entropy. They characterize the complexity of the orbit structures in a neighbourhood of ψ, with the time-zero deformations restricted to a finite dimensional subspace. Considering a sequence $V_1\subset V_2\subset V_3\ \cdots$ of successively larger finite dimensional subspaces, the construction of a global notion of entropy will depend on the existence of the corresponding inductive limit. No general statement can be made about the existence of these limits without specifying the nature of $\mathcal{A}$ and U_t.

The quantities $h^{(n)}(\psi,\epsilon)$ and $h^{(n)}(\psi)$ are sensitive to the diversity of dynamical evolution in the neighbourhood of ψ, and not to the state reduction nature of the measurement process, as $h(\mu,\mathcal{A})$ in Eq.(6.4). That these quantities are indeed closer to the corresponding classical notions is seen by the following simple estimate. The integration limits in (6.7) are defined by

$$x_i\le\frac{\epsilon}{2}\ \frac{1}{\sup\limits_{0\le t\le T}|\ \mathrm{Re}(e_i(t),\mathcal{A}\psi(t))|}$$

For a strictly positive covariance operator $\mathcal{B}$, the Gaussian measure becomes flat in the $\epsilon\to 0$ limit and

$$\mu(B_\epsilon^{(n)}(T,\psi)) \simeq \prod_{i=1}^{n} \frac{\epsilon}{\sup_{0\le t\le T} |\,\mathrm{Re}(e_i(t),\mathcal{A}\psi(t))|}$$

Then

$$h^{(n)}(\psi,\epsilon)=h^{(n)}(\psi) = \lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{n} \log \sup_{0\le t\le T} |\,\mathrm{Re}(e_i(t),\mathcal{A}\psi(t))|$$

The only terms contributing to the limit are those where $|\,\mathrm{Re}(e_i(t),\mathcal{A}\psi(t))|$ is exponentially unbounded in t. Comparing with Eq.(5.1) we conclude that $h^{(n)}(\psi,\epsilon)$ becomes the sum of the positive exponents associated to directions contained in V_n. This would be a quantum analogue of the Pesin formula. Notice however that it refers only to finite dimensional restrictions of the entropy and the exponents.

References and footnotes

[1] P. Bocchieri and A. Loinger, *Phys. Rev.* 107 (1957) 337.

[2] I. C. Percival, *J. Math. Phys.* 2 (1961) 235.

[3]N. S. Krylov, *"Works on the Foundations of Statistical Physics"*, Princeton Univ. Press, Princeton, New Jersey 1979.

[4] T. Hogg and B. A. Huberman, *Phys. Rev. Lett.* 48 (1982) 711 ; *Phys. Rev.* A28 (1983) 22.

[5] D. Ruelle, *Nuovo Cimento* 61A (1969) 655.

[6] W. O. Amrein and V. Georgescu, *Helv. Phys. Acta* 46 (1973) 635.

[7] K. Yajima and H. Kitada, *Ann. Inst. Henri Poincaré* A39 (1983) 145.

[8] V. Enss and K. Veselic, *Ann. Inst. Henri Poincaré* A39 (1983) 159.

[9] J. Bellissard, in *"Trends and Developments in the Eighties"*, S. Albeverio and P. Blanchard, Eds. ; World Scientific, Singapore, 1985.

[10] J. Howland, *Indiana Math. Jour.* 28 (1979) 471 ; *Ann. Inst. Henri Poincaré* A50 (1989) 309 and 325.

[11] K. Yajima, *J. Math. Soc. Japan* 29 (1977) 729.

[12] G. A. Hagedorn, M. Loss and J. Slawny, *J. Phys.* A19 (1986) 521.

[13] J. E. Avron and B. Simon, *J. Funct. Anal.* 43 (1981) 1.

[14] M. Combescure, *Ann. Inst. Henri Poincaré* 47 (1987) 63,451.

[15] P. Seba, Phys. Rev. A41 (1990) 2306.

[16] J. S. Howland, *Ann. Inst. Henri Poincaré* 49 (1989) 309,325.

[17] F. M. Izrailev and D. L. Shepelyanski, *Theor. Math. Phys.* 43 (1980) 553.

[18] G. Casati and I. Guarneri, *Commun. Math. Phys.* 95 (1984) 121.

[19] M. Combescure, *J. Stat. Phys.* 59 (1990) 679.

[20] L. Bunimovich, H. R. Jauslin, J. L. Lebowitz, A. Pellegrinotti and P. Nielaba, *"Diffusive energy growth in classical and quantum driven oscillators"*, Rutgers Univ. preprint (1990).

[21] M. Combescure, *"Recurrent versus diffusive dynamics for a kicked quantum system"*, Orsay preprint LPTHE 90/18.

[22] cf. M. Reed and B. Simon, *"Methods of Modern Mathematical Physics"*, vol. III, Academic Press, 1978.

[23] A remark of a related nature was made by Karner[24] for the Dirichlet quantum problem associated to the Fermi accelerator model.

[24] G. Karner, *Lett. Math. Phys.* 17 (1989) 329.

[25] A. N. Kolmogorov, *Dokl. Akad. Nauk. S.S.S.R.* 119 (1958) 861.

[26] J. Sinai, *Dokl. Akad. Nauk. S.S.S.R.* 124 (1959) 768.

[27]See for example R. Mañé, *"Ergodic Theory and Differentiable Dynamics"* ; Springer, Berlin, 1987.

[28] M. D. Srinivas, *J. Math. Phys.* 19 (1978) 1952.

[29] A. Connes, H. Narnhofer and W. Thirring, *Commun. Math. Phys.* 112 (1987) 691.

[30]H. Araki, *"Proceedings of VIII Int. Congress on Mathematical Physics"*, page 354, M. Mebkhout and R. Sénéor, Eds., World Scientific, Singapore, 1987 and references therein.

[31]M. Brin and A. Katok, in *"Geometrical Dynamics"*, Lecture Notes in Mathematics 1007, 30 , Springer , New York, 1983.

THE SECOND LAW OF THERMODYNAMICS: COMMENTS FROM ERGODIC THEORY.

Michael C. Mackey
Departments of Physiology and Physics
Centre for Nonlinear Dynamics
McGill University
Montréal, Québec
CANADA H3G 1Y6.

ABSTRACT. Recent results from ergodic theory are used to examine the possible dynamical foundations of nonequilibrium thermodynamics. Though ergodicity is necessary and sufficient to establish a unique state of thermodynamic equilibrium, much stronger dynamical properties (asymptotic periodicity and exactness) are needed to ensure that system entropy will change from its initial level. Asymptotic periodicity allows system entropy to evolve to a relative state of thermodynamic equilibrium in which the final entropy depends on system preparation. Exactness, a property that only irreversible (noninvertible) systems may have, is both necessary and sufficient for entropy to go to zero. Since all physical laws are formulated in terms of reversible dynamics, these results present a clear problem. Coarse graining, traces, factors, and perturbations are examined as possible sources of the experimentally observed behaviour of entropy.

0. INTRODUCTION.

The mathematical discipline of ergodic theory developed in response to mathematical questions raised by the seminal work of Boltzmann[1] and Gibbs[11] at the turn of the century. Though the first few decades of this century saw extensive interaction between ergodic theorists and statistical mechanicians, the fields have diverged somewhat in recent years leaving a number of important physical questions unanswered. This paper outlines a reconvergence of these two fields. A more extensive discussion of the issues raised here and examples of various points, as well as a proof of all results, can be found elsewhere[22,25,25a].

The central goal of this paper is to examine the primary issues that any successful statistical mechanics must address in deriving an understanding of the origin of the Second Law of thermodynamics. The Second Law of thermodynamics comes in so many forms that it is often confusing to understand precisely what a given author understands by the use of this term. To make matters explicit, we distinguish four versions. Let $S_{TD}(t)$ denote the thermodynamic entropy at time t.

The weakest form of the Second Law is the

0^{th} ORDER SECOND LAW. *$S_{TD}(t) = S_{TD}(t')$ for all times t, t', so the entropy difference $\Delta S = S_{TD}(t') - S_{TD}(t)$ satisfies $\Delta S \equiv 0$.*

In this form, the system entropy remains forever fixed at the value with which the system is prepared, be it by Nature or by an investigator.

The next strongest form is called the

1^{st} ORDER SECOND LAW. *$S_{TD}(t) \geq S_{TD}(t')$ for all times $t > t'$ Thus, with this form the system entropy may increase and $\Delta S \leq 0$.*

Following the 1^{st} order form we have the stronger assertion

2^{nd} ORDER SECOND LAW. *$S_{TD}(t) \geq S_{TD}(t')$ for all times $t > t'$ and at least one $\lim_{t\to+\infty} S_{TD}(t) = S^*_{TD}$ exists. Thus $\Delta S(t) = S_{TD}(t) - S^*_{TD} \leq 0$ and $\lim_{t\to+\infty} \Delta S(t) \leq 0$.*

In this case system entropy converges to a steady state value S^*_{TD} which may not be unique. If it is not unique it characterizes a metastable state.

The final, and strongest, form of the Second Law of thermodynamics is the most interesting.

3^{rd} ORDER SECOND LAW. *$S_{TD}(t) \geq S_{TD}(t')$ for all times $t > t'$ and there is a unique $\lim_{t\to+\infty} S_{TD} = S^*_{TD}$ for all initial system preparations. Under these circumstances, $\lim_{t\to+\infty} \Delta S(t) = 0$.*

In this case we know that the system entropy evolves to a unique maximum value irrespective of the way in which the system was prepared.

1. PRELIMINARIES.

In this section we introduce some basic concepts.

1.1. Thermodynamic Systems and Measure Spaces.

We first start with a set X which is going to be the **phase space** on which all of our dynamics operates. Whatever X is we are going to assume that it does not have any pathological properties. We let $\mathcal{A}$ denote a σ-**algebra** on X, and μ be a **measure** defined on the σ-algebra $\mathcal{A}$.

With the three concepts of a phase space X, a σ-algebra $\mathcal{A}$, and a measure μ we call the triple $(X, \mathcal{A}, \mu)$ a **measure space**. All of the measure spaces we consider will be σ-finite, and we associate a thermodynamic system with a σ-finite measure space through the following postulate.

POSTULATE A. A thermodynamic system is equivalent to a measure space.

1.2. Dynamics.

Consider a general thermodynamic system operating in a phase space X. On this phase space the temporal evolution of our system is described by a **dynamical law** S_t that maps points in the phase space X into new points, *i.e.*, $S_t : X \rightarrow X$, as time t changes. In general X may be a d-dimensional phase space, either finite or not, and therefore x is a d-dimensional vector. Time t may be either continuous ($t \in R$) or discrete (integer valued, $t \in Z$).

Two types of dynamics will be important in our considerations. First we introduce the concept of a **dynamical system** $\{S_t\}_{t \in R}$ (or, alternately, $t \in Z$ for discrete time systems) on a phase space X, which is simply any group of transformations $S_t : X \rightarrow X$ having the two properties: (1) $S_0(x) = x$; and (2) $S_t(S_{t'}(x)) = S_{t+t'}(x)$ for $t, t' \in R$. Dynamical systems are **invertible** or **reversible** since they may be run either forward or backward in time. All of the laws of classical and quantum physics are invertible and describe the behaviour of reversible systems.

The second type of dynamics that is important to distinguish are those of **semidynamical systems** $\{S_t\}_{t>0}$, which is any semigroup of transformations $S_t : X \rightarrow X$, *i.e.* (1) $S_0(x) = x$; and (2) $S_t(S_{t'}(x)) = S_{t+t'}(x)$ for $t, t' \in R^+$ (or N). In sharp contrast to dynamical systems, semidynamical systems are **noninvertible** or **irreversible** and may not be run backward in time in an unambigious fashion.

It is important not to confuse the issue of reversibility (or invertibility), which is a purely mathematical question, with the issue of dissipativeness, which is a physical question. In spite of the enormous significance of distinguishing between dynamical and semidynamical systems later, at this point no assumption is made concerning the invertibility or noninvertibility of the system dynamics.

1.3. Thermodynamic States and Densities.

In keeping with the ergodic theory approach adopted here we study the way in which the system dynamics operate on an infinite number of initial points. More specifically, we will examine the way in which the dynamics alter densities. If f is an L^1 function in the space X, *i.e.*, if $\int_X |f(x)|\, dx < \infty$, then f is a **density** if $f \in \{f \in L^1 : f \geq 0, \| f \| = 1\}$. As usual, $\| f \|$ denotes the L^1 norm of the function f, $\| f \| = \int_X |f(x)|\, dx$. The examination of the evolution of densities by system dynamics is equivalent to examining the behaviour of an infinite number of trajectories. This apparently simple assumption concerning the way in which systems operate on densities is so fundamental and important to the understanding of the foundations of thermodynamics that it is given a special status.

POSTULATE B. A thermodynamic system has, at any given time, a state characterized by a density f(x), not necessarily independent of time.

Given a density f then the **f-measure** $\mu_f(A)$ of the set A in the phase space X is defined by $\mu_f(A) = \int_A f(x)\, dx$, and f is called the **density** of the measure μ_f. The usual Lebesgue measure of a set A is denoted by $\mu_L(A)$, and the density of the Lebesgue measure is the uniform density, $f(x) = 1/\mu_L(X)$ for all points x in the phase space X. We always write $\mu_L(dx) = dx$.

Both Boltzmann and Gibbs, in their treatments of statistical mechanics, assumed they were dealing with systems of dimension $d = 2s$ whose dynamics were described by s position variables x_i and s momentum variables p_i.

Boltzmann considered the basic phase space to be a $2s$ dimensional space which is usually called μ **space**. He then considered the evolution of a large number N of identical particles, each with the same dynamics, in μ space. N is large and typically on the order of Avagadro's number, 6×10^{23}. The limiting case of $N \rightarrow \infty$ is the **thermodynamic limit** in which case the Boltzmann approach considers the evolution of a density in μ space.

Gibbs also considered N identical particles operating with these $2s$ dimensional dynamics in a phase space (commonly called the Γ **space**) of dimension $2sN$. He then considered an infinite number of copies of this original system, and gave this construct the name **ensemble**. Thus Gibbs studies the evolution of the ensemble density.

1.4. Boltzmann-Gibbs Entropy.

In his work Gibbs, assuming the existence of a system state density f on the phase space X, introduced the concept of the **index of probability** given by $\log f(x)$ where "log" denotes the natural logarithm. He then introduced a quantity $H(f)$ which is the negative of the phase space average of the index of

probability weighted by the density f, *i.e.*

$$H(f) = -\int_X f(x)\log f(x)\,dx.$$

This is now known as the **Boltzmann-Gibbs entropy** of a density f since precisely the same expression appears in Boltzmann's work (with the opposite sign) but the phase space is different for Boltzmann (μ space) and for Gibbs (Γ space). The Boltzmann Gibbs entropy is just the expectation of the negative of the index of probability, and is the only reasonable candidate for a theoretical analog of the empirical thermodynamic entropy. This is because the only function for the index of probability that gives the requisite additive property to make the entropy an extensive quantity is the logarithmic function, and that it is unique up to a multiplicative constant.

2. MAXIMAL ENTROPY PRINCIPLES.

2.1. Microcanonical Ensembles.

We may immediately understand the origin of the classical Gibbs microcanonical ensemble as reflecting a simple manifestation of extremal properties of the entropy. Consider a given space X with finite Lebesgue measure, $\mu_L(X) < \infty$ (forgo the normalization $\mu_L(X) = 1$ temporarily), and all possible densities f. Then the only density that maximizes the entropy is the (uniform) density of the Lebesgue measure of X. More precisely,

THEOREM 2.1. *When $\mu_L(x) < \infty$, the density that maximizes the Boltzmann-Gibbs entropy is the uniform density*

$$f_*(x) = \frac{1}{\mu_L(X)}. \tag{1}$$

For any other density $f \neq f_*$, $H(f) < H(f_*)$.

Notice that in this theorem there is no reference to the nature of the dynamics of the system generating the density. This is in sharp contrast to the usual approach in thermodynamics in which the dynamics are quite specifically used to argue for the plausibility of the microcanonical density (1). The fact that a generalization of this density appears in such a natural way merely illustrates the generality of both the density and the method used to obtain it, and that the existence of the density of the microcanonical ensemble is independent of the system dynamics.

2.2. Canonical Ensembles.

Even more interesting consequences can emerge from the extremal properties of entropy that offer insight into the basic foundation of thermodynamics of both classical and quantum systems. Namely,

THEOREM 2.2. *Assume that a nonnegative measurable function $\alpha(x)$ is given as well as the average $<\alpha>$ of that function over the entire space X, weighted by the density f:*

$$<\alpha> = \int_X \alpha(x) f(x)\,dx. \tag{2}$$

(Note that $<\alpha>$ is nonnegative and may be time dependent.) Then the maximum of the entropy $H(f)$, subject to the constraint (2), occurs for the density

$$f_*(x) = Z^{-1} e^{-\nu\alpha(x)} \tag{3}$$

where Z is defined by

$$Z = \int_X e^{-\nu\alpha(x)}\,dx, \tag{4}$$

and ν is implicitly determined from $<\alpha> = Z^{-1}\int_X \alpha(x) e^{-\nu\alpha(x)}\,dx$.

The choice of notation in (3) and (4) was intentional to draw the connection with the density of the Gibbs canonical ensemble, especially that Z corresponds to the partition function. It is quite easy to state and prove an obvious generalization of Theorem 2.2 that is applicable to systems in which there are multiple known averages $<\alpha_i>$.

2.3. The Thermodynamic Connection.

All of conventional equilibrium thermodynamics can be deduced from the density (3), by a proper association with thermodynamic quantities[25,25a], if we admit the fundamental assumption of thermodynamics that

POSTULATE C. There exists a one to one correspondence between states of thermodynamic equilibrium and states of maximum entropy.

If there is but one state of thermodynamic equilibrium that is attained regardless of the way in which the system is prepared then this is called a **globally stable equilibrium** and is associated with a globally stable state of maximal entropy (3^{rd} order form of the Second Law). If, however, there are multiple states of thermodynamic equilibrium, each corresponding to a state of locally maximal entropy and dependent on the initial preparation of the system, then we say that these are **local** or **metastable states of equilibrium** (2^{nd} order Second Law).

3. REVERSIBLE AND IRREVERSIBLE SYSTEMS.

From the perspective of the previous section, the central problem in thermodynamics is: **"How may one guarantee that the entropy of a system will increase to its maximum value regardless of the manner in which it was prepared?"** In this section we start our investigation of this question.

3.1. Markov Operators.

In every situation considered by theoretical physics, as developed to this point in time, the evolution of densities may be studied by the use of either the linear Markov or Frobenius-Perron operators. This is in spite of the fact that the underlying system dynamics responsible for the evolution of the density may be highly nonlinear.

The Frobenius-Perron operator, introduced in Section 4, is the most useful to describe the evolution of densities in systems for which the dynamics are totally deterministic, *i.e.* the dynamics evolve according to a very specific law that permits the accurate specification of a system state at any point in time.

The Frobenius-Perron operator is a special case of the more general Markov operator which may be used in the description of both deterministic and stochastic systems. Since the first results on reversibility and irreversibility that are of importance to an understanding of thermodynamics can be stated for Markov operators, we start with them.

Any linear operator $P^t : L^1 \to L^1$ that satisfies: (1) $P^t f \geq 0$; and (2) $\| P^t f \| = \| f \|$ for all $t \in R$ and $f \geq 0$, $f \in L^1$ is called a **Markov operator**. *If we restrict ourselves to only considering densities f, then any operator P which when acting on a density again yields a density is a Markov operator.* Any density f_* that satisfies $P^t f_* = f_*$ for all t is said to be a **stationary density** of the Markov operator P.

In precise analogy with the definitions of dynamical and semi- dynamical systems in the last section, we may discuss reversible and irreversible Markov operators. Given a Markov operator P^t, then P^t is a **reversible Markov operator** if: (1) $P^0 f = f$; and (2) $P^t(P^{t'} f) = P^{t+t'} f$ for all $t, t' \in R$ or Z. Clearly, allowing $t, t' \in R$ or Z is the origin of the reversibility. However, if property (2) of a reversible Markov operator is replaced by (2') $P^t(P^{t'} f) = P^{t+t} f$ for all $t, t' \in R^+$ or N, then P^t is an **irreversible Markov operator**.

3.2. Conditional Entropy.

If f and g are two densities such that supp $f \subset$ supp g [supp f denotes the **support** of f], then the **conditional entropy** of the density f with respect to the density g is

$$H_c(f|g) = -\int_X f(x) \log \left[\frac{f(x)}{g(x)} \right] dx. \tag{5}$$

The conditional entropy is always defined, *i.e.* H_c is finite or equal to $-\infty$. As is evident from the defining equation (5), $H_c(f|g)$ measures the deviation between the two densities f and g.

There are two important properties of $H_c(f|g)$: (1) Since f and g are both densities, it can be shown that $H_c(f|g) \leq 0$. It is only when $f = g$ that the equality holds; and (2) If g is the constant density of the microcanonical ensemble, *i.e.* $g = 1/\mu(X)$ throughout the phase space X, then $H_c(f|g) = H(f) - \log \mu(X)$. If the space X is normalized, then $g = 1$ and $H_c(f|1) = H(f)$. This illustrates how the conditional entropy is a generalization of the Boltzmann-Gibbs entropy.

From the definition it follows that $H_c(f|g) = H(f) + \int_X f(x) \log g(x)\, dx$. An elementary calculation using property (1) of $H_c(f|g)$ shows that the second term in the rewritten form of $H_c(f|g)$, with $f = g = f_*$, is just $-H(f_*)$ and that $H_c(f|f_*) = 0$ when $f = f_*$. These observations, in conjunction with our formulations of the 2^{nd} and 3^{rd} order forms of the Second Law, immediately suggest that the conditional entropy H_c can be interpreted as the entropy difference ΔS. For example, under the conditions of Theorem 2.2, the Boltzmann-Gibbs entropy $H(f)$ is maximized by the density f_* given by equation (3) and $H(f_*) = \log Z + \nu < \alpha >$. Thus, within the context of Theorems 2.1 and 2.2 we conclude that the conditional entropy will be zero whenever the Boltzmann-Gibbs entropy is at its maximum value.

With only the few tools developed so far and our identification of H_c with ΔS, the behaviour of the entropy of a sequence of densities $\{P^t f\}$ evolving under the action of a Markov operator may be examined. The first result is a weak (1^{st} order) form of the Second Law of thermodynamics stating that the conditional entropy is never decreasing. More precisely[32],

THEOREM 3.1. *If P^t is a Markov operator, then $H_c(P^t f|P^t g) \geq H_c(f|g)$ for and all densities f and g.*

Notice in this theorem that if $g = f_*$ is a stationary density of P^t so $P^t f_* = f_*$, then $H_c(P^t f|f_*) \geq H_c(f|f_*)$. Thus the conditional entropy with respect to a stationary density is always a nondecreasing function bounded above by $H_{max} = H_c(f_*|f_*) = 0$. In examining the behaviour of $H_c(P^t f|f_*)$ we therefore know that it has a limit as $t \to \infty$, though more information about P^t is required to define the limiting value.

The conclusions of Theorem 3.1 seem to be precisely the same as those reached by Boltzmann[1] in his pioneering work on the mechanical foundations of thermodynamics. However, things are not quite as transparent as this since to this point nothing has been said about the reversibility or irreversibility of the Markov operator P^t with respect to the behaviour of the entropy. This distinction is crucial for the limiting value of $H_c(P^t f|f_*)$ since the entropy for a reversible Markov operator is constant and determined by the way in which the system is prepared.

THEOREM 3.2. *If P^t is a reversible Markov operator, then the conditional entropy is constant for all times t, and equal to the value determined by the choice of the initial densities f and g. That is, $H_c(P^t f|P^t g) = H_c(f|g)$ for all t.*

From this theorem, in any system whose evolution of densities is described by a reversible Markov operator the entropy is forever fixed at a value determined by the initial state. Or, put another way, the entropy is uniquely determined

by the method of preparation of the system. A specialized form of the proof of Theorem 3.2 was used by Loschmidt[24] in his *Umkehreinwand* (objection based on time reversal) argument against the Boltzmann approach to statistical mechanics.

Thus, not too surprisingly, we conclude that *irreversibility in system dynamics*, as reflected in an evolution of densities *via* an irreversible Markov operator, *is necessary for the entropy to increase* as the system evolves. We cannot, however, assert that irreversibility is sufficient to guarantee this, and indeed it is not the case.

Based on much more specific assumptions, this result concerning the necessity of irreversibility was well known to Clausius[3] and Boltzmann[1], two of the founders of modern thermodynamic theory. How, then, did Boltzmann arrive at his conclusion that the entropy would increase to a maximum in a collection of particles moving under the action of (reversible) Hamiltonian dynamics? Both he and Clausius tried to circumvent this clear problem [the use of reversible (Hamiltonian) dynamics] by the addition of their *Stosszahlansatz* (molecular chaos) *postulate*. This reduces, quite simply, to a *postulate of irreversibility*.

4. ERGODICITY.

In the last section we asserted the necessity of irreversibility for increases in entropy to take place. However, the two interrelated questions of the existence of a unique state of thermodynamic equilibrium, and the global approach of the entropy to an absolute maximum, were not addressed. This section provides a necessary and sufficient criterion for the existence of a unique state of thermodynamic equilibrium as characterized by a unique stationary density f_*.

4.1. The Frobenius Perron Operator.

A transformation S_t is said to be **measurable** if $S_t^{-1}(A) \subset X$ for all $A \subset X$. Furthermore, given a density f_* and associated measure μ_*, a measurable transformation S_t is **nonsingular** if $\mu_*(S_t^{-1}(A)) = 0$ for all sets A such that $\mu_*(A) = 0$.

If S_t is a nonsingular transformation, then the unique operator $P^t : L^1 \to L^1$ defined by

$$\int_A P^t f(x)\,dx = \int_{S_t^{-1}(A)} f(x)\,dx \tag{6}$$

is called the **Frobenius-Perron operator** corresponding to S.

If f is a density, then equation (6) defining the Frobenius-Perron operator has a simple intuitive interpretation. Start with an initial density f and integrate this over a set B that will evolve into the set A under the action of the transformation S_t. However, the set B is $S_t^{-1}(A)$. This integrated quantity must be equal, since S_t is nonsingular, to the integral over the set A of the density obtained after one application of S_t to f. This final density is $P^t f$.

Given a density f and associated measure μ_f, then a measurable transformation S_t is said to be **f measure preserving** if $\mu_f(S_t^{-1}(A)) = \mu_f(A)$ for all sets A. Measure preserving transformations are necessarily nonsingular. Since the concept of measure preservation is not only dependent on the transformation but also on the measure, we alternately say that the measure μ_f is **invariant** with respect to the transformation S_t if S_t is f measure preserving.

It is possible to draw a connection between states of thermodynamic equilibrium, invariant measures and stationary densities of the Frobenius-Perron operator through the following theorem.

THEOREM 4.1. *Let S_t be a nonsingular transformation and P^t the Frobenius-Perron operator associated with S_t. Then there exists a state of thermodynamic equilibrium whose density f_* is a stationary density of P^t if and only if the measure $\mu_*(A) = \int_A f_*(x)dx$ is invariant with respect to S_t.*

In particular the density $f_* = 1$ of the microcanonical ensemble corresponds to a state of thermodynamic equilibrium if and only if the system dynamics preserve the Lebesgue measure. That is, systems preserving the Lebesgue measure may be appropriately described by the microcanonical ensemble. Of course it is important to realize that this theorem says nothing about either the uniqueness of this state of thermodynamic equilibrium or of the invariant measure corresponding to it.

4.2. Ergodicity.

We are ready to begin consideration of the characteristics S_t must have to guarantee the existence of a unique state of thermodynamic equilibrium that maximizes the entropy. The density maximizing the entropy should also be an equilibrium density, so our search is really one for the properties of S_t necessary to guarantee that a density f_* is a stationary density of the Frobenius-Perron operator corresponding to S_t, *i.e.* $P^t f_* = f_*$, and that f_* is unique.

We start by defining a few new terms, given, as usual, dynamics described by a transformation S_t. First, any set A such that $S_t^{-1}(A) = A$ is called an **invariant set**. Given a density f_* on a space X, any invariant set A such that $\mu_*(A) = 0$ or $\mu_*(X \setminus A) = 0$ is called **trivial**.

A nonsingular transformation S_t is said to be **f_* ergodic** if every invariant set A is a trivial subset of the phase space X, *i.e.* either $\mu_*(A) = 0$ or $\mu_*(X \setminus A) = 0$. If the phase space is finite and $f_*$is the uniform density of the microcanonical ensemble, then we say that S_t is **uniformly ergodic** instead of f_* ergodic. In the older physics and mathematics literature, ergodic systems were often called **metrically transitive** or **metrically indecomposable**.

The following result establishes a one to one correspondence between the uniqueness of a state of thermodynamic equilibrium and ergodicity.

THEOREM 4.2. *Let S_t be a nonsingular transformation and P^t the corresponding Frobenius Perron operator. There is a unique state of thermodynamic equilibrium with associated stationary density f_*, $P^t f_* = f_*$ if and only if S_t is f_* ergodic.*

What does this result, in conjunction with Theorem 4.1, tell us? First consider the microcanonical ensemble with its uniform density. Then a given dynamics S_t will be measure preserving with respect to the Lebesgue measure if and only if the uniform density of the microcanonical ensemble is a stationary density of the Frobenius-Perron operator corresponding to S_t. Furthermore, from Theorem 4.2 the uniform density of the microcanonical ensemble will be the unique stationary density of P^t if and only if the system S_t is uniformly ergodic. Hence, the existence of a unique state of thermodynamic equilibrium, characterized by the uniform density of the microcanonical ensemble which maximizes the Boltzmann-Gibbs entropy to zero, is totally dependent on the operation of a uniformly ergodic dynamics that preserves the Lebesgue measure!

In the more general case, the nonuniform density f_* of the canonical ensemble which maximizes the conditional entropy will be the unique density corresponding to a state of thermodynamic equilibrium if and only if it is the stationary density of the Frobenius-Perron operator corresponding to an f_* ergodic system S_t with respect to which the measure $\mu_*(A) = \int_A f_*(x)\,dx$ is invariant.

Thus in complete generality *ergodicity is necessary and sufficient to guarantee the existence of a unique state of thermodynamic equilibrium characterized by maximal entropy.* That this unique state exists is, of course, only half of the picture for we must also understand what kind of systems can evolve to that state.

To conclude we state one last theorem concerning conditions for the f_* ergodicity of a transformation S_t and thus, by our comments following Theorem 4.2, for the existence of a unique state of thermodynamic equilibrium.

THEOREM 4.3. *Let S_t be a nonsingular transformation and P^t the corresponding Frobenius Perron operator with stationary density $f_* > 0$ for all points in the phase space X. Then S_t is f_* ergodic if and only if $\{P^t f\}$ is Cesàro convergent to f_* for all densities f, i.e., if $\lim_{t\to\infty} \frac{1}{t}\sum_{k=0}^{t-1} < P^k f, g > = < f_*, g >$ in the discrete time case, or if $\lim_{T\to\infty} \frac{1}{T}\int_0^T < P^t f, g > dt = < f_*, g >$ in the continuous time case, for all bounded measurable functions g.*

Since Frobenius-Perron operators are specialized Markov operators, we extend the concept of ergodicity to Markov operators. Thus let P^t be a Markov operator and assume that P^t has a stationary density f_*. We will say that P^t is **f_* ergodic** if $\{P^t f\}$ is Cesàro convergent to f_* for all initial densities f.

5. ASYMPTOTIC PERIODICITY.

In this section we turn to an investigation of the fascinating property of asymptotic periodicity in the evolution of densities. This behaviour is the statistical analog for densities of the more common periodicity found in some time series. The existence of asymptotic periodicity implies a 2^{nd} order form of the Second Law.

5.1. Asymptotic Periodicity.

A Markov operator P^t is said to be **smoothing** if there exists a set A of finite measure, and two positive constants $k < 1$ and $\delta > 0$ such that for every set E with $\mu_L(E) < \delta$ and every density f there is some integer $t_0(f, E)$ for which $\int_{E\cup(X\setminus A)} P^t f(x)\,dx \le k$ for $t \ge t_0(f, E)$. This definition of smoothing just means that any initial density, no matter how small a region of the phase space X it is concentrated on, will eventually be smoothed out by P^t.

Smoothing operators are important because of the following result[19].

THEOREM 5.1. *Let P^t be a smoothing Markov operator. Then there is an integer $r > 0$, a sequence of nonnegative densities g_i and a sequence of bounded linear functionals λ_i, $i = 1, \ldots, r$, and an operator $Q : L^1 \to L^1$ such that for all densities f, Pf has the form*

$$Pf(x) = \sum_{i=1}^{r} \lambda_i(f) g_i(x) + Qf(x). \tag{7}$$

The densities g_i and the operator Q have the following properties: (1) The g_i have disjoint support (i.e. are mutually orthogonal), so $g_i(x)g_j(x) = 0$ for all $i \ne j$; (2) For each integer i there is a unique integer $\alpha(i)$ such that $Pg_i = g_{\alpha(i)}$. Furthermore, $\alpha(i) \ne \alpha(j)$ for $i \ne j$. Thus the operator P permutes the densities g_i; and (3) $\| P^t Q f \| \to 0$ as $t \to \infty$, $t \in N$, so Q can be viewed as a transient operator.

Notice from equation (7) that $P^{t+1}f$ may be immediately written in the form

$$P^{t+1}f(x) = \sum_{i=1}^{r} \lambda_i(f) g_{\alpha^t(i)}(x) + Q_t f(x), \qquad t \in N \tag{8}$$

where $Q_t = P^t Q$, $\| Q_t f \| \to 0$ as $t \to \infty$, and $\alpha^t(i) = \alpha(\alpha^{t-1}(i)) = \cdots$. The terms in the summation of (8) are just permuted by each application of P. Since r is finite, the sum $\sum_{i=1}^{r} \lambda_i(f) g_{\alpha^t(i)}(x)$ must be periodic with a period $T \le r!$. Further, as $\{\alpha^t(1), \ldots, \alpha^t(r)\}$ is just a permutation of $1, \ldots, r$ this summation takes the alternative form $\sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) g_i(x)$, where $\alpha^{-t}(i)$ is the inverse permutation of $\alpha^t(i)$.

This rewriting of the summation portion of (8) makes the effect of successive applications of P completely transparent. Each operation of P permutes the set of scaling coefficients associated with the densities $g_i(x)$ [remember that these densities have disjoint support]. Since the summation portion of (8) is periodic (with a period bounded above by $r!$), and $\| Q_t f \| \to 0$ as $t \to \infty$, we say that for any smoothing Markov operator the sequence $\{P^t f\}$ is **asymptotically periodic.**

One of the interesting interpretations of equation (8) is that *any asymptotically periodic system is quantized from a statistical point of view.* Thus if t

is large enough, which simply means that we have observed the system longer that its relaxation time so $\| Q_t f \| \simeq 0$, then $P^{t+1} f(x) \simeq \sum_{i=1}^{r} \lambda_i(f) g_{\alpha^t(i)}(x)$. Asymptotically, $P^t f$ is either equal to one of the basis densities g_i of the i^{th} pure state, or to a mixture of the densities of these states, each weighted by $\lambda_i(f)$. It is important to also realize that the limiting sequence $\{P^t f\}$ is, in general, dependent on the choice of the initial density f.

Asymptotically periodic Markov operators always have at least one stationary density given by

$$f_*(x) = \frac{1}{r} \sum_{i=1}^{r} g_i(x), \tag{9}$$

where r and the $g_i(x)$ are as in the Komornik-Lasota Theorem 5.1. It is easy to see that $f_*(x)$ is a stationary density, since by Property (2) of Theorem 5.1 we also have $P f_*(x) = \frac{1}{r} \sum_{i=1}^{r} g_{\alpha(i)}(x)$, and thus f_* is a stationary density of P^t. Hence, for any smoothing Markov operator the stationary density (9) is just the average of the densities g_i.

We close with a necessary and sufficient condition for the f_* ergodicity of a smoothing Markov operator.

THEOREM 5.2. *Let P be a smoothing, and thus asymptotically periodic, Markov operator working in a normalized measure space. Then P is ergodic if and only if the permutation $\alpha(i)$ of the Spectral Decomposition Theorem 5.1 is cyclical.*

Thus, cyclicity of the permutation $\alpha(i)$ is necessary and sufficient for the existence of a unique state of thermodynamic equilibrium characterized by the stationary density f_*.

5.2. Local Evolution of Entropy.

The fact that asymptotically periodic Markov operators have a stationary density given by (9) does not guarantee the uniqueness of this stationary density. Regardless of whether or not asymptotically periodic systems have unique stationary densities, they have the important property that their conditional entropy is an increasing function that approaches a maximum.

THEOREM 5.3. *Let P be an asymptotically periodic Markov operator with stationary density f_*. Then the conditional entropy $H_c(P^t f | f_*)$ of $P^t f$ with respect to f_* approaches a limiting value $H_{max}(f, f_*) \leq 0$, where*

$$H_{max}(f, f_*) = -\sum_i \int_X \lambda_i(f) g_i(x) \log \left\{ \frac{1}{f_*(x)} \sum_i \lambda_i(f) g_i(x) \right\} dx. \tag{10}$$

Note that if the stationary density f_* of P is given by (9), then the expression for $H_{max}(f, f_*)$ becomes even simpler. Namely, $H_{max}(f, f_*) = -\log r - \sum_{i=1}^{r} \lambda_i(f) \log \lambda_i(f)$ when we use the orthogonality of the densities $g_i(x)$. Since

$0 \leq \lambda_i(f) \leq 1$ for all i, we may also place a lower bound on $H_{max}(f, f_*)$ with $-\log r \leq H_{max}(f, f_*) \leq 0$.

This 2^{nd} order form of the Second Law of thermodynamics is the strongest result that we have yet encountered. It demonstrates that as long as the density in a discrete time system evolves under the operation of a Markov operator that is smoothing, the conditional entropy of that density converges to a maximum. However, there are two important facets of this evolution that should be recognized: (1) The convergence of the entropy is due to the fact that $\| Q_t f \| \to 0$ as $t \to \infty$ in the representation (8) of Theorem 5.1; and (2) The maximum value of the entropy, $H_{max}(f, f_*)$, as made explicit by the notation, is generally dependent on the choice of the initial density f and, thus, the method of preparation of the system.

6. MIXING.

In this section, we consider systems with irregular dynamical behaviors that are stronger than ergodic. Namely, we consider with dynamics described by f_* measure preserving transformations that have the property of strong, or Hopf, mixing. Systems with reversible mixing dynamics have entropies that are forever fixed by their mode of preparation. However, it is important to discuss mixing for two reasons. First there is a general misconception that mixing is sufficient to allow the evolution of entropy to a maximum. This is most certainly not the case in spite of the fact that mixing is necessary for the evolution of system entropy to a maximum. Secondly, as we will show in Section 8, if there is a certain imprecision in our knowledge of the values of the state variables in a mixing system then this is sufficient to cause the system entropy to evolve to its maximal value.

6.1. Mixing.

Gibbs[11] realized that ergodicity was inadequate to guarantee the approach of system entropy to equilibrium. As a consequence he qualitatively discussed a property stronger than ergodicity which is now called **(strong) mixing**. This was subsequently developed mathematically by Hopf[14], Koopman[20], and von Neumann[33].

Let S_t be an f_* measure preserving transformation operating on a finite normalized space. Then S_t is called **f_* mixing** if $\lim_{t \to \infty} \mu_*(A \cap S_t^{-1}(B)) = \mu_*(A)\mu_*(B)$ for all sets A and B. If f_* is the uniform density of the microcanonical ensemble, then we say that S_t is **uniformly mixing**. The defining relation for f_* mixing could equally well be written $\lim_{t \to \infty} \mu_*(A \cap S_t(B)) = \mu_*(A)\mu_*(B)$, whenever S_t is reversible (invertible).

It is a straightforward consequence of the definition that f_* mixing implies ergodicity. Furthermore, an f_* measure preserving transformation S_t, with associated Frobenius-Perron operator P^t and stationary density f_*, is f_* mixing if and only if the sequence $\{P^t f\}$ is weakly convergent to the density f_* for all

initial densities f. If $f_* = 1$, then S_t is uniformly mixing if and only if $\{P^t f\}$ is weakly convergent to the density of the microcanonical ensemble for all initial densities f. This is expressed more formally in:

THEOREM 6.1. *Let S_t be an ergodic transformation, with stationary density f_* of the associated Frobenius- Perron operator, operating in a phase space of finite f_* measure. Then S_t is f_* mixing if and only if $\{P^t f\}$is weakly convergent to f_*, i.e.,* $\lim_{t\to\infty} < P^t f, g >=< f_*, g >$.

In our subsequent discussion, we will call a Markov operator P^t with stationary density **f_* mixing** if $\{P^t f\}$ is weakly convergent to f_*.

Gibbs[11], Krylov[21], and many others have emphasized the importance of mixing for the understanding of thermodynamic behaviour. Indeed, at first one might think that this weak convergence of the sequence $\{P^t f\}$ to the density f_* of the canonical ensemble, or to the density $f_* = 1$ of the microcanonical ensemble, no matter what initial density f was chosen, would be exactly what is required to guarantee the approach of the entropy to its maximum. Such is not the case. It is most certainly true that mixing is necessary for this convergence of the entropy, but it is not sufficient as we show in Section 7.

6.2. Kolmogorov Automorphisms.

In this section, a concept that will be used later is briefly introduced, namely that of Kolmogorov automorphisms, or K automorphisms. We use the notation $S_t(\mathcal{A}) = \{S_t(A) : A \in \mathcal{A}\}$, $t = 0, \pm 1, \pm 2, \ldots$, where $\mathcal{A}$ is a sigma algebra. Then if S_t is an invertible transformation operating on a normalized space, and both S_t and S_{-t} are f_* measure preserving, S_t is said to be a **K-automorphism** if there is a sigma algebra $\mathcal{A}_0 \in \mathcal{A}$ such that: (1) $S_{-t}(\mathcal{A}_0) \in \mathcal{A}_0$; (2) The sigma algebra defined by $\cap_{t=0}^{\infty} S_{-t}(\mathcal{A}_0)$ is trivial in the sense that it only contains sets of f_* measure 0 or 1; and (3) The smallest sigma algebra containing $\cup_{t=0}^{\infty} S_t(\mathcal{A}_0)$ is identical to $\mathcal{A}$.

Kolmogorov automorphisms have behaviors stronger than mixing in that if a transformation is a K-automorphism then this also implies that it is f_* mixing[4,34]. The other property of K-automorphisms that is important for thermodynamic considerations is that since they are f_* measure preserving they have a unique stationary density f_*. However, since they are invertible,by Theorem 3.2 the entropy of a K-automorphism is identically equal to the initial entropy determined by the initial density with which the system is prepared.

7. ENTROPY EVOLUTION TO ITS MAXIMUM.

The results of the previous sections indicate that attention should be focused on extensions of the concepts of ergodicity, asymptotic periodicity, and mixing that may only occur in irreversible systems. Since we also know that increases in entropy need not culminate in the maximum value of the entropy (*e.g.* asymptotically periodic systems, Section 5), the essential question we must now face is:

Under what circumstances will the entropy of an irreversible system approach its maximum value of zero? This section provides a complete answer to this question.

7.1. Exactness.

If S_t is an f_* measure preserving transformation operating on a normalized phase space X, then S_t is said to be **f_* exact** if $\lim_{t\to\infty} \mu_*(S_t(A)) = 1$ for all sets A of nonzero measure. If f_* is the uniform density, $f_* = 1$, then we say that S_t is **uniformly exact**. The f_* exactness of a transformation implies that it is f_* mixing.

To understand the nature of exactness, it is first important to realize that *reversible systems can never be exact.* To see this, note that for a reversible f_* measure preserving transformation S_t we have $\mu_*(S_t(A)) = \mu_*(S_t^{-1}(S_t(A))) = \mu_*(A)$. Thus the definition of exactness is violated.

Exact systems are important for an understanding of how convergence to a stationary density f_* of the canonical ensemble may be reached in a way that is an extension of mixing. Specifically,

THEOREM 7.1. *If S_t is an f_* measure preserving transformation operating on a finite normalized phase space X and P^t is the associated Frobenius-Perron operator, then S_t is f_* exact if and only if* $\lim_{t\to\infty} \| P^t f - f_* \| = 0$, *i.e., $\{P^t f\}$ is strongly convergent to f_*, for all initial densities f.*

This theorem offers a necessary and sufficient condition for the exactness of S_t in complete analogy with the previously presented necessary and sufficient ergodicity and mixing conditions.

As with ergodicity and mixing, we extend the definition of exactness to Markov operators P^t with a stationary density f_*. Then we say P^t is **f_* exact** if $\{P^t f\}$ is strongly convergent to f_* for all initial densities f.[1] For Markov operators, f_* exactness implies f_* mixing implies f_* ergodicity. We close with a simple sufficient condition for the f_* exactness of asymptotically periodic Markov operators.

THEOREM 7.2. *Let P be a smoothing, and therefore asymptotically periodic, Markov operator. If $r = 1$ in the spectral decomposition (7) of $P^t f$, then P is f_* exact.*

7.2. The Second Law of Thermodynamics.

The main result of this section is a condition for the Second Law of thermodynamics to operate in its strongest possible (3^{rd} order) form. We consider a Markov operator P that has a stationary density f_* which is not necessarily constant, thus corresponding to the density of the canonical ensemble. Namely we have:

[1] Operators P^t that are f_* exact have been called strong Markov operators[27] and monotonic Markov operators[13].

THEOREM 7.3. *Let P^t be a Markov operator operating in a phase space X. Then the conditional entropy of $P^t f$ with respect to a density f_* goes to its maximal value of zero as $t \to \infty$,*

$$\lim_{t\to\infty} H_c(P^t f|f_*) = 0,$$

if and only if P^t is f_* exact.

Theorem 7.3 is remarkable in that it sets forth *necessary and sufficient criteria for the operation of the strongest form of the Second Law of thermodynamics*, namely for the entropy of a system to globally converge to its maximal value regardless of the way in which the system was prepared. The only requirement that the system must satisfy is that *the density must evolve under the action of an exact Markov operator.* If this operator is a Frobenius Perron operator then the dynamics must be f_* exact. Since f_* exactness implies f_* ergodicity, the state of thermodynamic equilibrium characterized by the density f_* is unique

Thus, the Boltzmann-Gibbs entropy will converge to its maximal value of zero if and only if the density of the microcanonical ensemble is a stationary density and the system evolves under the action of a uniformly exact Markov operator P! As before, *ergodicity of P guarantees that the uniform density of the microcanonical ensemble is the unique state of thermodynamic equilibrium, while the uniform exactness of P guarantees that the entropy will approach its maximum value of zero regardless of the way in which the system is prepared.* Hence, in general

$$P^t \text{ is } f_* \text{ exact} \Leftrightarrow \lim_{t\to\infty} H_c(P^t f|f_*) = 0.$$

These results point out a very interesting property of the entropy *vis /'a vis* the common notion that maximal entropy should be associated with maximal disorder, or minimal structure. Experimentally, what we *measure* is that the entropy of a system evolving in time goes to a maximum. Further, in the course of any experiment the *dynamics* are the ultimate selector of the proper f_* with respect to which the conditional entropy is "computed" by the system evolution. This state of maximal entropy, in turn, corresponds to a state of thermodynamic equilibrium, and in no way makes a judgment about whether this state is totally structureless ($f_* = 1$) or highly ordered. Any apparent inconsistency between a state of maximal entropy and a nonuniform f_* comes exclusively from the erroneous identification of $f_* = 1$ as the preferred state of thermodynamic equilibrium. This partially stems from the long historical preoccupation of trying to find a rational foundation for thermodynamics in the statistical mechanics of Hamiltonian systems which do preserve the Lebesgue measure and for which the attendant density $f_* = 1$ is a stationary density.

With the results of this section giving necessary and sufficient conditions for the approach of system entropy to a maximum, one might think that our quest for the dynamical foundations of thermodynamics and the functioning of

the second law was at an end. However, this is far from the reality of the situation as a moments reflection reveals.

Here it has been demonstrated that it is only through the operation of *irreversible* f_* exact systems that the entropy will increase to its maximal value (Theorem 7.3). Further, given the observation that dynamics are the ultimate determinant of the stationary density f_*, that this corresponds to a state of thermodynamic equilibrium, and that since states of thermodynamic equilibrium depend on a variety of parameters (temperature, pressure, etc.), we must conclude that the corresponding f_* must also depend on these parameters *as must the underlying dynamics.*

Given these results we are now faced with another problem since all of the laws of physics are framed in terms of reversible or invertible dynamical (as opposed to irreversible or noninvertible semidynamical) systems which are independent of these external parameters.

This dilemma seems to have at least two solutions. Either: (1) The laws of physics are at present incorrectly formulated. [Penrose[28] has argued quite lucidly and simply for this point of view, basing his thesis on CPT violation in K° meson decay. Fer[8] makes a similar point, basing his argument on the neglect of time delays in the usual formulations of physical laws. Gal-Or[9,10] and Davies[5] have extensively examined possible sources of time asymmetry in physics, primarily from a cosmological and electromagnetic perspective]; or (2) There is some effect, neglected to this point in our considerations, which alters the behaviour of a dynamical system to give rise to the observed behaviour.

The following sections are devoted to an exploration of the second of these possibilities, as the first involves a drastic restructuring of the entire formulation of classical and quantum physics.

8. COARSE GRAINING.

To this point, in calculating the entropy from the defining equations it has been assumed that the dynamical variables were known with infinite precision. As a consequence, the density f corresponding to a given thermodynamic state would also be known precisely. While this is the situation when an analytic form for the density is available, in the world of experiment the reality is that the density f (or, more usually, some functional of f) is either measured or estimated. Several consequences may ensue from this. The first and perhaps most obvious is that due to errors (arising, for example, from measurement impreciseness or numerical roundoff in computer experiments), f will not be known exactly but will be known only to some level of accuracy.

Alternately, it is entirely possible that Nature herself may have introduced an inherent graininess to phase space, rendering the absolute determination of dynamical variables, and thus densities, impossible. Many have suggested that there is an elementary fundamental volume in position-momentum space whose measure corresponds to Planck's constant. This would be entirely in keeping with

other apparently fundamental indivisible units in the real world. In this section we examine the consequences of imprecision in the measurement of dynamical variables by studying the properties of the coarse grained entropy.

8.1. Coarse Grained Entropy.

To examine the effect of imprecision in the measurement of dynamical variables on entropy calculations, we introduce the concept of the entropy of a coarse grained density, or more briefly, the **coarse grained entropy**. This concept seems to have been first qualitatively discussed by Gibbs[11], and quantified by Ehrenfest and Ehrenfest[7]. Denbigh and Denbigh[6] have considered aspects of the effects of coarse graining on the behaviour of entropy.

Coarse graining is carried out by first partitioning the phase space X (finite and normalized) into discrete cells A_i that satisfy $\cup_i A_i = X$ and $A_i \cap_{i \neq j} A_j = \emptyset$. Obviously, there is no unique way in which such a partition $\{A_i\}$ may be formed, but we require that the partition is **nontrivial** with respect to the Lebesgue measure μ_L so $0 < \mu_L(A_i) \leq \mu_L(X) = 1$ for all values of i. For every density f, within each cell A_i of this partition we denote the average of f over A_i by $< f >_i$,

$$< f >_i = \frac{1}{\mu(A_i)} \int_{A_i} f(x)\, dx, \tag{11}$$

so the density f coarse grained with respect to the partition A_i is given by

$$f^{cg}(x) = \sum_i < f >_i 1_{A_i}(x). \tag{12}$$

[Here, $1_A(x) = 1$ for $x \in A$ and $1_A(x) = 0$ when $x \notin A$.] Clearly $\sum_i < f >_i \mu_L(A_i) = 1$, and it is important to realize that f^{cg} is constant within each cell A_i, having the value given by (11).

Therefore, given a nontrivial partition A_i, a density f, and a coarse grained density f^{cg} defined by (11)-(12), then the Boltzmann-Gibbs entropy of the coarse grained density f^{cg} is given by

$$H(f^{cg}) = -\sum_i < f >_i \mu_L(A_i) \log < f >_i .$$

It is noteworthy that for any density f, the Boltzmann-Gibbs entropy of the coarse grained density f^{cg} may be greater than the entropy of f, or more specifically:

THEOREM 8.1. *For any density f and any nontrivial partition A_i of the phase space X, $H(f) \leq H(f^{cg})$.*

Thus, the effect of any imprecision in the estimation of a density f characterizing a system, no matter what the origin, will be to either increase the entropy of the estimated (coarse grained) density $H(f^{cg})$ above its actual value $H(f)$, or leave it unchanged.

Precisely analogously to the way in which the entropy of the coarse grained density was derived, it is easy to show that the conditional entropy of f^{cg} with respect to a second density g, also coarse grained with respect to the partition A_i, is given by $H_c(f^{cg}|g^{cg}) = -\sum_i < f >_i \mu_L(A_i) \log [< f >_i / < g >_i]$. It is equally easy to show that $H(f|g) \leq H_c(f^{cg}|g^{cg})$ for all densities f and g, and nontrivial partitions A_i of the phase space X.

Therefore, in general, *coarse graining of the phase space*, and the consequent coarse graining of a density, *will either increase the entropy or leave it equal to its value before coarse graining.*

In analogy with (12), the coarse grained $P^t f$ is given by $(P^t f(x))^{cg} = \sum_i < P^t f >_i 1_{A_i}(x)$ where $< P^t f >_i = [\mu_L(A_i)]^{-1} \int_{A_i} P^t f(x)\, dx$. It is important to realize that we are assuming that the Markov operator operates without any error on the density f, and that the coarse graining arises because of our inability to precisely measure dynamical variables, and consequently densities, for whatever reason. The converse situation in which we may measure densities with infinite precision, but the dynamics always work with some error are considered in Sections 10 and 11 where we consider system interactions with a heat bath.

Simple examples show that, for reversible systems, coarse graining: (1) Induces the entropy of the coarse grained density to approach the equilibrium entropy for both positive and negative times; (2) This approach may not be monotone: (3) The approach is not necessarily symmetric with respect to a reversal of time; and (4) The approach may be dependent on the partition chosen.

8.2. Coarse Graining of Mixing Systems.

Coarse graining has interested numerous authors since the concept was first introduced by Gibbs[11] with the observation that coarse graining of a mixing system should lead to an increase in the entropy to its maximal value. More specifically,

THEOREM 8.2. *If P^t is a reversible f_* mixing Markov operator with a unique stationary density f_*, and $\{A_i\}$ is a nontrivial partition of the phase space X, then $\lim_{t \to \pm\infty} (P^t f)^{cg} = f_*^{cg}$ for all initial densities f.*

As a consequence of this result and Theorem 7.3, we have

THEOREM 8.3. *If P^t is a reversible f_* mixing Markov operator with a unique stationary density f_* and $\{A_i\}$ is a nontrivial partition of the phase space X, then $\lim_{t \to \pm\infty} H_c((P^t f)^{cg} | f_*^{cg}) = 0$ for all initial densities f. [Notice that the entropy approach to zero is independent of the direction of time!]*

For uniformly mixing systems operating in a normalized finite space, it is an easy consequence of these results that after coarse graining of the phase space, $\{P^t f\}$ will approach the density of the microcanonical ensemble, and that the Boltzmann-Gibbs entropy will approach its maximum value of zero.

8.3. Summary.

Even setting aside the lack of irreversibility in the behaviour of the coarse grained entropy, it is important to realize that the rate of convergence of the entropy of the coarse grained densities that Theorem 8.3 guarantees will, in general, depend on the way in which the coarse graining of the phase space is carried out. Experimentally, if entropy increases to a maximum only because we have reversible mixing dynamics but there is coarse graining due to measurement imprecision, then the rate of convergence of the entropy (and all other thermodynamic variables) to equilibrium should become slower as measurement techniques improve. Such phenomena have not been observed. Thus, it seems likely that nontrivial coarse graining plays no role in determining thermodynamic behaviour, even if the coarse graining is externally imposed by Nature in the form of an inherent graininess or unitary cellularity of phase space.

9. TRACES AND FACTORS.

As an alternative to the coarse graining of the previous section, we now explore the consequences of a reversible dynamics in which not all dynamical variables are observable. Essentially this means that we have a dynamical system operating in an m-dimensional space, but are able to observe only $n < m$ of these variables. That is, we observe only a *trace* of its operation in an n-dimensional space because $(m - n)$ of the variables are hidden to us, *e.g.* because either we do not know about them, or do not have the technology to measure them.

9.1. Traces.

Let X and Y be two (topological Hausdorff) phase spaces, $F : Y \to X$ a given continuous function, and $S_t : Y \to Y$ a dynamical system operating in the phase space Y. A function $h : R \to X$ is the **trace** of the dynamical system if there is a point y in the space Y such that $h(t) = F(S_t(y))$ for all times t.

One is naturally led to wonder under what circumstances a trajectory can be viewed as the trace of a higher dimensional dynamical system. It is actually easy to give a surprising answer to a much more general question. *Every continuous trajectory (function) in a space X is the trace of a* **single** *dynamical system operating in a higher dimensional phase space Y*! More precisely, we have the following result.

THEOREM 9.1 ("GOD" THEOREM). *Let the phase space X be arbitrary. Then there is a second phase space Y, a dynamical system S_t operating in Y, and a continuous function $F : Y \to X$ such that every continuous trajectory $h : R \to X$ is the trace of S_t.*

There are important consequences for the behaviour of the entropy when one is considering the trace of a dynamical system. If we have a dynamical system S_t operating on Y, then the entropy is always identically equal to the entropy

of the initial density since it is impossible for the entropy of a reversible system to change (Theorem 3.2.). However, this may not be the case for the entropy of the density of a trace of a dynamical system, since the simple act of taking a trace of a dynamical system with time independent entropy may be sufficient to generate a system in which the entropy is increasing. Of course, in general we do not know what the limit of this increase may be, and the entropy may certainly approach a limit considerably less than its maximal value if, for example, the trace is asymptotically periodic.

9.2. Factors.

This leads us to discuss specific types of traces for which much more can be said about the behaviour of the entropy. To see how this works, we introduce the notion of a factor of a transformation with the aid of the following diagram.

$$\begin{array}{ccc} Y & \xrightarrow{S_t} & Y \\ {\scriptstyle F}\downarrow & & \downarrow{\scriptstyle F} \\ X & \xrightarrow[T_t]{} & X \end{array}$$

Let X and Y be two different phase spaces with normalized measures μ_{f_*} and μ_{g_*} and associated densities f_* and g_* respectively, and $T_t : X \to X$ and $S_t : Y \to Y$ be two measure preserving transformations. If there is a transformation $F : Y \to X$ that is also measure preserving, *i.e.*, if $\mu_{g_*}(F^{-1}(A)) = \mu_{f_*}(A)$ for all subsets A of the phase space Y, and such that $T_t \circ F = F \circ S_t$ (so the diagram commutes), then T_t is called a **factor** of S_t. From this definition *the trajectory of the factor T_t is a trace of the system S_t.*

The formal connection between these concepts and the behaviour of the entropy is furnished by the following theorem due to Rochlin[30].

THEOREM 9.2. *Every f_* exact transformation is the factor of a K-automorphism.*

The transformation F involved in our discussion of factors and traces is precisely what Misra *et al.*[27] and Prigogine[29] refer to as a projection operator in their work on the generation of irreversible behaviour from reversible dynamics. Theorem 9.2 serves as the analytic link in their work between reversible K-automorphisms and f_* exact transformations (or strong or monotonic Markov operators).

As noted in Section 6, since K-automorphisms are invertible their entropy is forever fixed at its initial value by Theorem 3.2. On the other hand, by Theorem 7.3 we know that the entropy of an f_* exact transformation, which by the above theorem is the factor of a K- automorphism, increases smoothly to its maximum value of zero irrespective of the initial density with which the system was prepared.

9.3. Coarse Graining and Traces.

There is a connection between the effects of coarse graining the phase space and taking the trace of a dynamical system. In nontrivial coarse graining, we lose *some* information about the exact values of the dynamical variables. One could easily imagine the situation in which we have m variables, of which n were measured with infinite precision, and $(m-n)$ of them were measured with some error. Thus the act of taking a trace is just a more extreme case of coarse graining in which we are not able to measure *any* of the $(m-n)$ variables, *i.e.* the partition is trivial.

Viewed from this perspective, it is surprising that there is such a difference between the results of a nontrivial coarse graining of the phase space (no induction of irreversible thermodynamic behaviour), and examining only the trace of a dynamical system operating in a higher dimensional space than our observations permit. However, it is precisely trivial coarse graining of a phase space in which the evolution of densities is governed by the Liouville equation that leads to the Boltzmann equation and its successful predictions of the behaviour of dilute gases. The Boltzmann equation describes the behaviour of a density evolving under the action of a factor of the original dynamics.

9.4. Summary.

Here we have introduced the concept of a trace, and shown that when a trace is a factor of a dynamical system, the entropy may increase. Even stronger results are available in some circumstances when the trace is taken from a K-automorphism, for then the trace may be f_* exact with an entropy that increases to its maximal value of zero. This and the previous section have presented two possible ways out of the clear problems associated with the necessity of an irreversible system for entropy to increase, and the fact that all of the laws of physics are formulated as reversible dynamical systems.

10. OPEN DISCRETE TIME SYSTEMS.

This section examines the consequences of having a discrete time deterministic transformation experience a perturbation from an outside source. Thus we are starting to consider open systems, the mathematical analogue of the interaction between a system and a "heat bath". Stochastically perturbed continuous time systems with dynamics described by ordinary differential equations are considered in the Section 11.

10.1. An Operator Equation.

Assume that, in general, a system evolves according to a given transformation $S(x_t)$. The qualifying phrase 'in general' means that the transition $x_t \to x_{t+1} = S(x_t)$ occurs with probability $(1-\epsilon)$. In addition, with probability ϵ, the value of x_{t+1} is uncertain. If $x_t = y$ is given, then, in this case, x_{t+1} may be considered as a random variable distributed with a density $K(x,y)$ which depends on y.

One interpretation of this process is that ϵ corresponds to the degree of coupling between the system under study and the heat bath. If this is the case, then the parameter ϵ can be thought of as a number related to the ratio of the fundamental frequency of operation of the basic deterministic system, F_D, to the frequency of the outside perturbation coming from the heat bath, F_P. Thus when $F_D \ll F_P$, $\epsilon \simeq 1$ and the system operates almost like a random walk, while with $F_D \gg F_P$, we have $\epsilon \simeq 0$ and the system evolves almost completely deterministically. We will refer to the situation when $0 < \epsilon < 1$ as 'loose coupling', while for $\epsilon = 1$, in which the influence of the external system is always experienced, we will speak of strong (or continuous) coupling. However, as will become clear in Section 10.3, precisely the same formulation may be interpreted in a totally different fashion.

Assume that the dynamics of our system operate in a phase space X (with positive measure, of course) which is some measurable subset of R^d, and that the dynamics S are nonsingular and have an associated Frobenius Perron operator P_S. Then the operator P describing the evolution of densities in this mixed system operating with both deterministic and perturbed elements is

$$Pf(x) = (1-\epsilon)P_S f(x) + \epsilon \int_X K(x,y) f(y)\, dy. \tag{13}$$

It is straightforward to show that (13) is a Markov operator.

Since for fixed y, $K(x,y)$ is a density it clearly satisfies $K(x,y) \geq 0$ and $\int_X K(x,y)\, dx = 1$. This condition, in conjunction with the requirement that K is measurable means that K is a **stochastic kernel**. Further, we will always assume that for every $\eta > 0$ there is a $\delta(\eta) > 0$ such that $\int_E K(x,y)\, dx \leq \eta$, for every y in X and subset E of X with $\mu_L(E) \leq \delta$.

10.2. Loosely Coupled Systems.

We are now in a position to state our main results concerning the behaviour of the entropy of a discrete time deterministic system coupled to a heat bath. In investigating the properties of the evolution of densities by the operator equation (13), and the consequent behaviour of the entropy of these densities, some mild restrictions on both the transformation S and the kernel K are required. First, assume that the deterministic transformation S satisfies

$$|S(x)| \leq a_0|x| + a_1 \tag{14}$$

throughout the phase space, where $a_0 < 1$ and a_1 are nonnegative constants. Secondly, it will be assumed that with $b_0 < 1$ and b_1 nonnegative constants,

$$\int_X |x| K(x,y)\, dx \leq b_0|y| + b_1. \tag{15}$$

This condition is automatically satisfied if the phase space X is finite, but if it is unbounded then (15) prevents divergence of the trajectories to infinity.

The first result guarantees the existence of at least one state of thermodynamic equilibrium and the evolution of the conditional entropy to a maximum, though not necessarily to zero, in the presence of noise. Thus this following result is equivalent to the 2^{nd} order formulation of the Second Law of thermodynamics.

THEOREM 10.1. *If S is a nonsingular transformation that satisfies (14) and K is a uniformly integrable stochastic kernel satisfying (15), then for $0 < \epsilon \leq 1$ the operator P given by (13) is smoothing, and thus asymptotically periodic.*

Therefore, for any closed system whose dynamics evolve through the action of a nonsingular transformation S satisfying (14), placing it in contact with a second system whose effect on the first is a perturbation characterized by a kernel K satisfying (15) leads automatically to a situation in which the resulting open system is asymptotically periodic *regardless of the nature of the original closed system S*. Further, since this procedure induces asymptotic periodicity we know that at least one state of thermodynamic equilibrium, characterized by a stationary density f_*, exists and that the conditional entropy $H(P^t f|f_*)$ is an increasing function with a limiting value given by $H_{max}(f, f_*)$ as defined in (10).

Under certain circumstances involving loose coupling to a heat bath, there are even stronger results concerning the behaviour of the entropy, corresponding to the 3^{rd} order formulation of the Second Law of thermodynamics. One such case is as follows.

Assume that the value of the perturbation to the system S coming from the heat bath (when it occurs) at time $(t+1)$ is independent of the value of x_t. Then the stochastic kernel $K(x,y)$ is independent of y and simply becomes $K(x,y) = g(x)$, where g is the density of the perturbations ξ_t. In this case, with the external perturbations independent of the state of the system S, the perturbations ξ_t could be interpreted as completely stochastic or as coming from another deterministic system. They could even be viewed as the trace of some deterministic system whose Frobenius-Perron operator has g as its unique stationary density. This is a slightly different situation from that explored in Section 9. There we considered the effect on the behaviour of the entropy of only examining the trace of a system. Now we are considering the situation which could be interpreted as the *perturbation of a system by the trace of another system.*

In this case, the operator equation (13) takes the simple form

$$Pf(x) = (1-\epsilon)P_S f(x) + \epsilon g(x). \tag{16}$$

There are some surprising consequences of making the loose coupling independent of the state of the system S. Namely

THEOREM 10.2. *If P is the operator defined by (16), then $\{P^t f\}$ is f_* exact.*

Thus, by the simple expedient of loosely coupling a system to a heat bath such that the system experiences perturbations that are independent of the state

of the system, there will be a unique state of thermodynamic equilibrium and the conditional entropy of the coupled system will globally converge to zero regardless of the nature of the original system S!

10.3. Strongly Coupled Systems.

A much different interpretation of this perturbation at random times of a deterministically operating system is possible and related to strong coupling between a deterministic system and an external source of noise. When $\epsilon = 1$ and $X = R^d$, then equation (13) takes the form

$$Pf(x) = \int_{R^d} K(x,y) f(y)\, dy. \tag{17}$$

In thinking about the interpretation of (17), consider the following. Take the quantities $\xi_0, \xi_1, \ldots$ to be d-dimensional random vectors and let the phase space X be R^d. Then for a given $\{\xi_t\}$ and a dynamics W of two variables, $W : R^d \times R^d \to R$, we may assume that the system goes from $x_t = y$ to $x_{t+1} = W(y, \xi_t)$. Let $K(x,y)$ be the density of $W(y,\xi_t)$. Then the density will always exist if $W(y,z)$ as a function of z is nonsingular. If this is the case, then equation (17) describes the evolution of the densities corresponding to $x_{t+1} = W(x_t, \xi_t)$. We can make this more formal through the following unpublished theorem initially formulated and proved by A. Lasota, J. Traple, and J. Tyrcha.

THEOREM 10.3. *Let $g : R^d \to R^d$ be a density and $K : R^d \times R^d \to R^d$ be a stochastic kernel. Then the (generally nonunique) function $W : R^d \times R^d \to R^d$ defined implicitly by $\int_0^{W(y,z)} K(r,y)\, dr = \int_0^z g(u)\, du$ defines a dynamical system $x_{t+1} = W(x_t, \xi_t)$ where the ξ_t are independent random variables with density g. This system has an evolution of densities described by $f_{t+1} = Pf_t$ where the operator P is given by $Pf(x) = \int_{R^d} K(x,y) f(y)\, dy$.*

10.4. Asymptotic Periodicity and Additive Perturbations.

Often the perturbations are additive, $W(y,z) = S(y) + z$, so we have $x_{t+1} = S(x_t) + \xi_t$. It is rather surprising that a dynamics of this form may also appear as the consequence of taking a factor or trace[25,25a].

If the sequence $\{\xi_t\}$ of random variables has a common density g, then it follows that $K(x,y) = g(x - S(y))$, and equation (17) becomes

$$Pf(x) = \int_{R^d} f(y) g(x - S(y))\, dy. \tag{18}$$

For the special case of additive noise, (18) can be derived independent of any assumption concerning the nonsingularity of S. Furthermore, in this case the condition given by equation (15) reduces to $m = \int_{R^d} |x| g(x)\, dx < \infty$. Thus we have an immediate corollary to Theorem 10.1 for systems with added noise[23].

COROLLARY 10.4. *If S (nonsingular or not) is a transformation operating in the phase space R^d, satisfies inequality (14), and experiences an additive perturbation whose density has a finite first moment, then the sequence $\{P^t f\}$, where P is the Markov operator defined by equation (18), is asymptotically periodic.*

Hence for all situations in which perturbations are *added* to a transformation S, the effect is to induce asymptotically periodic behaviour regardless of the nature of the original unperturbed dynamics S (remember that S may even be singular!). Because of this, we also know that *perturbations induce at least one state of thermodynamic equilibrium*, whose stationary density is given by equation (9), *and guarantee the approach of the conditional entropy to a maximum* (Theorem 5.3).

For some transformations, the induction of asymptotic periodicity by the addition of perturbations would not be at all surprising, *e.g.* the addition of a stochastic perturbation to a transformation with an exponentially stable periodic orbit gives asymptotic periodicity. The surprising content of Theorem 10.1 (and Corollary 10.4) is that even in a transformation that has aperiodic limiting behavior, additive perturbations will result in asymptotic periodicity.

10.5. f_* Exactness and Additive Perturbations.

Under certain circumstances there are even stronger results concerning the effects of additive perturbations. Namely, additive perturbations may induce f_* exactness with a consequent increase in the conditional entropy to its maximal value of zero corresponding to the strongest (3^{rd} order) form of the Second Law of thermodynamics.

10.6. Parametric Perturbations.

As another specific example, consider the case when $W(y,z) = zS(y)$ and $S > 0$, so $x_{t+1} = \xi_t S(x_t)$. Using Theorem 10.3, it is straightforward to show that the operator (17) takes the explicit form

$$Pf(x) = \int_{R^d} f(y) g\left(\frac{x}{S(y)}\right) \frac{dy}{S(y)}. \tag{19}$$

Horbacz[15,16] has considered the behaviour of this system when $S : R^+ \to R^+$. The flavor of her results are summarized in the following two theorems.

THEOREM 10.5. *Let the Markov operator P be defined by (19). Assume that g is a density, $0 < S(x) \leq \alpha x + \beta$, and $\alpha m < 1$ with $m = \int_0^\infty x g(x)\, dx$, where α and β are nonnegative constants. Then the sequence $\{P^t f\}$ is asymptotically periodic.*

We close with a second theorem concerning f_* exactness induced by multiplicative perturbations.

THEOREM 10.6. *If the Markov operator P defined by (19) satisfies the conditions of Theorem 10.5 and, in addition, $g(x) > 0$ then $\{P^t f\}$ is f_* exact.*

Theorems 10.5 and 10.6 illustrate the behaviors that may be induced by multiplicative perturbations in discrete time systems. A number of other results concerning asymptotic periodicity and f_* exactness induced by multiplicative perturbations may be proved, but rather than giving these the reader is referred to Horbacz[15,16].

10.7. Markov Operators and Deterministic Processes.

On several occasions we have emphasized that the interpretation of a given dynamics is not necessarily clear cut. In fact, given any Markov operator P it is always possible to construct a sequence of deterministic transformations $\{S_n\}$ such that the limiting value of $\{P^t_{S_n} f\}$ approximates $\{P^t f\}$ as closely as one likes. Results along this line have been published by Brown[2] and Kim[18], but we state this in the spirit of an unpublished result of Lasota.

THEOREM 10.7. *Let P be a given Markov operator operating in a finite normalized space $X = [0,1)$. Then there is a sequence of transformations $S_n : X \to X$ with Frobenius-Perron operators P_{S_n} such that $\lim_{n\to\infty} ||P_{S_n} f - Pf|| = 0$.*

The consequences of this theorem are extremely far reaching, for it tells us that *any Markov operator*, whether it arises from the influence of random or deterministic perturbations on a totally deterministic system or through the action of a completely unperturbed deterministic system, *can always be approximated by a totally deterministic system to any degree of accuracy.*

10.8. Summary.

In this section we have explored the effects of outside perturbations acting on a deterministic system with discrete time dynamics, and we have interpreted this as the coupling of the system to a "heat bath". These outside perturbations can be viewed as perturbations coming from another deterministic system, from the trace of a deterministic system, or as stochastic perturbations. Whatever their source, we have shown in a variety of situations that the effect of these perturbations may be to either induce asymptotic periodicity or f_* exactness. It is interesting that asymptotic periodicity or f_* exactness may be induced in systems that had absolutely no remarkable behaviour, including statistical behaviour, before they experienced the outside influences. Under certain circumstances it is not even necessary that the original system be nonsingular.

Thus, the effect of perturbing systems in this way has a very powerful influence on the behaviour of their entropy. If the perturbation induces asymptotic periodicity, then the entropy will increase to a local maximum whose value depends in a complicated way on the initial preparation of the system (Theorem 5.3). If, on the other hand, f_* exactness is induced, then the entropy will increase to its absolute maximal value of zero (Theorem 7.3).

11. OPEN CONTINUOUS TIME SYSTEMS.

Given the results for discrete time dynamics, it is natural to wonder if the perturbation of systems with continuous time dynamics (that is, placing them in contact with a heat bath) will yield analogous results concerning the entropy. The effects of perturbation on the entropy of systems with dynamics described by sets of ordinary differential equations is briefly considered in this section.

11.1. Stochastic Differential Equations.

Often, the dynamics of physical processes are formulated by a system of ordinary differential equations $dx_i/dt = F_i(x)$, $i = 1, \dots, d$ operating in some region of R^d with initial conditions $x_i(0) = x_{i,0}$. Here we examine the behaviour of the stochastically perturbed analog

$$\frac{dx_i}{dt} = F_i(x) + \sum_{j=1}^{d} \sigma_{ij}(x)\xi_j, \qquad i = 1, \dots, d \tag{20}$$

with the same initial conditions, where $\sigma_{ij}(x)$ is the amplitude of the stochastic perturbation and $\xi_j = dw_j/dt$ is a "white noise" term that is the derivative of a Wiener process. [A continuous process $\{w(t)\}_{t>0}$ is a one dimensional **Wiener process** if: (1) $w(0) = 0$; and (2) For all values of s and t, $0 \le s \le t$ the random variable $w(t) - w(s)$ has the Gaussian density $g(t-s, x) = [2\pi(t-s)]^{-1/2} \exp[-x^2/2(t-s)]$. In a completely natural manner this definition can be extended to say that the d-dimensional vector $w(t) = \{w_1(t), \cdots, w_d(t)\}_{t>0}$ is a **d-dimensional Wiener process** if its components are one dimensional Wiener processes.]

Equation (20) is a stochastic differential equation. As in the case of a nonperturbed system of ordinary differential equations, if the functions $F_i(x)$ and $\sigma_{ij}(x)$ satisfy Lipschitz conditions, then (20) has a unique solution[12].

11.2. The Fokker Planck Equation.

The density function $f(t,x)$ of the process $x(t)$ generated as the solution to the stochastic differential equation (20) is defined by $\text{prob}(x(t) \in B) = \int_B f(t,s)\,ds$. To guarantee the existence and differentiability of $f(t,x)$, we will assume the **uniform parabolicity condition**: $\sum_{i,j=1}^{d} a_{ij}(x)\lambda_i\lambda_j \ge \rho \sum_{i=1}^{d} \lambda_i^2$ where $\rho > 0$. If the $a_{ij}(x)$ satisfy the uniform parabolicity condition and if they and the $F_i(x)$ are continuous and C^3, then $f(t,x)$ exists and is differentiable.

Under the assumption that σ_{ij} and b_i are C^2, they and their derivatives up to second order are continuous for $t > 0$ and all $x \in R^d$, and that they and their first derivatives are bounded, the evolution equation for the density $f(t,x)$ is given by

$$\frac{\partial f}{\partial t} = -\sum_{i=1}^{d} \frac{\partial [F_i(x) f]}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 [a_{ij}(x) f]}{\partial x_i \partial x_j} \tag{21}$$

in the Ito interpretation of (20). This evolution equation is known as the **Fokker-Planck equation**. In the absence of the diffusion term ($a_{ij} \equiv 0$) it reduces to the **generalized Liouville equation**.

If the stochastic differential equation (20) has an initial condition $x(0)$ and an an associated initial density f_0, then the solution $f(t,x)$ of the Fokker-Planck equation satisfies $f(0,x) = f_0(x)$. Further, if the solution of the Fokker-Planck equation can be written in the form $f(t,x) = \int_{R^d} \Gamma(t,x,s) f_0 g(s)\,ds$, where the kernel Γ is independent of the initial density $f(0,x) = f_0(x)$ and $\lim_{t\to 0} f(t,x) = f(0,x) = f_0(x)$, then $f(t,x)$ is said to be a **generalized solution** of the Fokker-Planck equation. Under some standard regularity conditions on the coefficients of the Fokker-Planck equation the generalized solution is unique. Since f is a density, the generalized solution corresponds to the evolution of the system to a unique thermodynamic state.

From the expression for the generalized solution, a family $\{P^t\}_{t\ge 0}$ of integral operators can be defined by $P^0 f_0(x) = f_0(x)$, $P^t f_0(x) = \int_{R^d} \Gamma(t,x,s) f_0(s)\,ds$, where $f(0,x) = f_0(x)$. If the generalized solution is unique, then the operator P is a Markov operator. It is a rather simple demonstration that the unique stationary density f_* defined by $\lim_{t\to\infty} P^t f(x) = f_*(x)$ is given by the (unique) solution of the elliptic equation $-\sum_{i=1}^{d} \partial[F_i(x)f]/\partial x_i + \frac{1}{2}\sum_{i,j=1}^{d} \partial^2[a_{ij}(x)f]/\partial x_i \partial x_j = 0$

For the continuous time closed system without noise, the evolution of the Frobenius-Perron operator P^t is determined by the the generalized Liouville equation. When the very same system is subject to external white noise perturbations, then the evolution of the Markov operator $P^t f$ is governed by the Fokker-Planck equation (20) which is just the same as the generalized Liouville equation with the addition of the diffusion term.

11.3. The Behaviour of Entropy.

A closed continuous time system with dynamics described by ordinary differential equations (reversible) has an entropy that is absolutely constant and equal to the entropy of the initial density with which the system was prepared (Theorem 3.2).

We now examine the effects of perturbations on the entropy of these continuous time reversible systems. For one dimensional systems ($d = 1$), Rudnicki[31] (1991) has recently proved a very interesting necessary and sufficient condition for the f_* exactness, and consequent existence of 2^{nd} Law behaviour in its strongest form. We extend his result to the behaviour of the conditional entropy.

THEOREM 11.3. *Assume that $d = 1$ and (21) has a unique generalized solution. Then the Markov operator P^t whose evolution is governed by equation (21) is f_* exact and the corresponding conditional entropy $H_c(P^t f | f_*)$ approaches its maximal value of zero as $t \to \infty$ if and only if $\int_{-\infty}^{+\infty} \exp\left[-\int^x \frac{2F(z)}{\sigma^2(z)}\,dz\right] dx = \infty$.*

11.4. Phase Transitions and Perturbations.

A variety of studies[17] have shown that in systems of differential equations that display a Hopf bifurcation as a parameter is varied, when noise is added as in (20) the stationary density of the Fokker Planck equation has behavior like that encountered in phase transitions. Though general results are not available at this time, it appears that if, in the absence of noise the system has a supercritical Hopf bifurcation, then when noise is added the stationary density has behaviour like that seen in 2^{nd} order phase transitions. Alternately, if the unperturbed system has a subcritical Hopf bifurcation then the density of the perturbed system has 1^{st} order phase transition like behaviour[26].

REFERENCES.

1 Boltzmann, L. 1909. **Wissenschaftliche Abhandlungen.** (ed. Fr. Hasenohrl), Barth, Leipzig.

2 Brown, J.R. 1966. *Approximation Theorems for Markov operators.* Pacific J. Math. **16**, 13-23.

3 Clausius, R. 1879. **The Mechanical Theory of Heat.** Macmillan, London.

4 Cornfeld, I.P., Fomin, S.V., and Sinai, Ya. G. 1982. **Ergodic Theory.** Springer Verlag, New York.

5 Davies, P.C.W. 1974. **The Physics of Time Asymmetry.** University of California Press, Berkeley and Los Angeles.

6 Denbigh, K.G. and Denbigh, J.S. 1985. **Entropy in Relation to Incomplete Knowledge.** Cambridge University Press, Cambridge.

7 Ehrenfest, P. and Ehrenfest, T. 1959. **The Conceptual Foundations of the Statistical Approach in Mechanics.** Cornell University Press, Ithaca.

8 Fer, F. 1977. **L'irréversibilité, fondement de la stabilité du monde physique.** Gauthier-Villars, Paris.

9 Gal-Or, B. 1974. *Statistical mechanics, fallacy and a new school of thermodynamics*, in **Modern Developments in Thermodynamics** (ed. B. Gal-Or), pp 3-51, John Wiley and Sons, New York.

10 Gal-Or, B. 1981. **Cosmology, Physics, and Philosophy.** Springer Verlag, New York.

11 Gibbs, J.W. 1902. **Elementary Principles in Statistical Mechanics.** Yale University Press, New Haven.

12 Gikhman, I.I. and Skorokhod, A.V. 1969. **Introduction to the Theory of Random Processes.** Saunders, Philadelphia.

13 Goldstein, S., Misra, B., and Courbage, M. 1981. *On intrinsic randomness of dynamical systems.* J. Stat. Phys. **25**, 111-126.

14 Hopf, E. 1932. *Proof of Gibbs' hypothesis on the tendency toward statistical equilibrium.* Proc. Nat. Acad. Sci. U.S.A. **18** 333-340.

15 Horbacz, K. 1989. *Dynamical systems with multiplicative perturbations.* Ann. Pol. Math. **50**, 11-26.

16 Horbacz, K. 1989. *Asymptotic stability of dynamical systems with multiplicative perturbations.* Ann. Pol. Math. **50**, 209-218.

17 Horsthemke, W. and Lefever, R. 1984. **Noise Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology.** Springer Verlag, Berlin.

18 Kim, C. 1972. *Approximation theorems for Markov operators.* Z. Warscheinlichkeitstheorie verw. Geb. **21**, 207-214.

19 Komornik, J. and Lasota, A. 1987. *Asymptotic decomposition of Markov operators.* Bull. Polish. Acad. Sci. (Math.) **35** 321-327.

20 Koopman, B.O. 1931. *Hamiltonian systems and transformations in Hilbert space.* Proc. Nat. Acad. Sci. U.S.A. **17** 315-318.

21 Krylov, N.S. 1979. **Works on the Foundations of Statistical Physics.** Princeton University Press, Princeton.

22 Lasota, A., and Mackey, M.C. 1985. **Probabilistic Properties of Deterministic Systems.** Cambridge University Press, Cambridge.

23 Lasota, A., and Mackey, M.C. 1987. *Noise and statistical periodicity.* Physica **28D** 143-154.

24 Loschmidt, J. 1876. *Über den Zustand des Wärmegleichgewichtes eines Systeme von Körpern mit Rucksicht auf die Schwerkraft.* Wien. Ber. **73** 128-142.

25 Mackey, M.C. 1989. *The dynamic origin of increasing entropy.* Rev. Mod. Phys. **61**, 981-1016.

25a Mackey, M.C. 1991. **The Origin of Thermodynamic Behaviour.** Springer Verlag, New York. In press.

26 Mackey, M.C., Longtin, A. and Lasota, A. 1990. *Noise induced global asymptotic stability.* J. Stat. Phys. **60** 735-751.

27 Misra, B., Prigogine, I., and Courbage, M. 1979. *From deterministic dynamics to probabilistic descriptions.* Physica **98A** 1-26.

28 Penrose, R. 1979. *Singularities and time-asymmetry*, in **General Relativity** (eds. S.W. Hawking and W. Israel), pp 581-638, Cambridge University Press, Cambridge.

29 Prigogine, I. 1980. **From Being to Becoming: Time and Complexity in the Physical Sciences.** W.H. Freeman and Company, San Francisco.

30 Rochlin, V.A. 1964. *Exact endomorphisms of Lebesgue spaces.* Am. Math. Soc. Transl. (2) **39** 1-36.

31 Rudnicki, R. 1991. *Asymptotic stability in L^1 of parabolic equations.* Preprint, University of Silesia.

32 Voigt, J. 1981. *Stochastic operators, information, and entropy*, Commun. Math. Phys. **81** 31-38.

33 von Neumann, J. 1932. *Proof of the quasi-ergodic hypothesis*, Proc. Nat. Acad. Sci. U.S.A. **18** 70-82.

34 Walters, P. 1982. **Ergodic Theory-Introductory Lectures.** Lecture Notes in Mathematics **458** Springer Verlag, New York.

A FUNDAMENTAL RELATION BETWEEN STOCHASTICITY AND QUANTIZATION

L. de la Peña and A. M. Cetto

Instituto de Física, UNAM
Apartado postal 20-364
01000 México 20, D.F.
Mexico

After a brief review of the basic postulates of stochastic electrodynamics (SED), a set of stochastic dynamical equations for the charged particle bound by an arbitrary external force are derived, under the assumption that due to the action of the random field and radiation reaction the initial conditions of the particle are lost in the long run. A partial averaging leads to a statistical description for functions either in configuration or momentum space; to extend the description to general observables a set of probability amplitudes is required, and the rules for such extension are extracted from the stochastic equations of motion. This procedure leads to a statistical description of the SED system in Hilbert space, which fully coincides with Heisenberg's description of quantum mechanics. Some comments are added on the impossibility of extending the description to full phase space and on the relation of these results with Bell's theorem.

I. Introduction

In this couple of lectures we construct a route to the formalism of quantum mechanics and establish contact with non-relativistic quantum electrodynamics in the frame of a stochastic theory, namely stochastic electrodynamics, under certain physical conditions that define the so-called quantum regime. According to the results presented, quantum theories should be considered as a very condensed and powerful statistical description of the stochastic behaviour of matter in interaction with the zero point radiation field, and of the radiation field in interaction with matter, in this regime.

Stochastic electrodynamics (SED) cannot be considered a new theory; in fact the first proposals for it go back to 1953-4,[1] and one of its latest and most important rediscoveries comes from 1963.[2] However, the theory received little

attention and only much more recently did a sustained effort for its development take place; Refs. 3 and 4 contain two review articles referring to this period. As a glance to this literature shows, the theory had significative though partial successes, allowing even for the emergence of stochastic optics,[5] a newborn branch of theoretical physics intended to offer a fresh understanding of optical and related quantum phenomena. However, it became apparent along the development of SED that some of its crucial results lead to contradictions with quantum mechanics, in particular those referring to the atomic problem and the existence of stationary states.[3,4]

A careful consideration of the details of the traditional form of stochastic electrodynamics reveals the presence of a very weak point in its general development, namely, the use of perturbative methods to construct its fundamental statistical description by means of a Fokker-Planck-type equation. To carry out such development, the basic assumption is made that to zero order in its coupling to the vacuum field, the important motion is along a *classical* trajectory, which amounts to postulating that the role of the field is reduced to a perturbation of a classical system. But this assumption turns out to be untenable: the vacuum field and radiation reaction play in the long run a determinant role in the behaviour of the system, not amenable to a perturbative description. Starting from this viewpoint, a new non-perturbative form of SED is proposed here —a brief summary may be found elsewhere[6]—, in an attempt to determine the potentialities of stochastic electrodynamics once the perturbative hypothesis is discarded. The results obtained up to now, even if still partial and incomplete, are encouraging; more specifically, in this paper we intend to show that a statistical treatment of the mechanical part of the SED system (once the quantum regime has been attained) finds a natural description through the usual quantum mechanical formalism in Hilbert space. Thus the theory allows us to look from a new angle at many of the properties of quantum systems, as, *e.g.*, their description in terms of state vectors in a Hilbert space and the statistical nature of such description, in which the notion of particular trajectories has been eradicated, or the nonexistence of a unique phase space distribution. The companion paper is devoted to the study of the radiative corrections, thus making contact with some results of quantum electrodynamics.

II. The Dynamical Equations of SED in Stochastic Phase Space

A typical system in nonrelativistic SED is a charged, spinless point particle, subject to an external binding force plus the electromagnetic radiation field, which comprises the random zero-point field, the field radiated by the particle, and perhaps some additional, external field (for instance, thermal radiation). The hamiltonian of the complete system is therefore

$$H_T = \frac{1}{2m}(\mathbf{p} - \frac{e}{c}\mathbf{A})^2 + V(\mathbf{x}) + H_R, \tag{2.1}$$

where H_R stands for the hamiltonian of the total radiation field. The vector potential of the vacuum field alone is expressed in terms of normal modes as

$$\mathbf{A} = \sum_\alpha \left(\frac{2\pi\hbar c^2}{V\omega_\alpha}\right)^{1/2} \hat{\epsilon}_\alpha \left[a_\alpha \exp[-i(\omega_\alpha t - \mathbf{k}\cdot\mathbf{x})] + \text{c.c.}\right], \tag{2.2}$$

where V is the normalization volume (which is eventually taken as infinite), $\omega_\alpha = ck_\alpha$, $\hat{\epsilon}_\alpha$ is the polarization vector, and $\alpha = (n_1, n_2, n_3; \sigma)$ refers to the wave vector and polarization indices of every field mode ($\sigma = 1, 2$). The amplitudes $a_\alpha \equiv a_{n\sigma}$ are complex, independent, stochastic variables, with the following statistical properties:

$$\langle a_{n\sigma}\rangle = 0, \quad \langle a_{n\sigma}a_{n'\sigma'}\rangle = 0, \quad \langle a_{n\sigma}a^*_{n'\sigma'}\rangle = \frac{1}{2}\delta_{nn'}\delta_{\sigma\sigma'}. \tag{2.3}$$

A set of complex amplitudes $b_\alpha \equiv b_{n\sigma} = (2\hbar\omega_n)^{-\frac{1}{2}}(p_\alpha - i\omega_n q_\alpha)$ is introduced to express the total radiation field, so that the corresponding hamiltonian is

$$H_R = \frac{1}{2}\sum_{n,\sigma}(p^2_{n\sigma} + \omega_n^2 q^2_{n\sigma}) = \sum_{n,\sigma}\hbar\omega_n b^*_{n\sigma}b_{n\sigma}. \tag{2.4}$$

In absence of external radiation and when the field radiated by the particle is neglected, the b_α take their zero-point value, a_α. The statistical properties (2.3) are such that the average energy of each vacuum field mode is $\frac{1}{2}\hbar\omega$. The usual canonical procedure leads to the dynamical equations for particle and field variables, starting from Eq. (2.1); those for the particle variables are

$$\dot{\mathbf{x}} = [\mathbf{x}, H_T] = [\mathbf{x}, H] - \frac{e}{mc}\mathbf{A}$$
$$\dot{\mathbf{p}} = [\mathbf{p}, H_T] = [\mathbf{p}, H] - \frac{e}{c}x_j\nabla A_j,$$

where H is the mechanical hamiltonian (see Eq.(2.8)). Since we are specifically interested in the dynamical response of the particle to the vacuum field, we calculate the Poisson brackets with respect to the particle and field variables at the initial time $t = 0$ (when it is assumed that the two parts of the system start to interact), which are $\mathbf{x}_0, \mathbf{p}_0$ and q^0_α, p^0_α, respectively.[7] For functions of $\mathbf{x}$ and $\mathbf{p}$, the Poisson brackets transform into

$$[f, g] = [f, g]_0 + [f, g]_\alpha = [f, g]_0 - \frac{i}{\hbar}\langle f; g\rangle \tag{2.5}$$

where $[\cdot,\cdot]_0, [\cdot,\cdot]_\alpha$ denote Poisson brackets with respect to the initial particle and field variables, respectively. In the second equality we introduced the *poissonian*, *i.e.*, the Poisson bracket with respect to the zero-point field amplitudes:

$$\langle f; g\rangle \equiv \sum_\alpha \left(\frac{\partial f}{\partial a_\alpha}\frac{\partial g}{\partial a^*_\alpha} - \frac{\partial g}{\partial a_\alpha}\frac{\partial f}{\partial a^*_\alpha}\right). \tag{2.6}$$

With this procedure, the (exact) equations for the particle become:

$$\dot{\mathbf{x}} = [\mathbf{x}, H]_0 - \frac{i}{\hbar}\langle\mathbf{x}; H\rangle - \frac{e}{mc}\mathbf{A} \tag{2.7a}$$
$$\dot{\mathbf{p}} = [\mathbf{p}, H]_0 - \frac{i}{\hbar}\langle\mathbf{p}; H\rangle - \frac{e}{c}x_j\nabla A_j, \tag{2.7b}$$

where

$$H = \mathbf{p}^2/2m + V(\mathbf{x}) \tag{2.8}$$

and $\mathbf{A}$ refers to the complete radiation field. Now we introduce a physical hypothesis: we assume that in the long run the bound particle forgets its initial conditions. The reason for this can easily be understood by recalling that the accelerated particle radiates, and thus the energy of the motion associated to the initial trajectory is systematically being radiated, whereas the stochastic motion is permanently renewed. Thus in the long run, all traces of the initial motion will disappear (in absence of the stochastic term, the bound particle would fall to a point of stable equilibrium, with almost near independency of the initial conditions). Notice that the argument holds with minor changes also for the free particle; in this case, no radiation comes from the systematic motion (due to the absence of acceleration) and the only contribution to the equations of motion from the initial conditions comes from the initial momentum, which can be eliminated by passing to an appropriate inertial frame. When this situation is

reached, the Poisson brackets with respect to $\mathbf{x}_0, \mathbf{p}_0$ become negligible and thus the above equations reduce to

$$\dot{\mathbf{x}} = -\frac{i}{\hbar}\langle \mathbf{x}; H\rangle - \frac{e}{mc}\mathbf{A}, \tag{2.9a}$$

$$\dot{\mathbf{p}} = -\frac{i}{\hbar}\langle \mathbf{p}; H\rangle - \frac{e}{c}x_j \nabla A_j. \tag{2.9b}$$

Under the same argument, the Poisson brackets of the $\mathbf{x}$ and $\mathbf{p}$ variables:

$$[x_i, x_j] = [x_i, x_j]_0 - \frac{i}{\hbar}\langle x_i; x_j\rangle = 0,$$

$$[p_i, p_j] = [p_i, p_j]_0 - \frac{i}{\hbar}\langle p_i; p_j\rangle = 0,$$

$$[x_i, p_j] = [x_i, p_j]_0 - \frac{i}{\hbar}\langle x_i; p_j\rangle = \delta_{ij}$$

reduce to

$$\langle x_i; x_j\rangle = \langle p_i; p_j\rangle = 0, \quad \langle x_i; p_j\rangle = i\hbar\delta_{ij}. \tag{2.10}$$

These basic poissonians together with the dynamical equations (2.9) constitute what we call the poissonian equations of SED. They govern the dynamics of the mechanical system once the information about the initial conditions has been lost; this situation is what we call the *quantum regime*. Eqs. (2.10) exhibit the fundamental stochastic nature of the particle variables $\mathbf{x}$, $\mathbf{p}$ in this regime, since they depend essentially on the random field variables a, a^*. One may even neglect the interaction terms in the dynamical equations (2.9) and write in a first approximation:

$$\dot{\mathbf{x}} = -\frac{i}{\hbar}\langle \mathbf{x}; H\rangle, \quad \dot{\mathbf{p}} = -\frac{i}{\hbar}\langle \mathbf{p}; H\rangle, \tag{2.11}$$

As will be seen below, these equations lead under weak assumptions to the quantum mechanical description, and the omitted terms lead to the radiative corrections. In other words, in the present treatment quantum mechanics represents the zero-order approximation in a development in terms of powers of the electric charge.

The above equations exhibit the *essential* role played by the background field in connection with the dynamics of the particle. They evidently cannot be derived by a perturbative treatment of the effect of the field on the (otherwise classical) mechanical system; here we find an explanation for the failure of all

previous (perturbative) treatments of the atomic problem in SED. This point will be discussed more at length in the following lecture.

At this stage any explicit reference to the coupling to the radiation field has been lost, and the system admits a description in which the cause of the randomness is concealed. Thus an additional feature of the present description emerges, namely that Eqs. (2.11) turn out to be essentially independent of the electromagnetic nature of the vacuum field. Any vacuum field with similar properties should lead essentially to the same equations; this could be the basis for an explanation of the universality of quantum mechanics. A number of years ago, Santos[8] generalized the concept of SED and introduced the term 'stochastic theory' to refer to a more general theory that would take into account all existing vacuum fields, responsible for all interactions; he conjectured that these fields are essentially in equilibrium and thus have the same energy per normal mode. In the present terms, this would mean that the $\hbar$ appearing in the poissonian equations would be the same, irrespective of the origin of the fluctuations. (A more general viewpoint along this line has been discussed by Moore).[9] However, it should be noted that for particles with electromagnetic interactions, the electromagnetic vacuum seems to be sufficient not only to explain the quantum behaviour but also to obtain the radiative corrections, as shown in the companion paper.

III. Statistical Description in Configuration Space

Rather than directly study the predictions of the (stochastic) poissonian equations, our present intention is to establish connection with the usual quantum description. For this purpose we first develop a statistical treatment in configuration space for the SED system described by such equations. In order to simplify the notation we work in one-dimensional space; the generalization to more dimensions is self-evident.

Let us start by recalling that all dynamical variables, such as x, p, etc., are functions of the set $\{a_\alpha, a_\alpha^*\}$, which means that to every specific realization of these random variables there corresponds a well defined value of x, p, and so on. Let $W(a)$ be the probability distribution of all random field variables (of the a's, for short). Then the average value of any function F of the a's is given by

$$\langle F\rangle = \int F(a)W(a)da \tag{3.1}$$

where da stands for $\prod_\alpha da_\alpha da_\alpha^*$ and the integration is performed over the complete phase space of the a's. Now we introduce the probability density of the ensemble with $x = x(a)$:

$$\rho(x) = \int \delta(x - x(a))W(a)da. \tag{3.2}$$

With W normalized to unity, integration of Eq. (3.2) over x gives $\int \rho(x)dx = 1$; thus ρ coincides indeed with the probability density and can be used to calculate the average value of a function F of $x(a)$. Given any such function, its value in the *x-representation* will be given by

$$F(x)\rho(x) = \int \delta(x - x(a))F(x(a))W(a)da. \tag{3.3}$$

On integrating over all x space we get

$$\int F(x)\rho(x)dx = \int F(x(a))W(a)da, \tag{3.4}$$

which shows that the average of F over the ensemble is indeed the same as its average in x-space. Assuming the field itself to be ergodic, this is then an ergodic property of the ensemble described by the distribution in configuration space.

In the canonical treatment of the SED system, the variables x and p are dealt with in a symmetrical way, as can be seen, *e.g.*, in the poissonian equations (2.10, 11). However, in the above paragraphs we have broken this symmetry in favour of x, by choosing to develop a description in configuration space. One can of course proceed in an entirely analogous way to construct a *p-representation*, by introducing the probability density in p-space:

$$\pi(p) = \int \delta(p - p(a))W(a)da. \tag{3.5}$$

Then for any function F of $p(a)$,

$$F(p)\pi(p) = \int \delta(p - p(a))F(p(a))W(a)da \tag{3.6}$$

and the average value of F is obtained by integrating over p-space:

$$\int F(p)\pi(p)dp = \int F(p(a))W(a)da. \tag{3.7}$$

In what follows we shall make explicit use of either one of these representations. However, it is often irrelevant to specify which is being used, and in such cases

we shall speak of the *q-representation*, where q can stand for x, p or any other appropriate variable related to these through a transformation.

IV. Completing the Description in Configuration Space

Now we go back to coordinate space and try to complete the statistical description by expressing the canonical variable p and its functions in the x-representation. In the quantum regime x and p become functions of the random variables a, connected through the basic poissonians, Eqs. (2.10). In particular, the equation $\langle x; p\rangle = i\hbar$ shows that x and p have a different dependence on the a's, and thus for a fixed value of x, the p remains stochastic. It becomes therefore necessary to generalize Eq. (3.3) in order to represent functions of p in the x-representation. While the function $\rho(x)$ serves to calculate the moments of x, for the calculation of those of p at least a second function of x is required. A most appropriate way of introducing it is by writing $\rho(x)$ as the product of two functions,

$$\rho(x) = \exp(R + S)\exp(R - S) \tag{4.1}$$

with $R(x) = \ln\sqrt{\rho}$ a real function of x; for the time being, we have no argument to specify the real or complex character of S. In any case we may rewrite Eq. (4.1) in terms of two amplitudes of ρ:

$$\rho(x) = \psi_+\psi_- \tag{4.2a}$$

where

$$\psi_\pm = \exp(R \pm S). \tag{4.2b}$$

Now we must show that there exists indeed a function $S(x)$ by means of which one can calculate all the moments $\langle p^r\rangle$ averaged over the density $\rho(x)$, and consistent with the poissonian equations. For this purpose we introduce for the x-representation of p an operator $\mathcal{O}p_x(p) \equiv \widehat{p}$ acting upon one of the amplitudes, say ψ_-, such that

$$\psi_+\widehat{p}^r\psi_- = \int \delta(x - x(a))p^r(a)W(a)da \tag{4.3}$$

and the moments of p will then be given by integration over x:

$$\langle p^r\rangle = \int \psi_+ \hat{p}^r \psi_- dx = \int p^r(a) W(a) da. \tag{4.4}$$

To find the function S (if any) and the operator $\hat{p}$ that give a concrete meaning to these equations under the poissonian conditions (2.10), one must first answer a most important question, namely which is the x-representation of a poissonian. Since in general the poissonians are functions of x and p, they also become operators. Thus we write, for instance,

$$\psi_+ \mathcal{O}p_x(\langle x;p\rangle)\psi_- = \int \delta(x-x(a))\langle x;p\rangle W(a)da \tag{4.5}$$

and

$$\psi_+ \mathcal{O}p_x(\langle f(x);p\rangle)\psi_- = \int \delta(x-x(a))\langle f(x);p\rangle W(a)da \tag{4.6a}$$

$$\psi_+ \mathcal{O}p_x(\langle x;g(p)\rangle)\psi_- = \int \delta(x-x(a))\langle x;g(p)\rangle W(a)da \tag{4.6b}$$

for any smooth functions $f(x)$ and $g(p)$.

Now, from Eqs. (2.10) it follows that (for λ, λ' constants)

$$\langle \lambda f_1 + \lambda' f_2; p\rangle = \lambda\langle f_1;p\rangle + \lambda'\langle f_2;p\rangle, \quad \langle f;\lambda p\rangle = \lambda\langle f;p\rangle, \tag{4.7}$$

and so on. Moreover the dimensions of $\langle f;g\rangle$ are those of the product fg, and the interchange of f and g changes the sign of the poissonian:

$$\langle f;g\rangle = -\langle g;f\rangle. \tag{4.8}$$

In addition,

$$\langle f(x);p\rangle = i\hbar\partial f/\partial x, \tag{4.9a}$$

$$\langle x;g(p)\rangle = i\hbar\partial g/\partial p. \tag{4.9b}$$

By virtue of the linearity of the transformations (4.5,6), these properties should be transferred to the corresponding expressions in the x-representation.

Let us consider in particular a poissonian of the form $\langle x^r;p^s\rangle$; from the above properties, the x-representation $\mathcal{O}p_x(\langle x^r;p^s\rangle)$ must have the form

$$\mathcal{O}p_x(\langle x^r;p^s\rangle) = \sum_k A_{rs,k}(x^r\hat{p}^s)_k \tag{4.10}$$

where $(x^r\hat{p}^s)_k$ denotes the product $(x^{r_1}\hat{p}^{s_1}\ldots x^{r_N}\hat{p}^{s_N})$ for every possible partition $r_1+\ldots+r_N = r$ and $s_1+\ldots+s_N = s$, and $A_{rs,k}$ are numerical coefficients to be determined. The simplest cases may be readily solved. For $r = s = 1$ one gets:

$$\mathcal{O}p_x(\langle x;p\rangle) = A_1 x\hat{p} + A_2\hat{p}x.$$

Considering the symmetric roles of x and p and using Eq. (4.8):

$$\mathcal{O}p_x(\langle p;x\rangle) = A_1\hat{p}x + A_2 x\hat{p} = -A_1 x\hat{p} - A_2\hat{p}x.$$

Comparing terms we see that $A_1 = -A_2 \equiv A$. Now, according to $\langle x;p\rangle = i\hbar$, the value of the corresponding operator must be $i\hbar$, whence

$$A(x\hat{p} - \hat{p}x) = i\hbar. \tag{4.11}$$

The general solution of this equation is

$$A\hat{p} = -i\hbar\frac{\partial}{\partial x} + \gamma(x), \tag{4.12}$$

where $\gamma(x)$ is an unknown function of x. To determine A and γ we introduce this result into Eq. (4.4) with $r = 1$ and use Eqs. (4.2):

$$\langle p\rangle = \int \psi_+\hat{p}\psi_- dx = -\frac{i\hbar}{A}\int \psi_+\psi_-' dx + \frac{1}{A}\int \gamma(x)\rho(x)dx$$
$$= -\frac{i\hbar}{2A}\int \rho(\ln\psi_-/\psi_+)' dx + \frac{1}{A}\langle\gamma\rangle = \langle\frac{i\hbar}{A}S' + \frac{1}{A}\gamma\rangle.$$

The prime denotes derivation with respect to x. This result shows that both A and γ can be absorbed into the definition of the function S, through the transformation $(i\hbar S' + \gamma)/A \to \hbar S'$; using Eq. (4.2b) one thus arrives at

$$\mathcal{O}p_x(\langle x;p\rangle) = [x,\hat{p}] \tag{4.13}$$

$$\hat{p} = -i\hbar\frac{\partial}{\partial x} \tag{4.14}$$

$$\psi \equiv \psi_- = \exp(R+iS), \quad \psi^* = \psi_+ = \exp(R-iS), \tag{4.15}$$

where $[\cdot,\cdot]$ denotes a commutator and S is a real function of x, since $\langle p\rangle = \hbar\langle S'\rangle$ is real.

For $r = 2$ and $s = 1$, the solution of (4.10) gives

$$\mathcal{O}p_x(\langle x^2; p\rangle) = x^2\hat{p} - \hat{p}x^2 = [x^2, \hat{p}],$$

because the term $x\hat{p}x$ is excluded by the antisymmetry requirement (4.8). The same procedure leads more generally to

$$\mathcal{O}p_x(\langle f(x); p\rangle) = [f(x), \hat{p}] = i\hbar\frac{\partial f}{\partial x} \tag{4.16a}$$

$$\mathcal{O}p_x(\langle x; g(p)\rangle) = [x, g(\hat{p})] = i\hbar\frac{\partial g}{\partial \hat{p}} \tag{4.16b}$$

with $\hat{p}$ given by (4.14) and in accordance with Eqs (4.9).

Up to this point we have shown that if the density $\rho(x)$ is factored out according to Eqs. (4.2), then Eqs. (4.13)-(4.16) follow from the poissonian conditions as the rules for the description in the x-representation. On the other hand one may wonder whether these rules have the present set of poissonian equations as a necessary or unique background, or whether there can be in principle other (stochastic) equations that lead to the same (quantum) solution. For instance, the same solution could apparently be obtained if instead of $\langle x; p\rangle$ one considers any other bilinear, antisymmetric combination of x and p, say one involving second- instead of first-order derivatives. However, in such case the r.h.s. of Eqs. (4.9) would also contain second-order derivatives, and their x-representation would not be consistent with Eqs. (4.16). In conclusion, the structure of the stochastic equations leading to the quantum description is by no means arbitrary.

Now with Eq. (4.14) for $\hat{p}$ introduced into (4.4) the formula for the moments of p becomes

$$\langle p^r\rangle = \int \psi^*(-i\hbar\frac{\partial}{\partial x})^r\psi dx \tag{4.17a}$$

while the moments of x are of course simply

$$\langle x^r\rangle = \int \psi^* x^r \psi dx. \tag{4.17b}$$

There is, however, no recipe for the calculation of the average values of general phase space functions $f(x,p)$; the well-known problem of ordering of the operators arises. In particular, the properties of the poissonians do not determine uniquely the values of the coefficients $A_{rs,k}$ in Eq. (4.11) for $r, s \geq 2$. Take for

instance the case $r = s = 2$: the corresponding poissonian gives $\langle x^2; p^2\rangle = 4i\hbar xp$, and hence the same equality should hold among operators:

$$\mathcal{O}p_x(\langle x^2; p^2\rangle) = 4i\hbar\mathcal{O}p_x(xp),$$

but since there is no unique means to calculate $\langle xp\rangle$, $\mathcal{O}p_x(\langle x^2; p^2\rangle)$ remains undetermined. By choosing for instance

$$\mathcal{O}p_x(\langle x^2; p^2\rangle) = [x^2, \hat{p}^2] = 2i\hbar(x\hat{p} + \hat{p}x),$$

whence $\mathcal{O}p_x(xp) = \frac{1}{2}(x\hat{p} + \hat{p}x)$, one would be fixing an order of the operators for this case. More generally speaking, we see that the statistical method proposed does not allow to establish a rule of correspondence, *i.e.*, a rule that assigns a unique operator to every combination of x and p. The operator description in configuration space is therefore limited to a particular class of dynamical variables, namely those whose x-representation does not imply an ordering problem. This issue will be further discussed in Section VII.

V. The Passage to the p-representation

A statistical treatment entirely equivalent to that of last section can be developed in the p-representation. It is clear that in p-space, p is a c-variable, while x becomes an operator acting on a probability amplitude $\varphi(p)$. There is only a difference of sign in the operator, owing to the antisymmetry of $\langle x; p\rangle$. Since the procedure is completely parallel to the previous one, we just quote the end results:

$$\mathcal{O}p_p(\langle x; p\rangle) = [\hat{x}, p] \tag{5.1}$$

$$\mathcal{O}p_p(\langle f(x); p\rangle) = [f(\hat{x}), p] \tag{5.2a}$$

$$\mathcal{O}p_p(\langle x; g(p)\rangle) = [\hat{x}, g(p)] \tag{5.2b}$$

and

$$\hat{x} = i\hbar\frac{\partial}{\partial p}. \tag{5.3}$$

These operators act on the Hilbert space of the complex functions $\varphi(p)$, which are such that

$$\pi(p) = \varphi^*\varphi \tag{5.4}$$

and the moments of x and p are given by

$$\langle x^r \rangle = \int \varphi^* (i\hbar \frac{\partial}{\partial p})^r \varphi dp \tag{5.5a}$$

$$\langle p^r \rangle = \int \varphi^* p^r \varphi dp. \tag{5.5b}$$

Eqs. (4.17) and (5.5) are different expressions for the *same* quantities; hence the functions $\psi(x)$ and $\varphi(p)$ must be related. For instance, from (4.17a) and (5.5b) it follows that:

$$\int \varphi^* p^r \varphi dp = \int \psi^* (-i\hbar \frac{\partial}{\partial x})^r \psi dx. \tag{5.6}$$

The relation between ψ and φ can be readily established by resorting to the (second) Parseval theorem,[10] which states that for two functions $F(x)$ and $G(x)$ with respective Fourier transforms $\tilde{F}(k)$ and $\tilde{G}(k)$, the following equality holds:

$$\int \tilde{F}^*(k)\tilde{G}(k)dk = \int F^*(x)G(x)dx. \tag{5.7}$$

With $F(x) = \psi$ and $G(x) = (-i\hbar\partial/\partial x)^r \psi$, Eq. (5.6) has precisely the form of (5.7), whence $\varphi(p)$ is the Fourier transform of $\psi(x)$:

$$\varphi(p) = \tilde{\psi}(p) = (2\pi\hbar)^{-\frac{1}{2}} \int \psi(x) \exp(-ipx/\hbar) dx. \tag{5.8}$$

This is of course the well-known rule of transformation from the x- to the p-representation in quantum mechanics. The moments $\langle p^r \rangle$ are correctly given by Eq. (5.5b) and their values coincide with those obtained from Eq. (4.4) using for $\hat{p}$ the (quantum) expression Eq. (4.14); this confirms that the x-representation correctly gives all moments of x and of p.

A mathematical digression regarding the above procedure may be in order. The factorization of $\rho(x)$ which in Eq. (4.2) seems as a rather arbitrary step, finds mathematical support in an old theorem by Khinchin,[11] which states that a complex-valued function $M(k)$ is a moment generating (characteristic) function

if and only if it can be written in the form:

$$M(k) = \int g^*(y) g(y-k) dy, \tag{5.9}$$

with $g(y)$ a square-integrable function normalized to unity: $M(0) = \int g^*(y)g(y)dy = 1$ (the integrations are performed over the domain of y). Now, from this theorem it follows that the probability density of x can be written as follows:

$$\rho(x) = \frac{1}{2\pi} \int M(k) e^{-ikx} dk = \frac{1}{2\pi} \int g^*(y) g(y') e^{i(y'-y)x} dy dy'$$
$$= \tilde{g}^*(x)\tilde{g}(x),$$

which is precisely the form proposed in Eq. (4.15). Therefore, writing $\rho(x)$ as the product of the amplitude $\psi(x) = \tilde{g}(x)$ and its complex conjugate is equivalent to assuming the existence of a generating function for $\rho(x)$ —or for $\pi(p)$—, which is a weak assumption indeed. The Fourier transforms of these amplitudes serve then to construct the moment generating function $M(k)$, according to Eq. (5.9), and from this all moments of x are obtained by a Taylor series development, since $M(k)$ is the Fourier transform of $\rho(x)$:

$$M(k) = \sum_r \frac{(-k)^r}{r!} \int g^* \frac{\partial^r g}{\partial y^r} dy = \int \rho(x) e^{ikx} dx = \sum_r \frac{(ik)^r}{r!} \langle x^r \rangle.$$

In an entirely reciprocal form, the Fourier transforms of g and g^* serve to calculate the moments of k:

$$\sum_r \frac{(-s)^r}{r!} \int \tilde{g}^* \frac{\partial^r \tilde{g}}{\partial x^r} dx = \int M(k) e^{-iks} dk = \sum_r \frac{(-is)^r}{r!} \langle k^r \rangle,$$

whence $\hat{k} = -i(\partial/\partial x)$ and $[x, \hat{k}] = i$; with $\hat{p} = \hbar\hat{k}$ we thus recover Eq. (4.15) and $[x, \hat{p}] = i\hbar$. The Khinchin theorem is therefore directly applicable to our problem, with the Fourier variable k representing our physical variable p (up to a constant factor); this coincidence of the *mathematical* requirement $[x, \hat{k}] = i$ and the *physical* requirement $[x, \hat{p}] = i\hbar$ makes the present procedure highly suitable for SED (and for quantum theory, of course).

To conclude this section, we observe on comparing Eqs. (4.13, 16) with (5.1, 2) that the commutator expressions for the poissonians in the x- and p-representations coincide, and it is therefore unnecessary to specify the representation

when commutators are used. Thus, in the q-representation we write *e.g.*:

$$[\hat{x}, \hat{p}] = i\hbar. \tag{5.10}$$

Of course, the particular meaning of $\hat{x}$ and $\hat{p}$ depends on the specific representation under consideration.

VI. The Dynamical Equations

The poissonian equations of §II can now be transcribed to the q-representation. As for Eqs. (2.10), we already have the result of Eq. (5.10); in three-dimensional notation,

$$[\hat{x}_i, \hat{x}_j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}. \tag{6.1}$$

To transcribe the dynamical equations (2.11) we use the rules (5.2) rewritten in the q-representation, thus obtaining

$$i\hbar\dot{\hat{\mathbf{x}}} = [\hat{\mathbf{x}}, \widehat{H}], \quad i\hbar\dot{\hat{\mathbf{p}}} = [\hat{\mathbf{p}}, \widehat{H}] \tag{6.2}$$

where

$$\widehat{H} = \hat{\mathbf{p}}^2/2m + V(\hat{\mathbf{x}}), \tag{6.3}$$

according to Eq. (2.5), and $\hat{\mathbf{x}}, \hat{\mathbf{p}}, \widehat{H}$ are operators in the Hilbert space of the states of the system. Note that the passage to the q-representation has been straightforward because H does not contain products of $\mathbf{x}$ and $\mathbf{p}$ and the poissonians in Eqs. (2.11) are of the kind to which Eqs. (5.2) apply unambiguously. Thus we see that the Heisenberg equations are the q-representation of the SED system in the quantum regime under the radiationless approximation.

With the probability density $\rho(x)$ and the amplitudes $\psi(x)$, we have the necessary elements to go over to the Schrödinger picture in the x-representation. For this purpose we write ψ in the general form

$$\psi(x,t) = \sum c_n(t)\varphi_n(x,t), \tag{6.4}$$

where

$$\varphi_n = \theta_n(t)u_n(x) \tag{6.5}$$

are the solutions of the Heisenberg equations (6.2), written in terms of a complete basis of energy eigenfunctions $u_n(x)$, *i.e.*,

$$\widehat{H}_n u_n = E_n u_n, \tag{6.6}$$

and the functions θ_n are to be determined. Defining the matrix elements nm of $\hat{x}$ as

$$x_{nm} \equiv \int \varphi_n^* \hat{x} \varphi_m dx = \theta_n \theta_m \tilde{x}_{nm}, \tag{6.7}$$

where $\tilde{x}_{nm} \equiv \int u^* x u_m dx$, we obtain from Eqs. (6.2) and (6.6):

$$i\hbar(\dot{x})_{nm} = (E_m - E_n)x_{nm}.$$

But for a stationary ensemble, $(\dot{x})_{nm} = (d/dt)x_{nm}$, whence the solution of this equation gives $x_{nm} = \exp[(E_m - E_m)t/i\hbar]\tilde{x}_{nm}$. Comparing with Eq. (6.7), we get

$$\theta_n(t) = \exp(-iE_n t/\hbar)$$

and (6.4) becomes

$$\psi(\mathbf{x}, t) = \sum_n c_n(t)\exp(-iE_n t/\hbar)u_n(\mathbf{x}). \tag{6.8}$$

On the other hand, since particles are conserved, $\rho(\mathbf{x})$ must satisfy the continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{6.9}$$

where $\rho = \psi^*\psi$, and $\mathbf{v}(\mathbf{x})$ is the flow velocity given by by $\mathbf{p}(\mathbf{x})/m$. From Eqs. (4.15) and Eq. (4.16),

$$\mathbf{v} = \frac{\hbar}{m}\nabla S = -\frac{i\hbar}{2m}\left(\frac{\nabla\psi}{\psi} - \frac{\nabla\psi*}{\psi*}\right). \tag{6.10}$$

Introducing this and Eq. (6.8) into (6.9) one arrives after some simplification at

$$\sum_{m,n} u_n^* u_m \exp(i\omega_{nm}t)[d(c_n^* c_m)/dt] = 0.$$

It is easy to see (for instance, multiplying by an arbitrary $f(\mathbf{x})$ and integrating) that this equation implies necessarily $d(c_n^* c_m)/dt = 0$ for every n, m, whence all

c_n's are constant. Hence the general form for the probability amplitude is

$$\psi(\mathbf{x},t) = \sum_n c_n \exp(-iE_n t/\hbar) u_n(\mathbf{x}) \tag{6.11}$$

which, on using Eq. (6.6), is seen to be the solution of the Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = \widehat{H}\psi. \tag{6.12}$$

A statistical treatment of the quantum regime of the SED system has thus led us to the conventional descriptions of quantum mechanics. The state function ψ as it has been introduced here, clearly has a statistical character, and hence this theory unambiguously implies an ensemble interpretation of the Schrödinger equation. Also the averaging required to go over to the q-representation means that attention is given not to a single system but to a whole subensemble; thus, the notion of a trajectory followed by a single particle disappears from the present treatment. These are two of the most significant conceptual results of the theory, which at the same time point at important shortcomings of the description, avoidable probably only by a return to the a-representation.

VII. On the Phase Space Description

It is clear from the discussion of Sect. IV that the partial averaging involved in the transition to q-space has implied a loss of phase space information and hence the lack of a unique rule for calculating averages of arbitrary functions $f(x,p)$. However, at first glance it seems plausible to construct a statistical description for the SED system in phase space, by starting directly from the poissonian description and introducing the probability distribution:

$$R(x,p) = \int \delta(x - x(a))\delta(p - p(a))W(a)da \tag{7.1}$$

so that for any function F of x and p,

$$F(x,p)R(x,p) = \int \delta(x - x(a))\delta(p - p(a))F(a)W(a)da. \tag{7.2}$$

The integral of (7.2) over phase space would then give the average of $F(x,p)$. Integrating Eq. (7.1) over p-space and combining with (3.2), one obtains $\rho(x) = \int R(x,p)dp$, and alternatively, integrating over x one gets $\pi(p) = \int R(x,p)dx$,

which shows that the marginal distributions would be correctly obtained from such phase space distribution.

There is, however, no xp-representation for the poissonian equations. In such a representation the variables x and p must commute; the xp-representation of the poissonians must therefore be linear, homogeneous, antisymmetric forms involving x and p as c-numbers, not as operators acting on each other. The only solution that satisfies these requirements and in addition complies with Eqs. (2.10) and equations derived from them, like (4.9), is the transformation of the poissonians into Poisson brackets: $\langle f,g\rangle \to i\hbar[f,g]$, but under this rule of transformation the dynamical equations (2.10) become those of a classical hamiltonian system expressed in terms of Poisson brackets. Such a system of equations does not allow for a stochastic behaviour; the only possible statistical element in this case would be related to the initial conditions. Thus, the method used here to recover the quantum description from SED does not lead to a phase distribution.

Of course, a natural phase space description is contained in the original description of the SED system in a-space, but a different and probably more elaborate scheme will be necessary to get from it a (reduced) phase space representation for the particle —if at all possible—. What is clear is that in going to the q-representation the information about the x-p correlations is irretrievably lost, so that quantities like $\langle x^r p^s\rangle$ cannot be calculated without leaving quantum mechanics. This result is in agreement with Fine's theorem,[12] according to which for two noncommuting observables like $\hat{x}$ and $\hat{p}$ there is no joint probability distribution, even though there exists a well-defined joint distribution for x and p two random variables over a common probability space.

As discussed in Ref. 13, it may well be that the order of appearance of $\hat{x}$ and $\hat{p}$ in a dynamical function $f(x,p)$ depends on the meaning of this function or the way it is measured. —In fact, an entirely analogous observation can be made regarding the specificity of the order of appearance of photon creation and annihilation operators $\hat{a}^\dagger$, $\hat{a}$ in expressions related to absorption or emission of radiation, see Ref. 14—. Here we would have a physical explanation for the impossibility of finding a unique rule of correspondence and hence, phase space distribution, starting from the marginal distributions of quantum mechanics, even though x and p are simultaneously meaningful in the complete description in a-space.

This observation can throw some interesting light on the problem posed by the Bell theorem. As is well known, the Bell inequalities are very general conditions which any statistical, local system should obey;[15] however, quantum mechanics violates them. Without entering into the details, we can understand this as follows.

In quantum mechanics the Bell inequalities are applied to non-commuting observables, *i.e*, variables in the q-representation which have a direct relation to the possible outcomes of a specific measurement. To derive these inequalities, the correlation of two such observables is written in the form $\int A(\alpha,\lambda)B(\beta,\lambda)P(\lambda)d\lambda$ where λ is the set of hidden variables with probability density $P(\lambda)$ and α,β are parameters that specify the measurement being performed over a pair of dichotomic variables. Of course, in SED this expression for the correlation is entirely legitimate if $\{d\lambda\}$ is identified with $\{da\}$ and $P(\lambda)$ with $W(a)$; but in this case A and B are random variables in a-space, not observables in the q-representation which must be ordered in a definite way to correspond to specific measurements. Hence, quantum mechanical correlations and Bell correlations refer to different things, and only the latter are bound to satisfy Bell's inequalities, which are established for c-type variables. In this form, stochastic electrodynamics may be able to offer an answer to the long standing problem put forward by Bell.

References

1. Kalitsin N.S., JETP (URSS) **25**, 407 (1953); Adirovich E.I. and Podgoretskii M.I., JETP (URSS) **26**, 150 (1954); Braffort P., Spighel M. and Tzara C., C.R. Acad. Sci. (Paris) **239**, 157 (1954) (*erratum, ibid.* **239**, 925 (1954)); Braffort P. and Tzara C., *ibid.*, **239**, 1775 (1954). See also: Sokolov A.A. and Tumanov V.S., Zh. Eksp. i Teor. Fiz. **30**, 802 (1956); Soviet Phys. JETP **3**, 958 (1957); Schiller R. and Tesser H., Phys. Rev. **A3**, 2035 (1971).
2. Marshall T.W., Proc. Roy. Soc. (London) **276A**, 475 (1963).
3. Boyer T.H., in *Foundations of Radiation Theory and Quantum Electrodynamics*, edited by A. O. Barut, Plenum Press, N.Y., 1980. See also: Marshall T.W. and Claverie P., J. Math. Phys. **21**, 1819 (1980); Claverie P. and Soto F., J. Math. Phys. **23**, 753 (1982).

4. De la Peña L. in *Stochastic Processes Applied to Physics and Other Related Fields*, edited by B. Gómez *et al.*, World Scientific Publ., Singapore, 1983.
5. Marshall T.W. and Santos E., Ann. N. Y. Acad. Sci. **480**, 400 (1986); Europhys. Lett. **3**, 293 (1987); Found. Phys. **18**, 185 (1988); Phys. Rev. **A39**, 6271 (1989).
6. Cetto A.M. and de la Peña L., Found. Phys. Letters **4**, (1991)
7. De la Peña L. and Cetto A.M., Nuovo Cim. **92B**, 189 (1986). See also Rev. Mex. Fís. **37**, 17 (1991)
8. Santos E., *Stochastic theory versus quantum theory*, preprint GIFT 1975, Univ. of Santander.
9. Moore S.M., Physica **D6**, 211 (1983).
10. Butkov E., *Mathematical Physics*, Addison-Wesley, Mass., 1968, chapter 7.
11. Khinchin A., Bull. Univ. Moscow **1** (1937); Lukacs, E. *Characteristic Functions*, Ch. Griffin, London, 1970.
12. Fine A., J. Math. Phys. **23**, 1306 (1982).
13. Twareque S. and Doebner H.D., Phys. Rev. A **41**, 1199 (1990).
14. Cetto A.M. and de la Peña L., Found. Phys. **19**, 419 (1989).
15. See, *e.g.*, de Baere W., Adv. Electr. and Electron Phys. **68**, 245 (1986).

DETAILED BALANCE AND RADIATIVE CORRECTIONS IN STOCHASTIC ELECTRODYNAMICS

A. M. Cetto and L. de la Peña

Instituto de Física, UNAM
Apartado postal 20-364
01000 México 20, D.F.
Mexico

Starting from the results of the previous paper, we calculate the first-order correction to the position coordinates of the SED particle due to radiative effects, and we use this result to calculate the atomic absorption and emission rates and the shift of the energy levels; in all cases the theory gives exactly the formulas that one can derive using non relativistic quantum electrodynamics. The calculation leads also to a derivation of Planck's law within SED. Further, it is shown that the only spectrum for the background field that allows a state of equilibrium with matter is the one postulated by SED, namely, the zero-point field spectrum $\rho = \hbar\omega^3/2\pi^2c^3$. Finally, we show that the q-representation of the stochastic field coincides with the usual QED description of the radiation field in terms of creation and annihilation operators.

I. Introduction

In the preceding paper[1] we have shown that the new version of stochastic electrodynamics based on the poissonian equations leads in a quite natural form to the formalism of quantum mechanics, and assigns to it a well and uniquely defined statistical meaning, according to which the Heisenberg equations of motion correspond to the q-representation of the dynamical equations for the SED particle once the quantum regime has been attained. The transition from the poissonian equations to the statistical description in the q-representation was however rather formal and tends to hide the physical contents of the theory. This is of course undesirable, particularly if one is dealing with a theory one of whose main alleged advantages is physical transparency. We shall therefore begin by trying to recover a portion of the physical flavor of SED.

This second part of our work starts with a consideration about the general properties of the response of the (bound) particle to the field in the quantum regime, constructed on the basis of the results of Ref. 1. The rest of the paper is devoted to the study of some problems related to radiative corrections, that go beyond the scope of quantum mechanics into the domain of (non-relativistic, spinless) quantum electrodynamics; this extension is most natural here, because the fundamental equations of the theory contain the interaction of the particle with the fluctuating field from the very beginning. Of course, having reached the Heisenberg equations of motion, one is free to use the whole of the quantum mechanical formalism if desired; however, our intention is to get a fresh insight into the physics of the subject, and the use of direct methods based on the fundamental equations derived in the first part, is most appropriate for this purpose.

The most important radiative corrections are those which lead to an explanation and evaluation of atomic lifetimes. Since these are intimately related to electromagnetic transitions, we study the effects of a stochastic radiation field on the bound particle and calculate the Einstein A and B coefficients, to recover from them, using an already classical procedure that goes back to Einstein himself, the Planck distribution as the one that corresponds to the radiation field in thermal equilibrium.

By studying the conditions under which the atomic system may be in equilibrium with the vacuum field, we arrive at two important results: firstly, the spectral density of the zero-point field leading to equilibrium is uniquely determined as the one postulated in SED, namely, $\rho = \hbar\omega^3/2\pi^2c^3$, and secondly, equilibrium is attained at each separate field mode, not only globally for the whole energy interchange. Although such results are well known in quantum theory, they are nevertheless important for SED, since in previous attempts the theory has led to the Rayleigh-Jeans spectrum.[2] In addition, the theory predicts no spontaneous absorption, which is a highly nontrivial result, difficult to obtain without resorting to QED.

The same direct methods are also employed to calculate the effect of the radiative corrections on the atomic spectrum, *i.e.*, the Lamb shift; also here, the result coincides in detail with the predictions of QED. The paper ends with a discussion on the q-representation of the vacuum field, which is shown to corre-

spond to the description in terms of the creation and annihilation operators of QED.

II. The Linear Response to the Field

Once we have come to the conclusion that the SED system in the quantum regime is statistically described by quantum mechanics, we can use this knowledge to increase our understanding of the dynamical behaviour of the particle. Let us start by drawing a first general consequence that will be relevant to all the analysis that follows. As is well known, from the Heisenberg equations it follows that the matrix elements of x between two *stationary* states α, β, say, satisfy the differential equation (in one dimension, for simplicity):

$$\ddot{x}_{\alpha\beta} + \omega_{\alpha\beta}^2 x_{\alpha\beta} = 0 \tag{2.1}$$

with $\omega_{\alpha\beta}$ given by Bohr's formula

$$\omega_{\alpha\beta} = \frac{E_\alpha - E_\beta}{\hbar}, \tag{2.2}$$

which means that in quantum mechanics x behaves as a set of *independent* oscillators of frequencies $\omega_{\alpha\beta}$ and amplitudes $x_{\alpha\beta}$ (see also Eq. (6.7) of Ref.1). Note carefully that although each oscillator is linear, the equations that determine the amplitudes and frequencies are nonlinear, even for the harmonic oscillator (see for instance Eq. (2.5) below). Translated to the level of description of SED, since the independent elementary oscillators of the theory are the field variables $a_{\alpha\beta}(t)$ that oscillate with frequency $\omega_{\alpha\beta}$, this statistical behaviour is represented by the stochastic expression

$$x_\alpha^{(0)}(t) = \sum_\beta \tilde{x}_{\alpha\beta} a_{\alpha\beta}(t) + c.c. \tag{2.3}$$

where $\tilde{x}_{\alpha\beta}$ stands for the amplitude of the matrix element of x between states α and β, $x_{\alpha\beta} = \tilde{x}_{\alpha\beta} \exp\left(-i\omega_{\alpha\beta} t\right)$. In the following we shall omit for simplicity the index α referring to the state under consideration, and write

$$x^{(0)}(t) = \sum_\beta \tilde{x}_\beta a_\beta(t) + c.c.; \tag{2.3'}$$

below we shall comment on the superscript (0).

In writing Eq. (2.3) we are translating the already established quantum behaviour into a language appropriate to the stochastic description, without introducing further hypotheses. According to this equation, the SED system in equilibrium (in the so-called quantum regime) responds selectively and intensely to certain field modes —whose frequencies are precisely the frequencies of the possible transitions— and it is virtually insensitive to the rest of the spectrum, which merely contributes with a background "noise" to the motion. For the interpretation of Eq. (2.3) we recall that it has a statistical meaning only, expressing all what can possible happen to a system in quantum state α; the possibilities constitute a set of transitions to another state β, with statistical amplitudes $\tilde{x}_{\alpha\beta}$ and an associated change of energy $\hbar\omega_{\alpha\beta}$. It can be readily seen that Eq. (2.3) is the only possible expression —up to terms containing $|a|^2$, which is constant and hence dynamically uninteresting— that is consistent with the basic poissonian equation:

$$\langle x; p \rangle = i\hbar. \tag{2.4}$$

Specifically, introduction of Eq. (2.3) and its time derivative $p^{(0)} = m\dot{x}^{(0)}$ into Eq. (2.4) leads to

$$\sum_\beta \omega_\beta |\tilde{x}_\beta|^2 = \hbar/2m, \tag{2.5}$$

which is just the Thomas-Reiche-Kuhn sum rule of quantum mechanics. This result shows that the basic poissonian (2.4) is a very strong condition indeed —in fact, it is the stochastic version of the Dirac quantization rule.

Also, from (2.3) and recalling the statistical properties of the field variables:

$$\langle a_\beta \rangle = 0, \quad \langle a_\beta a_\gamma \rangle = \langle a_\beta^* a_\gamma^* \rangle = 0, \quad \langle a_\beta a_\gamma^* \rangle = \frac{1}{2}\delta_{\beta\gamma},$$

one obtains for the dispersions of x and p:

$$\sigma_x^2 = \sum |\tilde{x}_\alpha|^2, \quad \sigma_p^2 = m^2 \sum \omega_\alpha^2 |\tilde{x}_\alpha|^2.$$

Applying the Schwartz inequality to the product of these dispersions and using (2.5), one gets the result

$$\sigma_x^2 \sigma_p^2 \geq \hbar^2/4, \tag{2.6}$$

which assigns an unambiguous meaning to the Heisenberg inequality: in the quantum regime, the mechanical system is characterized by a stochastic behaviour of its dynamical variables; a lower limit to their dispersions is set by Eq. (2.6), independently of the specific state of motion, and this dispersion is a direct consequence of the effect of the fluctuating radiation field on the system.[3)]

The treatment of the poissonian dynamical equations, using the linear expression (2.3), is not as straightforward and presents some difficulties; a preliminary discussion of this problem can be found in Ref. 4. On the other hand, as we shall see in the following sections, Eq. (2.3) is remarkably useful for a simple derivation of results that normally require the use of QED.

III. Radiative Perturbation of the Quantum Mechanical Solution

The present theory provides a natural framework for the study of radiative effects on atomic systems. We recall that the starting point for the derivation of the Heisenberg equations of quantum mechanics was the set of poissonian dynamical equations written in the zero-order approximation:

$$\dot{\mathbf{x}} = -\frac{i}{\hbar}\langle \mathbf{x}; H\rangle, \qquad \dot{\mathbf{p}} = -\frac{i}{\hbar}\langle \mathbf{p}; H\rangle. \tag{3.1}$$

By recovering the neglected terms, which depend explicitly on the radiation field:

$$\dot{\mathbf{x}} = -\frac{i}{\hbar}\langle \mathbf{x}; H\rangle - \frac{e}{mc}\mathbf{A} \tag{3.2a}$$

$$\dot{\mathbf{p}} = -\frac{i}{\hbar}\langle \mathbf{p}; H\rangle - \frac{e}{c}x_j \nabla A_j \tag{3.2b}$$

we should therefore be in a position to calculate, via a perturbative procedure, the radiative corrections to the (unperturbed) quantum mechanical solution, to any order of approximation in the coupling constant e. We shall restrict our study to the lowest-order corrections.

With this purpose let us start by deriving a general formula for the first-order correction to the stochastic variable $\mathbf{x}(a)$. Consider a general dynamical equation of the form:

$$\ddot{x}_i = A_i + eB_i + \ldots \tag{3.3}$$

where A_i and B_i do not depend on e, and write the solution $\mathbf{x}(t)$ as a series in powers of e, $\mathbf{x} = \mathbf{x}^{(0)} + \mathbf{x}^{(1)} + \ldots$ Then the first-order term is a solution of the

equation $\ddot{x}_i^{(1)} = (\partial A_i/\partial x_j)x_j^{(1)} + eB_i$, with $\partial A_i/\partial x_j$ calculated at $\mathbf{x}^{(0)}$. Let us write it as

$$x_i^{(1)} = e\int_{-\infty}^{t} \mathcal{G}_{ij}(t,t')B_j(t')dt'; \tag{3.4}$$

then the Green function must satisfy the equation

$$\ddot{\mathcal{G}}_{ij} = \left(\frac{\partial A_i}{\partial x_k}\right)\mathcal{G}_{kj}$$

and the conditions $\mathcal{G}_{ij}(t,t) = 0$ and $\dot{\mathcal{G}}_{ij}(t,t')|_{t'=t} = \delta_{ij}$. The solution is

$$\mathcal{G}_{ij}(t,t') = m\left.\frac{\partial x_i(t)}{\partial p_j(t')}\right|_o$$

when $\mathbf{p} = m\dot{\mathbf{x}}$, as can be confirmed by direct substitution. Thus, $x_i^{(1)}$ becomes

$$x_i^{(1)} = em\int_{-\infty}^{t}\left.\frac{\partial x_i(t)}{\partial p_j(t')}\right|_o B_j(t')dt'. \tag{3.5}$$

Note that this expression does not depend explicitly on the force term A_i and hence its *form* is independent of the equations that govern the zero-order motion. Since in the quantum regime one can write

$$\frac{\partial x_i(t)}{\partial p_j(t')} = -\left[x_i(t), x_j(t')\right] = \frac{i}{\hbar}\langle x_i(t); x_j(t')\rangle, \tag{3.6}$$

with $x_i(t)$ given by Eq. (2.3) one gets

$$\frac{\partial x_i^{(0)}(t)}{\partial p_j^{(0)}(t')} = \frac{2}{\hbar}\delta_{ij}\sum_\beta |\tilde{x}_\beta^i|^2 \sin\omega_\beta(t-t'),$$

whence Eq. (3.5) transforms into

$$x_i^{(1)} = \frac{2e}{\hbar}\sum_\beta |\tilde{x}_\beta^i|^2 \int_o^\infty E_i(t-t')\sin\omega_\beta t' dt'. \tag{3.7}$$

This is a very useful formula for the first-order correction, since it is expressed in terms of the coefficients of the zero-order solution, $\tilde{x}_\beta$. In particular, we shall make use of it to derive general formulas for the (lowest-order) radiative corrections to the energy and the lifetime of the stationary states.

IV. Detailed Balance and the Spectrum of the Zero-point Field

From the stochastic Abraham-Lorentz equation of motion for the particle, that can be derived either from the Hamiltonian of the complete system (particle plus radiation) or else by combining the poissonian dynamical equations and using $\langle x_i; p_j \rangle = i\hbar\delta_{ij}$,

$$m\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + m\tau\dddot{\mathbf{x}} + e\mathbf{E} + \frac{e}{c}\dot{\mathbf{x}} \times \mathbf{B}, \tag{4.1}$$

it follows (after multiplying by $\dot{\mathbf{x}}$ and taking the average) that the mean energy of the particle varies according to

$$\frac{d\langle H\rangle}{d\,t} = m\tau\langle\dot{\mathbf{x}}\cdot\dddot{\mathbf{x}}\rangle + e\langle\dot{\mathbf{x}}\cdot\mathbf{E}\rangle. \tag{4.2}$$

Since $\tau = 2e^2/3mc^3$, the first term on the right hand side is at least of order e^2, whereas the second one may be of order e; hence, in equilibrium the contribution of lowest order in e from this second term should vanish. Writing $\mathbf{x}$ as a series in powers of e, $\mathbf{x} = \mathbf{x}^{(0)} + \mathbf{x}^{(1)} + \ldots$, we must have in equilibrium $\langle\dot{\mathbf{x}}^{(0)}\cdot\mathbf{E}\rangle = 0$ and Eq. (4.2) gives then to second order:

$$\frac{d\langle H\rangle}{dt} = m\tau\langle\dot{\mathbf{x}}^{(0)}\cdot\dddot{\mathbf{x}}^{(0)}\rangle + e\langle\dot{\mathbf{x}}^{(1)}\cdot\mathbf{E}\rangle. \tag{4.3}$$

Note that $\mathbf{x}$ given by Eq. (2.3) is indeed of zero order in e —this is the meaning of the superscript (0)—, and moreover, it satisfies the condition $\langle\dot{\mathbf{x}}^{(0)}\cdot\mathbf{E}\rangle = 0$, as a direct calculation shows (the term vanishes when integration over the wave vector $\mathbf{k}$ is performed, because of the isotropy of the vacuum field). As usual, we shall take $\mathbf{E}$ in the dipole approximation, $\mathbf{E} = \mathbf{E}(t)$. Then with $\dot{\mathbf{x}}^{(1)}$ given by the time derivative of Eq. (3.7), the absorption term of Eq. (4.3) becomes

$$\frac{\pi e^2}{\hbar}\sum_\beta \omega_\beta |\tilde{x}^i_\beta|^2 S(\omega_\beta),$$

where $S(\omega)$ is the power spectrum of the field, defined through

$$\langle E_i(t')E_j(t)\rangle = \delta_{ij}\int_o^\infty S(\omega)\cos\omega(t-t')d\omega. \tag{4.4}$$

On the other hand, from Eq. (2.3) one obtains for the radiation term:

$$-m\tau\sum_\beta \omega_\beta^4 |\tilde{x}^i_\beta|^2$$

and the energy rate equation becomes thus:

$$\frac{d\langle H\rangle}{d\,t} = \sum_\beta \left[\frac{\pi e^2}{\hbar}\omega_\beta S(\omega_\beta) - m\tau\omega_\beta^4\right]|\tilde{\mathbf{x}}_\beta|^2. \tag{4.5}$$

This equation shows not only that equilibrium can be attained between the atom and radiation, but also that detailed balance holds at every frequency if and only if the power spectrum is

$$S_o(\omega) = \frac{2\hbar\omega^3}{3\pi c^3} \tag{4.6}$$

(with $\omega > 0$); this is precisely the spectrum of the zero-point field. It has been possible to arrive at this universal result, irrespective of the details of the dynamical system, because the linear response implies a non-mixing of the field frequencies, and thus each mode can reach equilibrium independently of the others. We see therefore that the well-known property of detailed balance in quantum systems is closely connected to the linear response described by Eq. (2.3). Further, it becomes clear from this derivation that the failure of previous attempts to determine $S_o(\omega)$ from the detailed-balance condition —which, as mentioned before, had led to the Rayleigh-Jeans spectrum instead[2]— is associated to the use of a perturbative treatment of the dynamics, in which $\mathbf{x}^{(0)}$ is taken as the solution of the *classical* equations of motion instead of the poissonian equations. In this connection it is opportune to recall a recent paper by Cole,[5] in which the correct spectrum for the zero-point field is derived with the help of thermodynamic arguments involving van der Waals forces.

V. Spontaneous and Induced Electromagnetic Transitions

Let us now proceed to investigate the problem of absorption and emission of radiation by an atom (or any bound system) in an external electromagnetic field.[6] With this purpose we introduce Eq. (4.6) into the energy rate equation (4.5), and rewrite this in terms of the spectral energy density of the field, $\rho(\omega) = 3S(\omega)/4\pi$, obtaining

$$\frac{d\langle H\rangle}{d\,t} = \sum_\beta \frac{4\pi^2 e^2}{3\hbar}\omega_\beta\left[\mathrm{sign}(\omega_\beta)\rho(\omega_\beta) - \rho_o(\omega_\beta)\right]|\tilde{\mathbf{x}}_\beta|^2.$$

Since ω_β may be positive or negative, it is convenient to add an index to the coefficients $\tilde{x}$ to indicate the sign of the corresponding ω_β; thus,

$$\frac{d\langle H\rangle}{dt} = \sum_\beta \frac{4\pi^2e^2}{3\hbar}\omega_\beta \left[\rho(\omega_\beta) - \rho_o(\omega_\beta)\right] |\tilde{\mathbf{x}}_\beta^{(+)}|^2$$
$$- \sum_\beta \frac{4\pi^2e^2}{3\hbar}|\omega_\beta| \left[\rho(\omega_\beta) + \rho_o(\omega_\beta)\right] |\tilde{\mathbf{x}}_\beta^{(-)}|^2. \tag{5.1}$$

The terms of the first sum are different from zero only when the field is excited ($\rho > \rho_o$); they represent the energy absorption that takes the particle to a state of higher energy. The condition $\rho > \rho_o$ means that no spontaneous absorption can occur. The terms of the second sum are always negative: they represent the energy loss that occurs for $\omega_\beta < 0$, even in absence of external field, and refer to the processes of emission.

Note that for absorption to occur, one must have (in addition to $\rho > \rho_o$) at least one $|\tilde{\mathbf{x}}_\beta^{(+)}| \neq 0$ for some ω_β. On the other hand, if no $|\tilde{\mathbf{x}}_\beta^{(-)}|$ is different from zero (*i.e.*, if all ω_β are positive), no emission can occur; the system is then in its lowest energy state. Detailed balance holds in particular when both particle and field are in their ground state. According to these results, then, no self-ionization of the H atom occurs, as opposed to the predictions of the previous version of SED (see, *e.g.*, Ref. 7).

To obtain explicit formulas for the Einstein coefficients, we separate the absorption and emission terms in Eq. (5.1) as follows (omitting the index β, for simplicity):

$$\frac{d\langle H\rangle}{dt} = W_{\text{abs}} - W_{\text{emi}}, \tag{5.2}$$

with

$$W_{\text{abs}} = \frac{4\pi^2e^2}{3\hbar}\sum \omega(\rho(\omega) - \rho_o(\omega))|\tilde{\mathbf{x}}^{(+)}|^2, \tag{5.3a}$$
$$W_{\text{emi}} = \frac{4\pi^2e^2}{3\hbar}\sum \omega(\rho(\omega) + \rho_o(\omega))|\tilde{\mathbf{x}}^{(-)}|^2. \tag{5.3b}$$

In terms of the difference $\rho_e = \rho - \rho_o$ which represents the part of the spectral density which is due to an external field, Eqs. (5.3) become

$$W_{\text{abs}} = \sum \frac{4\pi^2e^2}{3\hbar}\omega\rho_e(\omega)|\tilde{\mathbf{x}}^{(+)}|^2, \tag{5.4a}$$
$$W_{\text{emi}} = \sum \frac{4\pi^2e^2}{3\hbar}\omega\rho_e|\tilde{\mathbf{x}}^{(-)}|^2 + \sum \frac{8\pi^2e^2}{3\hbar}\omega\rho_o|\tilde{\mathbf{x}}^{(-)}|^2. \tag{5.4b}$$

By comparing these expressions with the general formulas for the energy absorption and emission rates in terms of Einstein coefficients:

$$W_{\text{abs}} = \sum \Delta E(B^{(+)}\rho_e + A^{(+)}),$$
$$W_{\text{emi}} = \sum \Delta E(B^{(-)}\rho_e + A^{(-)}), \tag{5.5}$$

($A^{(+)}$ and $A^{(-)}$ are the probabilities per unit time of spontaneous absorption and emission, respectively, and $B^{(+)}$ and $B^{(-)}$ the corresponding probabilities for the transitions induced by the external field), we get the following results:

$$A^{(+)} = 0, \tag{5.6a}$$
$$A^{(-)} = \frac{8\pi^2e^2\omega}{3\hbar\Delta E}\rho_o|\tilde{\mathbf{x}}^{(-)}|^2, \tag{5.6b}$$
$$B^{(+)} = \frac{4\pi^2e^2\omega}{3\hbar\Delta E}|\tilde{\mathbf{x}}^{(+)}|^2, \tag{5.6c}$$
$$B^{(-)} = \frac{4\pi^2e^2\omega}{3\hbar\Delta E}|\tilde{\mathbf{x}}^{(-)}|^2. \tag{5.6d}$$

With ΔE given by Eq. (2.2), these equations coincide with the corresponding formulas of QED.[8)]

From Eqs. (5.3) we observe that in the absence of external field, the terms due to the vacuum field and to radiation reaction cancel each other in W_{abs} whereas they contribute an equal amount to W_{emi}, thus giving rise to the factor 2 that appears in the above formula $A^{(-)} = 2\rho_o B^{(-)}$. The question of whether the spontaneous transitions are due to the action of the vacuum field or to radiation reaction has been a matter of long discussions (a review on this subject is given in Ref. 9; a discussion of the physical meaning of the quantities $\rho+\rho_o$ and $\rho-\rho_o$ may be found in Ref. 10); here we have obtained a clearcut answer to this question.

Finally it is illustrative to rewrite the detailed-balance equation $W_{\text{abs}}(\omega) = W_{\text{emi}}(\omega)$ as a condition on the power spectrum:

$$S(\omega) = \frac{m\tau\hbar}{\pi c^2}|\omega|^3 \frac{|\tilde{\mathbf{x}}^{(+)}|^2 + |\tilde{\mathbf{x}}^{(-)}|^2}{|\tilde{\mathbf{x}}^{(+)}|^2 - |\tilde{\mathbf{x}}^{(-)}|^2}.$$

Here $|\tilde{\mathbf{x}}^{(+)}|^2$ and $|\tilde{\mathbf{x}}^{(-)}|^2$ refer to upward and downward transitions from the same state α and with the same (absolute) frequency ω. When the system is in an excited state, generally $\tilde{\mathbf{x}}^{(-)} \neq 0$ and $\tilde{\mathbf{x}}^{(+)} = 0$ for at least one frequency, which means that *no* power spectrum exists that can maintain a single excited atomic state in detailed balance with the radiation field. An important exception to this

rule is the harmonic oscillator (because $\tilde{\mathbf{x}}^{(-)}$ and $\tilde{\mathbf{x}}^{(+)}$ are different from zero for the same $\omega = \omega_o$): its nth excited state is in equilibrium with a field of power spectrum

$$S = S_o(\omega)\frac{|\tilde{\mathbf{x}}^{(+)}_{n,n+1}|^2 + |\tilde{\mathbf{x}}^{(-)}_{n,n-1}|^2}{|\tilde{\mathbf{x}}^{(+)}_{n,n+1}|^2 - |\tilde{\mathbf{x}}^{(-)}_{n,n-1}|^2} = S_o(\omega)(1+2n)$$

(to write the second equality we have used well-known quantum results); this expression corresponds to what is usually called a field state of n photons (of frequency $\omega = \omega_o$).

VI. On Planck's Distribution and the Meaning of Quantization

We have just seen that a single excited atomic state cannot be in equilibrium with a given radiation field. Now consider an ensemble of atoms in different states, in equilibrium with such field; some atoms will undergo absorption while others undergo emission. Let the average equilibrium population of state α be N_α and so on; the master equation for N_α is

$$\frac{dN_\alpha}{d\,t} = \sum_\beta \left[N_\beta P_{\beta\to\alpha} - N_\alpha P_{\alpha\to\beta}\right]$$

where $P_{\alpha\to\beta}$ is the probability per unit time of a transition $\alpha \to \beta$. Under the assumption of detailed equilibrium this equation should be satisfied term by term (assuming non-degenerate transition frequencies), hence $N_\beta P_{\beta\to\alpha} = N_\alpha P_{\alpha\to\beta}$. In our previous notation this equation reads

$$N_\beta W_{\text{emi}}(\beta\to\alpha) = N_\alpha W_{\text{abs}}(\alpha\to\beta)$$

because $W(\alpha\to\beta) = |\Delta E_{\alpha\beta}|P_{\alpha\to\beta}$. Introducing Eqs. (5.5) and (5.6) we obtain a general formula for the spectrum of the field that is in equilibrium with the populations N_α and N_β:

$$\rho = \rho_o + \rho_e = \rho_o\frac{N_\alpha + N_\beta}{N_\alpha - N_\beta}. \tag{6.1}$$

In particular, in thermodynamic equilibrium the atomic populations follow the Maxwell-Boltzmann distribution, so that $N_\alpha/N_\beta = \exp(E_\beta - E_\alpha)\beta$ with $\beta =$ $1/(kT)$; then Eq. (6.1) gives

$$\rho = \frac{\hbar\omega^3}{\pi^2c^3}\left[\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega}-1}\right] \tag{6.2}$$

with $\omega = |\omega_{\beta\alpha}| = |E_\beta - E_\alpha|/\hbar$. For other works concerning the derivation of Planck's law from SED, see the literature cited in Refs. 11, 12, and more recent work in Refs. 13-15.

In the above textbook calculation the Planck distribution followed from the demand of conservation of energy during transitions. We recall that in his original paper,[16] Einstein showed that this result for $\rho(\omega;T)$ follows also from the demand of conservation of momentum, if $\hbar\omega/c$ is the momentum exchanged with the field in a definite but arbitrary direction. A treatment similar to the above one,[4] considering the stepwise exchange of *angular* momentum, leads to the same results for A, B and ρ, though expressed in more detail. The general conclusion is that Planck's law is consistent with the (detailed) statistical balance of energy, linear momentum and angular momentum during atomic transitions.

According to the above discussion, in every transition only one mode of the radiation field is involved, and this mode absorbs or delivers the whole of the exchanged energy in an *elementary act of interaction* —as it used to be called, very appropriately—. This picture of a transition process as the result of such an elementary act contains the essence of the quantum description of the matter-field interaction, even though it is made in terms of continuous quantities. This approach helps us gain some understanding on the mechanism of matter quantization: imagine, to start with, an atomic system in its ground state, which is the stationary state in equilibrium with the zero-point field. Addition of an external field of frequency $\omega_{o\beta}$ can induce an excitation, and this excitation takes the system to a new state with precisely an extra energy $\hbar\omega_{o\beta}$. The quantum states are the stationary states in the so-called quantum regime, that are mutually connected by means of such transitions.

VII. The Lamb Shift

We now proceed to the calculation of the (lowest-order) radiative correction to the atomic energy, *i.e.* the Lamb shift, using the results of Section III for this purpose. In general, the radiative corrections to the energy of the atom (or any

bound system) will be given by the difference between the exact Hamiltonian

$$H(A) = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + V(\mathbf{x})$$

and the zero-order Hamiltonian

$$H(0) = \frac{1}{2m}\left(\mathbf{p}^{(0)}\right)^2 + V\left(\mathbf{x}^{(0)}\right).$$

With $\mathbf{x}$ and $\mathbf{p}$ expressed as a series in powers of e, one thus gets to second order:

$$\begin{aligned}\delta E &= \langle H(A) - H(0)\rangle \\ &= \left\langle \frac{(\mathbf{p}^{(1)})^2}{2m} - \frac{e}{mc}\mathbf{p}^{(1)}\cdot\mathbf{A} + \frac{e^2}{c^2}\mathbf{A}^2 + \frac{1}{2}x_i^{(1)}x_j^{(1)}\frac{\partial^2 V}{\partial x_i \partial x_j}\right\rangle \end{aligned} \tag{7.1}$$

since the terms $\mathbf{p}^{(0)}\cdot\mathbf{p}^{(1)}, \mathbf{p}^{(0)}\cdot\mathbf{A}$ and $\mathbf{x}^{(1)}\cdot\nabla V$ average to zero because of the isotropy of the vacuum field. Further,

$$\frac{1}{2m}\left(\mathbf{p}^{(1)}\right)^2 - \frac{e}{mc}\mathbf{p}^{(1)}\cdot\mathbf{A} = \frac{1}{2}\mathbf{p}^{(1)}\cdot\dot{\mathbf{x}}^{(1)} - \frac{e}{2mc}\mathbf{p}^{(1)}\cdot\mathbf{A},$$

whence Eq. (7.1) becomes

$$\delta E = \left\langle \frac{1}{2}\dot{\mathbf{x}}^{(1)}\cdot\mathbf{p}^{(1)} + \frac{1}{2}x_i^{(1)}x_j^{(1)}\frac{\partial^2 V}{\partial x_i \partial x_j}\right\rangle + \frac{e^2}{c^2}\langle\mathbf{A}^2\rangle - \frac{e}{2mc}\langle\mathbf{p}^{(1)}\cdot\mathbf{A}\rangle. \tag{7.2}$$

The first two terms on the right hand side combine to give rise to the Darwin term and they therefore represent a contribution to the fine structure of the atomic spectrum —a detailed study of these terms is under process—. The third term, which is sometimes called the free-particle Lamb shift, is independent of the dynamics; the last term is therefore the one responsible for the bound-particle contribution to the Lamb shift; *i.e.*, the Lamb shift proper is given by

$$\delta E_L = -\frac{e}{2mc}\langle\mathbf{p}^{(1)}\cdot\mathbf{A}\rangle. \tag{7.3}$$

To proceed with the calculation we introduce in this equation:

$$\mathbf{p}^{(1)} = m\dot{\mathbf{x}}^{(1)} + \frac{e}{e}\mathbf{A}$$

with $\dot{\mathbf{x}}^{(1)}$ given by the time derivative of Eq. (3.7):

$$\dot{x}_i^{(1)} = \frac{2e}{\hbar}\sum_\beta \omega_\beta |\tilde{x}_\beta^i|^2 \int_0^\infty E_i(t-t')\cos\omega_\beta t' dt', \tag{7.4}$$

whence

$$\delta E_L = \frac{e^2}{\hbar c^2}\sum_\beta \omega_\beta |\tilde{x}_\beta^i|^2 \int_0^\infty \langle\mathbf{A}(t')\cdot\dot{\mathbf{A}}(t-t')\rangle\cos\omega_\beta t' dt' - \frac{e^2}{2mc^2}\langle\mathbf{A}^2\rangle.$$

An integration by parts, taking into account the stationarity of the vacuum field, gives

$$\delta E_L = \frac{e^2}{\hbar c^2}\sum_\beta \omega_\beta |\tilde{x}_\beta^i|^2 \left[\langle\mathbf{A}^2\rangle - \omega_\beta \int_0^\infty \langle\mathbf{A}(t)\cdot\mathbf{A}(0)\rangle \sin\omega_\beta t\ dt\right] - \frac{e^2}{2mc^2}\langle\mathbf{A}^2\rangle.$$

The two terms containing $\langle A^2\rangle$ cancel each other by virtue of the sum rule (2.5). Further, from Eqs. (4.4) and (4.6) we have

$$\langle\mathbf{A}(t)\cdot\mathbf{A}(0)\rangle = \frac{2\hbar}{\pi c}\int_0^\infty \omega\cos\omega t\ d\omega, \tag{7.5}$$

and we are thus left with

$$\delta E_L = -\frac{2e^2}{\pi c^3}\sum_\beta \omega_\beta^2 |\tilde{x}_\beta^i|^2 \int_0^\infty \omega d\omega \int_0^\infty \cos\omega t \sin\omega_\beta t\ dt.$$

Integration over the time and the frequency variables (up to a cutoff frequency ω_c, so as to avoid the logarithmic divergence) leads to the result

$$\delta E_L = \frac{2e^2}{3\pi c^3}\sum_\beta |\tilde{\mathbf{x}}_\beta|^2 \omega_\beta^3 \ln\left|\frac{\omega_c}{\omega_\beta}\right|, \tag{7.6}$$

which is precisely Bethe's formula for the atomic Lamb shift.[17] It is interesting to note that the structure of this formula is essentially determined by the expression for the first-order correction $\mathbf{x}^{(1)}$ and the statistical properties of the vacuum field. The present result is of course applicable to any bound system, and hence it generalizes the results that had been obtained within SED for the harmonic oscillator (see the literature cited in Ref. 12 and more recent results in Ref. 18 and the last paper of Ref. 8). Also, notice that in the above theory one can easily substitute any spectrum $\rho(\omega)$ for $\rho_o(\omega)$, thus allowing the calculation of the effects of external fields on the Lamb shift; results along this line have been presented in Ref. 18.

VIII. On the Quantum Description of the Radiation Field

In the above calculations of the radiative corrections we have considered that the vector potential **A** is a function of the original stochastic variables a, a^*, which, however, have already been partly averaged in the transition from the poissonian equations to quantum mechanics. A full treatment of the radiative effects would actually require the use of the exact stochastic (poissonian) description and the transition to the complete q-representation for particle *and* field variables. A practical (though approximate) solution to this situation is obtained by working first with the zero-order equations in the q-representation, to later introduce the radiation terms as perturbations to these approximate equations; but then we must learn how to handle the radiation terms in this representation.

To zero order (neglecting the field radiated by the particle) the field fluctuations are independent of x and p and hence, this field cannot be represented in the q-space of the particle variables; so we introduce a new space for its representation. For this purpose we recall the canonical variables associated to the field mode α:

$$q_\alpha = (\hbar/2\omega_\alpha)^{1/2} i(a_\alpha - a_\alpha^*) \tag{8.1a}$$

$$p_\alpha = (\hbar\omega_\alpha/2)^{1/2}(a_\alpha + a_\alpha^*). \tag{8.1b}$$

Since $\langle a_\alpha; a_\alpha^* \rangle = 1$, the poissonian of the canonical variables is

$$\langle q_\alpha; p_\alpha \rangle = i\hbar. \tag{8.2}$$

The q_α- (or p_α-) representation can be constructed by following a procedure entirely similar to that used in the previous paper: for every mode α we introduce a probability (the index α will be omitted wherever this omission does not create confusion):

$$d\Gamma_q = \int_{dv_q} W(a, a^*) da da^*.$$

The probability density in q_α-space will then be given by $P_q dq = d\Gamma_q$, and for arbitrary functions of q_α we shall have

$$F(q) P_q dq = \int_{dv_q} F_q(a, a^*) W(a, a^*) da da^*.$$

The operator associated to the variable p_α in this representation, is obtained by introducing the poissonian condition (8.2); as a result one gets

$$\hat{p} = -i\hbar \frac{\partial}{\partial q} \tag{8.3}$$

and

$$\langle \hat{p}^r \rangle = \int \xi^* \left(-i\hbar \frac{\partial}{\partial q} \right)^r \xi dq,$$

where $\xi(q)$ is the probability amplitude associated to P_q:

$$P_q dq = \xi^*(q)\xi(q).$$

By combining Eqs. (8.1) and (8.3) we obtain for the q-representation of a_α and a_α^*:

$$\hat{a} = -i(2\hbar\omega)^{-1/2}(\hbar\partial/\partial q + \omega q) \tag{8.4a}$$

$$\hat{a}^\dagger = i(2\hbar\omega)^{-1/2}(-\hbar\partial/\partial q + \omega q). \tag{8.4b}$$

These operators satisfy the commutation relation:

$$[\hat{a}, \hat{a}^\dagger] = 1 \tag{8.5}$$

and in terms of them, the hamiltonian of the zero-point field mode α becomes

$$\widehat{H}_R = \frac{1}{2}(\hat{p}^2 + \omega^2 q^2) = \frac{1}{2}\hbar\omega(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}). \tag{8.6}$$

The vacuum field potential $\mathbf{A}_o$ has thus been transformed into an operator acting on the state vectors $\xi(q_\alpha)$:

$$\widehat{\mathbf{A}}_o = \sum \left(\frac{2\pi\hbar c^2}{V\omega_\alpha} \right)^{1/2} \hat{\boldsymbol{\epsilon}}_\alpha \left[\hat{a}_\alpha \exp[-i(\omega_\alpha t - \mathbf{k}_\alpha \cdot \mathbf{x})] + \hat{a}_\alpha^\dagger \exp[i(\omega_\alpha t - \mathbf{k}_\alpha \cdot \mathbf{x})] \right]. \tag{8.7}$$

The variable $\mathbf{x}$ appearing in the exponentials acts on the particle state vectors $\psi(\mathbf{x})$, hence it commutes with $\hat{a}$ and $\hat{a}^\dagger$. On the other hand, the field radiated by the particle is given by

$$\mathbf{A}_r = \frac{2\pi i e}{V} \sum \frac{1}{\omega_\alpha} \hat{\boldsymbol{\epsilon}}_\alpha \int_0^t \dot{\mathbf{x}}(t') \cdot \hat{\boldsymbol{\epsilon}}_\alpha \exp\left[-i\omega_\alpha (t - t') - \mathbf{k}_\alpha \cdot (\mathbf{x} - \mathbf{x}') \right] dt' + c.c.$$

This expression is a nonseparable function of both particle variables $\mathbf{x}$ and $\mathbf{p}$, hence we cannot transcribe it to the x-representation because of the problem of order of the operators. However, in a non-relativistic treatment like the present one it is legitimate to introduce the long-wavelength (or dipole) approximation, which implies neglecting the x-dependence in the exponentials; the formula for $\mathbf{A}_r$ is then simplified, and the transcription to the x-space is made possible, giving

$$\widehat{\mathbf{A}}_r = -\frac{2\pi i\epsilon}{V}\sum\frac{1}{\omega_\alpha}\hat{\boldsymbol{\epsilon}}_\alpha\int_0^t \hat{\mathbf{x}}(t')\cdot\hat{\boldsymbol{\epsilon}}_\alpha \exp\left[-i\omega_\alpha(t-t')\right]dt' + c.c. \qquad (8.8)$$

Now we have all the elements to transcribe the complete poissonian dynamical equations to their operator form:

$$i\hbar\dot{\hat{\mathbf{x}}} = [\hat{\mathbf{x}}, \widehat{H}] - \frac{ie\hbar}{mc}(\widehat{\mathbf{A}}_o + \widehat{\mathbf{A}}_r) \qquad (8.9a)$$

$$i\hbar\dot{\hat{\mathbf{p}}} = [\hat{\mathbf{p}}, \widehat{H}]. \qquad (8.9b)$$

In the second equation the radiative term has been omitted, since it is of order $(\dot{x}/c)$ and hence would give rise to a relativistic correction. Note that the Hilbert space of Eqs. (8.9) is now the product of spaces of the particle and the field state vectors.

Since $[\hat{\mathbf{x}}, \widehat{H}] = i\hbar\hat{\mathbf{p}}/m$ and $[\hat{\mathbf{p}}, \widehat{H}] = i\hbar\hat{\mathbf{F}}$, the time derivative of Eq. (8.9a) combined with (8.9b) gives

$$m\ddot{\hat{\mathbf{x}}} = \mathbf{F}(\hat{\mathbf{x}}) + e(\hat{\mathbf{E}}_o + \hat{\mathbf{E}}_r) \qquad (8.10)$$

where $\hat{\mathbf{E}}_{o,r} = -(1/c)(\partial\widehat{\mathbf{A}}_{o,r}/\partial t)$. With (8.8) for the radiated term one gets as usual the radiation reaction force $m\tau\dddot{\mathbf{x}}$ plus a mass correction. Renormalizing the mass, Eq. (8.10) becomes thus

$$m\ddot{\hat{\mathbf{x}}} = \mathbf{F}(\hat{\mathbf{x}}) + m\tau\dddot{\hat{\mathbf{x}}} + e\hat{\mathbf{E}}_o, \qquad (8.11)$$

which is the quantum version of the Abraham-Lorentz equation. Dalibard and collaborators[19] have shown in a very clear way (entirely within non-relativistic quantum electrodynamics) that Eq. (8.11) is the dynamical equation for the quantum system. We therefore conclude that the proposed method to deal with the radiative terms in the poissonian equations is equivalent to the introduction of second quantization into quantum theory. It is not surprising, then, that the

radiative predictions of the theory coincide with those of QED. The ensuing formalism is of course not able to furnish a description of the transient period before the onset of the quantum regime nor details of the orbital motion or the transitions.

Before closing this section we wish to recall that Eq. (8.11) was long ago proposed by Sokolov and Tumanov[20] as the mechanism of quantization of matter, with $\hat{\mathbf{E}}_o$ the second quantized radiation field acting on an otherwise classical particle. They showed for the harmonic oscillator that with $\hat{a}, \hat{a}^\dagger$ obeying the usual commutation rule (8.5), x and p in Eq. (8.11) become operators satisfying $[\hat{x}, \hat{p}] = i\hbar$; but the demonstration was not extended to more complex systems and thus the proposal had little impact despite its intrinsic value. It is to be noticed that in the present theory, on the other hand, all stochastic variables (for field and particle) have their own representation in q-space.

The authors wish to acknowledge stimulating discussions with F. Guerra and E. Santos on some aspects of this and the companion paper.

References

1. De la Peña L. and Cetto A.M., *A fundamental relation between stochasticity and quantization* (previous paper).
2. See, *e.g.*, Blanco R., Pesquera L. and Santos E., Phys. Rev. **D27**, 1254 (1983); *ibid.* **D29**, 2240 (1984); Santos E. in *Stochastic Processes Applied to Physics*, L. Pesquera and M.A. Rodríguez, eds., World Scientific, Singapore, 1985.
3. This meaning for the Heisenberg inequality has always been implicit in the postulates of SED; an explicit proposal in this direction can be found in Santos E., *Stochastic theory versus quantum theory*, GIFT preprint, 1975.
4. De la Peña L. and Cetto A. M., *Reformulation of stochastic electrodynamics for an explanation of quantum phenomena*, preprint IFUNAM 1989.
5. Cole D.C., Phys. Rev. **A42**, 1847 (1990).
6. A brief account of the material contained in this and the next section is given in Alcubierre M., Cetto A.M., de la Peña L. and Lozano N. S., *Detailed balance and radiative transitions according to the new approach to stochastic electrodynamics*, preprint IFUNAM 1988.

7. Marshall T.W. and Claverie P., J. Math. Phys. **21**, 1819 (1980); Claverie P., Pesquera L. and Soto F., Phys. Lett. **80A**, 113 (1980); Claverie P. and Soto F., J. Math. Phys. **23**, 753 (1982); Pesquera L., thesis, Université de Paris VI, 1980 (unpublished).
8. A derivation of the A coefficient for the SED harmonic oscillator can be found in Marshall T.W., Izv. Vuzov (Fizika) **11**, 34 (1968); more recent results can be seen in de la Peña L. and Cetto A.M., J. Math. Phys. **20**, 469 (1979); França H.M. and Marshall T.W., Phys. Rev. **A38**, 3258 (1988).
9. Milonni P. W., Phys. Rep. **25**:1, 1 (1976) and references therein. See also Milonni P. W., in *Foundations of Radiation Theory and Quantum Electrodynamics*, A. O. Barut, editor, Plenum Press, New York, 1980.
10. Cetto A.M. and de la Peña L., Found. Phys. **19**, 419 (1989).
11. Boyer T. H., in *Foundations of Radiation Theory and Quantum Electrodynamics*, edited by A. O. Barut, Plenum Press, New York, 1980.
12. De la Peña L., in *Stochastic Processes Applied to Physics and Other Related Fields*, edited by B. Gómez *et al.*, World Scientific, Singapore, 1983.
13. Theimer O., Phys. Rev. **D 4**, 1597 (1971); Theimer O. and Peterson P. R., Phys. Rev. **D 10**, 3962 (1974).
14. Boyer T. H., Phys. Rev. **D 27**, 2906 (1983); **29**, 1096 (1984).
15. Payen R., J. Physique **45**, 805 (1984).
16. Einstein A., Phys. Zeitschr. **18**, 121 (1917).
17. Bethe H.A., Phys. Rev. **72**, 339 (1947); reprinted in *Quantum Electrodynamics*, J. Schwinger, editor, Dover, New York, 1958.
18. Cetto A.M. and de la Peña L., Phys. Rev. **A 37**, 1952 (1988).
19. Dalibard J., Dupont-Roc J. and Cohen-Tannoudji C., *J. Physique* **43**, 1617 (1982).
20. Sokolov A.A. and Tumanov V.S., Soviet Physics JETP **3**, 958 (1957); see also Schiller R. and Tesser H., Phys. Rev. **A3**, 2035 (1975).

NELSON'S STOCHASTIC MECHANICS AS THE PROBLEM OF RANDOM FLIGHTS AND ROTATIONS

Piotr Garbaczewski

Institute of Theoretical Physics, University of Wrocław,
50-205 Wrocław, Poland

Abstract

We give a discussion of two important problems arising in the framework of Nelson's stochastic mechanics:
1) The phase-space implementation of Markovian diffusion in R^3
2) The relativistic description of rotational diffusions applicable to spin $\frac{1}{2}$ particles.

1 Stochastic mechanics as the problem of random flights

Although unusual according to current paradigms of quantum theory, the notion of a particle path may be given a well defined meaning in the context of configuration space motions. We can view it as a sample trajectory followed by the mass point undergoing a Markovian diffusion process in R^3 with dynamics constrained by the second Newton law in the (conditional) mean[1,2].

Leaving aside the problem to what extent a particular solution of the respective stochastic differential equation with constraint is capable of simulating a genuine point particle path in R^3, we are left with a persuasive representation of each Schrödinger wave function by a collection of sample propagation scenarios. They constitute a repeatable series of single particle propagation events (flight through) in a non-dissipative random medium (Nelson's background field). Each event is confined to the same duration interval, $[0,T] \ni t$ as example. While trying to understand this classically motivated stochastic model of quantum phenomena on physically deeper grounds, one is tempted[3] to reveal a random phase-space propagation, whose configuration space projection would imply stochastic mechanics. Apparently the physics of phase-space motions relates the non-zero energy-momentum transfer to each random scattering event along a particle path. This feature is irreducibly smoothed out while passing to the configuration space randomness.

The analysis of links between Einstein (configuration space) and Langevin (phase space) description of the Brownian motion was made in the expository Ref. 4: the large friction regime of the Ornstein-Uhlenbeck theory allows for Smoluchowski approximation (e.g. spatial diffusion) in case of a general external force.

On the opposite, attempts to derive stochastic mechanics in the framework of the so called stochastic electrodynamics [5,6] indicate that in the Markovian approximation of process with short correlation times [7] one should disregard friction to have eventually arrived at the Nelson's framework.

Our purpose is to view stochastic mechanics as a specific version of the general problem of random flights [8] and analyze the related Madelung fluid concept from this perspective. In fact this program has been pursued in [2] where forward and backward drifts for a Markovian diffusion in R^3 together with the acceleration formula were established as the conditional mean values evaluated over particles flying about a given point $\vec{x}$ in fixed (infinitesimal) time interval $[t-\Delta t, t+\Delta t]$.

Let us consider a Langevin problem for a mass m particle moving in a conservative field of force, while being subject to perturbations by random forces

$$\frac{d\vec{x}}{dt} = \frac{\vec{p}}{m}, \qquad \frac{d\vec{p}}{dt} = -\nabla V + \vec{\mathcal{F}}(t) \tag{1.1}$$

Here $\vec{\mathcal{F}}(t)$ represents a stochastic force for which we adopt the simplest white noise assumption about its statistical properties, via expectation values:

$$\langle \mathcal{F}_i(t) \rangle = 0 \quad \langle \mathcal{F}_i(t)\, \mathcal{F}_j(t') \rangle = \xi \delta_{ij} \delta(t-t') \tag{1.2}$$

We have totally disregarded any systematic friction effect, which in the standard Brownian motion theory is a must to have fulfilled the fluctuation - dissipation theorem for a system in equilibrium. The white noise amplitude ξ is left unspecified.

Let $\Phi(\vec{x},\vec{p},t)$ be a joint probability density that a particle with velocity $\frac{\vec{p}}{m}$ crosses $\vec{x}$ at time t.

By following a standard procedure [7-10] of expanding $\Phi(\vec{x}+\frac{\vec{p}}{m}\Delta t, \vec{p}+\vec{F}\,\Delta t, t+\Delta t)$ into the Taylor series, taking the expectation values (2), and disregarding higher expansion terms in the $\Delta t \to 0$ limit, we recover the following evolution equation for $\Phi(\vec{x},\vec{p},t)$:

$$\partial_t \Phi + \frac{\vec{p}}{m}\nabla_x \Phi + \vec{F}\nabla_p \Phi = \frac{\xi}{2}\Delta_p \Phi \tag{1.3}$$

with $\vec{F} = -\nabla V$. Given an initial distribution $\Phi_0(\vec{x},\vec{p}) = \Phi(\vec{x},\vec{p},0)$, we have a well defined Cauchy problem in hands.

Eq. (3) differs from the conventional Fokker-Planck equation characteristic for the Brownian motion, by the missing term $(\beta/m^2)\nabla_p(\vec{p}\Phi)$ on the right-hand-side, which would derive from the systematic friction correction $(-\beta/m)\vec{p}$ to (1).

For a random medium in thermodynamic equilibrium, the fluctuation - dissipation relationship $\xi = 2kT\beta$ would then hold true, while there is no reason for it in connection with[2],[3]. Accordingly, (3) should refers to statistical non-equilibrium phenomena, albeit not necessarily far away from the state of equilibrium.

It is worth mentioning that (3) is quite akin to the Boltzmann transport equation in the kinetic theory of gases and liquids [11]. However in virtue of the fundamentally different origin of randomness, Boltzmann two-body collision term in our case is replaced by the white noise diffusion term $\frac{\xi}{2}\Delta_p\Phi$. The Langevin problem (1),(2) by its very nature does not refer to any direct particle - particle interaction between the gas species (particles following sample paths of the process). By introducing the reduced (configuration space) probability distribution [5,12]

$$\rho(\vec{x},t) = \int \Phi(\vec{x},\vec{p},t)d\vec{p} \tag{1.4}$$

we can easily deduce from (3) the conservation laws for local (configuration space conditioned) moments of Φ. For example, these holds:

$$\begin{aligned} &\partial_t \rho + \frac{1}{m}\nabla(<\vec{p}>_x \rho) = 0 \\ &\partial_t(<p_i>_x \rho) + \frac{1}{m}\nabla_j(<p_i p_j>_x \rho) - F_i\rho = 0 \end{aligned} \tag{1.5}$$

where, for any $f = f(\vec{p})$:

$$< f >_x = \frac{1}{\rho(\vec{x},t)} \int f(\vec{p})\Phi(\vec{x},\vec{p},t)d\vec{p} \tag{1.6}$$

In virtue of momentum averaging, equations (5) refer to the phase-space ensemble (the totality of all samples) formed by single mass m particle flights. The local conservation laws (5) describe the mass and momentum balance in the local (spatial) flow representation of the stochastic process (1), (2). The local velocity of the spatial flow:

$$\vec{v}(\vec{x},t) = \frac{1}{m} < \vec{p} >_x = \frac{1}{\rho} \int \frac{\vec{p}}{m} \Phi(\vec{x},\vec{p},t)d\vec{p} \tag{1.7}$$

obeys the continuity equation:

$$\partial_t \rho = -\nabla(\vec{v}\rho) \tag{1.8}$$

By means of the standard decomposition[10,11]:

$$\frac{\vec{p}}{m} = \vec{v}(\vec{x},t) + \vec{\xi}(\vec{x},t)$$
$$\left\langle \vec{\xi}(\vec{x},t) \right\rangle_x = 0 \tag{1.9}$$

we can introduce the local pressure tensor P_{ij}:

$$\frac{1}{m^2} < p_i p_j >_x = v_i v_j + \frac{1}{\rho} P_{ij}$$
$$P_{ij}(\vec{x},t) = \rho < \xi_i \xi_j >_x \tag{1.10}$$

so that the momentum balance equation reads:

$$(\partial_t + \vec{v}\nabla)v_i = \frac{F_i}{m} - \frac{1}{\rho}\nabla_j P_{ij} \tag{1.11}$$

Upon assuming that $\Phi(\vec{x},\vec{p},t)$ has a property:

$$P_{ij} = P\delta_{ij} \quad P = \rho \left\langle \vec{\xi}^2 \right\rangle_x \tag{1.12}$$

we can give the equation (11) the familiar Euler form[10-12] of the transport equation for an ideal compressible fluid (or rather a gas):

$$\partial_t \vec{v} + (\vec{v}\nabla)\vec{v} = \frac{\vec{F}}{m} - \frac{1}{\rho}\nabla P \tag{1.13}$$

Equations (8), (13) are a consequence of (3), hence our gas is a purely mathematical construct: particles and particle paths comprising it have no simultaneous existence. The gas picture is exclusively a result of momentum averaging over alternative (sample flights) propagation scenarios.

By means of (1) we can generate concrete particle (sample) paths originating from a given initial phase-space point. The transport equation (13) does not refer to individual (single) particles but to (infinitesimal) fluid droplets propagating through R^3 . While a particle is acted upon by conservative and random faces, the droplet dynamics is dictated by conservative forces and the force due to stresses in the gas, which is purely statistical by origin.

It is a fundamental property of the gas [14] that all stresses in it are generated exclusively by momentum flows. In particular, the stress at any spatial point is determined by the gas pressure i.e. a force acting on the (infinitesimal) surface element at this point, in the normal direction. Such a force equals the momentum transfer through a surface element per unit of time. For a given (infinitesimal) droplet volume W we have:

$$\int_W \nabla P dv = \int_{\partial W} P d\vec{\sigma} \tag{1.14}$$

which is precisely the time rate of momentum transfer through the boundary ∂W of the volume W.

Let us assume that P is an osmotic [11] pressure so that particles are driven away from the main concentration in accordance with:

$$-\nu\nabla\rho = \vec{u}\rho \tag{1.15}$$

where $\vec{u} = \vec{u}(\vec{x},t)$ is called the osmotic velocity.

Remark 1: The well known Einstein's solution for the osmotic equilibrium of the gas in the gravitational field tells us that the osmotic pressure at the elevation x reads $P(x) = \rho(x)kT$, hence the osmotic force - $\nabla P = -kT\nabla\rho$ if balanced by gravitation: $-mg\rho - kT\nabla\rho = 0$ gives rise to $\rho(x) = \rho(0)exp(-mgx/kT)$. Let us observe that $\nabla P = \rho\nabla w$, where $w = -mgx$ is the osmotic potential, while $V = mgx$ is the gravitational one.

Motivated by this remark, let us assume that $\nabla P/\rho$ in (13) is a gradient field i.e. comes out from a certain potential. Then we have:

$$\nabla \times \frac{\nabla P}{\rho} = 0 \;\rightarrow\; \nabla P \times \nabla\rho = 0 \tag{1.16}$$

which establishes a functional relationship between P and ρ. By denoting $\nabla P = \rho\nabla w$ we realize that ∇w must be proportional to $\nabla\rho$ i.e. $\vec{u}$.

The transport equation (13) can be converted into the fundamental constraint equation (Newton second law in the mean) of Nelson's stochastic mechanics, if a very specific form is chosen for the osmotic potential $w = w(\vec{x},t)$:

$$w = -\frac{m}{2}\vec{u}^2 - \frac{\hbar}{2} div\, \vec{u} = -\frac{\hbar^2}{2m}\frac{\Delta\rho^{1/2}}{\rho^{1/2}} \tag{1.17}$$

which, in virtue of (15) gives rise to:

$$-\nabla w = m(\vec{u}\nabla)\vec{u} + \frac{\hbar}{2}\Delta\vec{u} \tag{1.18}$$

However, the assumption $m\nabla P = \rho\nabla w$ is in general incompatible with (18).

The above rather unsatisfactory situation is a direct consequence of (12), which corresponds to an ideal fluid feature that forces across any surface are normal to that surface. Certainly the (abstract) gas model of ours might not share this property. In fact, the kinetic theory of matter suggests that if we deal with a velocity field showing significant changes on short distances, then a diffusion of faster molecules may impart momentum to slower portions of the fluid, while the admixture (due to diffusion) of slow molecules would slow down the faster portions of the fluid (gas). We assume this property to be generated by osmotic flows and turn back to the general form (10), (11) of the momentum balance. Let us set the following relationship (equation of state) between the pressure tensor P_{ij} and the local particle density ρ:

$$P_{ij} = -\frac{\hbar}{2m}\,\rho\nabla_j\, u_i \tag{1.19}$$

hence $< \xi_i\xi_j >_s = \frac{\hbar}{2m}\,\nabla_j u_i$. Then, automatically (summation convention is implicit):

$$-m\nabla_j P_{ij} = \rho m(\vec{u}\,\nabla)u_i + \frac{\hbar}{2}\,\rho\Delta u_i = -\rho\nabla_i w \tag{1.20}$$

with $w = w(\vec{x}, t)$ defined previously by (17). The momentum transfer formula (14) should now be modified to show the anisotropy of momentum flows through ∂W:

$$\int_W \nabla_j P_{ij} dV = \int_{\partial W} P_{ij} n_j d\sigma \tag{1.21}$$

where $\vec{n}$ is the unit normal vector attached to $d\sigma$. In addition to the gradient flow definition (15), let us demand that the current velocity field $\vec{v}$ is irrotational as well. Then, by (17), (18) the system of equations (8), (11) acquires a canonical (Madelung) form:

$$\begin{aligned} \partial_t\vec{u} &= -\frac{\hbar}{2m}\,\Delta\vec{v} - \nabla(\vec{v}\vec{u}) \\ \partial_t\vec{v} &= \frac{\hbar}{2m}\,\Delta\vec{u} + \frac{1}{2}\nabla\vec{u}^2 - \frac{1}{2}\nabla\vec{v}^2 - \frac{1}{m}\nabla V \end{aligned} \tag{1.22}$$

whose link with Nelson's diffusions in R^3 is well established [1,4,17,18,2]. If supplied with the initial data $\vec{u}_0(\vec{x})$, $\vec{v}_0(\vec{x})$, equations (22) form a well defined Cauchy problem, whose solution is determined (uniquely in the simply connected area) by a solution of the Schrödinger equation. Let us notice, that with velocity fields $\vec{v}, \vec{u}$ in hands, we have completely determined the first and second moments of the joint probability distribution $\Phi(\vec{x}, \vec{p}, t)$ satisfying (3).

The major outcome or our discussion is a purely statistical derivation of the so called [15] quantum potential $-\frac{\hbar^2}{2m}\,\frac{\Delta\rho^{1/2}}{\rho^{1/2}}$, which indeed has nothing to do with forces affecting individual particles in their random (phase space) motion. It arises exclusively due to momentum flows in the statistical (ensemble) description of the gas of sample paths.

The interesting situation is here obtained by comparing the Langevin problem (1) with the Nelson's diffusion implementing (22):

$$\begin{aligned} &d\vec{X}(t) = (\vec{u} + \vec{v})(\vec{X}(t), t)dt + \sqrt{\frac{\hbar}{m}}\,d\vec{W}(t) \\ &\frac{\vec{F}}{m}(\vec{X}(t), t) = \frac{1}{2}(D_+D_- + D_-D_+)\vec{X}(t) = \\ &(\partial_t + \vec{v}\nabla)\vec{v} - \left(\frac{\hbar}{2m}\,\Delta + \vec{u}\nabla\right)\vec{u} = \\ &= \partial_t\vec{v} + \nabla\left(\frac{\vec{v}^2 - \vec{u}^2}{2} - \frac{\hbar}{2m}\,div\,\vec{u}\right) \end{aligned} \tag{1.23}$$

where $\vec{W}(t)$ is a normalized Wiener noise in R^3. Although (23) generates wildly random trajectories in R^3 they cannot be identified with the individual particle paths in the sense of (1) . The drift $\vec{b} = \vec{u} + \vec{v}$ entering (23) indicates the local flow tendency of propagation and comes from configuration space conditioned moments of the phase-space probability distribution $\Phi(\vec{x}, \vec{p}, t)$. Hence is definitely not the individual particle velocity consistent with (1). In virtue of $\vec{F} = -\nabla V$, as a consequence of (23) we get an analogue of the Bernoulli theorem for steady flows (stationary states of the related Schrödinger equation). Indeed $\partial_t\vec{v} = 0$ induces a conserved quantity

$$\frac{1}{2}\left(\vec{v}^2 - \vec{u}^2\right) - \frac{\hbar}{2m}\,div\,\vec{u} + \frac{V}{m} = E \tag{1.24}$$

For $\vec{u}$ and $\vec{v}$ given in the simply connected area where ρ has no zeroes, except for the boundaries where ρ is assumed to vanish, we easily recover [20] Nelson's energy localization condition (compute the mean value of (24) while integrating by parts):

$$\int d^3\vec{x}\;\rho(\vec{x}, t)\left[\frac{m}{2}\left(\vec{v}^2 + \vec{u}^2\right) + V\right] = E \tag{1.25}$$

2 The problem of random rotations in special relativity

We address the major unsolved problem in the framework of Nelson's stochastic mechanics and attempt to provide a description of relativistic spin 1/2 particles in terms of Markovian diffusions on S_3. Random rotations are here labelled by the proper time of a particle in relativistic motion and are continuously distributed

along a space-time trajectory followed by the particle in Minkowski space. The description of spin 1/2 in the framework of Nelson's stochastic mechanics [21,22] involves a harmonic analysis on the group $\mathcal{G}$ of rotations in R^3. For functions on a group the Hilbert space structure is induced by the invariant Haar measure on S_3. The scalar product reads[23,24]

$$(f_1, f_2) = \int dg\, f_1(g)\bar{f}_2(g) = \frac{1}{8\pi^2}\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \int_0^{2\pi} d\psi\, sin\theta\, f_1(g)\bar{f}_2(g) \qquad (2.1)$$
$$S \ni g = (\theta, \phi, \psi)$$

where θ, ϕ, ψ are the Euler angles locally parametrizing the group manifold in a given Cartesian frame of reference. Transformations $U_{g_1} f(g) = f(gg_1)$ are known to form an infinite unitary representation of the group of rotations in the Hilbert space $L^2(S_3)$, whose resolution into irreducible components gives rise to a four-dimensional space $\mathcal{H}_{1/2}$ characterizing spin 1/2. The familiar $SU(2)$ harmonics form an orthonormal basis system in it.

Let $G(t)$ be a random variable taking values in S_3, which undergoes a non-dissipative (rotation) Markovian diffusion[21,22]. The i-th $SU(2)$ harmonic describes the i-th state of stationary diffusion, i=1,2,3,4. For each state we can introduce a vector-valued function of this random variable $\vec{L} = \vec{L}(G(t))$, which following[21,22] may be attributed the role of an angular momentum induced by the rotation $G(t)$ in a given state of rotational diffusion. Let $e_i(g)$ denote the $SU(2)$ harmonic for the spin 1/2 case. Then $|e_i(g)|^2$ stands for the probability distribution of $G(t)$ in the state $e_i(g)$. We can evaluate the expectation values:

$$< \vec{L} >_i = \int \vec{L}(g)|e_i(g)|^2 dg = \epsilon_i \frac{\hbar}{2}\vec{s} \qquad (2.2)$$
$$\epsilon_i = \begin{cases} +1 & \text{i=1,2} \\ -1 & \text{i=3,4} \end{cases}$$

where $\vec{s}$ is the unit (spin polarization) vector in R^3 identifying the direction of the quantization axis in space. In the standard quantum mechanical lore it is the direction, spin projections on which are equal $\pm\hbar/2$.

In the Dankel's paper[21] the case of $\vec{s} = \vec{k}$ (z-direction in the Cartesian frame) was investigated. In virtue of [22] the mere change of Euler parametrization allows to consider quite arbitrary $\vec{s}$, eventually allowing for a smooth time dependence $\vec{s} = \vec{s}(t)$ characteristic for spin precession.

According to Ref. 22 rotational diffusions characterizing the spin 1/2 particle at rest involve the Cartesian system frame whose orientation relative to the laboratory frame is given by Euler angles $(\bar{\theta}, \bar{\phi}, \bar{\psi}) = \bar{g}$. They determine the quantization axis direction

$$R(\bar{g})\vec{k} = \vec{s} = (sin\bar{\psi}sin\bar{\theta}, cos\bar{\psi}sin\bar{\theta}, cos\bar{\theta}) \qquad (2.3)$$

All random fluctuations (e.g. rotations) are intrinsic to the system frame, and described in terms of another set of Euler angles $(\theta, \phi, \psi) = g$ referring to intrinsic rotation axes $\vec{e}_\theta$, $\vec{e}_\phi$, $\vec{e}_\psi$ in the system frame.

The discussion [21,22] of random rotations implementing spin 1/2 is purely non-relativistic and effectively confined to the system rest frame (inhomogeneous magnetic fields alter this picture[22]).

We denote by K' the inertial rest frame, in which the spin 1/2 rotational diffusion takes place. We admit furthermore that K' moves uniformly with the velocity $\vec{v}$, $|\vec{v}| < c$ relative to another inertial frame, and address the following relativistic problem:

> How does the K observer perceive the Markovian diffusion taking place in K'? Is it a stochastic diffusion process again?

Nonrelativistic intuitions of Ref. 22 suggest that we should first establish transformation properties of the polarization axes when passing from one inertial frame to another. The issue has been solved in the context of the Bargmann-Michel-Telegdi equation[25,26]. Specification of the rest frame polarization is known to determine the components of the polarization four-vector in any inertial frame. Indeed, if $\vec{s}$ is the rest frame polarization then its K frame image (via the Lorentz transformation Λ^{-1} taking $(c,0,0,0)$ into $(\gamma c, \gamma\vec{v})$ with $\gamma = (1-\beta^2)^{-1/2}$, $\vec{\beta} = \frac{\vec{v}}{c}$) is given by [26]

$$(0, \vec{s}) \rightarrow (S^0, \vec{S})$$
$$S^0 = \vec{\beta}\vec{S} = \gamma\vec{\beta}\vec{s} \qquad (2.4)$$
$$\vec{S} = \vec{s} + \frac{\gamma^2}{1+\gamma}(\vec{\beta}\vec{s})\vec{\beta}$$

while in reverse we have $\vec{s} = \vec{S} - \frac{\gamma}{\gamma+1}(\vec{\beta}\vec{S})\vec{\beta}$.

The normalized vector $\vec{s}$, $|\vec{s}| = 1$ has the unnormalized spatial image in K, since $(S^0)^2 - \vec{S}^2 = -1$. Nevertheless $\vec{S}$ properly identifies the polarization (quantization axis) direction as seen by another (e.g.K) inertial observer.

With a fixed Λ^{-1} in hands we have given a stationary group of the four-velocity vector $(\gamma c, \gamma\vec{v})$ as a subgroup of Lorentz transformations, which leave this vector in place (i.e. do not take this vector away from a given inertial frame). In K' this stability group coincides with the group of rotations while its isomorphic image in K is given[24] by

$$\mathcal{G} = \mathcal{G}(\vec{v}) = \Lambda^{-1}\mathcal{G}'\Lambda \qquad (2.5)$$

Unfortunately the action of $\mathcal{G}$ in K is a non-trivial transformation of four-vectors which modifies the length of their spatial component (unlike $\mathcal{G}'$ in K'). Let us consider the spatial part of the transformation (2.4) as a mapping in R^3. Vectors $\vec{S}$,$\vec{s}$, $\vec{v}$ are coplanar in R^3 and the map is realized by the alteration of the longitudinal

(along $\vec{v}$) component of $\vec{s}$ merely

$$\vec{s} \to \vec{S} \Rightarrow \quad \vec{s}_{\|} = \vec{\beta}(\vec{\beta}\cdot\vec{s})/\beta^2$$
$$\vec{S}_{\|} = \vec{s}_{\|} + \vec{\beta}\,(\vec{\beta}\cdot\vec{s})\gamma^2/(\gamma+1) \tag{2.6}$$

We are interested in the relative orientation of polarization axes along $\vec{s}$ and $\vec{S}$ respectively. This purely rotational output of the Lorentz transformation Λ^{-1} becomes isolated through:

$$cos\,\theta = \vec{s}\cdot\hat{S}, \quad \hat{S} = \vec{s}/|\vec{s}|, \quad \vec{s}\vec{S} = 1 + \tfrac{\gamma^2}{1+\gamma}\left(\vec{\beta}\vec{s}\right)^2$$
$$|\vec{S}| = \left\{1 + \left(\vec{\beta}\vec{s}\right)^2 \tfrac{\gamma^2}{1+\gamma}\left[\beta^2\tfrac{\gamma^2}{1+\gamma} + 2\right]\right\}^{1/2} \tag{2.7}$$

and θ is uniquely defined given $\vec{s}$ and $\vec{v}$.

Once in R^3, passing from $\vec{s}$ to $\vec{S}$ amounts to the rotation by an angle θ about the axis $\vec{s}\times\vec{\beta}$ (same as about $\vec{S}\times\vec{\beta}$)

$$g_\theta : \vec{s} \to \vec{S} = g_\theta \vec{s} \tag{2.8}$$

From now on directions $\vec{s}$ and $\vec{S}$ will be the fixed z-axis directions in the system frames located in K' and K respectively.

Given a unit vector $\vec{n}$, initially along the z-axis of the system frame in K'. Let us execute a rotation $\vec{n} \to g\vec{n}$. In view of (2.4) we have here

$$g\vec{n} \to \vec{N}_g = g\left[\vec{n} + \frac{\gamma^2}{1+\gamma}\left(\vec{\beta}'\vec{n}\right)\vec{\beta}'\right] = g\,\vec{N}_{\theta'} \tag{2.9}$$

where $\vec{\beta}' = g^{-1}\vec{\beta}$ induces a rotation angle v' about the axis $\vec{n}\times\vec{\beta}'$. Accordingly a unit vector $\hat{N}_g$ is recovered

$$\hat{N}_g = g\hat{N}_{\theta'} = gg_{\theta'}\vec{n} = gg_{\theta'}g_\theta^{-1}\hat{N} \tag{2.10}$$

where $\hat{N}$ was initially along the z-axis of the system frame located in K (i.e. parallel to $\vec{S}$). If we consider $\hat{N}_{g_1}$ and $\hat{N}_{g_2}$ then

$$\hat{N}_{g_1} = g_1 g_{1\theta'} g_\theta^{-1}\hat{N} = (g_1 g_{1\theta'})(g_2 g_{2\theta'})^{-1}\hat{N}_{g_2} \tag{2.11}$$

In the above g_θ was introduced as the rotation by θ about the spatial axis. Each element of the rotation group can be represented that way. Let g_i refers to the rotation axis $\hat{e}_i$ and angle α_i. By passing to $\vec{\mu}_i = \hat{e}_i\, tan\alpha_i$ we arrive at a particularly convenient representation [27] of spatial rotations by 3×3 matrices $R(\vec{\mu})$

$$R(\vec{\mu})_{ij} = [1/(1+\mu^2)]\left[\left(1-\mu^2\right)\delta_{ij} + 2\mu_i\mu_j - 2\epsilon_{ijk}\mu_k\right] \tag{2.12}$$

with the composition rule

$$R(\vec{\mu}')R(\vec{\mu}) = R(\vec{\mu}'')$$
$$\vec{\mu}'' = (\vec{\mu}' + \vec{\mu} + \vec{\mu}'\times\vec{\mu})/(1 - \vec{\mu}'\vec{\mu}) \tag{2.13}$$

allowing to attribute to each rotation g in K' a respective spatial rotation in K

$$g\vec{s} \leftrightarrow \hat{S}_g = (gg_{\theta'})g_\theta^{-1}\hat{S} = \tilde{g}g_\theta^{-1}\hat{S} \tag{2.14}$$

Let us notice that together with $\vec{S}$ we have automatically defined an orthogonal reference triad in K. The right screw convention for the vector product allows us to introduce the x and y axes analogues by $\vec{S}\times\vec{v}$ and $(\vec{S}\times\vec{v})\times\vec{S}$ respectively.

All rotations can be parametrized by Euler angles introduced in this frame. Respective parametrizations are not the same for $\{\tilde{g}\,g_\theta^{-1}\}$ and $\{\tilde{g}\}$. However, the very concept of the invariant integration on S_3 implies that a given g_0 displacement on a group. $g \to gg_0$ does not affect the integration formulas, hence the respective Euler parametrization. We have [24] $dg = d(gg_0)$ and $\int f(g)dg = \int f(gg_0)dg$. Consequently the effective image of the group of rotations in K' under the Lorentz transformation Λ^{-1} is

$$\mathcal{G}' \ni g \longrightarrow \tilde{\mathcal{G}} \ni \tilde{g} = g\cdot g_{\theta'} \tag{2.15}$$

which well fits with (2.11):

$$g_1\vec{n} = g_1 g_2^{-1}(g_2\vec{n}) \longrightarrow \hat{N}_{g_1} = \tilde{g}_1\tilde{g}_2^{-1}\hat{N}_{g_2} \tag{2.16}$$

If now g's represent random rotations about the $\vec{s}$ polarized frame in K', then $\tilde{g}$'s are their images as random rotations about the $\vec{S}$ polarized frame in K.

Irrespective of whether we refer to K' or K, the previous discussion and arguments of Ref. 23, Sect. 4 tell us that once we have fix the polarization axis direction (i.e. the z-axis of the orthonormal triad) then the induced Euler angle parametrization of S_3 allows us to determine the spin 1/2 $SU(2)$ harmonics as eigenfunctions of the Laplace-Beltrami operator on S_3

$$\begin{aligned}
&\vec{s} \longrightarrow g = (\theta,\phi,\psi) \quad \longrightarrow \quad \Delta_g f'(g) = (3/4)f'(g)\\
&\vec{S} \longrightarrow \tilde{g} = (\tilde{\theta},\tilde{\phi},\tilde{\psi}) \quad \longrightarrow \quad \Delta_{\tilde{g}} f(\tilde{g}) = (3/4)f(\tilde{g})\\
&\Delta_{\vec{g}} = \Delta_{(g\to\tilde{g})}\\
&\Delta_g = \frac{\partial^2}{\partial\theta^2} + cos\theta\,\frac{\partial}{\partial\theta} + \frac{1}{sin^2\theta}\left(\frac{\partial^2}{\partial\psi^2} + \frac{\partial^2}{\partial\phi^2}\right) - 2\,\frac{cot\theta}{sin\theta}\,\frac{\partial^2}{\partial\phi\partial\psi}
\end{aligned} \tag{2.17}$$

so that solutions acquire the same functional form, albeit with respect to entirely different parametrizations. Thereby showing up different polarization directions in R^3.

The change of the local S_3 parametrization from (θ, ϕ, ψ) to $(\tilde{\theta}, \tilde{\phi}, \tilde{\psi})$ implies a replacement of the angular momentum vector $\vec{L}(g)$ by $\tilde{\vec{L}}(\tilde{g})$

$$\vec{L}(g) = \alpha(g)\vec{e}_\phi + \beta(g)\vec{e}_\psi + \gamma(g)\vec{e}_\theta$$
$$\tilde{\vec{L}}(\tilde{g}) = \alpha(\tilde{g})\vec{e}_{\tilde{\phi}} + \beta(\tilde{g})\vec{e}_{\tilde{\psi}} + \gamma(\tilde{g})\vec{e}_{\tilde{\theta}} \tag{2.18}$$

where vectors $\vec{e}$ indicate directions about which rotations by respective angles are executed. Since $\vec{e}$'s are defined in the system frame (with z-axis given either by $\vec{s}$ or $\vec{S}$) the transformation from $\vec{e}_\theta$, $\vec{e}_\phi$, $\vec{e}_\psi$, to $\vec{e}_{\tilde{\theta}}$, $\vec{e}_{\tilde{\phi}}$, $\vec{e}_{\tilde{\psi}}$ is effected by the previously considered spatial rotation g_θ taking $\vec{s}$ into $\vec{S}$. Accordingly

$$\tilde{\vec{L}}(\tilde{g}) = (g_\theta \vec{L})(g \to \tilde{g}) \tag{2.19}$$

i.e. the change of arguments is accompanied by the overall rotation of $\vec{L}$. We have then

$$< \tilde{\vec{L}} >_i \; = \int \tilde{\vec{L}}(\tilde{g})|e_i(\tilde{g})|^2 d\tilde{g} = \epsilon_i \tfrac{\hbar}{2} \hat{S} = \epsilon_i \tfrac{\hbar}{2}\, g_\theta \vec{s} =$$
$$= g_\theta \int \vec{L}(\tilde{g})|e_i(\tilde{g})|^2 d\tilde{g} = g_\theta \int \vec{L}(g)|e_i(g)|^2 dg \tag{2.20}$$

Consequently four stationary states of rotational diffusion $\vec{e}_i(g)$, $g = (\theta, \phi, \psi)$ in K' can be mapped into four states of rotational diffusion again while their polarization $\vec{s}$ is taken over to $\vec{S}$.

This map we shall study in more detail in connection with solutions of the Dirac equation.

3 Rotational diffusions as seen by a relativistic observer: case of uniform motion

The description of a stochastic process is usually confined to a fixed time interval, which eventually might be extended to an arbitrary size. Let us choose $[0, T'] \ni t'$. A random variable $G(t')$ is represented by a rotational event g taking place at time t' while $\vec{x}'$ is a location of the rotating triad origin. Hence we deal with g at the space-time point $(ct', \vec{x}')$. By virtue of our previous considerations $G(t')$ induces a random variable $\tilde{G}(t')$ in K which refers to a rotation $\tilde{g}$ taking place at the space-time point $(ct, \vec{x})$ in K

$$\vec{x} = \vec{x}' + \vec{v}\left[(\gamma - 1)(\vec{v}\vec{x})/v^2 + \gamma t'\right]$$
$$t = \gamma\left(t' + \frac{\vec{v}\vec{x}'}{v^2}\right) \tag{3.1}$$

Since $\vec{x}'$ is fixed and the time label t' only is allowed to vary, we can rewrite (2.1) as $\vec{x} = \vec{x}'_0 + \gamma\vec{v}t'$, $t = \gamma(t'_0 + t')$. It tells us that the process is perceived in K as

taking place in the time interval $[\gamma t'_0, \gamma(t'_0 + T')]$ while the rotating triad origin is propagated uniformly with the velocity $\vec{v}$ from the spatial location $\vec{x} = \vec{x}'_0$ to $\vec{x}_T = \vec{x}'_0 + \gamma\vec{v}'T$. Here $\vec{x}'$ is associated with the origin of the rotating frame hence it is a valid assumption to consider $\vec{x}' = 0$ only. It yields

$$\vec{x} = \gamma\vec{v}t', \quad t = \gamma t' \quad t \in [0, \gamma T'] \tag{3.2}$$

If $\vec{v}$ is the speed of the particle relative to the reference frame in which the time interval is measure, the notion of proper time comes through $\Delta\tau = \Delta t/\gamma$.

Apparently it give rise to a proper time labeling of random variables

$$K' \to K \Rightarrow$$
$$G(t') \to \qquad \tilde{G}(t') = \tilde{G}(t/\gamma) = \tilde{G}(\tau)$$

under the assumption (3.2).

Remark 3: The notion of randomness automatically induces the notion of sampling: a repeatable processing confined to a fixed time interval. In particular the notion of sample paths[28] of a given stochastic process is of profound importance. In K' it amounts to representing the random propagation on S_3 by a collection of random trajectories: they are different realizations (samples) of a given random motion scenario in the time interval $[0, T']$ executed by the random variable $G(t')$.

Given a small surface on S_3 with the area Δg centered about the point (rotation) g. Let $e_i(g)$ be the i-th state of rotational diffusion. Then $|e_i(g)|^2\Delta g$ represents a probability with which rotations close to g are met along sample paths in the infinite sampling limit: then the frequency of an event approaches a probability of its occurrence.

The same process, but seen from another inertial frame K induces sample paths as rotational events which are continuously distributed along the relativistic path. The K' eigenvalue problem (2.17) refers to stationary solutions of the Schrödinger equation on S_3

$$i\hbar\partial_{t'} f'(g,t) = \left(-\hbar^2/2I\right)\Delta_g f'(g,t')$$
$$f'(g,t') = f'(g)exp(-mc^2t'\hbar) \tag{3.4}$$

where we set $I = 3\hbar^2/8mc^2$ to deal with spin 1/2. The respective eigenvalue problem in K is given by the Schrödinger equation on S_3 but with a proper time τ instead of t'

$$i\hbar\partial_\tau f(\tilde{g},\tau) = \left(-\hbar^2/2I\right)\Delta_{\tilde{g}} f(\tilde{g},\tau)$$
$$f(\tilde{g},\tau) = f(\tilde{g})exp(-mc^2\tau/\hbar) \tag{3.5}$$

Then (3.4) is related to the stochastic process with a random variable $G(t')$ while (3.5) induces $\tilde{G}(\tau)$.

On the other hand the manifestly covariant form $p_\mu x^\mu$ of mc^2t' emerges by setting $\vec{p} = \gamma m\vec{v}$, $p^0 = \gamma mc = E/c$ i.e. $E = (c^2\vec{p}^2 + m^2c^4)^{1/2}$, $\vec{v} = c\vec{p}/E$, $\vec{\beta} = \vec{p}/E$ which tells us that

$$mc^2t' = -\vec{p}\vec{x} + p^0ct = p_\mu x^\mu \tag{3.6}$$

with $\vec{x} = \vec{x}(\vec{x}', t')$, $t = t(\vec{x}', t')$ given by (2.1). It means however that $f'(g, t')$ should take in K the form

$$f(\tilde{g}, t) = f(\tilde{g})exp[-i(p^0ct - \vec{p}\vec{x})/\hbar] \tag{3.7}$$

which is characteristic for plane wave solutions of the Dirac equation. Except for the explicit $\tilde{g}$ dependence of the coefficient $f(\tilde{g})$.

At this point it is quite instructive to invoke an exhaustive discussion of Ref. 24 on the determination of the rest mass and spin of the particle in the context of relativistic invariant wave equations.

Usually one deals with arbitrary plane wave solutions and attempts to extract their rest frame properties. We have proceeded in reverse, while having a detailed rest frame picture (of random phenomena) in hands. Let us view (3.7) as an arbitrary plane wave i.e. allow t and $\vec{x}$ to take any value. We can always pass[5] to the rest frame of the wave and recover a corresponding stationary plane wave, which is (3.5) in our case. Indeed, the Lorentz transformation $\Lambda : (\gamma c, \gamma\vec{v}) \longrightarrow (c, 0, 0, 0)$ implies

$$\begin{aligned} &f(\tilde{g})exp(-ip_\mu x^\mu/\hbar) \to f'(g)exp(-imc^2t'/\hbar) \\ &mc^2t' = p'_0x'_0 \end{aligned} \tag{3.8}$$

for all $\vec{x}$ and t. Although the spatial image $\vec{x}'$ of $\vec{x}$ under Λ is not manifestly present in (3.8), it is implicitly there since in our framework spatial rotations g take place as rotations about this point.

The transformation Λ affects only the longitudinal component $\vec{x}'_{||}$ of $\vec{x}'$

$$\begin{aligned} &x_{||} = \gamma(x'_{||} + \vec{v}t') \\ &t = \gamma(t' + x'_{||}/v), \quad x_{||} = \vec{v}\vec{x}/|\vec{v}| \end{aligned} \tag{3.9}$$

while on the other hand $p_\mu x^\mu = p_0x_0 - |\vec{p}|x_{||}$ hence at each given time instant x_0/c the plane wave effectively describe a transversal plane (wave front) labeled by $x_{||}$. The formula (3.9) maps the $x'_{||}$ plane in K' into the $x_{||}$ plane in K. The space time location of the plane in K is uniquely defined by a corresponding time instant t' of the rest frame evolution. By (2.9) the one parameter family of wave fronts at rest (in K') is perceived in K as the one parameter family of traveling surface (planes): at $t' = 0$ we have $x^0_{||} = \gamma x'_{||}$ and $t_0 = \gamma x'_{||}/v$ to be compared with (3.9).

Given a plane $x^0_{||}$ at t_0. Each point of this transversal plane is uniquely mapped into the respective point of the image plane $x^T_{||}$ at time T by following the uniform motion path: paths do not intersect and the moving surface traces their flow in configuration space. In the framework of rotational diffusions, the plane wave thus accounts for all alternative (purely classical) motion scenarios to be followed by the origin of the rotating frame in the sampling series. Consequently they are much akin to the Hamilton waves of classical mechanics [29] and not at all to what we usually call travelling waves (Ref. 30 addresses the issue in more detail).

4 Dirac equation in stochastic mechanics: revival of some old ideas

The stochastic implementation [21,22,31,32] of the quantum spin 1/2 system involves four distinct states of rotational diffusion, which reflect[23] the existence of left and right representations of the $SU(2)$ on $\mathcal{H}_{1/2}$. While in K' we have

$$\begin{aligned} &(8\pi^2)^{1/2}e_1(g) = i\,cos\tfrac{\theta}{2}\,exp\tfrac{i}{2}(\psi+\phi) \sim d^{1/2}_{1/2\,1/2}(g), && <\vec{L}>_1 = \tfrac{\hbar}{2}\vec{k} \\ &(8\pi^2)^{1/2}e_2(g) = i\,sin\tfrac{\theta}{2}\,exp\tfrac{i}{2}(\phi-\psi) \sim d^{1/2}_{1/2-1/2}(g), && <\vec{L}>_2 = \tfrac{\hbar}{2}\vec{k} \\ &(8\pi^2)^{1/2}e_3(g) = isin\tfrac{\theta}{2}\,exp\tfrac{i}{2}(-\phi+\psi) \sim d^{1/2}_{-1/2\,1/2}(g), && <\vec{L}>_3 = -\tfrac{\hbar}{2}\vec{k} \\ &(8\pi^2)^{1/2}e_3(g) = i\,cos\tfrac{\theta}{2}\,exp\tfrac{i}{2}(-\phi-\psi) \sim d^{1/2}_{-1/2-1/2}(g), && <\vec{L}>_4 = -\tfrac{\hbar}{2}\vec{k} \end{aligned} \tag{4.1}$$

where d^s_{mn} is the standard notation for $SU(2)$ harmonics. The respective stochastic processes are determined by computing the angular velocity $\vec{\omega}(g)$ induced by the rotation g, and $\vec{\omega}(g)$ is a sum of the current $\vec{\omega}_v$ and osmotic $\vec{\omega}_u$ contributions behaving differently under time reversal. Namely $t' \to -t'$ implies $\vec{\omega}_v \to -\vec{\omega}_v$ while $\vec{\omega}_u \to \vec{\omega}_u$.

As a consequence (compare e.g. Sect. 4 of Ref. 3) we arrive at

$$\begin{aligned} &\vec{\omega}^1_v \to \vec{\omega}^4_v \quad \vec{\omega}^2_v \to \vec{\omega}^3_v \\ &\vec{\omega}^1_u \to \vec{\omega}^4_u \quad \vec{\omega}^2_u \to \vec{\omega}^3_u \end{aligned} \tag{4.2}$$

which amounts to the map

$$\begin{aligned} &e_1(g) \to e_4(g), \qquad e_2(g) \to e_3(g) \\ &f'(g, t') \to f'(g, -t') = \; e(g)exp(mc^2t'/\hbar) \end{aligned} \tag{4.3}$$

Remark 4: Four states of rotational diffusion (4.1) were introduced in connection with the forward propagation. Apparently the discussion[14] of how to describe effects of time reversal in stochastic mechanics as a forward propagation again, may be here adopted. Usually the reverse process is viewed as the random propagation in the backward direction, which allows to reproduce past (statistical) data of the process given the present. Hence as a mathematical artifice merely. It appears that in the case of spin 1/2 diffusions it is no longer the case. Arguments of Ref. 2 Sect. 1 tell us that for Markovian diffusions we can define a forward process which is an exact time reversal of another forward process, and diffusions underlying (4.1) provide us with explicit examples.

Let us recall that[21,22,33] the $SU(2)$ labeling of eigenfunctions (4.1) is provided by the eigenvalues of operators M_3, N_3 where $\vec{M}$ is the generator of left rotations while $\vec{N}$ is the (abnormal) generator of right rotations. We have [34] $\vec{M}^2 = \vec{N}^2 = -\hbar^2\Delta_g$ on the S_3 manifold, and $M_3 = -i\hbar\partial/\partial\phi$, $N_3 = -i\hbar\partial/\partial\psi$. Eigenvalues of M_3 correspond to expectation values $< \vec{L} >_i$ of the angular momentum (spin) arising due to the rotational diffusion.

The ordering (e_1, e_3, e_2, e_4) of the basis system refers to $(+,-,+,-)$ sequence of the M_3 eigenvalues while to $(+,+,-,-)$ for N_3. Analogously (e_2, e_4, e_1, e_3) refers to $(+,-,+,-)$ for M_3 and $(-,-,+,+)$ for N_3. In view of this, formulas (3.8), (4.3) give rise to two distinct evolution equations in K' which encompass the time reversal in a manifest way. Namely $e_j(g)exp(-imc^2t'/\hbar)$, $j = 1,3$ and $e_k(g)exp(imc^2t'/\hbar)$ $k = 2,4$ form a set of independent solutions for

$$i\hbar\partial_{t'}f'(g,t') = (2/\hbar)mc^2N_3'f'(g,t') \tag{4.4}$$

while $e_k(g)exp(-imc^2t'/\hbar)$, $k = 2,4$ and $e_j(g)exp(imc^2t'/\hbar)$ $j = 1,3$ for

$$i\hbar\partial_{t'}f'(g,t') = -(2/\hbar)mc^2N_3'f'(g,t') \tag{4.5}$$

The "positive energy" solutions of (4.4), (4.5) constitute the orthonormal set in $\mathcal{H}_{1/2}$. The prime refers to the rest frame Euler parametrization.

Remark 5: The above observation if combined with the previous Remark 2 lends weight to Barut's conjecture[35] that perhaps there is no real need to invoke the hole theory or the notion of backward propagation in time to describe antiparticles.

Let us address the question of how the rest frame evolution (respectively the eigenvalue problem for N_3) equations (4.4), (4.5) are seen in another Lorentz frame.

In accordance with the standard rules of the game[35] the Lorentz transformation $\Lambda : K \to K'$ should imply a non-unitary map in the function space replacing the K frame data by the K' ones

$$f'(x',g) = (T_\Lambda f)(x,\tilde{g}) \tag{4.6}$$

We shall investigate the outcome of (4.6) in the Hilbert space spanned by $e_i(\tilde{g})$ with the $L^2(S_3)$ scalar product $\int f_1(\tilde{g})f_2(\tilde{g})d(\tilde{g}) = (f_1, f_2)$ valid in K. Let us consider a transformation

$$\Lambda : K \to K' \Rightarrow f(\tilde{g}) = (T_\Lambda f)(\tilde{g}) \tag{4.7}$$

where $T_\Lambda = T_\Lambda(\vec{M}, \vec{N})$ is given by (this formula was first introduced to represent Lorentz transformations in Ref. 33)

$$T_\Lambda = cosh\,\frac{\eta}{2}\left[c + \frac{1}{E+mc^2}\left(\frac{2}{\hbar}\right)^2 N_1(\vec{p}\vec{M})\right] \tag{4.8}$$

with

$$cos\,\frac{\eta}{2} = \left(\frac{E+mc^2}{2mc^2}\right)^{1/2}, \quad p^0 = E/c = \left(\vec{p}^2 + m^2c^2\right)^{1/2} \tag{4.9}$$

In the above $\vec{N}$ and $\vec{M}$ are the previously defined differential operators on S_3 whose explicit form displays the local $(\tilde{\theta}, \tilde{\phi}, \tilde{\psi})$ parametrization.

By exploiting the formulas[24] valid for an irreducible representation of the group of rotations in $L^2(S_3)$

$$\begin{aligned}(M_1 \pm iM_2)\,d^s_{mn} &= \hbar\,[^3(s \mp m)(s \pm m + 1)]^{1/2}\,d^s_{m\pm1\,n} \\ (N_1 \mp iN_2)\,d^s_{mn} &= \hbar\,[(s \mp m)(s \pm m + 1)]^{1/2}\,d^s_{mn\pm1}\end{aligned} \tag{4.10}$$

and specializing them to spin 1/2 we obtain

$$\begin{aligned}N_1 d^{1/2}_{1/2n} &= (\hbar/2)d^{1/2}_{-1/2\,n}, \quad M_2 d^{1/2}_{1/2n} = -i(\hbar/2)d^{1/2}_{-1/2n} \\ N_1 d^{1/2}_{m1/2} &= (\hbar/2)d^{1/2}_{m-1/2}, \quad N_2 d^{1/2}_{m1/2} = -i(\hbar/2)d^{1/2}_{m-1/2}\end{aligned} \tag{4.11}$$

It entails an immediate evaluation of the action of T_Λ on any of e_i's. Let us introduce the notation

$$(e_1, e_3, e_2, e_4) = (\phi_1, \phi_2, \phi_3, \phi_4) \tag{4.12}$$

Then we arrive at

$$(T_\Lambda\phi_i)(\tilde{g}) = \sum_{k=1}^{4}(T_\Lambda)_{ik}\phi_k = \sum_{k=1}^{4} S^T_{ik}\phi_k \tag{4.13}$$

where one recognizes S^T to be a transposed bispinor transformation matrix (see (3.7) in Ref. 35))

$$\begin{aligned}S &= exp\left(-\frac{\eta}{2}\frac{\vec{\alpha}\vec{v}}{|\vec{v}|}\right) \qquad \gamma^i = \gamma^0\alpha_i \\ \psi'(x') &= \psi'(\Lambda x) = S(\Lambda)\psi(x) \quad = S(\Lambda)\psi(\Lambda^{-1}x')\end{aligned} \tag{4.14}$$

Accordingly the 4×4 matrix S comes out by evaluating matrix elements of the operator T_Λ in the $\{\phi_i(\tilde{g})\}$ rotational basis.

Let $w^r = w^r(\vec{p})$ be the r-th column of the matrix S. We can rewrite (4.13) as follows

$$(T_\Lambda\phi_r)(\tilde{g}) = \sum_{k=1}^{4} w^r_k(\vec{p})\phi_k(\tilde{g}) = \phi'_r(\vec{p},\tilde{g}) \tag{4.15}$$

The $L^2(S_3)$ orthonormality relations imply here (we use the bispinor normalization identity in the second step)

$$(\phi'_r(\epsilon_r\vec{p}), \phi'_{r'}(\epsilon_{r'}\vec{p})) \quad = w^{r'*}(\epsilon_{r'}\vec{p})w^r(\epsilon_r\vec{p}) = \frac{E}{mc^2}\,\delta_{rr'} \tag{4.16}$$

$$\epsilon_r = +1, \quad r = 1,2 \quad \epsilon_r = -1, \quad r = 3,4$$

and allow us to introduce a new orthonormal basis system in $\mathcal{H}_{1/2}$ encompassing effects of the Lorentz transformation $K \to K'$

$$e_1'(\tilde{g}) = \phi_1'(\vec{p},\tilde{g})\left(\tfrac{mc^2}{E}\right)^{1/2}, \quad e_3'(\tilde{g}) = \phi_2'(\vec{p},\tilde{g})\left(\tfrac{mc^2}{E}\right)^{1/2}$$
$$e_2'(\tilde{g}) = \phi_3'(-\vec{p},\tilde{g})\left(\tfrac{mc^2}{E}\right)^{1/2}, \quad e_4'(\tilde{g}) = \phi_4'(-\vec{p},\tilde{g})\left(\tfrac{mc^2}{E}\right)^{1/2} \tag{4.17}$$

where $(e_i', e_j') = \int d\tilde{g} e_i'(\tilde{g}) e_j'(\tilde{g}) = \delta_{ij}$ is a positive definite sesquilinear form.

More explicitly

$$e_1'(\tilde{g}) = \left(\frac{E+mc^2}{2E}\right)^{1/2}\left\{e_1(\tilde{g}) + \frac{c}{E+mc^2}[p_z e_2(\tilde{g}) + p_+ e_4(\tilde{g})]\right\}$$

$$e_3'(\tilde{g}) = \left(\frac{E+mc^2}{2E}\right)^{1/2}\left\{e_3(\tilde{g}) + \frac{c}{E+mc^2}[p_- e_2(\tilde{g}) - p_z e_4(\tilde{g})]\right\}$$

$$e_2'(\tilde{g}) = \left(\frac{E+mc^2}{2E}\right)^{1/2}\left\{e_2(\tilde{g}) + \frac{c}{E+mc^2}[-p_z e_1(\tilde{g}) - p_+ e_3(\tilde{g})]\right\} \tag{4.18}$$

$$e_4'(\tilde{g}) = \left(\frac{E+mc^2}{2E}\right)^{1/2}\left\{e_4(\tilde{g}) + \frac{c}{E+mc^2}[-p_- e_1(\tilde{g}) + p_z e_3(\tilde{g})]\right\}$$

$\vec{p} = (p_x, p_y, p_z)$

where $p_\pm = p_x \pm i p_y$ and $e_i(\tilde{g})$ $i = 1,2,3,4$ have the form (3.1) except for the replacement of θ, ϕ, ψ by $\tilde{\theta}, \tilde{\phi}, \tilde{\psi}$.

It was for the first time demonstrated in Ref. 33 that functions

$$f_j(\tilde{g}, x) = (T_\Lambda \phi_j) exp(-\epsilon_j i p_\mu x^\mu/\hbar) \tag{4.19}$$

solve the evolution equation

$$i\hbar\partial_t f_j(\tilde{g}, x) = \left\{\frac{2}{\hbar}mc^2 N_3 + \frac{4c}{\hbar^2}N_1[\vec{M}(-i\hbar\nabla]\right\} f_j(\tilde{g}, x) \tag{4.20}$$

Its matrix form in the $\{\phi_j(\tilde{g})\}$ basis is the familiar Dirac equation

$$[mc^2\beta + c\vec{\alpha}(-i\hbar\nabla)]\psi = i\hbar\partial_t\psi \tag{4.21}$$

The image of (4.5) under Λ is obtained through replacing m by $-m$ in the above.

The equation (4.20) is known to be Lorentz invariant, what then about (4.21). By setting

$$L^0 = I \quad L^i = -N_1 M_i \quad i = 1,2,3 \tag{4.22}$$

(4.20) can be cast in the manifestly covariant form

$$\frac{2}{\hbar}mc^2 N_3 f_j = i\hbar L^\mu \partial_\mu f_j \tag{4.23}$$

The standard (35) Lorentz covariance arguments demand (4.6) to be a map of the K frame data into the K' frame ones. Accordingly

$$N_3 = T_\Lambda N_3' T_\Lambda^{-1} \tag{4.24}$$

reflects merely the change of Euler parameters from $\tilde{\theta}, \tilde{\phi}, \tilde{\psi}$ to θ, ϕ, ψ as a result of the Lorentz transformation, while there holds

$$T_\Lambda L^\mu T_\Lambda^{-1} a^\nu{}_\mu \partial_\nu' f_i'(g, x') = L'^\nu \partial_\nu' f_j'(g, x') \tag{4.25}$$
$$a^\nu{}_\mu = \partial x'^\nu / \partial x^\mu$$

The K' frame version of (4.20)

$$\frac{2}{\hbar}mc^2 N_3' f_j'(g, x') = i\hbar L'^\nu \partial_\nu' f_j'(g, x') \tag{4.26}$$

reduces to (4.4) in case of plane wave solutions.

As a consequence of (4.6) we realize that stationary plane wave solutions $\phi_j(g) exp(-mc^2 t'/\hbar)$ of (4.4) are represented in terms of the K frame data by solutions (4.17) of the evolution equation (4.20). A serious problem comes here from the covariant normalization statement [35]

$$\bar{w}^r(\vec{p}) w^{r'}(\vec{p}) = \delta_{rr'}\epsilon_r \tag{4.27}$$

Before, the plane waves were found to refer to four distinct stochastic rotational processes in K'. Because of the improper (negative) normalization the $r = 3$ and 4 images of random motions in K' do not admit any reasonable probabilistic decoding in K, hence cannot be perceived as stochastic processes in K. It is the normalization identity (4.15) which allows to introduce an orthonormal basis system (4.16) with prospects for a correct probabilistic content (due to a positive normalization). Apparently (4.16) arises only if we consider a complete set of "positive energy" solutions of of both (4.4), (4.5). Both these evolution equations are indispensable for a covariant transformation of the orthonormal basis given in K' into an orthonormal basis in K. In fact formulas (4.16) identify these functions in $\mathcal{H}_{1/2}$ which provide us with the K frame image (via Lorentz transformation) of four distinct stochastic processes in K'. This map allows to perceive certain K' frame diffusions as genuine diffusion processes in the inertial frame K.
Since all $e_i(\tilde{g})$ solve the eigenvalue problem $\Delta_{\tilde{g}} f(\tilde{g}) = (3/4) f(\tilde{g})$ the basis functions (4.16) solve it as well. A complete stochastic decoding of (4.16) amount then to a

repetition of the Dankel's strategy[21] once e's are cast in the canonical (Madelung) form $e' = exp(R + iS)$. Now $|e'_i(\tilde{g})|^2$ represents the probability distribution of the i-the stationary diffusion as perceived in K. The respective random variable is labelled by the proper time.

We may now formulate a definite answer to the question raised in Section 2. What is perceived in K as a stochastic rotational diffusion again are not diffusions associated with the forward time development exclusively i.e. $e_i(g)exp(-imc^2t'/\hbar)$ for all $i = 1, 2, 3, 4$. The answer is positive if we turn over to the rest frame diffusions associated with evolutions $e_j(g)exp(-imc^2t'/\hbar)$ for $j = 1, 3$ and $e_k(g)$ $exp(imc^2t'/\hbar)$ for $k = 2, 4$. The backward evolution for $k = 2, 4$ can be represented as a forward evolution again by invoking the arguments of Ref. 2. but is irreducibly different from the one associated with $e_k(g)exp(-imc^2t'/\hbar)$.

Our analysis allows to associate diffusions on S_3 with plane wave solutions of the Dirac equation, which is possible due to the implicit validity of the proper time Schrödinger equation on S_3. We deal then with rotational fluctuations which are intrinsic to a particle in uniform motion. There is no essential difficulty in extending the arguments to cases covered by the semiclassical regime for solutions of the Dirac equation in the presence of external electromagnetic fields (inhomogeneities include). The proper time evolution governed by the Bargmann - Michel - Telegdi equation[25] amounts to a purely rotational diffusion process which is effected along a space-time trajectory of the particle. Motions show the same feature: randomness is exclusively intrinsic and does not affect the space-time path followed by the origin of the rotating frame (it is rather on the contrary that the spin precession is strongly path dependent).

The problem we have left aside at the moment is the probabilistic analysis of general wave packet solutions of the Dirac equation, where a non-trivial input of the random process affecting a particle velocity (extrinsic randomness) is expected to show up.

Since random paths of stochastic mechanics are quite akin to Feynman paths[2,17] it should in principle be possible to establish a unifying framework for an increasing number of path integral approaches to the description of Dirac particles in the non-Grassmann vein [32,38–49]. It especially pertains to random walk representations of the Dirac propagator[49–52] where one generally assumes that at each step of the random walk executed by the spinning particle in Minkowski space, its quantization axis is rotated by a certain angle. Compare e.g. our discussion of Section 2 where momentum change induces a well defined rotation of the polarization.

References to numerous relativisation attempts in the context of Nelson's stochastic mechanics can be found in the recent papers[53,54] see also [50,18]. A problem worth a deeper exploration in the presented probabilistic framework is a physical meaning of different notions of position invented for the Dirac particle and eventually this of Zitterbewegung which from our perspective is definitely not the intrinsic mechanism [55,56] implementing the electron spin.

References

[1] E. Nelson, Quantum Fluctuations, Princeton Univ. Press, Princeton 1985

[2] P. Garbaczewski, Phys.Lett. **A 147**, (1990)168

[3] E. Santos, private discussion

[4] E. Nelson, Dynamical Theories of the Brownian Motion, Princeton Univ. Press, Princeton, 1967

[5] L. de la Pena, A.M. Cetto, J.Math.Phys. **18** (1977)1612

[6] P. Claverie, S. Diner, Int.J.Quantum Chem., 12 Suppl.1, (1977)41

[7] M. Lax, Rev.Mod. Phys. **38**(1966)541

[8] S. Chandrasekhar, Rev.Mod.Phys. **15**(1943)1

[9] H.A. Kramers, Physica **7**(1940)284

[10] P. Résibois, M. De Leener, Classical Kinetic Theory of Fluids, Wiley, NY, 1977

[11] K. Huang, Statistical Mechanics, Wiley, NY 1963

[12] G. Wilemski, J.Stat.Phys. **14**(1976)153

[13] A. J. Chorin, J.E. Marsden, A Mathematical Introduction to Fluid Mechanics, Springer, Berlin 1979

[14] G.K. Batchelor, An Introduction to Fluid Dynamics, Cambridge University Press, Cambridge, 1967

[15] T. Misava, Nuovo Cim.**51B**(1986)1

[16] T. Misawa, Nuovo Cim. **59B**(1987)179

[17] P. Garbaczewski, Phys. Lett. **A143** (1990)85

[18] Ph Blanchard, Ph. Combe, W. Zheng, Mathematical and Physical Aspects of Stochastic Mechanics, Springer Lecture Notes in Physics, vol. 281, Berlin 1987

[19] D. Bohm, B.J. Hiley, Phys.Reports **172**(1989)94

[20] E. Nelson, Phys. Rev. **150** (1966)1079

[21] T. Dankel, Jr. Arch.Rat. Mech. Anal. **37**(1970)192

[22] P. Garbaczewski, Fortschr. Phys. **38**(1990) 447

[23] F. Guerra, R. Marra, Phys. Rev. Lett. **50** (1983)1715

[24] I.M. Gelfand, R.A. Milnos, Z. Ya. Shapiro, Representations of the Rotation and Lorentz Groups and Their Applicaations, Pergamon, NY, 1963

[25] V. Bargmann, L. Michel, V.L. Telegdi, Phys. Rev. Lett. **2** (1959) 435

[26] J.D. Jackson, Classical Electrodynamics, Wiley, NY, 1975

[27] M. Rivas, J. Math. Phys. **30** (1989), 318

[28] K. Ito, H.P. McKean Jr. Diffusion Processes and Their Sample Paths, Academic Press, NY, 1965

[29] L. Brillouin, Tensors in Mechanics and Elasticity, Chap. VIII, Academic Press, NY, 1965

[30] B. Diu, Eur. J. Phys. **1** (1980) 231

[31] L. de la Peña, J. Math. Phys. **12** (1971) 453

[32] L. Schulman, Phys. Rev. **176** (1968) 1558

[33] J.P. Dahl, Kon. Dansk. Vid. Selsk, Mat-fys. **39** (1977), No 12

[34] F. Bopp, R. Haag, Z. Naturforsch. **5a** (1950), 644

[35] J.D. Bjorken, S.D. Drell, Relativistic Quantum Mechanics, McGraw-Hill, NY, 1964

[36] A.O. Barut, Phys. Rev. Lett. **20** (1968) 893

[37] S.I. Rubinow, J.B. Keller, Phys. Rev. **131** (1963),2789

[38] P. Garbaczewski, Phys. Lett. **A73** (1979) 280

[39] P. Garbaczewski, J. Math. Phys. **23** (1983), 442

[40] P. Garbaczewski, Classical and Quantum Field Theory of Exactly Soluble Nonlinear Systems, World Scientific, Singapore, 1985

[41] A.O. Barut, I.H. Duru, Physics Reports **172** (1989), 1

[42] A.O. Barut, I.H. Duru, Phys. Rev. Lett. **53** (1984) 2355

[43] M. Stone, Nucl. Phys. **B314** (1989) 557

[44] M.S. Plyushchay, Phys. Lett. **B236** (1990) 291

[45] A. Yu. Alekseev, S.L. Shatashvili, Mod. Phys. Lett. **A3** (1988) 1551

[46] T. Kashiwa, Int. J. Mod. Phys. **A5** (1990) 375

[47] P. Orland, Int. J. Mod Phys. **A4** (1989) 3615

[48] K. Johnson, Ann. Phys. (NY) **192**(1989) 104

[49] T. Nakagomi, Progr. Theor. Phys. **81** (1989) 916

[50] T. Zastawniak, J. Math. Phys. **30** (1989) 1354

[51] T. Jacobson, Phys. Lett. B216 (1989) 150

[52] J. Ambjorn, B. Durhuus, T. Jonsson, Nucl. Phys. **B330** (1990) 509

[53] M. Serva, Ann. IHP, **49** (1988), 415

[54] G.F. De Angelis, J. Math. Phys. **31** (1990) 1408

[55] A.O. Barut, A.J. Bracken, Phys. Rev. **D23** (1981) 2454

[56] A.O. Barut, A.J. Bracken, J. Math. Phys. **26** (1985) 2515

EXCURSIONS IN STOCHASTIC MECHANICS AND THE QUANTUM MECHANICS OF BROWNIAN MOTION

Aubrey Truman,
Department of Mathematics
and Computer Science,
University College of Swansea,
Singleton Park,
SWANSEA SA2 8PP
U.K.

David Williams,
Department of Pure Mathematics
and Mathematical Statistics,
University of Cambridge,
16 Mill Street
CAMBRIDGE CB2 1SB
U.K.

ABSTRACT

We give a brief expository account of some of our results on Poisson-Lévy excursion measures, explaining how the spectra of certain 1-dimensional Schrödinger operators determine the number of excursions for 1-dimensional time-homogeneous diffusions.

INTRODUCTION

Let B be a BM($\mathbb{R}$) process starting at the origin O, normalised so that $\mathbb{E}(B(s)\,B(t)) = \min(s, t)$. Hinchin's law of the iterated logarithm states that

$$\mathbb{P}\left\{ \overline{\lim_{t\downarrow 0}} \frac{B(t)}{(2t \ln\ln t^{-1})^{1/2}} = 1 \right\} = 1 \ . \tag{1}$$

Since $-B$ is a BM($\mathbb{R}$) process with $B(0) = 0$, we obtain

$$\mathbb{P}\left\{ \overline{\lim_{t\downarrow 0}} \frac{B(t)}{(2t \ln\ln t^{-1})^{1/2}} = 1 \ , \ \underline{\lim_{t\downarrow 0}} \frac{B(t)}{(2t \ln\ln t^{-1})^{1/2}} = -1 \right\} = 1 \ . \tag{2}$$

Thus, B visits O infinitely often in any time interval $[0, \varepsilon)$, no matter how small $\varepsilon > 0$. (See Refs.(1), (2), (3)).

Also, almost surely, $s \mapsto B(s)$ is continuous, so $\{s > 0 : B(s) \gtrless 0\}$ is almost surely an open subset of the real line. As such, it can be decomposed into a countable

number of component intervals - the excursion intervals. Although we cannot discuss the first excursion, we can discuss the probability distribution of excursions of duration $\geq \varepsilon$, for any fixed $\varepsilon > 0$. (Ref.(4)).

Definition

$$\text{Local time at O, } L^0(t) = \lim_{h\downarrow 0} h^{-1}\, \text{Leb}\left\{ s \in [o, t] : B(s) \in \left(-\tfrac{h}{2}, \tfrac{h}{2}\right) \right\} = \int_0^t \delta(B(s))ds.$$

Definition

Subordinator, $\gamma^0(t) = \inf\{ s > 0 : L^0(s) > t\}$. γ^0 is the inverse function of L^0.

$\gamma^0(t)$ is how long you have to wait until L^0 equals t and we can identify the jumps in $\gamma^0(t)$ with the excursions upto L^0 equals t. Lévy proved that for each λ, $t > 0$,

$$\mathbb{E}_0\left\{ e^{-\lambda\gamma^0(t)} \right\} = \exp\left\{ -t \int_0^\infty (1 - e^{-\lambda s})\, d\nu_0(s) \right\} \ , \tag{3}$$

with

$$\frac{d\nu_0(s)}{ds} = (2\pi s^3)^{-1/2} \ , \quad s > 0 \ . \tag{4}$$

(See Refs.(1), (2), (3), (4), (5))

Setting

$\#(s, t)$ = number of excursions of duration s to (s + ds), upto L^0 equals t, comparing terms in λ above, we obtain

$$\mathbb{P}(\#(s, t) = N) = e^{-t d\nu_0(s)} \frac{(t d\nu_0(s))^N}{N!} \ , \quad \text{for } N = 0, 1, 2, \ldots \tag{5}$$

Thus, the probability distribution of the number of excursions of duration $\geq s$, upto the local time at O equals t, is Poisson, with mean $t\nu_0[s, \infty)$.

Also, Lévy showed that, if

$N_\varepsilon(t)$ = number of excursions of duration $\geq \varepsilon$, upto time t,

then, almost surely,

$$\lim_{\varepsilon\downarrow 0} \frac{N_\varepsilon(t)}{\nu_0[\varepsilon,\infty)} = \lim_{\varepsilon\downarrow 0} \frac{N_\varepsilon(t)}{\left(\frac{2}{\pi}\right)^{1/2}\varepsilon^{-1/2}} = L^0(t) \quad . \tag{6}$$

(See Itô and McKean Ref.(1)).

We seek to generalise the above to 1-dimensional time-homogeneous processes X in $\mathbb{R}$

$$dX(s) = b(X(s))ds + dB(s) \quad , \quad X(0) = a \quad , \tag{7}$$

with inaccessible boundary $\{-\infty, \infty\}$, with C^1 drift b. We shall assume that:-

(i) $(b^2(x) + b'(x))$ is bounded below,

(ii) $(b^2(x) + b'(x)) \uparrow \infty$ as $|x| \uparrow \infty$,

(iii) $\exp\left(2\int^x b(u)du\right) \in L^1(\mathbb{R})$.

Example (1-dimensional Ornstein-Uhlenbeck Process) (Hawkes and Truman Ref.(6))

Consider $dX(s) = -X(s)ds + dB(s)$, with $X(0) = 0$, $L^0(t) = \int_0^t \delta(X(s))ds$, $\gamma^0(t) = \inf\{s>0 : L^0(s) > t\}$. For each λ, $t>0$, we have

$$\mathbb{E}_0\left\{e^{-\lambda\gamma^0(t)}\right\} = \exp\left\{-t\int_0^\infty (1-e^{-\lambda s})\, d\nu_0(s)\right\} \tag{8}$$

with

$$\frac{d\nu_0(s)}{ds} = \frac{2}{\sqrt{\pi}} e^{-s}(1-e^{-2s})^{-3/2} \quad , \quad s>0 \quad . \tag{9}$$

We seek to generalise the above to initial conditions $X(0) = a \neq 0$ for more general drifts.

The above O-U process is the Nelson ground-state process for the harmonic oscillator Hamiltonian

$$H = 2^{-1}\left(-\frac{d^2}{dx^2} + x^2\right) \quad , \quad \rho_0(x) = e^{-x^2} \quad , \tag{10}$$

ρ_0 being the invariant density, $f_0 = \rho_0^{1/2} \in L^2$ being the ground-state. For each $a \in \mathbb{R}$, for the above X with drift b, define the local time at a

$$L^a(t) = \lim_{h\downarrow 0} h^{-1}\,\text{Leb}\left\{ s\in[o,t] : X(s) \in \left(a-\frac{h}{2}, a+\frac{h}{2}\right)\right\} \tag{11}$$

and $\gamma^a(t) = \inf\{s>0 : L^a(s) > t\}$ so that $X(\gamma^a(t)) = a$. Running the process upto time $\gamma^a(t)$ is equivalent to running the process upto the local time at a equals t, so we can discuss the number of upward or downward excursions which X makes, upto L^a equals t. Here we shall exploit the fact that X is the Nelson diffusion process corresponding to the ground-state, f_0 , of the quantum Hamiltonian

$$H = 2^{-1}\left(-\frac{d^2}{dx^2} + b^2(x) + b'(x)\right) \quad , \tag{12}$$

$f_0 = \rho_0^{1/2}$, ρ_0 being the invariant density, $\rho_0(x) = \exp\left(2\int^x b(u)du\right)$. We want H to be a self-adjoint operator on $L^2(\mathbb{R})$ so we assume that H is limit point at $\pm\infty$. Assumption (i) ensures that this condition is satisfied.

Then, setting

$\#^{\pm}(s, t)$ = number of $\begin{matrix}\text{upward}\\ \text{downward}\end{matrix}$ excursions of duration s to (s + ds), upto L^a equals t,

in Ref.(7), we proved that for bounded b

$$\mathbb{P}(\#^{\pm}(s,t) = N) = e^{-t d\nu_a^{\pm}(s)} \frac{(t d\nu_a^{\pm}(s))^N}{N!} \quad , \quad \text{for } N = 0, 1, 2, \ldots, \tag{13}$$

where for $s>0$

$$\frac{d\nu_a^{\pm}(s)}{ds} = \left(f_0, (H_{\pm}(a))^2 e^{-sH_{\pm}(a)} f_0\right)_{L^2} f_0^{-2}(a) \quad , \tag{14}$$

$H_{\pm}(a)$ being the Dirichlet Hamiltonians

$$H_{\pm}(a) = \lim_{\lambda\uparrow\infty} (H + \lambda\chi_{\mp}(a)) \quad , \tag{15}$$

$\chi_{\mp}(a)$ being the characteristic function of $\{x : x \lesseqgtr a\}$, respectively. Thus, for each $s > 0$, $\nu^{\pm}[s, \infty) = \left(f_0, H_{\pm}(a)\ e^{-sH_{\pm}(a)} f_0\right)_{L^2}$, so the quantum Hamiltonians $H_{\pm}(a)$ give the Poisson-Lévy excursion measures. Below we given an outline proof of this result in the limit point case when b is unbounded and investigate some of its consequences.

Assumptions (i) and (ii) guarantee that $H_{\pm}(a)$ are self-adjoint on a suitable domain $\mathcal{D}(H_{\pm}(a))$, with discrete spectrum $\{\alpha_n^{\pm}\}_{n=1,2,3\ldots}$ One way to find $\alpha_n^{\pm}$ is to utilise the limit point behaviour of $H = 2^{-1}\left(-\frac{d^2}{dx^2} + V(x)\right)$. This ensures that H has unique L^2 eigenfunctions near $\pm\infty$, $f_{\alpha}^{\pm}(\cdot)$, $\alpha > 0$, with ground-state eigenfunction $f_0(x) = \exp\left(\int^x b(u)du\right) \in L^2(\mathbb{R})$. Then $\alpha = \alpha_n^{\pm}\ (> 0)$ are the successive roots of

$$f_{\alpha}^{\pm}(a) = 0 \tag{16}$$

and for $x \gtrless a$ respectively we define

$$f_n^{\pm}(x) = f_{\alpha_n^{\pm}}^{\pm}(x)\ \left\{\int_{y \gtrless a} \left|f_{\alpha_n^{\pm}}^{\pm}(y)\right|^2 dy\right\}^{-1/2} \quad , \tag{17}$$

with, needless to say, $\mathcal{D}(H_{\pm}(a)) = \left\{f \in L^2(\mathbb{R}\pm) : \sum_1^{\infty} (\alpha_n^{\pm})^2 \left|(f_n^{\pm}, f)\right|^2 < \infty\right\}$.

Defining

$\#_{\varepsilon}^{\pm}(t)$ = number of $\begin{matrix}\text{upward}\\ \text{downward}\end{matrix}$ excursions from a of duration $\geq \varepsilon$, upto L^a equals t,

following Lévy and Itô and McKean, we prove that, almost surely,

$$\lim_{\varepsilon\downarrow 0} \frac{\#_{\varepsilon}^{\pm}(t)}{(f_0, H_{\pm}(a)\, e^{-\varepsilon H_{\pm}(a)} f_0)_{L^2}} = \lim_{\varepsilon\downarrow 0} \frac{\#_{\varepsilon}^{\pm}(t)}{\sum_1^{\infty} \alpha_n^{\pm}\, e^{-\varepsilon\alpha_n^{\pm}} \left|(f_0, f_n^{\pm})\right|^2} = \frac{t}{f_0^2(a)} \quad , \tag{18}$$

for each $a \in \mathbb{R}$. The point is that $e^{-\varepsilon H_{\pm}(a)} f_0 \in \mathcal{D}(H_{\pm}(a))$ for $\varepsilon > 0$ but not for $\varepsilon = 0$ and the divergence of the denominator gives the infinitesimal excursions which X makes from a. As we shall see the O-U process is of the above type.

NELSON'S STOCHASTIC MECHANICS AND DIFFUSIONS IN 1-DIMENSION

In this section we collect together some useful results connecting 1-dimensional time-homogeneous diffusions and Nelson's stochastic mechanics. In particular we calculate certain probability distributions in terms of quantum mechanical Hamiltonians. (See Refs.(7), (8), (9)).

Recall the Schrödinger equation for a unit mass quantum mechanical particle in the force field $-V'(x)$,

$$i\frac{\partial\Psi}{\partial t} = -2^{-1}\frac{\partial^2\Psi}{\partial x^2} + V(x)\Psi \quad , \tag{19}$$

$\Psi = \Psi(x, t)$ being the Schrödinger wave-function, $(x, t) \in \mathbb{R} \times \mathbb{R}_+$. Here the quantum mechanical particle density is $\rho = \Psi^*\Psi$, * being the complex conjugate.

Write $\Psi = e^{R+is}$, for real-valued R and S, and equate real and imaginary parts in above giving

$$\frac{\partial R}{\partial t} = \frac{\partial R}{\partial x}\frac{\partial S}{\partial x} + 2^{-1}\frac{\partial^2 S}{\partial x^2} \tag{20}(a)$$

and

$$\frac{\partial S}{\partial t} = 2^{-1}\left\{\left(\frac{\partial R}{\partial x}\right)^2 - \left(\frac{\partial S}{\partial x}\right)^2 + \frac{\partial^2 R}{\partial x^2}\right\} - V \quad . \tag{20}(b)$$

Nelson showed that Eq.(20)(a) is the forward Kolmogorov equation for the diffusion

$$dX(t) = b(X(t), t)dt + dB(t) \quad , \tag{21}$$

with drift $b = \frac{\partial}{\partial x}(R + S)$ and density $\rho = e^{2R}$. Nelson also showed that Eq.(20)(b) is equivalent to

$$2^{-1}(D_+D_- + D_-D_+)\,X(t) = -V'(X(t)) \quad , \tag{22}$$

$D_\pm$ being mean forward and backward time derivatives

$$D_\pm f(X(t), t) = \lim_{h\downarrow 0} \mathbb{E}\left\{ \frac{f(X(t \pm h), t \pm h) - f(X(t), t)}{\pm h} \,\middle|\, X(t) \right\} \quad . \tag{23}$$

The paths $s \mapsto X(s)$ are almost surely nowhere differentiable. Eq.(22) is the stochastic analogue of Newton's 2nd Law of Motion for the nowhere differentiable path X. (See Ref.(9)).

Specialize now to stationary states $\Psi_E(x, t) = e^{-iEt} f_E(x)$, where $f_E \in L^2(\mathbb{R})$ is C^2, with

$$(H - E) f_E(x) = 0 \quad , \tag{24}$$

$H = -2^{-1}\dfrac{d^2}{dx^2} + V(x)$, V being piecewise continuous, E being inf spec (H), $f_E > 0$ being the ground-state.

The corresponding Nelson diffusion is

$$dX(t) = \frac{d}{dX(t)} \ln\left(f_E(X(t))\right) dt + dB(t) \quad , \tag{25}$$

with drift $b = f_E'/f_E$ which we assume is Lipschitz. Because $f_E \in L^2$, $\int^y f_E^{-2}(x)\,dx$ diverges as $y \to \pm\infty$, which guarantees an inaccessible boundary $\{-\infty, \infty\}$. (Refs.(7) (11)).

A simple computation gives the generator L for the diffusion in terms of H,

$$L_x = 2^{-1}\frac{d^2}{dx^2} + \frac{f_E'(x)}{f_E(x)}\frac{d}{dx} = -f_E^{-1}(x)\,(H - E)\,f_E(x) \quad . \tag{26}$$

Assume now that V is bounded below, say $V > 0$. The semiboundedness and continuity of V ensure that the semigroup generated by H is given by a Feynman-Kac formula (Ref.(12)). Using the inaccessible boundary, the Girsanov-Cameron-Martin theorem gives the transition density for X.

$$p_t(x, y) = f_E^{-1}(x) \exp\left\{-t\,(H - E)\right\}(x, y)\, f_E(y) \quad . \tag{27}$$

Also, from the Feynman-Kac formula the semiboundedness of V gives

$$p_t(x, y) \le f_E^{-1}(x)\,(2\pi t)^{-1/2} \exp\left\{-\frac{(x - y)^2}{2t}\right\} f_E(y) \quad , \tag{28}$$

so far $\lambda > 0$

$$\int_0^\infty e^{-\lambda t}\, p_t(x, y)\, dt \stackrel{\text{def}}{=} \tilde{p}_\lambda(x, y) < \infty \quad . \tag{29}$$

Setting $\tau_x(a) = \inf\{s > 0 : X(s) = a \,|\, X(0) = x\}$, because X is a 1-dimensional time-homogeneous process, for any point a intermediate to x and y,

$$p_t(x, y) = \int_0^t \mathbb{P}(\tau_x(a) \in du) p_{t-u}(a, y) \quad . \tag{30}$$

Taking Laplace transforms, using joint continuity, gives

$$\mathbb{E}\left\{e^{-\lambda\tau_x(a)}\right\} = \tilde{p}_\lambda(x, a)/\tilde{p}_\lambda(a, a) \quad . \tag{31}$$

Thus, we obtain

$$\mathbb{E}\left\{e^{-\lambda\tau_x(a)}\right\} = f_E^{-1}(x)\,(H + \lambda - E)^{-1}(x, a)\, f_E(a)/(H + \lambda - E)^{-1}(a, a) \quad . \tag{32}$$

We need a stronger result. Define the Dirichlet Hamiltonians $H_\pm(a)$ by

$$H_\pm(a) = \lim_{\lambda\uparrow\infty} (H + \lambda\chi_\mp(a)) \quad , \tag{33}$$

$\chi_\mp(a)$ being the characteristic function of $\{x : x \lesseqgtr a\}$, respectively. Then we have:

PROPOSITION

For V as above, for $x \gtrless a$, respectively,

$$\mathbb{P}(\tau_x(a) > t) = f_E^{-1}(x)\left\{\exp\left(-t\,(H_\pm(a) - E)\right) f_E(x)\right\} \quad . \tag{34}$$

Outline Proof

The inaccessibility of the boundary enables us to use the Girsanov-Cameron-Martin theorem. The desired result follows from the Girsanov-Cameron-Martin theorem and the fact that for $x > a$

$$\mathbb{P}(\tau_x(a) > t) = \lim_{\lambda \uparrow \infty} \mathbb{E}_x \left\{ e^{-\lambda \int_0^t \chi_-(a)(X(s)ds)} \right\} , \qquad (35)$$

together with the corresponding result for $x < a$. //

We also need a beautiful result of Mark Kac. We drop the a label on $\chi_{\pm}(a)$.

Lemma

Let f_0 and V be piecewise continuous. Assume that $V > 0$ and that for all $x \in \mathbb{R}$ and some $\alpha > 0$

$$\int_{-\infty}^{\infty} |f_0(x+y)| e^{-\alpha|y|} dy < \infty \quad .$$

Define $u = u(x)$ by

$$u(x) = f_0^{-1}(x) \int_0^{\infty} e^{-\alpha t} \mathbb{E}_x \left\{ e^{-\int_0^t (V+\lambda\chi_-)(B(s))ds} f_0(B(t)) \right\} e^{Et} dt \quad . \qquad (36)$$

Then u is the piecewise C^2 solution of

$$\left(L_x - (\lambda\chi_-(x) + \alpha)\right) u(x) = -1 \quad , \qquad (37)$$

with $u(a+) = u(a-)$ and $u'(a+) = u'(a-)$, $L = -f_0^{-1}(H-E)f_0$, $H = -2^{-1}\frac{d^2}{dx^2} + V(x)$.

(A nice account of this is given in Karatzas and Shreve Ref.(13)).

We can now prove the main result of this section - the generalised arc-sine law. We assume that $H = \left(-2^{-1}\frac{d^2}{dx^2} + V(x)\right)$ is limit point at $\pm\infty$ (a natural assumption since we want H to be self-adjoint).

PROPOSITION (Batchelor and Truman Ref.(10))

For $L^{\pm}(t) = \text{Leb}\left\{ s \in [0,t] : X(s) \gtrless a \right\}$ for the above X

$$\int_0^{\infty} e^{-\alpha t} \mathbb{E}_a \left\{ e^{-\lambda_- L^-(t) - \lambda_+ L^+(t)} \right\} dt = \frac{\text{Disc}|_a (\alpha+\lambda)^{-1} \frac{d}{dx} \mathbb{E}\left\{ e^{-(\alpha+\lambda)\tau_x(a)} \right\}}{\text{Disc}|_a \frac{d}{dx} \mathbb{E}\left\{ e^{-(\alpha+\lambda)\tau_x(a)} \right\}} , \qquad (38)$$

for each $\alpha, \lambda_{\pm} > 0$, with $\lambda = \lambda(x) = \lambda_{\pm}$ for $x \gtrless a$, respectively.

Outline Proof

Since $L^+(t) = t - L^-(t)$, defining $f(\alpha, \lambda_1, \lambda_2) = \int_0^{\infty} \mathbb{E}_x \left\{ e^{-\lambda_- L^-(t) - \lambda_+ L^+(t)} \right\} e^{-\alpha t} dt$,

$$f(\alpha + \lambda_1, \lambda_1 - \lambda_2, 0) = f(\alpha, \lambda_1, \lambda_2) \quad ,$$

all arguments being positive. Thus, we only need prove above result for $\lambda_+ = 0$, $\lambda_- = \lambda > 0, \alpha > 0$.

Define

$$u(x) = \int_0^{\infty} e^{-\alpha t} \mathbb{E}_x \left\{ e^{-\lambda L^-(t)} \right\} dt = \int_0^{\infty} e^{-\alpha t} \mathbb{E}_x \left\{ e^{-\lambda \int_0^t \chi_-(X(s)ds)} \right) \right\} dt \quad . \qquad (39)$$

The Girsarov-Cameron-Martin theorem gives

$$u(x) = f_E^{-1}(x) \int_0^{\infty} e^{-\alpha t} \mathbb{E}_x \left\{ e^{-\int_0^t (V+\lambda\chi_-)(B(s))ds} f_E(B(t)) \right\} e^{Et} dt \quad , \qquad (40)$$

with $f_E \in L^2(\mathbb{R})$.

From Kac's result, u is the piecewise C^2 solution of

$$(L_x - \lambda\chi_-(x) - \alpha)u(x) = -1 \quad , \qquad (41)$$

with $u(a+) = u(a-)$ and $u'(a+) = u'(a-)$. The limit point hypothesis ensures that there is a unique C^1 solution of $L_x h(x) = \gamma h(x)$ bounded in a neighbourhood of $x = \pm\infty$. For, if there were 2 solutions, h_1 and h_2, there would be 2 linearly independent L^2

eigenfunctions of H at $\pm\infty$, $(h_1 f_E)$ and $(h_2 f_E)$. From above, the unique bounded C^1 eigenfunction of L_x is $\mathbb{E}\{e^{-\gamma\tau_x(a)}\}$.

Therefore, we write

$$u(x) = (\alpha+\lambda)^{-1} + A\,\mathbb{E}\{e^{-(\alpha+\lambda)\tau_x(a)}\} \quad , \; x < a \; ,$$
$$u(x) = \alpha^{-1} + B\,\mathbb{E}\{e^{-\alpha\tau_x(a)}\} \quad , \; x > a$$

and satisfying the boundary conditions gives

$$u(a) = \mathrm{Disc}\big|_a (\alpha+\lambda)^{-1}\frac{d}{dx}\mathbb{E}\{e^{-(\alpha+\lambda)\tau_x(a)}\} \Big/ \mathrm{Disc}\big|_a \frac{d}{dx}\mathbb{E}\{e^{-(\alpha+\lambda)\tau_x(a)}\} \; ,$$

with $\lambda = \lambda(x) = \lambda_\pm$ for $x \gtrless a$, $\lambda_- = \lambda$, $\lambda_+ = 0$ .//

All the above results are applicable to any 1-dimensional time-homogeneous diffusion

$$dX(s) = b(X(s))ds + dB(s) \quad , \tag{42}$$

with L^1 invariant density $\rho_0(x) = C\exp\left(2\int^x b(u)du\right)$, b being a Lipschitz drift. The point is that $f_0 = \rho_0^{1/2} = C^1 \exp\left(\int^x b(u)du\right)$ satisfies

$$\frac{f_0'}{f_0} = b \tag{43}$$

and differentiating w.r.t. x

$$\frac{f_0''}{f_0} - b^2 = b' \tag{44}$$

i.e.

$$\left(-2^{-1}\frac{d^2}{dx^2} + V(x) - E\right)f_0(x) = 0 \tag{45}$$

for $V(x) = 2^{-1}(b^2(x) + b'(x))$ and $E = 0$. By hypothesis $f_0 \in L^2$. (See Ref.(7)). Nelson's results give

$$2^{-1}(D_+D_- + D_-D_+)\,X(s) = -V'(X(s)) \quad , \tag{46}$$

for $V = 2^{-1}(b^2 + b')$, in the state with density ρ_0. (See Ref.(9)). In the next section we show that if V is bounded below $H = -2^{-1}\left(\frac{d^2}{dx^2}\right) + V(x)$ is limit point at $\pm\infty$.

SUBORDINATORS AND EXCURSIONS

Let X be the 1-dimensional time-homogeneous process X in $\mathbb{R}$, satisfying

$$dX(t) = b(X(t))dt + dB(t), \quad X(0) = a \; , \tag{47}$$

with a C^1 Lipschitz continuous drift b. Now let the corresponding potential $V = 2^{-1}(b^2 + b')$ be continuous and bounded below. Further assume that the invariant density $\rho_0(x) = C\exp\left(2\int^x b(u)du\right) \in L^1$ and is bounded. Some of these assumptions can be relaxed (e.g. to allow V to be piecewise continuous), but they remove a lot of technical difficulties here.

PROPOSITION

For the above V, $H = -2^{-1}\left(\frac{d^2}{dx^2}\right) + V(x)$ is limit point at $\pm\infty$.

Outline Proof

Recall the well-known sufficient conditions for V to be limit point at $+\infty$, namely the existence of a positive differentiable function M(x) with (i) $V(x) \geq -M(x)$, (ii) $\int_1^\infty (M(x))^{-1/2}\,dx = \infty$, (iii) $M'(x)/(M(x))^{3/2}$ bounded near $+\infty$. (See e.g. Ref.(14))

Set $M(x) = C$, for a constant $C > 0$ .//

For the local time at a, $L^a(t) = \lim_{h\downarrow 0} h^{-1}\,\mathrm{Leb}\left\{s \in [o, t] : X(s) \in \left(a - \frac{h}{2}, a + \frac{h}{2}\right)\right\}$, we set inverse $\gamma^a(t) = \inf\{s > 0 : L^a(s) > t\}$.

Evidently $X(\gamma^a(t)) = a$ and we can run X upto time $\gamma^a(t)$ i.e. until local time at a equals t. As before, $s \mapsto X(s)$ is almost surely continuous so $\{s > 0 : X(s) \gtrless a\}$ is almost surely an open subset of $\mathbb{R}$ which can be decomposed into its component excursion intervals. If we consider these excursions upto time $\gamma^a(t)$, then these excursions are characterised as the jumps in $\gamma^a(t)$. We set

$\#_\varepsilon(t)$ = number of jumps of duration $\geq \varepsilon$, upto local time at a equals t,

and

$N_\varepsilon(t)$ = number of jumps of duration $\geq \varepsilon$, upto time t,

i.e.

$$N_\varepsilon(t) = \#_\varepsilon(L^a(t)), \qquad \text{each } \varepsilon > 0 .$$

Subordination Property

The all-important property of $\gamma^a(t)$ is the fact that $\gamma^a(t)$ is a stopping time i.e. for probability space Ω, $\{\omega \in \Omega : \gamma^a(t)(\omega) < s\} \in \mathcal{F}_s$, $s > 0$, $\mathcal{F}_s$ being the usual filtration. This means that each $\omega \in \Omega$ starts afresh at $\gamma^a(t)$. Moreover, we have the subordination property

$$\gamma^a(t+s)(\omega) - \gamma^a(t)(\omega) = \gamma^a(s)(\theta_{\gamma^a(t)}(\omega)), \qquad s, t > 0, \qquad \omega \in \Omega, \qquad (48)$$

$\theta_{\gamma^a(t)}$ being the shift

$$\theta_{\gamma^a(t)}\,\omega(s) = \omega(s + \gamma^a(t)) - \omega(\gamma^a(t)) , \qquad \omega \in \Omega \qquad (49)$$

(See Rogers and Williams Ref.(4)).

The strong Markov property then gives the existence of a constant $\Psi(\lambda)(>0)$ for which

$$\mathbb{E}_a\left\{\exp\left(-\lambda\gamma^a(t)\right)\right\} = \exp\left\{-t\Psi(\lambda)\right\} , \qquad (50)$$

each $\lambda, t > 0$. We also have the beautiful proposition due to Lévy.

PROPOSITION

For the Markov diffusion process X, with bounded continuous resolvent kernel $(x, y) \mapsto \tilde{p}_\lambda(x, y)$,

$$\tilde{p}_\lambda(x, y) = \int_0^\infty e^{-\lambda s}\, p_s(x, y)\, ds \qquad (\lambda > 0) , \qquad (51)$$

for each $\lambda, t > 0$

$$\mathbb{E}_a\left\{\exp\left(-\lambda\gamma^a(t)\right)\right\} = \exp\left\{\frac{-t}{\tilde{p}_\lambda(a, a)}\right\} . \qquad (52)$$

Outline Proof

The key observation is that almost surely γ^a is the inverse of L^a. Thus, from above

$$(\Psi(\lambda))^{-1} = \int_0^\infty \mathbb{E}_a\left\{\exp\left(-\lambda\gamma^a(t)\right)\right\} dt = \mathbb{E}_a\left\{\int_0^\infty e^{-\lambda s}\, d\, L^a(s)\right\} . \qquad (53)$$

Therefore

$$\left(\Psi(\lambda)\right)^{-1} = \int_0^\infty e^{-\lambda s}\, d_s\left(\mathbb{E}_a(L^a(s))\right) \qquad (54)$$

and for $s > 0$

$$\mathbb{E}_a\left\{L^a(s)\right\} = \mathbb{E}_a\left\{\int_0^s \delta(X(u) - a)du\right\} = \int_0^s p_u(a, a)\, du \ .//$$

COROLLARY

For the Markov diffusion X, with bounded continuous resolvent kernel $\tilde{p}_\lambda(\cdot,\cdot)$ and invariant measure ρ_0, for each $\lambda, t > 0$,

$$\mathbb{E}_a\left\{\exp\left(-\lambda\gamma^a(t)\right)\right\} = \exp\left\{-\lambda t\, \rho_0^{-1}(a)\int_{-\infty}^\infty \rho_0(x)\, \mathbb{E}\left\{e^{-\lambda\tau_x(a)}\right\}dx\right\}, \qquad (55)$$

for each $a \in \mathbb{R}$.

Outline Proof

Recall that from the last section

$$\mathbb{E}\left\{e^{-\lambda\tau_x(a)}\right\} = \tilde{p}_\lambda(x, a)/\tilde{p}_\lambda(a, a) . \qquad (56)$$

Multiply both sides of this equation by $\rho_0(x)$. By Fubini's thorem

$$\int_{-\infty}^\infty \rho_0(x)\, dx\left\{\int_0^\infty e^{-\lambda s}\, p_s(x, a)ds\right\} = \int_0^\infty e^{-\lambda s}\, ds\left\{\int_{-\infty}^\infty \rho_0(x)\, p_s(x, a)dx\right\} = \lambda^{-1}\rho_0(a) \ .//$$

For $L^\pm(t) = \text{Leb}\{s \in [0, t] : X(s) \gtrless a\}$ we consider $\mathbb{E}_a\left\{\exp\left(-\lambda L^\pm(\gamma^a(t))\right)\right\}$, $\lambda > 0$. Then, because $L^\pm(\gamma^a(t))$ satisfy the subordination property, by the strong Markov property there exist positive constants $\Psi^\pm(\lambda)$ for which

$$\mathbb{E}_a\left\{\exp\left(-\lambda L^{\pm}(\gamma^a(t))\right)\right\} = \exp\left\{-t\,\Psi^{\pm}(\lambda)\right\} \quad . \tag{57}$$

For each $t>0$, $L^+(t)+L^-(t) = t$, so we obtain

$$L^+(\gamma^a(t)) + L^-(\gamma^a(t)) = \gamma^a(t) \tag{58}$$

and, by the independence of $L^{\pm}(\gamma^a(t))$, for each $\lambda>0$

$$\Psi^+(\lambda) + \Psi^-(\lambda) = \Psi(\lambda) \quad . \tag{59}$$

This suggests that

$$\Psi^{\pm}(\lambda) = \lambda\rho_0^{-1}(a)\int_{x\gtrless a}\rho_0(x)\,\mathbb{E}\left\{e^{-\lambda\tau_x(a)}\right\}dx \quad , \tag{60}$$

or from the last section

$$\Psi^{\pm}(\lambda) = \rho_0^{-1}(a)\int_0^{\infty} ds\,(1-e^{-\lambda s})\,\frac{d}{ds}\left(\rho_0^{1/2}, e^{-sH_{\pm}(a)}\rho_0^{1/2}\right)_{L^2} \quad . \tag{61}$$

If $H_{\pm}(a)$ is defined on a domain $\mathcal{D}(H_{\pm}(a)) \ni e^{-sH_{\pm}(a)}\rho_0^{1/2}$, each $s>0$, we obtain

$$\Psi^{\pm}(\lambda) = \int_0^{\infty}(1-e^{-\lambda s})\,d\nu^{\pm}(s) \quad , \tag{62}$$

with

$$\nu^{\pm}[s,\infty) = \rho_0^{-1}(a)\left(\rho_0^{1/2}, H_{\pm}(a)\,e^{-sH_{\pm}(a)}\rho_0^{1/2}\right)_{L^2} \quad . \tag{63}$$

PROPOSITION (Truman and Williams Refs.(7) and (8).)

For the above diffusion X, with a C^1 Lipschitz drift b, with a bounded invariant density $\rho_0(x) = \exp\left(2\int^x b(u)du\right) \in L^1$ and with a corresponding potential $V = 2^{-1}(b^2+b')$ bounded below, for each $\lambda, t>0$

$$\mathbb{E}_a\left\{\exp\left(-\lambda L^{\pm}(\gamma^a(t))\right)\right\} = \exp\left\{-t\int_0^{\infty}(1-e^{-\lambda s})\,d\nu^{\pm}(s)\right\} \quad , \tag{64}$$

where

$$\nu^{\pm}[s,\infty) = \rho_0^{-1}(a)\left(\rho_0^{1/2}, H_{\pm}(a)\,e^{-sH_{\pm}(a)}\rho_0^{1/2}\right)_{L^2} \quad , \tag{65}$$

$H_{\pm}(a)$ being the Dirichlet Hamiltonians on $L^2(\mathbb{R}_{\pm})$

$$H_{\pm}(a) = \lim_{\lambda\uparrow\infty}\left(-2^{-1}\frac{d^2}{dx^2} + V(x) + \lambda\chi_{\mp}(a)(x)\right) \quad , \tag{66}$$

with domains $\mathcal{D}(H_{\pm}(a)) \ni e^{-sH_{\pm}(a)}\rho_0^{1/2}$, $s>0$.

Outline Proof

To complete the proof we merely show that

$$\frac{\Psi^-(\alpha+\lambda_-)}{\Psi^+(\alpha+\lambda_+)} = \frac{-\frac{d}{dx}\,\mathbb{E}\left\{e^{-(\alpha+\lambda_-)\tau_x(a)}\right\}\Big|_{x=a_-}}{\frac{d}{dx}\,\mathbb{E}\left\{e^{-(\alpha+\lambda_+)\tau_x(a)}\right\}\Big|_{x=a_+}} \quad , \tag{67}$$

or equivalently

$$\frac{\Psi^-(\alpha+\lambda_-)}{\Psi^+(\alpha+\lambda_+)} = \frac{(\alpha+\lambda_-)\,\bar{\mathbb{E}}^-\left\{e^{-(\alpha+\lambda_-)\tau_x(a)}\right\}}{(\alpha+\lambda_+)\,\bar{\mathbb{E}}^+\left\{e^{-(\alpha+\lambda_+)\tau_x(a)}\right\}} \quad , \tag{68}$$

for $\bar{\mathbb{E}}_{\pm}(\cdot) = \int_{x\gtrless a} dx\,\rho_0(x)\,\mathbb{E}_x(\cdot)$.

We consider a Poisson process of + marks with a clock $(\lambda_+ + \alpha)L^+(t)$ and a Poisson process of - marks with a clock $(\lambda_- + \alpha)L^-(t)$. Let T_+ and T_- be the times of first + and - marks. Then

$$\mathbb{P}(T_+ < T_-) = \mathbb{P}\left(L^a(T_+) < L^a(T_-)\right) \quad . \tag{69}$$

But $L^a(T_{\pm})$ is exponential with rate $\Psi^{\pm}(\alpha+\lambda_{\pm}) \stackrel{\text{def}}{=\!=} q_{\pm}$ because $\mathbb{P}\left(L^a(T_{\pm}) < t\right)$ $= \mathbb{P}(T_{\pm} < \gamma^a(t)) = \mathbb{E}_a\left\{e^{-(\alpha+\lambda_{\pm})L^{\pm}(\gamma^a(t))}\right\}$. Thus, from above

$$\mathbb{P}(T_+ < T_-) = \int_0^{\infty} q_+ e^{-tq_+}\,dt\; e^{-tq_-} = q_+/(q_+ + q_-) \,. \tag{70}$$

Now colour + marks red with probability $\alpha/(\alpha+\lambda_+)$ and green with probability $\lambda_+/(\alpha+\lambda_+)$. Treating - marks the same, we obtain

$\mathbb{P}$ (first red mark precedes first green mark)

$$= \mathbb{P}(T_+ < T_-)\,\frac{\alpha}{(\lambda_+ + \alpha)} + \mathbb{P}(T_- < T_+)\,\frac{\alpha}{(\lambda_- + \alpha)}$$

$$= \frac{q_+}{(q_+ + q_-)}\,\frac{\alpha}{(\lambda_+ + \alpha)} + \frac{q_-}{(q_+ + q_-)}\,\frac{\alpha}{(\lambda_- + \alpha)} \quad . \qquad (71)$$

But above Poisson processes are equivalent to a Poisson process of red marks with a clock (αt) and a Poisson process of green marks with a clock $(\lambda_+ L^+(t) + \lambda_- L^-(t))$.

Therefore,

$$\mathbb{P}\text{ (first red mark precedes first green mark)} = \int_0^\infty \alpha e^{-\alpha t}\,dt\;\mathbb{E}_a\left(e^{-\lambda_+ L^+(t) - \lambda_- L^-(t)}\right).$$

Comparing this result with the Batchelor-Truman formula gives Eq.(68).//

An extra assumpton on V makes the spectral analysis of $H_\pm(a)$ easy.

PROPOSITION

If in addition $V(x) = 2^{-1}(b^2(x) + b'(x)) \uparrow \infty$ as $|x| \uparrow \infty$, the Dirichlet Hamiltonians $H_\pm(a)$ are self-adjoint on $\mathcal{D}(H_\pm(a))$, with discrete spectra $\{\alpha_n^\pm(a)\}$, where $\alpha = \alpha_n^\pm(a)$ are the successive positive roots of

$$f_\alpha^\pm(a) = 0 \quad , \qquad (72)$$

$f_\alpha^\pm(a)$ being the unique L^2 eigenfunctions near $\pm\infty$, respectively, of the differential operator, $H = -2^{-1}\frac{d^2}{dx^2} + V(x)$.

Outline Proof

This is a standard result of singular boundary value theory. (See e.g. p.524 in Ref.(15)). The point is that, setting $\{f_n^\pm\}_{n=1,2,\ldots} = \{f_{\alpha_n^\pm}^\pm\}_{n=1,2,\ldots}$, after normalising $\{f_n^\pm\}_{n=1,2,\ldots}$ is a complete o.n.b. for $L^2(\mathbb{R}^\pm)$ and $\mathcal{D}(H_\pm(a)) = \left\{ f \in L^2(\mathbb{R}^\pm) : \exists \text{ a finite } N \text{ with } f = \sum_1^N (f_n^\pm, f) f_n^\pm \right\}$ is a domain on which $H_\pm(a)$ is

essentially self-adjoint, $H_\pm(a)\,f = \sum_1^\infty \alpha_n^\pm\,(f_n^\pm, f) f_n^\pm)$, being the unique self-adjoint extension to $\mathcal{D}(H_\pm(a)) = \left\{ f \in L^2(\mathbb{R}^\pm) : \sum_1^\infty (\alpha_n^\pm)^2 |(f_n, f)|^2 < \infty \right\}$. (See Ref.(15)).//

In this case, after normalising $f_n^\pm$,

$$v_a^\pm[s, \infty) = \sum_1^\infty \alpha_n^\pm\, e^{-\alpha_n^\pm s}\, |(f_0, f_n^\pm)|^2\, f_0^{-2}(a) \quad , \qquad (73)$$

$f_0 = \rho_0^{1/2}$, and for excursions of X away from a

$$\mathbb{P}(\#_s^\pm(t) = N) = e^{-t v_a^\pm[s,\infty)}\,\frac{(t v_a^\pm[s, \infty))^N}{N!} \quad , \text{ for } N = 0, 1, 2, \ldots \qquad (74)$$

i.e. the spectrum of a 1-dimensional Schrödinger operator determines the distribution of excursions.

Example (Ornstein-Uhlenbeck Process)

Here $f_\alpha^\pm$ is given by

$$f_\alpha^\pm(x) = D_{\alpha - 1/2}(\pm x) = e^{-x^2/4}\,\frac{\Gamma(\frac{1}{2} + \alpha)}{2\pi i}\int_\gamma e^{\pm xs}\, e^{-s^2/2}\, s^{-\alpha - 1/2}\, ds \quad , \qquad (75)$$

γ being the contour starting just below the cut on the $-$ve reals at $-\infty$, looping once around the origin O, finishing just above the cut at $-\infty$. (See Abramowitz and Stegun Ref.(16) and Davies Ref.(17)).

Imitating Itô and McKean (Ref.(11)), we can also prove with the above assumptions:

PROPOSITION

With probability one,

$$\lim_{\varepsilon \downarrow 0} \frac{\#_\varepsilon^\pm(t)}{(\rho_0^{1/2},\, H_\pm(a)\, e^{-\varepsilon H_\pm(a)}\, \rho_0^{1/2})_{L^2}} = \frac{t}{\rho_0(a)} \quad , \qquad (76)$$

ρ_0 being the invariant density.

Outline Proof

Given that $\nu_a^+[\varepsilon, \infty)$ is continuous, $\varepsilon > 0$, and diverges to infinity as $\varepsilon \downarrow 0$, define

$$\nu_+^{-1}(s) = \inf \{ t > 0 : \nu_a^+[t, \infty) \leq s \} . \tag{77}$$

Then

$$\lim_{\varepsilon \downarrow 0} \frac{\#_\varepsilon^+(t)}{\nu_a^+[\varepsilon, \infty)} = \lim_{s \uparrow \infty} \frac{\#^+_{\nu_+^{-1}(s)}(t)}{s} .$$

But

$$\mathbb{E}_a \{ \#^+_{\nu_+^{-1}(s)}(t) / s \} = \frac{t \nu_a^+[\nu_+^{-1}(s), \infty)}{s} = \frac{ts}{s} = t . \tag{78}$$

The independent Poisson nature of $\#^+(s, t)$ and the law of large numbers then give, almost surely, for each $t > 0$

$$\lim_{\varepsilon \downarrow 0} \frac{\#_\varepsilon^+(t)}{\nu_a^+[\varepsilon, \infty)} = t , \tag{79}$$

from which Eq.(76) follows for the + sign. The downward excursions can be treated in the same way.//

The last result brings to mind the fact that the particle is diffusing in the force field $-2^{-1}(b^2 + b')'$ and it is the corresponding total quantum mechanical energy in the state $\rho_0^{1/2}$ which determines the number of infinitesimally short excursions.

Setting

$N_\varepsilon^\pm(t)$ = number of $\begin{matrix}\text{upward} \\ \text{downward}\end{matrix}$ excursions from a of duration $\geq \varepsilon$, upto time t,

we can prove that:

COROLLARY

With probability one, for the above diffusion X, with invariant density ρ_0,

$$\lim_{\varepsilon \downarrow 0} \frac{N_\varepsilon(t)}{(\rho_0^{1/2}, H_\pm(a)\, e^{-\varepsilon H_\pm(a)} \rho_0^{1/2})} = \lim_{\varepsilon \downarrow 0} \frac{N_\varepsilon(t)}{\sum_1^\infty \alpha_n^\pm e^{-\varepsilon \alpha_n^\pm} |(\rho_0^{1/2}, f_n^\pm)|^2} = \frac{L^a(t)}{\rho_0(a)} ,$$

$L^a(t)$ being the local time at a upto time t.

Outline Proof

We only need to prove that $\gamma^a (L^a(t)) = t$ for any $t > 0$ such that $X(t) = a$. This follows because from Equations (1) and (2) the closed set $\{ s \geq 0 : X(s) = a \}$ is perfect i.e. does not contain any points isolated from above. //

COROLLARY

If any of the limits below exists, given above hypotheses on b, almost surely, for X with $X(0) = a$,

$$\lim_{\varepsilon \downarrow 0} \frac{N_\varepsilon^+(t)}{N_\varepsilon^-(t)} = \lim_{\varepsilon \downarrow 0} \frac{\#_\varepsilon^+(t)}{\#_\varepsilon^-(t)} = \lim_{\varepsilon \downarrow 0} \frac{(\rho_0^{1/2}, H_+(a)\, e^{-\varepsilon H_+(a)} \rho_0^{1/2})_{L^2}}{(\rho_0^{1/2}, H_-(a)\, e^{-\varepsilon H_-(a)} \rho_0^{1/2})_{L^2}} , \quad t > 0 ,$$

ρ_0 being the invariant density.

It would be interesting to find an example in which this limit is different from unity.

We state without proof:

PROPOSITION

For the above X, with a C^2 Lipschitz drift b, invariant density $\rho_0(x) = \exp(2\int^x b(u)du) \in L^1(\mathbb{R})$, and with corresponding potential $V = 2^{-1}(b^2 + b')$ bounded below, having a bounded derivative, for each $a \in \mathbf{R}$,

$$\nu_a^\pm[\varepsilon, \infty) = (2\pi\varepsilon)^{-1/2} + 0(1) , \qquad \text{as } \varepsilon \downarrow 0 .$$

Thus, in this case,

$$\lim_{\varepsilon \downarrow 0} \frac{N_\varepsilon^+(t)}{N_\varepsilon^-(t)} = \lim_{\varepsilon \downarrow 0} \frac{\#_\varepsilon^+(t)}{\#_\varepsilon^-(t)} = 1 \ .$$

ACKNOWLEDGMENT

One of us, A.T., would like to thank the conference organisers, especially Professor Piotr Garbaczewski, for their kind invitation to visit Karpacz again and for the opportunity to present the above results to a wider audience.

REFERENCES

(1) Itô, K. and McKean, H.P., *Diffusion Processes and Their Sample Paths*, Springer, Berlin, 1965.

(2) McKean, H.P., *Stochastic Integrals*, Academic Press, New York, 1969.

(3) Williams, D., *Diffusions, Markov Processes and Martingales*, Volume 1, Foundations, John Wiley, Chichester, 1979.

(4) Rogers, L.C.G., and Williams, D., *Diffusions, Markov Processes and Martingales*, Volume 2, Itô Calculus, John Wiley, Chichester, 1987.

(5) Rogers, L.C.G., 'A Guided Tour Through Excursions', Bull. L.M.S.#91, Volume 21, Part 4, July 1989.

(6) Hawkes, J. and Truman, A., 'Statistics of Local Time and Excursions for the O-U Process', Swansea preprint, 1990.

(7) Truman, A. and Williams, D., 'A Generalised Arc-Sine Law', in *Diffusion Processes and Related Problems in Analysis*, Volume 1, Diffusions in Analysis and Geometry, editor Mark A. Pinsky, Birkhaüser, Boston, 1990.

(8) Truman, A. and Williams, D., 'Excursions and Itô Calculus in Nelson's Stochastic Mechanics', in *Recent Developments in Quantum Mechanics*, editors A. Boutet de Monvel et al., Kluwer Academic, Netherlands, 1991.

(9) Nelson, E., *Dynamical Theories of Brownian Motion*, Princeton University Press, Princeton, 1967.

(10) Batchelor, A. and Truman, A., 'Hitting, Killing and Capturing in Nelson's Stochastic Mechanics', in *Stochastic Processes, Physics and Geometry*, editors S. Albeverio, G. Casati et al., World Scientific, Singapore, 1990.

(11) Gihman, I.I. and Skorohod, A.V., *Stochastic Differential Equations*, Springer, Berlin, 1970.

(12) Simon, B., *Functional Integration and Quantum Physics*, Academic Press, New York, 1979.

(13) Karatzas, I. and Shreve, S.E., *Brownian Motion and Stochastic Calculus*, Springer, New York, 1988.

(14) Reed, M. and Simon, B., *Methods of Modern Mathematical Physics*, Volume II Fourier Analysis, Self-Adjointness, Academic Press, New York, 1975.

(15) Hille, E., *Lectures on Ordinary Differential Equations*, Addison Wesley, Reading (Mass), 1969.

(16) Abramowitz, M. and Stegun, I.A., *Handbook of Mathematical Functions*, Dover, New York, 1970.

(17) Davies, E.B., *Heat Kernels and Spectral Theory*, Cambridge University Press, Cambridge, 1989.

INVARIANT MEASURES FOR CLASSICAL AND QUANTUM FIELDS

Z. Haba

Institute of Theoretical Physics, University of Wroclaw, Wroclaw,Poland

A relation of classical mechanics to statistical mechanics and to the quantum field theory is a topic of continuous interest. In fact, a transition from one to another and the relation between these theories needs still thorough investigation. The passage from classical mechanics to classical equilibrium statistical mechanics goes through the notion of an invariant measure. The probabilistic interpretation of quantum mechanics is quite different from the probabilistic description of statistical mechanics. We must not expect that probabilistic aspects of quantum mechanics(e.g. the interference phenomena) can be derived from statistical mechanics. Nevertheless, some properties of quantum systems may have a description in terms of classical probability. By that we mean only that a probability measure is associated with the quantum system. It seems that the notion of an invariant measure may be a common ground for classical mechanics, (non-equilibrium) statistical physics and quantum mechanics.

We discuss the notion of an invariant measure in sec.1. In sec.2 we study in detail the relevance of a random perturbation for a statistical description of non-linear waves.

I. Some remarks on invariant measures in classical and quantum systems

Let $\mathcal{M}$ be a Riemannian manifold and μ a probability measure defined on $\mathcal{M}$. The measure μ is invariant under a flow

$$\xi_t: \mathcal{M} \to \mathcal{M} \tag{1}$$

if for any measurable set Γ $\mu(\xi_t(\Gamma)) = \mu(\Gamma)$.

It follows that an invariant measure defines a unitary one-parameter group in $L^2(d\mu)$

$$(T_t f)(x) = \exp iS(t,x)\, f(\xi_t(x)) \tag{2}$$

where S is an arbitrary real function, which must fulfil a composition law if T is to be a group. T_t could be considered as a quantum-mechanical evolution operator if we did not impose additional requirements on the "quantization of classical systems". Note, that from the invariance of the measure μ it follows that the correlation functions

$$E[\xi_t \ldots \xi_s] \tag{3}$$

are homogenous in time, what corresponds to the existence of a ground state of a quantum system.

Well-known examples of invariant measures are supplied by classical Hamiltonian systems. In such

a case the mesure μ is defined on the classical phase space. The conventional quantum mechanics is formulated in terms of wave functions defined on the configuration space. We must somehow restrict the set of wave functions in order to be in agreement with the conventional quantum mechanics (and with the experiment). In the case of the harmonic oscillator this is achieved by the restriction of wave functions to analytic functions. If $z=x+ip \to z$ expit is the classical flow and $d\mu = dzd\bar{z}\, \exp(-\frac{1}{2}z\bar{z})$ is the invariant measure then eq.(3) in fact produces the correlation functions

$$\langle z_t \ldots . \bar{z}_s \rangle \tag{4}$$

of the quantum harmonic oscillator up to the time ordering. We must not expect to derive the general quantum correlations, because these are non-commutative and the non-commutativity cannot follow from classical probability, but must be imposed by hand. We could generalize the construction (3)-(4) to the general coherent states defined on cosets of Lie groups. This is perhaps known to specialists in geometric quantization. However, we did perturbative calculations for general Hamiltonian systems in quantum mechanics and field theory and found (quite surprisingly) that up to the ordering of the quantum variables eqs.(3) (defined by a classical invariant measure) and (4) agree in each order of the perturbation expansion.

There is another class of quantum systems described by a classical invariant measure . Consider a Hamiltonian system linear in momenta

$$H = p^k \beta_k(x) \tag{5}$$

Then, the equations of motion read

$$\frac{dx}{dt} = \beta(x)$$

The unitary group (1) can be considered as a quantization of the Hamiltonian (5).

We would like to discuss next the conventional quantum mechanics in a some-what unconventional description. Let

$$H = \frac{p^2}{2} + U(x) \tag{6}$$

be the Hamiltonian and ψ_o its ground state

Denote

$$\overline{H} = \psi_o^{-1} H\, \psi_o \tag{7}$$

Then, the semigroup $\exp -t\overline{H}$ has the representation [1]

$$(\exp -t\overline{H}\, f)(x) = E_x[f(\xi_t)] \tag{8}$$

where ξ_t is the solution of the stochastic equation

$$d\xi_t = -\nabla \ln \psi_o\, dt + \sqrt{\hbar} db_t \tag{9}$$

where b_t is the Brownian motion.

The Hamiltonian $\overline{H}$ is selfadjoint in the Hilbert space $L^2(d\mu)$, where μ is the invariant measure for the process ξ_t defined by

$$\int d\mu(x) E_x[f(\xi_t)] = \int d\mu(x)\, f(x)$$

The invariant measure is expressed by the ground state $d\mu(x) = dx\ \psi_o^2(x)$.

A relation to the classical mechanics is not evident from eq.(9). There are two reasons for that. First, eq.(9) is of first order in time. Second, the classical mechanics is in real time, whereas eq.(9) in the imaginary time. However, we can get a relation to the classical mechanics if we first make (formally) the time real rewriting eq.(9) in the form

$$dq_t = -i\nabla \ln\psi_o\ dt + \sqrt{i\hbar}\, db_t \tag{10}$$

Then, through differentiation using the Ito formula and the definition of ψ_o as the ground state of H we get that q_t fulfills the randomly perturbed Hamiltonian equations

$$dq_t = \frac{\partial H}{\partial p} dt + i\sqrt{i\hbar}\ b_t \nabla\nabla \ln\psi_o\ dt$$

$$dp_t = -\frac{\partial H}{\partial q} dt + \sqrt{i\hbar}\, \frac{d^2 b}{dt^2} dt \tag{11}$$

where the classical Hamiltonian is defined in eq.(6).

From eq.(10) there follows the Feynman formula

$$(\exp it\overline{H}\ f)(x) = E[f(x+q_t)]$$

where q_t is the solution of eq.(10) starting from 0 and the expectation value is with respect to the standard Wiener measure.

"The quantum path" q_t is complex. However, this was to be expected, because the expectation values of the position operator in quantum mechanics in the Heisenberg picture are complex.

II. Ergodic behaviour of randomly perturbed non- linear waves

In this secton we are discussing a different topic (see our paper [2] for more details) : the statistical mechanics of non-linear waves. We discuss mathematical models of classical field theory, which are not ergodic, but a proper Markovian random perturbation of order ϵ makes them ergodic. For a definite random perturbation the solution of field equations is a stationary Markov process with the unique invariant measure μ_ϵ. With our choice of perturbation the measure μ_ϵ is also an invariant measure for the deterministic flow. The random perturbation selects one of classically allowed invariant measures. The memory of the random perturbation is preserved in the sense that the invariant measure depends on it also in the limit $\epsilon \to 0$. An invariant measure describes a large time asymptotics. For a finite time the limit $\epsilon \to 0$ of the Markov process does exist and coincides with the deterministic flow. Nevertheless, the limits $t\to\infty$ and $\epsilon \to 0$ are not exchangeable.

We outline a general formulation of the problem in the framework of infinite dimensional

Hamiltonian systems [3]. So, we consider an infinite dimensional symplectic Riemannian manifold $\mathcal{M}$ modelled on a Hilbert space H. We denote the symmetric bilinear form on $T\mathcal{M}$ by g , the symplectic form by ω and the exterior differentiation by D. Then, we assume

$$D\omega = 0 \tag{12}$$

g and ω define the corresponding operators $\bar{g}$ and $\bar{\omega}$ mapping $T\mathcal{M}^*$ into $T\mathcal{M}$. If $\mathcal{H}\in C^1(\mathcal{M})$ denotes the Hamiltonian then

$$v_{\mathcal{H}} = \bar{\omega}\, D\mathcal{H} \tag{13}$$

is the Hamiltonian vector field. $v_{\mathcal{H}}$ generates a flow ξ_t on $\mathcal{M}$

$$\frac{d\xi_t}{dt} = v_{\mathcal{H}}(\xi_t) \tag{14}$$

We introduce a dissipation v_D in order to have a well-defined long time asymptotics. We assume a dissipation , which shrinks the motion in the direction orthogonal to the Hamiltonian flow. If we take

$$v_D = \bar{g}\, D\mathcal{F} \tag{15}$$

for a certain $\mathcal{F}\in C^1(\mathcal{M})$ then

$$g(v_{\mathcal{H}}, v_D) = \{\mathcal{H},\mathcal{F}\} \tag{16}$$

$v_{\mathcal{H}}$ and v_D are orthogonal if their Poisson bracket is zero (i.e. $\mathcal{F}$ is a constant of motion).

We say that a measure μ on $\mathcal{M}$ is invariant under the flow $\xi_t(u)$ starting in $u\in\mathcal{M}$ at $t=0$ if for any bounded $f\in C(\mathcal{M})$

$$\int d\mu(u)\ (S_t f)(u) \equiv \int d\mu(u)\ f(\xi_t(u)) = \int d\mu(u)\ f(u) \tag{17}$$

or in a differential form

$$\int d\mu(u)\ \mathcal{A} f = 0 \tag{18}$$

where $\mathcal{A} = \frac{dS_t}{dt}\ S_t^{-1}$ is the generator of S_t.

A Markovian stochastic process ζ_t on $\mathcal{M}$ is constructed out of the basic one W_t-the Wiener process on H. It is defined as the Gaussian process with the covariance

$$E[(W_t,f)\ (W_s,h)] = \min(t,s)\ (f,h) \tag{19}$$

where $(\ ,\ )$ is the scalar product on H. We denote by $\|\ \|$ a measurable norm on H. The completion of H in $\|\ \|$ is a Banach space B. The triple $(\|\ \|,H,B)$ is called the abstract Wiener space . The Wiener process takes its values in B. The measurable norm $\|f\|^2$ is of the form (Tf,Tf), where T is a Hilbert-Schmidt operator. The stochastic perturbation of the Hamiltonian system (14) is defined by the stochastic differential equation

$$d\xi = \beta\, dt + \epsilon\, d\zeta \tag{20}$$

where ζ is constructed of W and

$$\beta = \bar{\omega} D\mathcal{H} - \tfrac{\kappa}{2}\bar{g} D\mathcal{F} \tag{21}$$

A stochastic process ξ_t defines a semigroup S_t on the domain $\mathcal{D}(S_t)\subset C(\mathcal{M})$

$$(S_t f)(u) = E[f(\xi_t(u))|\ t<\tau(\xi)] \equiv E_u[f(\xi_t)] \tag{22}$$

where $\xi_t(u)$ is the solution of the stochastic equation (20) starting in u at t=0 and stopped at the

explosion time $\tau(\xi)$ (if finite).

We say that μ is an invariant measure for ξ_t if for any bounded $f\in C(\mathcal{M}_b)$

$$\int d\mu(u)\, E_u[f(\xi_t)] = \int d\mu(u)\, f(u) \tag{23}$$

or in a differential form (infinitesimal invariance)

$$\int d\mu(u)\, \mathcal{A}f = 0 \tag{24}$$

where $\mathcal{A}$ is the generator of the semigroup S_t.

In order to show that the system (20) has a solution for any time we need the methods of Lyapunov. Let H_σ be a Sobolev space

Definition: V is a Lyapunov function on H_σ if

1) $V(\phi)$ is defined for $\phi\in H_\sigma$, continuous on H_σ and bounded by $\|\phi\|_\sigma^n$ for a certain $n>0$

2) $V(\phi)\geq\lambda(\|\phi\|_\sigma)$ for an increasing non-negative function λ.

Then, generalizing the method of ref.[4] we can show

Theorem 1

Let $\mathcal{A}$ be the generator of the process ξ_t. Assume that there exists a Lyapunov function $V\in\mathfrak{D}(\mathcal{A})$ defined on H_σ such that

$$\mathcal{A}V \leq -cV + r \tag{25}$$

for a certain $c>0$.

i) the solution ξ_t of the stochastic equation (20) can be continued to arbitrary time t

ii) there exists the unique invariant measure μ on $H_{-\sigma}$ determined by the transition function $P_t(\xi,\Gamma)$ ($\Gamma\subset H_{-\sigma}$) of the Markov process ξ_t

$$\lim_{t\to\infty} P_t(\xi,\Gamma) = \mu(\Gamma) \tag{26}$$

iii) the process ξ_t is uniquely ergodic

We apply this theorem to the nonlinear wave equation in one dimension. We can write the wave equation in one dimension in the form

$$\partial_t\partial_x\phi^j = \partial_j h(\phi) \tag{27}$$

where h is a local function of the vector $\phi = (\phi^1,\ldots,\phi^n)$.

In order to formulate eq.(27) in a rigorous mathematical setting we restrict first x to an interval [-r,r] , impose Dirichlet or periodic boundary conditions and then eventually let $r\to\infty$.

Eq.(27) can be expressed in the canonical form

$$\partial_t\phi^j = \{\phi^j,\mathcal{H}\} \tag{28}$$

where $\mathcal{H}=\int h\big(\phi(x)\big)dx$ and the Poisson bracket is

$$\{f,g\} = \tfrac{1}{2}\int dxdy\, \frac{\delta f}{\delta\phi^j(x)}\, \epsilon(x-y)\, \frac{\delta g}{\delta\phi^j(y)} \tag{29}$$

here $\epsilon(.)$ is an odd function such that $\epsilon(x)=1$ for $x>0$.

We shall treat eq.(27) and its stochastic perturbation in a particular realization of an abstract Wiener space. We introduce the Hilbert space $L^2(-r,r)$ of square integrable functions on the interval [-r,r]. We consider in L^2 the operator

$$T_{-\alpha}(x,y) \equiv G^\alpha(x,y) = \sum_{n\neq 0} \left(\tfrac{\pi}{r}|n|\right)^{-2\alpha} \exp i\tfrac{\pi}{r}n(x-y) \tag{30}$$

where $(\frac{\pi}{r}|n|)^2$ are the eigenvalues of $-\partial^2$ in $L^2(-r,r)$ (Dirichlet or periodic boundary conditions) . We complete $\mathfrak{D}(T_\alpha)$ in the norm $\|f\|_\alpha=|T_\alpha f|$. Then, we get the Hilbert space H_α. Its dual is $H_{-\alpha}$. The Brownian motion on $L^2(-r,r)$ is defined in eq.(19). The norm $\|.\|_{-\alpha}$ is measurable if $\alpha>\frac{1}{4}$. Then, for $B=H_{-\alpha}$ we have a realization of the abstract Wiener space $(L^2,B,\|.\|_{-\alpha})$.

The stochastic perturbation (20) of the wave equation (27) reads

$$d\phi = \{\phi,\mathcal{H}\}dt - \tfrac{\kappa}{2}P^2D\mathcal{F}dt + \epsilon PdW \tag{31}$$

where P^2 is the metric $\bar{g}$ in eq.(15).

We shall discuss $\mathcal{F}$ in eq.(31) of the form

$$\mathcal{F}_{a,b} = \tfrac{a}{2}\int(\partial\phi)^2 + b\mathcal{H} \tag{32}$$

The dissipation breaks the Lorentz invariance. If we let $\kappa=0$, then the stochastic equation

$$d\phi = \partial^{-1}D\mathcal{H}dt + \epsilon\partial^{-1}dW$$

is relativistic invariant and has a solution $\phi_t\in H_\sigma$ ($\sigma<\frac{1}{4}$). Relativistic solutions exist also in two-dimensional space, but their construction is much more difficult [5].

We show that eq.(31) has a non-exploding solution for a broad class of $\mathcal{H}$ and $\mathcal{F}$. For this purpose we apply the method of a Lyapunov function . Positively definite integrals of motion are usually good candidates for the Lyapunov function. By means of elementary calculations we can show that the generator K of space translations (i.e. $\{K,\phi\}=\partial\phi$)

$$K = \tfrac{1}{2}\int \partial\phi\partial\phi \tag{33}$$

and the potential energy $\mathcal{H}(\phi)$ are constants of motion of the flow (28) (Dirichlet or periodic boundary conditions). Then, we have

Theorem 2

Let μ_o be the Gaussian measure determined by the characteristic function

$$\int d\mu_o(\phi)\, \exp i(\phi,f) = \exp[-\tfrac{\epsilon^2}{2\kappa}(T_{-1}f,f)] \tag{34}$$

Assume that $\mathcal{H}$ is a sum of a polynomial and a trigonometric function then for each ϵ and κ μ_o is an (infinitesimally) invariant measure for the deterministic flow $\frac{d\phi}{dt}=\{\phi,\mathcal{H}\}$.

Using the invariance of μ_o under the deterministic flow we can show

Theorem 3

Let $\mathcal{F}=K$ in eq.(31) then ϕ_t is a stationary stochastic process with the unique invariant measure μ_o defined by eq.(34).

For more general dissipations

Theorem 4

Assume the Dirichlet boundary conditions for ∂^2. If $\mathcal{F}=\mathcal{F}_{1,b}$ in eq.(31) and $\mathcal{H}$ in $\mathcal{F}_{1,b}$ (32) is a sum of a polynomial of order $n\geq 2$ bounded from below and a trigonometric function, then the solution of the stochastic equation(31) is an ergodic stochastic process with the invariant probability measure

$$d\mu_r(\phi) = Z_r^{-1}\, d\mu_o(\phi)\, \exp[-\frac{b\kappa}{\epsilon^2}\int_{-r}^{r} h(\phi(x))\, dx] \tag{35}$$

where Z_r is the normalization factor and μ_o is defined in eq.(34). The weak limit μ_∞ exists and is orthogonal to μ_r.

Remark: There is a difficulty with the periodic boundary conditions. We would not be able to integrate over the zero mode $\int\phi(x)dx$ in $\mathcal{H}$. For this reason we restrict ourselves to Dirichlet boundary conditions in Theorem 4.

Assume that we have a stochastic process ϕ_t solving the stochastic equation

$$d\phi_t = \beta(\phi_t;\, \kappa(\epsilon))\, dt + \epsilon PdW \tag{36}$$

with the initial condition $\phi_o=\psi$.

Next, assume that we have an invariant measure μ resulting from Theorem1. Then, from Theorem 1 it follows that for any μ-integrable function F

$$\lim_{t\to\infty} E_\psi[\, F(\phi_t)] = \lim_{t\to\infty} \int P_t(\psi,d\phi)\, F(\phi) = \int d\mu(\phi)\, F(\phi) \tag{37}$$

Moreover, for such a function $F\in L^1(d\mu)$ it follows from the ergodic theorem that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t F(\phi_\tau)\, d\tau = \int d\mu(\phi)\, F(\phi) \tag{38}$$

with probability 1.

The invariant measures for the stochastic process ϕ_t are also invariant with respect to the deterministic Hamiltonian flow. However, for the deterministic system we cannot deduce any limiting behavior of the type (37)-(38) because there are many invariant measures i.e. the deterministic system is not ergodic. The stochastic perturbation selects one of many possible invariant measures of the Hamiltonian flow.

Now, we are going to switch off the perturbation i.e. $\epsilon\to 0$ and $\kappa(\epsilon)\to 0$. We wish to discuss what

happens with the stochastic process ϕ_t in this limit. In order to estimate the behavior for $\epsilon\to 0$ let us consider a quadratic form Q. Let χ_t be the deterministic Hamiltonian flow ($\kappa(\epsilon)= 0$) starting from $\chi_o=\psi$

$$\frac{d}{dt}\chi_t = \beta(\chi_t;\, 0\,) \tag{39}$$

Then , by means of the Ito calculus

$$dQ(\phi_t\text{-}\chi_t,\phi_t\text{-}\chi_t) = 2\, Q(d\phi_t\text{-}d\chi_t,\phi_t\text{-}\chi_t) + \epsilon^2 Q(PdW,PdW) \tag{40}$$

Hence, from eq.(36)

$$dQ(\phi_t\text{-}\chi_t,\phi_t\text{-}\chi_t) = 2\, Q\Big(\beta(\phi_t;\kappa) - \beta(\chi_t;0),\phi_t\text{-}\chi_t\Big)dt + 2\epsilon Q(\phi_t\text{-}\chi_t,PdW) + \frac{\epsilon^2}{2}TrPQP\, dt \tag{41}$$

So,

$$E_\psi[\phi_t\text{-}\chi_t,\phi_t\text{-}\chi_t] = 2\int_0^t ds\, E_\psi[Q\Big(\beta(\phi_s;\kappa) - \beta(\chi_s;0),\, \phi_s\text{-}\chi_s\Big)] + \frac{\epsilon^2}{2}\, t\, TrPQP \tag{42}$$

$$\beta(\phi_s;\kappa(\epsilon)\,) - \beta(\chi_s;0) = (D\beta,\phi_s\text{-}\chi_s) + o(\phi_s\text{-}\chi_s) + o(\kappa(\epsilon)\,) \tag{43}$$

for a differentiable β. We would like to conclude from eqs.(41)-(42) that

$$E_\psi[Q(\phi_t\text{-}\chi_t,\phi_t\text{-}\chi_t)] \leq \alpha(\epsilon,t) \tag{44}$$

where $\alpha(\epsilon,t)\to 0$ as $\epsilon\to 0$ i.e. the stochastic process ϕ_t tends to the deterministic flow χ_t in the square mean. We would be also interested whether we can let $t\to\infty$ in $\alpha(\epsilon,t)$. There is some difficulty in establishing the bound (44) in general. The problem is that the right hand side of eq.(42) might be large even if ϵ and $\phi_s\text{-}\chi_s$ were small , because of big fluctuations of ϕ. However, such fluctuations can be excluded (see ref.[6] for general results on the limit $\epsilon\to 0$). We restrict ourselves here to a proof of a mathematically simple case

Theorem 5

Assume $h(\phi) = \lambda\cos\alpha\phi + \frac{m^2}{2}\phi^2$ and $Q(\phi,\phi) = \|\phi\|_{\frac{1}{2}}^2$. Let ϕ_t be the stochastic process constructed in Theorem 4 then

$$E[\|\, \phi_t\text{-}\, \chi_t\|_{\frac{1}{2}}^2] \leq c_1 t\, (\, \epsilon^2 + \kappa(\epsilon)\,)\, \exp(c_2 t) \tag{45}$$

We cannot take the limit $t\to\infty$ in eq.(45). In fact, we shall show that the limits $\epsilon\to 0$ and $t\to\infty$ cannot be interchanged. When $\epsilon\to 0$ then from Theorem 5 we get

$$\lim_{\epsilon\to 0} E_\psi[\, F(\phi_t)] = F(\chi_t) \tag{46}$$

The limit $t\to\infty$ of $F(\chi_t)$ in eq.(46) in general does not exist. After taking the limit $t\to\infty$ the limit $\epsilon\to 0$ depends on the behavior of $\kappa(\epsilon)$.

Let $\mu_r^{\epsilon,\kappa}$ be the invariant measure of Theorem 4 we consider the limit

$$\lim_{\epsilon\to 0}\lim_{t\to\infty} E_\psi[F(\phi_t)] = \lim_{\epsilon\to 0}\int d\mu_r^{\epsilon,\kappa}(\phi)\, F(\phi) \tag{47}$$

First, if $\kappa(\epsilon)\epsilon^{-2}\to 0$ then in a formal sense $\mu_r^{\epsilon,\kappa}$ tends to the Lebesgue measure, which does not exist in infinite dimensions. In such a case the limit (47) depends sensitively on the form of F.

Theorem 6

Under the assumptions of Theorem 5 if $\kappa(\epsilon)\,\epsilon^{-2}\to 0$ then

$$\lim_{\epsilon\to 0}\int d\mu_r^{\epsilon,\kappa}(\phi)\ F(\phi) = \lim_{\epsilon\to 0}\int d\mu_m(\phi)\ F(\epsilon\kappa(\epsilon)^{-\frac{1}{2}}\phi) \tag{48}$$

(if either of the limits exists), here μ_m is the Gaussian measure with the covariance $(-\partial^2 + m^2)^{-1}$ (Dirichlet boundary conditions).

Remark: If F is a polynomial then the limit does not exist, if F is a trigonometric function then the limit is equal to zero.

In the remaining cases we have

Theorem 7

i) $\mu_r^{\epsilon,\kappa}$ is an invariant measure of the deterministic flow χ_t (39). The limit $\kappa(\epsilon)\epsilon^{-2}\to\theta^{-1}$=const exists. The invariant measure is determined by the dissipation $\kappa(\epsilon)$ D($K + b\mathcal{H}$)

ii) if $\kappa(\epsilon)\epsilon^{-2}\to\infty$ and $\mathcal{F}_{1,b}=K+b\mathcal{H}$ has the unique minimum $\phi_c$then

$$\lim_{\epsilon\to 0}\ \lim_{t\to\infty} E_\psi[\,F(\phi_t)] = F(\phi_c) \tag{49}$$

Remark: We may let $r=\infty$ in Theorem 7 (infinite volume limit)

In the case ii) the limit $\epsilon\to 0$ is a δ-measure concentrated on the solution of the equation

$$\partial_x{}^2\phi_c(x) = b\ h'(\phi_c(x)) \tag{50}$$

In general, ϕ_c is in no relation to the flow $\chi_t(x)$ (39) entering the limit $\epsilon\to 0$ in eq.(46) at time t, which solves the equation

$$\partial_t\partial_x\chi_t(x) = h'(\chi_t(x)) \tag{51}$$

with the initial condition $\chi_o=\psi$. However, if accidentally the initial condition $\psi=\phi_c$ then

$$\chi_t(x) = \phi_c(x+bt) \tag{52}$$

solves eq.(51).

It follows from the conservation of momentum $K=\frac{1}{2}\int\partial_x\phi\partial_x\phi$ and the potential energy $\mathcal{H}$ that the limits $t\to\infty$ and $\epsilon\to 0$ can be interchanged only if $\chi_t(x)=\psi(x)=\phi_c(x)$=const and $h'(\phi_c)=0$. In this sense the long time behavior of deterministic systems and systems with an arbitrarily small noise is different.

References

[1] S. Albeverio, R. Hoegh-Krohn and L. Streit, J.Math.Phys.18(1977)907

[2] Z. Haba, Ergodicity and invariant measures for some randomly perturbed classical fields, Wroclaw and BiBoS preprint,March 1991

[3] R. Chernoff and J.E. Marsden, Properties of Infinite Dimensional Hamiltonian Systems,Lect.Notes in Math., No.425,Springer,1974

[4] A. Lasota and M.C. Mackey, Probabilistic Aspects of Deterministic Systems, Cambridge,1985

[5] S.Albeverio,Z.Haba and F. Russo, in preparation

[6] A.D. Wentzel and M.I. Freidlin, Usp.Mat.Nauk 25,3(1970)(in Russian)

PART IV

FERMIONS: CLASSICAL, RANDOM, QUANTUM

REPRODUCING KERNEL SPACES AND RANDOM FIELDS FOR FERMIONS

J. Kupsch

Fachbereich Physik
Universität Kaiserslautern
D–6750 Kaiserslautern
Federal Republic of Germany

ABSTRACT

By a natural definition of coherent states the fermionic Fock space is represented as Hilbert space with reproducing kernel. A Bargmann–Fock type representation with a gaussian measure is possible. Moreover normal ordering leads to the Wiener isomorphism with an $\mathcal{L}^2$–space over a measure space which carries the physical degrees of freedom. The semigroup of a free evolution on the fermionic Fock space can be related to an Ornstein–Uhlenbeck process on the underlying measure space.

1. INTRODUCTION

The question for a probabilistic interpretation or a path–space interpretation of fermions is related to the question for coherent states. The basic starting point for these investigations was Klauder's article [1] from 1960. There have been many investigations in the meantime but there is nevertheless no final satifactory answer. In this lecture I shall present results which are related to the ideas of Klauder and which incorporate recent developments of a non–commutative stochastic calculus [2] [3].

The lecture is organized as follows. After a presentation of the basic notations in Sect. 2. a natural definition of coherent states is given in Sect. 3. which allows to represent both Fock spaces, the bosonic and the fermionic one, as Hilbert spaces with reproducing kernel. It is possible to map both Fock spaces into the same reproducing kernel space such that the normal ordering prescriptions for symmetric and for antisymmetric tensors coincide. As a consequence of this fact

the Wiener isomorphism between the bosonic Fock space with an $\mathcal{L}^2$–space can be transferred in Sect. 4. to an isomorphism between the fermionic Fock space and this $\mathcal{L}^2$ space (or a subspace of it). The fermions can therefore be represented by random fields in the basic measure space.

A positive one particle operator which may include an interaction with an external bosonic field induces semigroups on the bosonic and on the fermionic Fock space. In Sect. 5. it is shown that both semigroups correspond to Ornstein–Uhlenbeck processes. The (euclidean) evolution of fermions has therefore the same trajectory picture as that of bosons.

2. HILBERT SPACES AND TENSOR ALGEBRAS

In this Section the basic notations and normalizations used in this lecture are introduced. They follow essentially [4] and [5]. Let $\mathcal{Y}$ be a complex Hilbert space (of finite or countable infinite dimension) with inner product $(f|g)$ and with an antiunitary involution $f \to f^*$, $f^{**} = f$. This Hilbert space can be decomposed into two isomorphic subspaces, $\mathcal{Y} = \mathcal{V} \oplus \mathcal{V}^*$, which are mapped into each other by the involution, $\mathcal{V} \xrightarrow{*} \mathcal{V}^*$. Then the following two bilinear forms, $\omega_\epsilon(f,g)$: $\mathcal{Y} \times \mathcal{Y} \to \mathbb{C}$, $\epsilon = \pm 1$, are canonically related to this decomposition

$$\omega_\epsilon(f,g) = \begin{cases} (f^*|g) & \text{if } f \in \mathcal{Y},\ g \in \mathcal{V} \\ \epsilon(f^*|g) & \text{if } f \in \mathcal{Y},\ g \in \mathcal{V}^*. \end{cases} \tag{1}$$

The form $\omega_+(f,g)$ is symmetric and $\omega_-(f,g)$ is antisymmetric, $\omega_\epsilon(f,g) = \epsilon\, \omega_\epsilon(g,f)$, $f,g \in \mathcal{Y}$. The subspaces $\mathcal{V}$ and $\mathcal{V}^*$ are isotropic spaces of these forms.

As basic example we may take $\mathcal{Y} = \mathcal{L}^2(\mathbb{R}) \otimes \mathbb{C}^2$ with elements $f(s) = \begin{bmatrix} \varphi(s) \\ \psi(s) \end{bmatrix}$, where $\varphi(s)$ and $\psi(s)$ belong to $\mathcal{L}^2(\mathbb{R})$. The inner product is given by $(f|g) = \int f^+(s) g(s) ds$, the involution is $f^*(s) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \overline{f(s)}$, and the bilinear forms are $\omega_\epsilon(f,g) = \int f^T(s) \begin{bmatrix} 0 & \epsilon \\ 1 & 0 \end{bmatrix} g(s)\, ds$, $\epsilon = \pm 1$. The isotropic decomposition is $\mathcal{Y} = \mathcal{V} \oplus \mathcal{V}^*$ with $\mathcal{V} = \left\{ f(s) = \begin{bmatrix} \varphi(s) \\ 0 \end{bmatrix} \right\}$ and $\mathcal{V}^* = \left\{ f(s) = \begin{bmatrix} 0 \\ \psi(s) \end{bmatrix} \right\}$. The diagonal elements $f = f^*$ form a real Hilbert space

$$\mathcal{X}^{\mathbb{R}} = \{f | f = f^*\} = \left\{ f = \begin{bmatrix} \varphi \\ \overline{\varphi} \end{bmatrix}, \varphi \in \mathcal{L}^2(\mathbb{R}) \right\}.$$

This space is $\mathbb{R}$–isomorphic to the complex space $\mathcal{L}^2(\mathbb{R})$. Since $\mathcal{X}$ is the complexification of $\mathcal{X}^{\mathbb{R}}$, it may also be considered as a complexification of $\mathcal{L}^2(\mathbb{R})$.

The symmetric tensor product is denoted by $f \vee g$ and the antisymmetric tensor product (exterior product) is denoted by $f \wedge g$; for both tensor products we shall also write $f \circ g$. The space $\mathcal{A}_n^+(\mathcal{X})$ is the algebraic sum of all decomposable tensors $f_1 \vee \ldots \vee f_n$ and $\mathcal{A}_n^-(\mathcal{X})$ is the algebraic sum of all tensors $f_1 \wedge \ldots \wedge f_n$. The algebraic sum $\mathcal{A}^{\pm}(\mathcal{X}) = \bigoplus_n \mathcal{A}_n^{\pm}(\mathcal{X})$ is the algebra of symmetric/antisymmetric tensors. The tensor products are normalized to $\|f_1 \vee \ldots \vee f_n\|^2 = \text{per}\,((f_i|f_j))$ or $\|f_1 \wedge \ldots \wedge f_n\|^2 = \det((f_i|f_j))$.
The closure of the algebra $\mathcal{A}^{\pm}(\mathcal{X})$ with respect to this norm yields the Fock space $\mathcal{T}^{\pm}(\mathcal{X}) = \bigoplus_{n=0}^{\infty} \mathcal{T}_n^{\pm}(\mathcal{X})$ of all symmetric/antisymmetric tensors.

If A is a linear operator on $\mathcal{X}$ then $\Gamma_{\pm}(A)$ is an operator on the Fock space $\mathcal{T}^{\pm}(\mathcal{X})$ defined by $\Gamma_{\pm}(A)\, f_1 \circ \ldots \circ f_n = (Af_1) \circ (Af_2) \circ \ldots \circ (Af_n)$. The involution $f \to f^*$ on $\mathcal{X}$ is extended as usual to an involution on the Fock spaces by $(f_1 \circ \ldots \circ f_n)^* = f_n^* \circ f_{n-1}^* \circ \ldots \circ f_1^*$.

The (anti)symmetric bilinear form $\omega_{\pm}(f,g)$ can be represented as

$$\omega_{\pm}(f,g) = (\Omega_{\pm} | f \circ g) \tag{2}$$

with a tensor $\Omega_{\pm} \in \mathcal{T}_2^{\pm}(\mathcal{X}_{-1})$ where $\mathcal{X}_1 \subset \mathcal{X} \subset \mathcal{X}_{-1}$ is a triplet of Hilbert slaces with Hilbert– Schmidt embeddings. Eq. (2) is first derived for $f,g \in \mathcal{X}_1$; but since $\omega_{\pm}$ is continuous in the norm of $\mathcal{X}$, the right hand side is also meaningful for $f,g \in \mathcal{X}$.

The real subspace $\mathcal{X}^{\mathbb{R}} = \{f | f^* = f \in \mathcal{X}\}$ can be supplemented with the canonical (pro)measure ν:

$$\int e^{i(f|\xi)} d\nu(\xi) = e^{-\frac{1}{2}(f|f)} \quad \text{if } f \in \mathcal{X}^{\mathbb{R}} \tag{3}$$

which is the basis for a stochastic interpretation. In the infinite dimensional case ν is countable additive only on an extended space $\mathcal{X}_{-1}^{\mathbb{R}}$.

The complex space $\mathcal{X}$ – or more precisely its underlying real space – can also be provided with the corresponding (pro)measure μ which has the Fourier–Laplace transform

$$\int e^{(f|\xi) + \overline{(g|\xi)}} d\mu(\xi) = e^{(f|g)} \tag{4}$$

if $f,g \in \mathcal{X}$. If $\dim \mathcal{X} = \infty$ this measure is countable additive only on an extended space $\mathcal{X}_{-1}$.

3. FUNCTION SPACES

3.1 Reproducing Kernel Spaces

Following the properties of coherent states in [6] generalized coherent states are defined on both Fock spaces as a mapping from a subset $\mathcal{D} \subset \mathcal{X}$ which is a linear space over $\mathbb{R}$ into the Fock space $f \in \mathcal{D} \subset \mathcal{X} \longrightarrow E(f) \in \mathcal{T}^{\pm}(\mathcal{X})$, such that the linear span $\mathcal{L}\{E(f) | f \in \mathcal{D}\}$ is dense in $\mathcal{T}^{\pm}(\mathcal{X})$.

The mapping

$$F \in \mathcal{T}^{\pm}(\mathcal{X}) \longrightarrow \varphi_F(\xi) = (E(\xi)|F), \quad \xi \in \mathcal{D}, \tag{5}$$

then yields a space of functions on $\mathcal{D}$ denoted by $\mathcal{F}(\mathcal{D})$. The inner product on $\mathcal{F}(\mathcal{D})$ is defined as

$$(\varphi_F(\xi)\|\varphi_G(\xi)) := (F|G). \tag{6}$$

The coherent states E(f) are represented in $\mathcal{F}(\mathcal{D})$ by the functions

$$\mathcal{E}_f(\xi) = (E(\xi)|E(f)). \tag{7}$$

The space $\mathcal{F}(\mathcal{D})$ is a Hilbert space with the reproducing kernel

$$K(\xi,\eta) := (E(\xi)|E(\eta)) = \mathcal{E}_\eta(\xi). \tag{8}$$

For the proof one has to verify the following properties [7]:

i) $K(\xi,\eta) \in \mathcal{F}(\mathcal{D})$ for any $\eta \in \mathcal{D}$,

ii) $(K(\xi,\eta)\|\varphi(\xi)) = \varphi(\eta)$.

The first statement of obvious, the second follows for coherent states from (5)–(8) as $(K(\xi,\eta)\|\mathcal{E}_f(\xi)) = (E(\eta)|E(f)) = \mathcal{E}_f(\eta)$ and the fact that $\mathcal{L}\{\mathcal{E}_f(\xi)|f \in \mathcal{D}\}$ is dense in $\mathcal{F}(\mathcal{D})$.

Examples for the symmetric tensor algebra can easily be constructed with the exponential series. For any $f \in \mathcal{X}$ the exponential series

$\exp f = 1 + f + \frac{1}{2!} f \vee f + \ldots$ converges within $\mathcal{T}^+(\mathcal{X})$ with norm $\|\exp f\|^2 = e^{\|f\|^2}$.

There is essentially only one choice for a coherent state in $\mathcal{T}^+(\mathcal{X})$ which reproduces the algebraic structure in the simple way $E(f) \vee E(g) = E(f+g)$. This identity implies $E(f) \sim \exp(Af)$ with a linear operator A. In the following we take $A = I$. But even with this simple choice of the operator the space $\mathcal{D}$ is not fixed. There are essentially two cases which are based either on a complex space $\mathcal{D}$ or on a real space $\mathcal{D}$.

a) $\mathcal{D} = \mathcal{X}_1 \subset \mathcal{X}$ dense linear subset.

With respect to applications it is advantageous to assume that $\mathcal{X}_1 = T\mathcal{X}$ with a linear invertible Hilbert–Schmidt operator T. But to define a reproducing kernel space any dense linear subset $\mathcal{D} = \mathcal{X}_1 \subset \mathcal{X}$ with a reasonable topology is admitted, including $\mathcal{D} = \mathcal{X}$. The corresponding function spaces $\mathcal{F}(\mathcal{D})$ are equivalent.

The coherent states are then defined as usual

$$E(\xi) = \exp \xi \tag{9}$$

and the isomorphism (5) leads to the antiholomorphic functions

$\varphi_F(\xi) = (\exp \xi|F)$, $\xi \in \mathcal{X}$. The kernel function (8) is $K(\xi,\eta) = e^{(\xi,\eta)}$ and the coherent states are represented by the exponential functions $\mathcal{E}_f(\xi) = e^{(\xi|f)}$.

b) $\mathcal{D} = \mathcal{X}_1^{\mathbb{R}} \subset \mathcal{X}^{\mathbb{R}}$

Here $\mathcal{X}_1^{\mathbb{R}} = \mathcal{X}_1 \cap \mathcal{X}^{\mathbb{R}}$ is a dense linear subset of the real space $\mathcal{X}^{\mathbb{R}} = \{f|f^* = f\}$. The linear span $\mathcal{L}\{E(\xi), \xi \in \mathcal{X}_1^{\mathbb{R}}\}$ is still dense in $\mathcal{T}^+(\mathcal{X})$. The coherent states are again given by (9) but with ξ restricted to $\xi \in \mathcal{X}_1^{\mathbb{R}}$. (An alternative choice is $E(\xi) = \exp i\xi$.) The restriction to the real space $\mathcal{X}^{\mathbb{R}}$ can also be obtained by a restriction to the isotropic subspace $\mathcal{V}$ since $\xi \in \mathcal{X}^{\mathbb{R}}$ is equivalent to $\xi = \eta + \eta^*$ with an arbitrary element $\eta \in \mathcal{V}$. In this way all functions can be defined as functions on the complex space $\mathcal{V}_1 \subset \mathcal{V}$.

In the case b) the space $\mathcal{D}$ carries the physical degrees of freedom, whereas in case a) the space $\mathcal{D} = \mathcal{X}$ incorporates twice the degrees of freedom. That is compensated by the constraint of antiholomorphy for the functions $\varphi(\xi)$. Despite of the different base spaces $\mathcal{D}$ the algebra of functions is the same in all cases and the symmetric tensor product is always represented by the numerical multiplication $\varphi_{F\vee G}(\xi) = \varphi_F(\xi) \cdot \varphi_G(\xi)$ if $F,G \in \mathcal{A}^+(\mathcal{X})$.

The functions of $\mathcal{F}(\mathcal{X}_1^{\mathbb{R}})$ can be obtained from those of $\mathcal{F}(\mathcal{X}_1)$ by the restriction to the real subspace $\mathcal{X}_1^{\mathbb{R}}$. The functions of $\mathcal{F}(\mathcal{X}_1)$ can be revealed from $\mathcal{F}(\mathcal{X}_1^{\mathbb{R}})$ by a complexification of the base space and analytic continuation.

3.2 Bargmann–Fock Representation

The function space $\mathcal{F}(\mathcal{H}_1)$ can be represented as $\mathcal{L}^2$–space. Let μ be the gaussian measure (4) on an extension $\mathcal{H}_{-1}$, then $\varphi_F(\xi) = (\exp \xi|F)$ is a square integrable function on $\mathcal{H}_{-1}$ for all $F \in \mathcal{T}^+(\mathcal{H})$ and $F \in \mathcal{T}^+(\mathcal{H}) \longrightarrow \varphi_F(\xi) \in \mathcal{L}^2(\mathcal{H}_{-1},\mu)$ is an isometric mapping with $\int \overline{\varphi_F(\xi)}\, \varphi_G(\xi)\, d\mu(\xi) = (F|G)$. This result follows for $F = \exp f$ and $G = \exp g$ from (4). Since the linear hull of the coherent states is dense in the Fock space the general proof is obtained by a continuity argument.

The antiholomorphic and continuous functions $\varphi(\xi) \in \mathcal{F}(\mathcal{H}_1)$ determine uniquely square integrable functions on $\mathcal{H}_{-1}$, and the space $\mathcal{F}(\mathcal{H}_1)$ can be identified with $\mathcal{L}^2(\mathcal{H}_{-1},\mu)$. The arguments ξ of the functions are random fields in a gaussian probability space which incorporates twice the physical degrees of freedom.

3.3 Generalized Coherent States For The Antisymmetric Fock Space

The known examples of coherent states in $\mathcal{T}^-(\mathcal{H})$ can be constructed with the help of an isomorphism between $\mathcal{T}^-(\mathcal{H})$ and a subspace of $\mathcal{T}^+(\mathcal{H})$. Such mappings have been used by Friedrichs [8], Klauder [1] and more extensively by Garbaczewski and Rzewuski [9] [10].

Let Θ be an injective isometric mapping $\mathcal{T}^-(\mathcal{H}) \to \mathcal{T}^+(\mathcal{H})$ and $\hat{\Theta}$ be the inverse surjective mapping $\mathcal{T}^+(\mathcal{H}) \to \mathcal{T}^-(\mathcal{H})$ such that

$$\begin{cases} \hat{\Theta}\Theta = \mathrm{id} & \text{on } \mathcal{T}^-(\mathcal{H}) \\ \Theta\hat{\Theta} = P & \text{projection operator on } \mathcal{T}^+(\mathcal{H}). \end{cases} \tag{10}$$

Moreover, we shall assume $\Theta\, \mathcal{T}^-_n(\mathcal{H}) = \mathcal{T}^+_n(\mathcal{H})$ for $n = 0,1,2\ldots$. The essential statement for the construction of fermionic coherent states is: if $E_+(f)$ is a coherent state in $\mathcal{T}^+(\mathcal{H})$ then

$$E_-(f) = \hat{\Theta} E_+(f) \tag{11}$$

is a coherent state in $\mathcal{T}^-(\mathcal{H})$. This relation follows from the definition of coherent states in Sect. 3.1 and from the properties of $\hat{\Theta}$. The explicit construction of Θ depends on an ordering prescription on the basic Hilbert space $\mathcal{H}$. There are essentially two possibilities:

1) Let $\mathcal{H}$ be an arbitrary Hilbert space of finite or countable infinite dimension. Choose an ON basis $\{e_i\}$, $i = 1,2\ldots$. Then Θ is defined as

$$\begin{cases} \Theta 1 = 1 \quad \text{(vacuum sector)} \\ \Theta e_{i_1} \wedge \ldots \wedge e_{i_n} = e_{i_1} \vee \ldots \vee e_{i_n} \qquad \text{if} \quad i_1 < i_2 < \ldots < i_n. \end{cases} \tag{12}$$

The ordering is therefore given by the ordering of the basis. For $n = 2,3,\ldots$ the range of this mapping is a closed subspace of $\mathcal{T}^+_n(\mathcal{H})$ since the occupation number is bounded by one. The inverse mapping is given by

$$\begin{cases} \hat{\Theta} 1 = 1, \\ \hat{\Theta} e_{i_1} \vee \ldots \vee e_{i_n} = 0 \quad \text{if two of the indices coincide,} \\ \hat{\Theta} e_{i_1} \vee \ldots \vee e_{i_n} = e_{i_1} \wedge \ldots \wedge e_{i_n} \qquad \text{if} \quad i_1 < i_2 < \ldots < i_n \end{cases} \tag{13}$$

2) The second method applies to $\mathcal{L}^2$ spaces with a diffuse measure. As the simplest case take $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$. Then any tensor in $\mathcal{T}^+(\mathcal{H})$ can be represented by a totally symmetric function $f_+(s_1,\ldots,s_n) \in \mathcal{L}^2(\mathbb{R}^n)$, and any tensor in $\mathcal{T}^-(\mathcal{H})$ can be represented by a totally antisymmetric function $f_-(s_1,..,s_n) \in \mathcal{L}^2(\mathbb{R}^n)$. The isomorphism between $\mathcal{T}^-_n$ and $\mathcal{T}^-_+$ obtained by multiplication with the Friedrichs–Klauder signature function $\sigma(s_1,\ldots,s_n) = \prod_{i<j} \mathrm{sign}\,(s_j - s_i)$ as

$$\Theta\, f_-(s_1,\ldots,s_n) := \sigma(s_1,\ldots,s_n)\, f_-(s_1,\ldots,s_n) = f_+(s_1,\ldots,s_n). \tag{14}$$

By linearity and closure Θ is then defined as a bijective mapping between the Fock spaces.

The second method can be extended to more general $\mathcal{L}^2$–spaces of the type $\mathcal{L}^2(M,\lambda)$ where M is a topological space with a diffuse measure λ. The space M should allow a partial ordering $m_1 < m_2$ with the restriction that the product measure of all pairs (m_1, m_2) which cannot be ordered is zero. See ref. [5] for more details. Roughly speaking, this second method refers to an ordering of a continuous basis. The essential difference to the first method is that Θ comes out as a bijective mapping between the Fock spaces. There are, of course, also mixed types of the mapping Θ which naturally emerge if $\mathcal{H} = \mathcal{L}^2(M,\lambda)$ where the measure λ has atoms and a diffuse part.

If $\mathcal{H} = \mathcal{V} \oplus \mathcal{V}^*$ has the structure given in Sect. 2 the mapping Θ should respect the involution, i.e. $(\Theta F)^* = \Theta(F^*)$ for all $F \in \mathcal{T}^-(\mathcal{H})$ and $\Theta(F^* \wedge G) = (\Theta F^*) \vee \Theta G$ for all $F,G \in \mathcal{T}^-(\mathcal{V})$.

The Fock space $\mathcal{T}^-(\mathcal{H})$ can now be represented by the functions

$$F \in \mathcal{T}^-(\mathcal{H}) \longrightarrow \psi_F(\xi) = (E_-(\xi)|F) = (\exp \xi|\Theta F) \in \mathcal{F}(\mathcal{D}) \qquad (15)$$

with the alternatives for $\mathcal{D}$ given in Sect. 3.1. The algebraic structure of the antisymmetric tensor algebra can no longer be established in a simple way in the reproducing kernel space $\mathcal{F}(\mathcal{D})$ since the relation $E_-(f) \wedge E_-(g) = E_-(f+g)$ is only possible for vectors f and g which are ordered $f < g$ in the sense given above.

As example for the exterior multiplication of two functions (15) take $\psi_f(\xi) = (\xi|f)$ and $\psi_g(\xi) = (\xi|g)$ with $f,g \in \mathcal{H}$. Then an injection Θ of the type (12) yields a multiplication $\psi_f(\xi) \wedge \psi_g(\xi) = \psi_{f\wedge g}(\xi) = \sum_{i<j} (f_i g_j - f_j g_i)\, \xi_i \xi_j$ where $f_i = (e_i|f)$, $\xi_i = (e_i|\xi)$ etc. are components in the ON basis e_i, $i \in \mathbb{N}$. That is the multiplication of functions is modeled after the exterior multiplication of forms, a construction which also shows up in the fermionic integration of Gaveau and Schulman [11]. Examples with a bijection Θ of the type (14) can be found in [3] and [5], a more general construction is given in [12].

4. THE WIENER REPRESENTATION

Despite of its technical versatility the Bargmann–Fock representation has a delicate probabilistic interpretation because the measure space does not have the correct number of degrees of freedom. It is therefore essential that also the Wiener representation of the bosonic Fock space can be carried over to the fermionic Fock space. The problems of this Section lead to an investigation of normal ordering on the function space $\mathcal{F}(\mathcal{H}_1)$. For the technical details see [5].

The Wiener representation we use here is based on the choice b) of the coherent states in Sect. 3.1. Let $\mathcal{H}^{\mathbb{R}}$ be the restriction of the basic Hilbert space $\mathcal{H}$ to its real subspace. Then the white noise measure (3) on $\mathcal{H}^{\mathbb{R}}$ is a countable additive measure only on an extension $\mathcal{H}^{\mathbb{R}}_{-1}$. For all $F \in \mathcal{A}^+(\mathcal{H})$, the functions $\varphi_F(\xi) = (\exp \xi|F)$ are square integrable polynomials on $\mathcal{H}^{\mathbb{R}}_{-1}$ which satisfy $\int \overline{\varphi_F(\xi)}\, \varphi_G(\xi)\, d\nu(\xi) = (\mathcal{W}_+^{-1}F|\mathcal{W}_+^{-1}G)$. The operator $\mathcal{W}_+$ is the invertible mapping of normal ordering or Wick ordering on the algebra $\mathcal{A}^+(\mathcal{H})$. It is defined by the identity

$$(G|\mathcal{W}_+F) = (G \vee \exp(-\Omega_+)|F) \qquad (16)$$

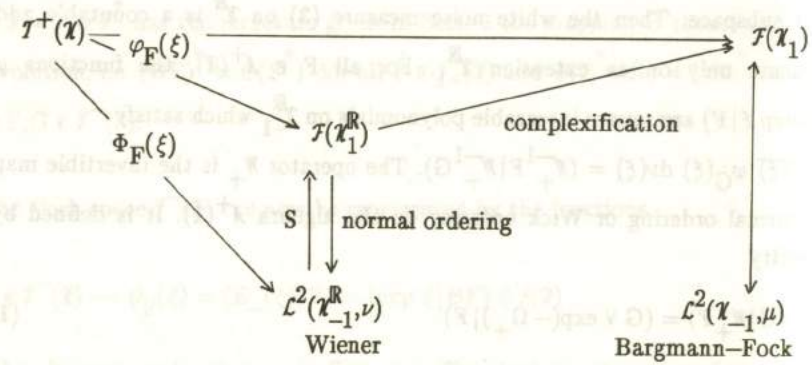
for all $F \in \mathcal{A}^+(\mathcal{H})$ and all $G \in \mathcal{A}^+(\mathcal{H})$. Here $\exp\Omega_+ = 1 + \Omega_+ + \frac{1}{2!}\,\Omega_+ \vee \Omega_+ + \ldots$ is the exponential series of the tensor Ω_+ defined by the fundamental bilinear symmetric form in (2). The mapping $\mathcal{W}_+$ induces the usual normal ordering for the polynomials $\varphi_F(\xi)$ with $F \in \mathcal{A}^+(\mathcal{H})$ by $\Phi_F(\xi) := \varphi_{\mathcal{W}_+F}(\xi)$. (These normal ordered polynomials Φ_F are usually denoted by $:\varphi_F(\xi):$.) The integrals of the normal ordered polynomials are then $\int \overline{\Phi_F(\xi)}\, \Phi_G(\xi)\, d\mu(\xi) = (F|G)$ and the mapping $F \in \mathcal{A}^+(\mathcal{H}) \longrightarrow \Phi_F(\xi) \in \mathcal{L}^2(\mathcal{H}^{\mathbb{R}}_{-1},\nu)$ can be extended to the Wiener isomorphism $\mathcal{T}^+(\mathcal{H}) \simeq \mathcal{L}^2(\mathcal{H}^{\mathbb{R}}_{-1},\nu)$.

The inverse of normal ordering is a nice integral operator [13]

$$(S\Phi)(\eta) = \int \Phi(\xi)\, e^{(\eta|\xi) - \frac{1}{2}(\eta|\eta)}\, d\nu(\xi) \qquad (17)$$

defined for arguments $\eta \in \mathcal{X}^{\mathbb{R}}$. If $F \in \mathcal{A}^{+}(\mathcal{X})$ then $(S\Phi_F)(\eta)$ agrees with $\varphi_F(\eta)$ for all $\eta \in \mathcal{X}^{\mathbb{R}}$. A closer inspection shows that S is an isometric mapping from $\mathcal{L}^2(\mathcal{X}^{\mathbb{R}}_{-1},\nu)$ onto the reproducing kernel space $\mathcal{F}(\mathcal{X}^{\mathbb{R}}_1)$. This isomorphism is successfully used in white noise analysis [13) 14)].

For the bosonic Fock space we have now established the following representations which are known for a long time



The new result, which we shall derive now, is that the whole diagram can be transferred to the fermionic case. We may start from any injection $\Theta\colon \mathcal{T}^{-}(\mathcal{X}) \to \mathcal{T}^{+}(\mathcal{X})$ to enter this diagram. But the situation is more subtle. On the algebra $\mathcal{A}^{-}(\mathcal{X})$ the normal ordering operator $\mathcal{N}_-$ can be defined in agreement with quantum field theory in the same way as $\mathcal{N}_+$ in (16) by

$$(G|\mathcal{N}_-F) = (G \wedge \exp(-\Omega_-)|F). \qquad (18)$$

Here the antisymmetric functional (2) mediates the contractions for normal ordering. Now it is possible to choose the mapping $\Theta\colon \mathcal{T}^{-}(\mathcal{X}) \to \mathcal{T}^{+}(\mathcal{X})$ in such a way that

$$\mathcal{N}_- = \hat{\Theta}\, \mathcal{N}_+\, \Theta \qquad (19)$$

is satisfied. This is the essential identity to transfer the whole diagram including the Wiener representation to the fermionic case. The isomorphism between $\mathcal{T}^{-}(\mathcal{X})$ and $\mathcal{L}^2(\mathcal{X}^{\mathbb{R}}_{-1},\nu)$ (or a subspace of it) is given by

$$F \in \mathcal{T}^{-}(\mathcal{X}) \to \Psi_F(\xi) = (\hat{\Theta} \exp \xi|\, \mathcal{N}_-F) = \Phi_{\Theta F}(\xi). \qquad (20)$$

The inverse of normal odering coincides again with the integral transform (17). A construction of a bijective mapping Θ between $\mathcal{T}^{-}(\mathcal{X})$ and $\mathcal{T}^{+}(\mathcal{X})$ for Hilbert spaces $\mathcal{X} = \mathcal{L}^2(\mu,\lambda)$ which satisfies (19) has been given in [5)]. Here we shall only consider the simpler version of an injection (12) for a finite dimensional space $\mathcal{X}$, $\dim \mathcal{X} = 2N$. The isotropic decomposition is given by the following arrangement of an ON basis $e_1,\ldots,e_{2N}$: $\mathcal{V} = \mathcal{L}\{e_1,e_3,\ldots e_{2N-1}\}$ and $\mathcal{V}^* = \mathcal{L}\{e_2,e_4,\ldots e_{2N}\}$ with an involution induced by $e^*_{2n-1} = e_{2n}$, $n = 1,2,\ldots,N$. The bilinear forms (2) are generated by the tensors

$$\Omega_+ = \sum_{n=1}^{N} e_{2n-1} \vee e_{2n} \in \mathcal{A}^{+}(\mathcal{X}) \quad \text{and} \quad \Omega_- = \sum_{n=1}^{N} e_{2n-1} \wedge e_{2n} \in \mathcal{A}^{-}(\mathcal{X}).$$

Then the injective mapping (12) obviously satisfies $\hat{\Theta}\Omega_+\Theta = \Omega_-$ and (19) is a consequence of (10), (16) and (18).

5. THE ORNSTEIN–UHLENBECK SEMIGROUP

Let $\mathcal{X} = \mathcal{V} \oplus \mathcal{V}^*$ be the complex Hilbert space of Sect. 2. and let A be an invertible positive operator on $\mathcal{X}$. Then e^{-At}, $t \geq 0$, is a semigroup on $\mathcal{X}$ which induces the semigroups $\Gamma_+(e^{-At})$ on $\mathcal{T}^{+}(\mathcal{X})$ and $\Gamma_-(e^{-At})$ on $\mathcal{T}^{-}(\mathcal{X})$. It is well known that $\Gamma_+(e^{-At})$ has an underlying trajectory picture given by an Ornstein–Uhlenbeck process, see e.g. [15)]. In the following this trajectory picture will be transferred to the fermionic case $\Gamma_-(e^{-At})$.

For completeness and to give the notations we first have to recapitulate the essential facts about the Ornstein–Uhlenbeck process in the bosonic case. To simplify some arguments we assume that A^{-1} is a bounded operator. Moreover A has to be real, i.e. $(Af)^* = Af^*$ for all f in the domain of A. Then A is also a positive operator with a bounded inverse on the real subspace $\mathcal{H}^{\mathbb{R}} = \{f | f^* = f \in \mathcal{H}\}$. The isomorphisms between $\mathcal{T}^{\pm}(\mathcal{H})$ and the reproducing kernel spaces $\mathcal{F}(\mathcal{D})$ or the Wiener representation $\mathcal{L}^2(\mathcal{H}^{\mathbb{R}}_{-1},\nu)$ are taken from the proceeding sections. (In the euclidean quantum field theory it is convenient to absorb the operator A, i.e. the hamiltonian, in the definition of the measure by a substitution $\xi \to \sqrt{A}\,\xi$, see e.g. [14] and [15]. Such a modification is also possible here, but it is not necessary.)

The semigroup $\Gamma_+(e^{-At})$ operates on the functions $\varphi_F(\xi) = (\exp \xi | F) \in \mathcal{F}(\mathcal{D})$ with $\mathcal{D} = \mathcal{H}_1$ or $\mathcal{H}_1^{\mathbb{R}}$ as

$$\varphi_F(\xi) \to (\exp \xi | \Gamma_+(e^{-At})F) = \varphi_F(e^{-At}\xi) \tag{21}$$

which induces the mapping $\mathcal{E}_f(\xi) \to \mathcal{E}_f(e^{-At}\xi) = \mathcal{E}_{e^{-At}f}(\xi)$ for the coherent states in $\mathcal{F}(\mathcal{H}_1^{\mathbb{R}})$. Normal ordering then yields the corresponding evolution on the space $\mathcal{L}^2(\mathcal{H}^{\mathbb{R}}_{-1},\nu)$

$$\Phi_F(\xi) \to \Phi_{\Gamma_+(e^{-At})F} = (\exp \xi | \mathcal{W}_+\, \Gamma_+(e^{-At})\, F). \tag{22}$$

For the coherent states of the Wiener representation $\mathcal{E}^W_f(\xi) \to \mathcal{E}^{(\xi|f) - \frac{1}{2}\|f\|^2}$ this relation reduces to

$$\mathcal{E}^W_f(\xi) \to \mathcal{E}^W_{e^{-At}f}(\xi). \tag{23}$$

This semigroup on $\mathcal{L}^2(\mathcal{H}^{\mathbb{R}}_{-1},\nu)$ will be denoted by $P_t\Phi$, $t \geq 0$. It can be represented by the Mehler formula of the Ornstein–Uhlenbeck semigroup, see [16], where a detailed investigation of the case A = I is given,

$$(P_t\Phi)(\xi) = \int \Phi\,(e^{-At}\xi - \sqrt{1-e^{-2At}}\,\eta\,)\,d\nu(\eta). \tag{24}$$

This integral transform has a well defined meaning at least for all finitely based functions (cylinder functions) Φ. Application of (24) to the coherent states $\mathcal{E}^W_f(\xi)$ leads to (23) and proves that the Ornstein–Uhlenbeck semigroup (24) is indeed the semigroup under consideration. The semigroup operator P_t is selfadjoint in $\mathcal{L}^2(\mathcal{H}^{\mathbb{R}}_{-1},\nu)$, which can be easily checked for the coherent states. This essential property implies that the measure ν is invariant against P_t. The semigroup determines a markovian stochastic process X_t on $\mathcal{H}^{\mathbb{R}}_{-1}$, the Ornstein–Uhlenbeck process, such that

$$\mathbf{E}_\xi\, \Phi(X_t) = (P_t\Phi)(\xi) \tag{25}$$

where $\mathbf{E}_\xi$ is the expectation value for all trajectories starting at $X_0 = \xi$.

This underlying trajectory picture can be transferred to the Bargmann–Fock representation. Let Z_t be the complex Ornstein–Uhlenbeck process on $\mathcal{H}_{-1}$ with invariant measure μ then $\mathbf{E}_\xi\varphi(Z_t)$ can again be calculated by the Mehler formula

$$\mathbf{E}_\xi\varphi(Z_t) = \int \varphi(e^{-At}\xi + \sqrt{1-e^{-2At}}\,\eta)\,d\mu(\eta) = \varphi(e^{-At}\xi). \tag{26}$$

The last identity follows since $\varphi(\xi)$ is an antiholomorphic function in $\mathcal{L}^2(\mathcal{H}_{-1},\mu)$. This result agrees with the semigroup (21). In this case of the Bargmann–Fock representation the underlying trajectory picture is highly non–unique. The generator of the semigroup $\varphi(e^{-At}\xi)$ is easily calculated. Since the antiholomorphic function φ is a harmonic function the diffusion term cannot be fixed. The choice of

the process (26) is possible, but other processes lead to the same semigroup evolution for antiholomorphic functions. The necessity of additional constraints is well known from the study of phase space functional integrals [17].

To apply these results to the fermionic case the mapping $\Theta: \mathcal{I}^-(\mathcal{X}) \to \mathcal{I}^+(\mathcal{X})$ has to fulfil in addition to the requirements of Sect. 3.3. and of eq. (19) the condition

$$\hat{\Theta}\, \Gamma_+(A)\Theta = \Gamma_-(A). \tag{27}$$

This condition can indeed be satisfied by an ordering prescription on the spectrum of A, at least for all pseudodifferential operators A. The spectral decomposition for these operators corresponds to an isomorphism $\mathcal{X} \hat{=} \mathcal{L}^2(\mathcal{M},\lambda)$ with a manifold $\mathcal{M}$ which allows a partial ordering, that is needed to construct the mapping Θ, see Sect. 3.3. A condition like (27) has also been found in the investigation of the boson–fermion correspondence in field theoretical models [10].

As a consequence of (27) the semigroup $\Gamma_-(e^{-At})$ operates on the fermionic functions in $\mathcal{F}(\mathcal{X}_1)$, see (15), as $\psi_F(\xi) = (\exp \xi | \Theta F) \longrightarrow$ $(\exp \xi | \Theta\, \Gamma_-(e^{-At})F) = (\exp \xi | \Gamma_+(e^{-At})\Theta F) = \psi_F(e^{-At}\xi)$, which reproduces (21). Moreover the condition (27) allows to derive the semigroup (22) also for the Wiener representation of the fermions: $\Psi_F = \Phi_{\Theta F} \to \Psi_{\Gamma_-(e^{-At})F} = \Phi_{\Gamma_+(e^{-At})\Theta F}$. These relations can again be generated by the Mehler formulas (26) or (24), respectively. It is therefore possible to describe the dynamics of the semigroup $\Gamma_-(e^{-At})$ by an Ornstein–Uhlenbeck process in both the Bargmann–Fock and the Wiener representation.

There is an important difference to the bosonic case. The injection Θ depends due to the additional requirement (27) on the semigroup under consideration. The explicit representation of the exterior algebra in the function spaces $\mathcal{F}(\mathcal{D})$ and $\mathcal{L}^2(\mathcal{X}^{\mathbb{R}}_{-1},\nu)$ depends therefore also on the semigroup. Any perturbation of the semigroup which violates (27) destroys the trajectory picture of the Ornstein–Uhlenbeck process. This problem is technically a disadvantage which reflects some problems indicated by Fröhlich and Osterwalder already in 1974 in their investigation of the gage space approach to euclidean quantum field theory of fermions [18]. On the other side it shows that a realization of the fermionic dynamics by a stochastic process is related to a rather unique definition of coherent states for fermions which agrees with the boson–fermion correspondence of [1] [9] and [10]. To what extend such a trajectory picture persists under a multi–fermion interaction is an open problem.

References

1. Klauder, J.R.: "The action option and a Feynman quantization of spinor fields in terms of ordinary c–numbers." *Ann. Phys. (N.Y.)* **11**, 123–168 (1960).
2. Hudson, R.L. and Parthasarathy, K.R.: "Unification of fermion and boson stochastic calculus." Commun. Math. Phys. 104 457–470 (1986).
3. Meyer, P.A.: "Eléments de probabilités quantiques I–V". In: *Sém. Prob.* XX., 186–312. VI–VIII: *Sém. Prob.* XXI., 34–80, IX–X: *Sém. Prob.* XXII., 101–128. Berlin/Heidelberg/New York: Springer (Lect. Notes in Math. **1204, 1247** and **1321**) 1986, 1987 and 1988.
4. Kupsch, J.: "Functional integration for euclidean fermi fields." In: M. Pawlowski and R. Raczka (eds.): *Gauge theories of fundamental interactions*, 425–451. (Proceedings of the XXXII semester in the S. Banach Int. Math. Center 1988) Singapore: World Scientific 1990.
5. Kupsch, J. "A probabilistic formulation of bosonic and fermionic integration." *Rev. Math. Phys.* to be published.
6. Klauder, J.R.: "Exponential Hilbert space: Fock space revisited." *J. Math. Phys.* 11, 609–630 (1970).
7. Meschkowski, H.: *Hilbertsche Räume mit Kernfunktion.* Berlin/ Göttingen/ Heidelberg: Springer 1962.
8. Friedrichs, K.O.: *Mathematical aspects of the quantum theory of fields.* New York: Interscience 1953.

9. Garbaczewski, P. and Rzewuski, J.: "On generating functionals for anti-symmetric functions and their applications in quantum field theory." *Reports Math. Phys.* 6, 431–444 (1974).
10. Garbaczewski, P.: *Classical and quantum field theory of exactly soluble nonlinear systems.* Singapore: World Scinetific 1985.
11. Gaveau, B. and Schulman, L.S.: "Dirac Equation Path integral: interpreting the Grassmann variables." *Nuovo Cimento* 11D, 31–51 (1989).
12. Lindsay, J.M. and Parthasarathy, K.R.: "Cohomology of power sets with applications in quantum probability." *Commun. Math. Phys.* 124, 337–364 (1989).
13. Kubo, T. and Takenaka, S.: "Calculus on Gaussian white noise, I–IV". *Proc. Japan Acad.* 56A, 376–380 (1980). 56A, 411–416 (1080). 57A, 433–437 (1981). 58A, 186–189 (1982).
14. Hida, T., Potthoff, J. and Streit, L.: "White noise analysis and applications." In: L. Streit (ed.): *Mathematics + physics, Lectures on recent results.* Vol 3. Singapore: World Scientific 1988.
15. Simon, B.: *The $P(\phi)_2$ euclidean (quantum) field theory.* Princeton: Univ. Press 1974.
16. Meyer, P.A.: "Note sur le processus d'Ornstein–Uhlenbeck." In: *Sém. Prob.* XVI., 95–132. Berlin/Heidelberg/New York: Springer (Lect. Notes in Math. 920) 1982.
17. Klauder, J.R.: "Path integrals." *Acta Physica Austriaca*, Suppl. XXII, 3–49 (1980).
18. Fröhlich, J. and Osterwalder, K.: "Is there a euclidean field theory for fermions?" *Helv. Phys. Acta* 47, 781–805 (1974).

Methods for Solving Path Integrals and QED from Classical Particle Trajectories

I.H.Duru
Univ. of Trakya
Mathematics Dept.
Edirne-Turkey
and
TÜBİTAK, Marmara Research Centre
P.K.74, 41470 Gebze, Turkey

Abstract

Path integration is a basic tool in the semiclassical or the full quantum study of field theories. In practice the treatment of a field theoretical problem very often effectively reduces to the path integration for a non-relativistic particle motion. With this motivation we present some general techniques employed for solving several quantum mechanical path integrals (both time independent and time dependent). At the end we review briefly the QED obtained by summing over the classical paths.

I. Introduction

In field theories whenever we have a two point function we can express it as a path integral. For example to obtain the vacuum (or a temperature) expectation value of a quantized field in a given environment one employs the Green function for the motion of the particle corresponding to the field in that environment [1]. Another example is the study of the vacuum decay for a given geometry or field, where one uses the Green function which propagates the particle corresponding to the produced quantum field [2].

In the semiclassical quantization of the non-linear systems, to quantize the fluctions around a given solitonic solution we employ path

integrals.

For most of the above cases, in practice the problem in hand is very often reduced to the development of a non-relativistic path integration [3]. Thus it is of interest to see what is our position in the exact solutions of the path integrals in the quantization of particle motions. With this motivation we will present several techniques used for solving the quantum mechanical path integrals. We first will deal with time-independent problems, then give two time-dependent examples.

At the end we briefly summarize the quantum electrodynamics obtained by the summation over the classical particle trajectories.

II. Time-independent path integrals

To present techniques for the general solution it is enough to consider one-dimensional quantum mechanical path integrals. For the motion of a mass μ in the potential $V(x)$, the probalitiy amplitued to go from the space-time point x_a,t_a to x_b, t_b is given by (with $\hbar=c=1$, $t_b-t_a=T$).

$$K(x_a,x_b;T) = \int DxDp_x \exp\left[i \int_{t_a}^{t_b} dt \left(p_x\dot{x} - \frac{p_x^2}{2\mu} - V(x)\right)\right] \qquad (1)$$

which is understood as the limit of the usual time-graded formula:

$$K(x_a,x_b;T) = \lim_{\substack{n\to\infty \\ \varepsilon\to 0}} \prod_{j=1}^{n} \int dx_j \prod_{j=1}^{n+1} \int \frac{dp_{xj}}{2\pi} \prod_{j=1}^{n+1} \exp\left[i\left(p_{xj}(x_j-x_{j-1}) - \frac{\varepsilon p_{xj}^2}{2\mu} - \varepsilon V(x_j)\right)\right] \qquad (2)$$

$$(n+1)\varepsilon = T; \quad x_o,t_o = x_a,t_a \; ; \quad x_{n+1}, t_{n+1} = x_b, t_b \; .$$

There are very few cases for which the above integrals can directly be executed:Free particle, harmonic oscillator [4], infinite square well potential [5], and the inverse square potential [6].

For most of the potentials the integrals in (2) are not soluable. We then employ canonical transformations to map the problem in hand into managable forms. Suppose we have

$$x = g(Q) \; , \; Q = g^{-1}(x) \; ; \; p_x = \frac{dg^{-1}}{dx} P \qquad (3)$$

The integral measure in (2) is not invariant under the above transformations for there is an extra momentum integration at point b. The Green function of (2) becomes:

$$K(x_a,x_b;T) = \left(\frac{dg^{-1}}{dx}\right)_b \int DQDP \exp\left[i \int_{t_a}^{t_b} dt \left(PQ - \left(\frac{dg^{-1}}{dx}\right)^2 \frac{P^2}{2\mu} V(g(Q))\right)\right] \qquad (4)$$

To eliminate the factor multiplying the kinetic energy, we reparametrize the dynamical variables in terms of a new "time" S given by

$$dt = (dg^{-1}/dx)^{-2} \, ds \qquad (5)$$

It can be enforced via the identity:

$$1 = \int_0^\infty dS \, \delta\left(T - \int_{S_a}^{S_b} ds \, (dg^{-1}/dx)^{-2}\right)(dg^{-1}/dx)_b^{-2}$$

$$= \int_0^\infty dS \int_{-\infty}^{\infty} \frac{dE}{2\pi} e^{iET} e^{-i\int_{S_a}^{S_b} ds \, E(dg^{-1}/dx)^{-2}} (dg^{-1}/dx)_b^{-2} \qquad (6)$$

with $S = S_b - S_a$. We then have

$$\kappa(x_a,x_b;T) = \int_0^\infty dS \int_{-\infty}^{\infty} \frac{dE}{2\pi} e^{iET} \left(d\bar{g}^{-1}/dx\right)_b^{-1}$$
$$\int DQDP \exp\left[i \int_{S_a}^{S_b} ds \left(PQ' - \frac{P^2}{2\mu} - \left(d\bar{g}^{-1}/dx\right)^{-2} V\text{-}E\left(d\bar{g}^{-1}/dx\right)^{-2}\right)\right] \quad (7)$$

where the over prime stands for derivative with respect to s. To obtain a symmetric expression with respect to the points x_a and x_b, we rewrite the Jacobian $J(Q_b) \equiv \left(d\bar{g}^{-1}/dx\right)_b^{-1}$ as:

$$J(Q_b) = \sqrt{J_b J_a} \exp \frac{1}{2} \ell n \frac{J_b}{J_a}$$
$$= \sqrt{J_b J_a} \exp i \int_{S_a}^{S_b} ds \, (-i) \frac{dJ/dQ}{Q} Q' \quad (8)$$

After inserting (8) into (7) and then translating P by $P \to P + i\,(dJ/dQ)/Q$, we get [7] :

$$\kappa(x_a,x_b;T) = \int_0^\infty dS \int_{-\infty}^{\infty} \frac{dE}{2\pi} e^{iET} \sqrt{J(Q_a)\, J(Q_b)}\ \kappa(Q_a,Q_b;S) \quad (9)$$

Here

$$\bar{\kappa}(Q_a,Q_b;S) = \int DQDP \exp\left[i \int_{S_a}^{S_b} ds \left(PQ' - \frac{P^2}{2\mu} - u(Q)\right)\right] \quad (10)$$

represents a motion in Q,s space-"time" with the potential

$$u(Q) = J^2(Q)\,(V(g(Q)) + E) - \frac{(dj/dQ)^2}{2\mu Q^2} \quad (11)$$

There is a vast literature concerning the transformation of the path integral measure and the ordering terms of the action. There are several other recipies which are differ from the one we use and all are claimed to be rigorous. Our preference is to employ the required canonical transformations in the continuum formulation instead of struggling with the details of the discrete formulae, and to determine the correct form of the ordering terms consulting with the Schrödinger picture when necessary.

Now to give a simple example to the method presented, we consider the Morse potential [8]:

$$V(x) = V_o(e^{-2ax} - 2e^{-ax}) \ , \ x \geq o \quad (12)$$

Appropriate coordinate and time transformations are given by

$$x = -\frac{2}{a} \ell nQ \ , \ Px = -\frac{a}{2} QP \ ; \ dt = ds/Q^2 \quad (13)$$

New dynamics in (Q,s) space-time are governed by the Hamiltonian

$$H(Q) = \frac{P^2}{2(4\mu/a^2)} + V_o Q^2 + \frac{E}{Q^2} - 2V_o \quad (14)$$

whose path integral was solved long time ago by expanding the short time propagator [6].

There exists an alternative method to solve (14) which can be used for many other problems [9]. Consider the potential

$$V(r) = ar^{-2} + br^2 \ , \ r \geq o \quad (15)$$

We introduce an "auxiliary" dynamic governed by the Hamiltonian

$$H_Q = (P_Q^2 + 1/4)/2 \ \mu r^2 \quad (16)$$

via the identity :

$$e^{-i\int_0^T dt\, a/r^2} = \int_{-\infty}^{\infty} dQ_b\, e^{-i\theta_b\sqrt{2\mu a + 1/4}}$$

$$\int D\theta Dp_\theta \exp[i\int_0^T dt\,(-p_\theta\dot{\theta} + \frac{p_\theta^2 + 1/4}{2\mu r^2})] \qquad (17)$$

Substituting the above identity into the path integral expression corresponding to the potential (15), we obtain

$$K(\vec{r}_a, \vec{r}_b; T) = (r_a r_b)^{1/2} \int_{-\infty}^{\infty} d\theta_b\, e^{\nu\theta_b}\, \bar{K}\,(r_a, \theta_a = 0\ ;\ r_b, \theta_b; T) \qquad (18)$$

where $\vec{r}$ is 2-dim- vector and $\nu = (1/2)\sqrt{1+8\mu a}$.

$$\bar{K}(r_a, o; r_b\theta_b\ ; T) = (r_a r_b)^{-1/2} \int D(r,\theta)\, D(p_r, p_\theta)$$

$$\exp[i\int_0^T dt(p_r\dot{r} - p_\theta\dot{\theta} - \frac{1}{2}(p_r^2 + \frac{p_\theta^2 + 1/4}{r^2} + ar^2)] \qquad (19)$$

The above "polar" coordinate path integral, with the term 1/4 in the angular part being the ordering contribution, is equivalent to the following Cartesian formula:

$$\bar{K} = \int D^2r\, D^2p \exp\,[i\int_0^T dt(p_x\dot{x} - p_y\dot{y} - \frac{1}{2\mu}(p_x^2 - p_y^2) - b(x^2 - y^2))] \qquad (20)$$

The Cartesian coordinates are given by $x = r\mathrm{ch}\theta$, $y = r\mathrm{sh}\theta$, $o \leq x < \infty$, $-\infty < y < \infty$ and the scalar product is defined as $\vec{r}_a \cdot \vec{r}_b = x_a x_b - y_a y_b$.

The techniques which have been discussed upto this point are also applicable to problems in more then one dimension. In fact the essence of these methods was first introduced for the H-atom path integration

[10]. There, after the introduction of a fourth degree of freedom, one employs a canonical transformation subject to a constraint. Then the adoption of a new "time" maps the problem into the four-dimensional harmonic oscillator, whose solution can be written from the known results.

Now, we like to mention another method which makes use of the underlying group symmetries of the problems. It is well known that many special functions we find in physics are related to the matrix elements of the representations of Lie groups [11]. This fact enables us to relate many quantum mechanical problems to the particle motions over the group manifolds. For example the quantization of the motion over the SU(2) space which is parametrized in terms of the Euler angles is utilized in the study of several potentials [12]: Consider the potential(Pöschl-Teller, 1938):

$$V = \frac{1}{2\mu}\,(\frac{A}{\sin^2 x} + \frac{B}{\cos^2 x}) \qquad (21)$$

By changing to θ by $\theta = 2x$, the corresponding Hamiltonian becomes

$$H = \frac{p_\theta^2}{2\mu} + \frac{\alpha^2+\beta^2-2\alpha\beta\cos\theta - 1/4}{2\mu\sin^2\theta} \qquad (22)$$

where α,β are constants defined in terms of A,B. The above expression is the Hamiltonian for the "SU(2) rotator" , with the angular momenta p_ϕ, p_ψ (ϕ,ψ = Euler angles) being fixed to be α,β. Introduction of an auxiliary dynamic in ϕ,ψ-space solves the problem [12]. The group SU(2) is also employed for Wood-Saxon, Rosen-Morse and Hulthén potentials [4]. It is obvious that, if the angle x in (21) is replaced by the hyperbolic one the group to be used becomes SU(1,1) [13].

The last method we like to present is the transformation to the Hamilton-Jacobi coordinates [14]. The idea is to transform to a null Hamiltonian problem by absorbing the potential in a new generalized momentum P (or a new generalized coordinate Q). For example we may set

$$p^2 = p_x^2/2\mu + V(x) \quad , \quad Q = \partial F_2/\partial p = \int^x dx \frac{4\mu p}{\sqrt{2\mu(p^2 - V(x))}} - 2pt \tag{23}$$

where F_2 is the generating function

$$F_2(x,p,t) = \int^x dx \sqrt{2\mu(p^2 - V(x))} - p^2 t \tag{24}$$

We can exacute trivally all the intermediate integrals, and arrive at an expression with an ordinary integral over the last momentum $dp_x \equiv dp_{x_{n+1}}$

$$K(x_b, x_a; T) = \int \frac{dp_x}{2\pi} \exp[i(F_2(b) - F_2(a))] \tag{25}$$

Here dp_x may be obtained from (23) in terms of dP with x being constant at the end points. The method has been used to solve the square potential in one dimension and the radial square-well potential in two dimensions. In most of the cases the integral (25) is not soluable; however, one may hope to employ numerical methods.

III. Time-dependent path integrals

Upto this point we have dealt with time-independent potentials. Now we will consider two types of time-dependent examples:

We first montion the harmonic oscillator potential with an arbitrary time-dependent frequency:

$$V(x,t) = \frac{\mu}{2} \omega^2(t)\, x^2 \tag{26}$$

Note that when one quantize the fluctuations around a solitonic solution for a nonlinear system one effectively comes across the above type potential with $\frac{\mu}{2}\omega^2(t) = \frac{1}{2} U''(x_{cl}(t))$ where U(x) is the original

nonlinear potential, $x_{cl}(t)$ is the solitonic solution and the over prime is a derivative with respect to x [15]. (26) is also equivalant to the free motion of a particle subject to time dependent boundary conditions [16]. Path integration of (26) is known [17]:

The canonical transformations to be used are

$$x = C(t)Q \quad , \quad P = Cp - \mu \dot{C} x \quad , \quad \frac{\ddot{C}}{C} = -\omega^2(t) \tag{27}$$

generated by

$$F_2(x,p,t) = \frac{x}{C(t)} P + \frac{\mu}{2} \frac{\dot{C}(t)}{C(t)} x^2 \quad .$$

Adoption of the above coordinates leads to

$$K(x_a, x_b; t_a, t_b) = \frac{\sqrt{i/2\pi}}{\sqrt{C_a C_b}} \sqrt{\frac{1}{\int_{t_a}^{t_b} dt/C^2}} \exp \frac{i\mu}{2} \left(\frac{\dot{C}_b}{C_b} x_b^2 - \frac{\dot{C}_a}{C_a} x_a^2\right) \exp \frac{i\mu}{2} \frac{(x_b/C_b - x_a/C_a)^2}{\int_{t_a}^{t_b} dt/C^2} \tag{28}$$

The second type of time-dependent potential that we are going to deal with is given by [18]

$$V(x,t) = V(x - f(t)) \tag{29}$$

where $f(t)$ is an arbitrary function of time. To solve the path integral for (29), we simply shift the coordinate by the canonical transformation

$$Q = t - f(t) \quad , \quad P = p - \mu f(t) \tag{30}$$

generated by

$$F_2(x,p,t) = (P + \mu\dot{f})(x-f) \ .$$

The problem is then mapped into the one for a linear potential with time dependent coefficient. The Green function is

$$\kappa(x_a,x_b;\, t_a,t_b) = \exp\{i\mu[\dot{f}_b(x_b-f_b)-\dot{f}_a(x_a-f_a)]\}\exp\Big(\frac{i\mu}{2}\int_{t_a}^{t_b} dt\ \dot{f}^2\Big)\bar{\kappa} \tag{31}$$

where

$$\bar{\kappa}(\theta_a,\theta_b;\, t_a,t_b) = \int D\theta DP\ \exp[i\int_{t_a}^{t_b} dt\ (PQ - \frac{p^2}{2\mu} - V(\theta) - \mu\ddot{f}\theta^2)] \tag{32}$$

There are several examples where the above path integral is exactly solvable[18]. On the other hand if the static problem $V(\theta)$ is exactly solvable we can treat the extra time dependent linear term as a perturbation:

We rewrite (32) by expanding the linear term of the action as

$$\bar{\kappa}(a,b) = \sum_{n=o}^{\infty} \frac{(-i)^n}{n!}\int DQDP\ (\int_{t_a}^{t_b} dt\ \mu\ddot{f}Q)^n\ \exp[i\int_{t_a}^{t_b} dt(P\dot{Q} - \frac{p^2}{2\mu} - V(\theta))] \tag{33}$$

After the familar manipulations we arrive at [1]

$$\bar{\kappa}(a,b) = \kappa^{(0)}(a,b) + \kappa^{(1)}(a,b) + \kappa^{(2)}(a,b) + \ldots . \tag{34}$$

where $\kappa^{(0)}$ is the propagator for potential $V(\theta)$, wheras the following terms represent propagations with several numbers of interactions via the perturbation term. For example

$$\kappa^{(1)}(a,b) = -i\int_{t_a}^{t_b} dt_1 \int_{-\infty}^{\infty} dQ(t_1)\kappa^{(0)}(Q_b t_b;\, Q_1 t_1)\ \mu\ddot{f}(t_1)Q(t_1)\kappa^{(0)}(Q,t_1;\, Q_b t_a) \tag{35}$$

is the propagator from a to b with one scattering at the spacetime point Q_1,t_1. From one interaction point to the next one the system propagates by the Green function of the static potential $V(\theta)$.

IV. Quantum Electrodynamics

The perturbative path integration presented at the end of last section was also employed in the recent formulation of QED in terms of the classical particle trajectories [19]. In this approach we obtain the Feynman graphs from a covariant action without the intermediary of quantized fields.

We start with the action for N particle interaction via the electromagnetic field A_μ

$$W = \sum_{k=1}^{N} (\int_{o}^{S_k} d\tau_k\ L_k^o - \int dx\ \overset{\mu}{J}_k(x)\ A_\mu(x)\ ) - \frac{1}{4}\int dx F_{\mu\nu}F^{\mu\nu} \tag{36}$$

where τ_k is the parameter time for k^{th} particle, L_k^o and $\overset{\mu}{J}_k(x)$ are the Lagrangian and the current respectively.

Using the Maxwell equations and assuming that the fields vanish at infinity, the last term in the above action becomes equal to the half of the second term with the opposite sign. Thus we write (1) as

$$W = \sum_{k=1}^{N} (\int_{o}^{S_k} d\tau_k L_k^o - \frac{1}{2}\int dx\ \overset{\mu}{J}_k(x)\ A_\mu(x)\ ) \tag{37}$$

Note the above Lagrangian can no longer be used to derive the equations of motions:

We then express the electromagnetic field in terms of the currents due to the charged particles in the theory as

$$A^{\mu}(x) = \int dy\ D(x-y) \sum_{k} \dot{j}^{\mu}_{k}(x) \tag{38}$$

with the gauge $A^{\mu}{}_{,\mu} = 0$. The Green function D(x-y) will be specified explicitly. The action then becomes

$$W = W_{o} + W_{int}. \tag{39}$$

where the interaction action is

$$W_{int} = -\frac{1}{2} \sum_{k,m=1}^{N} \int dx\ \dot{j}^{\mu}_{k}(x)\ D(x-y)\ \dot{j}^{k}_{\mu}(y)\ dy \tag{40}$$

in which the free electromagnetic field is eliminated and the N particles (scalar or spinor) interact via the Green function. It also includes self interactions (k=m terms). Note that our approach, in a sense resembles to Feynman-Wheeler absorber theory for the photons are always emitted to be absorbed.

Using the Fourier expansion of D(x-y), W_{int} becomes

$$W_{int} = \frac{1}{2} \sum_{\ell,m=1}^{N} \frac{1}{(2\pi)^4} \int dxdy\ \dot{j}^{\mu}_{\ell}(x)\ e^{-ikx}\ \dot{j}^{m}_{\mu}(y)\ e^{iky}\ \frac{dk}{k^2} \tag{41}$$

By expressing the current in terms of the world line of the particle

$$\dot{j}^{\ell}_{\mu}(x) = e \int d\tau_{\ell}\ \delta(x-x(\tau_{\ell}))\ \dot{j}^{\ell}_{\mu}(\tau) \tag{42}$$

and then integrating over dxdy we obtain an "action-at-a-distance" interaction in terms of the particle coordinates alone:

$$W_{int} = \frac{1}{2} \sum_{\ell,m=1}^{N} \frac{e_{\ell}e_{m}}{(2\pi)^4} \int d\tau''d\tau''\ \frac{dk}{k^2}\ M_{\ell}(k,\tau').M_{m}(-k,\tau'') \tag{43}$$

where

$$M^{\mu}_{\ell}(k,\tau) = \int \dot{j}^{\mu}_{\ell}(\tau)\ e^{-ikx_{\ell}(\tau)} \tag{44}$$

By expressing the interaction in terms of the classical particle coordinates we can write the propogator κ for the system of N particles as the phase space path integral

$$\kappa(a,b) = \prod_{r=1}^{N} \int_{o}^{\infty} dS_{r}\ \exp[-if(\mu)\ S_{r}] \int D_{r}(\cdot)\ \exp\left(i \int_{o}^{S_{r}} d\tau_{r}\ L_{r}\right)\ e^{iW_{int}} \tag{45}$$

Here (a,b) stands for the collection of the fixed initial and final coordinates. Each particle r has an arrival "time" S_r from a to b. We integrate over all possible S_r intervals with a measure $\exp[-if(\mu)S_r]$. The function $f(\mu)$ to be specified for scalar and spinor particles. The path integral measure for the r^{th} particle is defined by the "time"-sliced formulation as

$$D_{r}(\cdot) = \prod_{j=1}^{n+1} \frac{d^4p_j}{(2\pi)^4} \prod_{j=1}^{n} d^4x_j \prod_{j=1}^{n+1} d\xi_j \tag{46}$$

where $(d\xi_j)$ stands for the possible internal coordinates which to be specified for the spinning particles.

After introducing (43) into (45), we expand $\exp(iW_{int})$ in powers of $e_{\ell}e_{m}$ to obtain the perturbative treatment. The procedure is the direct generalization of the one given in the Feynman-Hibbs book. The details of this treatment which produces all the QED graphes are given in Ref.(16). Here we will only write down the free actions and currents employed in (45) and (43):

For the scalar particles we use

$$L_r^o = p.\dot{x} - \frac{1}{2}p^2 \quad , \quad f(\mu) = \mu^2 \quad , \quad \overset{r}{\underset{\mu}{j}}(\tau) = e\, p_\mu^r(\tau). \tag{47}$$

For the spinning particles we employ the classical electron theory introduced by Barut and Zanghi [20]:

$$L_r^o = -\frac{\lambda r}{2}(\dot{\bar{z}}_r z_r - \bar{z}_r \dot{z}_r) + p_\mu^r(\dot{x}_r^\mu - \bar{z}_r \gamma^\mu z_r) \quad , \quad f(\mu) = \frac{\mu}{\lambda r}$$

$$\overset{\mu}{\underset{r}{j}}(x) = e_r \bar{z}_r \gamma^\mu z_r\, \delta(x^\mu - x^\mu(\tau_r)) \tag{48}$$

Here (x_μ, p_μ) is the usual phase space, wheras $(z, i\bar{z})$ are another pair of conjugate (internal) spin variables; $z(\tau)$ is a c-number four-component spinor , $\bar{z} = \overset{+}{z}\gamma^o$, λ_r is a constant with dimension of action. In this formulation the 16-dimensional phase space $\Gamma = (x_\mu, p_\mu\ ;\ z, \bar{z})$ consists of the c-numbers only, thus it should be treated in a similar fashion to any classical system. The internal sector of the path integral measure for the r^{th} electron is given by

$$\Big(\prod_{j=1}^{n+1} d\xi j\Big)_r = \Big(\prod_{j=1}^{n+1} \frac{d\bar{z}\alpha_j^{(j)}}{2\pi}\Big)_r \Big(\prod_{j=1}^{n} i\lambda dz_{\beta j}^{(j)}\Big)_r$$

where α_j, β_j stand for the four internal indices at the j^{th} "time" division.

References

1) N.D.Birrel and P.C.W.Davies, Quantum Fields in Curved Space (Cambridge Univ.Press, Cambridge, 1982).

2) J.B.Hartle and S.W.Hawking, Phys.Rev. D13, 2188 (1976).

3) See for example I.H.Duru and N.Ünal, Phys.Rev. D34, 959 (1986).

4) R.P.Feynman and A.R.Hibbs, Quantum Mechanics and Path Integrals, Mc Graw-Hill, New York, 1965.

5) See for example V.N.Popov, Cern preprint, TH-2424-CERN

6) D.Peak and A.Inomata, J.Math.Phys. 10, 1422 (1969).

7) A justification of (9) can be found in I.H.Duru, Phys.Lett 119A, 163 (1986).

8) I.H.Duru, Phys.Rev. D28, 2689 (1983).

9) I.H.Duru, Phys.Lett. 112A, 421 (1985).

10) I.H.Duru and H.Kleinert, Phys.Lett. 84B, 185 (1979) ; and Forschr. Phys. 30, 401 (1982).

11) N.Ja.Vilenkin, Special Functions and the Theory of Group Representation, American Mathematical Society, Providence, Rhode Island, 1968 ; and J.D.Talman, Special Functions: A Group Theoretic Approach, Benjamin, New York, 1968.

12) I.H.Duru, Phys.Rev. D30, 2121 (1984).

13) C.Grosche, Annals of Phys. A38, 5906 (1988).

14) I.H.Duru and E.Keyman, Trieste preprint IC/80/129j and A.O.Barut and I.H.Duru, Phys.Rev A38, 5906 (1988).

15) See for example R.Rajaraman, Phys.Rep. 21C, 227 (1975).

16) A.Munier, J.R.Burgan, M.Feix and E.Fijalkow, J.Math, Phys. 22, 1219 (1980).

17) D.C.Khandekar and S.V.Lawande, Phys.Rep. C137, 115 (1986).

18) I.H.Duru, J.Phys.A:Math.Gen. 22, 4827 (1989).

19) A.O.Barut and I.H.Duru, Phys.Rep. 172C, 2 (1989).

20) A.O.Barut and N.Zanghi, Phys.Rev.Lett. 52, 2009 (1984).

CONSTRUCTION OF SOLUTIONS OF NONLINEAR D'ALEMBERT, MAXWELL, YANG-MILLS EQUATIONS BY SOLUTIONS OF NONLINEAR SPINOR EQUATIONS

Wilhelm FUSHCHICH

Institute of Mathematics, Academy of Sciences of the Ukraine, Repin Street 3, Kiev, USSR

ABSTRACT.

The connection between solutions of non-linear d'Alembert, Maxwell, Yang-Mills and Dirac equations is established. Multi-parameter families of exact solutions for these equations are adduced.

Many of us believe that all particles (fields) of spin $S = 0, 1, 3/2, 2, \ldots$ can be constructed by means of particles of spin 1/2. The main problem arising hence is how to realize this idea constructively if only in principle. It had been suggested in the well-known Heisenberg's program for creation of unified field theory to do this basing on quantization of the non-linear Dirac equation.

In this report I am going to speak about one approach to realization of the mentioned idea on the classical level. So we shall consider non-quantized non-linear wave equations. Our principal conclusion is the following: if we have exact solutions of non-linear spinor equations then by means of these solutions it is possible to construct solutions of the non-linear d'Alembert, Maxwell and Yang-Mills equations.

We are going to present below some results in this field which had been obtained in the Institute of Mathematics in Kiev by R.Zhdanov, W.Shtelen and myself in [1-6].

Let us consider the non-linear spinor equation

$$\gamma_\mu p^\mu \Psi + F(x, \Psi^*, \Psi)\Psi = 0. \qquad (1)$$

where $\Psi = \Psi(x) \equiv (\Psi_0, \Psi_1, \Psi_2, \Psi_3)$ is a four-component spinor, $x = (x_0 \equiv t, x_1, x_2, x_3) \in R(1,3)$, the four-dimensional Minkowski space, Ψ^* is the complex conjugate spinor, the γ_μ's are 4 x 4 Dirac matrices, F is an arbitrary smooth function,

$$p_0 = i\frac{\partial}{\partial x_0}, \quad p_j = -i\frac{\partial}{\partial x_j}, \quad \mu = \overline{0,3}, \quad j = 1,2,3. \qquad (2)$$

The plan of my report is the following:

1. The Poincaré and scale symmetry of the equation (1).
2. The Ansätze for the equation

$$\{\gamma_\mu p^\mu + \lambda(\bar\Psi\Psi)^{1/2k}\}\Psi = 0. \qquad (3)$$

3. The reduction of the spinor equations. The exact solutions.
4. Solutions of the d'Alembert, Maxwell, Yang-Mills equations.

1. The Symmetry Of The Equation (1).

Theorem 1. Eq.(1) is invariant with respect to the extended Poincaré group $\tilde P(1,3)$ that is the Poincaré group $P(1,3)$ supplemented with the group of scale transformations

$$x_\mu \to x'_\mu = x_\mu \exp(\theta),$$
$$\Psi \to \Psi'(x') = \Psi(x)\exp(k\theta),$$

where k and θ are arbitrary parameters, iff

$$F = F_1(z) + F_2(z)\gamma_4 + F_3(z)\gamma_\mu \bar\Psi \gamma_4 \gamma_\mu \bar\Psi + F_4(z) S^{\mu\nu}\bar\Psi \gamma_4 S_{\mu\nu}\Psi, \qquad (1.1)$$

where $z = \{\bar\Psi\Psi / \bar\Psi\gamma_4\Psi\}$, $\gamma_4 \equiv \gamma_0\gamma_1\gamma_2\gamma_3$,

$$S_{\mu\nu} = \frac{i}{4}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu),$$

$$F_i = (\bar\Psi\Psi)^{-1/2k} \tilde F_i(z), \quad i = 1,2,$$

$$F_{i+2} = (\bar\Psi\Psi)^{-\frac{(1+2k)}{2k}} \tilde F_{i+2}(z), \qquad (1.2)$$

where $\tilde F_1, \dots, \tilde F_4$ are arbitrary functions on z.

Theorem 2. Eq.(1) is invariant with respect to the conformal algebra $AC(1,3)$ with the basis operators

$$P_0 = i\frac{\partial}{\partial x_0}, \quad P_a = -i\frac{\partial}{\partial x_a},$$

$$J_{\mu\nu} = x_\mu P_\nu - x_\nu P_\mu + S_{\mu\nu}, \quad \mathcal{D} = x_\nu p^\nu + ik, \qquad (1.3)$$

$$K_\mu = 2x_\mu \mathcal{D} - x^2 P_\mu + 2S_{\mu\nu}x^\nu$$

iff

$$F_i = (\bar\Psi\Psi)^{1/3} \tilde F_i(z), \quad i = 1,2,$$

$$F_{i+2} = (\bar\Psi\Psi)^{-2/3} \tilde F_{i+2}(z). \qquad (1.4)$$

Note 1. In the case when $\tilde F_i(z) = 1$, $\tilde F_{i+2} = 0$ Eq.(1) coincides with the Dirac-Gürsey equation

$$\{\gamma_\mu p^\mu + \lambda(\bar\Psi\Psi)^{1/3}\}\Psi = 0, \qquad (1.5)$$

where λ is an arbitrary parameter.

Note 2. One of the simplest conformally-invariant equations which does not coincide with Eq.(1.5) is

$$\{\gamma_\mu p^\mu + \lambda\gamma_\mu(\bar\Psi\gamma_\mu\Psi)[(\bar\Psi\gamma_\alpha\Psi)(\bar\Psi\gamma^\alpha\Psi)]^{-1/3}\}\Psi = 0.$$

The adduced theorems can be proved with the classical S.Lie's method.

2. The Ansätze For The Spinor Equation (3).

We look for solutions of Eq.(3) in the form [1-3]

$$\Psi(x) = \mathcal{A}(x)\varphi(w), \qquad (2.1)$$

where $\mathcal{A}(x)$ is a 4 x 4 matrix, $\varphi(w)$ is a 4-component column function depending on three new independent variables $w = (w_1, w_2, w_3)$.

Definition. We say that the expression (2.1) is an ansatz for the equation (3) if the matrix $\mathcal{A}(x)$ and the variable w are defined in such a way that after substitution of (2.1) into (3) we get an equation for vector function $\varphi(w)$ depending on the variables w_1, w_2, w_3 only.

The ansatz (2.1) reduces the four-dimensional equation to a set of three-dimensional equations. This idea is used as the base of a method for construction of exact solutions for non-linear spinor equations. Thus in order to reduce Eq.(3) it is sufficient to describe the matrices $\mathcal{A}(x)$ and the variables $w = (w_1, w_2, w_3)$ with which (2.1) is an ansatz for this equation.

We find matrices $\mathcal{A}(x)$ and variables $w = (w_1, w_2, w_3)$ from the following system of first-order equations

$$(a^{\mu\nu}J_{\mu\nu} + b^\mu P_\mu + c\mathcal{D})\mathcal{A}(x) = 0,$$

$$(a^{\mu\nu}M_{\mu\nu} + b^\mu P_\mu + c\mathcal{D})w_\ell(x) = 0, \qquad (2.2)$$

$$\ell = 1,2,3, \quad \mu = \overline{0,3}, \quad M_{\mu\nu} = J_{\mu\nu} - S_{\mu\nu} = x_\mu p_\nu - x_\nu p_\mu.$$

In the case when the coefficients $a^{\mu\nu}$, b^μ, c do not depend on x, Ψ, $\partial\Psi/\partial x_\mu$ this system can be solved. E.g. if all coefficients except $a^{03} = 1$, $b_3 = b_2 = 1$

vanish then the equations for $A(x)$ and $w(x)$ take the form

$$(x_0\rho_3 - x_3\rho_0 + \tfrac{i}{2}\gamma_0\gamma_3)A(x) = 0,$$
$$\rho_1 A(x) = \rho_2 A(x) = 0, \qquad (2.3)$$
$$(x_0\rho_3 - x_3\rho_0)w(x) = 0, \quad \rho_1 w = \rho_2 w = 0.$$

This system has the following solution:

$$A(x) = \exp\{\tfrac{1}{2}\gamma_0\gamma_3 \ln(x_0+x_3)\}, \quad w(x) = x_0^2 - x_3^2. \qquad (2.4)$$

The ansatz (2.4) reduces the four-dimensional equation (3) to a system of ordinary differential equations.

Let us present one more solution of the system (2.2)

$$A(x) = (x_0 - x_2)^{-k}\exp\{\tfrac{1}{2a}\gamma_1(\gamma_2-\gamma_0)\ln(x_0 - x_2)\},$$
$$w_1 = (x_0^2 - x_1^2 - x_2^2)x_3^{-2}, \quad w_2 = (x_0 - x_2)x_3^{-2}, \qquad (2.5)$$
$$w_3 = a x_1 (x_0 - x_2)^{-1} - \ln(x_0 - x_2), \quad a \neq 0.$$

The ansatz (2.1), (2.5) reduces Eq.(3) to a three-dimensional system of differential equations for functions depending on the variables w_1, w_2, w_3. The large list of the Ansätze of the form (2.1) is given in [3,4].

One of the simplest Ansätze for the conformally-invariant Dirac-Gürsey equation takes the form [2,3]

$$\Psi = \gamma_\mu x^\mu f^2(x)\varphi(w),$$
$$\rho_\alpha\rho^\alpha f(x) = 0, \quad f(x) = (x_\alpha x^\alpha)^{-1}, \qquad (2.6)$$
$$w = \beta_\mu x^\mu (x_\alpha x^\alpha)^{-1}, \quad x_\alpha x^\alpha \neq 0, \ \beta_\mu\beta^\mu \neq 0,$$

where Ψ is a four-component vector-column.

3. The Reduced Equations.

Let us adduce some simplest partial differential equations obtained by means of the ansatz of the form (2.1):

$$ik\gamma_0\varphi - (\gamma_a - \gamma_0 w_a)\frac{\partial\varphi}{\partial w_a} = \lambda(\bar\varphi\varphi)^{1/2k}\varphi, \qquad (3.1)$$

$a = 1,2,3$, $A(x) = x_0^{-k}E$, $w_1 = x_1x_0^{-1}$, $w_2 = x_2x_0^{-1}$, $w_3 = x_3x_0^{-1}$.

where E is the unit matrix,

$$(\gamma_0 + \gamma_3)\frac{\partial\varphi}{\partial w_1} + \gamma_1\frac{\partial\varphi}{\partial w_2} + \gamma_2\frac{\partial\varphi}{\partial w_3} = \lambda(\bar\varphi\varphi)^{1/2k}\varphi, \qquad (3.2)$$

$A(x) = E$, $w_1 = x_0 + x_3$, $w_2 = x_1$, $w_3 = x_2$.

As a rule the reduced equations have essentially lower symmetry than the initial equations. However there is an exception. For example Eq.(3.2) is invariant under an infinite-dimensional algebra.

The simplest ordinary differential equations (ODE) obtained by reduction of partial differential equations (PDE) are of the form

$$i\gamma_2\frac{d\varphi}{dw} - \lambda(\bar\varphi\varphi)^{1/2k}\varphi = 0, \qquad \varphi = \begin{pmatrix}\varphi_0\\ \varphi_1\\ \varphi_2\\ \varphi_3\end{pmatrix}, \qquad (3.3)$$

$$i[w(\gamma_0+\gamma_3) + \gamma_0 - \gamma_3]\frac{d\varphi}{dw} + i(\gamma_0+\gamma_3)\varphi + \lambda(\bar\varphi\varphi)^{1/2k}\varphi = 0, \qquad (3.4)$$

$$2iw^{1/2}\gamma_2\frac{d\varphi}{dw} + \frac{i}{2}w^{-1/2}\varphi + \lambda(\bar\varphi\varphi)^{1/2k}\varphi = 0. \qquad (3.5)$$

The conformally invariant ansatz (2.6), (2.7) leads to the equation

$$\frac{d\varphi}{dw} = i\lambda(\beta_\alpha\beta^\alpha)^{-1}(\bar\varphi\varphi)^{1/3}(\gamma_\alpha\beta^\alpha)\varphi. \qquad (3.6)$$

Having solved the systems of ODE of the type (3.3) - - (3.6) we constructed families of exact solutions for the spinor equation (3). Let us adduce the simplest of these solutions.

$$\psi_1 = \frac{\gamma_0 x_0 - \gamma_1 x_1 - \gamma_2 x_2}{(x_0^2 - x_1^2 - x_2^2)^{3/2}} \exp\{i\lambda(\bar{\chi}\chi)^{1/2}\gamma_0 x_0 (x_0^2 - x_1^2 - x_2^2)^{-1}\}\chi, \quad (3.7)$$

where $k=1$, χ is a constant spinor.

$$\psi_2 = \gamma_a x_a (x_c x_c)^{-3/2} \exp\{-i\lambda(\bar{\chi}\chi)^{1/2}\gamma_1 x_1 (x_a x_a)^{-1/2}\}\chi, \quad (3.8)$$

$$\psi_3 = [(x_1+\phi_1)^2 + (x_2+\phi_2)^2]^{-1/4} \exp\{[-\tfrac{1}{2}(\dot{\phi}_1\gamma_1 + \dot{\phi}_2\gamma_2 + \phi_3\gamma_4](\gamma_0+\gamma_3)\} \exp\{-\tfrac{1}{2}\gamma_1\gamma_2 \operatorname{arctg}\frac{x_1+\phi_1}{x_2+\phi_2}\} \quad (3.9)$$

$$\times \exp\left\{\frac{2i\lambda k}{1-2k}(\bar{\chi}\chi)^{1/2k}\gamma_2[(x_1+\phi_1)^2 + (x_2+\phi_2)^2]^{\frac{2k-1}{4k}}\right\}\chi$$

where ϕ_1, ϕ_2, ϕ_3 are arbitrary functions on $x_0 + x_3$.

The conformally invariant ansatz (2.6) yields the solution

$$\psi_4 = \gamma_\alpha x^\alpha (x_\nu x^\nu)^{-2} \exp\{i\lambda k(\gamma_\alpha \beta^\alpha)\omega\}\chi, \quad (3.10)$$

$$\omega = \beta_\mu x^\mu (x_\alpha x^\alpha)^{-1}.$$

Using the property of invariance of Eq.(3) with respect to the extended Poincaré group $\tilde{P}(1,3)$ and the conformal group $C(1,3)$ (when $k=3/2$) it is possible by the adduced solutions to construct the multi-parameter families of exact solutions. If ψ_I is a solution of a conformally invariant spinor equation then a new solution ψ_{II} can be constructed by means of the formula

$$\psi_{II}(x) = \sigma^{-2}(x)[1 - (\gamma_\alpha x^\alpha)(\gamma_\nu \theta^\nu)]\psi_I(x'),$$

$$x'_\mu = [x_\mu - \theta_\mu (x_\alpha x^\alpha)]\sigma^{-1}(x), \quad \alpha,\mu = 0,1,2,3, \quad (3.11)$$

$$\sigma(x) = 1 - 2\theta_\alpha x^\alpha + (\theta_\alpha\theta^\alpha)(x_\nu x^\nu).$$

The conformally invariant family of solutions obtained with the formula (3.11) from the solutions (3.10) is of the form

$$\psi(x) = \sigma^{-1}(x)[1 - (\gamma_\alpha x^\alpha)\gamma_\nu\beta^\nu] \exp\{\frac{\theta}{2}(\gamma\cdot a)(\gamma\cdot b)$$

$$\times [b\cdot x - (b\cdot\theta)(x\cdot x)]\sigma^{-1}(x)\}\} \exp\{-\tfrac{i}{2}\lambda(\bar{\chi}\chi)^{1/3}$$

$$\times (\gamma\cdot a)[2(a\cdot x) - (a\cdot\theta)(x\cdot x)]\sigma(x) + \theta((b\cdot x) \quad (3.12)$$

$$- (b\cdot\theta)x^2)^2]\sigma^{-2}(x)\}\chi),$$

where parameters a_μ, b_μ, θ_μ are constants, $a\cdot b \equiv a_\alpha b^\alpha$, $\gamma\cdot b \equiv \gamma_\alpha b^\alpha$, $x^2 \equiv x_\alpha x^\alpha$, $\theta = (\theta_0^2 - \theta_1^2 - \theta_2^2 - \theta_3^2)^{1/2}$.

The multi-parameter family (3.12) is invariant with respect to the group $C(1,3)$ that is this family posesses the same symmetry as the spinor equation itself. Such non-generable families of solutions can be used to solve the problem of quantization of non-linear equations. Instead of quantization of non-linear equation it is possible to quantize a family of exact solutions of this equation. The quantization of exact solutions of non-linear equations can be carried out by means of the standard method e.g. using the Fourier integrals.

4. Solutions Of The D'Alembert, Yang-Mills And Maxwell Equations.

Let us construct a scalar from spinor solutions of Eq.(3) according to the formula

$$u = \bar{\psi}\psi \exp\{i\theta(x)\}, \quad (4.1)$$

where $\theta(x)$ is a phase. We shall consider the non-linear d'Alembert equation

$$\rho_\alpha \rho^\alpha u = \varkappa |u|^{2\tau} u, \tag{4.2}$$

where $\varkappa$ and τ are arbitrary real numbers.

The solution (3.7) of the Dirac equation (3) generates according to the formula (4.1) the following solution of the non-linear d'Alembert equation (4.2)($\tau = 1/2$):

$$u = const\,(x_0^2 - x_1^2 - x_2^2)^{-2}. \tag{4.3}$$

The solutions (3.8) of the Dirac equation yield the solutions of the equation (4.2) ($\tau = 1/2$) of the form

$$u = const\,(x_1^2 + x_2^2 + x_3^2)^{-2}. \tag{4.4}$$

The solution (3.9) of the Dirac equation yields the solution of the equation (4.2) ($\tau = 2$) of the form

$$u = const\,[(x_1 + \phi_1)^2 + (x_2 + \phi_2)^2]^{-1/2} \exp\{i\phi_0\}, \tag{4.5}$$

where ϕ_0, ϕ_1, ϕ_2 are arbitrary smooth functions on $x_0 + x_3$.

Thus the functions (4.3), (4.4), (4.5) constructed of solutions of the spinor equation (3) are solutions of the non-linear d'Alembert equation (4.2).

The Yang-Mills equation (the SU(2) equation) has the form

$$\Box \vec{Y}_\mu - \partial_\mu \partial^\nu \vec{Y}_\nu + e[(\partial_\nu \vec{Y}_\mu) \times \vec{Y}_\mu - 2(\partial^\nu \vec{Y}_\mu \times \vec{Y}_\nu + (\partial_\mu \vec{Y}_\nu \times \vec{Y}^\nu)] + e\vec{Y}^\nu \times \vec{Y}_\nu \times \vec{Y}_\mu) = 0, \tag{4.6}$$

where $\vec{Y}_\mu = \{Y_\mu^1, Y_\mu^2, Y_\mu^3\}$, $\mu, \nu = \overline{0, 3}$, $x \in R(4)$. Using the 't Hooft-Corrigan-Fairlie-Wilczek ansatz

$$eY_0^a = \mp \partial_a \ln u, \quad a = 1, 2, 3, \tag{4.7}$$

$$eY_j^a = (\varepsilon_{jan}\partial_n \pm \delta_{ja}\partial_0) \ln u.$$

Eq.(4.6) reduces to a conformally invariant scalar equation

$$\Box u + \lambda_1 u^3 = 0. \tag{4.8}$$

Knowing solutions of the scalar equation (4.8) by means of the formulas (4.7) we can construct exact solutions for the Yang-Mills equation.

Let us consider the conformally invariant Dirac-Gürsey equation

$$\{\gamma_\mu \rho^\mu + \frac{3}{2}(\lambda_1)^{1/2}(\bar\psi\psi)^{1/3}\}\psi = 0. \tag{4.9}$$

We shall construct a scalar field of spinors by means of the formula [6]

$$u = (\bar\psi\psi)^{1/3}. \tag{4.10}$$

The solution of the spinor equation (4.9) obtained in [7]

$$\psi = \frac{1}{4}(2\lambda_1)^{-3/4} \frac{i\gamma\cdot x + (x^2)^{1/2}}{(x^2)^{5/4}} \chi \tag{4.11}$$

generates the one-meron solution for the Yang-Mills equation

$$eY_0^a = \pm \frac{x_a}{x^2}, \quad eY_j^a = -\varepsilon_{jan}\frac{x_n}{x^2} \mp \delta_{aj}\frac{x_0}{x^2}. \tag{4.12}$$

The solutions of Eq.(4.9) [6]

$$\psi = \frac{1}{4}(2\lambda_1)^{-3/4}\left\{\frac{(a-b)^2}{(x-a)^2(x-b)^2}\right\}^{3/4}\left\{i\frac{\gamma\cdot x - \gamma\cdot a}{\sqrt{(x-a)^2}} + \left[1 - \frac{(\gamma\cdot x - \gamma\cdot a)(\gamma b - \gamma\cdot a)}{(a-b)^2}\right]\left(\frac{(a-b)^2}{(x-b)^2}\right)^{1/2}\right\}\chi \tag{4.13}$$

generates the double-meron solutions for the Yang-Mills equation:

$$e\,Y_0^n = \pm\left\{\frac{(x-a)_n}{(x-a)^2} + \frac{(x-b)_n}{(x-b)^2}\right\},$$

$$e\,Y_j^n = -\varepsilon_{jln}\left\{\frac{(x-a)_l}{(x-a)^2} + \frac{(x-b)_l}{(x-b)^2}\right\} \mp \delta_{nj}\left\{\frac{(x-a)_0}{(x-a)^2} + \frac{(x-b)_0}{(x-b)^2}\right\}. \tag{4.14}$$

The solutions of Eq.(4.9) [8]

$$\psi = (8/3\cdot\lambda_1)^{-3/4}\,\frac{i\gamma_\alpha x^\alpha + c}{(x^2+c^2)^2}\,\chi, \tag{4.15}$$

where c is an arbitrary constant, generates the instanton solution for the Yang-Mills equation

$$e\,Y_0^n = \pm\,\frac{2x_n}{x^2+c^2},$$

$$e\,Y_j^n = -\varepsilon_{jnl}\,\frac{2x_l}{x^2+c^2} \pm \delta_{nj}\,\frac{2x_0}{x^2+c^2}. \tag{4.16}$$

The solutions for electromagnetic field in terms of A_μ, $F_{\mu\nu}$ are constructed of solutions of the spinor Dirac-Gürsey equations acoording to the formulas

$$A_\mu = \bar\psi\gamma_\mu\psi, \tag{4.17}$$

$$F_{\mu\nu} = \frac{i}{2}\,\bar\psi(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)\psi. \tag{4.18}$$

Having used the conformally invariant ansatz for the spinor equation (1.5)

$$\psi = \frac{\gamma\cdot x}{(x_\alpha x^\alpha)^2}\,\varphi(w), \qquad w = \frac{B_\alpha x^\alpha}{x_\nu x^\nu} \tag{4.19}$$

we get Ansätze for conformally invariant electromagnetic field

$$A_\mu(x) = \frac{B_\mu(w)}{(x_\nu x^\nu)^k} - 2x_\mu\,\frac{x^\alpha B_\alpha(w)}{(x_\nu x^\nu)^{k+1}}, \tag{4.20}$$

$$F_{\mu\nu}(x) = \frac{f_{\mu\nu}(w)}{(x_\nu x^\nu)^k} - 2\,\frac{(x_\mu f_{\delta\nu} - x_\nu f_{\delta\mu})x^\delta}{(x_\alpha x^\alpha)^{k+1}}, \tag{4.21}$$

where B_μ, $f_{\mu\nu}$ are arbitrary functions on the invariant w.

Thus we have shown that it is possible to construct according to the formulas (4.1), (4.10), (4.20), (4.21) solutions for the d'Alembert, Yang-Mills, Maxwell equations of solutions of the spinor equation (3).

REFERENCES

1. Fushchich, W."On Symmetry and Exact Solutions of Some Many-Dimensional Equations of Mathematical Physics."- In the book: Theoretical-Algebraic Methods in Problems of Mathematical Physics, 4-23 (Kiev, 1983).
2. Fushchich, W. and Shtelen, W."On Some Exact Solutions of the Non-Linear Dirac Equation". J.Phys.A. 16, 271-277 (1983).
3. Fushchich, W. and Zhdanov, R. "Symmetry and Exact Solutions of Non-Linear Spinor Equations". Phys.Rep. 172, 124-174 (1989).
4. Fushchich, W., Shtelen, W. and Serov, N. "Symmetry Analysis and Exact Solutions of Non-Linear Equations of Mathematical Physics" (Kiev, Naukova Dumka, 1989 - Russian; Dordrecht, Kluwer, 1991 -English).
5. Fushchich, W."The Symmetries and Exact Solutions of Multidimensional Nonlinear Spinor Equations." Proc. of the Annual Seminar of the Canadien Math. Soc. on Lie Theory, Differential Equations and Representation

Theory. Editor V.Hussin (Montreal, 1990), 161-170.

6. Fushchich, W. and Shtelen, W."Merons and instantons as products of self-interaction of the Dirac-Gürsey spinor field". J.Phys.A. 23 , L517-520 (1990).

7. Kortel, F. "On some solutions of Gürsey's conformal-invariant spinor wave equations". Nuovo Cimento 4, 211-215 (1956).

8. Merwe P. "Classical solutions of the conformally invariant Gürsey equation". Phys.Lett. B. 106, 485-486 (1981).

On Schwinger Terms in (3+1) Dimensions

Edwin Langmann*
Institut für Theoretische Physik, Technische Universität
Graz, Austria

Abstract

Schwinger terms arise in current algebras due to regularisations required for a consistent construction of the currents. In (1+1)-dimensions one has to normal order, and the resulting Schwinger term is the well-known Kac-Peterson cocycle. In higher dimensions, an additional wave function renormalisation is necessary leading to operator valued Schwinger terms. A rigorous, non-perturbative construction of such Schwinger terms was given by Mickelsson and Rajeev [*Commun. Math. Phys.* **116**, 365 (1988)] in terms of determinant bundles over infinite dimensional Grassmannians.

We present an alternative construction of this Schwinger term by means of quasi-free second quantization of fermions. First, we review this formalism and the construction of current algebras in (1+1)-dimensions within this framework: gauge transformations correspond to unitarily implementable Bogoliubov transformations (BTS), and the currents can be obtained from the implementers of these BTS. It is argued that in higher dimensions, gauge transformations give rise to BTS which are not unitarily implementable. We propose an implementation of such BTS by quadratic forms which allows us to obtain current algebras in (3+1)-dimensions and the Mickelsson-Rajeev Schwinger term in a simple and natural way.

1 Introduction

We consider Dirac (or Weyl) fermions in $(d+1)$-dimensions, coupled to a static, external Yang-Mills field A with a gauge group G (e.g. a compact, semi-simple Lie group) and $\underline{g}$ the Lie algebra of G. Space M^d is assumed to be a compact, riemannian

*supported in parts by "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich"

spin manifold with a given spin structure, coordinates x^i, and $\mathrm{d} = \partial_i \mathrm{d}x^i$ the standard exterior derivative. The Yang-Mills field is a $\underline{g}$-valued 1-form on M^d: $A = A_i \mathrm{d}x^i$. On the one-particle level, we are given a Hilbert space $h = L^2(M^d) \otimes V$ of square integrable functions on M^d valued in a finite-dimensionsal complex vector space V (V is the tensor product of the space of spinors on M^d with the representation space for G and $\underline{g}$) and the time evolution is generated by the usual Dirac (or Weyl) Hamiltonian H_A which is a self-adjoint operator on h. Gauge transformations are given by smooth maps $U : M^d \to G; x \mapsto U(x)$,

$$H_A \mapsto U H_A U^{-1} = H_{A^U} \tag{1}$$

with $A^U = UAU^{-1} + U\mathrm{d}U^{-1}$, and the set $Map(M^d; G)$ of all gauge transformations is naturally a Lie group of unitary operators on h. Its (real) Lie algebra is the set $Map(M^d; \underline{g})$ of bounded, self-adjoint operators corresponding to generators of gauge transformations, i.e. smooth maps $X : M^d \to \underline{g}; x \mapsto X(x)$. Note that the group multiplication (the Lie bracket $[\cdot, \cdot]$) defined pointwise coincides with the operator product (the commutator) in h.

From this one-particle formalism, a multi-particle formalism can be constructed by means of second quantization. Currents correspond to the generators $J(X)$ of one-parameter families $U(t) = \mathrm{e}^{\mathrm{i}tX}$ $[X \in Map(M^d; \underline{g})]$ of gauge transformations on this second quantized level. In the "naive" quantization scheme corresponding to the free (Fock-Cook) representation of the fermion field algebra, these currents obey

$$[J(X), J(Y)] = J([X, Y]), \tag{2}$$

i.e. they give a representation of the Lie algebra $Map(M^d; \underline{g})$. However, this is not the representation of interest in quantum field theory. There is no ground state in this representation, and the multi-particle Hamiltonian corresponding to H_A is not bounded from below.

The physically relevant multi-particle formalism can be obtained by means of non-trivial representations of the fermion field algebra corresponding to "filling up the Dirac sea". It is known that the construction of the currents $\hat{J}(\cdot) = \hat{J}(\cdot; A)$ within such representations requires some renormalisation. Due to this, the commutator relations of the currents are modified to

$$[\hat{J}(X), \hat{J}(Y)] = \hat{J}([X, Y]) + c(X, Y; A) \tag{3}$$

where $c(X, Y; A)$ is the Schwinger term.

This situation is well understood in (1+1)-dimensions [1,2]: The currents can be renormalized by means of normal ordering, and the resulting Schwinger term is the Kac-Peterson cocycle which can be chosen independent of the external field A: $c(X, Y; A) = c(X, Y)$. The currents $\hat{J}(X)$ provide a unitary Hilbert space representation of a central extension of $Map(M^1; \underline{g})$ (which is essentially an affine Kac-Moody algebra). In fact, there are two different (rigorous) approaches to current algebras

in (1+1)-dimensions: One in terms of determinant bundles over infinite dimensional Grassmannians [1], and another one by means of quasi-free second quantization of fermions [2] which is closer to the quantum field theory.

The situation is more difficult in higher dimensions: normal ordering of the currents is not sufficient, and one has to perform an additional wave function renormalisation [3]. A generalisation of the Grassmannian-approach [1] to current algebras from (1+1)- to higher dimensions was developped by Mickelsson and Rajeev [3] (see also Ref. [4]). They obtained currents $\hat{J}(X)$ giving a representation of an abelian extension of $Map(M^1; \underline{g})$, and a Schwinger term $c(X, Y; A)$ depending on the external field A in an essential way. Moreover, they tried — without success — to find a Hilbert space $\mathcal{H}$ such that the currents are generators of one-parameter families of unitary operators on $\mathcal{H}$. Later on it was shown by Pickrell [5] that this is impossible: In higher dimensions, the current algebras of Mickelson and Rajeev do not allow for a faithful, unitary representation on a separable Hilbert space.

In this contribution we present an alternative approach to the Mickelsson-Rajeev current algebras by means of quasi-free second quantization of fermions [2]. In our opinion, this approach is much more tranparent (at least from a physicist's point of view), and it allows for a simple interpretation of "how Hilbert space is lost" in higher dimensions: In (1+1)-dimensions, the physically relevant represention of the field algebra corresponding to different external fields A are all unitarily equivalent, hence all gauge transformations can be unitarily implemented. This is not the case in higher dimensions. The idea is to extend the Hilbert space to a bundle of Hilbert spaces corresponding to the inequivalent representations of the fermion field algebras associated with different external fields A. The gauge transformations can be implemented on this bundle in a natural way, and this gives rise to the Mickelsson-Rajeev cocycle.

The plan of the contribution is as follows: In the next section we review the formalism of quasi-free second quantization of fermions and indicate why this formalism is sufficient for constructing current algebras in (1+1)- but not in higher dimensions. The extension of this formalism adequate for (3+1)-dimensions is presented in sect.3, and sect.4 contains a few comments.

2 Quasi-free Second Quantization

Fermion Field Algebras and Fock Spaces

In the spirit of the algebraic approach to quantum field theory [6], we may start with the fermion field algebra $\mathcal{A}(h)$ over h which is the C^*-algebra generated by operators $a^*(f)$, $f \in h$, such that the mapping $f \mapsto a^*(f)$ is linear and the following canonical anticommutator relations are fulfilled:

$$\{a(f), a^*(g)\} = (f, g)$$

$$\{a(f), a(g)\} = 0 \quad \forall f, g \in h \tag{4}$$

[$a(f) \equiv (a^*(f))^*$; $(\cdot,\cdot)$ denotes the scalar product in h and $\{\cdot,\cdot\}$ the anticommutator]. The free representation Π_1 of this algebra is given on the Fock space $\mathcal{F}(h)$ over h in terms of creation and annihilation operators $\breve{a}^{(*)}(f) \equiv \Pi_1(a^{(*)}(f))$ obeying

$$\breve{a}(f)\Omega = 0 \quad \forall f \in h \tag{5}$$

where Ω is the vacuum in $\mathcal{F}(h)$, and $\breve{a}^*(f) = \breve{a}(f)^*$ with $*$ denoting the Hilbert space adjoint [7]. For all unitary operators U on h, a unique unitary operator $\Gamma(U)$ on $\mathcal{F}(h)$ exists such that

$$\Gamma(U)\breve{a}(f)\Gamma(U)^* = \breve{a}(Uf) \quad \forall f \in h \tag{6}$$

and $\Gamma(U)\Gamma(V) = \Gamma(UV)$ for all unitary operators U and V on h [7]. Hence $\Gamma(\cdot)$ provides a unitary representation of the Lie group $Map(M^d; G)$. The currents obeying (2) are then given by $J(X) = (\mathrm{d}/\mathrm{d}(\mathrm{i}t))\Gamma(\mathrm{e}^{\mathrm{i}tX})\,|_{t=0}$. However, the representations Π_1 and $J(\cdot)$ are not satisfactory from a physical point of view: there is no ground state due to the fact that the multi-particle Hamiltonian $\mathrm{i}(\mathrm{d}/\mathrm{d}t)\Gamma(\mathrm{e}^{-\mathrm{i}tH_A})\,|_{t=0}$ is not bounded from below. [Mathematical speaking, the representation $J(\cdot)$ of $Map(M^d; \underline{g})$ is no heighest weight representation.] This deficiency can be remedied by introducing a quasi-free representation of the field algebra $\mathcal{A}(h)$. This gives a precise mathematical meaning to the idea of "filling up the Dirac sea".

In the following, we denote as $B_p(h)$ ($p \in \mathbb{N}$) the set of all bounded operators A on h obeying

$$\|A\|_p \equiv \left(\mathrm{tr}(A^*A)^{p/2}\right)^{1/p} < \infty \tag{7}$$

[$\mathrm{tr}(\cdot)$ is the trace in h]. Especially $B_1(h)$ and $B_2(h)$ are the trace-class and Hilbert-Schmidt operators on h, respectively. Note that $B_p(h)$ is an ideal (*Schatten ideal*) in the algebra of bounded operators on h, $B_1(h) \subseteq B_2(h) \subseteq B_3(h) \subseteq \cdots$, and that for $A \in B_p(h)$, $B \in B_q(h)$, the operator AB is in $B_r(h)$ with $\frac{1}{r} \leq \frac{1}{p} + \frac{1}{q}$ [3].

"Filling up the Dirac Sea"

For H_A the one-particle Hamiltonian, we introduce the grading operator $F = \mathrm{sign}(H_A)$ on h [with $\mathrm{sign}(x) = 1$ (-1) for $x \geq 0$ (< 0)]. Then $P_\pm = \frac{1}{2}(1 \pm F)$ are orthogonal projections onto the subspaces of positive/negative energy states in h. The *quasi-free representation* Π_F of $\mathcal{A}(h)$ on $\mathcal{F}(h)$ with $\Pi_F(a^{(*)}(\cdot)) \equiv \hat{a}^{(*)}(\cdot; F)$ is constructed such that

$$\hat{a}(f_+; F)\Omega = \hat{a}^*(f_-; F)\Omega = 0 \quad \forall f_\pm \in P_\pm h,$$

i.e. that the Fock vacuum Ω corresponds to the filled Dirac sea. It follows that

$$\hat{a}(f; F) \equiv \breve{a}(P_+ f) + \breve{a}^*(\overline{P_- f}) \quad \forall f \in h \tag{8}$$

(the bar denotes complex conjugation commuting with P_-) and $\hat{a}^*(f; F) = \hat{a}(f; F)^*$.

A unitary operator U on h is called *unitarily implementable in* Π_F if there is a unitary operator $\hat{\Gamma}(U; F)$ of $\mathcal{F}(h)$ such that

$$\hat{a}(Uf; F) = \hat{\Gamma}(U; F)\hat{a}(f; F)\hat{\Gamma}(U; F)^* \quad \forall f \in h, \tag{9}$$

and the well-known necessary and sufficient condition for this to be the case is $[U, F] \in B_2(h)$ [9]. Note that UV and U^{-1} are unitarily implementable in Π_F if U and V are.

One can easily see that this construction makes sense for *any* grading operator F on h. However, for $F = \mathrm{sign}(H_A)$, time evolution $\mathrm{e}^{-\mathrm{i}H_A t}$ ($t \in \mathbb{R}$) is unitarily implementable in Π_F, and it can be shown that the multi-particle Hamiltonian $\mathrm{i}(\mathrm{d}/\mathrm{d}t)\hat{\Gamma}(\mathrm{e}^{-\mathrm{i}H_A t}; F)\,|_{t=0}$ corresponding to H_A is bounded from below. Hence this is the "right" representation of the field algebra to use, and it is determined (up to unitary equivalence — see below) by the given one-particle Hamiltonian.

Let $\varepsilon = \mathrm{sign}(H_0)$ (no external field). Obviously, any operator $U = \mathrm{e}^{\mathrm{i}X}$ for X a bounded, self-adjoint operator on h obeying $[X, \varepsilon] \in B_2(h)$ is unitarily implementable in Π_ε. We denote the real Lie group [Lie algebra] of all such operators U [X] as $\boldsymbol{G}_2(h; \varepsilon)$ [$\boldsymbol{g}_2(h; \varepsilon)$]. $\boldsymbol{G}_2(h; \varepsilon)$ and $\boldsymbol{g}_2(h; \varepsilon)$ are normed algebras with the norm $\|\cdot\|_2$,

$$\|A\|_2 = \|P_+^0 A P_+^0\| + \|P_-^0 A P_-^0\| + \frac{1}{2}\|[A, \varepsilon]\|_2 \tag{10}$$

[$P_\pm^0 = \frac{1}{2}(1 \pm \varepsilon)$], and $\boldsymbol{g}_2(h; \varepsilon)$ is the Lie algebra of $\boldsymbol{G}_2(h; \varepsilon)$.

We introduce the set $Gr_2(h; \varepsilon)$ of all grading operators F on h which are of the form $U^*\varepsilon U$ for some $U \in \boldsymbol{G}_2(h; \varepsilon)$. Obviously all quasi-free representations Π_F with $F \in Gr_2(h; \varepsilon)$ are *unitarily equivalent* to Π_ε, i.e. there is a unitary operator $\mathcal{U}(F, \varepsilon)$ in $\mathcal{F}(h)$ such that

$$\hat{a}(f; F) = \mathcal{U}(F, \varepsilon)\hat{a}(f; \varepsilon)\mathcal{U}(F, \varepsilon)^* \quad \forall f \in h. \tag{11}$$

Moreover, $F \mapsto F^U \equiv U^*FU$ provides a representation of $\boldsymbol{G}_2(h; \varepsilon)$ in $Gr_2(h; \varepsilon)$, and the sets $\boldsymbol{G}_2(h; F)$ [resp. $\boldsymbol{g}_2(h; F)$] for $F \in Gr_2(h; \varepsilon)$ all coincide.

It is easy to see from (9) that for $F \in Gr_2(h; \varepsilon)$, $\hat{\Gamma}(\cdot; F)$ provides a projective representation of $\boldsymbol{G}_2(h; \varepsilon)$ with the operator product as group multiplication. However, there is another group multiplication — $\star$ — also making $\hat{\Gamma}(\cdot; F)$ a projective representation of $\boldsymbol{G}_2(h; \varepsilon)$: From (8) one can deduce

$$\hat{a}(Uf; F) = \Gamma(U)\hat{a}(f; F^U)\Gamma(U)^* \quad \forall f \in h \tag{12}$$

with $F^U = U^*FU$. Hence

$$\hat{a}(UVf; F) = \Gamma(U)\hat{a}(Vf; F^U)\Gamma(U)^* = \Gamma(U)\hat{\Gamma}(V; F^U)\Gamma(U)^*\hat{\Gamma}(U; F) \times \hat{a}(f; F)\hat{\Gamma}(U; F)^*\Gamma(U)\hat{\Gamma}(V; F^U)^*\Gamma(U)^*,$$

and

$$\hat{\Gamma}(U;F)\star\hat{\Gamma}(V;F)\equiv\Gamma(U)\hat{\Gamma}(V;F^U)\Gamma(U)^*\hat{\Gamma}(U;F)$$
$$=\hat{\chi}(U,V;F)\hat{\Gamma}(UV;F)\quad\forall F\in Gr_2(h;\varepsilon),\ U,V\in\boldsymbol{G}_2(h;\varepsilon)\tag{13}$$

follows with $\hat{\chi}(\cdots)$ some phase factor. One can easily see that $\star$ is associative, which is equivalent to

$$\hat{\chi}(U,V;F)\hat{\chi}(UV,W;F)=\hat{\chi}(V,W;F^U)\hat{\chi}(U,VW;F)\tag{14}$$

for all $U,V,W\in\boldsymbol{G}_2(h;\varepsilon)$, $F\in Gr_2(h;\varepsilon)$. As mentioned above, $\hat{\Gamma}(U;F)$ is determined by (9) only up to a phase factor. We may redefine it:

$$\hat{\Gamma}(U;F)\to\beta(U;F)\hat{\Gamma}(U;F)\tag{15}$$

with β some smooth function $\boldsymbol{G}_2(h;F)\times Gr_2(h;\varepsilon)\to S^1\equiv\{e^{i\varphi}|\varphi\in\mathbb{R}\}$. Then

$$\hat{\chi}(U,V;F)\to\delta\beta(U,V;F)\hat{\chi}(U,V;F)\tag{16}$$

with

$$\delta\beta(U,V;F)=\beta(U;F)\beta(V;F^U)/\beta(UV;F)\tag{17}$$

trivially obeying (14). From a mathematical point of view, eq. (14) is a *cocycle relation*, any function $\hat{\chi}$ satisfying it is a *cocycle*, and a cocycle of the form $\delta\beta$ (17) is a *coboundary of the group* $\boldsymbol{G}_2(h;F)$; a cocycle is non-trivial if it is no coboundary [8].

The explicit form of the operators $\hat{\Gamma}(U;F)$ was worked out by Ruijsenaars [9]. From his results one can deduce that the cocycle $\hat{\chi}$ is non-trivial and can be choosen F-independent [with an appropriate phase convention for the $\hat{\Gamma}(U;F)$ — see the next section]:

$$\hat{\chi}(U,V;F)=\left(\frac{\det([P_-^0V^*U^*P_-^0]^{-1}P_-^0V^*P_-^0U^*P_-^0)}{\det(P_-^0UP_-^0VP_-^0[P_-^0UVP_-^0]^{-1})}\right)^{1/2}\tag{18}$$

($[\cdots]^{-1}$ (inverse) and $\det(\cdots)$ (determinant) are understood as operatons in P_-^0h).

The currents can be defined as

$$\hat{J}(X;F)\equiv\frac{\mathrm{d}}{\mathrm{d}(it)}\hat{\Gamma}(e^{itX};F)\mid_{t=0},\tag{19}$$

and from

$$[\hat{J}(X;F),\hat{J}(Y;F)]=\frac{\mathrm{d}}{\mathrm{d}(it)}\frac{\mathrm{d}}{\mathrm{d}(is)}\hat{\Gamma}(e^{itX};F)\star\hat{\Gamma}(e^{isY};F)\star\hat{\Gamma}(e^{-itX};F)\star\hat{\Gamma}(e^{-isY};F)\mid_{s=t=0}\tag{20}$$

we obtain [note that $F=F(A)$]

$$[\hat{J}(X;F),\hat{J}(Y;F)]=\hat{J}([X,Y];F)+c(X,Y;F)\tag{21}$$

with the Schwinger term given by

$$c(X,Y;F)=\frac{\mathrm{d}}{\mathrm{d}(it)}\frac{\mathrm{d}}{\mathrm{d}(is)}\hat{\chi}(e^{itX}e^{isY};F)\hat{\chi}(e^{-itX}e^{-isY};F)\hat{\chi}(e^{itX}e^{isY}e^{-itX}e^{-isY};F)\mid_{s=t=0}.$$

Using (18), a simple calculation gives

$$c(X,Y;F)=\frac{1}{4}\mathrm{tr}([X,\varepsilon][Y,\varepsilon]\varepsilon)\tag{22}$$

which is obviously finite for all $X,Y\in\boldsymbol{g}_2(h;\varepsilon)$. This is the Kac-Peterson cocycle [2].

We remark that for operators $A\in\boldsymbol{g}_2(h;\varepsilon)$ of the form $A=\sum_{n\in\mathbb{N}}\lambda_nf_n(f_n,\cdot)$ [λ_n real numbers and $\{f_n\}_{n\in\mathbb{N}}$ an orthonormal system in h], the currents can be represented as $\hat{J}(A;\varepsilon)=:\sum_{n\in\mathbb{N}}\lambda_n\hat{a}^*(f_n;\varepsilon)\hat{a}(f_n;\varepsilon):$ where $:\cdots:$ denotes normal ordering [2].

Current Algebras in $(d+1)$-Dimensions

For simplicity we restrict ourselves to Yang-Mills fields corresponding to pure gauge transformations: $A=U\mathrm{d}U^{-1}$ with $U=e^{iX}$, $X\in Map(M^d;\underline{g})$. Let $\varepsilon=\mathrm{sign}(H_0)$ and $F=\mathrm{sign}(H_A)$. Then the operators $\hat{J}(\cdot;F)$ provide a representation of $Map(M^d;\underline{g})$ for every A if $Map(M^d;\underline{g})$ is a Lie subalgebra of $\boldsymbol{g}_2(h;\varepsilon)$. However, this is true only for $d=1$ *but not for* $d>1$: In fact, one can prove that $Map(M^d;\underline{g})$ is a Lie subalgebra of $\boldsymbol{g}_{2p}(h;\varepsilon)$ with $2p>d$, $p\in\mathbb{N}$, where $\boldsymbol{g}_{2p}(h;\varepsilon)$ is the normed Lie algebra of all bounded, self-adjoint operators X on h obeying $[X,\varepsilon]\in B_{2p}(h)$ with the norm $\|\cdot\|_{2p}$,

$$\|A\|_{2p}=\|P_+^0AP_+^0\|+\|P_-^0AP_-^0\|+\frac{1}{2}\|[A,\varepsilon]\|_{2p}\tag{23}$$

[$P_\pm^0=\frac{1}{2}(1\pm\varepsilon)$] [3]. The Lie group corresponding to $\boldsymbol{g}_{2p}(h;\varepsilon)$ is the normed Lie group $\boldsymbol{G}_{2p}(h;\varepsilon)$ of unitary operators U on h obeying $[U,\varepsilon]\in B_{2p}(h)$ with norm $\|\cdot\|_{2p}$. Moreover, we introduce the set $Gr_{2p}(h;\varepsilon)$ of all grading operators of the form $U^*\varepsilon U$ with $U\in\boldsymbol{G}_{2p}(h;\varepsilon)$ carrying a representation $F\mapsto F^U=U^*FU$ of $\boldsymbol{G}_{2p}(h;\varepsilon)$. In order to construct current algebras in $(d+1)$-dimensions, we must construct a representation of the Lie group $\boldsymbol{G}_{2p}(h;\varepsilon)$. This will be done in the next section for $p=2$ corresponding to $(3+1)$-dimensions ($d=3$).

3 Wave-function Renormalisation

We recall some basic facts about infinite dimensional and regularized determinants [3]: For A a linear operator on a Hilbert space h, the determinant $\det(1+A)$ exists if and only if $A\in B_1(h)$, and $\det((1+A)(1+B))=\det(1+A)\det(1+B)$ for all $A,B\in B_1(h)$. For $A\in B_2(h)$, $\det(1+A)$ and $\mathrm{tr}(A)$ both diverge; however,

$B_1(h)$ is dense in $B_2(h)$, and for any sequence $\{A_n\}$ in $B_1(h)$ converging to A in the $\|\cdot\|_2$-norm, the sequence $\det(1+A_n)\mathrm{e}^{-\mathrm{tr}(A_n)} \equiv \det_2(1+A_n)$ has a well-definined limit depending only on A — the *regularized determinant* $\det_2(1+A)$. Note that in this case $\det_2((1+A)(1+B)) \neq \det_2(1+A)\det_2(1+B)$.

The explicite form of the implementer of the Bogoliubov transformation (9) is [9]

$$\hat{\Gamma}(U;F) = N(U;F)E(U;F) \tag{24}$$

with $E(\cdots)$ an operator obeying $< \Omega, E(U;F)\Omega >= 1$ [$< \cdot,\cdot >$ is the inner product in the Fock space $\mathcal{F}(h)$] and

$$N(U;F) = \left(\det(P_-U^*P_-UP_-)\frac{\det(P_-^0U^*P_-^0P_-UP_-)}{\det(P_-U^*P_-P_-^0UP_-^0)}\right)^{1/2} \tag{25}$$

some normalisation constant. Moreover, one can show that

$$E(U;F)\star E(V;F) = \mathcal{E}(U,V;F)E(UV;F) \tag{26}$$

with the non-zero complex number

$$\mathcal{E}(U,V;F) = \det(P_-UVP_-[P_-UP_-U^*P_-UVU^*P_-U]^{-1}). \tag{27}$$

[We remark that our phase convention for the implementers differs from the one used in [9]; our choice gives the F-independent cocycle $\hat{\chi}$ (18).] One can show explicitely that the determinants in (25) and (27) all exist for $U,V \in \boldsymbol{G}_2(h;\varepsilon)$, i.e. that they are of the form $\det(A)$ with $(A-1) \in B_1(P_-^0h)$. However, for $U,V \in \boldsymbol{G}_4(h;\varepsilon)$ these operators A obey $(A-1) \in B_2(P_-^0h)$ and the numbers in (25) and (27) diverge.

The essential point is that the operator $E(U;F)$ exists not only for $U \in \boldsymbol{G}_2(h;\varepsilon)$, but it can be defined in the sense of quadratic forms whenever $[U,\varepsilon]$ is a compact operator [this observation is due to Ruijsenaars [10] (p345 *f*; see also his recent paper [11])] and especially for $U \in \boldsymbol{G}_4(h;\varepsilon)$. Hence *the non-existence of the implementers* $\hat{\Gamma}(U;F)$ *and* $\hat{\Gamma}(V;F)$ *and their product* $\hat{\Gamma}(U;F)\star\hat{\Gamma}(V;F)$ *for* $U,V \in \boldsymbol{G}_4(h;F)$ *is only due to the non-existence of the normalisation constants (25) and (27).* As $\boldsymbol{G}_2(h;\varepsilon)$ is dense in $\boldsymbol{G}_4(h;\varepsilon)$, we may try to *regularize* the implementers: $\hat{\Gamma}(U;F) \to \hat{\Gamma}_4(U;F)$, and the group product: $\star \to \hat{\star}$ in (9) by *replacing the determinants* $\det(\cdots)$ *in (25) and (27) by regularized determinants* $\det_2(\cdots)$. Hence

$$\hat{\Gamma}_4(U;F) \equiv N_4(U;F)E(U;F) = \beta_4^N(U;F)\hat{\Gamma}(U;F) \tag{28}$$

with $\beta_4^N \equiv N_4/N$,

$$N_4(U;F) = \left(\det_2(P_-U^*P_-UP_-)\frac{\det_2(P_-^0U^*P_-^0P_-UP_-)}{\det_2(P_-U^*P_-P_-^0UP_-^0)}\right)^{1/2}, \tag{29}$$

exists as quadratic form in Fock space $\mathcal{F}(h)$ for all $U \in \boldsymbol{G}_4(h;\varepsilon)$ and $F \in Gr_4(h;\varepsilon)$, i.e. the expressions

$$< \psi, \hat{\Gamma}_4(U;F)\varphi > \tag{30}$$

have a well-defined meaning for ψ and φ in a certain dense subset of $\mathcal{F}(h)$. However, in the case of $\mathcal{E}$ (27) we can not simply replace $\det(\cdots)$ by $\det_2(\cdots)$ as the resulting regularized product $\hat{\star}$ has to be associative. In fact, we must find a function $\beta_4^\star$: $\boldsymbol{G}_2(h;\varepsilon) \times Gr_2(h;\varepsilon) \to \mathbb{C}^\times$ such that

$$\mathcal{E}_4(U,V;F) = \delta\beta_4^\star(U,V;F)\mathcal{E}(U,V;F) = r_4(U,V;F)\det_2(P_-UVP_-[P_-UP_-U^*P_-UVU^*P_-U]^{-1}) \tag{31}$$

with $r_4(U,V;F)$ a complex number finite and non-zero for all $U,V \in \boldsymbol{G}_4(h;\varepsilon)$ and $F \in Gr_4(h;\varepsilon)$ [δ cf. (17)].

One can show by a straightforward calculation that for

$$\beta_4(U;F) = \exp\frac{1}{4}\mathrm{tr}(P_-^0FP_+^0UP_-^0U^*P_-^0 - P_-^0UP_-^0U^*P_+^0FP_-^0) \tag{32}$$

the regularized cocycle

$$\hat{\chi}_4(U,V;F) = \delta\beta_4(U,V;F)\hat{\chi}(U,V;F) \tag{33}$$

is finite for all $U,V \in \boldsymbol{G}_4(h;\varepsilon)$ and $F \in Gr_4(h;\varepsilon)$. Hence $\beta_4^\star = \beta_4/\beta_4^N$. Then

$$\hat{\Gamma}_4(U;F)\hat{\star}\hat{\Gamma}_4(V;F) = \hat{\chi}_4(U,V;F)\hat{\Gamma}_4(UV;F) \tag{34}$$

results with $\hat{\star}$ associative as $\hat{\chi}_4$ (33) obviously satisfies the cocycle relation (14).

We note that for $U,V \in \boldsymbol{G}_2(h;\varepsilon)$, both factors on the r.h.s. of (33) are finite, and $\hat{\chi}$ and $\hat{\chi}_4$ are equal up to a coboundary; however, for $U,V \in \boldsymbol{G}_4(h;\varepsilon)$, both factors on the r.h.s. of (33) diverge (do not exist) and only their product is finite. Moreover, one can easily show that $\hat{\chi}_4$ is a non-trivial cocycle for $\boldsymbol{G}_4(h;\varepsilon)$, i.e. that one can not choose a normalisation of the implementers and the group multiplication which gets rid of this cocycle.

From this, a representation of $\boldsymbol{g}_4(h;\varepsilon)$ by currents

$$\hat{J}_4(X;F) = \frac{\mathrm{d}}{\mathrm{d}(it)}\hat{\Gamma}_4(\mathrm{e}^{itX};F)\mid_{t=0} \tag{35}$$

can be obtained as in eq. (20): With

$$b_4(X;F) \equiv \frac{\mathrm{d}}{\mathrm{d}(it)}\beta_4(\mathrm{e}^{itX};F)\mid_{t=0} = -\frac{1}{16}\mathrm{tr}([X,\varepsilon][F,\varepsilon]) \tag{36}$$

and simple calculations, we obtain relations similar to (21) , with $\hat{J}$ replaced by $\hat{J}_4$ and $c(X,Y;F)$ by

$$c_4(X,Y;F) = c(X;Y;F) - \delta b_4(U,V;F) = \frac{1}{8}\mathrm{tr}([[X,\varepsilon],[Y,\varepsilon]](\varepsilon - F)) \tag{37}$$

where $\delta b_4(X,Y;F) = b_4([X,Y];F) + \delta_Y b_4(X;F) - \delta_X b_4(Y;F)$ and $\delta_X b_4(Y;F) = (\mathrm{d}/\mathrm{d}(it))b_4(Y;\mathrm{e}^{itX}F\mathrm{e}^{-itX})\mid_{t=0}$. This is the Mickelsson-Rajeev cocycle [3] [it is easily

seen that — in contrast to (22) — it is finite for all $X, Y \in \boldsymbol{g}_4(h;\varepsilon)$, $F \in Gr_4(h;\varepsilon)$]. Note that the Lie bracket $[\cdot,\cdot]$ is no longer just the commutator but some regularized version thereof. Moreover, the Lie bracket of a current with the Schwinger term is not zero

$$[\hat{J}_4(X;F), c_4(Y,Z;F)] = \delta_X c(Y,Z;F)$$

for all $X, Y, Z \in \boldsymbol{g}_4(h;\varepsilon)$. (For a discussion of this Schwinger term and its relation to Faddeev's operator anomaly of the Gauss law [8] we refer to Ref. [3].)

4 Some Comments

As mentioned above, the implementers $\hat{\Gamma}(\cdot;F)$ for each $F \in Gr_2(h;\varepsilon)$ provide a unitary representation of a central extension of the Lie group $\boldsymbol{G}_2(h;\varepsilon)$ on the Hilbert space $\mathcal{F}(h)$ with the operator product as group multiplication. For an interpretation of the group multiplication $\star$ (13), it is more natural to distinguish the Hilbert spaces $\mathcal{H}_F$ [each $\mathcal{H}_F$ a copy of $\mathcal{F}(h)$] corresponding to different quasi-free representations Π_F [$F \in Gr_2(h;F)$], and to consider the implementer $\hat{\Gamma}(U;F)$ as map from $\mathcal{H}_F$ to $\mathcal{H}_{FU}$ [cf. also eq. (12)]. From this point of view, $\hat{\Gamma}$ is a representation of a central extension of $\boldsymbol{G}_2(h;\varepsilon)$ on the fiber bundle with $Gr_2(h;\varepsilon)$ as base space and fibers $\mathcal{F}(h)$. However, this bundle is trival in the sense that all quasi-free representations Π_F for $F \in Gr_2(h;\varepsilon)$ are unitarily equivalent [this corresponds to the fact that the cocycle $\hat{\chi}$ (18) is F-independent], and

$$\hat{\Gamma}(U;F) \star \hat{\Gamma}(V;F) = \eta(U,V;F)\hat{\Gamma}(U;F)\hat{\Gamma}(V;F)$$

with $\eta : \boldsymbol{G}_2(h;\varepsilon) \times \boldsymbol{G}_2(h;\varepsilon) \times Gr_2(h;\varepsilon) \to S^1$ some function relating the two group multiplications.

$\hat{\Gamma}_4$ and eq. (34) can be interpreted as representation of an abelian extension of the Lie group $\boldsymbol{G}_4(h;\varepsilon)$ on a similar fiber bundle with base space $Gr_4(h;\varepsilon)$, but this Hilbert space bundle is non-trival as the quasi-free representations Π_F for $F \in Gr_4(h;\varepsilon)$ are *not unitarily equivalent*. Hence it is natural to think of $\hat{\Gamma}_4(U;F)$ as intertwiner of the unitarily inequivalent quasi-free representations Π_F and Π_{FU}, and to interpret the expression (30) as proportional to a transition amplitude between the vectors $\varphi \in \mathcal{H}_F$ and $\psi \in \mathcal{H}_{FU}$. Similarly the currents $\hat{J}_4$ represent an abelian extension of the Lie algebra $Map(M^3;\underline{g}) \subseteq \boldsymbol{g}_4(h;\varepsilon)$ on this non-trivial Hilbert space bundle.

In our opinion, our construction gives a precise, non-perturbative meaning to "wave-function renormalisation" in the context of external field problems. One can speculate that also in more general (non-linear) field theory models, renormalisation in Hilbert space might be equivalent to a well-definied calculus on some Hilbert space bundle.

In this contribution we restricted ourselves to $(3+1)$-dimensions; the generalisation of our considerations to arbitrary dimensions is straightforward [the only problem is to find the functions β_{2p} generalising β_4 (32)].

A more detailed version of this work will be published elsewhere [12].

Acknowledgement

I would like to thank A. Carey and H. Grosse for comments.

References

[1] Pressly A., Segal G.: *'Loop Groups'*, Oxford Math. Monographs, Oxford (1986)

[2] Carey A. L., Ruijsenaars S. N. M.: On fermion gauge groups, current algebras and Kac-Moody algebras. *Acta Appl. Mat.* **10**, 1 (1987)

[3] J. Mickelsson, and S. G. Rajeev, Current algebras in $d+1$- dimensions and determinant bundles ove infinite dimensional Grassmannians, *Commun. Math. Phys.* **116** 400 (1988).

[4] K. Fujii, and M. Tanaka, Universal Schwinger cocycle and current algebras in $(D+1)$- dimensions: geometry and physics, *Commun. Math. Phys.* **129**, 267 (1990).

[5] Pickrell D., On the Mickelsson-Faddeev extension and unitary representations, *Commun. Math. Phys.* **123**, 617 (1989).

[6] Haag R., and Kastler D.: An algebraic approach to quantum field theory, *J. Math. Phys.* **7**, 848 (1964).

[7] Bratteli O., Robinson D. W.: *Operator Algebras and Quantum Statistical Mechanics. Vol.2.* Berlin, Springer 1981.

[8] Faddeev L. D.: Operator anomaly for the Gauss law, *Phys. Lett.* **145B**, 81 (1984).

[9] Ruijsenaars S. N. M.: On Bogoliubov transformations for systems of relativistic charged particles, *J. Math. Phys.* **18**, 517 (1977).

[10] Ruijsenaars S. N. M.: Integrable quantum field theories and Bogoliubov transformations, *Ann. Phys.* **132**, 328 (1981).

[11] Ruijsenaars S. N. M.: Index formulas for generalized Wiener-Hopf operators and boson-fermion correspondence in $2N$ dimensions, *Commun. Math. Phys.* **124**, 553 (1989).

[12] Langmann E., in preparation.

ANYONS AND CHERN-SIMONS THEORY: A REVIEW

G.W. Semenoff*

Department of Physics, University of British Columbia, Vancouver, British Columbia, Canada V6T 2A6.

0. Contents

I. Introduction

In this series of Lectures I shall review both some old and some recent ideas about the possibility that quantum mechanical particles in two spatial dimensions have fractional statistics and how such systems are described by a 2+1-dimensional gauge field theory with a Chern-Simons term.

The idea that topological effects in a quantum mechanical system could lead to particles which have unusual statistics has a long history associated with the works of Skyrme, Finkelstein, Rubenstein and Williams [Skyrme, 1961; Finkelstein, 1966; Finkelstein and Rubenstein, 1968; Williams, 1970]. Particularly important in this regard are the works of Morette-De Witt and Laidlaw [Morette-De Witt and Laidlaw, 1971] who who associated exotic quantum statistics with multi-connectedness of the phase space and Leinaas and Myrlheim [Leinaas and Myrlheim, 1977] and Goldin, Menikoff and Sharp [Goldin, Menikoff and Sharp, 1981] who applied these ideas to two-dimensional spaces.

Independently, it was realized that in 2+1 spacetime dimensions it is possible that gauge theories and even gravity could have a gauge invariant, local mass term, the Chern-Simons three-form [Seigel, 1979; Schoenfeld, 1981; Jackiw, Deser and Templeton, 1982]. It was Wilczek and Zee [Wilczek and Zee, 1983] who first realized that the Chern Simons term was related to fractional statistics. They argued that the charged particles in a 2+1-dimenisonal U(1) gauge theory with a Chern-Simons term behave like anyons. There were subsequently many works which attempted to establish this

* This work is supported in part by the Natural Sciences and Engineering Research Council of Canada.

connection both in quantum mechanics [Polyakov, 1988; Tze, 1988] and in a quantum field theory [Semenoff, 1988; Matsuyama, 1989; Semenoff and Sodano, 1989; Forte and Jolicoeur, 1990].

The main purpose of the following lectures is to review the present status of these attempts. Anyons are commonly described as charged quantum mechanical particles coupled to a 'ficticious gauge field' which has a flat gauge connection and gives the correct holonomy to anyon states under adiabatic transport [Arovas, Schrieffer, Wilczek and Zee, 1985]. The flat gauge connection can be seen to be a solution of a Chern-Simons theory with point-sources. It is tempting to conclude that a Chern-Simons theory describes this situation.

I shall first discuss a coupled Chern-Simons theory and non-relativistic Schroedinger field theory in the continuum. The idea is that we can map a theory of nonrelativistic matter fields interacting with a U(1) gauge field and where the gauge field kinetic term in the action is a Chern-Simons term onto an equivalent theory with a quadratic 'free' action for the matter fields and with the boundary conditions that the wavefunctions are multi-valued. The latter boundary condition requires that, in the basis of states where the particles are localized at points, the holonomy of the wavefunction represents the braid group of the plane.This is the analog of the boson-fermion mapping constructed in two dimensions by Coleman and Mandelstam [Coleman, 1975; Mandelstam, 1975]. The essential element of the construction is a generalization of the Jordan-Wigner transformation [Jordan and Wigner, 1928] from 1+1 to 2+1-dimensions.

In order to generalize these ideas to a relativistic quantum field theory we find it necessary to go to a lattice regularization of matter-coupled Chern-Simons theory [Fradkin, 1989; Ambjorn and Semenoff, 1989; Luscher, 1989; Muller, 1990; Eliezer, Semenoff and Wu, 1990]. We shall find that the arguments presented in the continuum can be carried out in a slightly modified form. Bosonization can be carried out exactly. However, for other statistics parameters, the interactions do not decouple exactly. When the statistics parameter is a rational number, M/N, where M and N are integers, the anyon Hamiltonian has a Z_N gauge field which makes it covariant under changing the multi-valued operators. There appears to be no choice of operators for which this gauge field decouples.

As for physical applications of these ideas, there is one interesting rigorous result. It is possible to show that for a lattice Fermion-Chern-Simons theory when the coupling constant of the Chern-Simons term is such that the anyons of the model are Bosons, the ground state of the theory has long-range order and forms a charged superfluid. If it were coupled to the electromagnetic field, at least for weak enough coupling, it would be a superconductor. That Bosons should condense at zero temperature would be no surprise in 3+1 dimensions and is a slightly less trivial effect in 2+1 dimensions because of the more severe infrared properties there. The latter prevent this condensation at any finite temperature. However, it is highly nontrivial that we can give a rigorous proof that the ground state of a fermion-gauge field problem in 2+1-dimensions has a superconducting ground state. We shall attempt a detailed discussion of the robustness of this result in Section V.

II. Fractional Statistics

The topological reasoning which allows a system of quantum mechanical particles on a 2-dimensional space to have fractional statistics is similar to that which, in the standard situation of a quantum mechanical system with a multiply connected configuration space, allows for multi-valued wavefunctions and projective representations of symmetry groups [Morette-De Witt and Laidlaw, 1971;Lienaas and Myrlheim, 1979; Schulman, 1981; Wu, 1984]. The classical configuration space of a gas of N particles on the d-dimensional Euclidean space R^d is R^{Nd}. The quantum mechanical configuration space is the space on which the probability density

$$\rho(q_1,\ldots,q_N;t) = \psi^\dagger(q_1,\ldots,q_N;t)\psi(q_1,\ldots,q_N;t) \tag{2.1}$$

is defined. If the particles are indistinguishable, ρ is a symmetric function of its arguments. Furthermore we shall consider the cases where either for dynamical or topological reasons particles cannot occupy the same point and the probability density is zero whenever two positions coincide ($q_i = q_j$ for any i,j).

Thus, to construct the quantum configuration space we first take the classical space R^{Nd} and excise the diagonal subspace $D = \{(\vec{q}_1,\ldots\vec{q}_N)\,|\vec{q}_i = \vec{q}_j \text{ for any } i,j\}$ to get $R^{Nd} - D$. We then take into account the symmetry of $\rho(q_1,\ldots,q_N;t)$ by factoring out the permutation group to get

$$\mathcal{C}_N = \left(R^{Nd} - D\right)/S_N \tag{2.2}$$

The probability density (2.1) must be a single-valued function on this space. However, the wavefunction which is a probability amplitude is allowed to be a multi-valued function.

We assume that the configuration space is path-connected, $\Pi_0(\mathcal{C}) = 0$. Multi-valued functions can be defined on a space $\mathcal{C}$ which is not simply connected, i.e. which has nontrivial fundamental group, $\Pi_1(\mathcal{C}) \neq 0$ and are single-valued functions on the unviersal cover of the configuration space, $\hat{\mathcal{C}}$.

The universal covering space $\hat{\mathcal{C}}$ of $\mathcal{C}$ is constructed as follows: Take a base-point, $\mathcal{B} \in \mathcal{C}$ and consider the set of all curves going from the base-point to all other points $x \in \mathcal{C}$. Then, for each point, we identify all curves going there from the base-point which are in the same homotopy class. (Homotopy is an equivalence relation for these curves.) We denote these equivalence classes as $C[\mathcal{B},x]$. For all $x \in \mathcal{C}$, the equivalence classes $C[\mathcal{B},x]$ form the universal cover $\hat{\mathcal{C}}$ of $\mathcal{C}$.

Obviously, if $\Pi_1(\mathcal{C})$ is trivial, all curves associated with a given point are in the same homotopy class and the universal cover coincides with the space itself, ($\hat{\mathcal{C}} = \mathcal{C}$). If $\Pi_1(\mathcal{C})$ is nontrivial the universal cover $\hat{\mathcal{C}}$ is a principal fiber bundle with base $\mathcal{C}$ and structure group $\Pi_1(\mathcal{C})$. The bundle projection takes a homotopy class of curves to their endpoint $\pi : C[\mathcal{B},x] \to x$.

Multi-valued wavefunctions are sections of an associated line bundle* where the structure group of the principle bundle, $\Pi_1(\mathcal{C})$ acts through a one-dimensional unitary representation. They are

* Here we shall confine our attention to one-dimensional representations of the fundamental group.

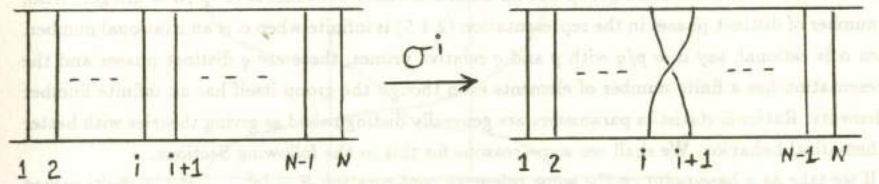
single-vlaued functions on $\hat{\mathcal{C}}$ and can therefore be though of as single-valued functions of homotopy classes of curves $C[\mathcal{B},x]$.

For $d \geq 3$, $\Pi_1(\mathcal{C}_N) = S_N$. This group has only two one-dimensional unitary representations, the trivial symmetric one and the anti-symmetric one. In the latter, a periodic motion on $\mathcal{C}_N$ is allowed to permute the points $q_1,\ldots,q_N$ and for an even permutation the wavefunction remains unchanged whereas for an odd permutation the wavefunction changes sign. The particles associated with these representations are Bosons and Fermions, respectively. The symmetric group also has higher dimensional representations with mixed symmetry which correspond to conventional parastatistical particles [Ohnuki and Kamefuchi, 1982].

A more interesting situation arises in $d = 2$ where

$$\Pi_1(\mathcal{C}_N) = B_N(R^2) \tag{2.3}$$

the N^{th} order braid group of the plane R^2 [Fox and Neuwirth, 1961; Fadell and Van Buskirk, 1962; Artin, 1947; Akutsu and Wadati, 1987; Akutsu, Deguchi and Wadati, 1987]. This is an infinite non-Abelian discrete group whose generators σ^i are the braiding operations for an ordered set of N adjacent lines:



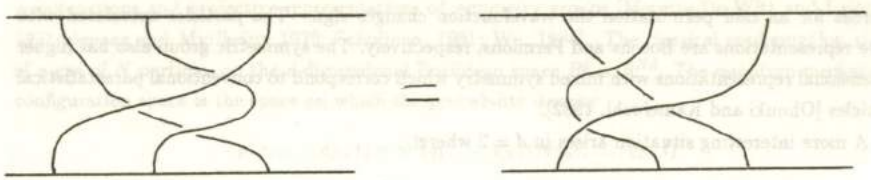
Consistency requires that the generators satisfy the structure relations

$$\sigma^i\sigma^j = \sigma^j\sigma^i \quad , \quad |i-j| > 1 \tag{2.4}$$

$$\sigma^i\sigma^{i+1}\sigma^i = \sigma^{i+1}\sigma^i\sigma^{i+1} \tag{2.5}$$

In general the wavefunction could be a section of a complex vector bundle which carries a higher dimensional representation of $\Pi_1(\mathcal{C})$. The resulting particles exhibit non-Abelian fractional statistics and are called non-Abelions.

The identity in (2.4) expresses the fact that braiding of separated lines commutes. Equation (2.5) is a result of the trivial equivalence of the braids



Braid group elements are products of the generators σ^i with an equivalence relation which identifies all products which can be obtained from each other using the relations (2.4) and (2.5).

Here, we shall be interested in the one-dimensional unitary representation of the braid group given by

$$\chi(\sigma^i) = e^{i\alpha\pi} \quad , \quad \chi((\sigma^i)^{-1}) = e^{-i\alpha\pi} \quad \forall i \tag{2.6}$$

Even for two particles the braid group has an infinite number of elements $(\sigma^1)^n, n =$ integer. Also, the number of distinct phases in the representation (2.4-5) is infinite when α is an irrational number. When α is rational, say $\alpha = p/q$ with p and q relative primes, there are q distinct phases and the representation has a finite number of elements even though the group itself has an infinite number of elements. Rational statistics parameters are generally distinguished as giving theories with better mathematical behavior. We shall see some reasons for this in the following Sections.

If we take as a base-point on $\mathcal{C}_N$ some reference configuration $B = (q_1^0, \ldots, q_N^0)$, a multi-valued function which carries the representation (2.6) is

$$\exp\left(-i\alpha\nu[C, q_1, \ldots, q_N]\right) = \exp\left(-i\alpha \int_0^1 ds \frac{d}{ds} \sum_{i<j} \theta(q_i(s) - q_j(s))\right)$$

$$= \exp\left(i\alpha \int_0^1 ds \sum_{i<j} \frac{(\dot{q}_i - \dot{q}_j) \times (q_i - q_j)}{(q_i - q_j)^2}\right) \tag{2.7}$$

where $\theta(q_i - q_j)$ is the angle between the vector $q_i - q_j$ and a reference direction* , $q_i(s=0) = q_i^0$

* This is the multi-valued function in R^2 which satisfies the equation

$$\nabla_i \theta(q) = -\epsilon_{ij} \nabla_j \ln|q| = -\epsilon_{ij} q_j / q^2$$

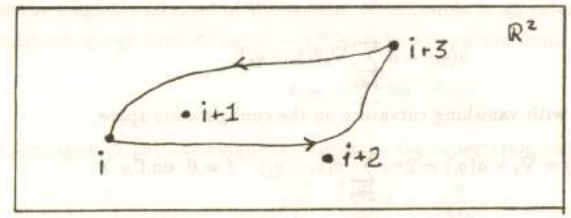
so that $\nabla^2 \theta(q) = 0$ and $\nabla \times \nabla \theta(q) = 2\pi\delta(q)$.

and $q_i(s=1) = q_i$ and $q_i(s) \neq q_j(s) \ \ \forall i, j$ and $s \in [0,1]$. Note that ν is a symmetric functional of $q_1(s), \ldots, q_N(s)$.

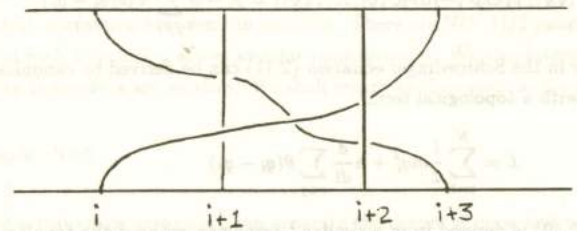
The curve-dependence of the multi-valued function is

$$\exp\left(-i\alpha\nu[C, q_1, \ldots, q_N]\right) = \exp\left(-i\alpha\nu[C - C', q_1, \ldots, q_N]\right) \exp\left(-i\alpha\nu[C', q_1, \ldots, q_N]\right) \tag{2.8}$$

In general the interpolation $C - C'$ could include a permutation of the points $q_1, \ldots, q_N$ (remember that all permutations of $q_1, \ldots, q_N$ are the same point on $\mathcal{C}_N$.) It is straightforward to verify that $\exp\left(-i\alpha\nu[C - C', q_1, \ldots, q_N]\right)$ is an element of the one-dimensional representation of the braid group (2.6). To see this, note that we can take a 'side view' of a periodic motion on $\mathcal{C}_N$ to get a braid. For example



coincides with the braid $(\sigma^{i+2})^{-1}\sigma^{i+1}\sigma^i\sigma^{i+1}\sigma^{i+2}$



In this case $\nu = 3\pi\alpha$. Reversing the orientation of the loop reverses the sign of ν. This is either done by reversing the orientation of the plane (parity) or reversing the sign of the loop parameter, i.e. tracing the loop in the opposite direction (time reversal). A parity or a time-reversal transformation are therefore equivalent to changing the sign of the statistics parameter α in the representation (2.6).

The multi-valued function (2.7) can be used as the generator of a canonical transformation which changes wavefunctions from single-valued functions to multi-valued functions. If $\psi(q_1, \ldots, q_N; t)$ is a symmetric single-valued function or an antisymmetric double-valued function on $\mathcal{C}_N$ then $\tilde{\psi}(q_1, \ldots, q_N; t)$ defined by

$$\tilde{\psi}_C(q_1, \ldots, q_N; t) = \exp\left(-i\alpha\nu[C; q_1, \ldots, q_N]\right) \psi(q_1, \ldots, q_N; t) \tag{2.9}$$

is a multi-valued function. Note that the probability measure (2.1) is invariant under this transformation and therefore remains single-valued.

If we consider a gas of 'free' anyons obeying the Schroedinger equation

$$i\frac{\partial}{\partial t}\tilde{\psi}_C(q_1,\ldots,q_N;t) = \sum_{i=1}^{N}\frac{p_i^2}{2m}\tilde{\psi}_C(q_1,\ldots,q_N;t) \tag{2.10}$$

(where $p_i \equiv -i\partial/\partial q_i$) with a multi-valued wavefunction, we could formulate the equivalent problem for a Bose gas with symmetric, single-valued wavefunctions or a Fermi gas with antisymmetric double-valued wavefunctions obeying the Schroedinger equation with a gauge field

$$i\frac{\partial}{\partial t}\psi(q_1,\ldots,q_N;t) = \sum_{i=1}^{N}\frac{1}{2m}\left(p_i - a(q_i)\right)^2\psi(q_1,\ldots,q_N;t) \tag{2.11}$$

Here

$$a(q_i) = \alpha\sum_{j\neq i}\nabla_i\theta(q_i-q_j) \tag{2.12}$$

is an external gauge field with vanishing curvature on the configuration space,

$$b(q_i) = \nabla_i \times a(q_i) = 2\pi\alpha\sum_{j\neq i}\delta(q_i-q_j) \quad (=0 \text{ on } \mathcal{C}_N) \tag{2.13}$$

We have used the identity

$$\exp\{i\alpha\nu[C,q_1,\ldots,q_N]\}\,p_i\exp\{-i\alpha\nu[C,q_1,\ldots,q_N]\} = p_i - \alpha\sum_{j\neq i}\nabla_i\theta(q_i-q_j) \tag{2.14}$$

The Hamiltonian operator in the Schroedinger equation (2.11) can be derived by canonical quantization of the Lagrangian with a topological term,

$$L = \sum_{i=1}^{N}\frac{1}{2}m\dot{q}_i^2 + \alpha\frac{d}{dt}\sum_{i<j}\theta(q_i-q_j) \tag{2.15}$$

and the Hamiltonian in (2.10) is derived from a similar Lagrangian without the topological term. Note that the topological term in (2.15) or, equivalently, the gauge field in (2.11) are not invariant under time reversal and parity transformations. Thus a theory of anyons of the type described here violates these discrete symmetries.

Since the gauge field (2.12) has vanishing Lie derivative,

$$\sum_j q_j \times \frac{\partial}{\partial q_j}a^m(q_i) + \epsilon^{mn}a^n(q_i) = 0 \tag{2.16}$$

The Lagrangian (2.15) is invariant under rotations of the coordinates. The Noether charge corresponding to this symmetry is the canonical angular momentum

$$j_c = \sum_i \delta q_i\frac{\delta L}{\delta\dot{q}_i} = \sum_i q_i \times p_i \tag{2.17}$$

The gauge invariant angular momentum operator

$$j = \sum_i q_i \times (p_i - a(q_i)) \tag{2.18}$$

is also conserved. Using the explicit form of the gauge field we obtain

$$j = \sum_i q_i \times p_i + \alpha\sum_{i\neq j}\frac{q_i\cdot(q_i-q_j)}{(q_i-q_j)^2} = \sum_i q_i\times p_i + \frac{\alpha}{2}N(N-1) \tag{2.19}$$

where in the last term, upon symmetrizing the summand, we find that it is independent of the coordinates. This fact was first discovered in a slightly different context by Hagen [Hagen, 1985].

The gauge invariant angular momentum j is the observable whose expectation value is proportional to the magnetic moment of the system. If we couple to an external magnetic field using the electromagnetic gauge field $A^m(q_i) = -\frac{1}{2}\epsilon^{mn}q_i^n B$ by adding the term

$$L_{\text{int.}} = \sum_i e\dot{q}_i\cdot A(q_i) \tag{2.20}$$

to the Lagrangian (2.15), the magnetic moment is the expectation value of the operator

$$\mu = -\frac{1}{\hbar}\frac{\partial}{\partial B}L = -\frac{e}{\hbar m}j \tag{2.21}$$

The first term on the right-hand-side of j in (2.19) generates canonical rotations of the coordinates in a wavefunction. The second term can be regarded as an induced spin carried by the statistical interaction between the particles. There are N(N-1)/2 gauge interactions between particles and each interaction stores angular momentum α. We also interpret this as meaning that the particles themselves are spinless. We shall return to this point in the next Section.

Example: N=2

For a two-anyon system we can separate the center of mass and relative coordinates,

$$R = \frac{1}{2}(q_1+q_2)\ ,\ \ q = q_1 - q_2 \tag{2.22}$$

In this case,

$$\mathcal{C}_2 = R^2\times(R^2-\{0\})/Z_2 \tag{2.23}$$

where we have excised the origin of the relative coordinates and identified $q\sim -q$. The eigenvalue problem for the Hamiltonian operator in (2.11) is

$$-\frac{1}{2m}\left(2\frac{\partial^2}{\partial Q^2} - \left(-i\frac{\partial}{\partial q} - 2a(q)\right)^2\right)\psi_E(Q,q) = E\psi_E(Q,q) \tag{2.24}$$

where

$$a^m(q) = -\alpha\epsilon^{mn}q^n/q^2 \tag{2.25}$$

In polar coordinates, $r = |q|$, $\phi = \arctan q_2/q_1$, where $a_r = 0$, $a_\phi = \alpha/r$ this equation reads

$$-\frac{1}{2m}\left(2\frac{\partial^2}{\partial Q^2}+\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r}+\frac{1}{r^2}\left(\frac{\partial}{\partial\phi}-i\alpha\right)^2\right)\psi_E(Q,r,\phi)=E\psi_E(Q,r,\phi) \tag{2.26}$$

The solutions of this equation should be single-valued, $\psi_E(Q,r,\phi+\pi)=\psi_E(Q,r,\phi)$ if the particles are Bosons and double-valued, $\psi_E(Q,r,\phi+\pi)=-\psi_E(Q,r,\phi)$ if the particles are Fermions. For Bosons the wavefunctions are

$$\psi_{Kkn}(Q,r,\phi)=\frac{e^{iKQ}}{2\pi}e^{2in\phi}j_{2|n+\alpha|}(kr) \tag{2.27}$$

where K is the wave-vector of the center of mass, k is the wave-vector for the relative motion, n is an integer and $j_{2|n+\alpha|}(kr)$ is the Bessel function of the first kind of order $2|n+\alpha|$. The spectrum is

$$E=\frac{K^2}{m}+\frac{k^2}{2m} \tag{2.28}$$

Near the origin, the Bessel function behaves as $J_{2|m+\alpha|}\sim r^{2|m+\alpha|}$, so that if the statistics parameter α is not an integer, there is no partial wave for which the wavefunction does not vanish at the origin. We can think of this behavior as due to a centrifugal barrier coming from the fractional shift in the angular momentum of the states, $n\to n+\alpha$. Alternatively, we could view it as enforcing the effective Pauli exclusion principle for an anyon gas. To see this, note that the multi-valued wavefunctions

$$\tilde\psi_{Kkn}(Q,r,\phi)=\frac{e^{iKQ}}{2\pi}e^{2i(n+\alpha)\phi}j_{2|n+\alpha|}(kr) \tag{2.29}$$

solve the eigenvalue problem

$$-\frac{1}{2m}\left(2\frac{\partial^2}{\partial Q^2}+\frac{\partial^2}{\partial q^2}\right)\tilde\psi_{Kkn}=E\tilde\psi_{Kkn} \tag{2.30}$$

with the same eigenvalues and carry the representation (2.5) of the braid group, $B_2(R^2)$,

$$\tilde\psi_{Kkn}(Q,r,\phi+\pi)=e^{i\alpha}\tilde\psi_{Kkn}(Q,r,\phi) \tag{2.31}$$

In (2.29) we should properly define ϕ as a line-integral, i.e. as $\nu[C,q_1,q_2]$ and $\tilde\psi$ should be labelled by the curve C. This notation is assumed to be implicit in (2.29) and (2.31).

Example: A Partially Solvable Many-Particle Model

When there are many particles it is convenient to parameterize R^2 by complex coordinates $z=x+iy$ and define $\partial_i\equiv\partial/\partial z_i$, $\bar\partial_i=\partial/\partial\bar z_i$. In this coordinate system the Hamiltonian which appears in the Schroedinger equation (2.11) is

$$-\frac{1}{m}\sum_i\left(D_i\bar D_i+\bar D_iD_i\right)$$

where $D_i=\partial_i-a(z_i)\equiv\partial_i-(a_2(q_i)+ia_1(q_i))/2$ and $\bar D_i=\bar\partial_i+\bar a(z_i)\equiv\bar\partial_i+(a_2(q_i)-ia_1(q_i))/2$. Instead of this Hamiltonian, we wish to consider the problem with Hamiltonian

$$h=-\frac{2}{m}\sum_i\bar D_iD_i=-\frac{1}{m}\sum_i\left(\bar D_iD_i+D_i\bar D_i\right)-\frac{1}{m}\sum_i\left(\bar D_iD_i-D_i\bar D_i\right)$$

$$=\sum_i-\frac{1}{m}\left(\bar D_iD_i+D_i\bar D_i\right)+\frac{1}{m}\sum_i\left(\partial_i\bar a(z^i)+\bar\partial_ia(z^i)\right)$$

$$=\sum_i-\frac{1}{m}\left(\bar D_iD_i+D_i\bar D_i\right)-\frac{2\pi\alpha}{m}\sum_{i\neq j}\delta^2(z_i-z_j) \tag{2.31'}$$

which differs from h by an interaction term which is the sum of magnetic fields at the sites of the particles* and constitutes a zero-range attractive force. In this model, h is the sum of operators of the form $\bar DD$. If we can find a state for which

$$D_i\psi=0\ ,\quad\forall i \tag{2.32}$$

then $h\psi=0$, i.e. ψ is the ground state and the ground state energy vanishes [Girvin, MacDonald, Fischer, Rey and Sethna, 1990]. As in the previous two-particle problem, the ground state here will be a continuum eigenstate at the minimum of the continuum spectrum and therefore it will not be normalizable. In order to make it normalizable we add an external magnetic field B. Then the gauge fields are

$$a(z_i)=-\frac{\alpha}{2}\sum_{j\neq i}\frac{1}{z_i-z_j}+\frac{1}{2}\bar z_iB\ ,\quad\bar a(z_i)=-\frac{\alpha}{2}\sum_{j\neq i}\frac{1}{\bar z_i-\bar z_j}+\frac{1}{2}z^iB \tag{2.33}$$

so that the interaction term in h is

$$h_{\rm int}=-\frac{2\pi\alpha}{m}\sum_{i\neq j}\delta^2(z^i-z^j)+NB \tag{2.34}$$

and the wavefunction satisfies

$$\left(\frac{\partial}{\partial z_i}+\frac{\alpha}{2}\sum_{j\neq i}\frac{1}{z_i-z_j}+\frac{1}{2}\bar z_iB\right)\psi_0=0\quad\forall i \tag{2.35}$$

This equation is solved by

$$\psi_0(z^1,\ldots,z^N;m_1,\ldots,m_N)=\frac{1}{\text{const.}}\prod_{i<j}|z^i-z^j|^\alpha\det|(\bar z^i)^{m_j}|e^{-\frac{B}{2}z\bar z} \tag{2.36}$$

where m_i are non-negative integers. Here, we have chosen the canonical statistics as Fermi statistics. This ground state is infinitely degenerate (actually any polynomial in $\bar z^i$ is allowed as a factor).

* This point interaction vanishes on the space where coincidence points $z+i=z_j$ are excluded.

We have chosen to resolve the degeneracy by also diagonalizing the canonical angular momentum operator j_c. Its eigenvalues are

$$j_c\psi_0(z^1,\ldots,z^N;m_1,\ldots,m_N)=\left(\sum_i m_i\right)\psi_0(z^1,\ldots,z^N;m_1,\ldots,m_N) \tag{2.37}$$

The state with minimal angular momentum is $\psi_0(z^1,\ldots,z^N;m_1=0,\ldots,m_N=N-1)$ with eigenvalue $-N(N-1)/2$. The gauge invariant angular momentum operator has eigenvalue $(\alpha-1)N(N-1)/2$. However, the total magnetic moment of this system is proportional to the derivative of the ground state energy by the external magnetic field B. This differs from the expectation value of the magnetic moment operator (2.21) because here the Hamiltonian contains nonminimal couplings to the magnetic field. Since the ground state energy vanishes for all values of B, the magnetic moment of the ground state of h given in (2.31)' is zero.

Note that we could make ψ_0 multi-valued and transform away the gauge field by the canonical transformation

$$\tilde{\psi}(z^1,\ldots,z^N)=\prod_{i<j}\left(\frac{z_i-z_j}{\bar{z}_i-\bar{z}_j}\right)^{\frac{q}{2}}\psi(z^1,\ldots,z^N) \tag{2.36}$$

This is the familiar Laughlin wavefunction used to describe the ground state of the fractionally quantized Hall effect system [Laughlin 1983]. The factor in the right-hand-side of (2.36) is multi-valued and should properly be defined by integration on a curve on $\mathcal{C}$, as is the function in (2.7).

Here we have solved the problem of a gas of anyons moving in a constant magnetic field when then have the particular Hamiltonian (2.31)'. Actually, it is possible to, at least formally, find the solution independently of the profile of the field configuration. This follows from an application of the Atiyah-Singer index theorem [Semenoff and Sodano, 1986]. It would be interesting to compare the correlation functions in the ground state with different profiles of the background magnetic field.

In the following Sections we shall examine a quantum field theoretical realization of these ideas.

III. Chern-Simons Theory and Fractional Statistics

In this Section we shall consider a second quantization of the many-particle problem which we discussed in Section II. The usual motivation for second quantization is to study the thermodynamic limit, $N\to\infty$. It is also useful in making perturbative and non-perturbative approximations. In the present case the second quantized theory which we obtain will be a theory of fermionic Schroedinger fields interacting with a U(1) Chern-Simons gauge theory. We shall find that for a fixed particle number the latter problem is completely equivalent to the many-particle quantum mechanics problem discussed in the previous Section. The bosonic analog of the classical theory has recently been found to have vortex solutions [Jackiw and Weinberg, 1990; Jackiw, Lee and Weinberg, 1990, Jackiw and Pi, 1990] which can be used, for example, for semiclassical studies anyonic matter.

We begin with operators which create and annihilate particles at a point q, $\psi^\dagger(q)$ and $\psi(q)$ respectively with the anti-commutator algebra*

$$\left\{\psi(q),\psi^\dagger(q')\right\}=\delta^2(q-q') \tag{3.1a}$$

$$\{\psi(q),\psi(q')\}=0=\left\{\psi^\dagger(q),\psi^\dagger(q')\right\} \tag{3.1b}$$

and the empty vacuum state $|0\rangle$ which satisfies $\psi(q)|0\rangle=0\ \ \forall q$. The state with N particles occupying positions $q_1,\ldots,q_N$

$$|q_1,\ldots,q_N\rangle=\psi^\dagger(q_1)\ldots\psi^\dagger(q_N)|0\rangle \tag{3.2}$$

is an anti-symmetric function of $q_1,\ldots,q_N$. These states form a complete (δ-function normalized) basis for the N-particle subspace of the Hilbert space of the second quantized theory. The N-particle wavefunction, $|\Psi(t)\rangle$, can be expanded in this basis as

$$|\Psi(t)\rangle=\frac{1}{N!}\int dq_1\ldots dq_N\psi(q_1,\ldots,q_N;t)\psi^\dagger(q_1)\ldots\psi^\dagger(q_N)|0\rangle \tag{3.3}$$

and

$$\psi(q_1,\ldots,q_N;t)=\langle q_1,\ldots,q_N|\Psi(t)\rangle \tag{3.4}$$

are the totally anti-symmetric N-particle wavefunctions of the quantum mechanics problem in Section II. On the states (3.2) the charge density operator

$$j^0(q)=\psi^\dagger(q)\psi(q) \tag{3.5}$$

has the eigenvalues

$$j^0(q)|q_1,\ldots,q_N\rangle=\sum_i\delta^2(q-q_i)|q_1,\ldots,q_N\rangle \tag{3.6}$$

with density concentrated at the discrete points $q_1,\ldots,q_N$.

The wavefunction (3.4) is an anti-symmetric function of the particle positions. The second-quantized Hamiltonian operator is

$$H=\int dq\psi^\dagger(q)\frac{1}{2m}\left(i\nabla+\hat{a}(q)\right)^2\psi(q) \tag{3.7}$$

where the gauge field is the operator

$$\hat{a}(q)=\alpha\int dq'\nabla\theta(q-q')\psi^\dagger(q')\psi(q') \tag{3.8}$$

The quantum mechanical problem of solving the second quantized Schroedinger equation

$$i\frac{\partial}{\partial t}|\Psi(t)\rangle=H|\Psi(t)\rangle \tag{3.9}$$

* Here we find it convenient to take the canonical statistics particles as Fermions.

is identical to solving the first-quantized theory in (2.21) where the wavefunctions of the latter are taken as completely anti-symmetric. The inequalities in the particle sums in the interaction term and also in the external gauge field in (2.21) are obtained in (3.7) by using the specific choice of operator ordering shown there together with the exclusion principle.

The gauge field operator in (3.8) satisfies the equations

$$\nabla \cdot \hat{a}(q) = 0 \tag{3.10}$$

$$\hat{b}(q) = \nabla \times \hat{a}(q) = 2\pi\alpha\psi^\dagger(q)\psi(q) \tag{3.11}$$

which are the gauge fixed solutions of a Chern-Simons gauge theory.

In fact, consider the Lagrangian

$$L = \int d^2q \left(\psi^\dagger (i\partial_t + a_0)\psi - \psi^\dagger \frac{1}{2m}(i\nabla + a)^2\psi - \frac{1}{4\pi\alpha}\epsilon^{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda\right) \tag{3.12}$$

which describes Fermionic matter coupled to a U(1) gauge field where the kinetic term for the gauge field is the Chern-Simons term. A model of this kind was first suggested by Hagen [Hagen, 1985] who argued the Chern-Simons coupling led to fractional spin. Canonical quantization of this theory [Semenoff 1988, Matsuyama 1989] yields nonvanishing (anti-)commutation relations,

$$\left\{\Psi(q), \Psi^\dagger(q')\right\} = \delta(q - q') \tag{3.13}$$

$$[\hat{a}_i(q), \hat{a}_j(q')] = 2\pi i\alpha\epsilon_{ij}\delta(q - q') \tag{3.14}$$

and the gauge constraint

$$\mathcal{G}(q) \equiv \hat{b}(q) - 2\pi\alpha\psi^\dagger(q)\psi(q) \sim 0 \tag{3.15}$$

which, using the commutation relations (3.13-14), generates the static gauge transformations of the dynamical variables,

$$\psi(q) \to \exp\left\{(i/2\pi\alpha)\int \chi\mathcal{G}\right\}\psi(q)\exp\left\{-(i/2\pi\alpha)\int \chi\mathcal{G}\right\} = e^{i\chi(q)}\psi(q) \tag{3.16}$$

$$\hat{a}_i(q) \to \exp\left\{(i/2\pi\alpha)\int \chi\mathcal{G}\right\}\hat{a}_i(q)\exp\left\{-(i/2\pi\alpha)\int \chi\mathcal{G}\right\} = \hat{a}_i(q) + \nabla_i\chi(q) \tag{3.17}$$

which is a symmetry of the Hamiltonian (3.7), i.e.

$$[\mathcal{G}(q), H] = 0 \tag{3.18}$$

The constraint (3.15) can be taken as a physical state condition which separates the gauge invariant states in the Hilbert space of the quantum theory

$$\mathcal{G}(q)\,|\Psi\rangle_{\text{phys}} = 0 \tag{3.19}$$

Alternatively, existence of a constraint which commutes with the Hamiltonian indicates that the set of dynamical variables used to describe the system is too large. The physical subset of the dynamical variables are equivalence classes where any two configurations which are related by a gauge transformation (3.17-18) are equivalent. A representative of these equivalence classes can be chosen by imposing an additional 'gauge fixing' constraint [Dirac, 1969]. Here it is convenient to choose

$$\nabla \cdot \hat{a}(q) = 0 \tag{3.20}$$

which is conjugate to the gauge constraint (3.15) and selects the transverse component of $a(q)$ as the representative of the gauge orbit of $a(q)$. The solution of (3.15) and (3.20) is the determined gauge field (3.8). Note that this solution is an operator which does not commute with ψ and $\psi^\dagger$. Therefore it is necessary to specify an operator ordering in the Hamiltonian (3.7). With the particular operator ordering in (3.7) we regain the quantum mechanical problem discussed in Section I. All operators in this model are to be ordered so that covariant derivatives of positively charged fields operate from the left

$$iD\psi \equiv (i\nabla + \hat{a})\psi \tag{3.21}$$

and negatively charged fields operate from the right,

$$(iD\psi)^\dagger \equiv \psi^\dagger(-i\overleftarrow{\nabla} + \hat{a}) \tag{3.22}$$

Thus we see that the gauge field carries no physical degrees of freedom and can be eliminated by gauge fixing and solving the gauge constraints. We must also specify an operator ordering for the resulting Hamiltonian. We shall see shortly that this gives quantum states with anyonic statistics. First consider the gauge invariant angular momentum operator,

$$J = \int dq\psi^\dagger(q)q \times (i\nabla + a(q))\psi(q) \tag{3.23}$$

That this is the correct generator can easily be seen by noting that it is the standard matter component of the gauge invariant angular momentum operator which would be constructed from the gauge invariant symmetric energy-momentum tensor of the non-relativistic Fermion field. Also, the Chern-Simons term in the action does not contribute to the energy-momentum tensor beacause it is a 3-form and therefore does not depend on the metric of the space. The gauge-invariant symmetric energy momentum tensor can be obtained from the action by taking a functional derivative with respect to a background metric tensor. To see the transformation which (3.23) generates, we substitute the determined gauge field $a(q)$ in equation (3.23),

$$\begin{aligned} J &= \int dq\psi^\dagger(q)q \times (i\nabla)\psi(q) + \frac{\alpha}{2}\int dqdq'\, q \times \frac{\partial}{\partial q}\theta(q - q')\psi^\dagger(q)\psi^\dagger(q')\psi(q')\psi(q) \\ &= J_c + \alpha\int dq \int dq' \frac{q\cdot(q - q')}{(q - q')^2}\left(\psi^\dagger(q)\psi(q)\psi^\dagger(q')\psi(q') - \delta(q - q')\psi^\dagger(q)\psi(q)\right) \\ &= J_c + \frac{\alpha}{2}Q(Q - 1) \end{aligned} \tag{3.24}$$

where to derive the last equality we have taken into account the fact that the operator part of the integrand in the second term is symmetric in q and q' and that the symmetrization of $q \cdot (q -$

$q')/(q-q')^2$ is 1/2. This result is very similar to (2.19), the canonical angular momentum operator is augmented by an operator whose expectation value is the total number of pairs of particles. The rotation of the Schroedinger field generated by J is,

$$e^{i\omega J}\psi(q)e^{-i\omega J} = e^{i\omega\frac{q}{2}Q(Q+1)}\psi(\Lambda(\omega)q)e^{-i\omega\frac{q}{2}Q(Q+1)}$$
$$= e^{-i\omega\alpha Q}\psi(\Lambda(\omega)q) \tag{3.25}$$

We see that the field ψ rotates canonically except for an operator-valued phase. The Schroedinger field operator does not give a faithful representation of the rotation group, but has an operator-valued cocycle [Semenoff and Sodano, 1989; Semenoff, 1989; Forte and Jolicoeur, 1991]. In particular, if we choose $\omega = 2\pi$ the spin-parity of ψ, defined by the change of its phase under a 2π rotation, is an operator

$$e^{2\pi iJ}\psi(q)e^{-2\pi iJ} = e^{-2\pi i\alpha Q}\psi(q) \tag{3.26}$$

The latter problem with operator-valued spin-parity is easily cured by a change of variables. We begin by noting that before gauge fixing, the operator $\psi(q)$ is not gauge invariant. A gauge invariant operator is constructed from it by taking the product with a Wilson line operator,

$$\phi_C(q) = \exp\left\{i\int_{\mathcal{B}}^{q} dl\cdot\hat{a}(l)\right\}\psi(q) \quad \phi_C^\dagger(q) = \psi^\dagger(q)\exp\left\{-i\int_{\mathcal{B}}^{q} dl\cdot\hat{a}(l)\right\} \tag{3.27}$$

where C is a curve which goes from the point q to a base-point $\mathcal{B}$ at infinity. The operators $\phi_C(q)$ and $\phi_C^\dagger(q)$ are invariant under gauge transformations where the gauge function has compact support. When we fix the gauge symmetry, we use the solution of the constraints (3.8) for the gauge field and (3.27) takes the form

$$\phi_C(q) = \exp\left\{i\alpha\int d^2q'\theta_C(q-q')\psi^\dagger(q')\psi(q)\right\}\psi(q)$$
$$\phi_C^\dagger(q) = \psi^\dagger(q)\exp\left\{-i\alpha\int d^2q'\theta_C(q-q')\psi^\dagger(q')\psi(q)\right\} \tag{3.28}$$

where

$$\theta_C(q-q') = \int_{\mathcal{B}}^{q} dl\cdot\nabla_l\theta(l-q') \tag{3.29}$$

is the multi-valued angle function defined by integration on C.

The states created from the vacuum by the operators in (3.28) are multi-valued,

$$\phi_{C_1}(q_1)\ldots\phi_{C_N}(q_N)|0\rangle = \exp\left(-i\alpha\nu[C;q_1,\ldots,q_N]\right)|q_1,\ldots,q_N\rangle \tag{3.30}$$

where $C = \sum C_i$, and carries the one-dimensional unitary representation of the braid group discussed in Section II.

It is straightforward to compute the spin-parity of the operator $\phi_C(q)$. In fact, we can now show that the angular momentum operator J now generates the rotation

$$e^{i\omega J}\phi_C(q)e^{-i\omega J} = \phi_{\Lambda C}(\Lambda(\omega)q) \tag{3.31}$$

where ΛC is the curve C rotated by angle ω about the origin. Since $\Lambda(2\pi)C = C$ the field ϕ_C has vanishing spin-parity. Thus the particles created by the multi-valued operators have fractional statistics but canonical spin [Semenoff, 1988].

Finally, it is possible to show that the Hamiltonian in terms of multi-valued operators has the form of a free Hamiltonian

$$H = \int dq\phi_C^\dagger(q)\frac{-\nabla^2}{2m}\phi_C(q) \tag{3.32}$$

Thus, we have traded holonomy for the gauge interaction of the Chern-Simons theory. Note that even though (3.32) appears to be a free Hamiltonian, the operators and the quantum states are multi-valued and finding the spectrum of H remains a highly nontrivial and as yet unsolved problem.

In this Section we have discussed a second quantized nonrelativistic field theory of anyons and have shown that it is described by matter fields with canonical statistics coupled to a Chern-Simons gauge theory. An interesting question is whether a relativistic field theory coupled to a Chern-Simons gauge theory would also describe a theory of relativistic anyons. The difficulty with transposing our results directly to relativistic field theory is that our analysis depended heavily on use of the empty vacuum $|0\rangle$ to define the anyonic states and the mapping from the Chern-Simons theory to the anyon theory. In a relativistic theory the empty vacuum is not the correct one to describe the theory. For example, for Fermions the empty vacuum is the one where both the positive and the negative energy states are empty. This is orthogonal and in fact very far away from the Dirac vacuum which is has the negative energy states occupied. Hence, in order to carry our proceedure over to relativistic field theory we need an ultraviolet regularization. In the following Sections we shall examine lattice regularizations of Chern-Simons theory.

IV. Lattice Chern-Simons Theory: The Minimal Model

In this Section we shall examine a particular form of lattice Chern-Simons theory which has a non-compact U(1) gauge symmetry. A lattice Chern-Simons-Maxwell theory has been constructed in the past by using a lattice generalization of exterior differential calculus to derive a gauge invariant 3-form with the properties of a Chern-Simons term [Frohlich and Marchetti, 1988]. To begin we work on a cubic Euclidean lattice with sites $x, y, \ldots$ which are superpositions of the lattice unit vectors $\hat{1}, \hat{2}, \hat{3}$ with integer coefficients, i.e. $x = n_1\hat{1} + n_2\hat{2} + n_3\hat{3}$. We denote generic basis vectors by $\mu, \nu, \ldots$. We define the forward difference operator d_μ as $d_\mu f(x) = f(x+\mu) - f(x)$ and the backward difference operator as $\hat{d}_\mu f(x) = f(x) - f(x-\mu)$. The Chern-Simons term constructed by Frohlich and Marchetti has the form

$$S_{CS} = \frac{1}{4\pi\alpha}\sum_{xy\mu\nu} A_\mu(x)K_{\mu\nu}(x-y)A_\nu(y) \tag{4.1}$$

where the covariance K is given by

$$K_{\mu\nu}(x-y) = \frac{1}{2}\epsilon_{\mu\nu\lambda}\left(d_\lambda\delta(x+\mu-y) + \hat{d}_\lambda\delta(x-\nu-y)\right) \tag{4.2}$$

We would like to consider a gauge theory where (4.1) is the only kinetic term for the gauge field A. However, this theory has the immediate difficulty that the covariance in (4.2) is degenerate. Aside from the usual zero eigenvalue which arises from the gauge invariance, the eigenvalues of the K are

$$K_{\pm}(p) = \pm\sqrt{\left(3 - \sum_1^3 \cos p_\mu\right)\left(1 + \cos \sum_1^3 p_\mu\right)} \tag{4.3}$$

where $p_\mu \in [-\pi, \pi]$ are the Bloch momenta. The eigenvalues vanish on the planes $\sum p_\mu = \pm\pi$. These planes have co-dimension 1 and the fact that $K_\pm(p)$ vanish there leads to logarithmic infrared divergences in the fourier transform of $K^{-1}_{\mu\nu}(p)$.

If we take the continuum time limit we obtain the action

$$S_{CS} = \frac{1}{4\pi\alpha}\int dt\left(\sum_x A_0(x,t)\left(B(x,t) + B(x - \hat{1} - \hat{2}, t)\right) + \sum_{xy}\dot{A}_i(x,t)K_{ij}(x-y)A_j(y,t)\right) \tag{4.4}$$

where we now denote the spatial directions by $i, j, \ldots$, the canonical form is

$$K_{ij}(x-y) = -\frac{1}{2}\epsilon_{ij}\left(\delta(x+i-y) + \delta(x-j-y)\right) \tag{4.5}$$

and $B(x,t) = \epsilon ij d_i A_j(x,t)$. If we attempt to invert K as

$$K^{-1}_{ij}(x) = 2\epsilon_{ij}\int \frac{d^2p}{(2\pi)^2}\frac{e^{-ip\cdot x}}{e^{ip_i} + e^{-ip_j}} \tag{4.6}$$

the integral has a logarithmic infrared divergence coming from the integration region near the lines $p_1 + p_2 = \pm\pi$. There is the related difficulty that, if we couple to a charge density by adding a term $A_0\rho$ to (4.4), Gauss Law is

$$2\pi\alpha\rho(x,t) = \frac{1}{2}\left(B(x,t) + B(x - \hat{1} - \hat{2}, t)\right) \tag{4.7}$$

This equation contains an extra constraint on the charge distribution

$$\sum_{x,y}(-1)^{x+y}\rho(x,y,t) = 0 \tag{4.8}$$

Here we shall try to circumvent this difficulty by considering the most general translationally invariant Chern-Simons term (here we work in continuum time),

$$S_{CS} = \frac{1}{4\pi\theta}\int dt\sum_{x,y}\left(2A_0(x,t)J_i(x-y)A_i(y,t) + \dot{A}_i(x,t)K_{ij}(x-y)A_j(y,t)\right) \tag{4.9}$$

and use the requirement of gauge invariance to determine the general form of the covariances J_i and K_{ij}. The gauge tranformation is $A_0(x,t) \to A_0(x,t) + \dot{\Lambda}(x,t)$, $A_i(x,t) \to A_i(x,t) + \Lambda(x+i,t) - \Lambda(x,t)$. In quantization of this theory the gauge fields have the commutator

$$\left[A_i(x), A_j(y)\right] = 2i\pi\alpha K^{-1}_{ij}(x-y)\ , \tag{4.10}$$

and so we shall also require that the covariance K_{ij} is invertable.

In momentum space

$$J_i(x) = \int_{\Omega_B}\frac{dk}{(2\pi)^2}e^{ik\cdot x}J_i(k) \tag{4.11}$$

$$K_{ij}(x) = \int_{\Omega_B}\frac{dk}{(2\pi)^2}e^{ik\cdot x}K_{ij}(k) \tag{4.12}$$

where the Brillouin zone is $\{\Omega_B : -\pi < k_i \le \pi\}$ and the Fourier transform of the forward and reverse difference operators is $d_i(k) = e^{-ik_i} - 1$ and $\hat{d}_i(k) = 1 - e^{ik_i} = -d_i^*(k)$, respectively. Gauge invariance requires

$$\sum_i d_i(k)J_i(k) = 0 \qquad \sum_i \hat{d}_i(k)\left(K_{ij}(k) - K_{ji}(-k)\right) = J_j(k) \tag{4.13}$$

The first equation is solved by

$$J_i(k) = \sum_j \epsilon_{ij}d_j J(k) \tag{4.14}$$

K has the decomposition [where $d_i^P = \epsilon_{ij}d_j$, etc.]

$$K_{ij}(k) = K_{NN}\hat{d}_i d_j + K_{PP}\hat{d}_i^P d_j^P + K_{PN}\hat{d}_i^P d_j + K_{NP}\hat{d}_i d_j^P \tag{4.15}$$

Without loss of generality we take $K_{ij}(k) = -K^*_{ji}(k)$. Then (4.13) implies $K_{NN}(k) = 0$, $K_{PP}(k)$ is arbitrary and

$$K_{NP}(k) = J(k)/2\hat{d}\cdot\hat{d}\ \ ,\ \ K_{PN}(k) = -J^*(k)/2d\cdot d \tag{4.16}$$

and $\det K = \frac{1}{4}J^*J$, i.e. $K_{ij}(k)$ is an invertible matrix whenever $|J(k)| \neq 0$. If, for example, $J = 1$ then

$$K_{ij}(k) = \hat{d}_i d_j^P/2\hat{d}\cdot\hat{d} - \hat{d}_i^P d_j/2d\cdot d + K_{PP}\hat{d}_i^P d_j^P \tag{4.17}$$

We shall see later on that it is exactly this choice, $J = 1$ that gives gauge invariant operators the correct holonomy so that particles are anyons. We will work with this choice from now on. Note that in this case the lattice Chern-Simons term is nonlocal in that it couples sites at large distances with a power-law cutoff of the coupling strength with distance. This is also seen in the inverse of K is

$$K^{-1}_{ij}(p) = 2d_i^P\hat{d}_j/\hat{d}^2 - 2d_i\hat{d}_j^P/d^2 + d_i\hat{d}_j K_{PP} \tag{4.18}$$

which, by (4.10), gives the gauge fields a non-local canonical structure. If, when this model is coupled to matter we are to have a Gauss Law like $B(x) \sim \rho(x)$, this non-locality is unavoidable. We can think of this as a generalization to lattice Chern-Simons theory of the Neilsen-Ninomiya theorem [Neilsen and Ninomiya, 1982] which forbids local lattice formulations of chiral fermions.

We couple fermionic matter fields with creation and annihilation operators $a^\dagger(x)$ and $a(x)$, respectively, which gauge transform as

$$a(x) \to e^{i\chi(x)}a(x)\ \ ,\ \ a^\dagger(x) \to a^\dagger(x)e^{-i\chi(x)} \tag{4.19}$$

and have the nonvanishing anti-commutator

$$\left\{a(x), a^\dagger(y)\right\} = \delta(x-y) \tag{4.20}$$

The generator of gauge transformations is

$$\mathcal{G}(x) = \frac{1}{2\pi\alpha}B(x) - a^\dagger(x)a(x) \tag{4.21}$$

Gauge invariance of the quantum theory is obtained either by choosing a subset of the states in the Hilbert space which are gauge invariant, i.e. which obey

$$\mathcal{G}(x)|\text{phys}\rangle = 0 \tag{4.22}$$

or by taking $\mathcal{G}(x) \sim 0$ as a constraint which, together with a gauge fixing condition, can be used to solve for the gauge field in terms of the matter charge density. In either case we construct operators which are invariant under all gauge transformations where the gauge functions have compact support by attatching Wilson line operators to the creation and annihilation operators for Fermions:

$$\alpha_C(x) = \mathcal{P}\exp\left(i\sum_{\text{links}\in C} A\right)a(x)\ ,\quad \alpha_C^\dagger(x) = a^\dagger(x)\mathcal{P}^\dagger\exp\left(-i\sum_{\text{links}\in C} A\right): \tag{4.23}$$

Here $\mathcal{P}$ and $\mathcal{P}^\dagger$ denote path ordering and anti-path ordering, respectively, of the Wilson line operators. It is necessary in this case since the components of the gauge field operator do not commute with each other even at separated distances. Generally, since the commuators of gauge fields are c-numbers, the difference between the path-ordered Wilson line operator and one with other orderings are some multi-valued c-number phases. There is some freedom in choosing this ordering. Here we choose path-ordering since it is natural for parallel transport of a charged field.

In order to study the holonomy of quantum states we begin with canonical quantization. The gauge fields have the commutator (4.10). In order to proceed it is useful to change variables so that one of the gauge field variables is the magnetic field operator

$$B(x) = d_i^P A_i(x) \tag{4.24}$$

and the other variable contains the longitudinal part of the gauge field

$$\lambda(x) = \frac{1}{2}\sum_y\left((x|\frac{1}{d^2}|y)d\cdot A(y) + \frac{1}{2}K_{PP}(x-y)B(y)\right) \tag{4.25}$$

Then, the nonvanishing commutator for these variables is

$$[\lambda(x), B(y)] = -2i\pi\alpha\delta(x-y) \tag{4.26}$$

We quantize using the functional Schroedinger picture [Dunne, Jackiw and Trugenberger, 1989]. We take states as functionals of the 'coordinate' variable $\lambda(x)$ and represent its canonical conjugate, the magnetic field, as a derivative operator

$$B(x) = 2\pi i\alpha\frac{\partial}{\partial\lambda(x)} \tag{4.27}$$

Plane wave states which are eigenstates of $B(x)$ are given by

$$B(x)\ \exp\left(-\frac{i}{2\pi\alpha}\sum_y\lambda(y)b(y)\right) = b(x)\ \exp\left(-\frac{i}{2\pi\alpha}\sum_y\lambda(y)b(y)\right) \tag{4.28}$$

Using these plane waves, we can construct a basis of gauge invariant states in the n-particle sector of the theory. Consider the state where all of the positions of the particles are fixed

$$|x_1,\dots x_N\rangle = a^\dagger(x_1)\dots a^\dagger(x_N)|0\rangle\otimes\exp\left(-i\sum_i\lambda(x_i)\right) \tag{4.29}$$

It can be checked that these states obey Gauss' Law

$$\mathcal{G}(x)|x_1\dots x_N\rangle = 0 \tag{4.30}$$

and form a complete basis for the physical subspace of the Hilbert space. Note that they are single-valued wavefunctions. Alternatively, using the gauge invariant multi-valued fields defined in (4.23) we can begin with the empty vacuum $|0\rangle\otimes 1$ and create a gauge invariant state

$$|x_1,\dots,x_N;C_1,\dots,C_N\rangle = \alpha_{C_1}^\dagger(x_1)\dots\alpha_{C_N}^\dagger(x_N)|0\rangle\otimes 1 \tag{4.31}$$

These states are equivalent to those in (4.29) since the Wilson line operator

$$W[C_{xy}] = \mathcal{P}\exp\left(i\sum_{\text{links}\in C} A\right) \tag{4.32}$$

where C_{xy} is an oriented curve following links of the lattice from site x to site y, creates magnetic flux at its endpoints.

$$B(w)W[C_{xy}] = W[C_{xy}]\left(B(w) + 2\pi\alpha\delta(x-w) - 2\pi\alpha\delta(y-w)\right) \tag{4.33}$$

Thus, if we consider a closed loop ℓ which is the boundary of a region D of the lattice, we have the identity

$$W[\ell]W[C_{xy}] = W[C_{xy}]W[\ell]\begin{cases} e^{2\pi i\alpha n} & y\in D,\ x\notin D \\ e^{2\pi i\alpha(n-m)} & xy\in D \text{ or } xy\notin D \\ e^{-2\pi i\alpha m} & x\in D,\ y\notin D \end{cases} \tag{4.34}$$

where n and m are the numbers of times ℓ links x and y respectively.

A charged particle is transported along path C_{xy} from x to y using the operator

$$T[C_{xy}] = a^\dagger(y)W[C_{xy}]a(x) \tag{4.35}$$

Fermi statistics guarantees that none of the sites through which the particle is transported are occupied. Transporting a particle around a loop effectively multiplies the state from the left by the Wilson loop operator

$$W[\ell]|x_1,\dots,x_N;C_1,\dots,C_N\rangle$$

$$= \exp(2\pi i\alpha(\text{net number of particles linked by } \ell)\,|x_1,\ldots,x_N;C_1,\ldots,C_N\rangle \tag{4.36}$$

Thus, the wavefunctions carry the one-dimensional representation of the braid group discussed in Section II. This would not be possible without the locality of the gauge generator in (4.21). It is interesting that the requirement of the local gauge generator and then gauge invariance of the Chern-Simons term are sufficient to give the gauge invariant states the correct holonomy to be anyons.

Attatching Wilson line operators to the creation and annihilation operators for charged particles as in (4.23) is, in Chern-Simons theory, a 2-dimensional version of the Jordan-Wigner transformation [Jordan and Wigner, 1928; Ambjorn and Semenoff, 1989; Fradkin, 1989]. This is more explicit in an alternative construction of gauge invariant operators given by Luscher and Muller [Luscher, 1989; V. Muller, 1990; Eliezer, Semenoff and Wu, 1990]. Here, we shall follow the presentation of Eliezer, Semenoff and Wu.

Consider the inverse of the lattice Laplace operator

$$g(x) = (x|-1/d\cdot\hat{d}|0) = \int_{\Omega_B}\frac{d^2p}{(2\pi)^2}(1-e^{ipx})1/d(p)\cdot\hat{d}(p) \tag{4.37}$$

and the function

$$f_i(x) = \hat{d}_i^P g(x) \tag{4.38}$$

which satisfies

$$-d_i^P f_i(x) = \delta(x) \tag{4.39}$$

Now, consider a continuous open curve which follows the links of the lattice from a base-point B to lattice site x and the function

$$\theta_C(x,y) = 2\pi\sum_{\text{sites } z\in C} l(z)\cdot f(z-y) \tag{4.40}$$

where $\vec{l}(z) = \pm\hat{i}$ for the oriented link going from site z to site $z\pm\hat{i}$, respectively, and we use $f_{-i}(z) = f_i(z-\hat{i})$ for negatively oriented links. θ_C is a function on the universal cover of the union of the lattice with its links and a multi-valued function of the lattice points. If we compare θ_C for two different curves C and C',

$$\theta_C(x,y) - \theta_{C'}(x,y) = \pm 2\pi\sum_{z\in\text{pl.}\in C-C'}\delta(z-y) \tag{4.41}$$

where the $\pm$sign is the orientation of $C-C'$, the sum is over plaquettes bounded by the closed curve $C-C'$ and the lattice site associated with a plaquette is the site in the lower left-hand corner. Using θ_C we construct the multi-valued operator which is the analog of a similar operator first used in the continuum [Semenoff, 1988; Semenoff and Sodano, 1989]

$$\mathcal{K}_C(x) = \exp\left\{i\alpha\sum_z\theta_C(x,z)a^\dagger(z)a(z)\right\} \tag{4.42}$$

We also consider the single-valued but gauge-variant operator

$$U(x) = \exp\left\{-i\sum_{ijz}\epsilon_{ij}f_i(x-z)A_j(z)\right\} \tag{4.43}$$

We form the anyon annihilation and creation operators

$$\phi_C(x) = U(x)\mathcal{K}_C(x)a(x)\ ,\ \phi_C^\dagger(x) = a^\dagger(x)\mathcal{K}_C^\dagger(x)U^\dagger(x) \tag{4.44}$$

where the charge density

$$\rho(x) = a^\dagger(x)a(x) = \phi_C^\dagger(x)\phi_C(x) \tag{4.45}$$

is independent of C. These operators have commutation relations

$$\phi_C(x)\phi_{C'}(y) = \phi_{C'}(y)\phi_C(x)\exp i\alpha\left(-\pi+2\pi\nu[C,C']\right) \tag{4.46}$$

$$\phi_C^\dagger(x)\phi_{C'}^\dagger(y) = \phi_{C'}^\dagger(y)\phi_C^\dagger(x)\exp -i\alpha\left(-\pi+2\pi\nu[C,C']\right) \tag{4.47}$$

$$\phi_C(x)\phi_{C'}^\dagger(y) - \phi_{C'}^\dagger(y)\phi_C(x)\exp i\alpha\left(-\pi+2\pi\nu[C,C']\right) = \delta(x-y) \tag{4.48}$$

where $\nu[C,C']$ is the signed intersection number of C and C'. To derive these, we have used the identity [Luscher, 1990; Eliezer, Semenoff and Wu, 1990]

$$\theta_C(x,y) - \theta_{C'}(y,x) = -\pi + 2\pi\nu[C,C'] + \\ +\frac{1}{2}(f_1(x-y)+f_2(x-y)+f_1(x-y+\hat{2})+f_2(x-y+\hat{1})) \tag{4.49}$$

which holds when $x\neq y$ and $\mathcal{B}$ is infinitely separated from x and y.* We have also used a specific choice of K_{PP}:

$$K_{PP}(p) = \frac{d\times\hat{d}}{d\cdot\hat{d}}\left(\frac{1}{d^2}+\frac{1}{\hat{d}^2}\right) \\ -\frac{1}{8\pi}(f_1(p)+f_2(p)+(1+d_2(p))f_1(p)+(1+d_1(p))f_2(p)) \tag{4.50}$$

Note that this choice of K_{PP} leads to a Chern-Simons term which has a nonlocal coupling between the (nondynamical) magnetic field and its time derivative. This term is down by one power of the lattice spacing and goes to zero in the continuum limit.

The anyon creation and annihilation operators are gauge invariant,

$$[\phi_C(x),\mathcal{G}(y)] = 0 = [\phi_C^\dagger(x),\mathcal{G}(y)] \tag{4.51}$$

so that with a ground state which obeys the physical state condition

$$\mathcal{G}(x)|0\rangle = 0 \tag{4.52}$$

* We note that setting $\mathcal{B}$ at infinity is compatible with finite volume if we consider lattice of interest as a bounded sublattice of an infinite lattice and $\mathcal{B}$ set at spatial infinity of the latter.

we have

$$\mathcal{G}(x)\phi^\dagger_{C_1}(y_1)\dots\phi^\dagger_{C_N}(y_N)|0\rangle = 0 \tag{4.53}$$

In addition, the wavefunction

$$\phi^\dagger_{C_1}(y_1)\dots\phi^\dagger_{C_N}(y_N)|0\rangle$$

changes by the phase $\pi\alpha + 2\pi\alpha\nu[C_i, C_{i+1}]$ when particle at (y_i, C_i) is interchanged with particle at (y_{i+1}, C_{i+1}). We denote this interchange by σ_i. It satisfies the relations $\sigma_i\sigma_j = \sigma_j\sigma_i$ when $|i-j| \geq 2$ and $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$ which are the structure relations of the N^{th} order braid group of the plane, $\mathcal{B}_N$, with presentation $\{\sigma^i\}$ which was discussed in Section II. The wavefunction therefore carries the unitary one-dimensional representation of $\mathcal{B}_N$ characteristic of the wavefunctions of anyons.

The variation of $\phi_C(x), \phi^\dagger_C(x)$ under change of C is

$$\phi_{C'}(x) = \exp 2\pi i\alpha \left\{ \sum_{z\in \text{pl.}\in C-C'} \phi^\dagger_C(z)\phi_C(z) \right\} \phi_C(x) \tag{4.54}$$

i.e. by a phase which is proportional to the charge in the region bounded by $C - C'$. In the number basis, where each basis vector corresponds to a distinct specified occupation number of all of the lattice sites, $\rho(z)$ is diagonal with eigenvalues 0 and 1 for each z. Since the charge in any region has integer eigenvalues, when α is a rational number, the phases in (4.54) take on a finite number of distinct values.

Thus, we see that our minimal, nonlocal lattice Chern-Simons theory leads to the correct holonomy for anyon wavefuctions. Even though the gauge fields have nonlocal canonical structure this may be an acceptable physical theory since the gauge fields contain no physical degrees of freedom. In fact we should think them as simply a kinematical mechanism for taking into account the fact that the charged particles in the theory are anyons.

A remaining question is whether the Chern-Simons gauge fields decouple in the lattice Hamiltonian. Let us consider the Hamiltonian

$$H = \sum_{x,i} a^\dagger(x+i)e^{iA_i(x)}a(x) + \text{h.c.} \tag{4.55}$$

In terms of multi-valued fields this becomes

$$H = \sum_{x,i} \phi_C(x+i)U_{CC'}\phi_{C'} + \text{h.c.} \tag{4.56}$$

where

$$U_{CC'} = \exp\left(2\pi i\alpha \sum_{\text{int.}C\cup C'\cup\text{link}} \rho(z)\right) \tag{4.57}$$

which is the exponential of $2\pi i\alpha$ times the charge operator summed over the sites inside the disc bounded by the union of the curves $-C$ from $x+i$ to $\mathcal{B}$, C' from $\mathcal{B}$ to x and the link between

x and $x+i$. This phase factor is necessary in order to make the hamiltonian single-valued, i.e. independent of the curves C and C'. Thus, the long-range interactions of the Chern-Simons term do not decouple on a lattice. There is, however, the special case there the statistics parameter α is a rational number. Then, since the charge density operator has integer eigenvalued, the phase U takes on only a finite number of different values. Indeed (4.56) can be thought of as a sort of Z_N gauge theory where N is the smallest integer for which $N\alpha$=integer [Eliezer, Semenoff and Wu, 1990]. A Z_N gauge field of this kind has previously appeared in the work of Hatsugai, Kohmoto and Wu [Hatsugai, Kohmoto and Wu, 1991]. In the special case there α is an integer, U is the unit matrix and the resulting Hamiltonian is the free Hamiltonian. When α is odd, the fields $\alpha_C(x)$ are bosons. Thus, we find that it is possible to use Chern-Simons theory for Fermion-Boson transmutation and the entire role of the Chern-Simons gauge field there is to transmute the statistics.

V. Lattice Chern-Simons Theory: A Parity Doubled Model

In the last Section we reviewed the 'minimal' formulation of a Hamiltonian lattice Chern-Simons theory which has one gauge field degree of freedom per link. We saw that in the simplest case where it led to topological holonomy for anyons, the Chern-Simons term was necessarily nonlocal in that it could not be written as a finite combination of lattice difference operations. In this Section we shall investigate the properties of a local, minimally doubled lattice Chern-Simons theory [Kantor and Susskind, 1990; Eliezer, Semenoff and Wu, 1991].

For convenience we shall deviate somewhat from the notational conventions of Section IV. We shall consider a square 2 dimensional spatial lattice with sites $x = n_1\hat{1} + n_2\hat{2}$ and basis vectors $\mu = \pm\hat{1}, \pm\hat{2}$ which is called the primary lattice. An oriented link from site x to site $x+\mu$ is denoted $[x,\mu]$. Similarly we consider the dual lattice with sites x^* located at the centers of plaquettes of the primary lattice. We take the convention that the link $[x,\mu]^*$ dual to $[x,\mu]$ is that link of the dual lattice which crosses $[x,\mu]$ and if $[x,\mu]$ is oriented in the forward direction, $[x,\mu]^*$ is oriented from right-to-left. Then we introduce gauge fields $A([x,\mu],t)$ associated with the link $[x,\mu]$, $\tilde{A}([x,\mu]^*,t)$ associated with $[x,\mu]^*$, $A(x,t)$ associated with sites and $\tilde{A}(y^*,t)$ associated with dual sites. These have the gauge transformations

$$A([x,\mu],t) \to A([x,\mu],t) + d_\mu\chi(x,t) \quad , \quad A(x,t) \to A(x,t) + \dot{\chi}(x,t) \tag{5.1}$$

$$\tilde{A}([x,\mu]^*,t) \to \tilde{A}([x,\mu]^*,t) + \tilde{d}_{\tilde{\mu}}\tilde{\chi}(x^*,t) \quad , \quad \tilde{A}(y^*,t) \to \tilde{A}(y^*,t) + \dot{\tilde{\chi}}(y^*,t) \tag{5.2}$$

where $\chi(x,t)$ and $\tilde{\chi}(y^*,t)$ are independent functions on the primary lattice and dual lattice respectively, $d_\mu f(x) = f(x+\mu) - f(x)$ and the dual lattice derivative $\tilde{d}_{\tilde{\mu}}\tilde{\chi}(y^*)$ is defined by the operation of taking the difference of $\tilde{\chi}(y^*)$ at the dual sites on the endpoints of $[x,\mu]^*$. Gauge fields have the property

$$A([x,-\mu],t) = -A([x-\mu,\mu],t) \quad , \quad \tilde{A}([x,-\mu]^*,t) = -\tilde{A}([x-\mu,\mu]^*,t) \tag{5.3}$$

The primary lattice gauge field has magnetic flux through a plaquette, i.e. a site of the dual lattice,

$$B(x^* = x + \hat{1}/2 + \hat{2}/2, t) = A([x+\hat{1},\hat{2}],t) - A([x,\hat{2}],t) - A([x+\hat{2},\hat{1}],t) + A([x,\hat{1}],t) \quad (5.4)$$

Similarly the magnetic flux due to the dual lattice field is located on a primary lattice site,

$$\tilde{B}(x) = \tilde{A}([x,\hat{1}]^*,t) + \tilde{A}([x,\hat{2}]^*,t) - \tilde{A}([x-\hat{1},\hat{1}]^*,t) - \tilde{A}([x-\hat{2},\hat{2}]^*,t) \quad (5.5)$$

A Chern-Simons term which is invariant under the both gauge transformations (1) and (2) when the gauge functions have compact support is [Kantor and Susskind, 1990; Eliezer, Semenoff and Wu, 1991]

$$S_{CS} = \frac{1}{\theta}\int dt \left(\sum_{x,\mu=\hat{1},\hat{2}} A([x,\mu],t)\frac{d}{dt}\tilde{A}([x,\mu]^*,t) + \sum_x A(x,t)\tilde{B}(x,t) - \sum_{x^*}\tilde{A}(x^*,t)B(x^*,t)\right) \quad (5.6)$$

Canonical quantization of the first order action (6) yields the canonical commutation relation

$$[A([x,\mu]), \tilde{A}([y,\nu]^*)] = i\theta\delta([x,\mu],[y,\nu]) \quad (5.7)$$

In two dimensions parity is defined as reflecting one of the coordinates, $(x_1', x_2') = (-x_1, x_2)$ and

$$A'([x,\hat{1}],t) = A([x',-\hat{1}],t) \quad , \quad A'([x,\hat{2}],t) = A([x',\hat{2}],t) \quad (5.8)$$

$$\tilde{A}'([x,\hat{1}]^*,t) = -\tilde{A}([x',-\hat{1}]^*,t) \quad , \quad \tilde{A}'([x,\hat{2}],t) = -\tilde{A}([x',\hat{2}]^*,t) \quad (5.9)$$

$$A'(x,t) = A(x',t) \quad , \quad \tilde{A}'(x^*,t) = -\tilde{A}(x'^*,t) \quad (5.10)$$

and the Chern-Simons action is invariant. The continuum limit of (5.6) is

$$L_{CS} = \frac{1}{\theta}\int d^2x \left(\epsilon_{ij}A_i(x)\frac{d}{dt}\tilde{A}_j(x) + A_0(x)\tilde{B}(x) - \tilde{A}_0(x)B(x)\right) \quad (5.11)$$

where A is a vector under parity and $\tilde{A}_\mu$ is a pseudo-vector so that the Chern-Simons cross-coupling is parity invariant. Such a model has been used to discuss a scenario for parity symmetric anyon superconductivity [Semenoff and Wiess, 1990].

This gauge theory is suited to couple to matter which resides either on the primary lattice or the dual lattice and which has transport confined to its respective lattice. Gauge invariant transport of a particle with charge g along an oriented curve C on the primary lattice or a particle with charge $\tilde{g}$ along an oriented curve $\tilde{C}$ of the dual lattice are accompanied by the Wilson line operators

$$W[C] = \exp ig \sum_{\text{links in C}} A([x,\mu]) \quad , \quad \tilde{W}[\tilde{C}] = \exp i\tilde{g} \sum_{\text{dual links in } \tilde{C}} \tilde{A}([x,\mu]^*) \quad (5.12)$$

If x_i, x_f are the initial and final endpoints of C and x_i^*, x_f^* the initial and final points of $\tilde{C}$ the operators in (5.12) gauge transform as

$$W[C] \to e^{ig\chi(x_f)}W[C]e^{-ig\chi(x_i)} \quad , \quad \tilde{W}[\tilde{C}] \to e^{i\tilde{g}\tilde{\chi}(x_f^*)}\tilde{W}[\tilde{C}]e^{-i\tilde{g}\tilde{\chi}(x_i)} \quad (5.13)$$

Also, for closed curves, they are gauge invariant and give the holonomy

$$W[\text{closed } C] = \exp ig \sum_{x^* \in \text{interior } C} B(x^*) \quad , \quad \tilde{W}[\text{closed } \tilde{C}] = \exp i\tilde{g} \sum_{x \in \text{interior } \tilde{C}} \tilde{B}(x) \quad (5.14)$$

and for either open or closed curves they have the commutator algebra

$$W[C]\,W[C'] = W[C']\,W[C] \quad , \quad \tilde{W}[\tilde{C}]\,\tilde{W}[\tilde{C}'] = \tilde{W}[\tilde{C}']\,\tilde{W}[\tilde{C}] \quad (5.15)$$

$$\tilde{W}[\tilde{C}]\,W[C] = W[C]\,\tilde{W}[\tilde{C}]\exp ig\tilde{g}\theta\nu[\tilde{C},C] \quad (5.16)$$

where $\nu[\tilde{C},C]$ is the number of right-handed intersections minus the number of left-handed intersections of $\tilde{C}$ and C. It follows that for closed loops C and $\tilde{C}$, $W[C]$ and $\tilde{W}[\tilde{C}]$ always commute.

We introduce annihilation and creation operators $a, a^\dagger$ for particles with charge g which reside on primary lattice sites and $\tilde{a}, \tilde{a}^\dagger$ for (fermionic) particles with charge $\tilde{g}$ which reside on dual sites with nonvanishing commutators anti-

$$\left\{a(x), a^\dagger(y)\right\} = \delta_{xy} \quad , \quad \left\{\tilde{a}(x^*), \tilde{a}^\dagger(y^*)\right\} = \delta_{x^*y^*} \quad (5.17)$$

The generators of gauge transformations are

$$\mathcal{G}_\chi = \sum_x \chi(x)\left(\frac{g}{2}[a^\dagger(x), a(x)] + \frac{1}{\theta}\tilde{B}(x)\right) \quad , \quad \tilde{\mathcal{G}}_{\tilde{\chi}} = \sum_{x^*}\tilde{\chi}(x^*)\left(\frac{\tilde{g}}{2}[\tilde{a}^\dagger(x^*), \tilde{a}(x^*)] + \frac{1}{\theta}B(x^*)\right) \quad (5.18)$$

Here, we have used a Dirac commutator ordering of the operators in the charge operator. This ordering is used in order to make charge conjugation invariance compatible with gauge invariance. Charge conjugation is defined by

$$a(x), a^\dagger(x) \to (-1)^{x_1+x_2}a^\dagger(x), (-1)^{x_1+x_2}a(x) \quad (5.19)$$

$$\tilde{a}(x^*), \tilde{a}^\dagger(x^*) \to (-1)^{x_1^*+x_2^*}\tilde{a}^\dagger(x^*), (-1)^{x_1^*+x_2^*}\tilde{a}(x^*) \quad (5.20)$$

$$A_\mu(x) \to -A_\mu(x) \quad , \quad \tilde{A}_\mu(x^*) \to -\tilde{A}_\mu(x^*) \quad (5.21)$$

The gauge transformation in (5.19) and (5.20) are to make the charge transport operators

$$T_C(x) = a^\dagger(x)e^{iA_i(x)}a(x) \quad (5.22)$$

$$\tilde{T}_C(x^*) = \tilde{a}^\dagger(x^*)e^{i\tilde{A}_i(x^*)}\tilde{a}(x^*) \quad (5.23)$$

Charge-conjugarion invariant.

Both of the gauge generators in (5.18) are odd under this transformation

$$\mathcal{G} \to -\mathcal{G} \quad , \quad \tilde{\mathcal{G}} \to -\tilde{\mathcal{G}} \quad (5.24)$$

Also, if we require parity symmetry, $\tilde{A}$ must couple to a pseudo-vector current. We see this in the gauge generator $\tilde{\mathcal{G}}$ in (5.18). The magnetic field of the primary lattice is a pseudo-scalar density

so, in order for $\tilde{\mathcal{G}}$ to have a well-defined transformation under parity the charge density on the dual lattice must be a pseudo-scalar. This is accomplished by combining the parity and charge conjugation transforms of $\tilde{a}(x^*)$:

$$\tilde{a}'(x^{*\prime}) = (-1)^{x_1^* + x_2^*}\tilde{a}^\dagger(x^*) \quad , \quad \tilde{a}^{\dagger\prime}(x^{*\prime}) = (-1)^{x_1^* + x_2^*}\tilde{a}(x^*) \tag{5.25}$$

With this transformation the charge transport operator in (5.23) is parity invariant.

Gauge invariance of a physical model with charged particles requires the constraints $\mathcal{G} \sim 0$ and $\tilde{\mathcal{G}} \sim 0$ which are imposed as physical state conditions. Gauge invariant [under transformations where the gauge functions have compact support] creation and annihilation operators are formed by taking a product of a creation or annihilation operator at a point x or x^* and the appropriate Wilson line operator with curve C or $\tilde{C}$ going from the point x or x^* to some point at infinity,

$$\alpha_C(x) = W[C]a(x) \quad , \quad \tilde{\alpha}_{\tilde{C}}(x^*) = \tilde{W}[\tilde{C}]a(x^*) \tag{5.26}$$

These are multi-valued operators with multi-valuedness characterized by the holonomy in (5.14). Thus, for example, when we transport the operator $\alpha_C(x)$ around a loop C' we obtain

$$\alpha_{C+C'}(x) = \exp\left(ig \sum_{x^* \in \text{interior}C'} B(x^*)\right)\alpha_C(x)$$

$$\sim \exp\left(ig\tilde{g}\theta \sum_{x^* \in \text{interior}C'} \tilde{a}^\dagger(x^*)\tilde{a}(x^*)\right)\alpha_C(x) \tag{5.27}$$

where the last weak equality holds for matrix elements with physical states. Thus the change in phase of the wavefunction upon transport of an α-particle around a loop is given by $\theta g\tilde{g}$ times the number of $\tilde{\alpha}$-particles whose positions are linked by the loop. If $\theta g\tilde{g}/2\pi$ is a rational number (say M/N where M and N are relative prime integers) this holonomy can only take on a finite number of values which are the elements of the one-dimensional representation of the discrete group Z_N by phases $\exp 2\pi i Mn/N$.

We quantize the gauge fields using the functional Schroedinger picture where we treat the primary gauge field $A([x,\mu])$ as a coordinate so the wavefunctionals depend on classical configurations of $A([x,\mu])$ and represent the commutator (5.7) by taking

$$\tilde{A}([x,\mu]^*) \equiv -i\theta\frac{\delta}{\delta A([x,\mu])} \tag{5.28}$$

Then the empty vacuum state is given by*

$$\Psi_0[A] = |0\rangle\prod_{x^*}\delta(B(x^*)) \tag{5.29}$$

* Note that, as is characteristic of constrained systems, this state is not normalizable. When taking the norm one should properly view $\delta^2(B(x^*))$ as $\delta(B(x^*))\times$(volume of gauge group).

where $|0\rangle$ is the state with no charged particles $a(x)|0\rangle = 0 = \tilde{a}(x^*)|0\rangle$. From the empty vacuum we can create the basis states for the n-particle and $\tilde{n}$-particle sector of the theory,

$$\alpha^\dagger_{C_1}(x_1)\ldots\alpha^\dagger_{C_n}(x_n)\tilde{\alpha}^\dagger_{\tilde{C}_1}(x_1^*)\ldots\tilde{\alpha}^\dagger_{\tilde{C}_{\tilde{n}}}(x_{\tilde{n}}^*)\Psi_0 \tag{5.30}$$

In this state, we transport a particle in a loop, ℓ, by successive operation from the left with the gauge invariant operator $T([x,\mu]) = a^\dagger(x+\mu)e^{iA([x,\mu])}a(x)$. This effectively multiplies the state from the left by the Wilson line operator $W[\ell]$. We can use lattice Stoke's theorem to show that $W[\ell]\Psi_0 = 0$ so that the only effect of this transport resides in the commutators of $W[\ell]$ with $\tilde{W}[\tilde{C}_i]$. It is easy to show that the net phase is $\exp(ig\tilde{g}\theta\times$ (number of tilde particles linked by ℓ)). Also, if the particle at x_1 is exchanged with the particle at x_2 by transporting particle 1 along C_{12} and particle 2 along C_{21} then the wavefunction aquires a phase $\exp(ig\tilde{g}\theta($ number of tilde-particles linked by $)C_{12} \oplus C_{21})$. It can be shown that if we transport a tilde-particle in a loop or exchange two tilde-particles the state changes by a phase $\exp(-ig\tilde{g}\theta($ number of non-tilde-particles linked by the loop $))$. Note that the sign of the phase is opposite to that for the non-tilde particle. Thus the wave-function carries a one-dimensional unitary representation of a two-color braid group where we are allowed to make braids by interchanging components of the same color. The representation gives a phase proportional to the number of components of the opposite color linked by a braid (weighted by the orientation of the braid). It has recently been shown that a similar braiding charcaterizes the solitons for the 2+1-dimensional CP^1 model [Bergeron, Douglas and Semenoff, 1991].

Thus, we can find a local, parity invariant Chern-Simons theory. Very much in analogy with the case of lattice fermions we see that it is difficult to fit both locality and chirality on a lattice. For more details on this subject, see [Kantor and Susskind, 1990] and [Eliezer, Semenoff and Wu, 1991].

VI. Discussion

We have given a detailed review of the project one might call anyonization, the mapping of a Chern-Simons theory coupled to fermions or bosons onto a theory of anyons. We find that if $1/4\pi\alpha$ is the coefficient of the Chern-Simons term the parameter α governs the statistics of particles.

In the special case that α is an even integer, we map a Chern-Simons-fermion system onto a system of free fermions. In the minimal lattice Chern-Simons model discussed in Section IV, in the special case that α is an odd integer, ϕ_C and $\phi_{C'}$ are bosonic and again independent of C and C'. With the identification found in [Anderson, John, Baskaran, Doucot and Liang, 1988] (see also [Fradkin, 1989; Ambjorn and Semenoff, 1989]),

$$\phi(x) = S^1(x) - iS^2(x) \quad , \quad \phi^\dagger(x) = S^1(x) + iS^2(x) \tag{6.1}$$

$$\rho(x) = S^3(x) - 1/2 \tag{6.2}$$

they form the $j = 1/2$ representation of the lattice SU(2) algebra

$$[S^a(x), S^b(y)] = i\epsilon^{abc}S^c(x)\delta(x-y) \ , \ \sum_{a=1}^{3} S^a(x)S^a(x) = \frac{3}{4} \tag{6.3}$$

and the Hamiltonian given in equation (4.56) is the Hamiltonian of the X-Y model *

$$H = -\sum_{x,i} \left(S^1(x+i)S^1(x) + S^2(x+i)S^2(x)\right) \tag{6.4}$$

It has been prove by [Kennedy, Lieb and Shastry, 1988] and [Kubo and Kishi, 1988] that, when the volume is infinite, the ground state of the X-Y model in two dimensions has long-range order with

$$\lim_{x\to\infty} < S^1(x)S^1(0) > \neq 0 \tag{6.5}$$

We can pull back this result to the fermion-Chern-Simons theory discussed in Section IV. It implies that the ground state has off-diagonal long range order,

$$\lim_{x\to\infty} < \psi^\dagger(x) \prod_{\text{links}} e^{iA_i}\psi(0) > \neq 0 \tag{6.6}$$

This implies that the fermion-Chern-Simons theory system is a charged superfluid. If it were coupled to the physical electromagnetic field in addition to the statistical gauge field, at least for sufficiently weak coupling, this system would then be a superconductor. I t was shown by [Kubo and Kishi, 1988] that the long-range order persists in the antiferromagnetic XXZ model formed by adding the coupling

$$H_{33} = g\sum_{x,i} S^3(x+i)S^3(x) \tag{6.7}$$

or, in fermionic variables,

$$H_{33} = g\sum_{x,i} \left(\psi^\dagger(x+i)\psi(x+i)\psi^\dagger(x)\psi(x)\right) - 2gQ + \text{const.} \tag{6.8}$$

with g sufficiently small.

This is a repulsive interaction and for g large enough it destroys the superfluid condensate. For $g \to \infty$ the ground state of $H + H_{33}$ is that of H_{33} which is the antiferromagnetic Ising model. The ground state of the latter does not have off-diagonal long range order and is not a superfluid but breaks symmetry under translation by one lattice spacing,

$$< \sum_x (-1)^{\sum_i x_i} S^3(x) > = < \sum_x (-1)^{\sum_i x_i} (\rho(x) - 1/2) > \neq 0 \tag{6.9}$$

* The discrete symmetries of Chern-Simons theory map onto the XY model as follows: $(CP)_{CS} \longleftrightarrow (CP)_{XY}$, and $(T)_{CS} \longleftrightarrow (T)_{XY}$. Parity, which is broken in general for Chern-Simons theory, is not broken in the case that α is an integer.

Here, antiferromagnetism of the XXZ model coincides with a commensurate charge-density wave state of the Chern-Simons-fermion model and is stable when the repulsive self-interaction of the fermions is strong enough [Ambjorn and Semenoff, 1988].

Finally, we note that at the level of the fundamental fields on the lattice, the possibility of anyonization seems to depend quite sensitively on the detailed form of the Chern-Simons term. We do not regard this as a fundamental difficulty with this formalism. In fact, in a realistic system one might imagine performing a sort of block spin construction of a coarse-grained theory, starting from the elementary lattice theory that we have discussed. This construction generates an effective Hamiltonian which has both relevant and irrelevant operators. We would expect that the Chern-Simons term in the effective Hamiltonian is modified. However, we still know that this system describes anyons.

References

Y. Akutsu and M. Wadati, J. Phys. Soc. Japan 56 (1987), 3039; J. Phys. Soc. Japan 59 (1987), 3034

Y. Akutsu, T. Deguchi and M. Wadati, J. Phys. Soc. Japan 56 (1987), 3464.

J. Ambjorn and G.W. Semenoff, Phys. Lett. 226B (1989), 107.

P. W. Anderson, S. John, G. Baskaran, B. Doucot and S.-D. Liang, Princeton preprint 1988.

D. Arovas, R. Schrieffer, F. Wilczek and A. Zee, Nucl. Phys. B251, [FS13] (1985), 117.

E. Artin, Ann. Math. 48 (1947), 101

M. Bergeron, G. W. Semenoff and R. R. Douglass, 'Canonical Quantization of the CP^1 Model in 2+1-dimensions, University of British Columbia Preprint, 1991.

S. Coleman, Phys. Rev. D11 (1975), 2088.

S. Deser, R. Jackiw and S. Templeton, Phys. Rev. Lett. 48 (1982), 475; and Ann. Phys. (N.Y.) 140 (1982), 372.

P. A. M. Dirac, *Lectures on Quantum Mechanics*, Yeshiva Press, New York, 1969.

G. Dunne, R. Jackiw and C. Trugenberger, Ann. Phys. (N.Y.) 194 (1989), 197.

D. Eliezer and G. W. Semenoff 'Comment on the Jordan-Wigner Transformation..' , University of British Columbia Preprint, 1991.

D. Eliezer, G. W. Semenoff and S. S. C. Wu, 'Anyonization and Superfluidity of Lattice Chern-Simons Theory', University of British Columbia preprint, 1990.

D. Eliezer, G. W. Semenoff and S. S. C. Wu, 'Novel Braid Structure in Dumbell Chern-Simons Theory', University of British Columbia preprint, 1990.

E. Fadell and J. Van Buskirk, Duke Math. 29 (1962), 243

D. Finkelstein, Jour. Math. Phys. 7 (1966), 1218.

D. Finkelstein and J. Rubenstein, Jour. Math. Phys. 9 (1968).

S. Forte and T. Jolicoeur, Nucl. Phys. B350 (1991), 589.
R. Fox and L. Neuwirth, Math. Scand. 10 (1961), 119
E. Fradkin, Phys. Rev. Lett. 63 (1989), 322.
J. Froehlich and P. Marchetti, Comm. Math. Phys. 116 (1988), 127.
S.Girvin in *The Quantum Hall Effect*, ed. by R. Prange and S. Girvin, Springer-Verlag 1987
S. Girvin, A. MacDonald, M. Fischer, S.-J. Rey and J. Sethna, Phys. Rev. Lett. 65 (1990), 1671.
G. Goldin, R. Menikoff and D. Sharp, Phys. Rev. Lett. 22 (1981), 1664.
F. D. M. Haldane, Phys. Rev. Lett. 61 (1988), 1029.
C. R. Hagen, Phys. Rev. D31 (1985), 2135.
C. R. Hagen, Phys. Rev. Lett. 63 (1989), 1025.
Y. Hatsugai, M. Kohmoto and Y.-S Wu, Phys. Rev. Lett. 66 (1991), 659.
R. Jackiw and S. Templeton, Phys. Rev. D24 (1981), 2291
R. Jackiw and S.-Y. Pi, Phys. Rev. Lett. 64, 2969 (1990); Phys. Rev. D42, 3500 (1990).
R. Jackiw and E. Weinberg, Phys. Rev. Lett. 64, 2344 (1990).
R. Jackiw, E. Weinberg and K.-Y. Lee, Phys. Rev. D42, 3488 (1990).
P. Jordan and E. Wigner, Zeit. Phys. 47 (1928), 631.
R. Kantor and L. Susskind, 'A Lattice Model of Fractional Statistics', Stanford Preprint, 1990.
T. Kennedy, E. Lieb, and S. Shastry, Phys. Rev. Lett. 61 (1988), 2582
K. Kubo and T. Kishi, Phys. Rev. Lett. 61 (1988), 2585.
R. Laughlin, Phys. Rev. Lett. 50 (1983), 1395.
J. Leinaas and J. Myrlheim, Nuovo Cimento B37 (1977), 1
M. Luscher, Nucl. Phys. B326 (1989), 557.
J. Lykken, J. Sonnenschein and N. Weiss, Phys. Rev. D42 (1990) 2161.
S. Mandelstam, Phys. Rev. D11 (1975), 3026.
T. Matsuyama, Phys. Lett. B228 (1989), 99.
C.Morette-De Witt and **M.Laidlaw**, Phys. Rev. D3 (1971), 1375
V.F. Muller, Z.Phys.C47:301-310,1990.
V.F. Muller, On the connection between euclidean and hamiltonian lattice field theories of vortices and anyons, Kaiserslautern U. preprint PRINT-91-0127, 1991.
Y. Nambu, Prog. Theor. Phys. 5 (1950), 1
H. B. Nielsen and M. Ninomiya, Nucl. Phys. B185 (1981), 20; Nucl. Phys. B193 (1981), 173.
A. J. Niemi and G. W. Semenoff, Phys. Rev. Lett. 51, 2077, (1983).
Y. Ohnuki and S. Kamefuchi, *Quantum Field Theory and Parastatistics*, Springer Verlag, Berlin, 1982.
A. M. Polyakov, Mod. Phys. Lett. A3 (1988), 325

J. Schoenfeld, Nucl. Phys. B185 (1981), 157
L.S. Schulman, *Techniques and Applications of Path Integration*, Wiley, New York, 1981.
G. W. Semenoff, Phys. Rev. Lett. 61 (1988), 517.
G. W. Semenoff, Proceedings of *Physics, Geometry and Topology*, Banff Summer School 1989, H.C. Lee, G. Kunstatter and R.L.Kobes ed.
G. W. Semenoff and P. Sodano, Phys. Rev. Lett. 57 (1986), 1195.
G. W. Semenoff and P. Sodano, Nucl. Phys. B328 (1989), 753.
G. W. Semenoff, P. Sodano and Yong-Shi Wu, Phys. Rev. Lett. 62 (1989), 715
G. W. Semenoff and N. Weiss, Phys. Lett. B250 (1990), 117.
G. W. Semenoff and L.C.R. Wijewardhana, Phys. Lett. B184, 397 (1987)
T.H.R. Skyrme, Proc. Roy. Soc. Lond. A260 (1961), 127.
C.-H. Tze, Int. Jour. Mod. Phys. A3 (1988), 1959
F. Wilczek and A. Zee, Phys. Rev. Lett. 51 (1983), 2250
J.G. Williams, Jour. Math. Phys. 11 (1970), 2611
Yong-Shi Wu, Phys. Rev. Lett. 52 (1984), 2103;

PART V

MISCELLANEOUS

CLASSICAL STRING WITH RIGIDITY

H. Arodź

Institute of Physics, Jagellonian University, Cracow, Poland

ABSTRACT. Some improvements in Lagrangian and Hamiltonian formulation of classical string with rigidity are presented. We also descibe tachyonic solutions of classical equations of motion for a point-like analogue of the string with rigidity and for the string itself.

Address: Reymonta 4 , 30-059 Cracow, Poland.

1. INTRODUCTION

We shall consider a string in Minkowski space-time. Evolution of the string is described by a 2-dimensional surface $\Sigma = (x^\mu(\tau,\sigma))$ in the space-time, called the world-sheet of the string. Here τ is the evolution parameter and $\sigma\in[0,\pi]$ is a parameter along the string. On Σ we have a metric induced by the ordinary metric in Minkowski space-time. For $(x^\mu)\in\Sigma$ we have

$$dx^\mu = x^\mu{}_{,a}\,du^a \qquad \text{and} \qquad ds^2 = dx^\mu dx_\mu = g_{ab}\,du^a du^b ,$$

where $a=0,1$ $(u^0=\tau\ ,\ u^1=\sigma)$, $(x^\mu{}_{,a}) = \left(\frac{\partial x^\mu}{\partial u^a}\right)$ are four-vectors tangent to Σ , and

$$g_{ab} = \frac{\partial x^\mu}{\partial u^a}\frac{\partial x}{\partial u^b}{}_\mu \qquad (1)$$

is the world-sheet metric tensor. We assume that $\dot{x}^2 \geqslant 0$ and $x'^2 \leqslant 0$,where $\dot{x} = (x^\mu{}_{,0})$, $x' = (x^\mu{}_{,1})$. Then

$$g = \det(g_{ab}) \leqslant 0 . \qquad (2)$$

The fact that $\dot{x}$ and x' can be light-like follows from equations of motion for the string. This factcomplicates a little bit formulation of the action principle for the string.[1] This difficulty is absent if we consider strings with massive ends[2], however in that case other problems appear.

The famous Nambu - Goto action functional for the string has the form

$$S_{N-G} = -\gamma \int \sqrt{-g}\, d\tau d\sigma , \qquad (3)$$

where $\gamma > 0$ is a constant. This action defines the Nambu - Goto string, either closed ($x^\mu(\tau,0) = x^\mu(\tau,\pi)$) or open with massless ends. Lot of research work has been done on classical and quantised Nambu - Goto string, for a review and references see, e.g. the recent books[3].

In 1986 an interesting modification of the Nambu - Goto action has been proposed.[4] It is given by the formula

$$S = -\gamma\int\sqrt{-g}\,d\tau d\sigma + \alpha\int\sqrt{-g}\,\Delta x^{\mu}\Delta x_{\mu}\,d\tau d\sigma \quad , \qquad (4)$$

where

$$\Delta x^{\mu} = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial u^{a}}\left(\sqrt{-g}\,g^{ab}\,\partial_{a}x^{\mu}\right) \qquad (5)$$

is the Laplacian of $x^{\mu}(\tau,\sigma)$, which is regarded as a field on the world-sheet Σ, with respect to the inner metric g_{ab} on Σ. g^{ab} present in formula (5) are components of the inverse metric tensor on Σ ($g^{ab}g_{bc}=\delta^{a}_{c}$). The coefficient α in (4) is a constant. The squared-Laplacian term on the r.h.s. of formula (4) is called the rigidity term. It can be related to an integral over Σ of extrinsic curvature of Σ [4]. It has been argued that the string with rigidity is less bizarre than the Nambu - Goto string [4]. This expectation has not been confirmed as yet. If we regard the string with rigidity as a fundamental object we have to cope with many difficulties, like indefinitness of energy of the string [5], or tachyons [6].

However, strings are interesting objects not only because of their potential for providing a new fundamental theory. Strings are useful for an approximate description of more complicated physical systems. In such a context the rigidity term is quite natural. Its presence implies that the string has some intrinsic stiffness (for $\alpha>0$, as it turns out), that it is resistant against bending without stretching, in contradistinction to the Nambu - Goto string. In the four-dimensional view, the presence of the rigidity term implies that the world-sheet Σ tends to be a rather smooth surface. Therefore, it is not a surprise that the rigidity term has appeared in a string approximation to dynamics of vortices [7], or when fermionic degrees of freedom of a superstring have been integrated out [8]. The rigidity term has also been considered in statistical theory of surfaces, in order to suppress contributions of surfaces with narrow spikes [9]. In these applications the string with rigidity

should be regarded as an effective model. As such, it merits an investigation of its properties. Such investigations have been undertaken by several authors [10),11].

In this lecture I would like to present main results of works [12),13),14] done in collaboration with A.Sitarz and P.Węgrzyn. We have found a new class of solutions to the equations of motion of the classical string with rigidity [12),13]. The new solutions are tachyonic in the sense that $P_{\mu}P^{\mu}<0$, where P^{μ} is the energy-momentum four-vector of the string (but no part of the string moves with a superluminal velocity). The classical Nambu - Goto string is not tachyonic - it becomes tachyonic after quantisation. The classical tachyonic solutions are of the "run-away" type, known to occur in theories with higher order derivatives.

We also have clarified some points in Lagrangian and Hamiltonian formulations of the string with rigidity [14]. In particular, we point out the importance of some boundary terms in the case of the open string with rigidity. These terms, when taken into account, cancel certain contributions to integrals of motion. The boundary terms are also responsible for a rather unexpected fact that the conserved energy of the string is not equal to the Hamiltonian of the string even if we eliminate constraints.

2. EQUATIONS OF MOTION AND CONSERVED QUANTITIES

Let us consider a general variation

$$x^{\mu}(\tau,\sigma) \to x^{\mu}(\tau,\sigma) + \delta x^{\mu}(\tau,\sigma) \quad . \qquad (6)$$

Under this variation the action

$$S = \int_{R} d^{2}u\, L(x^{\mu}, x^{\mu}_{,a}, x^{\mu}_{,oo}, x^{\mu}_{,o1}, x^{\mu}_{,11}) \qquad (7)$$

changes by

$$\delta S = \int_{R} d^{2}u\left[\frac{\partial L}{\partial x^{\mu}}\delta x^{\mu} + \frac{\partial L}{\partial x^{\mu}_{,a}}\delta x^{\mu}_{,a} + \frac{\partial L}{\partial x^{\mu}_{,oo}}\delta x^{\mu}_{,oo} + \frac{\partial L}{\partial x^{\mu}_{,o1}}\delta x^{\mu}_{,o1} + \frac{\partial L}{\partial x^{\mu}_{,11}}\delta x^{\mu}_{,11}\right]$$

where R is the rectangle $0 \leq \sigma \leq \pi$, $\tau_1 \leq \tau \leq \tau_2$. We have taken into account the fact that $x,_{01} = x,_{10}$. Integrating by parts we can cast δS into the form

$$\delta S = \int_R d^2u \left[R_\mu \delta x^\mu + \varepsilon^{ab} \partial_a Y_b + \partial_0 \partial_1 Z \right], \tag{8}$$

where

$$R_\mu = \partial_0^2 \left(\frac{\partial L}{\partial x,^\mu_{00}}\right) + \partial_0 \partial_1 \left(\frac{\partial L}{\partial x,^\mu_{01}}\right) + \partial_1^2 \left(\frac{\partial L}{\partial x,^\mu_{11}}\right) - \partial_a \left(\frac{\partial L}{\partial x,^\mu_{a}}\right) + \frac{\partial L}{\partial x^\mu}, \tag{9}$$

$$Y_0 = \left[-\frac{\partial L}{\partial x,^\mu_1} + \partial_a \left(\frac{\partial L}{\partial x,^\mu_{a1}}\right) \right] \delta x^\mu - \frac{\partial L}{\partial x,^\mu_{11}} \delta x,^\mu_1 \quad , \tag{10}$$

$$Y_1 = \left[\frac{\partial L}{\partial x,^\mu_0} - \partial_a \left(\frac{\partial L}{\partial x,^\mu_{0a}}\right) \right] \delta x^\mu + \frac{\partial L}{\partial x,^\mu_{00}} \delta x,^\mu_0 \quad , \tag{11}$$

$$Z = \frac{\partial L}{\partial x,^\mu_{01}} \delta x^\mu \ . \tag{12}$$

With the help of Stokes theorem we finally obtain

$$S = \int_R d^2u R_\mu \delta x^\mu + \int_{\partial R} Y_a du^a + Z(\tau_2,\pi) - Z(\tau_2,0) - Z(\tau_1,\pi) + Z(\tau_1,0) \ , \tag{13}$$

where ∂R is the boundary of the rectangle R. This form of δS contains the least number of derivatives of δx^μ. The remaining derivatives cannot be removed by any integrations by parts. Notice the "corner" terms (i.e.the Z-terms) on the r.h.s. of formula (13). In the case of closed string they cancel each other, however in the case of open string they give a contribution to integrals of motion.

Equations of motion for the open rigid string follow from the requirement $\delta S=0$ with respect to variations δx^μ obeying the following conditions

$$\delta x^\mu(\tau,\sigma) = 0 = \delta \dot{x}^\mu(\tau,\sigma) \quad \text{for } \tau=\tau_1,\tau_2 \ , \ \sigma \in [0,\pi] \ .$$

Then, it follows from formulae (8-11) that

$$R_\mu = 0 \ , \tag{14}$$

$$-\frac{\partial L}{\partial x,^\mu_1} + \partial_a \left(\frac{\partial L}{\partial x,^\mu_{1a}}\right) = 0 \quad \text{for } \sigma = 0,\pi \ , \tag{15}$$

$$\frac{\partial L}{\partial x,^\mu_{11}} = 0 \quad \text{for } \sigma=0,\pi \ . \tag{16}$$

Equations (15),(16) are usually called the boundary conditions. In the case of Lagrangian (4) equation (14) has the following form

$$(\gamma - \alpha \Delta x_\sigma \Delta x^\sigma) \Delta x_\mu + 2\alpha \left[\Delta(\Delta x_\mu) - g^{ab} x,^\sigma_a x,^\mu_b \Delta(\Delta x_\sigma) \right] - 4\alpha g^{ab} g^{cd} (\Delta x_\sigma),_b x,^\sigma_c \nabla_a x_{\mu,d} = 0, \tag{17}$$

where ∇_a denotes the covariant derivative with respect to the metric (1). One can show that the l.h.s. of Eq. (17) has identically vanishing projections on $x,^\mu_a$. These identities are Noether identities following from invariance of the action (4) under reparametrizations $(\tau,\sigma) \to (\tau',\sigma')$ of Σ.

In the case of action (4), in orthonormal coordinates (defined by the requirements $\dot{x}\, x' = 0$, $\dot{x}^2 + x'^2 = 0$), the boundary conditions (15),(16) have the following form

$$\gamma x'_\mu + 2\alpha \partial_1 (\Box x_\mu) = 0 \quad \text{for } \sigma=0,\pi \ , \tag{15'}$$

$$\Box x_\mu \equiv \ddot{x}_\mu - x''_\mu = 0 \quad \text{for } \sigma=0,\pi. \tag{16'}$$

It is easy to prove from here that

$$\dot{x}^2 = 0 \quad \text{for } \sigma = 0,\pi, \tag{18}$$

i.e. the ends of the open string with rigidity move with the velocity of light.

Conserved energy and momentum of the string follow from the fact that $\delta S = 0$ under variations $\delta x^\mu = \varepsilon^\mu$, where ε^μ are constant. In this case our "master formula" (13) gives

$$\delta S = P_\mu \varepsilon^\mu \Big|_{\tau=\tau_2} - P_\mu \varepsilon^\mu \Big|_{\tau=\tau_1} = 0 \ ,$$

where we have assumed that equations of motion (14-16) are satisfied , and

$$P_\mu = \int_0^\pi d\sigma\, p_\mu , \tag{19}$$

with

$$p_\mu = -\frac{\partial L}{\partial \dot{x}^\mu} + \partial_o\left(\frac{\partial L}{\partial \ddot{x}^\mu}\right) . \tag{20}$$

p_μ can be regarded as the density of energy-momentum of the string.

Similar calculation gives the following formula for the conserved angular momentum of the string

$$M_{\mu\nu} = \int_0^\pi d\sigma\,(x_\mu p_\nu - x_\nu p_\mu) + \int_0^\pi d\sigma\left(\frac{\partial L}{\partial x^\mu_{,a}}\, x_{\nu,a} - \frac{\partial L}{\partial x^\nu_{,a}}\, x_{\mu,a}\right).$$

We have presented our derivations of the equations of motion, energy-momentum and angular momentum because one can find in literature formulae which differ from the formulae we have obtained in the case of the open string.

3. HAMILTONIAN FORMULATION

In this Section we shall use the physical gauge, defined by the condition $x^o(\tau,\sigma) = \tau$, i.e. the evolution parameter is chosen to be the physical time. In this manner we avoid a discussion of constraints and of related complications in canonical formalism.

In this gauge the conserved energy P_o and the conserved momentum P_i of the string are given by the following formulae[13]

$$P_o = \int_0^\pi d\sigma\left\{\ddot{\vec{x}}\,\frac{\partial L}{\partial \ddot{\vec{x}}} + \dot{\vec{x}}\left[\frac{\partial L}{\partial \dot{\vec{x}}} - \partial_o\left(\frac{\partial L}{\partial \ddot{\vec{x}}}\right)\right] + \dot{\vec{x}}'\frac{\partial L}{\partial \vec{x}'} - L\right\}, \tag{21}$$

$$P_i = \int_0^\pi d\sigma\left[-\frac{\partial L}{\partial \dot{x}^i} + \partial_o\left(\frac{\partial L}{\partial \ddot{x}^i}\right)\right] . \tag{22}$$

On the other hand, the canonical momenta Π_{1i}, Π_{2i} and the Hamiltonian H are given by the formulae[11]

$$\Pi_{1i} = -\frac{\partial L}{\partial \dot{x}^i} + \partial_a\left(\frac{\partial L}{\partial \dot{x}^i_{,a}}\right) , \tag{23}$$

$$\Pi_{2i} = -\frac{\partial L}{\partial \ddot{x}^i} , \tag{24}$$

$$H = \int_0^\pi d\sigma\,(-\ddot{x}^i \Pi_{2i} - \dot{x}^i \Pi_{1i} - L) . \tag{25}$$

Comparing formulae (23-25) with (21-22) we see that $P_i \neq \int_0^\pi d\sigma\, \Pi_{1i}$, and, more interestingly,

$$H = P_o - \dot{x}^i \frac{\partial L}{\partial \dot{x}^i_{,1}}\Bigg|_{\sigma=0}^{\sigma=\pi} . \tag{26}$$

Thus, in general $H \neq P_o$, and H is not constant during the motion of the string, which is a surprising fact because the string is an isolated system.

Explanation of this fact is provided by the observation that in field-theoretical models defined on a strip, Hamilton equations of motion contain some boundary terms. In our case the fields are given by $x^i(t,\sigma)$, and the strip is defined by $-\infty < t < +\infty$, $0 \le \sigma \le \pi$. We have found[13] that for a generic dynamical variable of the form

$$\int_0^\pi d\sigma\, F(q_1^i, \dot{q}_1^i, q_1'^i, q_2^i, \dot{q}_2^i, \pi_1^i, \pi_2^i) ,$$

where $q_1^i \equiv x^i$, $q_2^i \equiv \dot{x}^i$, the Hamilton equations of motion have the following form

$$\frac{dF}{dt} = \{F, H\} + \left(\frac{\partial F}{\partial q_1'} - \partial_1\left(\frac{\partial F}{\partial q_1''}\right)\right)\dot{q}_1\Bigg|_{\sigma=0}^{\sigma=\pi} + \frac{\partial F}{\partial q_1''}\,\dot{q}_1'\Bigg|_{\sigma=0}^{\sigma=\pi} + \frac{\partial F}{\partial q_2'}\,\dot{q}_2\Bigg|_{\sigma=0}^{\sigma=\pi} , \tag{27}$$

where $\{\cdot,\cdot\}$ denotes the Poisson bracket. Thus, in general, one should not expect that $dH/dt = 0$. The Poisson bracket $\{H,H\}$

of course vanishes, but the boundary terms do not.

Notice that a boundary term is present also in the case of open Nambu - Goto string. From formula (27) we see that in this case

$$\frac{dF}{dt} = \{F,H\} + \frac{\partial F}{\partial q_1'^i}\dot{q}_1^i \Big|_{\sigma=0}^{\sigma=\pi} . \qquad (28)$$

4. POINT-LIKE COUNTERPART OF THE STRING WITH RIGIDITY

Let us consider the following Ansatz for the functions $x^\mu(\tau,\sigma)$:

$$x^\mu(\tau,\sigma) = \xi^\mu(\tau) + n^\mu\cdot\sigma , \qquad (29)$$

where n^μ is a constant four-vector such that $n^2=-1$ and $\dot{\xi}^\mu(\tau)n_\mu=0$. We also assume that $\dot{\xi}^2=1$ - this amounts to a gauge fixing for the reparametrizations $\tau\to\tau'(\tau)$. Notice that x^μ given by formula (29) obeys the requirements for the orthonormal coordinates: $\dot{x}\,x'=0$, $\dot{x}^2+x'^2=0$.

Inserting formula (29) into Eq.(17) we obtain the following equation for $\xi^\mu(\tau)$

$$2\alpha\left[\xi^{(4)}_\mu - (\xi^{(4)}_\sigma\dot{\xi}^\sigma)\dot{\xi}_\mu\right] + (\gamma + 3\alpha\ddot{\xi}^2)\ddot{\xi}_\mu = 0 , \qquad (30)$$

where $\xi^{(4)} \equiv d^4\xi/d\tau^4$. This equation is consistent with the conditions $\dot{\xi}n=0$ and $\dot{\xi}^2=1$.

In order to solve the string equations of motion by the Ansatz (29) we would have to obey not only Eq.(17) but also the boundary conditions (15′),(16′). It is easy to see that (15′), (16′) are satisfied only if $\gamma=0$ and $\ddot{\xi}=0$, which is not so interesting. The point-like counterpart of the string with rigidity is defined by its equation of motion, which by assumption has the form (30). We drop altogether the boundary conditions (15′),(16′).

Such a point particle has been considered by several authors.[15),12)] It has been shown that it possesses tachyonic trajectories, i.e. trajectories for which the conserved energy-momentum four-vector, given by the following formula

$$P_\mu = (\gamma + 3\alpha\ddot{\xi}^2)\dot{\xi}_\mu + 2\alpha\,\xi^{(3)}_\mu , \qquad (31)$$

is space-like, i.e. $P_\mu P^\mu<0$. The trajectories are of the "run-away" type. An example of the tachyonic trajectory can be explicitely given.[12)] For large $t=\xi^0$ the tachyonic trajectory $\vec{\xi}(t)$ approaches the following asymptotics

$$\vec{\xi}_{as}(t) = \mathrm{sign}(\alpha)\,\frac{\vec{P}}{|\vec{P}|}\,t . \qquad (32)$$

Notice a funny fact that for $\alpha<0$ the velocity $\dot{\vec{\xi}}$ is antiparallel to the conserved momentum $\vec{P}$. The velocity of the particle on the tachyonic trajectory is subluminal, $(\dot{\vec{\xi}}(t))^2<1$.

One can also show[12)] that the model, defined by the equation of motion (30), can be equivalently regarded as a family of ordinary, single point particle models. The family is parametrized by values of the conserved energy-momentum P_μ and of angular momentum $M_{\mu\nu}$ of the point-like counterpart of the string with rigidity. For any fixed values of P_μ and $M_{\mu\nu}$ we have the ordinary, single point particle which obeys Newton-Lorentz equation of motion

$$2\alpha\,\dddot{\xi}_\mu = F_{\mu\nu}(\xi)\,\dot{\xi}^\nu , \qquad (33)$$

where

$$F_{\mu\nu}(x) = M_{\mu\nu} + P_\mu x_\nu - P_\mu x_\nu . \qquad (34)$$

$F_{\mu\nu}(x)$ can be regarded as an external, time-dependent electromagnetic field (fictitious, of course).

The presence of tachyonic trajectories of the point-like analogue of the string with rigidity suggests that also the string can have tachyonic classical trajectories.

5. TACHYONIC TRAJECTORIES OF THE STRING WITH RIGIDITY

The classical equation of motion for the string with rigidity, i.e. Eq.(17), is very complicated. It is not known whether it can be linearized by an appropriate choice of coordinates τ,σ. The orthonormal coordinates which do the job in the case of Nambu - Goto string do not linearize Eq.(17).

It is easy to see that each solution of the equation of motion for the Nambu - Goto string,

$$\Delta x^{\mu} = 0 \tag{35}$$

also obeys Eq.(17). Only few explicit solutions of Eq.(17) such that they are not also solutions of Eq.(35) are known: a static circle of the radius $R_o=\sqrt{\alpha/\gamma}$ (for $\alpha>0$), expanding or shrinking circle, and an oblate loop rotating about its minor axis. These solutions have been found by E.Braaten, T.L.Curtright, G.I.Ghandour, C.B.Thorn and C.K.Zachos.[5),10)]

We shall again work using the physical gauge $\tau=x_o=t$. Then, the conserved energy and momentum are given by formulae (21),(22), with the Lagrangian L defined by formula (4). In the particular case of the following initial data

$$\vec{x}(0,\sigma)=\vec{f}(\sigma),\quad \dot{\vec{x}}(0,\sigma)=\ddot{\vec{x}}(0,\sigma)=0,\quad \dddot{\vec{x}}(0,\sigma)=\vec{h}(\sigma), \tag{36}$$

the conserved energy and momentum are given by

$$P_o = \int_0^{\pi} d\sigma\ (\vec{f}')^2\left\{\gamma+\frac{\alpha}{(\vec{f}')^2}\left[(\vec{f}')^2-\frac{(\vec{f}''\vec{f}')^2}{(\vec{f}')^2}\right]\right\}, \tag{37}$$

$$\vec{P} = 2\alpha\int_0^{\pi} d\sigma\ (\vec{f}')^2\left[\vec{h}-\frac{(\vec{f}'\vec{h})\vec{f}'}{(\vec{f}')^2}\right].$$

It follows from these formulae that if the component of $\vec{h}$ perpendicular to $\vec{f}'$ is sufficiently large, we can have arbitrarily large $|\vec{P}|$ while the energy P_o has the fixed value. Thus, $P_\mu P^\mu = P_o^2 - \vec{P}^2$ can be negative, what corresponds to the tachyon.

We shall consider the following Ansatz for the closed string with rigidity

$$\vec{x}(t,\sigma) = \begin{pmatrix} R(t)\cos\sigma \\ R(t)\sin\sigma \\ z(t) \end{pmatrix}, \tag{38}$$

where now $\sigma\in[0,2\pi]$ - for convenience we have changed the range of the parameter σ. For this Ansatz, $\ddot{\vec{x}}$ is perpendicular to $\vec{x}'$ what ensures that the tachyonic trajectory can be obtained with the help of the Ansatz (38). This Ansatz reduces the Lagrangian of the string with rigidity to the following form

$$L = -(\gamma R+\frac{\alpha}{R})\sqrt{1-\dot{R}^2-\dot{z}^2} - \frac{\alpha}{R}\frac{\dot{R}^2}{(1-\dot{R}^2-\dot{z}^2)^{1/2}}$$
$$-\frac{\alpha R}{(1-\dot{R}^2-\dot{z}^2)^{5/2}}\left[(1-\dot{z}^2)\ddot{R}^2+(1-\dot{R}^2)\ddot{z}^2+2\dot{R}\dot{z}\ddot{R}\ddot{z}\right] \tag{39}$$
$$-\frac{2\alpha}{(1-\dot{R}^2-\dot{z}^2)^{3/2}}\left[(1-\dot{z}^2)\ddot{R}+\dot{R}\ddot{z}\right].$$

The full set of equations of motion for the closed string reduces to the equations for R(t),z(t) obtained from Lagrangian (39) as Euler - Lagrange equations. This is due to the rotational symmetry of the Ansatz (38).

The equations for R(t),z(t) are of fourth order with respect to the time derivative. We find it more convenient to use Hamiltonian formulation of the system defined by Lagrangian (39). This formulation can be found with the help of the standard method.[16)] Then, after some calculations[14)] we obtain the following set of equations

$$2\alpha R\dddot{z} = M^{3/2}(p_1 t - E\cdot z + d_o),$$
$$2\alpha R\ddot{R} = M^{3/2}(L - E\cdot R + d_1), \tag{40}$$
$$\dot{L} = -2\gamma M^{1/2} + \frac{1-\dot{z}^2}{R^2}(L - E\cdot R + d_1) + \frac{1}{R}(E - p_1\dot{z} - \dot{L}\cdot\dot{R})$$
$$+\frac{\dot{R}\dot{z}}{R^2}(p_1 t - E\cdot z + d_o),$$

where L(t) is an auxilliary variable, $M\equiv 1-\dot{z}^2-\dot{R}^2$, and E, d_o, d_1, p_1

are constants determined by the initial data (36). Equations (40) are of the Newton type. They are quite regular, so we do not expect any problems with existence of solution for the given set of initial data.

Equations (40) are much simpler than the original equation (17). Nevertheless, they are fairly complicated too. In fact, we can make further progress only using numerical methods. Details of these computations will be given elsewhere.[14)]

6. ENDING REMARKS

One may ask about the significance of tachyonic trajectories of the string with rigidity. First, because the string is an approximation to a vortex in a field-theoretical model,[7)] we may expect that there exists a corresponding solution in the field-theoretical model. Thus, the dynamics of the string is an approximation to the dynamics of vortices. Because vortices are quite complicated configurations of fields, a direct investigation of their dynamics in the original field-theoretical model is almost impossible. In this context, the fact that our solution is tachyonic is not so important, because the corresponding field -theoretical solution can be nontachyonic. What matters here is that it is a new, non-trivial solution.

The fact that our solution is tachyonic might be of some importance while regarding the string with rigidity as a fundamental object. Then, such string should probably be quantised, and the quantised theory should have a physical mass spectrum. This requirement forbids tachyons in the quantised theory. If there is no quantisation which would make unimportant our classical negative-mass-squared trajectories, the model is not suitable for a new fundamental theory. Anyway, the quantisation procedure would also have to cure the model from the indefiniteness of the energy.[5)]

We think that the most interesting problem to be investigated in future is the correspondence between the vortices and the

strings. The string with rigidity is not a satisfactory model because of the indefiniteness of energy and the tachyons. The field-theoretical models which have vortex solutions do not have such flaws. Therefore, one may expect that there exists a better string model, more accurately modelling the dynamics of vortices. It is rather improbable that such a string model would be just the Nambu - Goto string. We think that in this context higher derivatives are unavoidable.

REFERENCES

[1] Salisbury D.C. and Sundermeyer K., Nucl. Phys. B191, 260 (1981).

[2] Barbashov B.M. and Nesterenko V.V., Introduction to the relativistic string theory. World Scientific, Singapore, 1990.

[3] Green M.B., Schwarz J.H. and Witten E., Superstring theory, vol. 1, 2. Cambridge University Press, 1987.
Brink L. and Henneaux M., Principles of string theory. Plenum Press, New York and London, 1988.
Kaku M., Introduction to superstrings. Springer, Berlin, 1988.

[4] Polyakov A.M., Nucl. Phys. B268, 406 (1986).

[5] Braaten E. and Zachos C.K., Phys. Rev. D35, 1512 (1987).

[6] Olesen P. and Yang S.K., Nucl. Phys. B283, 73 (1987).
Pisarski R.D., Phys. Rev. D38, 578 (1988).

[7] See, e.g. Barr S.M. and Hochberg D., Phys. Rev. D39, 2308 (1989); Silveira V., Phys. Rev. D41, 1914 (1990); Arodź H., Dobrowolski T. and Węgrzyn P., in preparation.
However, see also the paper by R. Gregory, Phys. Rev. D43, 520 (1991), where an approach is presented in which the rigidity term can appear only in higher order approximations, which is not so interesting.

[8] Kavalov A. and Sedrakyan A., Phys. Lett. B182, 33 (1986).

[9] There are many papers in this direction. See, e.g. Peliti L. and Leibler S., Phys. Rev. Lett. 54, 1690 (1985); Kleinert H., Phys. Lett. B174, 335 (1986); Ambjørn J., Acta Phys. Pol. B20, 313 (1989).

[10] Alonso F. and Espriu D., Nucl.Phys.B283, 393 (1987); Curtright T.L., Ghandour G.I., Thorn C.B. and Zachos C.K., Phys.Rev.Lett.57, 799 (1986) and Phys.Rev.D34, 3811 (1986).

[11] Nesterenko V.V. and Nguyen Suan Han, Int.J.of Mod.Phys.A3, 2315 (1988).

[12] Arodź H., Sitarz A. and Węgrzyn P., Acta Phys.Pol.B20, 921(1989).

[13] Arodź H., Sitarz A. and Węgrzyn P., preprint TPJU 1/91, January 1991.

[14] Arodź H., Sitarz A. and Węgrzyn P., in preparation.

[15] See, e.g. Pisarski R.D., Phys.Rev.D34, 670 (1986); Pavšič M., Phys.Lett.B205, 231 (1988); Nesterenko V.V., J.Phys A(Math&Gen) 22, 1673 (1989); Plyushchay M.S., Int.J.of Mod.Phys.A4, 3851 (1989); Grundberg J., Isberg J., Lindström U. and Nordström H., Phys.Lett.B231, 61(1989).

[16] Ostrogradski M., Mem.Acad.St.Petersburg 6, 385(1850).

The Unified Approach to Integrable Relativistic Equations: Soliton Solutions over Non-vanishing Backgrounds[1]

Barashenkov I.V.*, Getmanov B.S.*, Kovtun V.E.†

* *Joint Institute for Nuclear Research, LCTA, Head Post Office, P.O.Box 79, Moscow 101000, Russia*

† *Kharkov State University, Dzerjinski Sq.4, Kharkov, the Ukraine*

Abstract

The scheme for unified description of integrable relativistic massive systems provides an inverse scattering formalism that covers universally all $(1+1)$- dimensional models of this kind. In this work we construct the N-soliton solution (over an arbitrary background) for some generic system which is associated with the $sl(2,C)$ case of the scheme and whose reductions include the complex sine-Gordon equation, the massive Thirring model *etc.*, both in the Euclidean and Minkowskian spaces. Thus the N-soliton solutions for all these systems emerge in a unified form differing only in the type of constraints imposed on their parameters. In an earlier paper the case of the zero background was considered while here we concentrate on the case of the non-vanishing constant background i.e., on the *N-kink* solutions. Here the reduction to the new complexification of the sine-Gordon equation defined by

$$L=\frac{|\partial_\mu\phi|^2}{1-|\phi|^2}+|\phi|^2-\frac{1}{2}\frac{J_\mu^2}{|\phi|^2(1-|\phi|^2)}$$

with $J_\mu=i(\phi^*\partial_\mu\phi-\phi\partial_\mu\phi^*)$, appears to be the most interesting one as it exhibits decays and fusion of (subluminal) solitons.

[1] This is a preliminary version of an article to be submitted elsewhere.

1. Introduction

In recent years a considerable amount of interest has been paid to models of the $(1+1)$-dimensional Lorentz-invariant field theory possessing non-trivial integrability properties; a large number of integrable relativistic systems is known nowadays. On one hand, these systems serve as a nice laboratory for the more realistic $(3+1)$-dimensional theories and on the other one, the two-dimensional scalar and spinor equations are interesting in their own right, in connection mainly with the condensed matter applications. Since the suitable version of the Inverse Scattering method depends essentially on the structure of the associated linear system, each equation in these series was given, as a rule, an independent treatment. However, it would be highly desirable to have some unifying formalism which could allow one to analyse all the models from the common viewpoint. This sort of a formalism would be extremely helpful in revealing the hidden relations between systems that look unrelated superficially, and realizing the origin of distinctions in the dynamics (e.g. of solitons, particle-like singularities etc.) described by different relativistic equations. Lastly, it would provide the basis for the *classification* of integrable models of the field theory.

The crucial step in this direction was done presumably by Pohlmeyer [1] whose classification scheme covers such important classes of models as n-fields with values on spheres of arbitrary dimension. (He also revealed a natural connection between n-fields and unconstrained massive models, including the sine-Gordon model and its generalizations.) Another successful attempt to consider relativistic systems from the unified point of view belongs to Budagov and Takhtajan [2]. These authors have constructed an integrable generalization of the sine-Gordon model known as the "$U-V$" system.

An important advance was connected with the work of Zakharov and Mikhailov [3] which, in its turn, was based on the general Zakharov and Shabat's approach to the construction of integrable systems [4]. They have shown that each Lorentz-invariant integrable system can be represented as the compatibility condition for a pair of matrix linear equations depending rationally on a complex parameter λ. An important concept of gauge equivalence has also been introduced there, revealing the hidden correspondences between systems that were previously considered unrelated and allowing to group them into gauge-equivalent classes.

The authors of ref.[3] have limited themselves to the simplest case of the single pole λ-dependence of the linear problem matrices. In this case either the aforementioned $U-V$ systems emerge or the chiral fields (= free fields with values in homogeneous spaces.) It is fitting to note here that the density of the Hamiltonian of the chiral fields does not depend on coordinates and vanishes after the constant part subtraction. As it can be easily verified, this property is an attribute of *all* conformal-invariant (massless) systems in $(1+1)$-dimensions. It is nothing but the consequence of the special form of the energy - momentum conservation law which

looks like $\partial j_1/\partial\eta = 0$, $\partial j_2/\partial\xi = 0$ (where η and ξ are light-cone variables) and so j_1 and j_2 are arbitrary functions of ξ and η, respectively. In the case of an *integrable* system we can, following Pohlmeyer [1] [5] (see also [3] [6]) set $j_1(\xi) = C_1$, $j_2(\eta) = C_2$ and, breaking thereby the conformal invariance, reduce the system to the gauge-equivalent *massive* model, which does not suffer from the zero Hamiltonian drawback.[3] However, in this case another problem is met. Namely, it becomes nontrivial to demonstrate that the equations of motion follow from some least action principle, and this is a serious shortage from the field-theoretic viewpoint. (For example, the question of soliton masses remains open.) In the same way, it remains unclear whether the $U-V$ systems can be derived from a Lagrangian [2].

Another branch of generalizations was initiated by Mikhailov's work [7] who has found an integrable two-dimensional generalization of the classical periodic Toda chains. The problem of enumeration and classification of all possible $(1+1)$-dimensional relativistic generalizations of the Toda chains was solved by Mikhailov, Olshanetski and Perelomov [8] and Fordy and Gibbons [9]. In these papers the classification of nonlinear systems was reduced to the classification of the simple Lie algebras. Leznov and Saveliev obtained the solution of the analogous problem in the case of the finite unclosed (non-periodic) chains [10].

The spinor models were also immersed into the Inverse Scattering formalism. Here we refer to to the paper of Zakharov and Mikhailov [11] where the gauge equivalence was established between the chiral fields and the *massless* spinor systems (such as the Gross – Neveu, Vaks – Larkin – Nambu – Jona-Lasinio etc.) The integrability of several very important particular models in this series was discovered earlier by Neveu and Papanicolaou [6].

The present paper is concerned with the *massive* case. In this case the classification scheme was proposed by one of the authors [12], the so-called Unified Integrable Lorentz Fields (UNILOF) description scheme. This scheme covers all massive two-dimensional relativistic integrable systems known up until now, and allows to construct several new ones [12] [14]. The key point of the UNILOF scheme is the choice of the special, *triangular* gauge which retains the initial $\eta \leftrightarrow \xi$ symmetry of the linear problem and provides the unified description of both scalar and spinor models. In addition to the unification, this choice results in that all the emerging nonlinear equations are manifestly Lagrangian.

In nonlinear field theories, an especial role is played by particle-like solutions. Accordingly, Zakharov-Shabat's scheme of finding and classifying integrable systems was accompanied by the corresponding dressing procedure i.e., the method of obtaining multisoliton solutions which is the most suitable within this approach. In a similar way, we developed a natural dressing procedure for the UNILOF scheme [13] [14]. In those refs. we confined ourselves to the N-soliton solutions propagating over the *zero* background; the solitons' dynamics was found to be trivial. In the

[3] In [3] this procedure is misleadingly referred to as the solution of constraints.

present paper we construct multisoliton solutions over an *arbitrary* background.

Given some integrable equation, it seems natural to expect that the character of solitons' interaction does not depend on whether the background is zero or say, flat non-zero and is determined solely by the equation itself (i.e., by the type of dispersion and nonlinearity.) Surprisingly, our results refute this intuitively appealing idea. We show that, turning to the case of the non-zero backgrounds, the dynamics of solitons becomes non-trivial. Namely, in contrast to the solitons with the vanishing boundary conditions, the solitons travelling over e.g. flat non-zero background may interact *inelastically*.

As in ref.[14], we restrict ourselves to the sl(2,C) case of the UNILOF scheme here, and this fact makes the occurrence of the non-trivial dynamics even more unexpected. Indeed, the corresponding nonlinear system seems to be the first example of integrable equation associated with $rank = 1$ algebra, whose solitons interact inelastically[4]. Another restriction is that we exclude the so-called degenerate case when some of the linear problem matrices become singular [14]. This case pertains to the two-dimensional Toda lattices that were examined earlier [7], [8], [9].

The outline of the paper is as follows. In the next section we introduce the generic system of four first-order (or equivalently, two second-order) equations for 4 (respectively, 2) complex fields, which is the most general integrable system associated with the linear problem in question. (Hereafter this generic system will be referred to as the $\mathcal{G} = sl(2,C)$ system, or merely $\mathcal{G}$-system.) In Sec. 3 the multisoliton solution for the $\mathcal{G}$-system is derived. This solution corresponds to an arbitrary background and is explicit up to the defining of the seed Ψ-function, i.e., the solution of the linear problem pertaining to the background solution of the $\mathcal{G}$-system. In several physically important cases the seed Ψ-function can be found in a closed form thereby providing entirely explicit N-soliton solutions. Here we perform this computation for the constant non-zero background which is our main objective, and for the exponential background (Sec. 4, 5). (The need for the latter will be explained few lines below.)

Specifying the transformation properties of the fields entering in the $\mathcal{G}$-system with respect to the Lorentz transformations produces two models of the relativistic field theory. The first one is the model of 2 spinor fields while the second one comprises 2 complex scalar fields. Reducing the number of fields by imposing symmetry constraints, we are led to simpler systems which include, in particular, such important examples as the massive Thirring model and the complex sine-Gordon equation. Solutions of these reduced systems can be isolated from those of the $\mathcal{G}$-system by imposing suitable constraints on solitons' parameters. The problem of finding these constraints is called "the reduction problem"; we treat it in Sec. $6 \div 9$.

Sec.6 is devoted to the massive Thirring model; the reduction conditions appear in a straightforward way by using the standard algebraic technique. The interaction

[4]This result has been announced briefly in [15].

of solitons over the flat non-zero background turns out to be elastic — just as in the zero background case.

In Sec. 7 the complex sine-Gordon equation is dealt with, in the $(2+0)$-dimensional Euclidean space. The algebraic approach remains suitable here as well.

In the 8-th section we explore the same model (complex sine-Gordon) but this time in $(1+1)$ dimensions, i.e., in the Minkowskian space-time. This reduction cannot be defined by simply restricting the linear problem to some real form of $sl(2,C)$ algebra as it is done e.g. for the massive Thirring model. In ref.[14] the problem was cured by introducing an auxiliary gauge which induces rather non-obvious involution of the dressing matrix. This involution leads then to the necessary reduction conditions in a very straightforward way — for almost any choice of the seed solution. Curiously enough, but the sole exception is exactly the case of the flat non-zero background that we are most interested in! In this particular case the aforementioned involution degenerates. More precisely, the automorphism of the manifold $\{\Psi(\lambda)\}$ of the linear problem solutions $\Psi(\lambda)$ acquires a (removable) singularity.

In order to overcome this difficulty, we treat first a more general, generically non-singular case of the *exponential* background. Having constructed the N-soliton solution over this generic background, we then isolate the corresponding solution over the *constant* background (N-kink solution) – just as a suitable limiting case. Similarly to the massive Thirring model, the complex sine-Gordon kinks appear to interact *elastically*.

In Sec. 9 we proceed to another complexification of the sine-Gordon equation whose integrability was discovered in [14]. The Lagrangian is

$$L = \frac{|\partial_\mu \phi|^2}{1-|\phi|^2} - |\phi|^2 - \frac{1}{2}\frac{J_\mu^2}{|\phi|^2(1-|\phi|^2)} \tag{1.1}$$

where

$$J_\mu = i(\phi^*\partial_\mu\phi - \phi\partial_\mu\phi^*). \tag{1.2}$$

We refer to this model as the "$O(1,1)$ sine-Gordon", for the reason that will be explained in Sec. 9. (For the same reason the conventional complex sine-Gordon, defined by

$$L = \frac{|\partial_\mu \phi|^2}{1-|\phi|^2} - |\phi|^2, \tag{1.3}$$

is sometimes called the "$O(2)$ sine-Gordon".)

This reduction is accomplished according to the standard prescription; however, the result appears to be quite unexpected. Namely, we shall show that even the simplest solution of this model which corresponds to the single pole of $\Psi(\lambda)$, describes an inelastic process: decay or fusion of (subluminal) solitons (Sec. 10).

Finally, in Sec.11 several concluding remarks are made.

2. The $\mathcal{G}$-system

As we have already mentioned, we restrict ourselves here to the simplest case of the UNILOF scheme, defined by the choice of the $sl(2,C)$ algebra and the linear dependence on λ (spectral parameter) and λ^{-1}. That is, the linear problem is

$$i\Psi_\eta = (\lambda U_2 + U_0)\Psi$$
$$i\Psi_\xi = (\lambda^{-1}V_2 + V_0)\Psi, \tag{2.1}$$

where U, V and Ψ are 2×2 matrix-valued functions of η and ξ, and without loss of generality U_2, V_2, U_0, V_0 can be considered to be traceless [14]. The light cone variables η and ξ are defined in the standard way:

$$\eta = \frac{t+x}{2}, \qquad \xi = \frac{t-x}{2}. \tag{2.2}$$

The integrability conditions of 2.1 are

$$i\partial_\xi U_0 - i\partial_\eta V_0 + [U_2, V_2] + [U_0, V_0] = 0, \tag{2.3}$$
$$i\partial_\eta V_2 + [V_2, U_0] = 0, \tag{2.4}$$
$$i\partial_\xi U_2 + [U_2, V_0] = 0, \tag{2.5}$$

which are 3 equations for 4 matrix-valued unknowns. The corresponding ambiguity is a manifestation of the well-known gauge-invariance of (2.1). More precisely, the set (2.1) is covariant [4] [3] under the gauge transformation

$$U_2 = g\tilde{U}_2 g^{-1}, \qquad V_2 = g\tilde{V}_2 g^{-1},$$
$$U_0 = g\tilde{U}_0 g^{-1} + ig_\eta g^{-1}, \qquad V_0 = g\tilde{V}_0 g^{-1} + ig_\xi g^{-1},$$
$$\Psi = g\tilde{\Psi}, \tag{2.6}$$

where $g = g(\eta, \xi; \lambda) \in SL(2,C)$. The central idea of the UNILOF scheme consists in fixing the *triangular* gauge: $(U_2)_{21} = (V_2)_{12} = 0$. Assuming then U_2 and V_2 to be non-singular and excluding thereby what we call the degenerate case[5], eqs.(2.4) and (2.5) imply $(U_0)_{12} = (V_0)_{21} = 0$. Finally, rescaling the coordinates and fields [14], the linear problem matrices can be cast to the following ultimate form:

$$U_2 = \begin{pmatrix} 1/2 & q_1 \\ 0 & -1/2 \end{pmatrix}, \qquad U_0 = \begin{pmatrix} F/2 & 0 \\ q_2 & -F/2 \end{pmatrix},$$
$$V_2 = \begin{pmatrix} 1/2 & 0 \\ q_4 & -1/2 \end{pmatrix}, \qquad V_0 = \begin{pmatrix} G/2 & q_3 \\ 0 & -G/2 \end{pmatrix}. \tag{2.7}$$

Here $q_{1,2,3,4}$ and F, G are complex fields; the reason for the chosen notation will become obvious below.

[5] The degenerate case (zero diagonal in U_2 and V_2) corresponds to the two-dimensional Toda lattices that were analysed earlier [7] – [10].

Now the integrability conditions (2.3-2.5) can be written componentwise as

$$iq_{1\xi} + q_3 - Gq_1 = 0 \tag{2.8}$$
$$iq_{2\xi} - q_4 + Gq_2 = 0 \tag{2.9}$$
$$iq_{3\eta} + q_1 - Fq_3 = 0 \tag{2.10}$$
$$iq_{4\eta} - q_2 + Fq_4 = 0 \tag{2.11}$$
$$iF_\xi - iG_\eta + 2(q_1q_4 - q_2q_3) = 0. \tag{2.12}$$

This system possesses a "residual" gauge invariance :

$$(q_1, q_3) = e^{-i\theta}(\tilde{q}_1, \tilde{q}_3), \qquad (q_2, q_4) = e^{i\theta}(\tilde{q}_2, \tilde{q}_4),$$
$$F = \tilde{F} + \theta_\eta, \qquad G = \tilde{G} + \theta_\xi, \tag{2.13}$$

which may be used to identify F and G. Namely, eqs.(2.8-2.12) imply

$$(F + q_1q_2)_\xi = (G + q_3q_4)_\eta,$$

whence

$$F + q_1q_2 = \pi_\eta, \qquad G + q_3q_4 = \pi_\xi \tag{2.14}$$

for some $\pi(\eta, \xi)$. Recovering π from here, we can perform the transformation (2.13) with $\theta = \pi$ so as to accomplish[6]

$$F = -q_1q_2, \qquad G = -q_3q_4. \tag{2.15}$$

The equations (2.8-2.11), (2.15) provide the generic integrable system associated with the $sl(2,C)$ case of the UNILOF scheme; we call it "$\mathcal{G}$-system". For the later convenience, we rewrite it in a self-contained way:

$$iq_{1\xi} + q_3(1 + q_1q_4) = 0, \qquad iq_{2\xi} - q_4(1 + q_2q_3) = 0,$$
$$iq_{3\eta} + q_1(1 + q_2q_3) = 0, \qquad iq_{4\eta} - q_2(1 + q_1q_4) = 0. \tag{2.16}$$

The $\mathcal{G}$-system is manifestly Lagrangian, with the Lagrangian density being

$$\mathcal{L} = iq_2q_{1\xi} + iq_4q_{3\eta} + q_1q_4 + q_2q_3 + q_1q_2q_3q_4 + (c.c.). \tag{2.17}$$

Several important models of relativistic field theory are contained in (2.16) as its reductions, and we shall isolate N-soliton solutions of these models from the generic N-soliton solution of the $\mathcal{G}$-system.

[6] More generally, we could identify $F = -\omega_1 q_1 q_2,, G = -\omega_2 q_3 q_4$ where $\omega_{1,2}$ are any two constants satisfying $\omega_1 + \omega_2 = 2$. However, the nonlinear systems emerging in this way are in general non-lagrangian [14] and so we do not discuss them here. Soliton solutions for these systems can be obtained in exactly the same way as those for the system (2.16) but the resulting formulas (noninteger or even irrational powers of rational functions of exponentials) are too cumbersome even in the zero background case [14].

3. The Dressing Procedure

Formally, the $\mathcal{G}$-system itself looks like a special case of some more general system whose associated linear problem (2.1) involves *full* matrices U_2, V_2, U_0, V_0 (though in actual fact it is gauge-equivalent to this latter system). It is just the solution of the latter system that the inverse scattering technique povides, and we have yet to extract the solution for the $\mathcal{G}$-system from it. In order to facilitate this extraction, we shall represent the linear problem (2.1-2.5) as the Z_2 reduction of the generic *quadratic* bundle.

Namely, substituting λ^2 for λ in (2.1) [this clearly does not influence the integrability conditions (2.3)] and making the gauge transformation (2.6) with the matrix $g = diag\{\lambda^{1/2}, \lambda^{-1/2}\}$, the linear system (2.1) is converted to the following one [14]:

$$i\Psi_\eta = (\lambda^2 A_2 + \lambda A_1 + A_0)\Psi \equiv A(\lambda)\Psi \tag{3.1}$$

$$i\Psi_\xi = (\lambda^{-2} B_2 + \lambda^{-1} B_1 + B_0)\Psi \equiv B(\lambda)\Psi. \tag{3.2}$$

Here

$$A_2 = B_2 = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} = \frac{1}{2}\sigma_3, \tag{3.3}$$

$$A_1 = \begin{pmatrix} 0 & q_1 \\ q_2 & 0 \end{pmatrix}, \qquad B_1 = \begin{pmatrix} 0 & q_3 \\ q_4 & 0 \end{pmatrix}, \tag{3.4}$$

$$A_0 = \frac{1}{2} F\sigma_3, \qquad B_0 = \frac{1}{2} G\sigma_3. \tag{3.5}$$

Since the matrices of potentials in (3.1-3.2) satisfy

$$\sigma_3 A(-\lambda)\sigma_3 = A(\lambda), \qquad \sigma_3 B(-\lambda)\sigma_3 = B(\lambda), \tag{3.6}$$

the combination $\sigma_3\Psi(-\lambda)\sigma_3$ solves the same eqs.(3.1-3.2) as $\Psi(\lambda)$ does, so that normalizing $\Psi(\lambda)$ properly we can arrange[7] that

$$\sigma_3\Psi(-\lambda)\sigma_3 = \Psi(\lambda). \tag{3.7}$$

This symmetry appears to be crucial for the isolation of solutions of the $\mathcal{G}$-system.

To construct the soliton solutions we adopt the idea of the dressing method [4], [3]. Suppose we are given some background solution $q_1^{(0)}(\eta,\xi), ..., q_4^{(0)}(\eta,\xi)$ of eqs.(2.8-2.11), (2.15). Denote the corresponding matrices (3.4), (3.5) through $A_1^{(0)}$, $B_1^{(0)}$, $A_0^{(0)}$, and $B_0^{(0)}$, and let $\Psi_{(0)}$ stand for the pertaining solution of the linear problem (3.1,3.2). The Ψ-function corresponding to the new solution $q_1, ..., q_4$ describing N solitons propagating over the given background $q_1^{(0)}, ...q_4^{(0)}$, is obtained then as

$$\Psi = \chi\Psi_{(0)} \tag{3.8}$$

where "the dressing matrix" χ is meromorphic in λ, together with its inverse. Choosing Ψ_0 and Ψ to satisfy (3.7) we have

$$\sigma_3\chi(-\lambda)\sigma_3 = \chi(\lambda) \tag{3.9}$$

whence

$$\chi(\lambda) = R\tilde{\chi}(\lambda), \tag{3.10}$$

$$\tilde{\chi}(\lambda) = 1 + \sum_{i=1}^{N} \frac{P^i}{\lambda - \nu_i} - \sum_{i=1}^{N} \frac{\sigma_3 P^i \sigma_3}{\lambda + \nu_i}, \tag{3.11}$$

and

$$\chi^{-1}(\lambda) = \tilde{\chi}(\lambda)^{-1} R^{-1} = \left(1 + \sum_{i=1}^{N} \frac{Q^i}{\lambda - \mu_i} - \sum_{i=1}^{N} \frac{\sigma_3 Q^i \sigma_3}{\lambda + \mu_i}\right) R^{-1}. \tag{3.12}$$

Here $R = \chi_{|\lambda=\infty}$ is usually called the normalization matrix. Actually R may be looked upon as a matrix of some gauge transformation, so that the dressing procedure consists of the dressing with the "intermediate" dressing matrix $\tilde{\chi}$ followed by the gauge transformation with the matrix R. With respect to this transformation, the "potentials" A and B transform as

$$A_0 = \tilde{A}_0 + iR_\eta R^{-1}, \qquad A_1 = R\tilde{A}_1 R^{-1},$$
$$B_0 = \tilde{B}_0 + iR_\xi R^{-1}, \qquad B_1 = R\tilde{B}_1 R^{-1}. \tag{3.13}$$

Taking (3.9) at the point $\lambda = \infty$, we have that $\sigma_3 R\sigma_3 = R$ and so R is necessarily diagonal. Furthermore, it may be inferred from (3.1-3.2) by Liouville's formula that $\det\Psi$ is (η,ξ)-independent so that we can normalize Ψ (or Ψ_0) in such a way that $\det R = 1$. Hence,

$$R = \begin{pmatrix} r(\eta,\xi) & 0 \\ 0 & r^{-1}(\eta,\xi) \end{pmatrix}. \tag{3.14}$$

Using this in (3.13) we obtain, componentwise:

$$q_1 = r^2\tilde{q}_1, \quad q_2 = r^{-2}\tilde{q}_2, \quad q_3 = r^2\tilde{q}_3, \quad q_4 = r^{-2}\tilde{q}_4 \tag{3.15}$$

$$F = \tilde{F} + 2ir_\eta r^{-1}, \qquad G = \tilde{G} + 2ir_\xi r^{-1}. \tag{3.16}$$

[Here $\tilde{q}_{1,2,3,4}$ and $\tilde{F}, \tilde{G}$ are elements of the matrices A_0, B_0, A_1, B_1 whose structure is the same as in eqs. (3.4-3.5).] Comparing to (2.13) we see that R is nothing but

[7] In general, we have $\sigma_3\Psi(\lambda)\sigma_3 = \Psi(\lambda)H(\lambda)$ where $H(\lambda)$ is some non-singular coordinate-independent matrix. However, it can be shown that (at least) for the physically meaningful backgrounds we can substitute $\Psi \to \tilde{\Psi} = \Psi M$ where the coordinate-independent matrix M is chosen so that $\tilde{\Psi}$ satisfy (3.7).

the aforementioned "residual" gauge transformation. Accordingly, we can, without loss of generality, "dress" the background with the help of the intermediate dressing matrix χ only. However, the result will be a solution of the non-gauged system (2.8-2.12) so that to obtain a solution of the gauged system (2.16) we have yet to find the gauge transformation (2.13) (that is, the matrix R) accomplishing the condition (2.15).

We shall be concerned with the generic situation of $\nu_i \neq \pm\mu_k$. In this case the identity $\tilde{\chi}\tilde{\chi}^{-1} = 1$ requires that

$$P^i\tilde{\chi}^{-1}(\nu_i) = \tilde{\chi}(\mu_i)Q^i = 0, \tag{3.17}$$

$i = 1, ..., N$. It is convenient to represent the degenerate matrices P^i and Q^j as

$$P^i = \vec{x}^{\,i} \otimes \vec{t}^{\,i}, \qquad Q^i = \vec{s}^{\,i} \otimes \vec{y}^{\,i}, \tag{3.18}$$

where $\vec{x}^{\,i}$ and $\vec{s}^{\,i}$ are two-component vector-columns,

$$\vec{x}^{\,i} = \begin{pmatrix} x_1^i \\ x_2^i \end{pmatrix}, \qquad \vec{s}^{\,i} = \begin{pmatrix} s_1^i \\ s_2^i \end{pmatrix}$$

while $\vec{t}^{\,i}$ and $\vec{y}^{\,i}$ denote two-component vector-rows: $\vec{t}^{\,i} = (t_1^i, t_2^i)$, $\vec{y}^{\,i} = (y_1^i, y_2^i)$. The direct products in (3.18) should be understood in the standard way, e.g.

$$P^i = \begin{pmatrix} x_1^i t_1^i & x_1^i t_2^i \\ x_2^i t_1^i & x_2^i t_2^i \end{pmatrix}$$

(note that no summation is assumed over i). The components of the above $4N$ two-component vectors can be re-arranged to form eight N-component vector-columns $|\, x_A >$, $|\, y_A >$, $|\, s_A >$ and $|\, t_A >$, $A = 1, 2$. For example,

$$|\, x_1 >= \begin{pmatrix} x_1^1 \\ x_1^2 \\ \vdots \\ x_1^N \end{pmatrix}.$$

Also it is useful to introduce the transposed vector-rows $< x_A\,|$, $< y_A\,|$, $< s_A\,|$ and $< t_A\,|$, e.g. $< x_1\,| = (x_1^1, x_1^2, ..., x_1^N)$. Now feeding (3.18) into (3.17) yields

$$\begin{aligned} 2a_1 \,|\, x_1 >=|\, s_1 >, \qquad & 2a_2 \,|\, x_2 >=|\, s_2 >, \\ 2 < y_1 \,|\, a_2 = - < t_1 \,|, \qquad & 2 < y_2 \,|\, a_1 = - < t_2 \,|, \end{aligned} \tag{3.19}$$

where a_1 and a_2 are $N \times N$ matrices with elements

$$a_1^{ij} = \frac{\nu_j s_1^i t_1^j + \mu_i s_2^i t_2^j}{\nu_j^2 - \mu_i^2}, \qquad a_2^{ij} = \frac{\mu_i s_1^i t_1^j + \nu_j s_2^i t_2^j}{\nu_j^2 - \mu_i^2}. \tag{3.20}$$

Eqs.(3.19) imply

$$\begin{aligned} |\, x_1 >= \frac{1}{2}a_1^{-1} \,|\, s_1 >, \qquad & |\, x_2 >= \frac{1}{2}a_2^{-1} \,|\, s_2 >, \\ < y_1 \,|= -\frac{1}{2} < t_1 \,|\, a_2^{-1}, \qquad & < y_2 \,|= -\frac{1}{2} < t_2 \,|\, a_1^{-1}. \end{aligned} \tag{3.21}$$

In terms of $\tilde{\chi}$ eq.(3.11), the linear problem (3.1-3.2) takes the form

$$i\tilde{\chi}_\eta\tilde{\chi}^{-1} + \tilde{\chi}\left(A_0^{(0)} + \lambda A_1^{(0)} + \lambda^2 A_2^{(0)}\right)\tilde{\chi}^{-1} = \tilde{A}_0 + \lambda\tilde{A}_1 + \lambda^2\tilde{A}_2 \tag{3.22}$$

$$i\tilde{\chi}_\xi\tilde{\chi}^{-1} + \tilde{\chi}\left(B_0^{(0)} + \lambda^{-1}B_1^{(0)} + \lambda^{-2}B_2^{(0)}\right)\tilde{\chi}^{-1} = \tilde{B}_0 + \lambda^{-1}\tilde{B}_1 + \lambda^{-2}\tilde{B}_2. \tag{3.23}$$

Representing $\tilde{\chi}'\tilde{\chi}^{-1}$ as $-\tilde{\chi}(\tilde{\chi}^{-1})'$, and requiring that residues of the l.h.s. at $\lambda = \mu_i$ vanish, we obtain

$$\begin{aligned} \tilde{\chi}(\mu_i)\{-i\vec{s}_\eta^{\,i} + A^{(0)}(\mu_i)\vec{s}^{\,i}\} = 0 \\ \tilde{\chi}(\mu_i)\{-i\vec{s}_\xi^{\,i} + B^{(0)}(\mu_i)\vec{s}^{\,i}\} = 0. \end{aligned} \tag{3.24}$$

Comparing this to (3.1) we get a particular solution:

$$\vec{s}^{\,i} = \Psi_{(0)}(\mu_i)\vec{n}^{\,i}, \tag{3.25}$$

where $\vec{n}^{\,i} = \begin{pmatrix} n_1^i \\ n_2^i \end{pmatrix}$ is an arbitrary constant vector-column. Following [16], it may be proved that this is, in fact, the *general* solution. In the same way,

$$\vec{t}^{\,i} = \vec{m}^{\,i}\Psi_{(0)}^{-1}(\nu_i), \tag{3.26}$$

where $\vec{m}_i = (m_1^i, m_2^i)$ is an arbitrary constant vector-row.

Next, let us express the "potentials" $\tilde{A}$ and $\tilde{B}$ in terms of $\tilde{\chi}$. Expanding the l.h.s. of (3.22,3.23) in the Laurent series at $\lambda = \infty$, we can express e.g. $\tilde{A}_1$ and $\tilde{B}_0$:

$$\tilde{A}_1 = A_1^{(0)} + \lim_{\lambda\to\infty}\left\{\lambda\tilde{\chi}(\lambda)A_2^{(0)}\tilde{\chi}^{-1}(\lambda)\right\}, \tag{3.27}$$

$$\tilde{B}_0 = B_0^{(0)}, \tag{3.28}$$

while the simplest expressions for $\tilde{B}_1$ and $\tilde{A}_0$ arise by expanding at $\lambda = 0$:

$$\tilde{B}_1 = \tilde{\chi}(0)B_1^{(0)}\tilde{\chi}^{-1}(0) + \lim_{\lambda\to 0}\left\{\lambda^{-1}\tilde{\chi}(\lambda)B_2^{(0)}\tilde{\chi}^{-1}(\lambda)\right\}, \tag{3.29}$$

$$\tilde{A}_0 = A_0^{(0)} + i(\tilde{\chi}_\eta\tilde{\chi}^{-1})_{|\lambda=0}. \tag{3.30}$$

In addition, we record here two identities that will appear useful below. First, evaluate the trace of the square of eq.(3.22):

$$tr\left\{(A^{(0)})^2 - (\tilde{\chi}^{-1}\tilde{\chi}_\eta)^2 + 2iA^{(0)}\tilde{\chi}^{-1}\tilde{\chi}_\eta\right\} = tr\tilde{A}^2. \tag{3.31}$$

Expanding (3.31) at $\lambda = \infty$, we find, as a coefficient at λ^2:

$$tr\{(A_1^{(0)})^2 + 2A_2A_0^{(0)}\} = tr\{\tilde{A}_1^2 + 2A_2\tilde{A}_0\},$$

or, componentwise,

$$-F^{(0)} = \tilde{F} + 2\tilde{q}_1\tilde{q}_2. \tag{3.32}$$

In a similar way we obtain the second identity:

$$-G^{(0)} + 2ip_\xi p^{-1} = \tilde{G} + 2\tilde{q}_3\tilde{q}_4, \tag{3.33}$$

where $p(\eta, \xi)$ is the element of the matrix $\tilde{\chi}_{|\lambda=0}$:

$$\tilde{\chi}_{|\lambda=0} = const \times \begin{pmatrix} p(\eta,\xi) & 0 \\ 0 & p^{-1}(\eta,\xi) \end{pmatrix}. \tag{3.34}$$

[The fact that $\tilde{\chi}_{|\lambda=0}$ is diagonal, is a mere consequence of eq.(3.9), and that it has the structure (3.34) follows from that $\det\tilde{\chi}$ does not depend on η and ξ.]

Now we have to specify the gauge transformation (3.13) to ensure the condition (2.15). In principle, this can be accomplished by recovering $\pi(\eta, \xi)$ from (2.14) but then the ultimate solutions would include exponentials of quadratures. Fortunately, in the pure solitonic case that we deal with here, eq.(2.14) appears to be unnecessary, for the matrix R can be found in a more direct way, and in a closed form[8]. Indeed, from (3.15) and (2.15) it ensues that

$$\tilde{F} = -\tilde{q}_1\tilde{q}_2 - 2ir_\eta r^{-1}, \qquad \tilde{G} = -\tilde{q}_3\tilde{q}_4 - 2ir_\xi r^{-1}, \tag{3.35}$$

while eliminating $\tilde{q}_1\tilde{q}_2$ and $\tilde{q}_3\tilde{q}_4$ by virtue of eqs.(3.32, 3.33), we have

$$(F^{(0)} - \tilde{F})/2 = 2ir_\eta r^{-1}, \qquad (G^{(0)} - \tilde{G})/2 = 2ir_\xi r^{-1} + ip_\xi p^{-1}. \tag{3.36}$$

On the other hand, eqs.(3.30) and (3.28) read, componentwise:

$$(\tilde{F} - F^{(0)})/2 = ip_\eta p^{-1}, \qquad \tilde{G} - G^{(0)} = 0. \tag{3.37}$$

Comparing to (3.36) produces the logarithmic derivatives of r:

$$2r_\eta r^{-1} = -p_\eta p^{-1}, \qquad 2r_\xi r^{-1} = -p_\xi p^{-1} \tag{3.38}$$

[8]The key point here is clearly to use the trace identities (3.32,3.33) and this idea traces back to the work of David et.al. on the Thirring model [17]. This is not the unique way, however; we could alternatively utilize a completely different approach that was used in the case of the zero background [14]. In this latter approach the necessary identities arise from the comparison of two sets of expressions for F and G, obtained by expanding eqs.(3.22,3.23) at $\lambda = 0$ and ∞, respectively.

whence, up to an unessential constant,

$$r^2 = p^{-1}. \tag{3.39}$$

The quantity p is computed in the Appendix; by (A4) we have then

$$r^2(\eta, \xi) = \det a_1 / \det a_2. \tag{3.40}$$

So, writing componentwise (3.27) and (3.29),

$$\tilde{q}_1 = q_1^{(0)} - 2\sum_{i=1}^{N} P_{12}^i, \qquad \tilde{q}_2 = q_2^{(0)} + 2\sum P_{21}^i,$$

$$\tilde{q}_3 = p^2 q_3^{(0)} + 2p\sum \nu_i^{-2} P_{12}^i, \qquad \tilde{q}_4 = p^{-2} q_4^{(0)} - 2p^{-1}\sum \nu_i^{-2} P_{21}^i,$$

and using (3.15) and (3.40), we arrive finally at the closed expressions for the N-soliton solutions of the $\mathcal{G}$-system (2.16):

$$q_1 = (\det a_1/\det a_2)(q_1^{(0)} - < t_2 \mid a_1^{-1} \mid s_1 >), \tag{3.41}$$

$$q_2 = (\det a_2/\det a_1)(q_2^{(0)} + < t_1 \mid a_2^{-1} \mid s_2 >), \tag{3.42}$$

$$q_3 = \left(\frac{\det a_2}{\det a_1}\right) q_3^{(0)} + \prod_{i=1}^{N}\left(\frac{\nu_i}{\mu_i}\right) < \nu^{-2} t_2 \mid a_1^{-1} \mid s_1 >, \tag{3.43}$$

$$q_4 = \left(\frac{\det a_1}{\det a_2}\right) q_4^{(0)} - \prod_{i=1}^{N}\left(\frac{\nu_i}{\mu_i}\right) < \nu^{-2} t_1 \mid a_2^{-1} \mid s_2 > . \tag{3.44}$$

The latter two expressions can be symmetrized by means of the identities

$$\det a_2 < t_2\nu \mid a_2^{-1} \mid s_1\mu^{-1} > = \prod_{i=1}^{N}(\nu_i\mu_i) \det a_1 < t_2\nu^{-2} \mid a_1^{-1} \mid s_1 >, \tag{3.45}$$

$$\det a_1 < t_1\nu \mid a_1^{-1} \mid s_2\mu^{-1} > = \prod_{i=1}^{N}(\nu_i\mu_i) \det a_2 < t_1\nu^{-2} \mid a_2^{-1} \mid s_2 > . \tag{3.46}$$

(The proof of these is similar to the proof of eq.(A4) of the Appendix, and is given in [14].) By virtue of (3.45,3.46) eqs.(3.43, 3.44) take the form

$$q_3 = (\det a_2/\det a_1)(q_3^{(0)} + < t_2\nu^{-1} \mid a_2^{-1} \mid s_1\mu^{-1} >), \tag{3.47}$$

$$q_4 = (\det a_1/\det a_2)(q_4^{(0)} + < t_1\nu^{-1} \mid a_1^{-1} \mid s_2\mu^{-1} >). \tag{3.48}$$

[We remind here that the notation $< t_1\nu^{-1} \mid$ is understood as $(t_1^1\nu_1^{-1}, t_1^2\nu_2^{-1}, ..., t_1^N\nu_N^{-1})$].

Solution (3.41-3.42), (3.47-3.48) has the drawback of that in order to visualize it (for example, to plot the profile by means of the computer), one has to invert the $N \times N$ matrices a_1 and a_2. However, this problem can be circumvented. Invoking eq.(A3) of the Appendix, the solutions acquire the form

$$q_1 = q_1^{(0)}\frac{\det b_1}{\det a_2}, \qquad q_2 = q_2^{(0)}\frac{\det b_2}{\det a_1}, \qquad q_3 = q_3^{(0)}\frac{\det b_3}{\det a_1}, \qquad q_4 = q_4^{(0)}\frac{\det b_4}{\det a_2}, \tag{3.49}$$

where

$$b_1 = a_1 - \frac{1}{q_1^{(0)}} \mid s_1 >< t_2 \mid, \qquad b_2 = a_2 + \frac{1}{q_2^{(0)}} \mid s_2 >< t_1 \mid,$$

$$b_3 = a_2 + \frac{1}{q_3^{(0)}} \mid s_1\mu^{-1} >< t_2\nu^{-1} \mid, \qquad b_4 = a_1 - \frac{1}{q_4^{(0)}} \mid s_2\mu^{-1} >< t_1\nu^{-1} \mid . \quad (3.50)$$

[We remind that e.g. $\mid s_1\mu^{-1} >< t_2\nu^{-1} \mid$ denotes the matrix whose (ij)-th element is $s_1^i\mu_i^{-1}t_2^j\nu_j^{-1}$]. The representation (3.49-3.50) is valid only for non-zero backgrounds, $q_i^{(0)} \neq 0$. In the case of the zero background, the determinant representation also exists [14], but involves matrices of the dimension $(N+1)\times(N+1)$.

4. The seed Ψ-function: exponential background

The N-soliton solutions (3.49) are closed up to the computing of the seed Ψ-function (i.e., the function $\Psi_{(0)}$ corresponding to the background solution) which defines the coordinate dependence of $\vec{s}$ and $\vec{t}$ [see eqs.(3.25), (3.26)]. In several instances $\Psi_{(0)}$ can be found explicitly. Physically, the most interesting situation corresponds to the flat backgrounds, $q_i^{(0)} \equiv const$. In the preceding paper [14] the zero background case was discussed, $q_i^{(0)} \equiv 0$, and here our goal is the flat *non-zero* background. However, in view of certain difficulties with isolating solutions for some of the reductions in this case, we choose to solve a more general problem. Namely, we shall obtain $\Psi_{(0)}$ corresponding to the *exponential* background which contains the flat background as a limiting case.

The said exponential solution is

$$q_1^{(0)} = C_1e^z, \qquad q_2^{(0)} = C_2e^{-z}, \qquad q_3^{(0)} = C_3e^z, \qquad q_4^{(0)} = C_4e^{-z}, \quad (4.1)$$

where

$$z = i(\alpha\eta + \beta\xi) \quad (4.2)$$

and $C_1, ..., C_4, \alpha, \beta$ are constants. Feeding (4.1, 4.2) into (2.16) and assuming $\alpha, \beta \neq 0$, we have

$$\alpha = C_1/C_3 + C_1C_2, \quad (4.3)$$

$$\beta = C_4/C_2 + C_3C_4, \quad (4.4)$$

$$C_1C_4 = C_2C_3. \quad (4.5)$$

The flat background corresponds to $\alpha = \beta = 0$. In this case eqs.(4.3-4.5) simplify to

$$C_1C_4 = C_2C_3 = -1. \quad (4.6)$$

It is advantageous to treat the solution (4.1) as a gauge-transformed *constant* solution. Indeed, letting

$$\Psi_{(0)} = exp(z\sigma_3/2)\hat{\Psi}_{(0)} \quad (4.7)$$

where as before $\Psi_{(0)}$ stands for the background Ψ-function [i.e., the solution of the linear problem (3.1, 3.2) pertaining to the background potentials (4.1)], $\hat{\Psi}_{(0)}$ satisfies the linear system with *constant* coefficients:

$$i\partial_\eta\hat{\Psi}_{(0)} = (\lambda^2A_2 + \lambda\hat{A}_1^{(0)} + \hat{A}_0^{(0)})\hat{\Psi}_{(0)} \equiv \hat{A}^{(0)}\hat{\Psi}_{(0)} \quad (4.8)$$

$$i\partial_\xi\hat{\Psi}_{(0)} = (\lambda^{-2}B_2 + \lambda^{-1}\hat{B}_1^{(0)} + \hat{B}_0^{(0)})\hat{\Psi}_{(0)} \equiv \hat{B}^{(0)}\hat{\Psi}_{(0)} \quad (4.9)$$

where $A_2 = B_2 = \frac{1}{2}\sigma_3$,

$$\hat{A}_1^{(0)} = \begin{pmatrix} 0 & C_1 \\ C_2 & 0 \end{pmatrix}, \qquad \hat{B}_1^{(0)} = \begin{pmatrix} 0 & C_3 \\ C_4 & 0 \end{pmatrix}, \quad (4.10)$$

$$\hat{A}_0^{(0)} = \frac{1}{2}(\alpha - C_1C_2)\sigma_3, \qquad \hat{B}_0^{(0)} = \frac{1}{2}(\beta - C_3C_4)\sigma_3. \quad (4.11)$$

Note that this is the same linear problem as the one for the flat background, $q_1^{(0)} \equiv C_1, ..., q_4^{(0)} \equiv C_4$; the only distinction is that in the latter case we would have $A_0^{(0)} = -\frac{1}{2}C_1C_2\sigma_3$, $B_0^{(0)} = -\frac{1}{2}C_3C_4\sigma_3$ instead of (4.11).

The solution of (4.8,4.9),

$$\hat{\Psi}_{(0)} = exp\left\{-i(\hat{A}^{(0)}\eta + \hat{B}^{(0)}\xi)\right\}M \quad (4.12)$$

is unique up to multiplying by an arbitrary (η, ξ)-independent nonsingular matrix M on the right. Choosing this matrix properly, $\hat{\Psi}_{(0)}$ can be rewritten in the following convenient form:

$$\hat{\Psi}_{(0)} = M\exp\{i\sigma_3 f\}, \quad (4.13)$$

where

$$f(\eta, \xi; \lambda) = (a\eta + b\xi)/2, \quad (4.14)$$

$$a = [(\lambda^2 + C_1/C_3)^2 + 4C_1C_2\lambda^2]^{1/2}, \quad (4.15)$$

$$b = [(\lambda^{-2} + C_3/C_1)^2 + 4C_3C_4\lambda^{-2}]^{1/2}, \quad (4.16)$$

and in eqs. (4.15, 4.16) we should choose only those branches of the square root that satisfy

$$ab = C_3C_1^{-1}\lambda^2 + C_1C_3^{-1}\lambda^{-2} + 4C_1C_4 + 2. \quad (4.17)$$

Here the matrix M is given by

$$M = \frac{1}{\sqrt{1 + J^2(\lambda)C_1^{-1}C_2^{-1}}}\begin{pmatrix} 1 & J(\lambda)/C_2 \\ -J(\lambda)/C_1 & 1 \end{pmatrix}, \quad (4.18)$$

where we have introduced the notation

$$J(\lambda)=\frac{a+\lambda^2+C_1/C_3}{2\lambda}. \tag{4.19}$$

(Actually the class of suitable matrices is somewhat larger, but the choice (4.18) is quite sufficient for our purposes).

The proof of that $\hat{\Psi}_{(0)}$ eq.(4.12) coincides with eq.(4.13) if M is taken as in eq.(4.18), is trivial provided we use the representation

$$\exp(i\sum_{n=1}^{3} f_n\sigma_n)=\cos f\cdot 1+i\sin f\cdot\sum_{n=1}^{3}\frac{f_n}{f}\sigma_n, \tag{4.20}$$

where f_n are complex quantities, σ_n $(n=1,2,3)$ are the standard Pauli matrices, and $f=(f_1^2+f_2^2+f_3^2)^{1/2}$. Using (4.20), we can rewrite $\exp\{-i(\hat{A}^{(0)}\eta+\hat{B}^{(0)}\xi\}$ in eq.(4.12), as

$$\cos f\cdot\mathbf{1}+\frac{i\sin f}{a}\cdot\begin{pmatrix}\lambda^2+C_1/C_3 & -2C_1\lambda\\ -2C_2\lambda & \lambda^2+C_1/C_3\end{pmatrix}, \tag{4.21}$$

with f as in (4.14) and a as in (4.15). Multiplying this expression by M eq.(4.18) on the right, it is straightforward to verify that it is equal to $M\exp(i\sigma_3 f)$. Q.E.D.

For the latter convenience, we write out $\Psi_{(0)}$ in a self-contained way:

$$\Psi_{(0)}=e^{z\sigma_3/2}Me^{i\sigma_3 f}. \tag{4.22}$$

Here z is given by (4.2), f by (4.14), and M by (4.18). The formula (4.22) provides us with explicit expressions for $\vec{s}^{\,i}$ and $\vec{t}^{\,i}$, eqs.(3.25, 3.26). Consequently, eqs.(3.41-3.44), (3.47-3.48), and (3.49) become *closed, explicit* N-soliton solutions.

5. Simplified N-soliton formulas: the case of the flat background

In the case of the flat background, $q_i^{(0)}\equiv C_i$, the N-soliton formulas simplify considerably. In this case the constants C_i obey $C_1C_4=C_2C_3=-1$, z eq.(4.2) is equal to zero, and the seed Ψ-function simplifies to

$$\Psi_{(0)}=Me^{if\sigma_3}, \tag{5.1}$$

with

$$M(\lambda)=\frac{1}{\sqrt{1+C_3C_4\lambda^2}}\begin{pmatrix}1 & -C_3\lambda\\ C_4\lambda & 1\end{pmatrix}, \tag{5.2}$$

and

$$f(\lambda)=\frac{1}{2}\left\{(\lambda^2+C_1C_2)\eta-(\lambda^{-2}+C_3C_4)\xi\right\}. \tag{5.3}$$

Let us turn to the N-soliton solutions (3.49). Using $\Psi_{(0)}$ eq.(5.1) in the expressions for $\vec{s}^{\,i}$ and $\vec{t}^{\,i}$, the N-soliton formulas acquire the following simplified form:

$$q_1=C_1\frac{\det b_1}{\det a_2},\qquad q_2=C_2\frac{\det b_2}{\det a_1},\qquad q_3=C_3\frac{\det b_3}{\det a_1},\qquad q_4=C_4\frac{\det b_4}{\det a_2} \tag{5.4}$$

where the matrices $a_1, a_2, b_1, ..., b_4$ have the "shortened", three-term structure:

$$a_1^{ij}=\tilde{a}_1^{ij}/r_{ij},\qquad a_2^{ij}=\tilde{a}_2^{ij}/r_{ij},\qquad b_1^{ij}=\tilde{b}_1^{ij}/r_{ij},\qquad ...\qquad b_4^{ij}=\tilde{b}_4^{ij}/r_{ij} \tag{5.5}$$

with $r_{ij}=\nu_j^2-\mu_i^2$ and

$$\begin{aligned}
\tilde{b}_1^{ij}&=\nu_jq_jn_1^im_1^je^{i\{f(\mu_i)-f(\nu_j)\}}+C_3^2C_1^{-1}\mu_i\nu_jr_{ij}n_1^im_1^je^{-i\{f(\mu_i)+f(\nu_j)\}}\\
&\quad+\mu_ip_in_2^im_2^je^{i\{f(\nu_j)-f(\mu_i)\}},\\
\tilde{b}_2^{ij}&=\mu_ip_in_1^im_1^je^{i\{f(\mu_i)-f(\nu_j)\}}+C_3C_1^{-2}\mu_i\nu_jr_{ij}n_1^im_2^je^{i\{f(\mu_i)+f(\nu_j)\}}\\
&\quad+\nu_jq_jn_2^im_2^je^{i\{f(\nu_j)-f(\mu_i)\}},\\
\tilde{b}_3^{ij}&=\nu_j^2\mu_i^{-1}p_in_1^im_1^je^{i\{f(\mu_i)-f(\nu_j)\}}+C_3^{-1}\mu_i^{-1}\nu_j^{-1}r_{ij}n_1^im_2^je^{i\{f(\mu_i)+f(\nu_j)\}}\\
&\quad+\mu_i^2\nu_j^{-1}q_jn_2^im_2^je^{i\{f(\nu_j)-f(\mu_i)\}},\\
\tilde{b}_4^{ij}&=\mu_i^2\nu_j^{-2}q_jn_1^im_1^je^{i\{f(\mu_i)-f(\nu_j)\}}+C_1\mu_i^{-1}\nu_j^{-1}r_{ij}n_2^im_1^je^{-i\{f(\mu_i)+f(\nu_j)\}}\\
&\quad+\nu_j^2\mu_i^{-1}p_in_2^im_2^je^{i\{f(\nu_j)-f(\mu_i)\}},\\
\tilde{a}_1^{ij}&=\nu_jp_in_1^im_1^je^{i\{f(\mu_i)-f(\nu_j)\}}+C_1^{-1}r_{ij}n_1^im_2^je^{i\{f(\mu_i)+f(\nu_j)\}}\\
&\quad+\mu_iq_jn_2^im_2^je^{i\{f(\nu_j)-f(\mu_i)\}},\\
\tilde{a}_2^{ij}&=\mu_iq_jn_1^im_1^je^{i\{f(\mu_i)-f(\nu_j)\}}+C_3r_{ij}n_2^im_1^je^{-i\{f(\mu_i)+f(\nu_j)\}}\\
&\quad+\nu_jp_in_2^im_2^je^{i\{f(\nu_j)-f(\mu_i)\}},
\end{aligned} \tag{5.6}$$

$p_i=1-C_3C_1^{-1}\mu_i^2$, $q_j=1-C_3C_1^{-1}\nu_j^2$. In eqs.(5.5-5.6) $f(\lambda)$ is given by (5.3) while $\nu_i, \mu_i, n_1^i, n_2^i, m_1^i, m_2^i$ denote arbitrary complex constants. In the subsequent sections we shall find the restrictions on these constants so as to assure that the solutions (5.4) verify the reductions of the $\mathcal{G}$-system.

6. The Massive Thirring Model

This model is defined by the Lagrangian

$$L=i\bar{\psi}\gamma^\mu\partial_\mu\psi+\bar{\psi}\psi+\frac{1}{4}(\bar{\psi}\gamma_\mu\psi)^2. \tag{6.1}$$

Here ψ stands for the two-dimensional spinor field,

$$\psi = \begin{pmatrix} u \\ v \end{pmatrix}$$

(and should not be confused with the Ψ-function introduced above). Other notations are standard [14]. The corresponding equations of motion read

$$i\gamma_\mu \partial^\mu \psi + \psi + \frac{1}{2}(\bar{\psi}\gamma_\mu \psi)\gamma^\mu \psi = 0, \tag{6.2}$$

or, componentwise,

$$\begin{aligned} iu_\xi + v + |v|^2 u &= 0 \\ iv_\eta + u + |u|^2 v &= 0, \end{aligned} \tag{6.3}$$

where η and ξ are given by eq.(2.2).

The integrability of this model has been discovered by Aref'eva [18] and Mikhailov [19]. The appropriate inverse scattering formalism has been developed in [20], [21]. The N-soliton solution over the zero background is also available, see [13], [17], [14].

The system (6.3) emerges from the $\mathcal{G}$-system (2.16) by identifying

$$q_1 = q_2^* = u, \qquad q_3 = q_4^* = v \tag{6.4}$$

whereby the potential matrices A_0, A_1, B_0, and B_1 in the linear problem (3.1-3.2) are hermitean. (In particular, the background solution should satisfy

$$C_1 = C_2^*, \qquad C_3 = C_4^*.) \tag{6.5}$$

Consequently, given a solution $\Psi(\lambda)$ of eqs.(3.1-3.2), $(\Psi^{-1}(\lambda^*))^\dagger$ is a solution as well and so we can identify

$$\Psi(\lambda) = \left(\Psi^{-1}(\lambda^*)\right)^\dagger H(\lambda), \tag{6.6}$$

where $H(\lambda)$ is some non-singular coordinate-independent matrix. In particular, the seed Ψ-function eq.(5.1) satisfies

$$\left(\Psi_0^{-1}(\lambda^*)\right)^\dagger = \Psi_0(\lambda). \tag{6.7}$$

As for the dressing matrix , $\chi = \Psi\Psi_{(0)}^{-1}$, eqs.(6.6) and (6.7) imply

$$\chi^\dagger(\lambda^*)\chi(\lambda) = \Psi_{(0)}(\lambda)H(\lambda)\Psi_{(0)}^{-1}(\lambda) \equiv \tilde{H}(\lambda). \tag{6.8}$$

The left-hand side of eq.(6.8) is a rational function of λ while the right-hand side exhibits essential singularities at $\lambda = 0$ and ∞. These are removed iff H is diagonal which implies, in turn, that $\tilde{H}$ is equal merely to MHM^{-1} and so does not depend

on η and ξ. Next, it is quite straightforward to demonstrate that $\tilde{H}$ does not depend on λ either (see Lemma 3.1 in [14]), i.e., it is just a constant matrix. Letting $\lambda \to \infty$ in (6.8) we observe that $\tilde{H}$ is diagonal,

$$\tilde{H} = R^\dagger R. \tag{6.9}$$

Now if we compute the product MHM^{-1} with M as in (5.2), we shall see that the off-diagonal part of this product is proportional to $(H_{11} - H_{22})$ where H_{11} and H_{22} are the elements of the matrix H. Hence, $H_{11} = H_{22}$ holds true, H and $\tilde{H}$ are proportional to the identity matrix and, in view of eq.(6.9), $\tilde{H} = 1$.

Thus the reduction to the massive Thirring model is accomplished by the involution

$$\chi(\lambda) = \left(\chi^{-1}(\lambda^*)\right)^\dagger. \tag{6.10}$$

Identifying the location of poles and the corresponding residues on the left-hand side of eq.(6.10) with those on the right-hand side, we obtain $\nu_i = \mu_i^*$, and $\vec{t}^{\,i} = (\vec{s}^{\,i})^\dagger$. Invoking then eqs.(3.25), (3.26) and (6.7), we arrive finally at

$$\nu_i = \mu_i^*, \qquad m_1^i = \left(n_1^i\right)^*, \qquad m_2^i = \left(n_2^i\right)^*, \tag{6.11}$$

$i = 1, ..., N$. This is the solution of the reduction problem.

For the reader's convenience, we reproduce here the N-soliton solution of the massive Thirring model (6.3) in a self-contained way. Using the constraints (6.11) in eqs.(5.4)-(5.6), we have

$$u = q_1 = C_1 \frac{\det b_1}{\det a_2}, \qquad v = q_3 = -\frac{1}{C_1^*}\frac{\det b_3}{\det a_1}, \tag{6.12}$$

where

$$a_1^{ij} = \tilde{a}_1^{ij}/r_{ij}, \qquad b_1^{ij} = \tilde{b}_1^{ij}/r_{ij}, \qquad b_3^{ij} = \tilde{b}_3^{ij}/r_{ij},$$

$$\begin{aligned} \tilde{b}_1^{ij} &= \nu_j q_j n_1^i n_1^{j*} e^{i\{f(\nu_i)^* - f(\nu_j)\}} + \nu_i^* \nu_j r_{ij} (C_1^* |C_1|^2)^{-1} n_2^i n_1^{j*} e^{-i\{f(\nu_i)^* + f(\nu_j)\}} \\ &\quad + \nu_i^* q_i^* n_2^i n_2^{j*} e^{i\{f(\nu_j) - f(\nu_i)^*\}}; \\ \tilde{b}_3^{ij} &= (\nu_j^2/\nu_i^*) q_i^* n_1^i n_1^{j*} e^{i\{f(\nu_i)^* - f(\nu_j)\}} - (\nu_i^* \nu_j)^{-1} r_{ij} C_1^* n_1^i n_2^{j*} e^{i\{f(\nu_i)^* + f(\nu_j)\}} \\ &\quad + (\nu_i^*)^2 \nu_j^{-1} q_j n_2^i n_2^{j*} e^{i\{f(\nu_j) - f(\nu_i)^*\}}; \\ \tilde{a}_1^{ij} &= \nu_j q_i^* n_1^i n_1^{j*} e^{i\{f(\nu_i)^* - f(\nu_j)\}} + r_{ij} C_1^{-1} n_1^i n_2^{j*} e^{i\{f(\nu_i)^* + f(\nu_j)\}} \\ &\quad + \nu_i^* q_j n_2^i n_2^{j*} e^{i\{f(\nu_j) - f(\nu_i)^*\}}; \end{aligned} \tag{6.13}$$

$\tilde{a}_2 = (\tilde{a}_1)^\dagger$, $q_j = 1 + \nu_j^2/|C_1|^2$, $r_{ij} = \nu_j^2 - (\nu_i^*)^2$, and

$$f(\lambda) = \frac{1}{2}\left\{(\lambda^2 + |C_1|^2)\eta - (\lambda^{-2} + |C_1|^{-2})\xi\right\}. \tag{6.14}$$

The one-soliton solution arises as a trivial particular case of (6.12), (6.13) – one should merely drop the indices "i" and "j" in (6.13), and omit the "det" symbols in (6.12). The propagation of the soliton is governed by the exponential $e^{i\{f(\nu)^*-f(\nu)\}}$ and its inverse, while the exponential $e^{i\{f(\nu)^*+f(\nu)\}}$ describes the internal (spatial and temporal) oscillations. Recalling that $\eta=(t+x)/2$ and $\xi=(t-x)/2$, we have

$$i\{f(\nu)^*-f(\nu)\}=\frac{i}{4}\left(\frac{1}{\nu^{*2}}-\frac{1}{\nu^2}\right)\left\{(|\nu|^4-1)x+(|\nu|^4+1)t\right\}$$

and so the velocity of the soliton is $v=\frac{|\nu|^4+1}{|\nu|^4-1}$ which is clearly greater than 1 in absolute value. Consequently, the solitons of the massive Thirring model propagating over the flat background, are tachyons.

Now if we are interested in subluminal solitons, we should interchange x and t in the definition of the model and, simultaneously, in solutions. (In terms of the light cone variables, this implies the substitution $\xi\to-\xi$). Instead of the faithful Thirring model, eqs.(6.1)-(6.2), we shall have the "modified Thirring model", with the Lagrangian

$$L=i\bar\psi\gamma^5\gamma^\mu\partial_\mu\psi+\bar\psi\psi+\frac{1}{4}(\bar\psi\gamma_\mu\psi)^2 \tag{6.15}$$

and the equations of motion

$$\begin{aligned} iu_\xi-v-|v|^2u&=0\\ iv_\eta+u+|u|^2v&=0. \end{aligned}\tag{6.16}$$

[In eq.(6.15), we have introduced $\gamma^5=\gamma^0\gamma^1=\begin{pmatrix}-1&0\\0&1\end{pmatrix}$.]

The N-soliton solution to this model is again given by eqs.(6.12- 6.13), with the only reservation that eq.(6.14) should be substituted with

$$f(\lambda)=\frac{1}{2}\left\{(\lambda^2+|C_1|^2)\eta+(\lambda^{-2}+|C_1|^{-2})\xi\right\}.$$

Turning to a single soliton, we see that

$$i\{f(\nu)^*-f(\nu)\}=\frac{i}{4}\left(\frac{1}{\nu^2}-\frac{1}{\nu^{*2}}\right)\left\{(|\nu|^4+1)x+(|\nu|^4-1)t\right\}$$

so that the soliton's velocity, $v=\frac{|\nu|^4-1}{|\nu|^4+1}$ is subluminal. Passing to the soliton's rest frame (which is accomplished by the Lorentz transformation $\eta\to|\nu|^{-2}\eta$, $\xi\to|\nu|^2\xi$), we have

$$\begin{aligned} u&=\frac{qe^{i\alpha+\sin 2\alpha\cdot z}+q^*e^{-i\alpha-\sin 2\alpha\cdot z}+2\sinh 2\alpha\cdot e^{3\beta-iz}}{qe^{-i\alpha+\sin 2\alpha\cdot z}+q^*e^{i\alpha-\sin 2\alpha\cdot z}+2\sinh 2\alpha\cdot e^{\beta-iz}},\\ v&=-\frac{q^*e^{3i\alpha+\sin 2\alpha\cdot z}+qe^{-3i\alpha-\sin 2\alpha\cdot z}-2\sinh 2\alpha\cdot e^{-\beta+iz}}{q^*e^{i\alpha+\sin 2\alpha\cdot z}+qe^{-i\alpha-\sin 2\alpha\cdot z}+2\sinh 2\alpha\cdot e^{\beta+iz}}. \end{aligned}\tag{6.17}$$

Here $q=1+\nu^2$, $z=(\cosh 2\beta+\cos 2\alpha)t-\sinh 2\beta(x-x_0)$, x_0 arbitrary constant, and we have parametrized ν as $\nu=e^{\beta+i\alpha}$ and set, for simplicity, $C_1=1$. The solution (6.17) is a kind of a stationary wobbling kink: away from a quescent finite neighbourhood of $x=0$ (of the size $\sim 2/\sin 2\alpha$), the solution is static, flat and approaches different values as $x\to\pm\infty$ ($u\to e^{\pm 2i\alpha}$, $v\to -e^{\pm 2i\alpha}$) whereas inside this neighbourhood, the solution is oscillatory.

7. Complex sine-Gordon in the Euclidean space

To pass to the Euclidean space, we substitute z for η, and z^* for ξ in the $\mathcal{G}$-system (2.16), and in its N-soliton solution (5.4-5.6). Here z and z^* are the Laplace coordinates on the (x,y) plane:

$$z=\frac{x+iy}{2},\qquad z^*=\frac{x-iy}{2}. \tag{7.1}$$

Since we did not impose any reality assumptions on η and ξ, the solution (5.4-5.6) clearly survives this transition.

The reduction to the complex sine-Gordon equation is defined by the identification[9]

$$q_1=-q_4^*,\qquad q_2=-q_3^*. \tag{7.2}$$

This constraint takes the $\mathcal{G}$-system (2.16) to

$$\begin{aligned} i\partial q_1/\partial z^*-q_2^*(1-|q_1|^2)&=0\\ i\partial q_2^*/\partial z-q_1^*(1-|q_2|^2)&=0, \end{aligned}\tag{7.3}$$

whence, eliminating q_2^* and denoting

$$q_1=\phi \tag{7.4}$$

we get

$$\frac{\partial^2\phi}{\partial z\partial z^*}+\phi(1-|\phi|^2)+\frac{\partial\phi}{\partial z}\frac{\partial\phi}{\partial z^*}\frac{\phi^*}{1-|\phi|^2}=0, \tag{7.5}$$

or, in terms of the laboratory coordinates,

$$\Delta\phi+\phi(1-|\phi|^2)+\frac{(\nabla\phi)^2\phi^*}{1-|\phi|^2}=0. \tag{7.6}$$

[9] In the case of the *zero* background [14], there were *four* combinations. Namely, we could substitute $\eta\to z$, $\xi\to\varepsilon z^*$ and identify $q_1=\tau q_4^*$, $q_2=\varepsilon\tau q_3^*$ with $\varepsilon,\tau=\pm1$. However, as one can check easily, only the choice $\varepsilon=1$, $\tau=-1$ is compatible with the flat *non-zero* background (more precisely, with the identity $C_1C_4=C_2C_3=-1$).

This is the complex sine-Gordon equation in the Euclidean domain [14]. The corresponding Lagrangian is

$$L = \frac{|\nabla\phi|^2}{1-|\phi|^2} - |\phi|^2. \tag{7.7}$$

(Here $\Delta \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$, and $\nabla \equiv \vec{i}\partial/\partial x + \vec{j}\partial/\partial y$).

It is worthwhile to mention here that the equivalent first-order system (7.3) is interesting in its own right. Indeed, it is nothing but the euclidean version of the massive Thirring model [14].

The reduction is accomplished according to the standard prescription. First of all, eqs.(7.2) as applied to the seed constant solution, yield

$$C_1 = -C_4^*, \qquad C_2 = -C_3^*. \tag{7.8}$$

Next, the linear problem matrices satisfy $B_1^\dagger = -A_1$, $B_0^\dagger = +A_0$ and so

$$\Psi(\lambda) = \left(\Psi^{-1}(-1/\lambda^*)\right)^\dagger H(\lambda). \tag{7.9}$$

In particular, the seed Ψ-function eq.(5.1-5.2) obeys

$$\Psi_{(0)}(\lambda) = \left(\Psi_{(0)}^{-1}(-1/\lambda^*)\right)^\dagger H_{(0)}(\lambda) \tag{7.10}$$

with

$$H_{(0)}(\lambda) = \frac{1}{\sqrt{C_3C_4}} \cdot \begin{pmatrix} 0 & -C_3 \\ C_4 & 0 \end{pmatrix}. \tag{7.11}$$

For the dressing matrix these imply

$$\chi^\dagger(-1/\lambda^*)\chi(\lambda) = \Psi_{(0)}(\lambda)H_{(0)}^{-1}(\lambda)H(\lambda)\Psi_{(0)}^{-1}(\lambda). \tag{7.12}$$

Following the standard route described in the previous section we obtain that the right-hand side of eq.(7.12) is equal to $\alpha \cdot \mathbf{1}$, where α is a real constant, so that we have

$$\chi(\lambda) = \alpha\left(\chi^{-1}(-\frac{1}{\lambda^*})\right)^\dagger \tag{7.13}$$

whence

$$\nu_i = -1/\mu_i^* \tag{7.14}$$

and $\vec{t}^{\,i} = (\vec{s}^{\,i})^\dagger$. (The latter equality holds modulo an unessential constant factor which cancels in the solutions). Using then (7.10), (7.11) we find

$$m_1^i = C_2 n_2^{i*}, \qquad m_2^i = -C_1 n_1^{i*}. \tag{7.15}$$

Eqs.(7.14),(7.15) furnish the ultimate form of the constraints isolating solutions of the complex sine-Gordon and the massive Thirring model in the euclidean domain.

Let us dwell on the 1-soliton solution. (Here we confine ourselves to the complex sine-Gordon case.) Using (7.14) in (5.4-5.6), we obtain

$$\phi = C_1 \frac{e^{Y-Y_0+i(2\alpha-\omega)} - \cos(X - X_0 + 2i\beta)}{e^{Y-Y_0} + \cos(X - X_0)} \tag{7.16}$$

where X and Y is a pair of independent (though non-orthogonal) coordinates on the (x,y)-plane:

$$X = \sinh 2\beta\{\cos 2\alpha \cdot x - \sin 2\alpha \cdot y\}$$
$$Y = \{\cosh 2\beta \sin 2\alpha + \sin\omega\}x + \{\cosh 2\beta \cos 2\alpha + \cos\omega\}y,$$

$\alpha, \beta, \omega, X_0$, and Y_0 are real quantities defined by $\mu = e^{\beta+i\alpha}$, $C_1C_2 = e^{i\omega}$, $(C_1/\mu + C_2^*\mu)n_1^*(n_2^*\sinh 2\beta)^{-1} = e^{Y_0+iX_0}$. The solution (7.16) is obviously periodic in X and singular along the curves $X - X_0 = \pm \arccos e^{Y-Y_0} + 2\pi k$, $k \in Z$.

8. Complex sine-Gordon in the Minkowski space

Now let us turn back to the light-cone variables, η and ξ. It is useful[10] to make the substitution $\xi \to -\xi$ in the equations of the $\mathcal{G}$-system, eqs.(2.16). Setting

$$q_1 = -q_4^*, \tag{8.1}$$

expressing q_2 and q_3 by means of the first and fourth equation in (2.16):

$$q_2 = -iq_{1\eta}^*(1-|q_1|^2)^{-1}, \qquad q_3 = iq_{1\xi}(1-|q_1|^2)^{-1}, \tag{8.2}$$

and substituting into the second and third, we have the complex sine-Gordon equation:

$$\phi_{\eta\xi} - \phi(1-|\phi|^2) + \frac{\phi^*\phi_\xi\phi_\eta}{1-|\phi|^2} = 0, \tag{8.3}$$

where $\phi \equiv q_1$. The corresponding Lagrangian is

$$L = \frac{\phi_\eta\phi_\xi^*}{1-|\phi|^2} + (|\phi|^2 - 1). \tag{8.4}$$

The integrability of this model has been established in [1], [24] and [25].

[10] Unless this change was made, we would arrive at *tachyonic* solutions of the complex sine-Gordon.

In ref.[14] the N-soliton solution over the zero background was constructed, both for eq.(8.3) and for the complex **sinh**-Gordon,

$$\phi_{\eta\xi} - \phi(1+|\phi|^2) - \frac{\phi^*\phi_\eta\phi_\xi}{1+|\phi|^2} = 0. \tag{8.5}$$

However, it is elementary to check that eq.(8.5) does not admit *non-zero* constant solutions and so we shall be dealing solely with the **sine** case here, eq.(8.3).

As it was demonstrated in [14], the reduction conditions (8.1-8.2) lead to rather non-obvious automorphism of the manifold of fundamental solutions of the linear problem (3.1-3.2):

$$\Psi(\lambda) \to G^{-1}(\lambda)\left(\Psi^{-1}(\lambda^*)\right)^\dagger \tag{8.6}$$

with

$$G(\lambda) = \frac{1}{\sqrt{1-|\phi|^2}}\begin{pmatrix} \lambda & \phi \\ \phi^* & \lambda^{-1} \end{pmatrix}. \tag{8.7}$$

Now if we take $\phi = C = const$ as a seed solution for eq.(8.3), we will necessarily have $|C|^2 = 1$ and so the corresponding matrix G is singular. Consequently, the flat non-zero solution cannot be "dressed" directly – as we have dressed the zero vacuum, for example [14]. Instead, we shall first dress an *exponential* solution and then, taking an appropriate limit, will recover the N-soliton solution over the *flat* background.

So, consider the exponential solution of the form (4.1-4.5). Imposing the condition $q_1^{(0)} = -\left(q_4^{(0)}\right)^*$ yields

$$C_1 = -C_4^*, \qquad C_1C_2 = C_1^*C_2^* \tag{8.8}$$

and z eq.(4.2) becomes pure imaginary:

$$z = i(|C_1|^2 - 1)\left(\frac{C_2}{C_1^*}\eta - \frac{C_1^*}{C_2}\xi\right). \tag{8.9}$$

Now the involution (8.6) may be written as

$$G^{-1}(\lambda)\left(\Psi^{-1}(\lambda^*)\right)^\dagger = \Psi(\lambda)H^{-1}(\lambda), \tag{8.10}$$

where $H(\lambda)$ is some non-singular coordinate-independent matrix. For the background Ψ-function, eq.(8.10) reads

$$G_{(0)}^{-1}(\lambda)\left(\Psi_{(0)}^{-1}(\lambda^*)\right)^\dagger = \Psi_{(0)}(\lambda)H_{(0)}^{-1}(\lambda), \tag{8.11}$$

where

$$G_{(0)}(\lambda) = \frac{1}{\sqrt{1-|C_1|^2}}\begin{pmatrix} \lambda & C_1e^z \\ C_1^*e^{-z} & \lambda^{-1} \end{pmatrix} \tag{8.12}$$

and $H_{(0)}$ stands for the corresponding H-matrix. Combining (8.10) and (8.11) yields for $\chi = \Psi\Psi_{(0)}^{-1}$:

$$K(\lambda) \equiv \chi^\dagger(\lambda^*)G(\lambda)\chi(\lambda) = \tag{8.13}$$

$$= G_{(0)}(\lambda)\Psi_{(0)}(\lambda)H_{(0)}^{-1}(\lambda)H(\lambda)\Psi_{(0)}^{-1}(\lambda). \tag{8.14}$$

Analogously to Sec. 6 it may be shown that $H_{(0)}^{-1}H$ is diagonal. The following simple consideration shows then that $K(\lambda)$ cannot have poles at $\lambda = \nu_i$ and ν_i^*.

Take the residue of $K(\lambda)$ at $\lambda = \nu_i$, say:

$$res\{K(\lambda), \nu_i\} = \chi^\dagger(\nu_i^*)G(\nu_i)R\vec{x}^i \otimes \vec{m}^i\Psi_{(0)}^{-1}(\nu_i). \tag{8.15}$$

On the other hand, eq.(8.14) implies

$$res\{K(\lambda), \nu_i\} = G_{(0)}(\nu_i)\Psi_{(0)}(\nu_i)\rho_i\Psi_{(0)}^{-1}(\nu_i), \tag{8.16}$$

where $\rho_i = res\{H_{(0)}^{-1}(\lambda)H(\lambda), \nu_i\}$ is a diagonal matrix. Comparing (8.15) to (8.16) we have that

$$\chi^\dagger(\nu_i^*)G(\nu_i)R\vec{x}^i \otimes \vec{m}^i = G_{(0)}(\nu_i)e^{z\sigma_3/2}M_{(0)}(\nu_i)\rho_i e^{i\sigma_3 f(\nu_i)}. \tag{8.17}$$

Now it is straightforward to see that the both sides of (8.17) depend on η, ξ in completely different ways. For example, we can multiply both sides by an arbitrary constant row-vector $\vec{c}$ on the left to obtain

$$\vec{u} = \vec{v} \tag{8.18}$$

where

$$\vec{u}(\eta,\xi) \equiv \tilde{c}(\eta,\xi)\vec{m}^i, \qquad \tilde{c}(\eta,\xi) \equiv \vec{c}\cdot\chi^\dagger(\nu_i^*)G(\nu_i)R\vec{x}^i,$$
$$\vec{v}(\eta,\xi) \equiv \vec{c}\cdot G_{(0)}(\nu_i)e^{z\sigma_3/2}M_0(\nu_i)\rho_i e^{i\sigma_3 f(\nu_i)}$$

[In (8.18) we have suppressed the superscripts for simplicity.] The vector $\vec{u}$ depends on the coordinates through the scalar function $\tilde{c}(\eta,\xi)$ so that

$$u_1/u_2 = m_1^i/m_2^i = const.$$

On the contrary, the dependence of v_1/v_2 on the coordinates is explicit and nontrivial. The contradiction is resolved only if $\tilde{c}(\eta,\xi) = 0$ for any $\vec{c}$ which implies, in turn, that

$$res\{K(\lambda), \nu_i\} = 0,$$

$i = 1, .., N$. In the same way, $res\{K(\lambda), \nu_i^*\} = 0$. Thus, only three terms survive in the expansion of $K(\lambda)$ over residues:

$$K(\lambda) = K_+\lambda + K_0 + K_-\lambda^{-1}. \tag{8.19}$$

The coefficients $K_0, K_\pm$ are easily computed by the definition of K, eq.(8.13). So we have

$$K(\lambda) = \frac{1}{\sqrt{1-|\phi|^2}} \left|\frac{\det a_1}{\det a_2}\right| \begin{pmatrix} \lambda & C_1 e^z \\ C_1^* e^{-z} & \kappa\lambda^{-1} \end{pmatrix} \tag{8.20}$$

with $\kappa = \prod_{i=1}^N |\frac{\mu_i}{\nu_i}|^2$. [Note that the determinant of $K(\lambda)$,

$$\det K(\lambda) = \frac{1}{1-|\phi|^2} \left|\frac{\det a_1}{\det a_2}\right|^2 (\kappa - |C_1|^2) \tag{8.21}$$

is λ-independent.] Alternatively, we may compute $\det K$ directly from (8.13):

$$\det K(\lambda) = \det K(\infty) = \det \chi^\dagger(\infty) \cdot (\det G)_{|\infty} \cdot \det \chi(\infty) = 1. \tag{8.22}$$

Then eq.(8.21) yields

$$\frac{1}{1-|\phi|^2} \left|\frac{\det a_1}{\det a_2}\right|^2 = \frac{1}{(\kappa - |C_1|^2)}$$

and $K(\lambda)$ acquires the following simple form:

$$K(\lambda) = \frac{1}{\sqrt{\kappa - |C_1|^2}} \begin{pmatrix} \lambda & C_1 e^z \\ C_1^* e^{-z} & \kappa\lambda^{-1} \end{pmatrix}. \tag{8.23}$$

Now let us return to eq.(8.13):

$$\left(\chi^{-1}(\lambda^*)\right)^\dagger = G(\lambda)\chi(\lambda). \tag{8.24}$$

Comparing the location of poles and the corresponding residues at l.h.s. to those at r.h.s., we obtain

$$\mu_i = \nu_i^*, \tag{8.25}$$

$$\vec{t}^i = \left(\vec{s}^i\right)^\dagger K(\nu_i). \tag{8.26}$$

So by virtue of eq.(8.25), $\kappa = 1$ and we observe that

$$K(\lambda) = G_{(0)}(\lambda). \tag{8.27}$$

Now recalling that $\vec{t}^i = \vec{m}^i \Psi_{(0)}^{-1}(\nu_i)$ and $(\vec{s}^i)^\dagger = (\vec{n}^i)^\dagger \Psi_{(0)}^\dagger(\mu_i)$, using (8.11) and(8.27), eq.(8.26) yields

$$\vec{m}^i = (\vec{n}^i)^\dagger H_{(0)}(\nu_i). \tag{8.28}$$

$H_{(0)}(\lambda)$ can be computed by virtue of eq.(8.11):

$$H_0(\lambda) = \Psi_{(0)}^\dagger(\lambda^*) G_{(0)}(\lambda) \Psi_{(0)}(\lambda) = \frac{1}{\sqrt{1-|C_1|^2}} \times \frac{1}{1 + C_1^{-1} C_2^{-1} J^2}$$
$$\times \begin{pmatrix} \lambda - 2J + J^2\lambda^{-1}|C_1|^{-2} & 0 \\ 0 & \lambda^{-1} + 2C_1^* C_2^{-1} J + \lambda J^2 |C_2|^{-2} \end{pmatrix}. \tag{8.29}$$

[We remind that $J = J(\lambda)$ is given by eq.(4.20)].

Substituting formulas (8.25), (8.28), and (8.29) into eq.(3.49) we arrive at the N-soliton solution of the complex sine-Gordon equation over the *exponential* background. The problem is now to find the limit of this solution corresponding to the *flat* background. To this end, let us define an infinitesimal parameter

$$\epsilon \equiv 1 - |C_1|^2. \tag{8.30}$$

(The flat background corresponds to $|C_1|^2 = 1$ and thereby to $\epsilon = 0$.) Eq.(4.5) implies then

$$C_2 C_3 = \epsilon - 1. \tag{8.31}$$

Expanding $J(\lambda)$ in terms of ϵ, we find

$$J(\lambda) = \lambda \left(1 - \frac{C_1 C_2}{\lambda^2 + C_1 C_2}\epsilon\right) + O(\epsilon^2) \tag{8.32}$$

and so eq.(8.29) is replaced with

$$H_0(\lambda) = \begin{pmatrix} \frac{\lambda}{1+\lambda^2/(C_1C_2)}\epsilon^{1/2} + O(\epsilon^{3/2}) & 0 \\ 0 & \frac{1+\lambda^2/(C_1C_2)}{\lambda}\epsilon^{-1/2} + O(\epsilon^{1/2}) \end{pmatrix}. \tag{8.33}$$

Using this in eq.(8.28) we can express m_1 and n_2 through n_1^* and m_2^*, respectively:

$$m_1^i = \frac{\nu_i}{1 + \nu_i^2/(C_1C_2)} \epsilon^{1/2} (n_1^i)^* \equiv \gamma_i \epsilon^{1/2} (n_1^i)^*$$
$$n_2^i = \frac{\nu_i^*}{1 + \nu_i^{*2}/(C_1C_2)} \epsilon^{1/2} (m_2^i)^* \equiv \gamma_i^* \epsilon^{1/2} (m_2^i)^*. \tag{8.34}$$

Now let us forget about the reduction for a moment and consider the N-soliton solutions (3.49), more precisely, the matrices $a_1, a_2, b_1, ..., b_4$. As we saw in Sec. 5, passing from the exponential to the flat background, the matrix elements acquire

the simplified three-term structure [eq.(5.5-5.6)]. In other words, if we expanded $a_1, ..., b_4$ in terms of ϵ, the $O(\epsilon^0)$ term would be given by eqs.(5.5-5.6). What is also important is that these matrices are analytic functions of ϵ and so terms of order $\epsilon^{1/2}$ cannot emerge in the expansion. Symbolically,

$$a_2^{EXP}(n_1^i, n_2^i, m_1^i, m_2^i) = a_2^{FLAT}(n_1^i, n_2^i, m_1^i, m_2^i) + O(\epsilon),$$
$$b_1^{EXP}(n_1^i, n_2^i, m_1^i, m_2^i) = b_2^{FLAT}(n_1^i, n_2^i, m_1^i, m_2^i) + O(\epsilon), \qquad (8.35)$$

and so on.

Now let us invoke the reduction conditions (8.34) and express m_1^i and n_2^i through n_1^i and m_2^i. The quantities we are interested in, are the "reduced" a_2^{EXP} and b_1^{EXP}, i.e.,

$$a_2^{EXP}(n_1^i, \gamma_i^* \epsilon^{1/2} m_2^{i*}, \gamma_i \epsilon^{1/2} n_1^{i*}, m_2^i), \quad b_1^{EXP}(n_1^i, \gamma_i^* \epsilon^{1/2} m_2^{i*}, \gamma_i \epsilon^{1/2} n_1^{i*}, m_2^i)$$

However, using (8.35),we have

$$a_2^{EXP}(n_1^i, \gamma_i^* \epsilon^{1/2} m_2^{i*}, \gamma_i \epsilon^{1/2} n_1^{i*}, m_2^i) =$$
$$= a_2^{FLAT}(n_1^i, \gamma_i^* \epsilon^{1/2} m_2^{i*}, \gamma_i \epsilon^{1/2} n_1^{i*}, m_2^i) + O(\epsilon) \qquad (8.36)$$

and a similar relation for b_1. These equations imply that when $\epsilon \to 0$, the reduced form of a_2^{EXP} and b_1^{EXP} coincides, up to the higher order terms, with the reduced form of a_2^{FLAT} and b_1^{FLAT}, respectively.

This observation simplifies the calculation drastically. Indeed, we want to find the *flat*-background limit of the complex sine-Gordon multisoliton solution over the *exponential* background. To do so, one can substitute eq.(8.34) into the (non-reduced) N-soliton formula (3.49) (thereby arriving at sine-Gordon's solitons over the exponential background) and then send $\epsilon \to 0$. However, eq.(8.36) tells us that it is quite sufficient to substitute the restrictions (8.34) into the much simpler formulas (5.5-5.6) (*already corresponding to the flat background!*) and then take $\epsilon \to 0$ limit.

Both b_1 and a_2 eq.(5.5-5.6) comprise three terms which are proportional to $n_1^i m_1^j$, $n_2^i m_1^j$, and $n_2^i m_2^j$, respectively. The first and the third terms are proportional to $\epsilon^{1/2}$ and will dominate the second one which is $O(\epsilon)$. Consequently, both $\det b_1$ and $\det a_2$ are multiples of $(\epsilon^{1/2})^N$ which cancels in $\det b_1 / \det a_2$. Finally, the N-soliton solution of the complex sine-Gordon equation (8.3) over the flat non-zero background

$$\phi^{(0)}(\eta, \xi) \equiv e^{i\theta_0} \qquad (8.37)$$

($\theta_0 =$ real const), is given by

$$\phi = e^{i\theta_0} \prod_{k=1}^{N} \frac{1}{|\nu_k|^2} \frac{\det b_1}{\det \hat{a}_2} \qquad (8.38)$$

with

$$b_1^{ij} = \frac{\nu_j^2 n_1^i n_1^{j*} e^{\zeta_{ij}} + \nu_i^{*2} m_2^{i*} m_2^j e^{-\zeta_{ij}}}{\nu_j^2 - \nu_i^{*2}}, \qquad \hat{a}_2^{ij} = \frac{n_1^i n_1^{j*} e^{\zeta_{ij}} + m_2^{i*} m_2^j e^{-\zeta_{ij}}}{\nu_j^2 - \nu_i^{*2}},$$

$$\zeta_{ij} = i\{f(\nu_i^*) - f(\nu_j)\} = \frac{i}{2}\left\{(\nu_i^{*2} - \nu_j^2)\eta + (1/\nu_i^{*2} - 1/\nu_j^2)\xi\right\}. \qquad (8.39)$$

In particular, the 1-soliton solution is

$$\phi = e^{i\theta_0}\{\cos 2\alpha + i \sin 2\alpha \cdot \tanh(\zeta - \zeta_0)\}, \qquad (8.40)$$

where

$$\zeta = \sin 2\alpha (e^{-2\beta}\eta - e^{2\beta}\xi) = \sin 2\alpha (\cosh 2\beta \cdot x - \sinh 2\beta \cdot t),$$

$\zeta_0 = \ln|m_2/n_1|$, and α, β are defined by $\nu = e^{-\beta + i\alpha}$. Eq.(8.40) is nothing but the (subluminal) kink moving with the velocity $v = \tanh 2\beta$.

9. O(1,1) sine-Gordon

In this section we again make a substitution in the equations of motion of the $\mathcal{G}$-system (2.16). Namely, we substitute $i\eta$ for η and $(-i\xi)$ for ξ. Expressing then q_2 and q_3 via eqs.(2.16.4) and (2.16.1), respectively,

$$q_2 = \frac{q_{4\eta}}{1 + q_1 q_4}, \qquad q_3 = \frac{q_{1\xi}}{1 + q_1 q_4}, \qquad (9.1)$$

substituting (9.1) into the remaining two equations and defining

$$\phi^+ \equiv -q_4, \qquad \phi^- \equiv q_1, \qquad (9.2)$$

we have

$$\phi^-_{\eta\xi} + \phi^-(1 - \phi^+\phi^-) + \phi^+ \frac{\phi^-_\eta \phi^-_\xi}{1 - \phi^+\phi^-} = 0,$$
$$\phi^+_{\eta\xi} + \phi^+(1 - \phi^+\phi^-) + \phi^- \frac{\phi^+_\eta \phi^+_\xi}{1 - \phi^+\phi^-} = 0. \qquad (9.3)$$

If we consider ϕ^+ and ϕ^- to be *real* quantities, eqs.(9.3) are equations of motion for another scalar model [14] with the Lagrangian

$$L = \frac{\phi^+_\eta \phi^-_\xi}{1 - \phi^+\phi^-} + 1 - \phi^+\phi^-. \qquad (9.4)$$

We have named this model "O(1,1) sine-Gordon equation" [14] and the reason is that $O(1,1)$ is its group of internal symmetry. This fact becomes evident if we replace ϕ^+, ϕ^- with their sum and difference: $\phi^\pm \equiv u_1 \pm u_2$. Then (9.4) is

$$L = \frac{u_{1\xi} u_{1\eta} - u_{2\xi} u_{2\eta}}{1 - (u_1^2 - u_2^2)} + 1 - (u_1^2 - u_2^2). \qquad (9.5)$$

There is a close resemblance between (9.5) and the complex sine-Gordon (8.4) whose symmetry group is $U(1)$ or, equivalently, $O(2)$. The similarity becomes even more striking if we reformulate the models in the Regge-Lund-like notation [23] [24]. Assume

$$u_1^2 \geq u_2^2 \tag{9.6}$$

or, equivalently,

$$\phi^+\phi^- \geq 0. \tag{9.7}$$

Defining the Regge-Lund variables α and β,

$$u_1 = \cos\alpha\cosh\beta, \qquad u_2 = \cos\alpha\sinh\beta, \tag{9.8}$$

the $O(1,1)$ sine-Gordon model (9.5) is taken to

$$L = \alpha_\eta\alpha_\xi - \cot^2\alpha\cdot\beta_\eta\beta_\xi + \sin^2\alpha. \tag{9.9}$$

On the other hand, the Regge-Lund form of the complex sine-Gordon (8.4) is [23] [24]:

$$L = \alpha_\eta\alpha_\xi + \cot^2\alpha\cdot\beta_\eta\beta_\xi - \sin^2\alpha. \tag{9.10}$$

Lastly, we shall provide the *complex* formulation of the $O(1,1)$ sine-Gordon equation. Keeping in mind the condition (9.6), we introduce a complex field $\phi = \rho e^{i\theta}$ with ρ and θ defined as

$$u_1 = \rho\cosh\theta, \qquad u_2 = \rho\sinh\theta. \tag{9.11}$$

Eq.(9.5) passes then to

$$L = \frac{\phi_\eta\phi_\xi}{1-|\phi|^2}\cdot\frac{\phi^*}{\phi} + 1 - |\phi|^2 + c.c., \tag{9.12}$$

or, in the covariant notation,

$$L = \frac{|\partial_\mu\phi|^2}{1-|\phi|^2} + 1 - |\phi|^2 - \frac{1}{2}\frac{J_\mu^2}{|\phi|^2(1-|\phi|^2)}. \tag{9.13}$$

Here $J_\mu = i(\phi^*\partial_\mu\phi - \phi\partial_\mu\phi^*)$.

As we have seen, the reduction to $O(1,1)$ sine-Gordon is defined by the requirement that all fields $q_1, ..., q_4$ be real. This implies the following involution:

$$\Psi(\lambda) = \Psi^*(\lambda^*)H(\lambda). \tag{9.14}$$

[We remind that the substitution $i\partial_\eta \to \partial_\eta$, $i\partial_\xi \to -\partial_\xi$ was made in the linear system (3.1-3.2)]. For example, the seed Ψ-function (5.1) corresponding to the flat non-zero background satisfies

$$\Psi_{(0)}(\lambda) = \Psi^*_{(0)}(\lambda^*). \tag{9.15}$$

(Here we assumed that $C_1, ..., C_4$ are real.) Accordingly, we have for $\chi = \Psi\Psi_{(0)}^{-1}$:

$$\chi^{-1}(\lambda)\chi^*(\lambda^*) = \Psi_{(0)}(\lambda)H^{-1}(\lambda)\Psi_{(0)}^{-1}(\lambda). \tag{9.16}$$

In complete analogy with Sec.6 this equation can be used to demonstrate that, up to an unimportant sign factor, $H = 1$. Thus

$$\chi(\lambda) = \chi^*(\lambda^*), \qquad \chi^{-1}(\lambda) = \left(\chi^{-1}(\lambda^*)\right)^*, \tag{9.17}$$

whence

$$\nu_i^* = \delta_i\nu_{(i)}, \qquad \mu_i^* = \gamma_i\mu_{[i]}; \tag{9.18}$$

$$(s_1^i)^* = s_1^{[i]}, \qquad (s_2^i)^* = \gamma_i s_2^{[i]},$$
$$(t_1^i)^* = t_1^{(i)}, \qquad (t_2^i)^* = \delta_i t_2^{(i)}; \tag{9.19}$$

$$R = R^*. \tag{9.20}$$

In eqs.(9.18-9.19) γ_i and δ_i are (uncorrelated) sign factors, $\gamma_i = \pm 1$, $\delta_i = \pm 1$ ($i = 1, ..., N$). Next, $(.)$ denotes a permutation of indices: $\{1, ..., N\} \to \{(1), ..., (N)\}$. For example, for the permutation $\{1, 2, 3, 4, ..., N\} \to \{4, 7, 11, 1, ..., N\}$ we would have: $(1) = 4$, $(2) = 7$, $(3) = 11$, $(4) = 1$,..., $(N) = N$. The square brackets, $[\cdot]$ denote an independent permutation: $\{1, ..., N\} \to \{[1], ..., [N]\}$. The permutations $(\cdot)$ and $[\cdot]$ are not arbitrary but involutive, i.e. should satisfy $((i)) = [[i]] = i$ for any $i \in \{1, ..., N\}$.

Eqs.(9.18) imply that the poles of $\chi(\lambda)$ can be either located on the real or imaginary axis, or form pairs symmetric w.r.t. the real or imaginary axis. The same applies to $\chi^{-1}(\lambda)$. It turns out that the locations of poles of χ and χ^{-1} are not completely independent, however. Indeed, it is easy to check that (9.18 -9.19) imply for the matrices (3.20):

$$(a_1^{ij})^* = \delta_j a_1^{[i](j)}, \qquad (a_2^{ij})^* = \gamma_i a_2^{[i](j)},$$

whence, by eq.(3.40)

$$(r^2)^* = \left(\frac{\det a_1}{\det a_2}\right)^* = \prod_{i=1}^N(\gamma_i\delta_i)\cdot\frac{\det a_1}{\det a_2} = \prod_{i=1}^N(\gamma_i\delta_i)\cdot r^2.$$

Consequently, in order that r be real [which is required by eq.(9.20)] it is necessary that

$$\prod_{i=1}^N\gamma_i\delta_i = 1. \tag{9.21}$$

Finally, eq.(9.19) yields for m and n:

$$(n_1^i)^* = n_1^{[i]}, \qquad (n_2^i)^* = \gamma_i n_2^{[i]},$$
$$(m_1^i)^* = m_1^{(i)}, \qquad (m_2^i)^* = \delta_i m_2^{(i)} \tag{9.22}$$

The restrictions (9.18) and (9.22) where γ_i and δ_i obey the constraint (9.21), isolate the N-soliton solution of the $O(1,1)$ sine-Gordon equation.

10. 1-soliton solution of O(1,1) sine-Gordon

Let us consider the $N = 1$ case in some more detail. In this case we have that m_1 and n_1 are real while μ, ν, m_2, and n_2 are either all real or all imaginary. Let us put, for simplicity, $C_1 = -C_4 = 1$. (Actually there is no loss of generality in this choice since equations (9.3) are invariant w.r.t. the transformation $\phi^+ \to a\phi^+$, $\phi^- \to a^{-1}\phi^-$ with a real constant.) Eqs.(5.4-5.6) reduce then to

$$\phi^- = q_1 = \frac{\nu(1-C_3\nu^2)n_1m_1e^{2z_1} + \mu(1-C_3\mu^2)n_2m_2e^{2z_2} + C_3^2\mu\nu(\nu^2-\mu^2)n_2m_1}{\mu(1-C_3\nu^2)n_1m_1e^{2z_1} + \nu(1-C_3\mu^2)n_2m_2e^{2z_2} + C_3(\nu^2-\mu^2)n_2m_1}$$
$$\phi^+ = -q_4 = \frac{\frac{\mu^2}{\nu}(1-C_3\nu^2)n_1m_1e^{2z_1} + \frac{\nu^2}{\mu}(1-C_3\mu^2)n_2m_2e^{2z_2} + \frac{1}{\mu\nu}(\nu^2-\mu^2)n_2m_1}{\mu(1-C_3\nu^2)n_1m_1e^{2z_1} + \nu(1-C_3\mu^2)n_2m_2e^{2z_2} + C_3(\nu^2-\mu^2)n_2m_1} \tag{10.1}$$

with C_3 real. In eq.(10.1) we have introduced important variables z_1 and z_2:

$$z_1 = if(\mu) = -\frac{1}{2}\{(\mu^2 - C_3^{-1})\eta + (\mu^{-2} - C_3)\xi\} = \sinh(\alpha-\beta_1)(\cosh r_1 \cdot x - \sinh r_1 \cdot t),$$
$$z_2 = if(\nu) = -\frac{1}{2}\{(\nu^2 - C_3^{-1})\eta + (\nu^{-2} - C_3)\xi\} = \sinh(\alpha-\beta_2)(\cosh r_2 \cdot x - \sinh r_2 \cdot t),$$

where β_1, β_2, α, r_1, and r_2 are defined as follows:

$$\mu = e^{\beta_1}, \qquad \nu = e^{\beta_2}, \qquad C_3 = e^{-2\alpha}, \tag{10.2}$$
$$r_1 = -(\beta_1 + \alpha), \qquad r_2 = -(\beta_2 + \alpha). \tag{10.3}$$

Solution (10.1) admits several non-trivial degenerate cases. For convenience, let us write it as

$$\phi^\pm = \frac{A_1^\pm e^{2z_1} + A_2^\pm e^{2z_2} + A_3^\pm}{\tilde{A}_1 e^{2z_1} + \tilde{A}_2 e^{2z_2} + \tilde{A}_3} = \tag{10.4}$$
$$= \frac{A_1^\pm e^{2z_3} + A_2^\pm + A_3^\pm e^{-2z_2}}{\tilde{A}_1 e^{2z_3} + \tilde{A}_2 + \tilde{A}_3 e^{-2z_2}}, \tag{10.5}$$

where constants $A^\pm_{1,2,3}$ and $\tilde{A}_{1,2,3}$ are defined in an obvious way,

$$z_3 = z_1 - z_2 = \sinh(\beta_2 - \beta_1)(\cosh r_3 \cdot x - \sinh r_3 \cdot t), \tag{10.6}$$

and

$$r_3 = -(\beta_1 + \beta_2). \tag{10.7}$$

Now let $m_2 = 0$ (and hence $A_2^\pm = \tilde{A}_2 = 0$). Eq.(10.4) becomes then

$$\phi_1^\pm = \frac{(A_1^\pm/\tilde{A}_3)e^{2z_1} + A_3^\pm/\tilde{A}_3}{(\tilde{A}_1/\tilde{A}_3)e^{2z_1} + 1} =$$
$$= e^{\pm(\alpha-\beta_2)}\{\cosh(\alpha-\beta_1) \mp \sinh(\alpha-\beta_1)\tanh\tilde{z}_1\} \tag{10.8}$$

where $\tilde{z}_1 = z_1 + \frac{1}{2}\ln(\tilde{A}_1/\tilde{A}_3)$. If we let $n_1 = 0$ (and so $A_1^\pm = \tilde{A}_1 = 0$), eq.(10.4) reduces to

$$\phi_2^\pm = e^{\pm(\alpha-\beta_1)}\{\cosh(\alpha-\beta_2) \mp \sinh(\alpha-\beta_2)\tanh\tilde{z}_2\} \tag{10.9}$$

with $\tilde{z}_2 = z_2 + \frac{1}{2}\ln(\tilde{A}_2/\tilde{A}_3)$. The third interesting limiting case arises if we set $n_2 = \epsilon\tilde{n}_2$, $m_1 = \epsilon\tilde{m}_1$ and then let $\epsilon \to 0$. [This amounts to setting $A_3^\pm = \tilde{A}_3 = 0$ in (10.5)]. This gives

$$\phi_3^\pm = \cosh(\beta_1 - \beta_2) \pm \sinh(\beta_1 - \beta_2)\tanh\tilde{z}_3 \tag{10.10}$$

with $\tilde{z}_3 = z_3 + \frac{1}{2}\ln(\tilde{A}_1/\tilde{A}_2)$.

Let us restrict ourselves to the case of **real** μ, ν, m_2 and n_2; the imaginary case turns out to be equivalent. Assume also that $C_3 > 0$ and that α, β_1, β_2 are real. The solutions ϕ_1, ϕ_2 and ϕ_3 describe then three subluminal kinks moving with rapidities[11] r_1, r_2 and r_3, respectively. Assume we that C_3 is < 0, the quantities r_1 and r_2 would have a non-zero imaginary part equal to $\frac{\pi}{2}$ and so the kinks ϕ_1 and ϕ_2 would travel with *superluminal* velocities

$$v_{1,2} = \tanh r_{1,2} = \coth(\mathrm{Re}\, r_{1,2}).$$

However, since the tachyonic case has been discussed in detail in a separate paper [15], we shall not dwell on this possibility here.

The general 3-term solution (10.4-10.5) describes an interaction of ϕ_1, ϕ_2 and ϕ_3. Let us set for definiteness $r_3 > 0 \geq r_1 > r_2$ and demonstrate that this interaction is the decay of the soliton ϕ_1 into two solitons, ϕ_2 and ϕ_3 (Fig. 1).

It is straightforward to see from eq.(10.4) that the soliton ϕ_1 is observed when $\tilde{z}_1 \sim 0$ and $\tilde{z}_2 \to -\infty$ while ϕ_2 is observed when $\tilde{z}_2 \sim 0$ and $\tilde{z}_1 \to -\infty$. Eq.(10.5) tells us, on the other hand, that ϕ_3 is observed when $\tilde{z}_3 \sim 0$ and $\tilde{z}_2 \to +\infty$. Now the location of the centre of ϕ_1 is determined by $\tilde{z}_1 = 0$, i.e. $x = \tanh r_1 \cdot t + x_1^0$. On this line we have

$$\tilde{z}_2 = \sinh(\alpha-\beta_2)\frac{\sinh(r_1 - r_2)}{\cosh r_1}t + const. \tag{10.11}$$

[11] We remind that rapidity (r) is the quantity related to the velocity v by $v = \tanh r$. In relativistic kinematics, it is more convenient to use rapidities rather than velocities because under the Lorentz transformations rapidities merely add.

Since $\alpha - \beta_2 = r_3 - r_1 > 0$, $\tilde{z}_2 \to -\infty$ as $t \to -\infty$ and so ϕ_1 is observed only at large negative times. Analogously, ϕ_2 is seen only for large *positive* t. As for ϕ_3, the location of its center is defined by $\tilde{z}_3 = 0$, i.e. $x = \tanh r_3 \cdot t + x_3^0$. On this line

$$\tilde{z}_2 = \sinh(\alpha - \beta_2)\frac{\sinh(r_3 - r_2)}{\cosh r_3}t + const \qquad (10.12)$$

which $\to +\infty$ as $t \to +\infty$, and so ϕ_3 also appears only for large $t > 0$.

This process is shown in Fig. 1 where we have plotted the function $\phi^-(x,t)$ eq.(10.1) for some particular choice of r_1, r_2 and r_3.

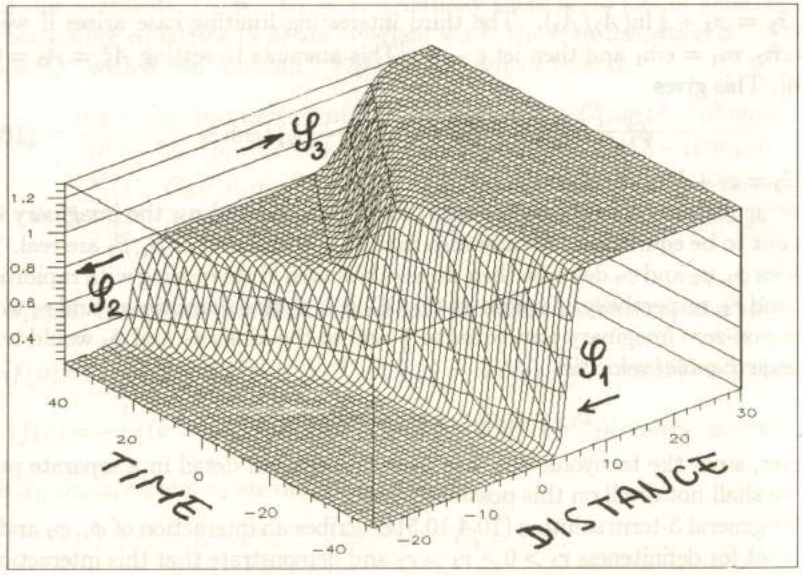


Fig. 1
The decay $\phi_1 \to \phi_2 + \phi_3$ described by eq.(10.1) for $\nu = 0.9$, $\mu = 0.7$, and $C_3 = 0.4$. (The ϕ^--component is plotted.)

It is easy to verify that for C_3, μ, and ν positive, the "interactive" solution eq.(10.1) satisfies the condition (9.7): $\phi^+\phi^- \geq 0$. So we may transform it both to the Regge-Lund-like variables (9.8), and to the complex field $\phi = \rho e^{i\theta}$ with ρ and θ as in (9.11). An important consequence of this fact is that both the Regge-Lund-like and complex-field formulations of the O(1,1) sine-Gordon model, i.e.,

$$L = (\partial_\mu \alpha)^2 - \cot^2\alpha(\partial_\mu\beta)^2 + \sin^2\alpha$$

and

$$L = \frac{|\partial_\mu\phi|^2}{1-|\phi|^2} + 1 - |\phi|^2 - \frac{1}{2}\frac{J_\mu^2}{|\phi|^2(1-|\phi|^2)},$$

respectively, exhibit nontrivial interaction of solitons (decays and fusion).

11. Concluding Remarks

a) On the Thirring kinks, Sec.6. We would like to underline once more that solitons of the massive Thirring model propagating over the flat non-zero background, are tachyons. It is the model (6.15) (differing from the Thirring model in the presence of the γ^5 factor in the kinetic term) that possesses normal, subluminal kinks.

b) On the complex sine-Gordon in the Euclidean space, Sec.7. The complex sine-Gordon equation in the Euclidean space turns out *not* to possess regular solitons over the flat non-zero background. It is in contrast to the case of the zero background where there is a rich spectrum of regular solitons [14].

c) On the nontrivial interaction of solitons and the O(1,1) sine-Gordon model, Sec.9-10. In this paper we have confined ourselves to the case when all the three interacting solitons travel slower than light, $|v| < 1$. However, the $O(1,1)$ sine-Gordon model admits more general solutions describing e.g. the decay of a subluminal soliton into a pair of a *sub*luminal and *super*luminal soliton. This possibility has been analysed in a separate paper [15]. An interesting by-product of these studies was an unexpected application to the mathematical theory of tachyons.

It is also worth mentioning here that the mechanism of the described nontrivial interaction of solitons is different from previously considered mechanisms which were realized in (more complicated) systems that are integrable on algebras of rank ≥ 2 (see e.g. [2], [26]). In our case (rank=1) the creation and annihilation of solitons becomes possible thanks to the fact that the group of internal symmetry in eq.(9.5) is non-compact. With regards to such simple one- or two-component relativistic models, it is appropriate to note that studies of inelastic interactions have been conducted (numerically) only within the frame of nonintegrable systems.

Our final remark here is that the $O(1,1)$ sine-Gordon equation is transformable into some classical spinor model. Namely, defining complex u and v by

$$\mathrm{Re}\,u = \frac{1}{2}\phi^+, \qquad \mathrm{Im}\,u = \frac{1}{2}\frac{\phi_\eta^-}{1-\phi^+\phi^-}$$
$$\mathrm{Re}\,v = -\frac{1}{2}\phi^-, \qquad \mathrm{Im}\,v = -\frac{1}{2}\frac{\phi_\xi^+}{1-\phi^+\phi^-}, \qquad (11.1)$$

the system (9.3) is taken to the first order system [14] with the Lagrangian

$$L = iu_\xi u^* + iv_\eta v^* + uv^* + u^*v + L_{int},$$
$$L_{int} = -\tfrac{1}{2}(u^2 - u^{*2})(v^2 - v^{*2}). \qquad (11.2)$$

Having the substitution (11.1) in mind, the N-soliton solution for this spinor model is given by eqs.(5.4)-(5.6) with the constraints (9.18), (9.22), and (9.21). Unlike the massive Thirring model, the system (11.2) possesses the non-trivial interaction of solitons.

Acknowledgment

One of the authors (Barashenkov) wishes to thank Dr. Ziemowit Popowicz and his colleagues in Wroclaw University for their kind hospitality in Silezia.

A Appendix: the Closed Expression for $\tilde{\chi}_{|\lambda=0}$

Here we compute the matrix (3.34),

$$\tilde{\chi}_{|\lambda=0} = const \times \begin{pmatrix} p(\eta,\xi) & 0 \\ 0 & p^{-1}(\eta,\xi) \end{pmatrix}. \tag{A1}$$

Let us begin with listing several useful identities. First of all, from eqs.(3.17) it follows that

$$s_1^i t_1^j = a_1^{ij}\nu_j - \mu_i a_2^{ij}, \qquad s_2^i t_2^j = a_2^{ij}\nu_j - \mu_i a_1^{ij}$$

or, in the matrix form,

$$\begin{aligned} |s_1><t_1| &= a_1 <\nu| - |\mu> a_2, \\ |s_2><t_2| &= a_2 <\nu| - |\mu> a_1. \end{aligned} \tag{A2}$$

Another important identity is

$$\det(a + |u><v|) = \det a + <v|\mathcal{A}|u>, \tag{A3}$$

where a is a non-singular $N\times N$ matrix; $\mathcal{A}$ is the augmented matrix, $\mathcal{A} = \det a \cdot a^{-1}$; $<v|$ and $|u>$ are N-dimensional row- and column-vector, respectively: $<v| = (v^1, ..., v^N)$, and

$$|u>= \begin{pmatrix} u^1 \\ \vdots \\ u^N \end{pmatrix}.$$

Consider now $(\tilde{\chi}_{11})_{|\lambda=0}$:

$$(\tilde{\chi}_{11})_{|\lambda=0} = 1 - 2\sum_{i=1}^{N} \nu_i^{-1} P_{11}^i = 1 - 2\sum_{i=1}^{N} \nu_i^{-1} t_1^i x_1^i = 1 - 2 <\nu^{-1} t_1 | x_1>.$$

Using (3.18), this can be rewritten as $1 - 2 <\nu^{-1} t_1 | a_1^{-1} | s_1>$. Now let us invoke eq.(A3):

$$(\tilde{\chi}_{11})_{|\lambda=0} = (\det a_1)^{-1}(\det a_1 - <t_1\nu^{-1}|\mathcal{A}_1|s_1>) = (\det a_1)^{-1} \det(a_1 - |s_1><t_1\nu^{-1}|).$$

Finally, making use of (A2), this transforms to

$$(\tilde{\chi}_{11})_{|\lambda=0} = \frac{\det(|\mu> a_2 <\nu^{-1}|)}{\det a_1} = \prod_{i=1}^{N} \left(\frac{\mu_i}{\nu_i}\right) \frac{\det a_2}{\det a_1},$$

In the same way,

$$(\tilde{\chi}_{22})_{|\lambda=0} = \prod_{i=1}^{N} \left(\frac{\mu_i}{\nu_i}\right) \frac{\det a_1}{\det a_2},$$

so that comparing to (A1) we arrive finally at

$$p(\eta,\xi) = \det a_2 / \det a_1. \tag{A4}$$

References

[1] K. Pohlmeyer. Commun.Math.Phys. **46** 207 (1976)

[2] A.S. Budagov and L.A. Takhtajan. DAN SSSR **235** 805 (1977); LOMI Notices **77** 24 (1978)

[3] V.E. Zakharov, A.V. Mikhailov ZhETF **74** 1953 (1978) [translation in Sov.Phys.JETP **47** 1017 (1979)]

[4] V.E. Zakharov and A.B. Shabat. Funct.Anal.Prilozh. **13** 13 (1979) [translation in Funct.Anal.Appl. **13** 166 (1979)]

[5] K. Pohlmeyer, K.-H. Rehren. J.Math.Phys. **20** 2628 (1979)

[6] A. Neveu and N. Papanicolaou. Commun.Math.Phys. **58** 31 (1978)

[7] A.V. Mikhailov. Physica **D3** 73 (1981)

[8] A.V. Mikhailov, M.A. Olshanetsky and A.M. Perelomov. Commun.Math.Phys. **79** 473 (1981)

[9] A.P. Fordy and J. Gibbons. Commun.Math.Phys. **77** 21 (1980)

[10] A.N. Leznov and M.V. Saveliev. Phys.Elem.Part.Atom.Nucl. **12** 125 (1981)

[11] V.E. Zakharov and A.V. Mikhailov. Commun. Math. Phys. **74** 21 (1980)

[12] B.S. Getmanov. In: Proc. III Int. Symp. on Statistical Mechanics, Dubna, JINR D17-84-407, pp.212 and 217 (1984)

[13] I.V. Barashenkov and B.S. Getmanov. In: Proc. III Int. Symp. on Statistical Mechanics, Dubna, JINR D17-84-407, p.37 (1984)

[14] I.V. Barashenkov and B.S. Getmanov. Commun.Math.Phys. **112** 423 (1987)

[15] I.V. Barashenkov, B.S. Getmanov, and V.E.Kovtun. Phys.Lett. **A128** 18 (1988); Yad.Fiz. **48** 886 (1988) [translation in Sov.J.Nucl.Phys. **48** 565 (1988)]

[16] J. Harnad, Y. Saint-Aubin, S. Shnider. Commun.Math.Phys. **93** 33 (1984)

[17] D. David et.al. Lett.Math.Phys. **8** 27 (1984)

[18] I.Ya. Aref'eva. Teor.Mat.Fiz. **26** 306 (1976)

[19] A.V. Mikhailov. Pis'ma ZhETF **23** 356 (1976) [translation in JETP Lett. **23** 321 (1976)]

[20] E.A. Kuznetsov and A.V. Mikhailov. Teor.Mat.Fiz. **30** 303 (1977)

[21] D.J. Kaup and A.C. Newell. Lett.Nuovo Cim. **20** 325 (1977)

[22] D. David J.Math.Phys. **25** 3424 (1984)

[23] F. Lund, and T. Regge. Phys.Rev. **D14** 1524 (1976)

[24] F. Lund. Phys.Rev.Lett. **38** 1175 (1977); Ann.Phys. **115** 251 (1978)

[25] B.S. Getmanov. Pis'ma ZhETF **25** 132 (1977) [translation in JETP Lett. **25** 119 (1977)]

[26] V.E. Zakharov and A.V. Mikhailov. Pis'ma ZhETF. **27** 47 (1978) [translation in JETP Lett. **27** 42 (1978)]

Λ-MODEL: A UNIVERSAL CLASS OF KINK-BEARING SYSTEMS

J. A. Hołyst[#*] and H. Benner[*]

[#] Institute of Physics, Warsaw Technical University, Koszykowa 75, 00-662 Warsaw, Poland

[*] Institut für Festkörperphysik, Technische Hochschule Darmstadt, 6100 Darmstadt, Hochschulstraße 6, Federal Republic of Germany

ABSTRACT

We present a continuous family of kink-bearing potentials $V^{(\Lambda)}(\varphi)$ for a one-dimensional nonlinear Klein-Gordon equation. The form of these potentials has been reconstructed from the Pöschl-Teller potential which has been assumed to describe the interaction between kinks and low amplitude harmonic solutions (phonons). For special values of the characteristic parameter Λ the presented model can be reduced to the well known sine-Gordon, ϕ^4, Eshelby and double-quadratic models.

1. INTRODUCTION

There exists a certain class of quasi-one-dimensional solids[1)] (magnets, conductors etc) which above their three-dimensional ordering temperature[2-3)] behave like systems of isolated chains of spins or ions. Special solutions of the nonlinear equation of motion of such chains can have the form of large-amplitude solitary waves, kinks or (if the system is completely integrable) solitons[4)]. From a thermodynamical point of view[4-6)] so-called topological solutions i.e. kinks (or kink-solitons) can be treated as a new type of elementary excitations forming a collective gas, together with low-amplitude harmonic excitations such as magnons or phonons (for

simplicity, we use the latter name for all kinds of harmonic excitations). The nonlinear interference[7] between different quasi-particles in this gas gives rise to mutual phase shifts and kink-phonon bound states. Usually, one assumes that the Hamiltonian H_{1D} of a single chain is known and the potential U_{K-Ph} describing kink-phonon scattering[6] can be calculated from this Hamiltonian. One can ask, however, the following question: Suppose that we know the interaction potential U_{K-Ph} between collective excitations (i.e. kinks and phonons), can we find the corresponding Hamiltonian of a single chain H_{1D} ? We shall show that the answer is positive and that we obtain an interesting family of Hamiltonians H_{1D} when starting from a special kink-phonon scattering potential $U^{(\Lambda)}(x)$ which is of the Pöschl-Teller type.

2. INVERSE POTENTIAL METHOD

We consider the classical Hamiltonian of a one-dimensional nonlinear Klein-Gordon model

$$H_{1D} = A \int dx \left\{ \frac{1}{2} [\partial\varphi(x,t)/\partial t)]^2 + \frac{1}{2} c_0^2 [\partial\varphi(x,t)/\partial x]^2 + \omega_0^2 V(\varphi) \right\}, \quad (1)$$

where A, c_0 and ω_0 are characteristic constants. We assume that the local potential $V(\varphi)$ has at least two degenerate minima φ_1 and φ_2 ($> \varphi_1$), and we put $V(\varphi_1) = V(\varphi_2) = 0$, $d^2V/d\varphi^2|_{\varphi_1,\varphi_2} = 1$. A static kink solution $\varphi_k(x)$ connecting the ground states φ_1 and φ_2 can be obtained[6] from the first integral of the equation of motion for the Hamiltonian (1), i.e.

$$d\varphi_k(x)/dx = \pm\, d^{-1}[2\, V(\varphi_k)]^{1/2} . \quad (2)$$

Here, $d = c_0/\omega_0$ is the natural length of the model. Considering small oscillations around the kink $\varphi_k(x)$, i.e. assuming

$$\varphi(x,t) = \varphi_k(x) + f(x)\, \exp(-i\omega t) \quad (3a)$$

where $|f(x)| \ll 1$, one gets[6] a Schrödinger-like equation

$$-d^2 f(x)/dx^2 + U_{K-Ph}(x)\, f(x)/d^2 = (\omega/c_0)^2\, f(x) \quad (3b)$$

with a "potential" $U_{K-Ph}(x) = [d^2V(\varphi)/d\varphi^2]_{\varphi=\varphi_k(x)}$ describing the effects of kink-phonon interactions. Now, we shall solve an, in a certain sense, "inverse problem" i.e. we shall find the potential $V(\varphi)$ assuming that the potential $U_{K-Ph}(x)$ is known. We stress here that our approach should not be confused with the well known "inverse scattering method" which is used to solve the equations of motion of completely integrable systems[8].

Let us assume that the potential $U_{K-Ph}(x)$ is given. One cannot choose it completely arbitrarily because from the properties of the model (1) it follows that[6]

(a) the potential $U_{K-Ph}(x)$ has to be bounded from below;

(b) $\lim_{x\to\pm\infty} U_{K-Ph}(x) = 1$;

(c) the ground state function $f_1(x)$ should correspond to the Goldstone mode of the system (1), i.e. $\omega_1^2 = 0$ is the lowest eigenvalue of Eq.(3b).

We perform our reconstruction of the potential $V(\varphi)$ using the fact[6] that the Goldstone mode $f_1(x)$ is proportional to the spatial derivative of a kink solution $\varphi_k(x)$. Assuming that this kink is centered at x = 0 we can write

$$\varphi_k(x) = \varphi_0 + \frac{\Delta\varphi}{\Delta F} F(x) , \quad (4)$$

where $F(x) = \int_0^x f_1(x)\, dx$, $\Delta F = F(+\infty) - F(-\infty)$, $\Delta\varphi = \varphi_2 - \varphi_1$ and $\varphi_0 = \varphi_2 - F(+\infty)\, \frac{\Delta\varphi}{\Delta F}$ [we use the solution with the plus sign in Eq.(2)]. Since the ground state function $f_1(x)$ is nodeless, the function F(x) is strictly monotonous. Combining Eqs. (2) and (4) we obtain the potential $V(\varphi)$ as the following superposition[9]:

$$V(\varphi) = \frac{1}{2}\left\{d\,\frac{\Delta\varphi}{\Delta F}\,f_1 \circ F^{inv}\left[(\varphi - \varphi_0)\,\frac{\Delta F}{\Delta\varphi}\right]\right\}^2 . \qquad (5)$$

Here, $F^{inv}(y)$ represents the inverse function of $F(x)$, and the monotony of $F(x)$ implies that $F^{inv}(y)$ is well defined. The equation of motion for model (1) now reads

$$\partial^2\varphi(x,t)/\partial t^2 - c_0^2\,\partial^2\varphi(x,t)/\partial x^2 + c_0^2\frac{\Delta\varphi}{\Delta F}\,f_1' \circ F^{inv}\left[(\varphi - \varphi_0)\frac{\Delta F}{\Delta\varphi}\right] = 0 \qquad (6)$$

where $f_1'(x) = df_1(x)/dx$. The creation energy of the kink $\varphi_k(x)$ can be written as

$$E_k = A\,c_0^2 \int_{-\infty}^{+\infty} [d\varphi_k(x)/dx]^2 dx = A\left[c_0\,\frac{\Delta\varphi}{\Delta F}\right]^2 \int_{-\infty}^{+\infty} [f_1(x)]^2 dx . \qquad (7)$$

3. Λ - MODEL

Now we shall illustrate the whole procedure by using a Pöschl-Teller potential $U^{(\Lambda)}(x)$, i.e. we assume

$$U_{K\text{-}Ph}(x) = U^{(\Lambda)}(x) = 1 - \frac{\Lambda + 1}{\Lambda\,\cosh^2[x/(\Lambda d)]} . \qquad (8)$$

It is easy to check that for the potential (8) the conditions (a - c) below Eq.(3b) are fulfilled. Since the unnormalised ground state for this potential[10] is $f_1^{(\Lambda)}(x) = \cosh^{-\Lambda}[x/(\Lambda d)]$, a kink solution of this model (we will call it "the Λ-model") can be derived from Eq.(4)

$$\varphi_k^{(\Lambda)}(x) = \varphi_1 + \Delta\varphi\; I\left[\frac{\Lambda}{2}, \frac{\Lambda}{2}, \frac{1+\tanh[x/(\Lambda d)]}{2}\right] , \qquad (9)$$

where $I(a,b,z) \equiv I_z(a,b)$ denotes the incomplete β function. If the parameter Λ is a natural number, Eq.(9) reduces to a finite series of powers of hypergeometric functions[9]. Inserting Eq.(9) into (5) we now derive the general form of $V^{(\Lambda)}(\varphi)$

$$V^{(\Lambda)}(\varphi) = 2\left[\frac{\Delta\varphi}{\Lambda\; B(\Lambda/2, \Lambda/2)}\right]^2 \left[I^{inv}\left(\frac{\Lambda}{2}, \frac{\Lambda}{2}, \frac{\varphi - \varphi_1}{\Delta\varphi}\right) I^{inv}\left(\frac{\Lambda}{2}, \frac{\Lambda}{2}, \frac{\varphi_2 - \varphi}{\Delta\varphi}\right)\right]^{\Lambda} \qquad (10)$$

Here, $B(a,b)$ is the beta function and $I^{inv}(a,b,y)$ stands for the inverse function of the incomplete beta function $y = I_z(a,b)$ with respect to the argument z. We stress here that $V^{(\Lambda)}(\varphi)$ is only defined for $\varphi \in [\varphi_1, \varphi_2]$, so one is free to extend the potential outside this regime provided that the conditions (a - c) below Eq.(3b) are not violated. It is important to note that Eq.(10) defines a potential $V^{(\Lambda)}(\varphi)$ for any $\Lambda > 0$. Thus, by fixed values of φ_1 and φ_2 we have a well defined family of potentials $V^{(\Lambda)}(\varphi)$, and on variation of Λ we can continuously pass between all members of this family[9]. Some special cases of Eq.(10) will be discussed below. The creation energy of the kink (9) is given by

$$E_k^{(\Lambda)} = (\Delta\varphi)^2\, A\, c_0\, \omega_0\, 2^{4\Lambda-2}\, \pi^{-2}\, \frac{\left[\Gamma\left(\frac{\Lambda + 1}{2}\right)\right]^4}{\Gamma(2\Lambda + 1)} . \qquad (11)$$

A direct physical interpretation of the parameter Λ is related to the number of bound state solutions. For $\Lambda \in (\ell\text{-}1, \ell]$, $\ell \in \mathbb{N}$, $\ell \neq 0$, Eq.(3b) has exactly ℓ bound state solutions, i.e. solutions with eigenfrequencies $\omega < \omega_0$. This means for our model that there are ℓ different phonon modes which are trapped by the kink (9). The frequencies of these modes are given by

$$\omega_{b,n}^2 = \omega_0^2\left[1 - \left(1 - \frac{n}{\Lambda}\right)^2\right] , \qquad (12)$$

where n= 0, 1, 2,...(ℓ-1). If we write $\Lambda = \ell - 1 + \varepsilon$ where $0 < \varepsilon \leq 1$, then the frequency of the highest bound state is $\omega_{b,max}^2 = \omega_0^2\,(1-\varepsilon^2/\Lambda^2)$. Thus, the square value of ε (the fractal part of Λ) is proportional to the distance between the square frequencies of the highest bound state ($\omega_{b,max}^2$) and of long wave phonons (ω_0^2). When Λ is continuously diminished in such a way that $\varepsilon \to 0_+$, then $\omega_{b,max}$ tends to ω_0 (except for the case $\ell = 1$), which means that this highest

trapped mode migrates into a band of non-localized phonon modes. Using the results of Ref. 10 and different properties of the hypergeometric function[11] we arrive at the following expression for the (unnormalized) bound-state modes

$$f_{b,n}^{(\Lambda)}(x) = (-z)^{(1\pm1)/4}(1-z)^{(\Lambda+1)/2}\frac{d^m}{dz^m}\left[z^{m\mp1/2}(1-z)^{m-\Lambda-1/2}\right], \quad (13)$$

where $z = -\sinh^2[x/(\Lambda d)]$. Upper signs refer to even-parity bound states ($n = 2m$) and lower signs to odd-parity bound states ($n = 2m+1$), $m = 0, 1, 2\ldots$ and $n < \Lambda$.

Besides the discrete bound states defined by Eqs. (12) and (13) there are also continuum states (i.e. extended phonon states) which are affected by the presence of a kink (9) due to a phase shift[6] $\Delta^{(\Lambda)}(k)$. Since, in general, the kinks in our model are not transparent for phonons one has to consider the phase shift for even and odd parity states separately[6, 10]. This is a standard problem of quantum dynamics and one can find the detailed results in Refs. 10 and 12. After some algebra we get the compact form of the phase shift for any natural value of Λ as

$$\Delta^{(\Lambda)}(k) = \pi\Lambda\frac{k}{|k|} - 2\sum_{l=1}^{\Lambda}\operatorname{arctg}(k\Lambda d/l). \quad (14)$$

Another parameter characterizing the kink-phonon collision is the reflection coefficient $|R|^2$ which can be written as

$$|R^{(\Lambda)}(k)|^2 = \left[1 + \frac{\sinh^2(\pi k\Lambda d)}{\sin^2(\pi\Lambda)}\right]^{-1}. \quad (15)$$

From (15) we see that in the present model the kinks (9) are transparent to phonons only if Λ is a natural number, because only in this case we have $|R^{(\Lambda=n)}(k)|^2 = 0$. (The case of long-wave phonons has to be excluded since the double limit $\lim_{\substack{\Lambda\to n\\ k\to 0}} |R^{(\Lambda)}(k)|^2$ does not exist.) For non-integer values of Λ the reflection coefficient $|R^{(\Lambda)}(k)|^2$ shows an exponential decay with respect to Λ which is modulated by a periodic factor (see Fig. 1).

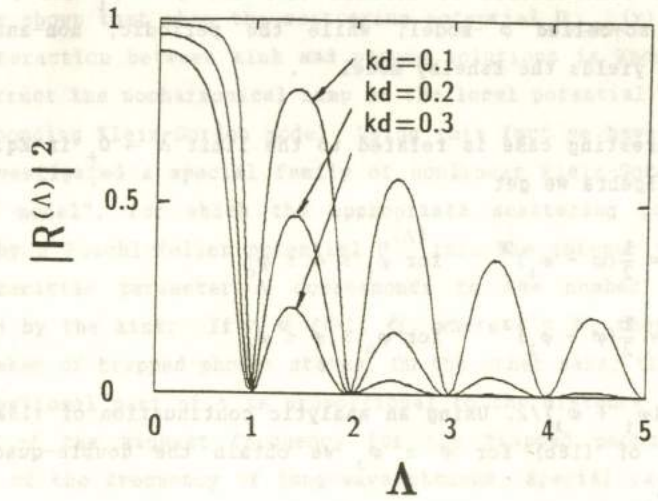


Fig. 1 Reflection coefficient $|R^{(\Lambda)}(k)|^2$ as a function of Λ for different wave vectors of scattered phonons.

Now, we discuss special cases[9] of potential (10) where the incomplete beta function can be expressed in terms of elementary functions.

For $\Lambda = 1$ we have $I\left(\frac{1}{2},\frac{1}{2},z\right) = \frac{2}{\pi}\arcsin\sqrt{z}$. Thus, from Eq. (10) we get

$$V^{(\Lambda=1)}(\varphi) = \frac{(\Delta\varphi)^2}{4\pi^2}\left\{1-\cos\left[\frac{2\pi(\varphi-\varphi_1)}{\Delta\varphi}\right]\right\}. \quad (16)$$

Combining this expression with an analytic continuation of the r.h.s. for $\varphi < \varphi_1$ and $\varphi > \varphi_2$, we obain the potential of the sine-Gordon model.

For $\Lambda = 2$ we have $I(1,1,z) = z$, thus

$$V^{(\Lambda=2)}(\varphi) = \frac{(\varphi - \varphi_1)^2 (\varphi - \varphi_2)^2}{2(\Delta\varphi)^2} . \qquad (17)$$

The analytic continuation of formula (17) for $\varphi < \varphi_1$ and $\varphi > \varphi_2$ yields the so-called ϕ^4-model, while the periodic, non-analytic continuation yields the Eshelby model[13].

Another interesting case is related to the limit $\Lambda \to 0_+$ in Eq.(10). After some algebra we get

$$V^{(\Lambda\to 0+)}(\varphi) = \frac{1}{2}(\varphi - \varphi_1)^2 \qquad \text{for } \varphi_1 \le \varphi \le \varphi_0 \qquad (18a)$$

$$V^{(\Lambda\to 0+)}(\varphi) = \frac{1}{2}(\varphi - \varphi_2)^2 \qquad \text{for } \varphi_0 \le \varphi \le \varphi_2 \qquad (18b)$$

where $\varphi_0 = (\varphi_1 + \varphi_2)/2$. Using an analytic continuation of (18a) for $\varphi \le \varphi_1$ and of (18b) for $\varphi \ge \varphi_2$ we obtain the double-quadratic model [6, 13].

Now, let us consider the case $\Lambda = 4$ (the case $\Lambda = 3$ is less interesting because it is not possible to express $V^{(\Lambda=3)}(\varphi)$ in terms of elementary functions[9]). After some algebra[9] the potential (10) can be simplified to the following form (we assumed $\varphi_1 = -1$ and $\varphi_2 = 1$)

$$V^{(\Lambda=4)}(\varphi) = \frac{9}{128}\left[1 - 4\cos^2\left(\frac{\pi + \arccos\varphi}{3}\right)\right]^4 . \qquad (19)$$

The kink solution corresponding to the potential (19) reads

$$\varphi_k^{(\Lambda=4)}(x) = \varphi_0 + \frac{\Delta\varphi}{2}\left[\tanh(x/D_\Lambda) + \frac{1}{2}\tanh(x/D_\Lambda)\,\mathrm{sech}^2(x/D_\Lambda)\right] . \qquad (20)$$

where $D_\Lambda = 4d$. The creation energy of the kink (20) is given by $E_k^{(\Lambda=4)} = (9/70)\, A\, c_0\omega_0(\Delta\varphi)^2$. There occur four phonon bound states for the kink (20) with the eigenfrequencies $\omega_{b,n}/\omega_0 = 0,\ \sqrt{7}/4,\ \sqrt{12}/4,\ \sqrt{15}/4$, and the corresponding eigenfunctions can also be written explicitly[9].

4. CONCLUDING REMARKS

In this lecture we have considered classical, one-dimensional, kink-bearing systems described by nonlinear Klein-Gordon equations. We have shown that when the scattering potential $U_{K\text{-}Ph}(x)$ describing the interaction between kink and phonon solutions is known, one can reconstruct the nonharmonical hump of the local potential $V(\varphi)$ of the corresponding Klein-Gordon model. Using this fact we have introduced and investigated a special family of nonlinear Klein-Gordon models, "the Λ-model", for which the appropriate scattering potential is given by a Pöschl-Teller potential $U^{(\Lambda)}(x)$. The integer part of the characteristic parameter Λ corresponds to the number of phonons trapped by the kink: If $\Lambda \in (\ell\text{-}1, \ell]$, where $\ell \in \mathbb{N}$, then ℓ is just the number of trapped phonon states. On the other hand, the square of the fractional part of Λ is proportional to the distance between the square of the highest frequency for the trapped phonons and the square of the frequency of long-wave phonons. Special cases of this family correspond to the double-quadratic, sine-Gordon, ϕ^4 and Eshelby models. The fact that our model is well defined for any positive value of Λ also enables us to treat models that are intermediate - between the double-quadratic and sine-Gordon model or between the sine-Gordon and ϕ^4 model.

ACKNOWLEDGMENTS

We are grateful to Professor Andrzej Sukiennicki for his continuous interest in this investigation. This work was partly financed by special funds of the Deutsche Forschungsgemeinschaft as a Project of SFB 185 and it was also supported by the Polish Ministry of National Education under the Project P/04/160/90-2.

REFERENCES

1) Mikeska, H.J., J. Magn. Magn. Mater.13, 35 (1979); Kagoshima, S., Nagasawa, H., and Sambongi, T., "One-dimensional Conductors", Springer Series in Solid-State Sciences, Vol. 72, (Springer, Berlin, 1988).

2) Hołyst, J.A. and Sukiennicki, A., J. Phys. C18, 2411 (1985); Phys. Rev.B38, 6975 (1988).

3) Hołyst, J.A., Z. Phys. B74, 341 (1989); Benner, H., Hołyst, J.A., and Löw, J., Europhys. Lett. 11, 383 (1991).

4) Bishop, A.R., Krumhansl, J.A., and Trullinger, S.E., Physica D1, (1980); "Nonlinearity in Condensed Matter", edited by Bishop, A.R., Campbell, D.K., Kumar, P. and Trullinger, S.E., Springer Series in Solid-State Sciences, Vol. 69 (Springer-Verlag, Berlin 1987).

5) Krumhansl, J.A. and Schrieffer, J.R., Phys. Rev. B11, 3535 (1975).

6) Currie, J.F., Krumhansl, J.A., Bishop, A.R., and Trullinger, S.E., Phys. Rev. B22, 477 (1980).

7) Benner, H., Seitz, H., Wiese, J., and Boucher, J.P., J. Magn. Magn. Mater. 45, 354 (1984); Benner, H., Wiese, J., Geick R., and Sauer H., Europhys. Lett.3, 1135 (1987).

8) Gardner, C.S., Greene, J.M., Kruskal M.D., and Miura R.M., Phys. Rev. Lett. 19, 1095 (1967); Eilenberger, G., "Solitons: Mathematical Methods for Physicists", Springer Series in Solid-State Sciences, Vol. 19, (Springer, Berlin, 1981).

9) Hołyst J.A. and Benner H., Phys. Rev. B, in press (1991).

10) Flügge, S., "Practical Quantum Mechanics", (Springer, Berlin, 1971).

11) "Handbook of Mathematical Functions", edited by Abramowitz, M. and Stegun, I.A., (Dover Publ., New York, 1970).

12) Landau, L.D. and Lifshitz, E.M., "Quantum Mechanics", (Pergamon, New York, 1977).

13) Sasaki, K., Progress of Theoret. Phys. 67, 464 (1982).

DYNAMICAL SYMMETRIES AND FINITE DIMENSIONAL NONLINEAR SYSTEMS: INTEGRABILITY AND SEPARABILITY

M. LAKSHMANAN and M. SENTHIL VELAN

Centre for Nonlinear Dynamics, Department of Physics, Bharathidasan University, Tiruchirapalli 620 024, India

ABSTRACT

We indicate a general method to locate integrable as well as separable nonlinear dynamical systems with finite degrees of freedom, with two degrees of freedom as an example. The method is to directly integrate the dynamical symmetries using geometric theory of first order linear partial differential equations and obtain the local invariants which can be identified with the required involutive integrals of motion.

1. INTRODUCTION

Some of the standard finite degrees of freedom systems in classical dynamics are the Kepler system, isotropic oscillator, etc. In recent times with the advent of soliton theory and the Painlevé singularity structure analysis, many new integrable systems have been discovered.[1,2] Apart from their intrinsic mathematical interest, they are also of great physical importance. For example, they divide the parameter regimes into integrable and non-integrable regions leading to regular and chaotic motions. For these reasons it is of importance to develop procedures to isolate such systems systematically.

Recently it was shown[3,4] that invariance analysis of nonlinear dynamical systems under extended Lie symmetries

can lead to a systematic identification of integrable systems. Considering for example, two degrees of freedom systems with the Lagrangian and Hamiltonian respectively of the form

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - V(x,y), \quad H = \frac{1}{2}(p_x^2 + p_y^2) + V(x,y), \tag{1.1a}$$

we have the equations of motion

$$\ddot{x} = \alpha_1(x,y) \equiv -\frac{\partial V}{\partial x}, \quad \ddot{y} = \alpha_2(x,y) \equiv -\frac{\partial V}{\partial y}. \tag{1.1b}$$

Then we can consider the one parameter continuous group of symmetry transformations $(t,x,y,\dot{x},\dot{y}) \to (T,X,Y,\dot{X},\dot{Y})$ which leave (1.1b) invariant, with their infinitesimals taking the form

$$T = t + \varepsilon\xi(t,x,y,\dot{x},\dot{y}), \quad X = x + \varepsilon\eta_1(t,x,y,\dot{x},\dot{y}),$$

$$Y = y + \varepsilon\eta_2(t,x,y,\dot{x},\dot{y}), \qquad \varepsilon << 1 \tag{1.2}$$

so that

$$\dot{X} = \dot{x} + \varepsilon(\dot{\eta}_1 - \dot{x}\dot{\xi}), \quad \dot{Y} = \dot{y} + \varepsilon(\dot{\eta}_2 - \dot{y}\dot{\xi}). \tag{1.3}$$

Thus the infinitesimal generator of the symmetry group becomes

$$E = \xi\frac{\partial}{\partial t} + \eta_1\frac{\partial}{\partial x} + \eta_2\frac{\partial}{\partial y} + (\dot{\eta}_1 - \dot{x}\dot{\xi})\frac{\partial}{\partial \dot{x}}$$

$$(\dot{\eta}_2 - \dot{y}\dot{\xi})\frac{\partial}{\partial \dot{y}}. \tag{1.4}$$

The associated invariance condition for the equation of motion (1.1b) under this extended vector field may be written as[5]

$$\ddot{\eta}_1 - \dot{x}\ddot{\xi} - 2\dot{\xi}\alpha_1 = E\{\alpha_1\},$$

$$\ddot{\eta}_2 - \dot{y}\ddot{\xi} - 2\dot{\xi}\alpha_2 = E\{\alpha_2\}. \tag{1.5}$$

Methods for solving the invariance condition (1.5) and locating the integrable choices have been discussed in detail elsewhere.[3,6]

After obtaining the infinitesimal symmetries, to prove the integrability one has to find the required involutive integrals of motion. If a given set of symmetries is of Noether type, then one can find the integrals of motion through Noether's theorem. However one can find the required involutive integrals of motion also by looking for a local invariant satisfying the condition[7,8]

$$E\{U\} = \xi\frac{\partial U}{\partial t} + \eta_1\frac{\partial U}{\partial x} + \eta_2\frac{\partial U}{\partial y} + (\dot{\eta}_1 - \dot{x}\dot{\xi})\frac{\partial U}{\partial \dot{x}} +$$

$$(\dot{\eta}_2 - \dot{y}\dot{\xi})\frac{\partial U}{\partial \dot{y}} = 0. \tag{1.6}$$

Equation (1.6) is a first order linear partial differential equation in five independent variables, solving which we can obtain the required involutive integrals of motion.

In this contribution, we will briefly explain how the required invariants can be obtained by analysing directly (1.6) and also how separable coordinates can be obtained. Fuller details and extensions to higher dimensions will be published elsewhere.[9]

2. GEOMETRIC THEORY FOR FIRST ORDER LINEAR PARTIAL DIFFERENTIAL EQUATION[10]

To begin with let us consider a first order partial differential equation involving two independent variables,

$$P(x,y,z)\frac{\partial z}{\partial x} + Q(x,y,z)\frac{\partial z}{\partial y} = R(x,y,z). \qquad (2.1)$$

The solution of eq. (2.1) can be represented geometrically as an integral surface in the (x,y,z) space:

$$I(x,y,z) = z(x,y) - z = 0. \qquad (2.2)$$

Now the tangent to the integral surface (2.2) at any point in the (x,y,z) space is obviously given by

$$\frac{\partial I}{\partial x}\,dx + \frac{\partial I}{\partial y}\,dy + \frac{\partial I}{\partial z}\,dz = 0$$

or

$$p\,dx + q\,dy = dz, \qquad (2.3)$$

where $p = \frac{\partial z}{\partial x}$ and $q = \frac{\partial z}{\partial y}$. Equation (2.3) represents an equation of the tangent plane at any point. We can easily check the vector **N** = (p,q,-1) is normal to the integral surface I. Similarly we can also prove that the normal **N** is perpendicular to the vector **M** = (P,Q,R), since

$$\mathbf{N}.\mathbf{M} = (p,q,-1).(P,Q,R) = Pp + Qq - R = 0 \qquad (2.4)$$

so that **M** lies in the tangent plane.

Now we consider the characteristic curve C (parametrised by s) associated with the tangent vector **M**. In general the tangent vector at a point **r** on C is given by

$$\frac{d\mathbf{r}}{ds} = \frac{dx}{ds}\,\mathbf{i} + \frac{dy}{ds}\,\mathbf{j} + \frac{dz}{ds}\,\mathbf{k}. \qquad (2.5)$$

So the vector **M** is proportional to $(\frac{d\mathbf{r}}{ds})$ from which follows the characteristic equation

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dz}{R}. \qquad (2.6)$$

Eq. (2.6) is only a pair of independent, first order ordinary differential equations and that the general solution can be expressed in terms of two independent functions $G(x,y,z) = C_1$ and $g(x,y,z) = C_2$ each containing an arbitrary or free parameter C_i, i = 1,2. These solutions constitute a two parameter family of space curves that generate integral surfaces satisfying the first order partial differential eq. (2.1).

3. METHOD OF CONSTRUCTING INVARIANTS FOR THE SYMMETRIES[11)]

Using the above theory, we can easily show that solving the homogeneous linear partial differential equation (1.6) is equivalent to solving the following system of characteristic equation,

$$\frac{dt}{\xi} = \frac{dx}{\eta_1} = \frac{dy}{\eta_2} = \frac{d\dot{x}}{(\dot{\eta}_1 - \dot{x}\dot{\xi})} = \frac{d\dot{y}}{(\dot{\eta}_2 - \dot{y}\dot{\xi})}. \qquad (3.1)$$

System (3.1) obviously admits four functionally independent invariants $U_i(t,x,y,\dot{x},\dot{y})$, i = 1,2,3,4, which often in practice turn out to be quite complicated to find. Using the previous results, we can treat U_i's as integral (hyper) surfaces and the tangentials through the points $(t,x,y,\dot{x},\dot{y})$ satisfy the condition

$$dU_i = \frac{\partial U_i}{\partial t}\,dt + \frac{\partial U_i}{\partial x}\,dx + \frac{\partial U_i}{\partial y}\,dy + \frac{\partial U_i}{\partial \dot{x}}\,d\dot{x} + \frac{\partial U_i}{\partial \dot{y}}\,d\dot{y} = 0. \qquad (3.2)$$

As we have already seen the tangential directions to the integral curves of the characteristic equation (3.1) through the point $(t,x,y,\dot{x},\dot{y})$ are also tangential directions

to these surfaces. Hence

$$\xi \frac{\partial U_i}{\partial t} + \eta_1 \frac{\partial U_i}{\partial x} + \eta_2 \frac{\partial U_i}{\partial y} + (\dot{\eta}_1 - \dot{x}\dot{\xi}) \frac{\partial U_i}{\partial \dot{x}} +$$

$$(\dot{\eta}_2 - \dot{y}\dot{\xi}) \frac{\partial U_i}{\partial \dot{y}} = 0. \qquad (3.3)$$

To find U_i $(i = 1,2,3,4)$ we try to find functions P_i, Q_i, R_i, S_i and T_i such that

$$\xi P_i + \eta_1 Q_i + \eta_2 R_i + (\dot{\eta}_1 - \dot{x}\dot{\xi})S_i + (\dot{\eta}_2 - \dot{y}\dot{\xi})T_i = 0,$$

$$i = 1,2,3,4. \qquad (3.4)$$

with the property

$$P_i = \frac{\partial U_i}{\partial t}, \quad Q_i = \frac{\partial U_i}{\partial x}, \quad R_i = \frac{\partial U_i}{\partial y},$$

$$S_i = \frac{\partial U_i}{\partial \dot{x}}, \quad T_i = \frac{\partial U_i}{\partial \dot{y}}, \qquad (3.5)$$

so that $P_i\, dt + Q_i\, dx + R_i\, dy + S_i\, d\dot{x} + T_i\, d\dot{y}$ is an exact differential dU_i.

For the system (1.1), we can choose without loss of generality $\xi = 0$. Then the two invariants U_1 and U_2 for the characteristic eq. (3.1) satisfying (3.4) can be given immediately with the following choices:

(1) $P_1 = 0$, $Q_1 = -\alpha_1 = -\frac{\partial L}{\partial x}$, $R_i = -\alpha_2 = -\frac{\partial L}{\partial y}$,

$$S_1 = \dot{x} = \frac{\partial L}{\partial \dot{x}}, \quad T_1 = \dot{y} = \frac{\partial L}{\partial \dot{y}} \qquad (3.6)$$

so that U_1 is the Hamiltonian for all the integrable cases.

(2) $P_2 = 0$, $Q_2 = -\dot{\eta}_1$, $R_2 = -\dot{\eta}_2$, $S_2 = \eta_1$, $T_2 = \eta_2$ (3.7)

so that U_2 is the second integral of motion in all the integrable cases.

To prove the above, we may proceed as follows. Since the Hamiltonian in (1.1) is conserved, we have $E\{H\} = 0$. Therefore the choice (3.6) follows. Choice (3.7) obviously satisfies equation (3.4) and it leads to the necessary second integral of motion in all the cases we have studied.[9)]

4. EXAMPLE

Now we apply the above method to construct the integrals of motion for the following Lagrangian[12)]

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - \left[A(x^2 + 4y^2) + B(x^4 + 12x^2y^2 + 16y^4) + \frac{C}{x^2} + D\left(\frac{1}{x^4} + \frac{4y^2}{x^6}\right)\right]. \qquad (4.1)$$

The corresponding equation of motion is

$$\ddot{x} = -\left[2Ax + 4Bx^3 + 24Bxy^2 - \frac{2C}{x^3} - \frac{4D}{x^5} - \frac{24Dy^2}{x^7}\right] = \alpha_1,$$

$$\ddot{y} = -\left[8Ay + 24Bx^2y + 64By^3 + \frac{8Dy}{x^6}\right] = \alpha_2. \qquad (4.2)$$

Solving the invariance condition (1.5) with (4.2) we get the following infinitesimal symmetries[6)]

$$\xi = 0, \quad \eta_1 = -2y\dot{x} + x\dot{y}, \quad \eta_2 = x\dot{x}. \qquad (4.3)$$

Now to find the required invariants we have to solve the characteristic equation associated with the infinitesimal symmetries (4.3). This can be done by choosing

the functions P_i, Q_i, R_i, S_i and T_i as follows. We choose

(1) $P_1 = 0$, $Q_1 = -\alpha_1$, $R_1 = -\alpha_2$, $S_1 = \dot{x}$, $T_1 = \dot{y}$, (4.4)

where α_1 and α_2 are as given in (4.2). Then we get the first invariant

$$U_1 = H = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + A(x^2 + 4y^2) + B(x^4 + 12x^2y^2 + 16y^4) + \frac{C}{x^2} + D(\frac{1}{x^4} + \frac{4y^2}{x^6}) , \quad (4.5)$$

which is the Hamiltonian.

(2) $P_2 = 0$, $Q_2 = -\dot{\eta}_1$, $R_2 = -\dot{\eta}_2$, $S_2 = \eta_1$, $T_2 = \eta_2$

so that we get a second invariant,

$$U_2 = (x\dot{y} - y\dot{x})\dot{x} + 2Ax^2y + 4Bx^4y + 8Bx^2y^3 - \frac{4Dy}{x^4} - \frac{2Cy}{x^2} - \frac{8Dy^3}{x^6} , \quad (4.6)$$

which is the required second involutive integral of motion for the Lagrangian (4.1).

5. SEPARABLE COORDINATES

In this section we will show that from the generalized symmetries we can also find the suitable coordinate system in which either the equation of motion or the Hamilton-Jacobi equation becomes separable, whenever the generalized symmetries are linear in the velocities. Since the separability is associated with the coordinate transformations, we will consider the corresponding part of the characteristic equation discussed earlier, namely

$$\frac{dx}{\eta_1} = \frac{dy}{\eta_2} . \quad (5.1)$$

For example in our case

$$\frac{dx}{-2y\dot{x} + x\dot{y}} = \frac{dy}{x\dot{x}} . \quad (5.2)$$

Rewriting (5.2), we get

$$x\left[(\frac{dy}{dx})^2 - 1\right] - 2y\frac{dy}{dx} = 0 , \quad (5.3)$$

which on integration gives

$$2y = cx^2 - \frac{1}{c} , \quad (5.4)$$

where c is an arbitrary constant. Rewriting (5.4) we get

$$\frac{x^2}{\eta^4} - \frac{2y}{\eta^2} = 1 , \quad (5.5)$$

where $c = 1/\eta^2$, which can be parametrized as

$$x = \xi\eta , \quad y = \frac{1}{2}(\xi^2 - \eta^2) . \quad (5.6)$$

Then the Hamilton-Jacobi equation takes the separable form

$$S_\xi^2 + S_\eta^2 + 2A(\xi^6 + \eta^6) + 2B(\xi^{10} + \eta^{10}) + 2C(\frac{1}{\xi^2} + \frac{1}{\eta^2}) + 2D(\frac{1}{\xi^6} + \frac{1}{\eta^6}) - 2E(\xi^2 + \eta^2) = 0, \text{ E = energy.} \quad (5.7)$$

6. CONCLUSIONS

In this paper we have indicated a general method to find the integrals of motion and separable coordinates associated with nonlinear dynamical systems involving two degrees of freedom. This method can also be extended to

higher dimensional systems as well. The details are given in refs.[9)]

Acknowledgement. This work has been supported by the Council of Scientific and Industrial Research, India through a project grant. M.L. wishes to thank Dr Z. Popowicz and Prof. J. Zagrodzinski for their excellent hospitality in Poland.

REFERENCES

1. Perelomov, A. M., "Integrable Systems of Classical Mechanics and Lie Algebras" (Birkhäuser Verlag, Basel, 1990).
2. Ramani, A., Grammaticos, B., and Bountis, T., Phy. Rep. **180**, 159 (1989).
3 Sahadevan, R. and Lakshmanan, M., J. Phys. A: Math. Gen. **19**, L947 (1986).
4. Lakshmanan, M. and Sahadevan, R., J. Math. Phy. **32**, 75 (1991).
5. Lutzky, M., J. Phys. A: Math. Gen. **12**, 973 (1979).
6. Senthil Velan, M. and Lakshmanan, M., in "Symmetries and Singularity Structures: Integrability and Chaos in Nonlinear Dynamical Systems," Research Reports in Physics, (eds.) Lakshmanan, M. and Daniel, M. (Springer-Verlag, Heidelberg, 1990).
7. Olver, P. J., "Applications of Lie Groups to Differential Equations" (Springer-Verlag, New York, 1986).
8. Bluman, G. W. and Kumei, S., "Symmetries and Differential Equations" (Springer-Verlag, New York, 1989).
9. Lakshmanan, M. and Senthil Velan, M. (Preprints 1991) to be published.
10. Street, R. L., "Analysis and Solutions of Partial Differential Equations" (Brooks/Cole Publishing Company, Monterey, 1973).
11. Sneddon, I., "Elements of Partial Differential Equations" (McGraw-Hill, Singapore, 1957).
12. Grammaticos, B., Dorizzi, B., Ramani, A. and Hietarinta, J., Phys. Lett. **A109**, 81 (1985).

ON THE ENTROPIC PROPERTIES OF QUANTUM DYNAMICAL SYSTEMS

W.A. Majewski

Institute of Theoretical Physics and Astrophysics

Gdańsk University, Gdańsk, Poland

ABSTRACT

A discussion of entropic properties of quantum dynamical systems is given. In particular, a notion of quantum chaos based on entropic properties of quantum system is discussed. Examples are used to illustrate the theory.

One of the aims of the theory of quantum dynamical systems is to help us to understand the basic questions of statistical physics. Among the other things this theory provides us with models, K-systems, which show how randomness can result from deterministic time evolution. The key quantity for investigating the chaotic behaviour of classical dynamical systems is the **Kolmogorov-Sinai invariant** - the dynamical entropy $h(\rho)$. Let me recall that $h(\rho)$ is related to sensitivity to initial conditions. In particular, for smooth dynamical classical systems one has **Pesin's formula**

$$h(\rho)=\sum \textit{ positive characteristic exponents}$$

However, applications of the theory of classical dynamical systems to physical problems are limited because Nature, at least on a microscopic level, is described by quantum laws.

In recent years the theory of quantum dynamical systems has made a considerable progress. In particular, the quantum analogs of *K-S* invariant, **Q.D.E.** (quantum dynamical entropy), were defined[1-2]. As this quantity has a number of promising properties I am interested in the following questions:

Q1. *Can one gain as much information about the quantum system from* **Q.D.E.** *as it was possible in the theory of classical dynamical systems?*

and

Q2. *Is* **Q.D.E.** *related to chaotic properties of a system?*

Up till now only partial answers are available (see review[3]). Let me recall that in the quantum case, one has no phase space, no trajectories and hence no Liapunov exponents in the usual sense. Therefore, **I shall consider the positivity of Q.D.E. as an indicator of chaotic properties of dynamics**. The aim of this lecture is to present models supporting this definition. Moreover, these models clearly show that **Q.D.E.** may be very useful for studying the chaotic properties of quantum dynamical systems. Consequently, these results constitute further partial answers to **Q1** and **Q2**.

REMARK: I would like to stress that I use the term *chaotic* in the sense of ergodic theory. Let me recall that in quantum optics[4], a (energy) level statistics is often identified with a quantum chaos. However, it seems that in the context of quantum optics the usage of the term *chaotic* has not a convincing justification. Moreover, it is as yet not clear, what is the relationship between an energy level statistics and chaotic behaviour of a system.

I. THE MODIFIED JAYNES-CUMMINGS MODEL[5]

I intend to investigate a physical system consisting of a two level atom S interacting with a propagating light beam R. The main idea for the modification of Jaynes-Cummings model is taken from the theory of kicked dynamics. Let me recall that the kicked dynamics provides simple models that classically show a chaotic behaviour, e.g. periodically kicked rotator[6]. So I shall study equations of motion of Jaynes-Cummings model in which the interaction of an atom with the light beam has been modified in order to describe alternations of the field. In other words, a sequence of modifications of field and its interaction with an atom produces a "discrete" process. This process can be considered as a good candidate for a description of an atom interacting with a "propagating" light beam. The interaction hamiltonian H_{SR} between λ-mode and the atom S is taken in RWA[7]. An application of the open-system theory for the reduction procedure of the dynamics[8] gives the following reduced time evolution of S:

Case 1: at the instant $(k-1)\tau$ the initial state of the light is given by n-photon state

$$\rho_s((k-1)\tau+t) = \sum_{l=1}^{3} V_l \rho((k-1))V_l^*$$

where ρ denotes a density matrix of the atom S, $t \in (0,\tau]$, V_l are operators on the Hilbert space of S, $t\in(0,\tau]$, $k=1,2,...$; τ denotes a time unit which is equal to a interaction time of an atom with a photon.

Case 2: at the instant $(k-1)\tau$ the initial state of the light is described by a coherent state

$$\phi(z) = \exp(-\frac{1}{2}|z|^2) \sum_{n=0}^{\infty} \frac{z^n}{(n!)^{1/2}} |n\rangle$$

$$\rho_s((k-1)\tau+t) = \exp(-\frac{1}{2}|z|^2) \sum_{n=0}^{\infty} W_n \rho_s((k-1)\tau) W_n^*$$

where $\tau \in (0,\tau]$, $z \in \mathbb{C}$ and W_n are operators on $\mathcal{H}_S$, depending on z,n,t.

The atom S is a finite system. Hence, **Q.D.E.**, introduced by Connes et al[1], h^Q is equal to 0. On the other hand, the **Benatti-Narnhofer argument**[9] clearly shows the existence of such states of S for which the von Neumann entropy is not monotonic. Finally, computing the Lindblad

entropy[2] h^L for this model we get

$$h^L(\text{case 2}) \geq h^L(\text{case 1}) \geq 0$$

Thus, summing up the results for such a finite system we have: h^Q vanishes; the von Neumann entropy is not applicable for description of chaotic properties. h^L entropy exhibits some basic properties of *K-S* invariant[2]. Moreover, if one combines the sensitivity of $h^L(\rho)$ with respect to the initial state of the field with the concavity of the entropy function h^L then it is clear that, at least for some cases, there is an additional production of entropy. It should be stressed that this phenomenon (the additional production of entropy) can occur at each instant $k\tau$. Moreover, for large time t, one ought to sum up this additional production. Thus, **this form of "reactualizing" the initial conditions and a slight chaotic behaviour of the model for each interval $[k\tau,(k+1)\tau)$ can produce the chaotic behavior of the system for large time.** Thus h^L can be used as a quantity describing chaotic behaviour of the considered model.

II. INFINITE FERMI SYSTEM

The power of **Q.D.E.** h^Q appears only for infinite systems. Therefore, my next model is an infinite system of fermions. The C^*-field algebra $\mathcal{A}$ of system is generated by even polynomials in creation operators $a^*(f)$ and the annihilation operators $a(g)$. Let ω_β be the Gibbs grand canonical equilibrium state, Π_ω the GNS representation given by ω_β. I consider the following dynamical systems:

$$(\Pi_\omega(\mathcal{A})'', \tau_x^\pi, \Omega_\omega) \qquad x\in\mathbf{R}$$
$$(\Pi_\omega(\mathcal{A})'', T_t, \Omega_\omega) \qquad t\in\mathbf{R}^+$$

where τ_x^π denotes action of the group $\mathbf{R}^\nu$ of space translations in the representation Π_ω, Ω_ω is a cyclic vector defining the state ω_β, and T_t is the dynamical semigroup given by

$$T_t a = \int_{\mathbf{R}^3} d\mu_t(x)\, \tau_x^\pi(a)$$

$$d\mu_t(x) = (4\pi t)^{-3/2} \exp\{-\frac{x^2}{4t}\}\, d^3x$$

Results[10,11]

$$h^Q(\tau_x^\pi) = |x| \int_{\mathbf{R}} \frac{dk}{2\pi} S(\rho(k))$$

where

$$S(x) = -x\ln x-(1-x)\ln(1-x), \qquad \omega_\beta(a^*(f)a(g)) = \int dk \overline{f(k)} g(k)\rho(k)$$

$$h^Q(T_t) = 0$$

Moreover, one can prove that the entropic dimension for the latter system is equal to 0. This last result means that one can consider the dynamics T_t as a very regular one.

III. INFINITE BOSE SYSTEM (A VERSION OF FORD-KAC-MAZUR MODEL[12])

Let $\mathfrak{A}$ be the CCR algebra over one dimensional Hilbert space $\mathbb{C}$, i.e. $\mathfrak{A}$ is generated by $\{W(z), z\in\mathbb{C}\}$ where

$$W(z)^* = W(-z), \qquad W(z)W(z') = e^{-\frac{i}{2}\,\mathrm{Im}(z,z')} W(z+z')$$

Further, let Ψ_β denote the Gibbs state. The GNS representation associated with $(\mathfrak{A},\Psi_\beta)$ will be denoted by $(\Pi(\mathfrak{A}),\mathcal{H},\Omega_\beta)$. I shall consider a semigroup time evolution τ_t, which one can interpret as describing the diffusion of a quantum particle in a harmonic well[13,14]

$$\tau_t : W_\pi(z) \longrightarrow W_\pi(e^{-\lambda t}z)\, \exp\{-\tfrac{1}{4}Q_\beta |z|^2 [1-\exp(-2\lambda t)]\}$$

where λ is a positive fixed constant, $t>0$, $W_\pi(\cdot) = \Pi_0 W(\cdot)$. τ_t has an ultra weak extension to a completely positive map on $\mathfrak{M} = \Pi(\mathfrak{A})''$. I shall denote the extension by the same letter τ. Thus, the following dynamical system has been defined

$$(\mathfrak{M}, \tau_t, \Omega_\beta)$$

For this system, $h_\omega^Q(\tau) = 0$ as well as the entropic dimension[12] of the diffusion τ_t is equal to 0. Therefore, this evolution can be considered as a very regular one (in the sense of ergodic theory).

IV. CONCLUSIONS AND REMARKS

(i) In order to describe a chaotic behaviour of finite systems one can use generalizations of *K-S* invariant, e.g. h^L. However, it should be stressed that h^Q is useless quantity for such (finite) systems.

(ii) T_t (see **II**) is constructed as a "mixture" of translations. Thus, we can see how dramatically **Q.D.E.** h^Q changes under the special mixture of transformations (cf. **II**).

(iii) Considering spin-boson models one can make the conjecture that **Q.D.E.** h^Q is equal to 0 for a large class of dynamical semigroups satisfying detailed balance equation. Thus, one can conclude that h^Q fits very well for the description of ergodic properties of quantum dynamical systems.

This work has been partially supported by M.E.N. grant.

References

[1] A. Connes, H. Narnhofer and W. Thirring, *Commun. Math. Phys.* **112**, 691 (1987)

[2] G. Lindblad, *Quantum ergodicity and chaos* in: Fundamental aspects of quantum theory; V. Gorini and A. Frigerio (eds.) *New York: Plenum Press 1986*

[3] H. Narnhofer, *Dynamical Entropy in Quantum Theory* UWThPh-1988-27

[4] Progress of Theoretical Physics, Supplement No 98 (1989)

[5] W.A. Majewski, *On the origin of chaos in Jaynes-Cummings models,* Gdańsk (1990)

[6] G. Casati, B.V. Chirikov and F.M. Izrailev and J. Ford *in* Lecture Notes in Physics **93**, *Springer N.Y.* (1979)

[7] W.H. Louisell, *Quantum Statistical Properties of Radiation,* New York (1973)

[8] E.B. Davies, *Quantum theory of open system, Academic Press* (1976)

[9] F. Benatti and H. Narnhofer, *Lett. Math. Phys.* **15**, 325 (1988)

[10] H. Narnhofer and W. Thirring, Lett. Math. Phys. **14**, 89 (1987)

[11] W.A. Majewski and M. Kuna, Lett. Math. Phys. **20**, 337 (1990)

[12] W.A. Majewski, *On the entropic properties of the Ford-Kac-Mazur model*, Gdańsk (1990)

[13] G.W. Ford, M. Kac and P. Mazur, *J. Math. Phys.* **6**, 504 (1965)

[14] G.G. Emch, *Non-equilibrium Quantum Statistical Mechanics, in* Current Problems in elementary particles and physics, ed. P. Urban, *Springer Verlag* (1976)

ON THE CHOICE OF INVARIANT VARIABLES FOR TWO-PARTICLE PROCESSES

V.R.Garsevanishvili
Mathematical Institute of the Academy of Sciences
of the Georgian SSR, Tbilisi, USSR

M.B.Sheftel
North-West Politechnical Institute, Leningrad, USSR

A two-particle scattering amplitude can be represented as a function on the hyperboloid using the transformation from the Mandelstam invariants s, t, u to new invariant variables. In this paper we consider the stereographic projection from the hyperboloid onto the plane, leading to the new invariant variables. In the physical region the scattering amplitude expressed in terms of new variables is a function defined on the plane. The interior of the unit circle on the plane corresponds to the physical region of the direct channel, whereas its exterior to the physical regions of cross-channels.

1. Introduction

The choice of variables plays an important role in the analysis of various characteristics of physical systems. A wrong choice of variables leeds, for example, to kinematic singularities and additional kinematic conditions (see, for example, /1/), whereas their adequate choice makes it possible to find additional symmetries of the problem under consideration (see, for example, /2/).

If we restrict our consideration to the analysis of the scattering amplitude in the physical region, then the Mandelstam variables s, t, u are inconvenient, since the boundary of the physical region has a complicated form in these variables and the variation limits of anyone of them depend on the values of the other variables. This, in particular, makes it difficult to study the amplitude by means of expansions in complete sets of functions in the physical region.

The two-particle scattering amplitude can be represented as a function on the hyperboloid, using the transformation from the Mandelstam invariants to the new invariant variables, namely to the coordinates on the hyperboloid /3/. The new variables have independent variation limits, thus allowing one to expand the amplitude in a complete set of functions on the hyperboloid.

In this paper we consider the stereographic projection from the hyperboloid onto the equatorial plane /4/. In this case instead of the coordinates on the hyperboloid there arise new invariant variables called coordinates on the plane. The physical scattering amplitude is the function defined on the plane; in these case the interior of the unit circle corresponds to the physical region of the direct channel and its exterior to the physical regions of cross-channels. All points of the new plane are physical: to each point of the physical region in the Mandelstam plane there corresponds in the one-to-one manner a point of the Poincare plane.

If it is further assumed to which class of functions on the plane the physical scattering amplitude belongs, then this assumption will be a dynamic requirement from which we obtain corollaries for the observed quantities. For this reason here we are concerned only with the kinematic part of the problem connected with the mapping onto the Poincare plane and with the choice of variables.

2. The Scattering Amplitude as a Function on the Hyperboloid

For the two-particle reaction

$$a_1 + a_2 \to a_3 + a_4$$

with equal masses the scattering amplitude can be represented as a function on the hyperboloid /3/

$$x_0^2 - x_1^2 - x_2^2 = 1 \qquad (1)$$

Here we shall choose the following specific relation on the variables x_0, x_1, x_2 with the Mandelstam variables s, t, u in the direct s-channel:

$$x_0 = \frac{\sqrt{s}}{2m}, \quad x_1 = \frac{1}{m}\sqrt{\frac{tu}{s-4m^2}}, \quad x_2 = \frac{u-t}{2m\sqrt{s-4m^2}} \qquad (2)$$

(The main results are not related with the choice of this specific parametrization). The crossing-transformation is obtained by the analytical continuation of the coordinates x_0, x_1, x_2:

$$x_0' = \varepsilon i x_0 = \frac{\varepsilon\sqrt{-s}}{2m}, \quad x_1' = x_1, \quad x_2' = i x_2 = \frac{u-t}{2m\sqrt{4m^2-s}} \qquad (3)$$

$\varepsilon = \pm 1$ for the u- or the t-channel, respectively, and the cross-channels are mapped onto the hyperboloid:

$$x_2'^2 - x_1'^2 - x_0'^2 = 1 \qquad (4)$$

The right sheet of hyperboloid (4) with $x_2' > 0$ corresponds to the u-channel and the left sheet with $x_2' < 0$ to the t-channel.

Note that only the points of the upper hyperboloid (1), satisfying the condition

$$x_0 \geqslant 1, \quad x_1 \geqslant 0 \tag{5}$$

correspond to the s-channel, whereas the points of hyperboloid (4), satisfying the conditions

$$x'_0 \geqslant 0, \quad x'_1 \geqslant 0, \quad x'_2 \geqslant 1 \qquad \text{(u-channel)} \tag{6}$$

and

$$x'_0 \leqslant 0, \quad x'_1 \geqslant 0, \quad x'_2 \leqslant -1 \qquad \text{(t-channel)} \tag{7}$$

correspond to the u- and t-channels.

3. Stereographic Projection from the Hyperboloid onto the Plane

The hyperboloid geometry can be realized on the plane [4,5]. It is obtained by means of the stereographic projection of the upper sheet of the hyperboloid from the lower sheet vertex onto the coordinate plane $x_0 = 0$.

Let us first consider the upper sheet of hyperboloid (1), corresponding to the s-channel. The stereographic projection is given by the formulas

$$x = \frac{x_1}{x_0 + 1}, \quad y = \frac{x_2}{x_0 + 1} \tag{8}$$

where x, y are the Cartesian coordinates on the plane $x_0 = 0$.

The inverse transformation has the form

$$x_0 = \frac{1 + x^2 + y^2}{1 - x^2 - y^2}, \quad x_1 = \frac{2x}{1 - x^2 - y^2}, \quad x_2 = \frac{2y}{1 - x^2 - y^2} \tag{9}$$

Formulas (8), (9) establish the one-to-one correspondence between the points of the upper sheet of hyperboloid (1) corresponding to the s-channel (condition (5)) and the points of the right half of the unit circle on the Poincare plane with the centre at the origin

$$x^2 + y^2 \leqslant 1, \quad x \geqslant 0 \tag{10}$$

Let us now consider the stereographic projection of hyperboloid (4), corresponding to cross-channels, onto its equatorial plane $x'_2 = 0$:

$$x' = \frac{x'_1}{\varepsilon x'_2 + 1}, \quad y' = \frac{x'_0}{\varepsilon x'_2 + 1} \tag{11}$$

$$\varepsilon = \pm 1 \tag{12}$$

Here x', y' are the Cartesian coordinates on the plane $x'_2 = 0$. In this case the physical region of the u-channel is mapped onto quadrant I of the unit circle on the plane $x'_2 = 0$.

$$x'^2 + y'^2 \leqslant 1, \quad x' \geqslant 0, \quad y' \geqslant 0 \tag{13}$$

and the physical region of the t-channel is mapped onto quadrant IV of the same circle

$$x'^2 + y'^2 \leqslant 1, \quad x' \geqslant 0, \quad y' \leqslant 0 \tag{14}$$

(see conditions (6), (7)).

Next we map the interior of the unit circle on the plane $x'_2 = 0$ onto its exterior and combine the planes $x'_2 = 0$ and $x_0 = 0$. The direct channel now occupies the interior of the unit circle on the right half-plane, whereas cross-channels occupy the exterior of the circle on the same half-plane.

The interior of the circle is mapped onto its exterior by transforming the inversion

$$z = \frac{1}{z'} \tag{15}$$

where

$$z = x + iy, \quad z' = x' + iy' \tag{16}$$

and the new coordinates x, y are defined as follows:

$$x = \frac{x'}{x'^2 + y'^2}, \quad y = -\frac{y'}{x'^2 + y'^2} \tag{17}$$

In this case the physical region of the u-channel is mapped onto the exterior of the unit circle in quadrant IV:

$$x^2 + y^2 \geqslant 1, \quad x \geqslant 0, \quad y \leqslant 0 \tag{18}$$

and the region of the t-channel is mapped onto the exterior of the circle in quadrant I:

$$x^2 + y^2 \geqslant 1, \quad x \geqslant 0, \quad y \geqslant 0 \tag{19}$$

Taking into account transformations (11) and (17), for cross-channels the relation of the variables x, y with the coordinates on hyperboloid (4) takes the form

$$x = \frac{x'_1}{\varepsilon x'_2 - 1}, \quad y = -\frac{\varepsilon x'_0}{\varepsilon x'_2 - 1} \tag{20}$$

where the quantity

$$x'_0 = i x_0 \tag{21}$$

is now equally defined for the t- and u-channels. (The quantity x'_0 defined by (3) differed in the sign for the u- and t-channels).

Let us now combine the planes $x_0 = 0$ and $x'_2 = 0$ and call the obtained plane the z-plane:

$$z = x + iy \tag{22}$$

The physical region of the direct channel is mapped onto the right half of the unit circle, whereas the cross-channels are mapped onto the exterior of the semi-circle in the right half-plane by means of transformations (8) and (20), respectively. All points of the right z-half-plane are physical.

4. Transformation of z-Half-Plane onto the w-Plane to the Universal Variable w

The new invariant variables x, y are defined by formulas (8), (20) in a different manner for different channels. To obtain the variable w which is universal for all channels we must carry out the transformation of the z-half-plane onto the w-plane transforming the right semi-circle into the unit circle. This transformation leaves three points of the unit circle fixed:

$$(x=1, y=0), \quad (x=0, y=1), \quad (x=0, y=-1) \tag{23}$$

and transforms the point $(x=0, y=0)$ onto the point with the coordinates $(U=-1, V=0)$. These conditions define the desired transformation as

$$w = \frac{z^2+2z-1}{-z^2+2z+1} \tag{24}$$

where

$$w = U+iV \tag{25}$$

The Cartesian coordinates U, V in the w-plane are related with the coordinates x, y in the z-plane by the relations

$$U = \frac{4x^2-(x^2+y^2-1)^2}{[(x-1)^2+y^2-2]^2+8y^2}, \tag{26a}$$

$$V = \frac{4y(x^2+y^2+1)}{[(x-1)^2+y^2-2]^2+8y^2}. \tag{26b}$$

Formulas (8) and (20) make it possible to relate the coordinates in the w-plane with the coordinates on the hyperboloid:

$$U = \frac{x_1^2-1}{(x_1+1)^2+2x_2^2}, \tag{27a}$$

$$V = \frac{2x_2x_0}{(x_1+1)^2+2x_2^2}. \tag{27b}$$

for the direct channel and

$$U = \frac{x_1'^2-1}{(x_1'^2-1)^2+2x_0'^2}, \tag{28a}$$

$$V = \frac{-2x_2'x_0'}{(x_1'^2-1)^2+2x_0'^2}. \tag{28b}$$

for the cross-channels. Using relations (3) between the coordinates on hyperboloids (1) and (4), we can express U, V in terms of x_0, x_1, x_2 for all channels:

$$U = \eta \frac{x_1^2-1}{(x_1+1)^2+2x_2^2}, \tag{29a}$$

$$V = \eta \frac{2x_2x_0}{(x_1+1)^2+2x_2^2}. \tag{29b}$$

where

$$\eta = \operatorname{sign}[(x_1+1)^2+2x_2^2] \tag{30}$$

Since

$$[(x_1+1)^2+2x_2^2] = -[(x_1'-1)^2+2x_0'] \tag{31}$$

we have

$$\eta = \pm 1 \tag{32}$$

for the direct channel and cross-channels, respectively.

Thus formulas (29) define the transformation from the coordinates on the hyperboloid to the new invariant variables U, V which are the Cartesian coordinates in the w-plane and universal for all channels. To the physical region of the direct channel there corresponds the interior of the unit circle in the w-plane with the centre at the origin. To the physical region of the u- and t-channels there corresponds the exterior of the unit circle in the upper and the lower half-plane, respectively.

Transformation (29) from the hyperboloid onto the w-plane does not depend on the parametrization of the hyperboloid through Mandelstam's invariants, i.e., it does not depend on the relation between x_0, x_1, x_2 and s, t, u.

5. The Correspondence Between the Physical Region in the Mandelstam Plane and the z-Plane

Here a detailed consideration will be given to only one parametrization of the hyperboloid defined by formulas (2) and most closely related with the energy and the scattering angle in the centre-of-mass system of the s-channel. Then formulas (2) and (8) define the transformation of the physical Mandelstam region of the s-channel onto the right half of the unit circle in the z-plane

$$x = \frac{2\sqrt{tu}}{(\sqrt{s}+2m)\sqrt{s-4m^2}}, \tag{33a}$$

$$y = \frac{u-t}{(\sqrt{s}+2m)\sqrt{s-4m^2}}. \tag{33b}$$

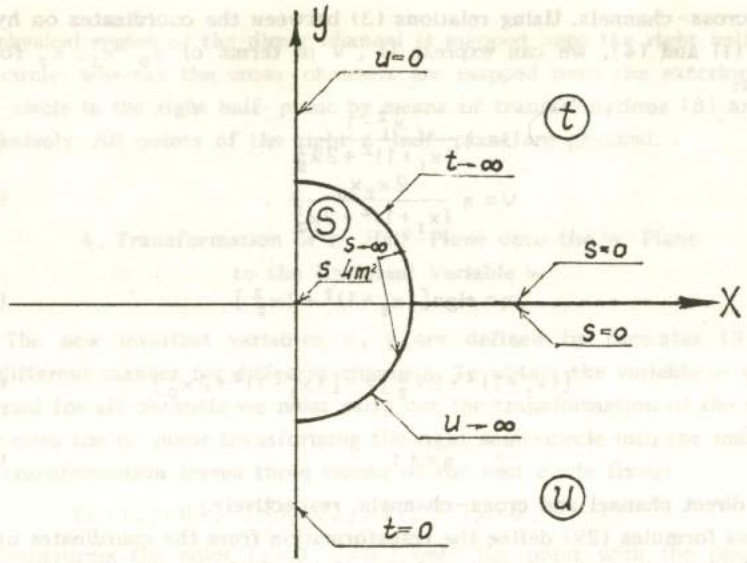


Fig. 1. z-plane

The lines (see Fig. 1) of the fixed energy in the z-plane are the right halves of the circles of the fixed radius:

$$r^2 = x^2 + y^2 = \frac{\sqrt{s} - 2m}{\sqrt{s} + 2m} \leqslant 1 \tag{34}$$

The points of the unit semi-circle correspond to asymptotic energies

$$r \to 1 \qquad \text{when } s \to \infty \tag{35}$$

for various fixed values of the scattering angle ϑ_s in the centre-of-mass system (c.m.s.) of the s-channel. The lines of the fixed ϑ_s are rays coming from the origin of the coordinates on which the polar angle φ is fixed

$$\sin\varphi = \frac{y}{\sqrt{x^2+y^2}} = \frac{u-t}{\sqrt{s-4m^2}} = -\cos\vartheta_s \tag{36}$$

or

$$\varphi = \vartheta_s - \frac{\pi}{2} \tag{37}$$

where

$$0 \leqslant \vartheta_s \leqslant \pi\,, \qquad -\frac{\pi}{2} \leqslant \varphi \leqslant \frac{\pi}{2} \tag{38}$$

The transformation of the physical regions of the cross-channels in the

Mandelstam plane onto the exterior of the unit circle in the right z-half-plane is defined by formulas (3), (20):

$$x = \frac{\sqrt{-tu}}{\varepsilon\frac{u-t}{2} - m\sqrt{4m^2-s}}\,, \tag{39a}$$

$$y = \frac{\frac{1}{2}\sqrt{s(s-4m^2)}}{\frac{u-t}{2} - \varepsilon m\sqrt{4m^2-s}}\,. \tag{39b}$$

where $y \geqslant 0$ for the t-channel and $y \leqslant 0$ for the u-channel. The new variables x, y which are the Cartesian coordinates on the z-plane are defined by formulas (33) and (39) in a different manner in the direct channel and in cross-channels. Therefore it is more convenient to consider the crossing-transformation in the w-plane where the Cartesian coordinates U, V are the variables which are universal for all channels.

6. The Correspondence Between the Physical Region in the Mandelstam Plane and the w-Plane

Using formulas (2) and (29), we obtain the correspondence between the points of the physical region in the Mandelstam plane and the w-plane:

$$U = \eta\frac{tu - m^2(s-4m^2)}{(m\sqrt{s-4m^2} + \sqrt{tu})^2 + \frac{1}{2}(u-t)^2}\,, \tag{40a}$$

$$V = \eta\frac{\frac{1}{2}(u-t)\sqrt{s(s-4m^2)}}{(m\sqrt{s-4m^2} + \sqrt{tu})^2 + \frac{1}{2}(u-t)^2}\,. \tag{40b}$$

where

$$\eta = \pm 1 \tag{41}$$

for the direct channel and cross-channels, respectively.

In terms of the variables U, V the physical scattering amplitude is a function defined on the plane. The exterior of the unit circle with the centre at the origin corresponds to the physical region of the s-channel, its exterior in the upper half-plane corresponds to the t-channel, the exterior of the circle in the lower half-plane to the u-channel. The arc of the unit circle in quadrants I and IV corresponds to the asymptotic energies in each of the channels for various values of the scattering angle. In particular, the point ($U=0$, $V=-1$) corresponds to the forward scattering ($t=0$) when $s \to \infty$, whereas the point ($U=0$, $V=1$) to the backward scattering when $s \to \infty$. The arc of the unit circle in quadrant II corresponds to the backward scattering ($u=0$)

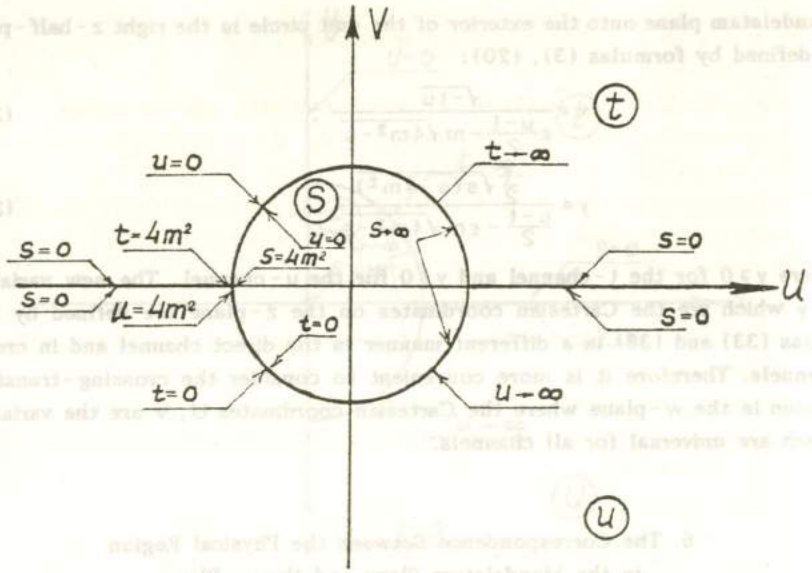


Fig. 2. w-plane

and in quadrant III it corresponds to the forward scattering ($t = 0$) for various energies. The point ($U = -1$, $V = 0$) corresponds to the threshold of any channel. This correspondence can be seen more clearly in Fig. 2.

7. Conclusion

For the physical values of the Mandelstam invariants s, t, u the variables $U(s,t,u)$ and $V(s,t,u)$, which we have introduced and which are defined as Cartesian coordinates in the w-plane, take real values and the set of their values fills up the entire w-plane.

If the scattering amplitude is considered as a function of two complex variables s and t, then the w-plane will correspond to that region of the 4-dimensional manifold of real quantities $\mathrm{Re}\, s$, $\mathrm{Im}\, s$, $\mathrm{Re}\, t$, $\mathrm{Im}\, t$, where $\mathrm{Im}\, s = \mathrm{Im}\, t = 0$. The consideration of the scattering amplitude as a function on the w-plane may be helpful in studying the cross-symmetry and problems of the asymptotic behaviour of various characteristics of the elastic scattering.

The authors wish to express their sincere thanks to A.D.Donkov, V.G.Kadyshevsky, M.A.Mestvirishvili, V.A.Meshcheryakov, R.M.Mir-Kasimov, I.S.Shapiro, D.I.Shirkov, N.B.Skachkov, L.A.Slepchenko, Yu.A.Smorodinsky, A.N.Tavkhelidze, I.T.Todorov, V.S.Vladimirov for interesting discussions.

References

1. G.Cohen-Tannoudji, A.Morel, N.Navelet. Ann.Phys., 46, 239 (1968).
   E.Leader. Phys.Rev., 166, 1599 (1968).
2. V.A.Fock. Zc.Phys., 98, 145 (1935).
   N.N.Bogolyubov. UMZh, 2, 3 (1950), see also Selected Works, v.2., Kiev (1970).
   M.Toller. Nuovo Cim., 53A, 671 (1968).
3. P.Vinternitz, Ya.A. Smorodinski, M.B.Sheftel.YaF, 7, 1325 (1968).
4. H.Poincare, Les geometries non-euclidiennes, Oeuvres, Paris Gauthier-Villars, 1951.
5. B.A.Rosenfeld. Non-Euclidean spaces, "Nauka", Moscow, 1969.
   I.M.Gelfand, M.I.Graev, N.Ya.Vilenkin. Generalized functions, v.5., Fizmatgiz, Moscow, 1962.

LIST OF PARTICIPANTS

E. ALDROVANDI
SISSA

Strada Costiera 11
34014 TRIESTE
ITALY

M. ANTONOWICZ
Uniwersytet Warszawski
Instytut Fizyki Teoretycznej
ul.Hoza 69
00-681 Warszawa
POLAND

I. BAKAS
University of Maryland
Department of Physics
College Park
20 742 Maryland
U S A

M. BLASZAK
Uniwersytet A.Mickiewicza
Instytut Fizyki
ul.Grunwaldzka 6
60 -780 Poznan
POLAND

A. BOGUSZ
Uniwersytet M.Sklodowskiej
Zaklad Fizyki
pl.M.S.Sklodowskiej 1
20 - 031 Lublin
POLAND

A. BOROWIEC
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

J. BURZLAFF
Dublin City University
Ollscoil Chathair Bhaile
Atha Cliath
9 Dublin
IRELAND

A.M. CETTO
Universidad Nacional Autonoma
de Mexico
Institute de Fisica
01 000 Mexico
MEXICO

M.RUIZ ALTABA
CERN

CH-1211 Geneva
SWITZERLAND

H. ARODZ
Uniwersytet Jagiellonski
Instytut Fizyki
ul. Reymonta 4
30-059 KRAKOW 16
POLAND

I. BARASHENKOV
Joint Institute for Nuclear
Research JINR
Dubna
101 000 Moskva
U S S R

A. BOBYK
Uniwersytet M.Sklodowskiej
Zaklad Fizyki
pl.M.C.Sklodowskiej 1
20 - 031 Lublin
POLAND

L. BONORA
S I S S A
Strada Costiera 11

34014 Triesta
ITALY

M. BOZEJKO
Uniwersytet Wroclawski
Instytut Matematyki
pl.Grunwaldzki 7
50-210 Wroclaw
POLAND

G. CASATI
Universida degli studi di
Milano ,Dipart. di Fisica
via Celoria 16
20 133 Milano
ITALY

A. CRUMEY
Centre for Nonlinear Studies
Depart. of App.Math.Studies

LS2 9JT Leeds
GREAT BRITAN

I.H. DURU
Research Council of Turkey
Marmara Scient. and Industrial
Research Center.
41470 Gebze-Kocaeli
TURKEY

A. FORDY
Centre for Nonlinear Studies
Depart. of App.Math.Studies

LS2 9JT Leeds
GREAT BRITAN

Z. GALASIEWICZ
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

V. GARSEVANISHVILI
Academy of Science of the
Georgian SSR
Z.Rukhadze 1
380093 Tbilisi
U S S R

R. GIELERAK
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

M. GORZELANCZYK
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

D. GRECH
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

Z. HABA
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

P. DYOMKIN
Vilnius University
Departmentof Physics
Sauletekio al.9
232054 Vilnius
LITHUANIA

W. FUSHCHIH
Institute of Mathematics
Repin Street 3

4 Kiev
U S R R

P. GARBACZEWSKI
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

F. GESZTESY
University of Missouri
Department of Mathematics
Columbia
65211 Missouri
U S A

J. GORA
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

M. GOZDZ
Uniwersytet M.Sklodowskiej
Zaklad Fizyki
pl.M.C.Sklodowskiej 1
20 -031 Lublin
POLAND

P. GUSIN
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

W. HANN
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

E. HELLWIG
Tech.Univ.Berlin
Inst.f.Strahlungs-u Kernphysik
Hardenberstr.36
D-1000 Berlin
Germany

A. ICHA
Institute of Oceanology
Polish Academy of Science
ul.Powstancow Warszawy 55
81-967 Sopot
POLAND

A. JADCZYK
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

C. JUSZCZAK
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

M. KEYL
Tech.Univ.Berlin
Inst.f.Strahlungs-u.Kernphysik
Hardenberstr.36
D-1000 Berlin
GERMANY

A. KORNAKIEWICZ
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

J. KUPSCH
Univeritat Kaiserslautern
Fachbereich Physik
E.Schrodinger str.46
D-6750 Kaiserslautern
R F G

E. LANGMANN
Technische Universtat Graz
Inst. fur Theor. Physik
Petergasse 16
A-8010 Graz
AUSTRIA

J. HOLYST
Politechnika Warszawska
Instytut Fizyki
Koszykowa 75
00 602 Warszawa
POLAND

A. ISAEV
Joint Institute for Nuclear
Research
JINR Dubna
101 000 Moskov
U S S R

B. JANCEWICZ
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

W. KARWOWSKI
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

M. KLIMEK
Wyzsza Szkola Pedagogiczna
Katedra Fizyki
al.J.Pawla II
42-219 Czestochowa
Poland

J. KRAWCZYNSKI
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

M. LAKSHMANNAN
Bharathidasan University
Department of Physics
Centre for Nonlinear Dynamics
620 024 Tiruchirapali
INDIA

Z. LIPINSKI
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

J. LOEFFELHOLZ
Karl-Mark-University
Sektion Physik
Karl-Marx-Platz 10-11
D-7010 LEIPZIG
GERMANY

M. MAAMACHE
Universite De Nice-Sophia
Laboratorie De Phisi.Theor.
Parc Valrose
06034 Nice
France

A. MAJEWSKI
Uniwersytet Gdanski
Instytut Fizylki Teoretycznej
ul.Wita Stwosza 57
80 -952 Gdansk
POLAND

R. VILELA MENDES
CERN

CH-1211 Geneva 23
SWITZERLAND

T. MIWA
Kyoto University
Research Institute for
Mathematical Sciences
603 Kyoto
JAPAN

W. MULAK
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

W. OEVEL
Loughborough University
Department of Mathematical
Sciences
LE11 8TU Loughborough
Great Britan

A. PODOLSKA
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

J. LUKIERSKI
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

M. MACKEY
McGill University
Centre for Nonlinear Dynamics
3655 Drummond Street
H3G 1Y6 Montreal
CANADA

P. MATHIEU
Universite Laval
Department de Physique

G1K 7P4 Laval
CANADA

P. MINNAERT
Universite de Bordeaux I
Laboratorie de Phys.Theorique
Rue du Solarium
F-33170 Gradignan
France

M. MOZRZYMAS
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

G. OEVEL
Universitat Paderborn
Fachbereich 17
Mathematik-Informatik
D-4790 Paderborn
R F G

L. de la PENA
Universidad Nacional Autonoma
de Mexico
Instituto de Fisica
20 D.F. Mexico
MEXICO

Z. POPOWICZ
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

D. PROROK
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

J. ROZANSKI
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

P. RUSEK
Politechnika Wroclawska
Instytut Fizyki
ul.Wybrzeze Wyspianskiego
50-510 Wroclaw
Poland

S. SAITO
Tokyo Metropolitan University
Department of Physics
2-1-1 Fukazawa
158 Tokyo
JAPAN

A. SEMIKHATOV
Academy of Svciences of USSR
P.N.Lebedev Inst.of Physics
Leninskij pr.53
117 924 Moskva
U S S R

I. STRACHAN
University of Durham
Department of Mathematics
South Road
DH1 3LE Durham
Great Britan

L. TOLSTOY
Academy of Sciences of USSR
L.Landau Inst.for Theor.Phys.
ul.Kosygin 2
117 334 Moskva
U S S R

J. ZAGRODZINSKI
Polska Akademia Nauk
Instytut Fizyki
al.Lotnikow 32
02-668 Warszawa
POLAND

M. ROGATKO
Uniwersytet M.Sklodowskiej
Zaklad Fizyki
pl.M.C.Sklodowskiej 1
20 -031 Lublin
POLAND

S. RUDNICKI
Uniwersytet Gdanski
Instytut Fizyki Teoretycznej
ul.Wita Stwosza 57
80 - 952 Gdansk
Poland

J. RZEWUSKI
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

G. SEMENOFF
University of British Columbia
Department of Physics
6224 Agriculture Road
BC V6T 2A6 Vancover
CANADA

J. SOBCZYK
Uniwersytet Wroclawski
Instytut Fizyki Teoretycznej
ul.Cybulskiego 36
50 - 205 Wroclaw
Poland

J. TAFEL
Uniwersytet Warszawski
Instytut Fizyki Teorettycznej
ul.Hoza 69
00-681 Warszawa
POLAND

A. TRUMAN
University College of Swansea
Singleton Park

SA2 8PP Swansea
Great Britan

W. ZAKRZEWSKI
University of Durham
Mathematical Department
South Road
DH1 3LE Durham
Great Britan