

---

# **MEASUREMENTS IN QUANTUM MECHANICS**

---

Edited by **Mohammad Reza Pahlavani**

**INTECHWEB.ORG**

## **Measurements in Quantum Mechanics**

Edited by Mohammad Reza Pahlavani

### **Published by InTech**

Janeza Trdine 9, 51000 Rijeka, Croatia

### **Copyright © 2012 InTech**

All chapters are Open Access distributed under the Creative Commons Attribution 3.0 license, which allows users to download, copy and build upon published articles even for commercial purposes, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications. After this work has been published by InTech, authors have the right to republish it, in whole or part, in any publication of which they are the author, and to make other personal use of the work. Any republication, referencing or personal use of the work must explicitly identify the original source.

As for readers, this license allows users to download, copy and build upon published chapters even for commercial purposes, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

### **Notice**

Statements and opinions expressed in the chapters are those of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

**Publishing Process Manager** Maja Bozicevic

**Technical Editor** Teodora Smiljanic

**Cover Designer** InTech Design Team

First published February, 2012

Printed in Croatia

A free online edition of this book is available at [www.intechopen.com](http://www.intechopen.com)

Additional hard copies can be obtained from [orders@intechweb.org](mailto:orders@intechweb.org)

Measurements in Quantum Mechanics, Edited by Mohammad Reza Pahlavani

p. cm.

ISBN 978-953-51-0058-4

# INTECH

open science | open minds

**free** online editions of InTech  
Books and Journals can be found at  
**[www.intechopen.com](http://www.intechopen.com)**



---

# Contents

---

## **Preface IX**

- Chapter 1 **Measurement in Quantum Mechanics: Decoherence and the Pointer Basis 1**  
Anu Venugopalan
- Chapter 2 **Time as Quantum Observable, Canonical Conjugated to Energy 17**  
Vladislav S. Olkhovsky, Erasmo Recami and Sergei P. Maydanyuk
- Chapter 3 **Order of Time Derivatives in Quantum-Mechanical Equations 57**  
Jan Jerzy Sławianowski
- Chapter 4 **Theory of "Weak Value" and Quantum Mechanical Measurements 75**  
Yutaka Shikano
- Chapter 5 **Generalized Non-Relativistic Supersymmetric Quantum Mechanics 101**  
Thomas L. Markovich, Mason T. Biamonte, Eric R. Bittner and Donald J. Kouri
- Chapter 6 **A Statistical Derivation of Non-Relativistic Quantum Theory 141**  
Ulf Klein
- Chapter 7 **The Quantum Completeness Problem 175**  
Carsten Held
- Chapter 8 **Quantum Correlations in Successive Spin Measurements 197**  
Ali Ahanj
- Chapter 9 **The Gluon Emission Model for Vector Meson Decay 229**  
D. White

- Chapter 10 **Vector Correlations in Collision of Atom and Diatomic Molecule** 249  
Xian-Fang Yue
- Chapter 11 **Schrödinger Transform of Image: A New Tool for Image Analysis** 269  
Liantang Lou, Hua Zeng, Jipeng Xiong,  
Lingling Li and Wenliang Gao
- Chapter 12 **Entanglement in Two and Three Quantum Mechanical Systems** 293  
L. M. Arévalo Aguilar, M. M. Méndez Otero  
and P. Mastranzo Ortega
- Chapter 13 **Quantum Spin Coherent Electron Correlation** 309  
Peter G. Burton
- Chapter 14 **Toward Automatic Regularization for Feynman Loop Integrals in Perturbative Quantum Field Theory** 333  
E. de Doncker and F. Yuasa





---

## Preface

---

The volume *Measurements in Quantum Mechanics* explains different aspects of measurements in quantum mechanical systems. Understanding of quantum mechanics requires a careful definition of measurement. The subject of measurement is a basic part of the problem of the interpretation of quantum mechanics, for which there is currently no consensus. It is a postulate of quantum mechanics that all measurements have an associated observable. The quantum state of a system is a mathematical object that fully describes the quantum systems. From the inception of quantum mechanics, the concept of measurement proved to be a source of difficulties that found a concrete expression in the Einstein-Bohr debates, out of which both the Einstein Podolsky Rosen paradox and the Schrödinger's cat paradox developed. In brief, the difficulties stemmed from an apparent conflict between several principles of the quantum theory of measurement. In particular, the linear dynamics of quantum mechanics seemed to conflict with the postulate that during measurement a non-linear collapse of the wave packet occurred. The measurement problem is not just an interpretational difficulty internal to quantum mechanics. It raises broader issues as well, such as the philosophical debate between two different ideas. Quantum mechanics can be presented through concrete examples. I believe that most physicists learn through specific examples, test the obtained results through measurements and then find it easy to generalize. Spin measurements do not need sophisticated experimental devices, so the first postulate devoted through Estern-Gerlakh experiments of spin for measurements theory. Non-relativistic quantum mechanics has experienced a phase of turbulent development in the past few years. Quantum computing, quantum teleportation, quantum cryptography, and quantum information are typical buzzwords that reflect these developments. In spite of the enormous success of quantum mechanics to predict a wide range of physical phenomena, the quantum measurement problem through quantum mechanics is accepted as the fundamental theory of nature and gives satisfactory predictions to deduce the results of measurements.

This volume contains fourteen chapters that created by a group of invited authors from all over the world. A brief outline of the book is as follows:

Chapter one reviews the coherent and decoherent systems, which is the main goal of measurement theory. Time and energy as measurable quantities and their relations

through uncertainty principal are discussed in chapter two. The different nature of time and energy as a parameter or observable in various parts of physics is also reviewed in this chapter. Chapter three deals with the order of time derivative in different branches of relativistic and nonrelativistic quantum mechanics. Dirac and Clain-Gordon equations, together with a group representation of subject, are also studied phenomenologically in this chapter. Chapter four covers the relation between the weak value theorem and the measurements in quantum mechanics. Supersymmetric quantum mechanics is presented in chapter five. Electronic structure of multielectron systems and their relation with the measurement is also discussed in this chapter. Chapter six relates to the statistical derivation of quantum mechanics, the probability to measure energy as a conjugate observable of time and the minimal Fisher information theory. Quantum completeness problem and its relation to quantum measurements are discussed in chapter seven. Chapter eight presented the quantum correlations in successive spin measurements. Application of multimeasurements results to spin  $\frac{1}{2}$  systems and well-known Bell inequality are also discussed in this chapter. Spin coherent measurement, electron correlations in quantum chemistry and their application in gravitational bach ground are reported in chapter nine. Chapter ten deals with vector correlations in the collision of atom and diatomic molecule. Different types of mesons and their decay modes are also discussed in this chapter. Vector correlations in collision of atoms and diatomic molecules, their collision cross section and different angular distributions are reported in chapter eleven. Chapter twelve discusses the entanglement in two and three quantum mechanical systems and their relations with quantum mechanical measurements. Basic conventions of quantum chemistry and their relation with measurements in multi-electron systems, together with applications of this new model to maximize the intensity of images are reported in chapter thirteen. The last chapter deals with the Feynman loop integral in perturbative quantum field theory and their application in measurements of new type of colliders.

This book is a result of collaboration of the international community and I thank all the authors. I especially want to thank Ms. Maja Bozicevic for her valuable assistance through the publishing process and the InTech publishing team for their efforts in preparing the book for publication.

**M. R. Pahlavani**

Head of Nuclear Physics Department,  
Mazandaran University, Mazandaran, Babolsar,  
Iran





# Measurement in Quantum Mechanics: Decoherence and the Pointer Basis

Anu Venugopalan

*University School of Basic and Applied Sciences, GGS Indraprastha University,  
Sector 16C, Dwarka, New Delhi-110075  
India*

## 1. Introduction

Quantum Mechanics represents one of the greatest triumphs of the scientific enterprise of the twentieth century. The stunning success of quantum theory has led to many revolutionary inventions and its extraordinary concepts describe the heart of several important real world applications like transistors and lasers. The theory makes accurate predictions about a wide range of physical phenomena and has historically withstood the tests and scrutiny of every experimental investigation. However, inspite of the fact that quantum theory is widely regarded by the scientific establishment as the fundamental theory of nature and is immensely successful and useful, its conceptual framework makes many predictions which are difficult to comprehend "classically". The theory is, paradoxically, powerful and confusing at the same time. Quantum theory's unusual predictions originate from its basic formalism which involves concepts like probability amplitudes and the linear superposition principle (Dirac, 1947). The quantum view appears abstract and counterintuitive and at odds with classical perceptions. Many of the conceptual problems of quantum mechanics are encompassed in what is known as the *quantum measurement problem* (Peres, 1986; von Neumann, 1932; Wheeler & Zurek, 1983). Conventionally, the measurement paradox is supposedly 'resolved' by forcing a notion of a sudden collapse of the state vector of the system being measured. However, the nature of this mechanism is at odds with the basic tenets of quantum mechanics and hence may lie outside its realm thus questioning the validity of the theory itself. Closely related with the problem of measurement in quantum mechanics is the question of its connection with the emergence of classicality and the elusive boundary between quantum and classical worlds. What is the connection between the 'classical' and the 'quantum' worlds? Is there a definite relationship? Are classical mechanics and quantum mechanics two mutually exclusive incompatible theories or are they two aspects of the same underlying philosophy? Classical objects are eventually composed of elements of the microworld which can be described quantum mechanically. So, how and where can there be a boundary between the two worlds? In the following section we begin by introducing the quantum measurement problem. In the next section we will discuss attempts to understand and explain away the underlying paradoxes in quantum theory as highlighted in the quantum measurement problem and the question of the quantum-classical connection. From among the various explanations that seek a resolution to the conceptual problems of quantum mechanics, we focus on the 'environment induced decoherence theory' - an approach that employs the methods developed by several authors to analyse the quantum mechanics of

a system in interaction with its environment(Zeh, 1970). The central idea of this approach is that classicality is an emergent property triggered in open systems by their environments and it is the influence of environmental interactions that explains the perceived outcomes of quantum measurements(Zurek, 1981). We illustrate this approach through some specific system-apparatus models and highlight some key results of other researchers and ours. Following this, the next section will address a specific aspect of the decoherence theory, i.e., the notion of a 'preferred basis' or a 'pointer basis'. Our experience of the classical world suggests that unlike quantum systems, which are allowed to exist in all possible states, classical systems only exist in a few select states. The decoherence approach demonstrates that such states are singled out by the environment from a larger quantum menu. These special states are the 'preferred basis', also referred to as the 'pointer states' in a quantum-measurement-like scenario. What is the 'preferred' or 'pointer' basis? This question is examined via specific system-apparatus models and answered through some key results of our work and that of other researchers. Some of these results show that the 'pointer states' could emerge independent of the initial state of the apparatus. In the light of several advances in technology and high precision experiments, many of the questions relating to the conceptual problems of quantum mechanics are no longer merely 'academic' in nature. These questions and theories like environment-induced decoherence now offer themselves to experimental tests. Many recent experiments have provided important insights into the role of the environment in bringing about classicality. The decoherence theory is strengthened by these spectacular observations and there is no doubt that this approach has provided many important insights into the actual mechanism of the loss of quantum coherence. However, many researchers believe that many of the conceptual problems of quantum mechanics are still unresolved and the decoherence explanation is not adequate. In the concluding section of this chapter we summarize the main ideas that are presented in it and also highlight some of the difficulties and unresolved issues and their implications.

## 2. Measurement in quantum mechanics

Though quantum theory is widely accepted as the fundamental theory of nature and is immensely successful and satisfying, its conceptual framework makes predictions which are difficult to comprehend "classically". Many of these conceptual problems are encompassed in what is known as the quantum measurement problem. While the basic formalism of quantum mechanics was developed between 1925 and 1927, the standard interpretation of quantum measurement is attributed to von Neumann's theory presented in his book in 1932 (von Neumann, 1932). The quantum mechanical description of a system is contained in its wave function or state vector  $|\psi\rangle$  which lives in an abstract "Hilbert space". The dynamics of the wavefunction is governed by the Schrödinger equation

$$i\hbar \frac{d}{dt}|\psi\rangle = H|\psi\rangle. \quad (1)$$

Here  $H$  is the Hamiltonian of the system and the equation is linear, deterministic and the time evolution governed by it is unitary. Dynamic variables or *observables* are represented in quantum mechanics by *linear Hermitian operators* which act on the state vector. An operator  $\hat{A}$ , corresponding to the dynamical quantity  $A$  is associated with *eigenvalues*  $a_i$  and corresponding *eigenvectors*  $\{\alpha_i\}$  which form a *complete orthonormal set*. Any arbitrary state vector,  $|\psi\rangle$  can, in general, be represented as a linear superposition of these eigenvectors:

$$|\psi\rangle = \sum c_i |\alpha_i\rangle. \quad (2)$$

A basic postulate of quantum mechanics regarding *measurement* is that any measurement of the quantity  $A$  can only yield *one* of the eigenvalues,  $a_i$ s, but the result is not definite in the sense that different measurements for the quantum state  $|\psi\rangle$  can yield different eigenvalues. However, quantum theory predicts only that the *probability* of obtaining eigenvalue  $a_i$  is  $|c_i|^2$ . An additional postulate of quantum mechanics is that the measurement of an observable  $A$ , which yields one of the eigenvalues  $a_i$  (with probability  $|c_i|^2$ ) culminates with the *reduction* or *collapse* of the state vector  $|\psi\rangle$  to the eigenstate  $|\alpha_i\rangle$ . This means that every term in the linear superposition vanishes, except one. This reduction is a *non unitary process* and hence in complete contrast to the unitary dynamics of quantum mechanics predicted by the Schrödinger equation and this is where the crux of the conceptual difficulties encountered in quantum theory lies. These two stages of quantum measurement are captured in the well-know von Neumann model through two distinct processes - first, where the system and apparatus interact through linear unitary Schrödinger evolution via an appropriate interaction Hamiltonian, and second - the nonlinear, indeterministic collapse (von Neumann, 1932). In this sense, the idea of measurement is very different from what we understand for classical systems. Classical systems are independent from measurements - the act of measurement does not disturb the state of the system or its 'properties'. In the language of quantum mechanical wave functions, the von Neumann measurement scheme can be illustrated as follows:

Measurements are described by treating both the system and the measuring apparatus as quantum objects. Let the quantum system be in the superposition state  $|\psi_S\rangle = \sum_n c_n |\psi_{Sn}\rangle$ , where  $|\psi_{Sn}\rangle$  are the eigenstates of the operator that needs to be measured. For a measurement to be affected, the measured system described by  $|\psi_S\rangle$  needs to interact with the measuring apparatus described by  $|\phi_A\rangle$ , so that the total wave function before the interaction is  $|\psi_S\rangle|\phi_A\rangle$ . During the interaction of the system and the apparatus, the unitary evolution realizes the following transition from the initial to the final total wave function:

$$|\psi_S\rangle|\phi_A\rangle \rightarrow \sum_n c_n |\psi_{Sn}\rangle|\phi_{An}\rangle \quad (\text{measurement of the first kind}). \quad (3)$$

Here  $|\phi_{An}\rangle$  are orthonormal states of the measuring apparatus. This unitary evolution is referred to as *premeasurement*. The transition

$$|\psi_S\rangle \rightarrow \sum_n |c_n|^2 |\psi_{Sn}\rangle\langle\psi_{Sn}| \quad (4)$$

is often referred to as *the wave function collapse*. The final density operator corresponding to the system is calculated as  $\sum_n |c_n|^2 |\psi_{Sn}\rangle\langle\psi_{Sn}|$ . This density operator describes an *ensemble* of system states, which, after the measurement will be found in the state  $|\psi_{Sn}\rangle$  with probability  $|c_n|^2$ . The transition

$$|\psi_S\rangle \rightarrow \sum_n |c_n|^2 |\psi_{Sn}\rangle\langle\psi_{Sn}| \rightarrow |\psi_{Sn}\rangle, \quad (5)$$

corresponds to an additional selection of a subensemble by means of observation. In measurements of the second kind, the unitary evolution during the interaction of the system and measuring apparatus is described as:

$$|\psi_S\rangle|\phi_A\rangle \rightarrow \sum_n c_n |\chi_{Sn}\rangle|\phi_{An}\rangle, \quad (6)$$

in which the states  $|\chi_{Sn}\rangle$  of the system are determined by the nature of the interaction between system and measuring apparatus. As in the case of measurements of the first kind, the

final state of the system will be  $|\chi_{S_n}\rangle$  with probability  $|c_n|^2$ . The concept that quantum mechanics does not yield an objective description of microscopic reality but deals only with probabilities (as illustrated in the measurement process) is an essential part of the Copenhagen interpretation of quantum mechanics which is regarded as the "standard" interpretation of quantum mechanics. The von Neumann measurement scheme is in tune with the Copenhagen interpretation of quantum mechanics which was one of the first attempts to understand quantum mechanics, initiated by Niels Bohr, and supported by Werner Heisenberg, Max Born and others. The von Neumann scheme described above would typically involve a coupling between the microscopic system and a 'macroscopic' apparatus (meter), resulting in states like (3) and (6), called *entangled states* which are uniquely quantum mechanical states for the composite. The term *entanglement* was coined by Schrödinger and describes a correlated state that is "not separable" (Schrödinger, 1935). Today, entanglement is considered one of the most defining concepts in quantum mechanics - a uniquely quantum mechanical possibility with no classical analogue. In quantum information and quantum computation, entanglement is viewed as a resource for computing tasks that can be performed faster or in a more secure way than is classically possible and there are intensive experimental efforts to create entangled states in the laboratory. The entangled state describing the system-apparatus, as in (3) above should contain one-to-one correlations between the states of the system,  $\{|\psi_{S_n}\rangle\}$ , and the states of the apparatus  $\{|\phi_{A_n}\rangle\}$ , so that a read out of the apparatus or 'meter' states gives information about the states of the system. Consider a simple example of a two-level system for which the entangled system-apparatus state after the measurement interaction should look like

$$|\psi_S\rangle|\phi_A\rangle \rightarrow |\psi_{S1}\rangle|\phi_{A1}\rangle + |\psi_{S2}\rangle|\phi_{A2}\rangle. \quad (7)$$

Such an entangled state is like a two-particle *superposition state*. The problem with such an entangled state is that it seems to allow the 'meter' (apparatus) to exist in a coherent superposition of the two states  $|\phi_{A1}\rangle$  and  $|\phi_{A2}\rangle$  which could be macroscopically distinct - a situation hard to reconcile with classical intuition. Historically, it was the Einstein, Podolsky, Rosen (EPR) paper (Einstein et al., 1935) which first highlighted the problem of quantum entanglement. In response to the EPR work, Schrödinger posed a thought experiment which is now famously known as Schrödinger's Cat paradox (Schrödinger, 1935). Schrödinger's cat is the unfortunate victim of a nasty contraption where the decay of a radioactive atom triggers a device which kills the cat. The quantum mechanical description of this scenario demands that a superposition state of 'decayed' and 'not decayed' for the atom lead to an entangled state of the kind (7) for the atom-cat composite with the cat being in a superposition state of 'dead' and 'alive'. This amounts to interpreting the quantum state of the cat as being in a coherent superposition of 'dead' and 'alive' states - a situation which is completely at odds with our familiar classical perceptions. Schrödinger's Cat paradox is often presented as an illustration of the conceptual problems of quantum mechanics. It is worth mentioning at this point that the density matrix is a convenient formal tool to compare and contrast quantum and classical systems in terms of probabilities. Some of the conceptual problems of quantum measurement become more transparent when analyzed in this language. It can be easily seen that the density matrix corresponding to the entangled state (7) is

$$\begin{aligned} \hat{\rho}_{S+A} = & |\psi_{S1}\rangle\langle\psi_{S1}||\psi_{A1}\rangle\langle\psi_{A1}| + |\psi_{S2}\rangle\langle\psi_{S2}||\psi_{A2}\rangle\langle\psi_{A2}| \\ & + |\psi_{S1}\rangle\langle\psi_{S2}||\psi_{A1}\rangle\langle\psi_{A2}| + |\psi_{S2}\rangle\langle\psi_{S1}||\psi_{A2}\rangle\langle\psi_{A1}|. \end{aligned} \quad (8)$$

While (8) represents a perfectly legitimate solution of the Schrödinger equation, the physical interpretation in the usual language of probabilities leads to difficulties. While the diagonal

elements ( the first and second terms) can be easily interpreted as probabilities corresponding to the system being in state  $|\psi_{S1}\rangle$  or  $|\psi_{S2}\rangle$  (with the corresponding correlations with the apparatus states  $|\phi_{A1}\rangle$  and  $|\phi_{A2}\rangle$ , respectively), the off-diagonal elements represented by the third and fourth terms are difficult to interpret classically in terms of probabilities. In order to make 'classical' sense, the density matrix corresponding to the pure state ensemble described by the entangled state (8) must reduce to a *statistical mixture* which is *diagonal* in some basis with appropriate system-apparatus correlations. Such a *mixed* density matrix would look like

$$\rho_{mixed} \sim |\psi_{S1}\rangle\langle\psi_{S1}||\psi_{A1}\rangle\langle\psi_{A1}| + |\psi_{S2}\rangle\langle\psi_{S2}||\psi_{A2}\rangle\langle\psi_{A2}|. \quad (9)$$

Several interpretations of quantum mechanics seek to explain this  $\rho_{pure} \rightarrow \rho_{mixed}$  transition (von Neumann's irreversible 'reduction' process) and a resolution to the mechanism for the apparently nonunitary 'collapse' in a quantum measurement (Wheeler & Zurek, 1983; Zurek, 1991). In recent years, the decoherence approach (Joos et al., 2003) has been widely discussed and accepted as the mechanism responsible for this transition. The central idea of this approach has been that 'classicality' is an emergent property of systems interacting with an environment. The theory also predicts that in a quantum measurement, the apparatus will have correlations with the system in a set of 'preferred states' (Joos et al., 2003; Zurek, 1981; 1991) selected by the environment. In the next two sections we describe the progress made in adopting this approach to explain the mechanism for the perceived outcomes of a quantum measurement as well as the emergence of classicality from an underlying quantum world. The strength of this approach lies in the fact that it provides a reasonably satisfying explanation *within* the realm of quantum mechanics.

### 3. Decoherence

In the previous section, we have seen that there is a serious interpretational problem with the way quantum mechanics deals with the act of measurement. In particular, the problem lies in von Neumann's postulate of an irreversible reduction process which takes the quantum superposition to a statistical mixture which is supposedly classically interpretable and meaningful. However, the non-unitary nature of this reduction is at odds with the inherent unitary nature of the Schrodinger equation, implying, somehow that the mechanism seems to lie *outside* the realm of quantum mechanics. From among the various explanations that seek a resolution to the conceptual problems of quantum mechanics, in this section we focus on the 'environment induced decoherence theory'. As pointed out in the previous section, the problem lies with the *off-diagonal* elements of the density matrix describing the entangled state of the system-apparatus composite, (8). These off-diagonal elements are the signatures of *quantum correlations*. In quantum mechanics, wave functions evolve according to the Schrödinger equation which is linear and deterministic and this evolution is unitary in nature. Unitary evolutions ensure that eigenvalues are preserved. There is no way that some terms of the density matrix can vanish in the course of a unitary evolution. How, then, can 'classical behaviour' (as discussed above) ever emerge from this substrate of the quantum world where entanglements and coherences are ubiquitous and inevitable? How can a pure state density matrix become a "classically interpretable" statistical mixture? The decoherence approach seeks to answer this problem by providing a mechanism which leads to the loss of quantum coherence, thus bringing about the much desired  $\rho_{pure} \rightarrow \rho_{mixed}$  transition, and hence, classicality. The answer lies in realizing the fact that the Schrödinger equation driving unitary evolutions is strictly applicable only to *completely isolated systems*. In reality, we know that macroscopic systems are *almost never isolated* from their surroundings but are known to

be constantly interacting with a complex environment. In a measurement like situation, the apparatus is almost always a macroscopic object from which one reads out the measured property of the system. In fact, the apparatus is not only considered macroscopic, it is also regarded as 'classical' in its dynamics. How does the apparatus (which starts off as a quantum object) end up appearing 'classical'?

The 'decoherence' explanation is that it is the influence of the environment that makes a quantum system appear 'classical'. The environment 'washes away' quantum coherence ('decoherence'), leaving behind a system which looks and behaves like a classical object of our cherished commonsense world. The system no longer constitutes a closed system but an 'open system' which is coupled to a large number degrees of freedom which constitute the environment. However, one is always monitoring only a few degrees of freedom, which are of relevance. Technically, this amounts to 'tracing' over all other degrees of freedom. This tracing over has the effect of causing the transition:

$$\rho_{\text{pure}} \rightarrow \rho_{\text{mixed}}. \quad (10)$$

An illuminating and popular paradigm for understanding decoherence is the phenomenon of Brownian motion which describes the motion of a particle suspended in a liquid. Such a suspended particle, when examined closely, is seen to bounce around in a random, irregular, 'zig-zag' fashion. Einstein showed that this behaviour is exactly what should be expected if the suspended particle is being repeatedly 'kicked' by other unseen smaller particles. The random motion of the suspended particle can be statistically explained by taking into account its interaction with a large number of particles which constitute the reservoir of liquid molecules or the 'environment'. When we see Brownian motion, we are only focussing on the dynamics of the suspended particle and do not monitor each and every particle of the environment. Mathematically, we trace over all the degrees of freedom of the environment and look only at the reduced system -the suspended particle. As a consequence, the tagged particle is found to show a dynamics that contains dissipation (a steady loss of energy or relaxation) and diffusion (the random zig-zag motion). *Quantum Brownian motion* describes a similar situation at the quantum mechanical level (Agarwal, 1971; Caldeira & Leggett, 1983). A simple example by Zurek (Zurek, 1991) illustrates this point (see Figure 1). Here an initial pure state constructed as a coherent superposition of two spatially separated Gaussian wavepackets decoheres into a statistical mixture (diagonal density matrix) when its dynamics incorporates the coupling to a large number of environmental degrees of freedom. While the pure state density matrix of the system (Fig. 1(a)) has both diagonal and off-diagonal elements, it can be seen that after a certain time, impacted by the environmental influence, the off-diagonal elements of the reduced system are diminished to give us a statistical mixture (Fig1(b)).

Let us now look at the quantum measurement situation through this approach. The microscopic system couples to a macroscopic apparatus, which in turn is interacting with a large number of degrees of freedom which constitutes the environment. Schrödinger's equation is applied to the entire closed system of system-apparatus-environment. Hamiltonian evolution drives this closed system from an initial uncoupled state into a gigantic entangled state containing all the degrees of freedom. A tracing over all the environmental degrees of freedom salvages the reduced system-apparatus combine from this mess. After a characteristic time, the apparatus, impacted by the environment, appears classical in its dynamics. Thus, the environment causes a general quantum state to decay into a statistical mixture of "pointer states" which can be understood and interpreted as classical probability distributions. This, in essence, is the approach of the decoherence theory to explain the

emergence of classicality and the perceived outcomes of quantum measurements (Zurek, 1991). In the following subsections we highlight studies done on two measurement models where the outcome of the system-apparatus interaction is explained by the decoherence approach.

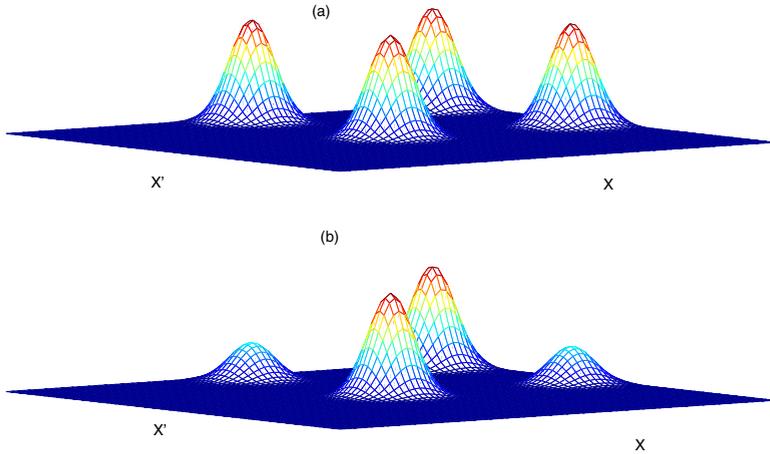


Fig. 1. (a) Density matrix of a superposition of two Gaussian wave packets (b) Density matrix after the off-diagonal elements have been partially washed away by decoherence

### 3.1 Decoherence and the Stern-Gerlach measurement

Venugopalan et al (Venugopalan et al., 1995a,b) first analysed a Stern-Gerlach-like measurement model through the decoherence approach. Their analysis shows that decoherence would bring out the desired 'classical-like' outcome in such a measurement scenario. Consider a simple model of measurement consisting of a free particle with spin (for simplicity, consider a two-state system or spin-1/2 which could represent a qubit). Here the spin degrees of freedom represent the system and the position and momentum degrees represent the apparatus. Let us first look at the bare system without the inclusion of additional environment degrees of freedom. The system and apparatus are coupled by a Stern-Gerlach measurementlike interaction such that the trajectory of the particle (position and momentum degrees) correlates with the spin states. The Hamiltonian describing this model is:

$$H = \lambda\sigma_z + \frac{p^2}{2m} + \epsilon z\sigma_z. \quad (11)$$

While the first two terms represent the self Hamiltonians of the system and apparatus, respectively, the last term is the interaction Hamiltonian.  $z$  and  $p$  denote the position and momentum of the particle of mass  $m$ ,  $\lambda\sigma_z$  the Hamiltonian of the system and  $\epsilon$  the product of the field gradient and the magnetic moment of the particle. The most general initial state for the system-apparatus combine can be written as

$$\psi = \{a|\uparrow\rangle + b|\downarrow\rangle\} \otimes \phi(z). \quad (12)$$

This is a product state of the most general spin state for the system and an arbitrary state,  $\phi(z)$  for the apparatus (free particle).  $|\uparrow\rangle$  and  $|\downarrow\rangle$  are the eigenstates of  $\sigma_z$ . Following the measurement interaction governed by Hamiltonian evolution, this initial state becomes an *entangled state* between the system and apparatus. The density matrix of the entangled state, when the initial state, (12), undergoes a Hamiltonian evolution via (11) will be of the form

$$\begin{aligned} \rho_{S+A} = & |a|^2 |\uparrow\rangle\langle\uparrow| \rho_{\uparrow\uparrow}(z, z', t) + |b|^2 |\downarrow\rangle\langle\downarrow| \rho_{\downarrow\downarrow}(z, z', t) \\ & + ab^* |\uparrow\rangle\langle\downarrow| \rho_{\uparrow\downarrow}(z, z', t) + a^*b |\downarrow\rangle\langle\uparrow| \rho_{\downarrow\uparrow}(z, z', t). \end{aligned} \quad (13)$$

Here  $\rho_{\uparrow\uparrow}$  and  $\rho_{\downarrow\downarrow}$  correspond to the diagonal elements (in spin) of the density matrix ( $\rho_d$ ) for the apparatus which could correlate with up and down spin states of the system and  $\rho_{\uparrow\downarrow}$  and  $\rho_{\downarrow\uparrow}$  correspond to the off-diagonal elements ( $\rho_{od}$ ), and  $\rho(z, z', t) = \langle z|\rho|z'\rangle$ . The specific form of  $\rho_d$  and  $\rho_{od}$  and the system correlations they would (or would not) contain depends on the initial state of the apparatus. When they contain one-to-one system-apparatus correlations, the states corresponding to  $\rho_d$  would be candidate *pointer states*. For the purpose of illustration we look at the situation when the apparatus starts off in an initial state which is a Gaussian wave packet as was first analyzed by Venugopalan et al. The initial system-apparatus state is thus given as

$$\{a|\uparrow\rangle + b|\downarrow\rangle\} \otimes \phi(z) = \{a|\uparrow\rangle + b|\downarrow\rangle\} \otimes \frac{1}{\sqrt{\sigma\sqrt{\pi}}} \exp\left\{\frac{-z^2}{2\sigma^2}\right\}. \quad (14)$$

Here  $\sigma$  is the width of the wavepacket. The wavepacket in quantum mechanics is often viewed as the most "nearly classical" state and is known to exhibit many striking classical properties and hence is a reasonably good choice for the initial state of the apparatus. Following Hamiltonian evolution via (11), the system-apparatus composite ends up in an entangled state whose density matrix is of the form (13). One examining the detailed form of the complete density matrix representing this entangled state, one can identify the parts of the "apparatus", i.e.,  $\rho_{\uparrow\uparrow}$  and  $\rho_{\downarrow\downarrow}$  that correlate with the up and down spins, and it can be shown that the diagonal elements of  $\rho_{\uparrow\uparrow}$  and  $\rho_{\downarrow\downarrow}$  are

$$\rho_{\uparrow\uparrow}, \rho_{\downarrow\downarrow} \rightarrow \rho_d(z, t) = \frac{2}{\sigma} \sqrt{\frac{\pi}{N(t)}} \exp\left\{-\frac{4}{\sigma\sqrt{N(t)}} \left(z \mp \frac{\epsilon t^2}{2m}\right)^2\right\}, \quad (15)$$

in the position representation and

$$\rho_{\uparrow\uparrow}, \rho_{\downarrow\downarrow} \rightarrow \rho_d(p, t) = 2\sigma\sqrt{\pi} \exp\left\{-4\sigma^2 \left(p \mp \frac{\epsilon t}{\hbar}\right)^2\right\}, \quad (16)$$

in the momentum representation. It must be kept in mind, though, that the density matrix represents a pure state which has 'nonlocal' quantum correlations both in the spin space and the position and momentum space. In (15) & (16) above, we are looking at the *diagonal* elements of the position and momentum space density matrix and these show system-apparatus correlations. The up and down spin states of the system correlate with a Gaussian wavepackets centered round  $-\frac{\epsilon t^2}{2m}$  and  $+\frac{\epsilon t^2}{2m}$ , in the position space, respectively, and around  $-\frac{\epsilon t}{\hbar}$  and  $+\frac{\epsilon t}{\hbar}$  in the momentum space. This pure entangled state of the system and the apparatus is akin to a 'Schroödinger cat state' which contains one-to-one correlations between the system and 'macroscopic' apparatus states with all quantum coherences intact. Since the dynamics, as governed by the Hamiltonian (11), is purely unitary, there is no

dissipation/decoherence involved, and the state remains a pure entangled state with all its non-diagonal elements ( $\rho_{\uparrow\downarrow}, \rho_{\downarrow\uparrow}$ ) as well as the spatial nonlocality. The establishment of system-apparatus correlations, therefore, *is not enough to affect a measurement as the off diagonal elements of the density matrix have not vanished* and we do not have the desired *mixed state* density matrix.

Now let us look at the situation where we include the interaction with the environment. A commonly used model to describe the environment is to consider it as a reservoir of quantum oscillators, each of which interacts with the apparatus in our case and is described by the Hamiltonian

$$H = \lambda\sigma_z + \frac{p^2}{2m} + \epsilon z\sigma_z + \sum_j \frac{P_j^2}{2M_j} + \frac{M_j^2\Omega_j^2}{2} \left( Z_j - \frac{C_j z}{M_j\Omega_j^2} \right)^2. \quad (17)$$

Here  $z$  and  $p$  denote the position and momentum of the particle (apparatus) of mass  $m$ .  $\lambda\sigma_z$  is the Hamiltonian of the system and  $\epsilon$  is the strength of the system-apparatus coupling as mentioned earlier. The last term represents the Hamiltonian for the bath of oscillators (environment) and the apparatus-environment interaction.  $Z_j$  and  $P_j$  are the position and momentum coordinates of the  $j$ th harmonic oscillator of the bath,  $C_j$ s are the coupling strengths and  $\Omega_j$ s are the frequencies of the oscillators comprising the bath. Note that the coupling of the apparatus with the many environmental degrees of freedom is a coordinate-coordinate coupling. The dynamics for the closed universe of the system, apparatus and environment is governed by the Hamiltonian evolution via (17) and the Schrödinger equation. In the decoherence approach, a tracing over all the degrees of freedom of the environment results in an equation describing the dynamics of the *reduced density matrix* of the system-apparatus combine. The reduced density matrix evolves according to a master equation which is obtained by solving the Schrödinger equation for the entire universe of the system, apparatus and the environment and then tracing over the environment degrees of freedom. Several authors like Zeh and Zurek, among others, have worked extensively on the decoherence approach using the master equation for the reduced density matrix. The master equation for this kind of model of the environment was first derived separately by Caldeira and Leggett, Agarwal, Dekker (Dekker, 1977) and others in the context of quantum Brownian motion and is a popular equation for the study of open quantum systems. For our purpose, we deal with the master equation for the system-apparatus-environment composite described by (17). The master equation for the density matrix, corresponding to the four elements of spin space ( $\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$ ) (see (13)) is

$$\frac{\partial \rho_{ss'}(z, z', t)}{\partial t} = \left[ -\frac{\hbar}{2im} \left( \frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial z'^2} \right) - \gamma(z - z') \left( \frac{\partial}{\partial z} - \frac{\partial}{\partial z'} \right) - \frac{D}{4\hbar^2} (z - z')^2 + \frac{i\epsilon(zs - z's')}{\hbar} + \frac{i\lambda(s - s')}{\hbar} \right] \rho_{ss'}(z, z', t), \quad (18)$$

where  $s, s' = +1$  (for  $\uparrow$ ) or  $-1$  (for  $\downarrow$ ). Here  $\gamma$  is the Langevin friction coefficient and  $D$  has the usual interpretation of the diffusion coefficient.  $\gamma$  and  $D$  are related to the parameters of the Hamiltonian of the total system. Without going into further details of the master equation, it suffices to point out that the equation naturally separates into three distinct terms, namely (i) a term describing the von Neumann equation which can be derived from the Schrödinger equation and thus represents the pure quantum evolution, (ii) a term that causes dissipation, which can be understood as a steady loss in energy or a relaxation process, and (iii) a term causing diffusion, which can be understood as the fluctuations or zig-zag

movement seen in Brownian motion. The initial work of Zeh, Zurek and subsequent work where the decoherence approach has been applied to typical quantum-measurement-like situations like the Stern Gerlach experiment have shown that in the dynamics governed by the master equation, coherent quantum superpositions persist for a very short time. They are rapidly destroyed by the action of the environment. In fact, results show that the larger the quantum superposition, the faster the decoherence. For truly macroscopic superpositions such as Schrödinger's cat, decoherence occurs on such a short time scale that it is impossible to observe these quantum coherences. As mentioned above, the simple example by Zurek (Fig.1) illustrates this point. Several other authors have studied decoherence in other systems and their calculations are seen to contain two main features which can be seen as signatures of decoherence: (a) the decoherence time, over which the superpositions decay is much much shorter than any characteristic time scale of the system (the thermal relaxation time,  $\gamma^{-1}$ , and (b) the decoherence time varies versely as the square of a quantity that indicates the 'size' of the quantum superposition (e.g.  $\frac{\epsilon t^2}{m}$  in (15) is the "separation" in position space of the two Gaussians correlating with up and down spin states).

Venugopalan et al(Venugopalan et al., 1995a;b) have shown that if the evolution of the initial state (14) is via the mater equation, (18), there is a one-to-one correlation established between the system states (spin) and the apparatus states (position and momentum degrees of freedom). Further, the coupling with the environmental degrees of freedom causes decoherence of the pure density matrix of the entangled state into a statistical mixture. For more details of the exact solutions, the interested reader may see(Kumar, 1998; Venugopalan et al., 1995a;b; Venugopalan, 1997). The density matrix for an initial system-apparatus state described by (14), evolving via the master equation (18) has the form

$$\begin{aligned} \rho_{S+A} = & |a|^2 |\uparrow\rangle\langle\uparrow| \rho_{\uparrow\uparrow}(z, z', t) + |b|^2 |\downarrow\rangle\langle\downarrow| \rho_{\downarrow\downarrow}(z, z', t) \\ & + \left[ ab^* |\uparrow\rangle\langle\downarrow| \rho_{\uparrow\downarrow}(z, z', t) + a^* b |\downarrow\rangle\langle\uparrow| \rho_{\downarrow\uparrow}(z, z', t) \right] e^{-\alpha t^3}. \end{aligned} \quad (19)$$

Here the off-diagonal elements of the density matrix (last two terms) contain a multiplicative factor of the form  $e^{-\alpha t^3}$  which causes the decay of these terms to zero over a characteristic time making the density matrix diagonal in spin space. Thus, we are left with a mixed state density matrix which can be interpreted in terms of classical probabilities. It also turns out that the spatial nonlocality of the diagonal components  $\rho_{\uparrow\uparrow}$  and  $\rho_{\downarrow\downarrow}$  disappear over time and these are rendered completely diagonal in momentum space. The density matrix, thus ends up being completely diagonal with perfect correlations between the spin component value and the average momentum of the particle(Kumar, 1998; Venugopalan et al., 1995a). Thus, the inclusion of environmental interaction has destroyed the quantum corelations (signified by the off-diagonal elements of the density matrix) and rendered the reduced system-apparatus combine into a statistical mixture. This, in effect, explains the mechanism leading to the "collapse" in a quantum measurement, as well as the description of the way classicality emerges from an underlying quantum substrate.

### 3.2 Decoherence and the harmonic oscillator apparatus

In the previous section we have seen a measurement-like scenario where the apparatus was represented by the continuous position and momentum degrees of freedom of the particle who's spin was measured. Venugopalan(Venugopalan, 2000) has looked at the same problem where the apparatus is a harmonic osillator. The interaction of a quantum system (spin-1/2)

with a macroscopic quantum apparatus (harmonic oscillator) coupled to a bath of harmonic oscillators was analysed. The Hamiltonian for this system would be

$$H = \lambda\sigma_z + \frac{1}{2}m\omega^2 z^2 + \epsilon z\sigma_z + \sum_j \frac{P_j^2}{2M_j} + \frac{M_j^2 \Omega_j^2}{2} \left( Z_j - \frac{C_j z}{M_j \Omega_j^2} \right)^2. \quad (20)$$

Here  $\omega$  is the frequency of the oscillator and all the other symbols have the same meaning as explained in the previous section. The master equation for this system can be written exactly in the same way as for the free particle in the previous section. The initial system-apparatus state is considered to be a product state of any arbitrary state of the apparatus and a general superposition state of the spin-1/2 state of the system, of the form

$$\psi = \{a|\uparrow\rangle + b|\downarrow\rangle\} \otimes \phi(z). \quad (21)$$

Exact solutions of the master equation show that the reduced density matrix of the system-apparatus combine decoheres to a statistical mixture where up and down spins eventually correlate with pointer states of the apparatus. The strength of this analysis is that unlike in the previous section where the initial state of the apparatus was considered to be a Gaussian wavepacket, no particular initial state of the harmonic oscillator was chosen. Venugopalan shows that for the zero temperature bath the system-apparatus combine ends up with spin-apparatus correlations in the *coherent states of the harmonic oscillator* for arbitrary initial states of the apparatus. For a high temperature bath, pointer states are Gaussian distributions or *generalized coherent states* (Venugopalan, 2000). For both cases, the off-diagonal elements in spin-space decohere over a time scale which goes inversely as the square of the "separation" between the "pointers". The "statistical mixture" into which the density matrix decoheres looks like

$$\rho_{S+A} \sim |a|^2 |\uparrow\rangle\langle\uparrow| \rho_{\uparrow\uparrow}(z, z', t) + |b|^2 |\downarrow\rangle\langle\downarrow| \rho_{\downarrow\downarrow}(z, z', t) \quad (22)$$

where the apparatus states are

$$\rho_{\uparrow\uparrow} = \sqrt{\frac{m\omega}{\pi\hbar}} \exp\left\{ -\frac{m\omega}{\hbar} \left( R + \frac{\epsilon}{m\omega^2} \right)^2 - \frac{m\omega r^2}{4\hbar} \right\}, \quad (23)$$

$$\rho_{\downarrow\downarrow} = \sqrt{\frac{m\omega}{\pi\hbar}} \exp\left\{ -\frac{m\omega}{\hbar} \left( R - \frac{\epsilon}{m\omega^2} \right)^2 - \frac{m\omega r^2}{4\hbar} \right\}. \quad (24)$$

with  $R = (z + z')/2$ , and  $r = (z - z')$ . This is nothing but the density matrix corresponding to a *coherent state*,  $|\alpha\rangle$ , of a harmonic oscillator with zero mean momentum and mean positions  $= \pm\epsilon/m\omega^2$ , and

$$|\alpha|^2 = \frac{m\omega}{2\hbar} \left( \frac{\epsilon}{m\omega^2} \right). \quad (25)$$

The above result is for a zero temperature bath. For a high temperature bath, *generalized coherent states* constitute the pointer states. Exact results also demonstrate in an unambiguous way that the pointer states in this measurement model emerge independent of the initial state of the apparatus. For details of the analysis and exact solutions, the reader is referred to (Venugopalan, 2000). For both the zero temperature and high temperature cases analysed, the exact solutions for this model demonstrate the two main signatures of the decoherence mechanism in a quantum measurement, namely, (a) the decoherence time is much smaller than the thermal relaxation time, and (b) the decoherence time is inversely proportional to the

square of the "separation" between the two "pointers" that correlate with the system states. Thus, one again, one can see that the inclusion of the environmental degrees of freedom takes a pure entangled system-apparatus state to a statistical mixture which lets us interpret the measurement in terms of classical probabilities as well as lets us explain the emergence of classical correlations from an underlying quantum substrate.

#### 4. Decoherence and the pointer basis

Finally, we focus on a specific aspect of the decoherence theory, i.e., the notion of a 'preferred basis' or a 'pointer basis'. Our experience of the classical world suggests that unlike quantum systems, which are allowed to exist in all possible states, classical systems only exist in a few select states which are singled out by the environment from a larger quantum menu (Joos et al., 2003; Zurek, 1981). These special states are the *preferred basis*, also referred to as the "pointer states" in a quantum-measurement-like scenario. In a measurement-like scenario the pointer basis should also be understood as those states of the apparatus in which correlations with the system states are eventually established (irrespective of the initial states of the apparatus). In spite of many theoretical studies on the decoherence approach, not many systems have been analysed with the aim of predicting what the pointer basis should be in a given situation. Let us look at the issue of the pointer basis a little more closely by once again referring to the two quantum measurement examples discussed in the previous sections: (i) the spin-1/2 system with the free particle as an apparatus in the Stern-Gerlach-like interaction described in section 3.1 (ii) the spin-1/2 system with the harmonic oscillator as an apparatus described in section 3.2.

What is the preferred basis in a given scenario, and what decides in which pointer basis the system-apparatus will be finally established? For simplified models where the self Hamiltonian of the system has either been ignored or considered co-diagonal with the interaction Hamiltonian, the pointer variable has been shown to be the one which commutes with the interaction Hamiltonian (Zurek, 1981). However, in more general situations where all terms are included, as in the two examples discussed here, the various parts of the Hamiltonian may not commute. In such situations it is not obvious what decides the preferred basis. For example, in the case of the spin system and free particle apparatus with Stern-Gerlach like interaction discussed above, the coordinate-coordinate coupling indicates that the position basis is intuitively expected to emerge as the preferred basis. However, as we have seen, this is contrary to the conclusion of Venugopalan et al (Kumar, 1998; Venugopalan, 1994; Venugopalan et al., 1995a) in their analysis of the Stern-Gerlach measurement model where the spin components eventually correlate with distributions which are completely diagonal in the momentum basis and only approximately diagonal in the position basis (Kumar, 1998).

In the literature, the preferred basis has been variously described as the one in which the final state density matrix becomes diagonal or that set of basis states which are characterized by maximum stability or a minimum increase in linear or statistical entropy, decided by a *predictability sieve* (Zurek et al., 1993). Using the Markovian Master equation for a harmonic oscillator coupled to a heat bath and the criterion of the predictability sieve Zurek argues that coherent states emerge as the preferred basis. This is in tune with the results for the harmonic oscillator apparatus model where we have seen that pointer states end up being coherent states or generalized coherent states, *irrespective* of the initial state in which the apparatus starts off. This result also agrees with a study by Tegmark and Shapiro (Tegmark & Shapiro, 1999) where they show that generalized coherent states tend to be produced naturally when

one looks at the reduced Wigner distributions of infinite systems of coupled oscillators as  $t \rightarrow \infty$ . Paz and Zurek (Paz & Zurek, 1999) have investigated decoherence in the limit of weak interaction with the environment and show that the eigenstates of energy emerge as pointer states. Roy and Venugopalan have also obtained the exact solutions of the Master equation for a harmonic oscillator and a free particle in a compact factorizable form and have shown that the density matrix diagonalizes in the energy basis which is number states for the oscillator and momentum states for the free particle for arbitrary initial conditions (Roy & Venugopalan, 1999). It is intuitive that the pointer states should naturally be a consequence of the interplay between the various components of the total Hamiltonian and one should also expect them to be independent of the initial state of the system/apparatus.

The studies mentioned above do shed light on the nature of the preferred basis but are inadequate and there is a need to analyze more systems. In particular, it is important to look at systems like the harmonic oscillator apparatus model which is fairly generic and exact solutions make it an interesting candidate to explore experimentally in the context of decoherence and quantum measurements. Also, this example indicates that it seems pertinent to look at a system-apparatus-environment like scenario for measurement to analyse the issue of the pointer basis and the states singled out by the environment.

Finally, we would like to refer to a recent work of Venugopalan (Venugopalan, 2011) which indicates that there is a need to look at the bare system-apparatus interaction in a measurement like scenario more carefully to get insights into what states would eventually end up as pointer states and be selected in the event of environmental influence. Venugopalan has looked at a simple one dimensional model for the system-apparatus interaction where the system is a spin-1/2 particle, and its position and momentum degrees constitutes the apparatus, like the Stern-Gerlach model discussed above. An analysis involving only unitary Schrödinger dynamics illustrates the nature of the correlations established in the system-apparatus entangled state. It is shown that even in the absence of any environment-induced decoherence, or any other measurement model, certain initial states of the apparatus - like localized Gaussian wavepackets - are preferred over others, in terms of the establishment of measurement-like one-to-one correlations in the pure system-apparatus entangled state. This result indicates that perhaps there already exist special states of the apparatus within the quantum menu, and it is these that end up being ultimately selected by the environment as the preferred states.

## 5. Conclusions

The central theme of the decoherence approach has been to explain, within the realm of quantum theory, the appearance of classicality in the macroscopic, familiar physical world. The strength of the theory has been often claimed to be the fact that it is within the realm of quantum mechanics, and uses the rules of quantum theory to explain the emergence of classicality. In the case of quantum measurement, decoherence is believed to "mimic" wave function collapse. This is achieved via the transition from a pure state density matrix to a statistical mixture with 'classically' meaningful terms. In a sense, decoherence explains the washing away of quantum coherences and the emergence of a state which makes classical sense. In this chapter we have seen this happen in the two measurement models considered above - one with free particle as an apparatus and one with a harmonic oscillator as the apparatus. We have seen that the environmental influence is crucial in not only destroying the quantum coherences, but also is selecting a special state or a preferred basis. A criticism levelled against the decoherence approach is often that it does not explain the fact that *only*

one of the mixed outcomes is actually observed, nor does it allow us to predict exactly which one will be observed. In spite of these criticisms, it is generally accepted that the decoherence explanation has certainly provided valuable insights into the actual mechanism of the loss of quantum coherences. In recent years, the predictions of the decoherence theory have been tested in several spectacular experiments which put the theory on a firm footing. Of these, two experiments are particularly noteworthy as they have succeeded in monitoring the decoherence mechanism, i.e., the actual transition from a pure entangled state to a statistical mixture. Moreover, the experiments also give a quantitative estimate of the decoherence time. These experiments have done no less than create Schrödinger-cat-like entangled states in the laboratory and seen them transform into classically recognizable objects under the influence of environmental coupling. Among the first successful attempts is an experiment by Brune et al. (Brune et al., 1996) at the Ecole Normale Supérieure in Paris. Using Rubidium atoms and high technology superconducting microwave cavities, Brune et al. created a superposition of quantum states involving radiation fields. The superposition was the equivalent of a "system+measuring apparatus" situation in which the "meter" was pointing simultaneously towards two different directions. This is a Schrödinger-Cat-like entangled state. Through a series of ingenious "atom-interferometry" experiments, Brune et al. managed to not only "read" this pure state but also monitored the decoherence phenomenon as it unfolded, transforming this superposition state to a statistical mixture. Besides providing a direct insight into the role of the environment in a quantum measurement process, their experiment also confirmed the basic tenets of the decoherence theory. The two main signatures of the decoherence theory were clearly observed in this classic experiment. The environment in this experiment are the "modes" of the electromagnetic field in the cavity. At the National Institute of Standards and Technology, Boulder, Colorado, the group headed by Wineland (Monroe et al., 1996) created a Schrödinger-Cat-like state using a series of laser pulses to entangle the internal (electronic) and the external (motional) states of a Beryllium ion in a "Paul trap". The motion of this trapped ion couples to an electric field which changes randomly, thus simulating an environment. Monroe et al. call this environment an engineered reservoir whose state and coupling can be controlled. Through their measurements, Wineland et al. have successfully demonstrated the two important signatures of the decoherence mechanism. Thus one can see that the qualitative and quantitative predictions of the decoherence theory are experimentally tested. These tests would be highly relevant to all experimental implementations of the novel ideas of quantum information and computation as decoherence would ruin the functioning of devices which use uniquely quantum mechanical effects for information processing. The above two experiments, along with others have provided important insights into the role of the environment in bringing about classicality and thus decoherence theory is strengthened by these spectacular observations. What, then, is the final verdict of the decoherence theory? Has it resolved the conceptual problems of quantum mechanics? There are many who believe that the conceptual problems of quantum mechanics are still unresolved and decoherence does not answer many issues. At the end of the day we can say that the decoherence explanation takes away some of the mystery from the idea of 'wave function collapse' and provides a conventional mechanism to explain the appearance of a classical world. Many physicists find this a satisfactory explanation and there is no doubt that the experiments discussed clearly show how decoherence washes away quantum coherences providing a fairly convincing evidence for explaining the absence of Schrödinger's Cats in the real world. For all practical purposes, the decoherence explanation finds favour as a satisfactory settlement of the quantum measurement problem.

## 6. References

- Agarwal, G. S., Brownian Motion of a Quantum Oscillator, *Phys. Rev. A* 4, 739-747 (1971)
- Brune, M., Hagley, E., Dreyer, J., Maître, X., Maali, A., Wunderlich, C. , Raimond, J. M. & Haroche, S., Observing the Progressive Decoherence of the "Meter" in a Quantum Measurement, *Phys. Rev. Lett.* 77, 4887-4890 (1996)
- Caldeira, A. O. & Leggett, A. J., Path integral approach to quantum Brownian motion, *Physica A: Statistical and Theoretical Physics* Volume 121, Issue 3, pp 587-616 (1983)
- Dekker, H., Quantization of the linearly damped harmonic oscillator, *Phys. Rev. A* 16, pp 2126-2134,(1977)
- Dirac, P. A. M. *The Principles of Quantum Mechanics*, p 36 (Oxford University Press, Oxford), (1947).
- Einstein, A., Podolsky, B. & Rosen, N., Can Quantum-Mechanical Description of Physical Reality be Considered Complete?. *Physical Review* 47 (10) 777-780, (1935).
- Joos, E., Zeh, H. D., Kiefer, C. , Giulini, D., Kupsch, J. and Stamatescu, I.-O.: *Decoherence and the Appearance of a Classical World in Quantum Theory*, second edition, (Springer 2003).
- Kumar, D., Measurement Problem in Quantum mechanics, *Pramana-Journal of Physics*, 51 (5), pp 567-575 (1998)
- Monroe, C., Meekhof, D. M., King & B. E., Wineland, D. J., A Schrödinger cat superposition state of an atom *Science* 272, 1131 (1996).
- Paz, J. P & Zurek, W. H., Quantum limit of decoherence: Environment induced superselection of energy eigenstates, *Phys. Rev. Lett.* 82, 5181-5185 (1999)
- Peres, A., When is a quantum measurement?, *American Journal of Physics*, Volume 54, Issue 8, pp. 688-692 (1986).
- Roy, S. M. & Venugopalan, A., Exact Solutions of the Caldeira-Leggett Master Equation: A Factorization Theorem For Decoherence *arXiv:quant-ph/9910004v1* (1999)
- Schödinger, E. , Discussion of Probability Relations Between Separated Systems, *Proceedings of the Cambridge Philosophical Society*, 31, 555-563 (1935).
- Schödinger, E., Die gegenwärtige Situation in der Quantenmechanik *Naturwissenschaften* 23, 807-812 (1935); 823-828; 844-849 (1935).
- Tegmark, M. & Shapiro, H. S., Decoherence produces coherent states: An explicit proof for harmonic chains *Phys. Rev. E* 50, 2538-2547 (1994)
- Venugopalan, A., Preferred basis in a measurement process, *Physical Review A*, 2742 (1994)
- Venugopalan, A., Kumar, D. & Ghosh, R. Analysis of the Stern Gerlach measurement *Physica A* 220, p568 (1995);
- Venugopalan, A.,Kumar, D, & Ghosh, R., The Quantum measurement process: Nature of the Apparatus, *Current Science* 68 (1), 62 (1995)
- Venugopalan, A., Decoherence and Schrodinger Cat states in a Quantum measurement, *Physical Review A* 56 (5), 4307 (1997)
- Venugopalan, A., Pointer States Via measurement in a Quantum measurement, *Physical Review A* 61, 012102 (2000)
- Venugopalan, A., The Coming of a Classical World *Resonance: Journal of Science Education*, Vol 9, 10, September (Springer)(2004)
- Venugopalan, A., Preferred states of the apparatus, *Pramana-Journal of Physics* (Springer) (in press) (2011)
- von Neumann, J. *Mathematische Grundlagen der Quanten Mechanik* (Springer-Verlag, Berlin, 1932)(English translation : *Mathematical Foundations of Quantum Mechanics*. Robert T. Beyer. Princeton University Press, Princeton (1955))

- Wheeler, J. A. and Zurek, W. H., eds. *Quantum Theory and Measurement*, Princeton University Press (1983)
- Zeh, H. D., On the interpretation of measurement in quantum theory, *Foundations of Physics*, Volume 1, Issue 1, pp.69-76, (1970)
- Zurek, W. H., Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? *Phys. Rev. D* 24, 1516 - 1525 (1981)
- Zurek, W.H., Decoherence and the Transition from the Quantum to the Classical *Physics Today*, 36-44 (October 1991).
- Zurek, W. H., Habib,S. & Paz, J. P. Coherent states via decoherence, *Phys. Rev. Lett.*70, 1187-1190 (1993).

# Time as Quantum Observable, Canonical Conjugated to Energy

Vladislav S. Olkhovsky<sup>1</sup>, Erasmo Recami<sup>2</sup> and Sergei P. Maydanyuk<sup>3</sup>

<sup>1</sup>*Institute for Nuclear Research, National Academy of Sciences*

<sup>2</sup>*Facoltà di Ingegneria, Università statale di Bergamo, Bergamo  
and INFN-Sezione di Milano, Milan*

<sup>3</sup>*Institute for Nuclear Research, National Academy of Sciences*  
<sup>1,3</sup>*Ukraine*  
<sup>2</sup>*Italy*

## 1. Introduction

Time, as well as 3-position, sometimes is a parameter, but sometimes is an observable that in quantum theory would be expected to be associated with an operator. However, almost from the birth of quantum mechanics (cf., e.g., Ref.(Pauli, 1926; 1980)), it is known that time cannot be represented by a selfadjoint operator, except in the case of special systems (such as an electrically charged particle in an infinite uniform electric field)<sup>1</sup>. The list of papers devoted to the problem of time in quantum mechanics is extremely large (see, for instance, Refs. (Aharonov et al., 1998; Atmanspacher & Amann, 1998; Blanchard P & Jadczyk, 1996; Busch et al., 1994; Delgado, 1999; Egusquiza & Muga, 1999; Giannitrapani, 1997; Gózdź A & Dębicki, 2007; Grot et al., 1996; Holevo, 1978; 1982; Kijowski, 1997; Kobe et al., 1994; Kochański & Wo'dkievicz, 1999; Leo'n, 1997; Muga et al., 1999; Olkhovsky & Recami, 1968; 1969; 1970; Olkhovsky, 1973; Olkhovsky et al., 1974; Olkhovsky, 1984; 1990; 1992; Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky, 1998; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007; Olkhovsky, 2009; 2011; Recami, 1976; 1977; Srinivas & Vijayalakshmi, 1981; Toller, 1999; Wang & Xiong, 2007), and references therein). The same situation had to be faced also in quantum electrodynamics and, more in general, in relativistic quantum field theory (see, for instance, Refs.(Olkhovsky & Recami, 1968; 1969; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007)).

As to quantum mechanics, the very first relevant articles are probably Refs. (Holevo, 1978; 1982; Olkhovsky & Recami, 1968; 1969; 1970; Olkhovsky, 1973; Olkhovsky et al., 1974; Olkhovsky, 1984; 1990; 1992; 1998; Recami, 1976; 1977), and refs. therein. A second set of papers on time in quantum physics (Aharonov et al., 1998; Atmanspacher & Amann, 1998; Blanchard P & Jadczyk, 1996; Busch et al., 1994; Delgado, 1999; Egusquiza & Muga, 1999; Giannitrapani, 1997; Gózdź A & Dębicki, 2007; Grot et al., 1996; Kijowski, 1997; Kobe et al.,

---

<sup>1</sup> This is a consequence of the semi-boundedness of the continuous energy spectra from below (usually from zero). Only for an electrically charged particle in an infinite uniform electric field, and other very rare special systems, the continuous energy spectrum is not bounded and extends over the whole axis from  $-\infty$  to  $+\infty$ . It is curious that for systems with continuous energy spectra bounded from above and from below, the time operator is however selfadjoint and yields a discrete time spectrum.

1994; Kochański & Woździejewicz, 1999; Leoń, 1997; Muga et al., 1999; Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007; Srinivas & Vijayalakshmi, 1981; Toller, 1999; Wang & Xiong, 2007) appeared in the nineties, stimulated partially by the need of a consistent definition for the tunneling time. It is noticeable, and let us stress it right now, that this second set of papers seems however to have ignored Naimark's theorem (Naimark, 1940), which had previously constituted (directly or indirectly) an important basis for the results in Refs. (Holevo, 1978; 1982; Olkhovsky & Recami, 1968; 1969; 1970; Olkhovsky, 1973; Olkhovsky et al., 1974; Olkhovsky, 1984; 1990; 1992; 1998; Recami, 1976; 1977), moreover, all the papers (Aharonov et al., 1998; Atmanspacher & Amann, 1998; Blanchard P & Jadczyk, 1996; Busch et al., 1994; Grot et al., 1996; Kobe et al., 1994; Leoń, 1997; Srinivas & Vijayalakshmi, 1981) attempted at solving the problem of time as a quantum observable by means of formal mathematical operations performed *outside* the usual Hilbert space of conventional quantum mechanics. Let us recall that Naimark's theorem states (Naimark, 1940) that the non-orthogonal spectral decomposition of a hermitian operator *can be approximated* by an orthogonal spectral function (which corresponds to a selfadjoint operator), in a weak convergence, *with any desired accuracy*.

The main goal of the first part of the present paper is to justify the use of time as a quantum observable, basing ourselves on the properties of the hermitian (or, rather, maximal hermitian) operators for the case of continuous energy spectra: cf., e.g., the Refs. (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; 2004; Olkhovsky & Recami, 2007)).

The question of time as a quantum-theoretical observable is conceptually connected with the much more general problem of the four-position operator and of the canonically conjugate four-momentum operator, both endowed with an hermitian and an anti-hermitian part, for relativistic spin-zero particles: This problem is analyzed in the second part of the present paper.

In the third part of this work, it is shown how non-hermitian operators can be meaningfully and extensively used, for instance, for describing *unstable states* (decaying resonances). Brief mentions are added of the cases of quantum dissipation, and of the nuclear optical potential.

## **2. Time operator in non-relativistic quantum mechanics and in quantum electrodynamics**

### **2.1 On *Time* as an observable in non-relativistic quantum mechanics for systems with continuous energy spectra**

The last part of the above-mentioned list (Aharonov et al., 1998; Atmanspacher & Amann, 1998; Blanchard P & Jadczyk, 1996; Busch et al., 1994; Delgado, 1999; Egusquiza & Muga, 1999; Giannitrapani, 1997; Gózdź A & Dębicki, 2007; Grot et al., 1996; Kijowski, 1997; Kobe et al., 1994; Kochański & Woździejewicz, 1999; Leoń, 1997; Muga et al., 1999; Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007; Toller, 1999; Wang & Xiong, 2007), of papers, in particular Refs. (Aharonov et al., 1998; Atmanspacher & Amann, 1998; Blanchard P & Jadczyk, 1996; Delgado, 1999; Egusquiza & Muga, 1999; Giannitrapani, 1997; Gózdź A & Dębicki, 2007; Grot et al., 1996; Kijowski, 1997; Kobe et al., 1994; Kochański & Woździejewicz, 1999; Leoń, 1997; Muga et al., 1999; Toller, 1999; Wang & Xiong, 2007), appeared in the nineties, devoted to the problem of Time in non-relativistic quantum mechanics, essentially because of the need to define the tunnelling

time. As remarked, those papers did not refer to the Naimark theorem<sup>2</sup> (Naimark, 1940) which had mathematically supported, on the contrary, the results in (Holevo, 1978; 1982; Olkhovsky & Recami, 1968; 1969; 1970; Olkhovsky, 1973; Olkhovsky et al., 1974; Olkhovsky, 1984; 1990; 1992; 1998; Recami, 1976; 1977), and afterwards in (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007). Indeed, already in the seventies (in Refs. (Olkhovsky & Recami, 1968; 1969; 1970; Olkhovsky, 1973; Olkhovsky et al., 1974; Recami, 1976; 1977) while more detailed presentations and reviews can be found in (Olkhovsky, 1984; 1990; 1992; 1998) and independently in (Holevo, 1978; 1982)), it was proven that, for systems with continuous energy spectra, Time is a quantum-mechanical observable, canonically conjugate to energy. Namely, it had been shown the time operator

$$\hat{t} = \begin{cases} t, & \text{in the time } (t)\text{-representation, (a)} \\ -i\hbar \frac{\partial}{\partial E}, & \text{in the energy } (E)\text{-representation (b)} \end{cases} \quad (1)$$

to be not selfadjoint, but hermitian, and to act on square-integrable space-time wave packets in the representation (1a), and on their Fourier-transforms in (1b), once point  $E = 0$  is eliminated (i. e., once one deals only with moving packets, excluding any *non-moving* rear tails and the cases with zero fluxes)<sup>3</sup> In Refs.(Olkhovsky, 1984; 1990; 1992; 1998) and (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007), the operator  $\hat{t}$  (in the  $t$ -representation) had the property that any averages over time, in the one-dimensional (1D) scalar case, were to be obtained by use of the following *measure* (or weight):

$$W(t, x) dt = \frac{j(x, t) dt}{\int_{-\infty}^{+\infty} j(x, t) dt}, \quad (2)$$

where the the flux density  $j(x, t)$  corresponds to the (temporal) probability for a particle to pass through point  $x$  during the unit time centered at  $t$ , when traveling in the positive  $x$ -direction. Such a measure is not postulated, but is a direct consequence of the well-known probabilistic *spatial* interpretation of  $\rho(x, t)$  and of the continuity relation  $\partial\rho(x, t)/\partial t + \text{div}j(x, t) = 0$ . Quantity  $\rho(x, t)$  is, as usual, the probability of finding the considered moving particle inside a unit space interval, centered at point  $x$ , at time  $t$ .

Quantities  $\rho(x, t)$  and  $j(x, t)$  are related to the wave function  $\Psi(x, t)$  by the ordinary definitions  $\rho(x, t) = |\Psi(x, t)|^2$  and  $j(x, t) = \Re[\Psi^*(x, t) (\hbar/i\mu) \Psi(x, t)]$ . When the flux density  $j(x, t)$  changes its sign, quantity  $W(x, t) dt$  is no longer positive-definite and, as in Refs.(Olkhovsky, 1984; Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007), it acquires the physical meaning of a probability density *only* during those partial time-intervals in which the flux

<sup>2</sup> The Naimark theorem states in particular the following(Naimark, 1940): The non-orthogonal spectral decomposition of a maximal hermitian operator can be approximated by an orthogonal spectral function (which corresponds to a selfadjoint operator), in a weak convergence, with any desired accuracy.

<sup>3</sup> Such a condition is enough for operator (1a,b) to be a *hermitian*, or more precisely a *maximal hermitian*[2–8] *operator* (see also (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007)); but it can be dispensed with by recourse to bilinear forms (see, e.g., Refs.(Recami, 1976; 1977; Recami et al., 1983) and refs. therein), as we shall see below.

density  $j(x, t)$  does keep its sign. Therefore, let us introduce the *two* measures (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; 2004; Olkhovsky & Recami, 2007) by separating the positive and the negative flux-direction values (that is, the flux signs)

$$W_{\pm}(t, x) dt = \frac{j_{\pm}(x, t) dt}{\int_{-\infty}^{+\infty} j_{\pm}(x, t) dt} \quad (3)$$

with  $j_{\pm}(x, t) = j(x, t)\theta(\pm j)$ .

Then, the mean value  $\langle t_{\pm}(x) \rangle$  of the time  $t$  at which the particle passes through position  $x$ , when traveling in the positive or negative direction, is, respectively,

$$\langle t_{\pm}(x) \rangle = \frac{\int_{-\infty}^{+\infty} t j_{\pm}(x, t) dt}{\int_{-\infty}^{+\infty} j_{\pm}(x, t) dt} = \frac{\int_0^{+\infty} \frac{1}{2} [G^*(x, E) \hat{t} v G(x, E) + v G^*(x, E) \hat{t} G(x, E)] dE}{\int_0^{+\infty} v |G(x, E)|^2 dE}, \quad (4)$$

where  $G(x, E)$  is the Fourier-transform of the moving 1D wave-packet

$$\begin{aligned} \Psi(x, t) &= \int_0^{+\infty} G(x, E) \exp(-iEt/\hbar) dE = \\ &= \int_0^{+\infty} g(E) \varphi(x, E) \exp(-iEt/\hbar) dE \end{aligned}$$

when going on from the time to the energy representation. For free motion, one has  $G(x, E) = g(E) \exp(ikx)$ , and  $\varphi(x, E) = \exp(ikx)$ , while  $E = \mu \hbar^2 k^2 / 2 = \mu v^2 / 2$ . In Refs. (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; 2004; Olkhovsky & Recami, 2007), there were defined the mean time *durations* for the particle 1D transmission from  $x_i$  to  $x_f > x_i$ , and reflection from the region  $(x_i, +\infty)$  back to the interval  $x_f \leq x_i$ . Namely

$$\langle \tau_T(x_i, x_f) \rangle = \langle t_+(x_f) \rangle - \langle t_+(x_i) \rangle \quad (5)$$

and

$$\langle \tau_R(x_i, x_f) \rangle = \langle t_-(x_f) \rangle - \langle t_+(x_i) \rangle, \quad (6)$$

respectively. The 3D generalization for the mean durations of quantum collisions and nuclear reactions appeared in (Olkhovsky, 1984; 1990; 1992; 1998). Finally, suitable definitions of the averages  $\langle t^n \rangle$  on time of  $t^n$ , with  $n = 1, 2 \dots$ , and of  $\langle f(t) \rangle$ , quantity  $f(t)$  being any analytical function of time, can be found in (Olkhovsky & Recami, 2007; 2008), where single-valued expressions have been explicitly written down.

The two canonically conjugate operators, the time operator (1) and the energy operator

$$\hat{E} = \begin{cases} E, & \text{in the energy (E-) representation, (a)} \\ i\hbar \frac{\partial}{\partial t}, & \text{in the time (t-) representation (b)} \end{cases} \quad (7)$$

do clearly satisfy the commutation relation (Olkhovsky & Recami, 2007; 2008; Recami, 1976; 1977)

$$[\hat{E}, \hat{t}] = i\hbar. \quad (8)$$

The Stone and von Neumann theorem (Stone, 1930), has been always interpreted as establishing a commutation relation like (8) for the pair of the canonically conjugate operators (1) and (7), in both representations, for selfadjoint operators only. However, it can be generalized for (maximal) hermitian operators, once one introduces  $\hat{t}$  by means of the *single-valued* Fourier transformation from the  $t$ -axis ( $-\infty < t < \infty$ ) to the  $E$ -semiaxis ( $0 < E < \infty$ ), and utilizes the properties (Akhiezer & Glazman, 1981; D ter Haar, 1971) of the “(maximal) hermitian” operators: This has been shown, e.g., in the last one of Refs. (Olkhovsky & Recami, 1968; 1969) as well as in Refs. (Olkhovsky & Recami, 2007; 2008).

Indeed, from eq. (8) the uncertainty relation

$$\Delta E \Delta t \geq \hbar/2 \quad (9)$$

(where the standard deviations are  $\Delta a = \sqrt{Da}$ , quantity  $Da$  being the variance  $Da = \langle a^2 \rangle - \langle a \rangle^2$ , and  $a = E, t$ , while  $\langle \dots \rangle$  denotes the average over  $t$  with the measures  $W(x, t) dt$  or  $W_{\pm}(x, t) dt$  in the  $t$ -representation) can be derived also for operators which are simply hermitian, by a straightforward generalization of the procedures which are common in the case of *selfadjoint* (canonically conjugate) quantities, like coordinate  $\hat{x}$  and momentum  $\hat{p}_x$ . Moreover, relation (8) satisfies (Olkhovsky & Recami, 2007; 2008) the Dirac “correspondence” principle, since the classical Poisson brackets  $\{q_0, p_0\}$ , with  $q_0 = t$  and  $p_0 = -E$ , are equal to 1. In Refs. (Olkhovsky, 1973; Olkhovsky et al., 1974; Olkhovsky, 1984; Recami, 1976; 1977), and (Olkhovsky & Recami, 2007; 2008), it was also shown that *the differences*, between the mean times at which a wave-packet passes through a *pair* of points, obey the Ehrenfest correspondence principle.

As a consequence, one can state that, for systems with continuous energy spectra, the mathematical properties of (maximal) hermitian operators, like  $\hat{t}$  in eq. (1), are sufficient for considering them as quantum observables. Namely, the *uniqueness* (Akhiezer & Glazman, 1981) of the spectral decomposition (although not orthogonal) for operators  $\hat{t}$ , and  $\hat{t}^n$  ( $n > 1$ ), guarantees the “equivalence” of the mean values of any analytical function of time when evaluated in the  $t$  and in the  $E$ -representations. In other words, such an expansion is equivalent to a completeness relation, for the (approximate) eigenfunctions of  $\hat{t}^n$  ( $n > 1$ ), which *with any accuracy* can be regarded as orthogonal, and corresponds to the actual eigenvalues for the continuous spectrum. These approximate eigenfunctions belong to the space of the square-integrable functions of the energy  $E$  (cf., for instance, see, for instance Refs. (Olkhovsky, 1984; 1990; 1992; 1998; Olkhovsky & Recami, 2007; Recami, 1976; 1977) and refs. therein).

From this point of view, there is no *practical* difference between selfadjoint and maximal hermitian operators for systems with continuous energy spectra. Let us repeat that the mathematical properties of  $\hat{t}^n$  ( $n > 1$ ) are enough for considering time as a quantum mechanical observable (like energy, momentum, space coordinates, etc.) *without having to introduce any new physical postulates*.

It is remarkable that von Neumann himself (Von Neumann, 1955), before confining himself for simplicity to selfadjoint operators, stressed that operators like our time  $\hat{t}$  may represent physical observables, even if they are not selfadjoint. Namely, he explicitly considered the example of the operator  $-i\hbar \partial/\partial x$  associated with a particle living in the right semi-space bounded by a rigid wall located at  $x = 0$ ; that operator is not selfadjoint (acting on wave packets defined

on the positive  $x$ -axis) only, nevertheless it obviously corresponds to the  $x$ -component of the observable *momentum* for that particle: See Fig.1.

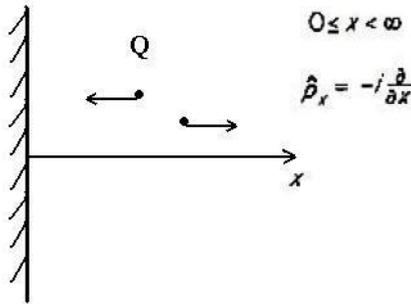


Fig. 1. For a particle  $Q$  free to move in a semi-space, bounded by a rigid wall located at  $x = 0$ , the operator  $-i\partial/\partial x$  has the clear physical meaning of the particle momentum  $x$ -component even if it is *not* selfadjoint (cf. von Neumann (Von Neumann, 1955), and Ref. (Recami, 1976; 1977)): See the text.

At this point, let us emphasize that our previously assumed boundary condition  $E \neq 0$  can be dispensed with, by having recourse (Olkhovsky & Recami, 1968; 1969; Recami, 1976; 1977) to the *bi-linear* hermitian operator

$$\hat{t} = \frac{-i\hbar}{2} \overset{\leftrightarrow}{\frac{\partial}{\partial E}} \quad (10)$$

where the meaning of the sign  $\leftrightarrow$  is clear from the accompanying definition

$$(f, \hat{t} g) = \left( f, -\frac{i\hbar}{2} \frac{\partial}{\partial E} g \right) + \left( -\frac{i\hbar}{2} \frac{\partial}{\partial E} f, g \right).$$

By adopting this expression for the time operator, the algebraic sum of the two terms in the r.h.s. of the last relation results to be automatically zero at point  $E = 0$ . This question will be exploited below, in Sect. 3 (when dealing with the more general case of the four-position operator). Incidentally, such an “elimination” (Olkhovsky & Recami, 1968; 1969; Recami, 1976; 1977) of point  $E = 0$  is not only simpler, but also more physical, than other kinds of elimination obtained much later in papers like (Egusquiza & Muga, 1999; Muga et al., 1999).

In connection with the last quotation, let us for briefly comment on the so-called *positive-operator-value-measure* (POVM) approach, often used or discussed in the second set of papers on time in quantum physics mentioned in our Introduction. Actually, an analogous procedure had been proposed, since the sixties (Aharonov & Bohm, 1961), in some approaches to the quantum theory of measurements. Afterwards, and much later, the POVM approach has been applied, in a simplified and shortened form, to the time-operator problem in the case of one-dimensional free motion: for instance, in Refs. (Delgado, 1999; Egusquiza & Muga, 1999; Giannitrapani, 1997; Gózdź A & Dębicki, 2007; Kijowski, 1997; Kobe et al., 1994; Kochański & Wołkievicz, 1999; Leoń, 1997; Muga et al., 1999; Srinivas & Vijayalakshmi, 1981; Toller, 1999; Wang & Xiong, 2007) and especially in (Egusquiza & Muga, 1999; Muga et al., 1999). These papers stated that a generalized decomposition of unity (or “POV measure”) could be obtained from selfadjoint extensions of the time operator inside an extended Hilbert

space (for instance, adding the negative values of the energy, too), by exploiting the Naimark dilation-theorem(Naimark, 1943): But such a program has been realized till now only in the simple cases of one-dimensional particle free motion.

By contrast, our approach is based on a different Naimark's theorem(Naimark, 1940), which, as already mentioned above, allows a much more direct, simple and general –and at the same time non less rigorous– introduction of a quantum operator for Time. More precisely, our approach is based on the so-called *Carleman theorem*(Carleman, 1923), utilized in Ref.(Naimark, 1940), about approximating a hermitian operator by suitable successions of “bounded” selfadjoint operators: That is, of selfadjoint operators whose spectral functions do weakly converge to the non-orthogonal spectral function of the considered hermitian operator. And our approach is applicable to a large family of three-dimensional (3D) particle collisions, with all possible Hamiltonians. Actually, our approach was proposed in the early Refs. (Olkhovsky & Recami, 1968; 1969; 1970; Olkhovsky, 1973; Olkhovsky et al., 1974; Olkhovsky, 1984; Recami, 1976; 1977) and in the first one of Ref. (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995), and applied therein for the time analysis of quantum collisions, nuclear reactions and tunnelling processes.

## 2.2 On the momentum representation of the Time operator

In the continuous spectrum case, instead of the  $E$ -representation, with  $0 < E < +\infty$ , in eqs.(1)–(4) one can also use the  $k$ -representation(Holevo, 1978; 1982), with the advantage that  $-\infty < k < +\infty$ :

$$\Psi(x, t) = \int_{-\infty}^{+\infty} g(k) \varphi(x, k) \exp(-iEt/\hbar) dk \quad (11)$$

with  $E = \hbar^2 k^2 / 2\mu$ , and  $k \neq 0$ .

For the extension of the momentum representation to the case of  $\langle t^n \rangle$ , with  $n > 1$ , we confine ourselves here to refer the reader to the papers (Olkhovsky & Recami, 2007; 2008).

## 2.3 An alternative weight for time averages (in the cases of particle dwelling inside a certain spatial region)

We recall that the weight (2) [as well as its modifications (3)] has the meaning of a probability for the considered particle to pass through point  $x$  during the time interval  $(t, t + dt)$ . Let us follow the procedure presented in Refs. (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007) and refs. therein, and analyze the consequences of the equality

$$\int_{-\infty}^{+\infty} j(x, t) dt = \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx \quad (12)$$

obtained from the 1D continuity equation. One can easily realize that a second, alternative weight can be adopted:

$$dP(x, t) \equiv Z(x, t) dx = \frac{|\Psi(x, t)|^2 dx}{\int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx} \quad (13)$$

which possesses the meaning of probability for the particle to be located (or to sojourn, i. e., to *dwell*) inside the infinitesimal space region  $(x, x + dx)$  at the instant  $t$ , independently of its motion properties. Then, the quantity

$$P(x_1, x_2, t) = \frac{\int_{x_1}^{x_2} |\Psi(x, t)|^2 dx}{\int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx} \quad (14)$$

will have the meaning of probability for the particle to dwell inside the spatial interval  $(x_1, x_2)$  at the instant  $t$ .

As it is known (see, for instance, Refs.(Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; 2004; Olkhovsky & Recami, 2007) and refs. therein), the *mean dwell time* can be written in the *two* equivalent forms:

$$\langle \tau(x_i, x_f) \rangle = \frac{\int_{-\infty}^{+\infty} dt \int_{x_i}^{x_f} |\Psi(x, t)|^2 dx}{\int_{-\infty}^{+\infty} j_{\text{in}}(x_i, t) dt} \quad (15)$$

and

$$\langle \tau(x_i, x_f) \rangle = \frac{\int_{-\infty}^{+\infty} t j(x_f, t) dt - \int_{-\infty}^{+\infty} t j(x_i, t) dt}{\int_{-\infty}^{+\infty} j_{\text{in}}(x_i, t) dt}, \quad (16)$$

where it has been taken account, in particular, of relation (12), which follows — as already said — from the continuity equation.

Thus, in correspondence with the two measures (2) and (13), when integrating over time one gets *two* different kinds of time distributions (mean values, variances...), which refer to the particle traversal time in the case of measure (2), and to the particle dwelling in the case of measure (13). Some examples for 1D tunneling are contained in Refs.(Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; 2004; Olkhovsky & Recami, 2007).

## 2.4 Extension of the notion of Time as a quantum-theoretical observable for the case of photons

As is known (see, for instance, Refs. (Akhiezer & Berestetzky, 1959; Olkhovsky et al., 2004; Schweber, 1961)), in first quantization the single-photon wave function can be probabilistically described in the 1D case by the wave-packet<sup>4</sup>

$$A(\mathbf{r}, t) = \int_{k_0} \frac{d^3k}{k_0} \chi(\mathbf{k}) \varphi(\mathbf{k}, \mathbf{r}) \exp(-ik_0 t), \quad (17)$$

<sup>4</sup> The gauge condition  $\text{div}A = 0$  is assumed.

where, as usual,  $A(\mathbf{r}, t)$  is the electromagnetic vector potential, while  $\mathbf{r} = \{x, y, z\}$ ,  $\mathbf{k} = \{k_x, k_y, k_z\}$ ,  $k_0 \equiv \omega/c = \varepsilon/\hbar c$ , and  $k \equiv |\mathbf{k}| = k_0$ . The axis  $x$  has been chosen as the propagation direction. Let us notice that  $\chi(\mathbf{k}) = \sum_{i=y,z} \chi_i(\mathbf{k}) e_i(\mathbf{k})$ , with  $e_i e_j = \delta_{ij}$ , and  $x_i, x_j = y, z$ , while  $\chi_i(\mathbf{k})$  is the probability amplitude for the photon to have momentum  $\mathbf{k}$  and polarization  $e_j$  along  $x_j$ . Moreover, it is  $\varphi(\mathbf{k}, \mathbf{r}) = \exp(ik_x x)$  in the case of plane waves, while  $\varphi(\mathbf{k}, \mathbf{r})$  is a linear combination of evanescent (decreasing) and anti-evanescent (increasing) waves in the case of “photon barriers” (i.e., band-gap filters, or even undersized segments of waveguides for microwaves, or frustrated total-internal-reflection regions for light, and so on). Although it is not easy to localize a photon in the direction of its polarization (Akhiezer & Berestezky, 1959; Schweber, 1961), nevertheless for 1D propagations it is possible to use the space-time probabilistic interpretation of eq. (17), and define the quantity

$$\rho_{\text{em}}(x, t) dx = \frac{S_0 dx}{\int S_0 dx}, \quad S_0 = \iint s_0 dy dz \quad (18)$$

( $s_0 = [\mathbf{E}^* \cdot \mathbf{E} + \mathbf{H}^* \cdot \mathbf{H}]/4\pi$  being the energy density, with the electromagnetic field  $\mathbf{H} = \text{rot } \mathbf{A}$ , and  $\mathbf{E} = -1/c \partial \mathbf{A}/\partial t$ ), which represents the probability density of a photon to be found (localized) in the spatial interval  $(x, x + dx)$  along the  $x$ -axis at the instant  $t$ ; and the quantity

$$j_{\text{em}}(x, t) dt = \frac{S_x dt}{\int S_x(x, t) dt}, \quad S_x(x, t) = \iint s_x dy dz \quad (19)$$

( $s_x = c \Re[\mathbf{E}^* \wedge \mathbf{H}]_x / 8\pi$  being the energy flux density), which represents the flux probability density of a photon to pass through point  $x$  in the time interval  $(t, t + dt)$ : in full analogy with the probabilistic quantities for non-relativistic particles. The justification and convenience of such definitions is self-evident, when the wave-packet group velocity coincides with the velocity of the energy transport; in particular: (i) the wave-packet (17) is quite similar to wave-packets for non-relativistic particles, and (ii) in analogy with conventional non-relativistic quantum mechanics, one can define the “mean time instant” for a photon (i.e., an electromagnetic wave-packet) to pass through point  $x$ , as follows

$$\langle t(x) \rangle = \int_{-\infty}^{+\infty} t J_{\text{em},x} dt = \frac{\int_{-\infty}^{+\infty} t S_x(x, t) dt}{\int_{-\infty}^{+\infty} S_x(x, t) dt}.$$

As a consequence [in the same way as in the case of equations (1)–(2)], the form (1) for the time operator in the energy representation is valid also for photons, with the same boundary conditions adopted in the case of particles, that is, with  $\chi_i(0) = \chi_i(\infty)$  and with  $E = \hbar c k_0$ . The energy density  $s_0$  and energy flux density  $s_x$  satisfy the relevant continuity equation

$$\frac{\partial s_0}{\partial t} + \frac{\partial s_x}{\partial x} = 0 \quad (20)$$

which is Lorentz-invariant for 1D spatial propagation (Olkhovsky et al., 2004; Olkhovskiy & Recami, 2007) processes.

## 2.5 Introducing the analogue of the “Hamiltonian” for the case of the Time operator: A new hamiltonian approach

In non-relativistic quantum theory, the Energy operator acquires (cf., e.g., Refs. (Olkhovsky, 1990; 1992; 1998; Olkhovsky & Recami, 2007)) the *two* forms: (i)  $i\hbar \frac{\partial}{\partial t}$  in the  $t$ -representation, and (ii)  $\hat{H}(\hat{p}_x, \hat{x}, \dots)$  in the hamiltonian formalism. The “duality” of these two forms can be easily inferred from the Schrödinger equation itself,  $\hat{H}\Psi = i\hbar \frac{\partial \Psi}{\partial t}$ . One can introduce in quantum mechanics a similar duality for the case of *Time*: Besides the general form (1) for the Time operator in the energy representation, which is valid for any physical systems in the region of continuous energy spectra, one can *express the time operator also in a “hamiltonian form”*, i.e., in terms of the coordinate and momentum operators, by having recourse to their commutation relations. Thus, by the replacements

$$\begin{aligned} \hat{E} &\rightarrow \hat{H}(\hat{p}_x, \hat{x}, \dots), \\ \hat{t} &\rightarrow \hat{T}(\hat{p}_x, \hat{x}, \dots), \end{aligned} \quad (21)$$

and on using the commutation relation [similar to eq. (3)]

$$[\hat{H}, \hat{T}] = i\hbar, \quad (22)$$

one can obtain (Rosenbaum, 1969), given a specific ordinary Hamiltonian, the corresponding explicit expression for  $\hat{T}(\hat{p}_x, \hat{x}, \dots)$ .

Indeed, this procedure can be adopted for any physical system with a known Hamiltonian  $\hat{H}(\hat{p}_x, \hat{x}, \dots)$ , and we are going to see a concrete example. By going on from the coordinate to the momentum representation, one realizes that the *formal* expressions of *both* the hamiltonian-type operators  $\hat{H}(\hat{p}_x, \hat{x}, \dots)$  and  $\hat{T}(\hat{p}_x, \hat{x}, \dots)$  *do not change*, except for an obvious change of sign in the case of operator  $\hat{T}(\hat{p}_x, \hat{x}, \dots)$ .

As an explicit example, let us address the simple case of a free particle whose Hamiltonian is

$$\hat{H} = \begin{cases} \hat{p}_x^2 / 2\mu, & \hat{p}_x = -i\hbar \frac{\partial}{\partial x}, & \text{in the coordinate representation (a)} \\ p_x^2 / 2\mu. & & \text{in the momentum representation (a)} \end{cases} \quad (23)$$

Correspondingly, the Hamilton-type *time operator*, in its symmetrized form, will write

$$\hat{T} = \begin{cases} \frac{\mu}{2} (\hat{p}_x^{-1} x + x \hat{p}_x^{-1} + i\hbar; \hat{p}_x^{-2}), & \text{in the coordinate representation (a)} \\ -\frac{\mu}{2} (p_x^{-1} \hat{x} + \hat{x} p_x^{-1} + i\hbar / p_x^2), & \text{in the momentum representation (b)} \end{cases} \quad (24)$$

where

$$\hat{p}_x^{-1} = \frac{i}{\hbar} \int dx \dots, \quad \hat{x} = i\hbar \frac{\partial}{\partial p_x}.$$

Incidentally, operator (24b) is equivalent to  $-i\hbar \frac{\partial}{\partial E}$ , since  $E = p_x^2 / 2\mu$ ; and therefore it is also a (maximal) *hermitian* operator. Indeed, by applying the operator  $\hat{T}(\hat{p}_x, \hat{x}, \dots)$ , for instance, to a plane-wave of the type  $\exp(ikx)$ , we obtain the same result in both the coordinate and the momentum representations:

$$\hat{T} \exp(ikx) = \frac{x}{v} \exp(ikx) \quad (25)$$

quantity  $x/v$  being the free-motion time (for a particle with velocity  $v$ ) for traveling the distance  $x$ .

On the basis of what precedes, it is possible to show that the wave function  $\Psi(x, t)$  of a quantum system satisfies the two (dual) equations

$$\hat{H}\Psi = i\hbar \frac{\partial \Psi}{\partial t} \text{ and } \hat{T}\Psi = t\Psi. \quad (26)$$

In the energy representation, and in the stationary case, we obtain again *two* (dual) equations

$$\hat{H}\varphi_\varepsilon = \varepsilon\varphi_\varepsilon \text{ and } \hat{T}\varphi_\varepsilon = -i\hbar \frac{\partial \varphi_\varepsilon}{\partial \varepsilon}, \quad (27)$$

quantity  $\varphi_\varepsilon$  being the Fourier-transform of  $\Psi$ :

$$\varphi_\varepsilon = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} \Psi(x, t) e^{i\varepsilon t/\hbar} dt. \quad (28)$$

It might be interesting to apply the two pairs of the last dual equations also for investigating tunnelling processes through the quantum gravitational barrier, which appears during inflation, or at the beginning of the big-bang expansion, whenever a quasi-linear Schrödinger-type equation does approximately show up.

## 2.6 Time as an observable (and the time-energy uncertainty relation), for quantum-mechanical systems with discrete energy spectra

For describing the time evolution of non-relativistic quantum systems endowed with a purely *discrete* (or a continuous *and discrete*) spectrum, let us now introduce wave-packets of the form (Olkhovsky, 1990; 1992; 1998; Olkhovsky & Recami, 2007; 2008):

$$\psi(x, t) = \sum_{n=0} g_n \varphi_n(x) \exp[-i(\varepsilon_n - \varepsilon_0)t/\hbar], \quad (29)$$

where  $\varphi_n(x)$  are orthogonal and normalized bound states which satisfy the equation  $\hat{H}\varphi_n(x) = \varepsilon_n \varphi_n(x)$ , quantity  $\hat{H}$  being the Hamiltonian of the system; while the coefficients  $g_n$  are normalized:  $\sum_{n=0} |g_n|^2 = 1$ . We omitted the non-significant phase factor  $\exp(-i\varepsilon_0 t/\hbar)$  of the fundamental state.

Let us first consider the systems whose energy levels are separated by intervals admitting a maximum common divisor  $D$  (for ex., harmonic oscillator, particle in a rigid box, and spherical spinning top), so that the wave packet (29) is a periodic function of time possessing as period the Poincaré cycle time  $T = 2\pi\hbar/D$ . For such systems it is possible (Olkhovsky, 1990; 1992; 1998; Olkhovsky & Recami, 2007; 2008) to construct a *selfadjoint* time operator with the form (in the time representation) of a saw-function of  $t$ , choosing  $t = 0$  as the initial time instant:

$$\hat{t} = t - T \sum_{n=0}^{\infty} \Theta(t - [2n + 1]T/2) + T \sum_{n=0}^{\infty} \Theta(-t - [2n + 1]T/2). \quad (30)$$

This periodic function for the time operator is a linear (increasing) function of time  $t$  within each Poincaré cycle: see Fig.2.

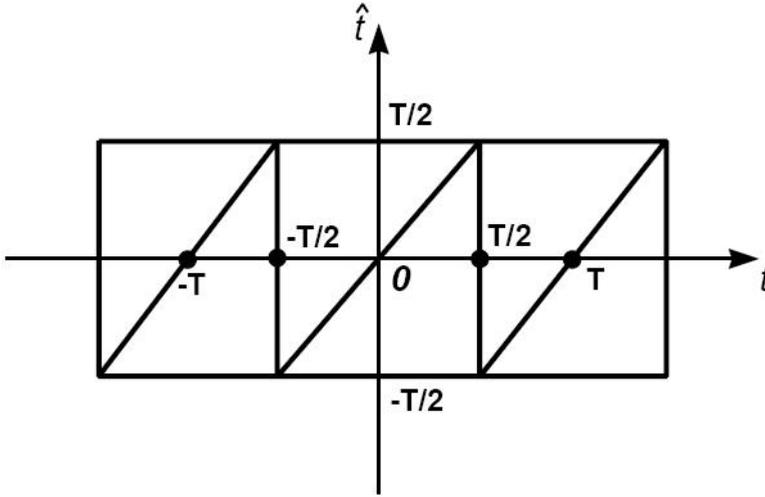


Fig. 2. The periodic saw-tooth function for the time operator in the case of quantum mechanical systems with *discrete* energy spectra: Namely, for the case of eq. (30).

The commutation relations of the Energy and Time operators, now both selfadjoint, acquires in the case of discrete energies and of a periodic Time operator the form

$$[\hat{E}, \hat{t}] = i\hbar \left\{ 1 - T \sum_{n=0}^{\infty} \delta(t - [2n + 1]T) \right\}, \quad (31)$$

wherefrom the uncertainty relation follows in the new form

$$(\Delta E)^2 (\Delta t)^2 = \hbar^2 \left[ 1 - \frac{T |\psi(T/2 + \gamma)|^2}{\int_{-T/2}^{T/2} |\psi(t)|^2 dt} \right], \quad (32)$$

where it has been introduced a parameter  $\gamma$ , with  $-T/2 < \gamma < T/2$ , in order to assure that the r.h.s. integral is single-valued (Olkhovsky & Recami, 2007; 2008).

When  $\Delta E \rightarrow 0$  (that is, when  $|g_n| \rightarrow \delta_{nm}$ ), the r.h.s. of eq. (32) tends to zero too, since  $|\psi(t)|^2$  tends to a constant value. In such a case, the distribution of the time instants at which the wave-packet passes through point  $x$  becomes flat within each Poincaré cycle. When, by contrast,  $\Delta E \gg D$  and  $|\psi(T + \gamma)|^2 \ll (\int_{-T/2}^{T/2} |\psi(t)|^2 dt)/T$ , the periodicity condition may become inessential whenever  $\Delta t \ll t$ . In other words, our uncertainty relation (32) transforms into the ordinary uncertainty relation for systems with continuous spectra.

In more general cases, for excited states of nuclei, atoms and molecules, the *energy-level intervals, for discrete and quasi-discrete (resonance) spectra, are not multiples of a maximum common divisor*, and hence the Poincaré cycle is not well-defined for such systems. Nevertheless, even for those systems one can introduce an approximate description (sometimes, with any desired degree of accuracy) in terms of Poincaré quasi-cycles and a quasi-periodical evolution; so that for sufficiently long time intervals the behavior of the wave-packets can be associated with a *periodical motion (oscillation)*, sometimes — e.g., for very narrow resonances — with any desired

accuracy. For them, when choosing an approximate Poincaré-cycle time, one can include in one cycle as many quasi-cycles as it is necessary for the demanded accuracy. Then, with the chosen accuracy, a *quasi-selfadjoint time operator* can be introduced.

### 3. Multiple internal reflections approach in description of tunneling

#### 3.1 Tunneling in consideration of multiple internal reflections of waves between internal boundaries

An approach for description of one-dimensional motion of a non-relativistic particle above a barrier on the basis of multiple internal reflections of stationary waves relatively boundaries has been studied in number of papers and is known (see (Anderson, 1989; Fermor, 1966; McVoy et al., 1967) and references therein). Tunneling of the particle under the barrier was described successfully on the basis of multiple internal reflections of the wave packets relatively boundaries (approach was called as *method of multiple internal reflections* or *method MIR*, see Refs. (Maydanyuk et al., 2002a;b; Maydanyuk, 2003; Olkhovsky, 2000)). In such approach it succeeded in connecting: 1) continuous transition of solutions for packets after each reflection, total packets between the above-barrier motion and the under-barrier tunneling; 2) coincidence of transmitted and reflected amplitudes of stationary wave function in each spatial region obtained by approach MIR with the corresponding amplitudes obtained by standard method of quantum mechanics; 3) all non-stationary fluxes in each step, are non-zero that confirms propagation of packets under the barrier (i. e. their “tunneling”). In frameworks of such a method, non-stationary tunneling obtained own interpretation, allowing to study this process at interesting time moment or space point. In calculation of phase times this method turns out to be enough simple and convenient (Cardone et al., 2006). It has been adapted for scattering of the particle on nucleus and  $\alpha$ -decay in the spherically symmetric approximation with the simplest radial barriers (Maydanyuk et al., 2002a; Maydanyuk, 2003; Olkhovsky, 2000) and for tunneling of photons (Cardone et al., 2006; Maydanyuk et al., 2002a). However, further realization of the MIR approach meets with three questions. 1) *Question on effectiveness*. The multiple reflections have been proved for the motion above one rectangular barrier and for tunneling under it (Anderson, 1989; Cardone et al., 2006; Maydanyuk et al., 2002a). However, after addition of the second step it becomes unclear how to separate the needed reflected waves from all their variety in calculation of all needed amplitudes. After obtaining exact solutions of the stationary amplitudes for two arbitrary rectangular barriers (Maydanyuk, 2003; Olkhovsky, 2000), it becomes unclear how to generalize such approach for barriers with arbitrary complicate shape. So, *we come to a serious unresolved problem of realization of the approach of multiple reflections in real quantum systems with complicated barriers*, and clear algorithms of calculation of amplitudes should be constructed. 2) *Question on correctness*. Whether is interference between packets formed relatively different boundaries appeared? Whether does this come to principally different results of the approach of multiple internal reflections and direct methods of quantum mechanics? Note that such interference cannot be appeared in tunneling through one rectangular barrier and, therefore, it could not visible in the previous papers. 3) *Question on uncertainty in radial problem*. Calculations of half-lives of different types of decays based on the semiclassical approach are prevailing today. For example, in Ref. (Buck et al., 1993) agreement between experimental data of  $\alpha$ -decay half-lives and ones calculated by theory is demonstrated in a wide region of nuclei from  $^{106}\text{Te}$  up to nucleus with  $A_d = 266$  and  $Z_d = 109$  (see Ref. (Denisov & Ikezoe, 2005) for some improved approaches). In review (Sobiczewski & Pomorski, 2007) methodology of calculation of half-lives for

spontaneous-fission is presented (see eqs. (21)–(24) in p. 321). Let us consider proton-decay of nucleus where proton penetrates from the internal region outside with its tunneling through the barrier. At the same boundary condition, reflected and incident waves turn out to be defined with uncertainty. How to determine them? The semiclassical approach gives such answer: *according to theory, in construction of well known formula for probability we neglect completely by the second (increasing) item of the wave function inside tunneling region* (see Ref. (Landau & Lifshitz, 1989), eq. (50.2), p. 221). In result, equality  $T^2 + R^2 = 1$  has no any sense (where  $T$  and  $R$  are coefficients of penetrability and reflection). Condition of continuity for the wave function and for total flux is broken at turning point. So, we do not find reflection  $R$ . We do not suppose on possible interference between incident and reflected waves which can be non zero. The penetrability is determined by the barrier shape inside tunneling region, while internal and external parts do not take influence on it. The penetrability does not dependent on depth of the internal well (while the simplest rectangular well and barrier give another exact result). But, the semiclassical approach is so prevailing that one can suppose that it has enough well approximation of the penetrability estimated. It turns out that in fully quantum approach to determine the penetrability through the barrier (constructed on the basis of realistic potential of interaction between proton and daughter nucleus) then one can obtain answer “no”. Fully quantum penetrability is a function of new additional independent parameters, it can achieve essential difference from semiclassical one (at the same boundary condition imposed on the wave function). This will be demonstrated below.

### 3.2 Tunneling of packet through one-dimensional rectangular step

Let us consider a problem of tunneling of a particle in a positive  $x$ -direction through an one-dimensional rectangular potential barrier (see Fig. 3). Let us label a region I for  $x < 0$ , a region II for  $0 < x < a$  and a region III for  $x > a$ , accordingly. In standard approach, with

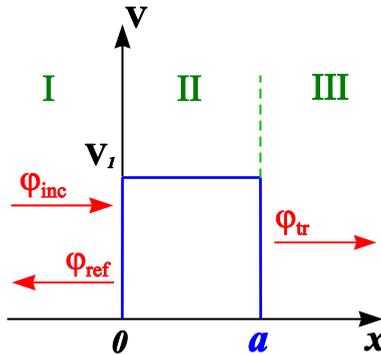


Fig. 3. Tunneling of the particle through one-dimensional rectangular barrier

energy less than the barrier height the tunneling evolution of the particle is described using a non-stationary propagation of WP

$$\psi(x, t) = \int_0^{+\infty} g(E - \bar{E}) \varphi(k, x) e^{-iEt/\hbar} dE, \quad (33)$$

where stationary WF is:

$$\varphi(x) = \begin{cases} e^{ikx} + A_R e^{-ikx}, & \text{for } x < 0; \\ \alpha e^{\xi x} + \beta e^{-\xi x}, & \text{for } 0 < x < a; \\ A_T e^{ikx}, & \text{for } x > a; \end{cases} \quad (34)$$

and  $k = \frac{1}{\hbar} \sqrt{2mE}$ ,  $\xi = \frac{1}{\hbar} \sqrt{2m(V_1 - E)}$ ,  $E$  and  $m$  are the total energy and mass of the particle, accordingly. The weight amplitude  $g(E - \bar{E})$  can be written in a form of gaussian and satisfies to a requirement of the normalization  $\int |g(E - \bar{E})|^2 dE = 1$ , value  $\bar{E}$  is an average energy of the particle. One can calculate coefficients  $A_T$ ,  $A_R$ ,  $\alpha$  and  $\beta$  analytically, using a requirements of a continuity of WF  $\varphi(x)$  and its derivative on each boundary of the barrier. Substituting in eq. (33) instead of  $\varphi(k, x)$  the incident  $\varphi_{inc}(k, x)$ , transmitted  $\varphi_{tr}(k, x)$  or reflected part of WF  $\varphi_{ref}(k, x)$ , defined by eq. (34), we receive the incident, transmitted or reflected WP, accordingly.

We assume, that a time, for which the WP tunnels through the barrier, is enough small. So, the time necessary for a tunneling of proton through a barrier of decay in proton-decay of a nucleus, is about  $10^{-21}$  seconds. We consider, that one can neglect a spreading of the WP for this time. And a breadth of the WP appears essentially more narrow on a comparison with a barrier breadth. Considering only sub-barrier processes, we exclude a component of waves for above-barrier energies, having included the additional transformation

$$g(E - \bar{E}) \rightarrow g(E - \bar{E})\theta(V_1 - E), \quad (35)$$

where  $\theta$ -function satisfies to the requirement

$$\theta(\eta) = \begin{cases} 0, & \text{for } \eta < 0; \\ 1, & \text{for } \eta > 0. \end{cases}$$

The method of multiple internal reflections considers the propagation process of the WP describing a motion of the particle, sequentially on steps of its penetration in relation to each boundary of the barrier (Anderson, 1989; Fermor, 1966; McVoy et al., 1967). Using this method, we find expressions for the transmitted and reflected WP in relation to the barrier. At the first step we consider the WP in the region I, which is incident upon the first (initial) boundary of the barrier. Let us assume, that this package transforms into the WP, transmitted through this boundary and tunneling further in the region II, and into the WP, reflected from the boundary and propagating back in the region I. Thus we consider, that the WP, tunneling in the region II, is not reached the second (final) boundary of the barrier because of a terminating velocity of its propagation, and consequently at this step we consider only two regions I and II. Because of physical reasons to construct an expression for this packet, we consider, that its amplitude should decrease in a positive  $x$ -direction. We use only one item  $\beta \exp(-\xi x)$  in eq. (34), throwing the second increasing item  $\alpha \exp(\xi x)$  (in an opposite case we break a requirement of a finiteness of the WF for an indefinitely wide barrier). In result, in the region II we obtain:

$$\psi_{tr}^1(x, t) = \int_0^{+\infty} g(E - \bar{E})\theta(V_1 - E)\beta^0 e^{-\xi x - iEt/\hbar} dE, \text{ for } 0 < x < a. \quad (36)$$

Thus the WF in the barrier region constructed by such way, is an analytic continuation of a relevant expression for the WF, corresponding to a similar problem with above-barrier

energies, where as a stationary expression we select the wave  $\exp(ik_2x)$ , propagated to the right.

Let us consider the first step further. One can write expressions for the incident and the reflected WP in relation to the first boundary as follows

$$\begin{aligned}\psi_{inc}(x, t) &= \int_0^{+\infty} g(E - \bar{E})\theta(V_1 - E)e^{ikx - iEt/\hbar} dE, & \text{for } x < 0, \\ \psi_{ref}^1(x, t) &= \int_0^{+\infty} g(E - \bar{E})\theta(V_1 - E)A_R^0 e^{-ikx - iEt/\hbar} dE, & \text{for } x < 0.\end{aligned}\quad (37)$$

A sum of these expressions represents the complete WF in the region I, which is dependent on a time. Let us require, that this WF and its derivative continuously transform into the WF (36) and its derivative at point  $x = 0$  (we assume, that the weight amplitude  $g(E - \bar{E})$  differs weakly at transmitting and reflecting of the WP in relation to the barrier boundaries). In result, we obtain two equations, in which one can pass from the time-dependent WP to the corresponding stationary WF and obtain the unknown coefficients  $\beta^0$  and  $A_R^0$ .

At the second step we consider the WP, tunneling in the region II and incident upon the second boundary of the barrier at point  $x = a$ . It transforms into the WP, transmitted through this boundary and propagated in the region III, and into the WP, reflected from the boundary and tunneled back in the region II. For a determination of these packets one can use eq. (33) with account eq. (35), where as the stationary WF we use:

$$\begin{aligned}\varphi_{inc}^2(k, x) &= \varphi_{tr}^1(k, x) = \beta^0 e^{-\xi x}, & \text{for } 0 < x < a, \\ \varphi_{tr}^1(k, x) &= A_T^0 e^{ikx}, & \text{for } x > a, \\ \varphi_{ref}^2(k, x) &= \alpha^0 e^{\xi x}, & \text{for } 0 < x < a.\end{aligned}\quad (38)$$

Here, for forming an expression for the WP reflected from the boundary, we select an increasing part of the stationary solution  $\alpha^0 \exp(\xi x)$  only. Imposing a condition of continuity on the time-dependent WF and its derivative at point  $x = a$ , we obtain 2 new equations, from which we find the unknowns coefficients  $A_T^0$  and  $\alpha^0$ .

At the third step the WP, tunneling in the region II, is incident upon the first boundary of the barrier. Then it transforms into the WP, transmitted through this boundary and propagated further in the region I, and into the WP, reflected from boundary and tunneled back in the region II. For a determination of these packets one can use eq. (33) with account eq. (35), where as the stationary WF we use:

$$\begin{aligned}\varphi_{inc}^3(k, x) &= \varphi_{ref}^2(k, x), & \text{for } 0 < x < a, \\ \varphi_{tr}^3(k, x) &= A_R^1 e^{-ikx}, & \text{for } x < 0, \\ \varphi_{ref}^3(k, x) &= \beta^1 e^{-\xi x}, & \text{for } 0 < x < a.\end{aligned}\quad (39)$$

Using a conditions of continuity for the time-dependent WF and its derivative at point  $x = 0$ , we obtain the unknowns coefficients  $A_R^1$  and  $\beta^1$ .

Analyzing further possible processes of the transmission (and the reflection) of the WP through the boundaries of the barrier, we come to a deduction, that any of following steps can be reduced to one of 2 considered above. For the unknown coefficients  $\alpha^n$ ,  $\beta^n$ ,  $A_T^n$  and  $A_R^n$ , used in expressions for the WP, forming in result of some internal reflections from the

boundaries, one can obtain the recurrence relations:

$$\begin{aligned}\beta^0 &= \frac{2k}{k+i\zeta}, \quad \alpha^n = \beta^n \frac{i\zeta - k}{i\zeta + k} e^{-2\zeta a}, \quad \beta^{n+1} = \alpha^n \frac{i\zeta - k}{i\zeta + k}, \\ A_R^0 &= \frac{k-i\zeta}{k+i\zeta}, \quad A_T^n = \beta^n \frac{2i\zeta}{i\zeta + k} e^{-\zeta a - ika}, \quad A_R^{n+1} = \alpha^n \frac{2i\zeta}{i\zeta + k}.\end{aligned}\quad (40)$$

Considering the propagation of the WP by such way, we obtain expressions for the WF on each region which can be written through series of multiple WP. Using eq. (33) with account eq. (35), we determine resultant expressions for the incident, transmitted and reflected WP in relation to the barrier, where one can need to use following expressions for the stationary WF:

$$\begin{aligned}\varphi_{inc}(k, x) &= e^{ikx}, \quad \text{for } x < 0, \\ \varphi_{tr}(k, x) &= \sum_{n=0}^{+\infty} A_T^n e^{ikx}, \quad \text{for } x > a, \\ \varphi_{ref}(k, x) &= \sum_{n=0}^{+\infty} A_R^n e^{-ikx}, \quad \text{for } x < 0.\end{aligned}\quad (41)$$

Now we consider the WP formed in result of sequential  $n$  reflections from the boundaries of the barrier and incident upon one of these boundaries at point  $x = 0$  ( $i = 1$ ) or at point  $x = a$  ( $i = 2$ ). In result, this WP transforms into the WP  $\psi_{tr}^i(x, t)$ , transmitted through boundary with number  $i$ , and into the WP  $\psi_{ref}^i(x, t)$ , reflected from this boundary. For an independent on  $x$  parts of the stationary WF one can write:

$$\begin{aligned}\frac{\varphi_{tr}^1}{\exp(-\zeta x)} &= T_1^+ \frac{\varphi_{inc}^1}{\exp(ikx)}, \quad \frac{\varphi_{ref}^1}{\exp(-ikx)} = R_1^+ \frac{\varphi_{inc}^1}{\exp(ikx)}, \\ \frac{\varphi_{tr}^2}{\exp(ikx)} &= T_2^+ \frac{\varphi_{inc}^2}{\exp(-\zeta x)}, \quad \frac{\varphi_{ref}^2}{\exp(\zeta x)} = R_2^+ \frac{\varphi_{inc}^2}{\exp(-\zeta x)}, \\ \frac{\varphi_{tr}^1}{\exp(-ikx)} &= T_1^- \frac{\varphi_{inc}^1}{\exp(\zeta x)}, \quad \frac{\varphi_{ref}^1}{\exp(-\zeta x)} = R_1^- \frac{\varphi_{inc}^1}{\exp(\zeta x)},\end{aligned}\quad (42)$$

where the sign “+” (or “-”) corresponds to the WP, tunneling (or propagating) in a positive (or negative)  $x$ -direction and incident upon the boundary with number  $i$ . Using  $T_i^\pm$  and  $R_i^\pm$ , one can precisely describe an arbitrary WP which has formed in result of  $n$ -multiple reflections, if to know a “path” of its propagation along the barrier. Using the recurrence relations eq. (40), the coefficients  $T_i^\pm$  and  $R_i^\pm$  can be obtained.

$$\begin{aligned}T_1^+ &= \beta^0, \quad T_2^+ = \frac{A_T^n}{\beta^n}, \quad T_1^- = \frac{A_R^{n+1}}{\alpha^n}, \\ R_1^+ &= A_R^0, \quad R_2^+ = \frac{\alpha^n}{\beta^n}, \quad R_1^- = \frac{\beta^{n+1}}{\alpha^n}.\end{aligned}\quad (43)$$

Using the recurrence relations, one can find series of coefficients  $\alpha^n, \beta^n, A_T^n$  and  $A_R^n$ . However, these series can be calculated easier, using coefficients  $T_i^\pm$  and  $R_i^\pm$ . Analyzing all possible

“paths” of the WP propagations along the barrier, we receive:

$$\begin{aligned}
\sum_{n=0}^{+\infty} A_T^n &= T_2^+ T_1^- \left( 1 + \sum_{n=1}^{+\infty} (R_2^+ R_1^-)^n \right) = \frac{i4k\zeta e^{-\zeta a - ika}}{F_{sub}}, \\
\sum_{n=0}^{+\infty} A_R^n &= R_1^+ + T_1^+ R_2^+ T_1^- \left( 1 + \sum_{n=1}^{+\infty} (R_2^+ R_1^-)^n \right) = \frac{k_0^2 D_-}{F_{sub}}, \\
\sum_{n=0}^{+\infty} \alpha^n &= \alpha^0 \left( 1 + \sum_{n=1}^{+\infty} (R_2^+ R_1^-)^n \right) = \frac{2k(i\zeta - k)e^{-2\zeta a}}{F_{sub}}, \\
\sum_{n=0}^{+\infty} \beta^n &= \beta^0 \left( 1 + \sum_{n=1}^{+\infty} (R_2^+ R_1^-)^n \right) = \frac{2k(i\zeta + k)}{F_{sub}},
\end{aligned} \tag{44}$$

where

$$\begin{aligned}
F_{sub} &= (k^2 - \zeta^2)D_- + 2ik\zeta D_+, \\
D_{\pm} &= 1 \pm e^{-2\zeta a}, \\
k_0^2 &= k^2 + \zeta^2 = \frac{2mV_1}{\hbar^2}.
\end{aligned} \tag{45}$$

All series  $\sum \alpha^n$ ,  $\sum \beta^n$ ,  $\sum A_T^n$  and  $\sum A_R^n$ , obtained using the method of multiple internal reflections, coincide with the corresponding coefficients  $\alpha$ ,  $\beta$ ,  $A_T$  and  $A_R$  of the eq. (34), calculated by a stationary methods. Using the following substitution

$$i\zeta \rightarrow k_2, \tag{46}$$

where  $k_2 = \frac{1}{\hbar} \sqrt{2m(E - V_1)}$  is a wave number for a case of above-barrier energies, expression for the coefficients  $\alpha^n$ ,  $\beta^n$ ,  $A_T^n$  and  $A_R^n$  for each step, expressions for the WF for each step, the total eqs. (44) and (45) transform into the corresponding expressions for a problem of the particle propagation above this barrier. At the transformation of the WP and the time-dependent WF one can need to change a sign of argument at  $\theta$ -function. Besides the following property is fulfilled:

$$\left| \sum_{n=0}^{+\infty} A_T^n \right|^2 + \left| \sum_{n=0}^{+\infty} A_R^n \right|^2 = 1. \tag{47}$$

### 3.3 Exact solutions for wave function for tunneling through radial barrier composed from arbitrary number of potential steps

Now we shall come to radial problem (Maydanyuk & Belchikov, 2011). Let us assume that starting from some time moment before decay the nucleus could be considered as system composite from daughter nucleus and fragment emitted. Its decay is described by a particle with reduced mass  $m$  which moves in radial direction inside a radial potential with a barrier. We shall be interesting in the radial barrier of arbitrary shape, which has successfully been approximated by finite number  $N$  of rectangular steps:

$$V(r) = \begin{cases} V_1, & \text{at } R_{\min} < r \leq r_1 & \text{(region 1),} \\ V_2, & \text{at } r_1 \leq r \leq r_2 & \text{(region 2),} \\ \dots & \dots & \dots \\ V_N, & \text{at } r_{N-1} \leq r \leq R_{\max} & \text{(region N),} \end{cases} \tag{48}$$

where  $V_i$  are constants ( $i = 1 \dots N$ ). We define the first region 1 starting from point  $R_{\min}$ , assuming that the fragment is formed here and then it moves outside. We shall be interesting

in solutions for above barrier energies while the solution for tunneling could be obtained after by change  $i \zeta_i \rightarrow k_i$ . A general solution of the wave function (up to its normalization) has the following form:

$$\psi(r, \theta, \varphi) = \frac{\chi(r)}{r} Y_{lm}(\theta, \varphi), \quad (49)$$

$$\chi(r) = \begin{cases} e^{ik_1 r} + A_R e^{-ik_1 r}, & \text{at } R_{\min} < r \leq r_1 & \text{(region 1),} \\ \alpha_2 e^{ik_2 r} + \beta_2 e^{-ik_2 r}, & \text{at } r_1 \leq r \leq r_2 & \text{(region 2),} \\ \dots & \dots & \dots \\ \alpha_{n-1} e^{ik_{N-1} r} + \beta_{N-1} e^{-ik_{N-1} r}, & \text{at } r_{N-2} \leq r \leq r_{N-1} & \text{(region } N-1), \\ A_T e^{ik_N r}, & \text{at } r_{N-1} \leq r \leq R_{\max} & \text{(region } N), \end{cases} \quad (50)$$

where  $\alpha_j$  and  $\beta_j$  are unknown amplitudes,  $A_T$  and  $A_R$  are unknown amplitudes of transmission and reflection,  $Y_{lm}(\theta, \varphi)$  is spherical function,  $k_i = \frac{1}{\hbar} \sqrt{2m(E - V_i)}$  are complex wave numbers. We shall be looking for solution for such problem in approach of multiple internal reflections (we restrict ourselves by a case of orbital moment  $l = 0$  while its non-zero generalization changes the barrier shape which was used as arbitrary before in development of formalism MIR and, so, is absolutely non principal).

According to the method of multiple internal reflections, scattering of the particle on the barrier is considered on the basis of wave packet consequently by steps of its propagation relatively to each boundary of the barrier (the most clearly idea of such approach can be understood in the problem of tunneling through the simplest rectangular barrier, see (Cardone et al., 2006; Maydanyuk et al., 2002a; Maydanyuk, 2003) where one can find proof of this fully quantum exactly solvable method, one can analyze its properties). Each step in such consideration of propagation of the packet will be similar to one from the first  $2N - 1$  steps, independent between themselves. From analysis of these steps recurrent relations are found for calculation of unknown amplitudes  $A^{(n)}$ ,  $S^{(n)}$ ,  $\alpha^{(n)}$  and  $\beta^{(n)}$  for arbitrary step  $n$ , summation of these amplitudes are calculated. We shall be looking for the unknown amplitudes, requiring wave function and its derivative to be continuous at each boundary. We shall consider the coefficients  $T_1^\pm, T_2^\pm, T_3^\pm \dots$  and  $R_1^\pm, R_2^\pm, R_3^\pm \dots$  as additional factors to amplitudes  $e^{\pm i k x}$ . Here, bottom index denotes number of the boundary, upper (top) signs “+” and “-” denote directions of the wave to the right or to the left, correspondingly. At the first, we calculate  $T_1^\pm, T_2^\pm \dots T_{N-1}^\pm$  and  $R_1^\pm, R_2^\pm \dots R_{N-1}^\pm$ :

$$\begin{aligned} T_j^+ &= \frac{2k_j}{k_j + k_{j+1}} e^{i(k_j - k_{j+1})r_j}, & T_j^- &= \frac{2k_{j+1}}{k_j + k_{j+1}} e^{i(k_j - k_{j+1})r_j}, \\ R_j^+ &= \frac{k_j - k_{j+1}}{k_j + k_{j+1}} e^{2ik_j r_j}, & R_j^- &= \frac{k_{j+1} - k_j}{k_j + k_{j+1}} e^{-2ik_{j+1} r_j}. \end{aligned} \quad (51)$$

Using recurrent relations:

$$\begin{aligned} \tilde{R}_{j-1}^+ &= R_{j-1}^+ + T_{j-1}^+ \tilde{R}_j^+ T_{j-1}^- \left( 1 + \sum_{m=1}^{+\infty} (\tilde{R}_j^+ R_{j-1}^-)^m \right) = R_{j-1}^+ + \frac{T_{j-1}^+ \tilde{R}_j^+ T_{j-1}^-}{1 - \tilde{R}_j^+ R_{j-1}^-}, \\ \tilde{R}_{j+1}^- &= R_{j+1}^- + T_{j+1}^- \tilde{R}_j^- T_{j+1}^+ \left( 1 + \sum_{m=1}^{+\infty} (R_{j+1}^+ \tilde{R}_j^-)^m \right) = R_{j+1}^- + \frac{T_{j+1}^- \tilde{R}_j^- T_{j+1}^+}{1 - R_{j+1}^+ \tilde{R}_j^-}, \\ \tilde{T}_{j+1}^+ &= \tilde{T}_j^+ T_{j+1}^+ \left( 1 + \sum_{m=1}^{+\infty} (R_{j+1}^+ \tilde{R}_j^-)^m \right) = \frac{\tilde{T}_j^+ T_{j+1}^+}{1 - R_{j+1}^+ \tilde{R}_j^-}, \end{aligned} \quad (52)$$

and selecting as starting the following values:

$$\tilde{R}_{N-1}^+ = R_{N-1}^+, \tilde{R}_1^- = R_1^-, \tilde{T}_1^+ = T_1^+, \quad (53)$$

we calculate successively coefficients  $\tilde{R}_{N-2}^+ \dots \tilde{R}_1^+, \tilde{R}_2^- \dots \tilde{R}_{N-1}^-$  and  $\tilde{T}_2^+ \dots \tilde{T}_{N-1}^+$ . At finishing, we determine coefficients  $\beta_j$ :

$$\beta_j = \tilde{T}_{j-1}^+ \left( 1 + \sum_{m=1}^{+\infty} (\tilde{R}_j^+ \tilde{R}_{j-1}^-)^m \right) = \frac{\tilde{T}_{j-1}^+}{1 - \tilde{R}_j^+ \tilde{R}_{j-1}^-}, \quad (54)$$

the amplitudes of transmission and reflection:

$$A_T = \tilde{T}_{N-1}^+, A_R = \tilde{R}_1^+ \quad (55)$$

and corresponding coefficients of penetrability  $T$  and reflection  $R$ :

$$T_{MIR} = \frac{k_n}{k_1} |A_T|^2, R_{MIR} = |A_R|^2. \quad (56)$$

We check the property:

$$\frac{k_n}{k_1} |A_T|^2 + |A_R|^2 = 1 \quad \text{or} \quad T_{MIR} + R_{MIR} = 1, \quad (57)$$

which should be the test, whether the method MIR gives us proper solution for wave function. Now if energy of the particle is located below then height of one step with number  $m$ , then for description of transition of this particle through such barrier with its tunneling it shall need to use the following change:

$$k_m \rightarrow i \zeta_m. \quad (58)$$

For the potential from two rectangular steps (with different choice of their sizes) after comparison between the all amplitudes obtained by method of MIR and the corresponding amplitudes obtained by standard approach of quantum mechanics, we obtain coincidence up to first 15 digits. Increasing of number of steps up to some thousands keeps such accuracy and fulfillment of the property (57). This is important test which confirms reliability of the method MIR. So, we have obtained full coincidence between all amplitudes, calculated by method MIR and by standard approach of quantum mechanics, and that is way we generalize the method MIR for description of tunneling of the particle through potential, consisting from arbitrary number of rectangular barriers and wells of arbitrary shape.

### 3.4 Analysis of the proton-decay for ${}_{73}^{157}\text{Ta}$ , ${}_{75}^{161}\text{Re}$ , ${}_{77}^{167}\text{Ir}$ and ${}_{83}^{185}\text{Bi}$

Today, there are a lot of modern methods able to calculate half-lives, which have been studied experimentally well. So, we have a rich theoretical and experimental material for analysis. We shall use these nuclei:  ${}_{73}^{157}\text{Ta}$ ,  ${}_{75}^{161}\text{Re}$ ,  ${}_{77}^{167}\text{Ir}$  for  $l = 0$ , and  ${}_{53}^{109}\text{I}$ ,  ${}_{55}^{112}\text{Cm}$ ,  ${}_{69}^{147}\text{Tm}$  for  $l \neq 0$ . Such a choice we explain by that they have small coefficient of quadruple deformation  $\beta_2$  and at good approximation can be considered as spherical. We shall study proton-decay on the basis of leaving of the particle with reduced mass from the internal region outside with its tunneling through the barrier. This particle is supposed to start from  $R_{\min} \leq r \leq r_1$  and move outside ( $r_1$  is defined in eq. (1)). Using technique of the  $T_j^\pm$  and  $R_j^\pm$  coefficients in eqs. (51)–(53), we calculate total amplitudes of transmission  $A_T$  and reflection  $A_R$  by eqs. (55),

the penetrability coefficient  $T_{MIR}$  by eqs. (56). We check the found amplitudes, coefficients  $T_{MIR}$  and  $R_{MIR}$  comparing them with corresponding amplitudes and coefficients calculated by standard approach of quantum mechanics. We restrict ourselves by eq. (50) for  $F_1$  and find width  $\Gamma$  by eq. (48) and half-live  $\tau_{MIR}$  by eq. (52). We define the penetrability  $T_{WKB}$  by eq. (49), calculate  $\Gamma$ -width and half-live  $\tau_{WKB}$  by eqs. (48) and (52).

### 3.4.1 Dependence of the penetrability on the starting point

The first interesting result which we have obtained is *essential dependence of penetrability on the position of the first region where we localize the wave incident on the barrier*. In particular, we have analyzed how much the internal boundary  $R_{min}$  takes influence on the penetrability. Taking into account that width of each interval is 0.01 fm, we consider left boundary  $R_{min}$  of the first interval as a *starting point* (with error up to 0.01 fm), from here proton begins to move outside and is incident on the internal part of the barrier in the first stage of the proton decay. In the Fig. 4 [left panel] one can see that half-live of the proton decay of  $^{157}_{73}\text{Ta}$  is changed essentially at displacement of  $R_{min}$ . So, we establish *essential dependence of the penetrability on the starting point  $R_{start}$ , where the proton starts to move outside by approach MIR*. At

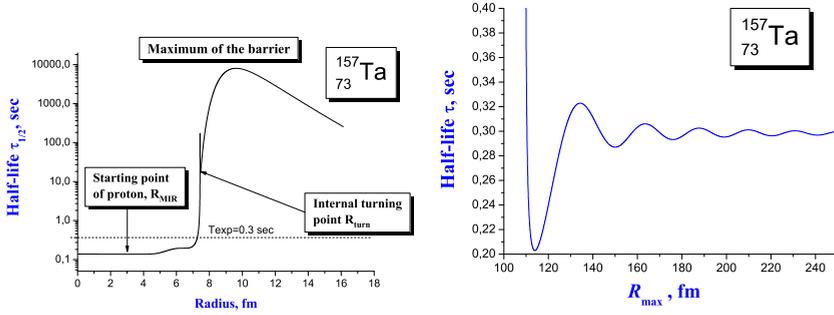


Fig. 4. Proton-decay for the  $^{157}_{73}\text{Ta}$  nucleus: the left panel is for dependence of the half-life  $\tau_{MIR}$  on the starting point  $R_{min}$ , the right panel is for dependence of the half-life  $\tau_{MIR}$  on  $R_{max}$  (here, we use  $R_{form} = 7.2127$  fm where calculated  $\tau_{MIR}$  at  $R_{max} = 250$  fm coincides with experimental data  $\tau_{exp}$  for this nucleus). In all calculations factor  $F$  is the same.

$R_{form} = 7.2127$  fm this dependence allows us to achieve very close coincidence between the half-live calculated by the approach MIR and experimental data.

### 3.4.2 Dependence of the penetrability on the external region

The region of the barrier located between turning points  $R_2$  and  $R_3$  is main part of the potential used in calculation of the penetrability in the semiclassical approach (up to the second correction), while the internal and external parts of this potential do not take influence on it. Let us analyze whether convergence exists in calculations of the penetrability in the approach MIR if to increase the external boundary  $R_{max}$  ( $R_{max} > R_3$ ). Keeping width of each interval (step) to be the same, we shall increase  $R_{max}$  (through increasing number of intervals in the external region), starting from the external turning point  $R_3$ , and calculate the corresponding penetrability  $T_{MIR}$ . In Fig. 4 [central panel] one can see how the penetrability is changed for  $^{157}_{73}\text{Ta}$  with increasing  $R_{max}$ . Dependence of the half-life  $\tau_{MIR}$  on  $R_{max}$  is shown in the next figure 4 [right panel]. One can see that the method MIR gives convergent values for the penetrability and half-life at increasing of  $R_{max}$ . From such figures we find that *inclusion of the external region into calculations changes the half-life up to 1.5 times* ( $\tau_{min} = 0.20$  sec is the

Parent nucleus				Half-live-values, sec		
Nucleus	$Q$ , MeV	Orbit	$S_p^{\text{th}}$	$\tilde{\tau}_{WKB}$	$\tilde{\tau}_{MIR}$	$\tau_{\text{exp}}$
$^{157}_{73}\text{Ta}_{84}$	0.947	$2s_{1/2}$	0.66	$2.813 \cdot 10^{-1}$	$2.789 \cdot 10^{-1}$	$3.0 \cdot 10^{-1}$
$^{161}_{75}\text{Re}_{84}$	1.214	$2s_{1/2}$	0.59	$2.720 \cdot 10^{-4}$	$2.673 \cdot 10^{-4}$	$3.7 \cdot 10^{-4}$
$^{167}_{77}\text{Ir}_{90}$	1.086	$2s_{1/2}$	0.51	$5.85 \cdot 10^{-2}$	$5.84 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$
$^{109}_{53}\text{I}_{56}$	0.829	$1d_{5/2}$	0.76	$3.937 \cdot 10^{-6}$	$3.992 \cdot 10^{-6}$	$1.0 \cdot 10^{-4}$
$^{112}_{55}\text{Cm}_{57}$	0.823	$1d_{5/2}$	0.59	$3.526 \cdot 10^{-5}$	$3.539 \cdot 10^{-5}$	$5.0 \cdot 10^{-4}$
$^{147}_{69}\text{Tm}_{78}$	1.132	$1d_{5/2}$	0.79	$7.911 \cdot 10^{-5}$	$7.796 \cdot 10^{-5}$	$3.6 \cdot 10^{-4}$

Table 1. Experimental and calculated half-lives of some proton emitters. Here,  $S_p^{\text{th}}$  is theoretical spectroscopic factor,  $\tau_{WKB}$  is half-life calculated by in the semiclassical approach,  $\tau_{MIR}$  is half-life calculated by in the approach MIR,  $\tilde{\tau}_{WKB} = \tau_{WKB}/S_p^{\text{th}}$ ,  $\tilde{\tau}_{MIR} = \tau_{MIR}/S_p^{\text{th}}$ ,  $\tau_{\text{exp}}$  is experimental data. Values for  $S_p^{\text{th}}$ ,  $\tau_{\text{exp}}$  are used from Table IV in Ref. (Aberg et al., 1997) (see p. 1770); in calculations for each nucleus we use:  $R_{\text{min}} = 0.11$  fm,  $R_{\text{max}} = 250$  fm; number of intervals in region from  $R_{\text{min}}$  to maximum of the barrier is 10000, from maximum of the barrier to  $R_{\text{max}}$  is 10000.

minimal half-life calculated at  $R_3 \leq R_{\text{max}} \leq 250$  fm,  $\tau_{\text{as}} = 0.30$  sec is the half-life calculated at  $R_{\text{max}} = 250$  fm, error =  $\tau_{\text{as}}/\tau_{\text{min}} \approx 1.5$  or 50 percents). So, *error in determination of the penetrability in the semiclassical approach (if to take the external region into account) is expected to be the same as a minimum on such a basis.*

### 3.4.3 Results of calculations of half-lives in our and semiclassical approaches

As we have demonstrated above, the fully quantum calculations of the penetrability of the barrier for the proton decay give us its essential dependence on the starting point. In order to give power of predictions of half-lives calculated by the approach MIR, we need to find recipe able to resolve such uncertainty in calculations of the half-lives. So, we shall introduce the following hypothesis: *we shall assume that in the first stage of the proton decay proton starts to move outside the most probably at the coordinate of minimum of the internal well.* If such a point is located in the minimum of the well, the penetrability turns out to be maximal and half-life minimal. So, as criterion we could use minimum of half-live for the given potential, which has stable basis. We should take into account that the half-lives obtained before are for the proton occupied ground state while it needs to take into account probability that this state is empty in the daughter nucleus. In order to obtain proper values for the half-lives we should divide them on the spectroscopic factor  $S$  (which we take from (Aberg et al., 1997)), and then to compare them with experimental data. Results of such calculations and experimental data for some proton emitters are presented in Table 1. To complete a picture, we add half-lives calculated by the semiclassical approach to these data.

## 4. On four-position operators in quantum field theory, in terms of bilinear operators

In this Section we approach the *relativistic* case, taking into consideration — therefore — the space-time (four-dimensional) “position” operator, starting however with an analysis of the 3-dimensional (spatial) position operator in the simple relativistic case of the Klein-Gordon equation. Actually, this analysis will lead us to tackle already with non-hermitian operators.

Moreover, while performing it, we shall meet the opportunity of introducing bilinear operators, which will be used even more in the next case of the full 4-position operator.

Let us recall that in Sect.2.1 we mentioned that the boundary condition  $E \neq 0$ , therein imposed to guarantee (maximal) hermiticity of the time operator, can be dispensed with just by having recourse to bilinear forms. Namely, by considering the bilinear hermitian operator(Recami, 1976; 1977; Recami et al., 1983)  $\hat{t} = (-i\hbar \overset{\leftrightarrow}{\partial} / \partial E) / 2$ , where the sign  $\leftrightarrow$  is defined through the accompanying equality  $(f, \hat{t}g) = \left(f, -\frac{i\hbar}{2} \frac{\partial}{\partial E} g\right) + \left(-\frac{i\hbar}{2} \frac{\partial}{\partial E} f, g\right)$ .

#### 4.1 The Klein-Gordon case: Three-position operators

The standard position operators, being hermitian and moreover selfadjoint, are known to possess real eigenvalues: i.e., they yield a *point-like* localization. J. M. Jauch showed, however, that a point-like localization would be in contrast with “unimodularity”. In the relativistic case, moreover, phenomena so as the pair production forbid a localization with precision better than one Compton wave-length. The eigenvalues of a realistic position operator  $\hat{z}$  are therefore expected to represent space *regions*, rather than points. This can be obtained only by having recourse to non-hermitian (and therefore non-selfadjoint) position operators  $\hat{z}$  (a priori, one can have recourse either to non-normal operators with commuting components, or to normal operators with non-commuting components). Following, e.g., the ideas in Ref. (Gallardo et al., 1967b;c; Ka’lnay, 1966; Ka’lnay & Toledo, 1967; Olkhovsky et al., 1967), we are going to show that the mean values of the *hermitian (selfadjoint) part* of  $\hat{z}$  will yield a mean (point-like) position (Baldo & Recami, 1969; Recami, 1970), while the mean values of the *anti-hermitian (anti-selfadjoint) part* of  $\hat{z}$  will yield the sizes of the localization region(Olkhovsky & Recami, 1968; 1969).

Let us consider, e.g., the case of relativistic spin-zero particles, in natural units and with metric  $(+ - - -)$ . The position operator  $i \nabla_{\mathbf{p}}$ , is known to be actually non-hermitian, and may be in itself a good candidate for an extended-type position operator. To show this, we want to split (Gallardo et al., 1967b;c; Ka’lnay, 1966; Ka’lnay & Toledo, 1967; Olkhovsky et al., 1967) it into its hermitian and anti-hermitian (or skew-hermitian) parts.

Consider, then, a vector space  $V$  of complex differentiable functions on a 3-dimensional phase-space(Recami et al., 1983) equipped with an inner product defined by

$$(\Psi, \Phi) = \int \frac{d^3 \mathbf{p}}{p_0} \Psi^*(\mathbf{p}) \Phi(\mathbf{p}) \quad (59)$$

quantity  $p_0$  being  $\sqrt{\mathbf{p}^2 + m_0^2}$ . Let the functions in  $V$  satisfy moreover the condition

$$\lim_{R \rightarrow \infty} \int_{S_R} \frac{dS}{p_0} \Psi^*(\mathbf{p}) \Phi(\mathbf{p}) = 0 \quad (60)$$

where the integral is taken over the surface of a sphere of radius  $R$ . If  $U : V \rightarrow V$  is a differential operator of degree one, condition (60) allows a definition of the transpose  $U^T$  by

$$(U^T \Psi, \Phi) = (\Psi, U \Phi) \text{ for all } \Psi, \Phi \in V, \quad (61)$$

where  $U$  is changed into  $U^T$ , or vice-versa, by means of integration by parts.

This allows, further, to introduce a *dual representation*(Recami et al., 1983) ( $U_1, U_2$ ) of a *single* operator  $U_1^T + U_2$  by

$$(U_1\Psi, \Phi) + (\Psi, U_2\Phi) = (\Psi, (U_1^T + U_2)\Phi). \quad (62)$$

With such a dual representation, it is easy to split any operator into its hermitian and anti-hermitian parts

$$(\Psi, U\Phi) = \frac{1}{2} \left( (\Psi, U\Phi) + (U^*\Psi, \Phi) \right) + \frac{1}{2} \left( (\Psi, U\Phi) - (U^*\Psi, \Phi) \right). \quad (63)$$

Here the pair

$$\frac{1}{2} (U^*, U) \equiv \overset{\leftrightarrow}{U}_h, \quad (64)$$

corresponding to  $(1/2) (U + U^{*T})$ , represents the hermitian part, while

$$\frac{1}{2} (-U^*, U) \equiv \overset{\leftrightarrow}{U}_a \quad (65)$$

represents the anti-hermitian part.

Let us apply what precedes to the case of the Klein-Gordon position-operator  $\hat{z} = i \nabla_p$ . When

$$U = i \frac{\partial}{\partial p_j} \quad (66)$$

we have(Olkhovsky & Recami, 1968; 1969)

$$\begin{aligned} \frac{1}{2} (U^*, U) &= \frac{1}{2} \left( -i \frac{\partial}{\partial p_j}, i \frac{\partial}{\partial p_j} \right) \equiv \frac{i}{2} \overset{\leftrightarrow}{\frac{\partial}{\partial p_j}}, \quad (a) \\ \frac{1}{2} (-U^*, U) &= \frac{1}{2} \left( i \frac{\partial}{\partial p_j}, i \frac{\partial}{\partial p_j} \right) \equiv \frac{i}{2} \overset{\leftrightarrow}{\frac{\partial}{\partial p_j}}. \quad (b) \end{aligned} \quad (67)$$

And the corresponding *single* operators turn out to be

$$\begin{aligned} \frac{1}{2} (U + U^{*T}) &= i \frac{\partial}{\partial p_j} - \frac{i}{2} \frac{p_j}{p^2 + m_0^2}, \quad (a) \\ \frac{1}{2} (U - U^{*T}) &= \frac{i}{2} \frac{p_j}{p^2 + m_0^2}. \quad (b) \end{aligned} \quad (68)$$

It is noteworthy(Olkhovsky & Recami, 1968; 1969) that, as we are going to see, operator (68a) is nothing but the usual Newton-Wigner operator, while (68b) can be interpreted (Gallardo et al., 1967b;c; Ka'lnay, 1966; Ka'lnay & Toledo, 1967; Olkhovsky et al., 1967; Olkhovsky & Recami, 1968; 1969; Toller, 1999) as yielding the sizes of the localization-region (an ellipsoid) via its average values over the considered wave-packet.

Let us underline that the previous formalism justifies from the mathematical point of view the treatment presented in papers like (Baldo & Recami, 1969; Gallardo et al., 1967b;c; Ka'lnay, 1966; Ka'lnay & Toledo, 1967; Olkhovsky et al., 1967; Recami, 1970). We can split(Olkhovsky & Recami, 1968; 1969) the operator  $\hat{z}$  into two *bilinear* parts, as follows:

$$\hat{z} = i \nabla_p = \frac{i}{2} \overset{\leftrightarrow}{\nabla}_p + \frac{i}{2} \overset{\leftrightarrow(+)}{\nabla}_p \quad (69)$$

where  $\Psi^* \overset{\leftrightarrow}{\nabla}_p \Phi \equiv \Psi^* \nabla_p \Phi - \Phi \nabla_p \Psi^*$  and  $\Psi^* \overset{\leftrightarrow(+)}{\nabla}_p \Phi \equiv \Psi^* \nabla_p \Phi + \Phi \nabla_p \Psi^*$ , and where we always referred to a suitable (Baldo & Recami, 1969; Gallardo et al., 1967b;c; Ka'lnay, 1966; Ka'lnay & Toledo, 1967; Olkhovsky et al., 1967; Recami, 1970; 1976; 1977; Recami et al., 1983) space of wave packets. Its hermitian part (Baldo & Recami, 1969; Gallardo et al., 1967b;c; Ka'lnay, 1966; Ka'lnay & Toledo, 1967; Olkhovsky et al., 1967; Recami, 1970)

$$\hat{x} = \frac{i}{2} \overset{\leftrightarrow}{\nabla}_p, \quad (70)$$

which was expected to yield an (ordinary) point-like localization, has been derived also by writing explicitly

$$(\Psi, \hat{x} \Phi) = i \int \frac{d^3 p}{p_0} \Psi^*(p) \nabla_p \Phi(p) \quad (71)$$

and imposing hermiticity, i.e., imposing the reality of the diagonal elements. The calculations yield

$$\Re(\Phi, \hat{x} \Phi) = i \int \frac{d^3 p}{p_0} \Phi^*(p) \overset{\leftrightarrow}{\nabla}_p \Phi(p), \quad (72)$$

suggesting to adopt just the Lorentz-invariant quantity (70) as a bilinear hermitian position operator. Then, on integrating by parts (and due to the vanishing of the surface integral), we verify that eq. (70) is equivalent to the ordinary Newton-Wigner operator:

$$\hat{x}_h \equiv \frac{i}{2} \overset{\leftrightarrow}{\nabla}_p \equiv i \nabla_p - \frac{i}{2} \frac{\mathbf{p}}{p^2 + m^2} \equiv \mathbf{N} - \mathbf{W}. \quad (73)$$

We are left with the (bilinear) anti-hermitian part

$$\hat{y} = \frac{i}{2} \overset{\leftrightarrow(+)}{\nabla}_p \quad (74)$$

whose *average values* over the considered state (wave-packet) can be regarded as yielding (Baldo & Recami, 1969; Gallardo et al., 1967b;c; Ka'lnay, 1966; Ka'lnay & Toledo, 1967; Olkhovsky et al., 1967; Recami, 1970; 1976; 1977; Recami et al., 1983) the sizes of an ellipsoidal localization-region.

After the digression associated with eqs.(69)–(74), let us go back to the present formalism, as expressed by eqs.(59)–(68).

In general, the extended-type position operator  $\hat{z}$  will yield

$$\langle \Psi | \hat{z} | \Psi \rangle = (\alpha + \Delta\alpha) + i(\beta + \Delta\beta), \quad (75)$$

where  $\Delta\alpha$  and  $\Delta\beta$  are the mean-errors encountered when measuring the point-like position and the sizes of the localization region, respectively. It is interesting to evaluate the commutators ( $i, j = 1, 2, 3$ ):

$$\left( \frac{i}{2} \overset{\leftrightarrow}{\frac{\partial}{\partial p_i}}, \frac{i}{2} \overset{\leftrightarrow(+)}{\frac{\partial}{\partial p_j}} \right) = \frac{i}{2 p_0^2} \left( \delta_{ij} - \frac{2 p_i p_j}{p_0^2} \right), \quad (76)$$

wherefrom the noticeable “uncertainty correlations” follow:

$$\Delta\alpha_i \Delta\beta_j \geq \frac{1}{4} \left| \left\langle \frac{1}{p_0^2} \left( \delta_{ij} - \frac{2 p_i p_j}{p_0^2} \right) \right\rangle \right|. \quad (77)$$

## 4.2 Four-position operators

It is tempting to propose as *four-position operator* the quantity  $\hat{z}^\mu = \hat{x}^\mu + i\hat{y}^\mu$ , whose hermitian (Lorentz-covariant) part can be written

$$\hat{x}^\mu = -\frac{i}{2} \frac{\overleftrightarrow{\partial}}{\partial p_\mu}, \quad (78)$$

to be associated with its corresponding “operator” in four-momentum space

$$\hat{p}_h^\mu = +\frac{i}{2} \frac{\overleftrightarrow{\partial}}{\partial x_\mu}. \quad (79)$$

Let us recall the proportionality between the 4-momentum operator and the 4-current density operator in the chronotopical space, and then underline the canonical correspondence (in the 4-position and 4-momentum spaces, respectively) between the “operators” (cf. the previous subsection):

$$m_0 \hat{p} \equiv \hat{p}_0 = \frac{i}{2} \frac{\overleftrightarrow{\partial}}{\partial t} \quad (a) \quad (80)$$

$$m_0 \hat{j} \equiv \hat{p} = -\frac{i}{2} \frac{\overleftrightarrow{\partial}}{\partial \mathbf{r}}, \quad (b)$$

and the operators

$$\hat{t} \equiv -\frac{i}{2} \frac{\overleftrightarrow{\partial}}{\partial p_0} \quad (a) \quad (81)$$

$$\hat{x} \equiv \frac{i}{2} \frac{\overleftrightarrow{\partial}}{\partial \mathbf{p}}, \quad (b)$$

where the four-position “operator” (81) can be considered as a 4-current density operator in the energy-impulse space. Analogous considerations can be carried on for the anti-hermitian parts (see the last one of Refs.(Olkhovsky & Recami, 1968; 1969)).

Finally, by recalling the properties of the time operator as a maximal hermitian operator in the non-relativistic case (Sec.2.1), one can see that the relativistic time operator (81a) (for the Klein-Gordon case) is also a selfadjoint bilinear operator for the case of continuous energy spectra, and a (maximal) hermitian linear operator for free particles [due to the presence of the lower limit zero for the kinetic energy, or  $m_0$  for the total energy].

## 5. Decoherence (without instantaneous wave-function collapse<sup>5</sup>)

In this paper we want to show, within the density matrix formalism, that a simple way to get decoherence is through the introduction of a “quantum” of time (or rather of a *chronon*): thus replacing the differential Liouville–von Neumann equation with a finite-difference version of it. In this way, one is given the possibility of using a very simple quantum equation to describe the decoherence effects due to dissipation, and of partially solving the measurement-problem in quantum mechanics (avoiding any recourse to the wave-function collapse). *Namely*, the mere introduction (not of a “time-lattice”, but simply) of the “chronon” allows us to go on from differential to finite-difference equations; and in particular to write down the

<sup>5</sup> This section is developed by Erasmo Recami.

Schrödinger equation (as well as the Liouville–von Neumann equation) in three different ways: “retarded”, “symmetrical”, and “advanced”. One of such three formulations — the *retarded* one — describes in an elementary way a system which is exchanging (and losing) energy with the environment. In its density-matrix version, indeed, it can be easily shown that all non-diagonal terms go to zero very rapidly.

Let us refer, in particular, to the theory of the *chronon* as proposed by P. Caldirola. Let us recall that such an interesting “finite difference” theory, forwards — at the classical level — a solution for the motion of a particle endowed with a non-negligible charge in an external electromagnetic field, overcoming all the known difficulties met by Abraham-Lorentz’s and Dirac’s approaches (and even allowing a clear answer to the question whether a free falling charged particle does or does not emit radiation), and — at the quantum level — yields a remarkable mass spectrum for leptons.

It is easy to compare one another the *new* representations of Quantum Mechanics (QM) resulting from it, in the Schrödinger, Heisenberg and density-operator (Liouville–von Neumann) pictures, respectively.

For each representation, three (*retarded*, *symmetric* and *advanced*) formulations are possible, which refer either to times  $t$  and  $t - \tau_0$ , or to times  $t - \tau_0/2$  and  $t + \tau_0/2$ , or to times  $t$  and  $t + \tau_0$ , respectively. It is interesting to notice that, when the chronon tends to zero, the ordinary QM is obtained as the limiting case of the “symmetric” formulation only; while the “retarded” one does naturally appear to describe QM with friction, i. e., to describe *dissipative* quantum systems (like a particle moving in an absorbing medium).

In this sense, *discretized* QM is much richer than the ordinary one. Here, we want to pay attention to the fact that, when applying the density matrix formalism to the solution of the *measurement problem* in QM, interesting results are met, as, for instance, a natural explication of the “decoherence” due to dissipation: which seem to reveal the power of discretized (in particular, *retarded*) QM.

### 5.1 On discretized Quantum Mechanics

Let us approach our eventual application of the discretization procedures for a possible solution of the measurement problem in Quantum Mechanics, without having to make recourse to the reduction (wave-packet instantaneous collapse) postulate. Namely, let us focus our attention, now, on the consequences for QM of the introduction of a chronon. In QM, time will still be a continuous variable, but the evolution of the system along its world line will be regarded as discontinuous. In analogy with the electron theory in the *non-relativistic* limit, one has to substitute the corresponding finite-difference expression for the time derivatives; e. g.:

$$\frac{df(t)}{dt} = \frac{f(t) - f(t - \Delta t)}{\Delta t}, \quad (82)$$

where proper time is now replaced by the local time  $t$ . The chronon procedure can then be applied to obtain the finite-difference form of the Schrödinger equation. As for the electron case, there are three different ways to perform the discretization, and three “Schrödinger equations” can be obtained:

$$i \frac{\hbar}{\tau} [\Psi(x, t) - \Psi(x, t - \tau)] = \hat{H} \Psi(x, t), \quad (83)$$

$$i \frac{\hbar}{2\tau} [\Psi(x, t + \tau) - \Psi(x, t - \tau)] = \hat{H} \Psi(x, t), \quad (84)$$

$$i \frac{\hbar}{\tau} [\Psi(x, t + \tau) - \Psi(x, t)] = \hat{H} \Psi(x, t), \quad (85)$$

which are, respectively, the *retarded*, *symmetric* and *advanced* Schrödinger equations, all of them transforming into the (same) continuous equation when the fundamental interval of time (that can now be called just  $\tau$ ) goes to zero.

Since the equations are different, the solutions they provide are also fundamentally different. In the classical theory of the electron the symmetric equation represented a non-radiating motion, providing only an approximate description of the motion (without taking into account the effects due to the self fields of the electron). However, in the quantum theory it plays a fundamental role. In the discrete formalism too, the symmetrical equation constitutes the only way to describe a bound non-radiating particle.

However, the solutions of the *retarded* (and *advanced*) equations show a completely different behaviour. For a Hamiltonian explicitly independent of time, the solutions have a general form given by

$$\Psi(x, t) = \left[ 1 + i \frac{\tau}{\hbar} \hat{H} \right]^{-t/\tau} f(x) \quad (86)$$

and, expanding  $f(x)$  in terms of the eigenfunctions of  $\hat{H}$ :

$$\hat{H} u_n(x) = W_n u_n(x), \quad (87)$$

that is, writing  $f(x) = \sum_n c_n u_n(x)$ , with  $\sum_n |c_n|^2 = 1$ , one can obtain that

$$\Psi(x, t) = \sum_n c_n \left[ 1 + i \frac{\tau}{\hbar} W_n \right]^{-t/\tau} u_n(x). \quad (88)$$

The norm of this solution is given by

$$|\Psi(x, t)|^2 = \sum_n |c_n|^2 \exp(-\gamma_n t) \quad (89)$$

with

$$\gamma_n = \frac{1}{\tau} \ln \left( 1 + \frac{\tau^2}{\hbar^2} W_n^2 \right) = \frac{W_n^2}{\hbar^2} \tau + O(\tau^3), \quad (90)$$

where it is apparent that the damping factor depends critically on the value  $\tau$  of the chronon. This *dissipative* behaviour originates from the character of the *retarded* equation; in the case of the electron, the retarded equation possesses intrinsically dissipative solutions, representing a radiating system. The Hamiltonian has the same status as in the ordinary (continuous) case: It is an observable, since it is a hermitian operator and its eigenvectors form a basis of the state space. However, as we have seen, the norm of the state vector is not constant any longer, due to the damping factor. An opposite behaviour is observed for the solutions of the advanced equation, in the sense that they increase exponentially.

## 5.2 Discretized (retarded) Liouville equation, and a solution of the measurement problem: Decoherence from dissipation

Suppose we want to measure the dynamical variable  $R$  of a (microscopic) object  $\mathcal{O}$ , by utilizing a (macroscopic) measuring apparatus  $\mathcal{A}$ .

In the discrete case the interaction is embedded in the Hamiltonian  $\hat{H}$ , with the following consequences. Let us consider the energy representation, where  $|n\rangle$  are the states with defined

energy:  $H|n\rangle = E_n|n\rangle$ ). Since the time evolution operator is a function of the Hamiltonian, and commutes with it, the basis of the energy eigenstates will be a basis also for this operator. The discretized (*retarded*) Liouville-von Neumann equation is

$$\frac{\rho(t) - \rho(t - \tau)}{\tau} = -i\mathcal{L}\rho(t), \quad (91)$$

which reduces to the LvN equation when  $\tau \rightarrow 0$ . The essential point is that, following e.g. a procedure similar to Bonifacio's, one gets in this case a *non-unitary* time-evolution operator:

$$V(t, 0) = \left[ 1 + \frac{i\tau\mathcal{L}}{\hbar} \right]^{-t/\tau}, \quad (92)$$

which, as all non-unitary operators, does not preserve the probabilities associated with each of the energy eigenstates (that make up the expansion of the initial state in such a basis of eigenstates). We are interested in the time instants  $t = k\tau$ , with  $k$  an integer.

Thus, the time-evolution operator (92) takes the initial density operator  $\rho^{\text{in}}$  to a final state for which the non-diagonal terms decay exponentially with time; namely, to

$$\rho_{rs}^{\text{fin}} = \langle r|V(t, 0)|s\rangle = \rho_{rs}^{\text{in}} [1 + iw_{rs}\tau]^{-t/\tau}, \quad (93)$$

where

$$w_{rs} \equiv \frac{1}{\hbar} (E_r - E_s) \equiv \frac{1}{\hbar} (\Delta E)_{rs}. \quad (94)$$

Expression (93) can be written

$$\rho_{rs}(t) = \rho_{rs}(0) e^{-\gamma_{rs}t} e^{-iv_{rs}t}, \quad (95)$$

with

$$\gamma_{rs} \equiv \frac{1}{2\tau} \ln(1 + w_{rs}^2\tau^2), \quad (96)$$

$$v_{rs} \equiv \frac{1}{\tau} \tan^{-1}(w_{rs}\tau). \quad (97)$$

One can observe, indeed, that the non-diagonal terms tend to zero with time, and that the larger the value of  $\tau$ , the faster the decay becomes. Actually, the chronon  $\tau$  is now an interval of time related no longer to a single electron, but to the whole system  $\mathcal{O} + \mathcal{A}$ . If one imagines the time interval  $\tau$  to be linked to the possibility of distinguishing two successive, different states of the system, then  $\tau$  can be significantly larger than  $10^{-23}$  sec, implying an extremely faster damping of the non-diagonal terms of the density operator.

Thus, the reduction to the diagonal form occurs, provided that  $\tau$  possesses a finite value, no matter how small, and provided that  $w_{nm}\tau$ , for every  $n, m$ , is not much smaller than 1; where

$$w_{nm} = (E_n - E_m)/\hbar \quad (98)$$

are the transition frequencies between the different energy eigenstates (the last condition being always satisfied, e. g., for non-bounded systems).

It is essential to notice that *decoherence has been obtained above, without having recourse to any statistical approach, and in particular without assuming any "coarse graining" of time*. The reduction to the diagonal form illustrated by us is a consequence of the discrete (retarded) Liouville-von Neumann equation only, once the inequality  $w_{nm}\tau \ll 1$  is *not* verified.

Moreover, the measurement problem is still controversial even with regard to its mathematical approach: In the simplified formalization introduced above, however, we have not included any consideration beyond those common to the quantum formalism, allowing an as clear as possible recognition of the effects of the introduction of a chronon.

## 6. Non-hermitian Hamiltonians and microscopic quantum dissipation

Various different approaches are known, aimed at getting dissipation — and possibly decoherence — within quantum mechanics. First of all, the simple introduction of a “chronon” (see, e.g., Refs. (Caldirola et al., 1978; Caldirola, 1979; Caldirola & Montaldi, 1979; Caldirola & Lugiato, 1982; Caldirola, 1983; Farias & Recami, 2007)) allows one to go on from differential to finite-difference equations, and in particular to write down the quantum theoretical equations (Schrödinger’s, Liouville-von Neumann’s, etc.) in three different ways: symmetrical, retarded, and advanced. The retarded “Schrödinger” equation describes in a rather simple and natural way a dissipative system, which exchanges (loses) energy with the environment. The corresponding non-unitary time-evolution operator obeys a semigroup law and refers to irreversible processes. The retarded approach furnishes, moreover, an interesting way for solving the “measurement problem” in quantum mechanics, without any need for a wave-function collapse: see Refs.(Bonifacio, 1983; Bonifacio & Caldirola, 1983; Farias & Recami, 2007; Ghirardi & Weber, 1984; Recami & Farias, 2002). The chronon theory can be regarded as a peculiar “coarse grained” description of the time evolution.

Let us stress that it has been shown that the mentioned discrete approach can be replaced with a continuous one, at the price of introducing a *non-hermitian* Hamiltonian: see, e.g., Ref.(Casagrande & Montaldi, 1977).

Further relevant work can be found, for instance, in papers like (Caldirola, 1941; Janussis et al., 1980; 1981a;b; 1982a;b;c; 1984; 1991; 1995; Mignani, 1983) and refs. therein.

Let us add, at this point, that much work is still needed for the description of time irreversibility at the microscopic level. Indeed, various approaches have been proposed, in which new parameters are introduced (regulation or dissipation) into the microscopic dynamics (building a bridge, in a sense, between microscopic structure and macroscopic characteristics). Besides the Caldirola-Kanai(Caldirola, 1941; Kanai, 1948) Hamiltonian

$$\hat{H}_{CK}(t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} e^{-\gamma t} + V(x) e^{\gamma t} \quad (99)$$

(which has been used, e.g., in Ref.(Angelopoulon et al., 1995)), other rather simple approaches, based of course on the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = \hat{H} \Psi(x, t), \quad (100)$$

and adopting a microscopic dissipation defined via a coefficient of extinction  $\gamma$ , are for instance the following ones:

A) *Non-linear (non-hermitian) Hamiltonians*

$$\hat{H}_{nl} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) + \hat{W}, \quad (101)$$

with “potential” operators  $\hat{W}$  of the type:

1. Kostin’s operator (see Ref.(Kostin, 1972)):

$$\hat{W}_K = -\frac{i\hbar}{2m} \left\{ \frac{\ln \Psi}{\Psi^*} - \left\langle \frac{\ln \Psi}{\Psi^*} \right\rangle \right\}; \quad (102)$$

2. Albrecht's operator (see Ref.(Albrecht, 1975)):

$$\hat{W}_A(x) = \langle p \rangle (x - \langle x \rangle), \quad (103)$$

where  $\langle \rangle$  is the averaging produced over  $|\Psi(x)|^2$ ;

3. Ref.(Hasse, 1975):

$$\hat{W}_H(x) = \frac{1}{4} [x - \langle x \rangle, p + \langle p \rangle]_+, \quad (104)$$

where  $[A, B]_+$  is the anticommutator:  $[A, B]_+ = AB + BA$ .

B) *Linear (non-hermitian) Hamiltonians:*

1. Ref.(Gisin, 1982):

$$\hat{H}_G = (1 - i\gamma) \hat{H} + i\gamma \langle \hat{H} \rangle; \quad (105)$$

2. Ref.(Exner, 1983):

$$\hat{H}_E = \hat{H} + i\hat{W}(x) - i\langle \hat{W}(x) \rangle. \quad (106)$$

One might recall also the important, so-called "microscopic models"(Caldeira & Leggett, 1983), even if they are not based on the Schrödinger equation.

All such proposals are to be further investigated, and completed, since they have not been apparently exploited enough, till now. Let us remark, just as an example, that it would be desirable to take into deeper consideration other related phenomena, like the ones associated with the "Hartman effect" (and "generalized Hartman effect") (Aharonov, 2002; Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; 2002; 2004; 2005; Recami, 2004), in the case of tunneling with dissipation: a topic faced in few papers, like (Nimtz et al., 1994; Raciti & Salesi, 1994).

As a small contribution of ours, in the Appendix we present a scheme of iterations (successive approximations) as a possible tool for explicit calculations of wave-functions in the presence of dissipation, by using as an example the simple Albrecht's potential. Our scheme may be useful, in any case, for the investigation of possible violations of the Hartman effect, as well as for analyzing a few irreversible phenomena. See the Appendix.

At last, let us incidentally recall that two generalized Schrödinger equations, introduced by Caldirola (Caldirola, 1941; 1976a;b; 1977) in order to describe two different dissipative processes (behavior of open systems, and the radiation of a charged particle) have been shown — see, e.g., Ref.(Mignani, 1983)) — to possess the same algebraic structure of the Lie-admissible type(Santilli, 1983).

## 7. Some conclusions

1. We have shown that the Time operator (1), hermitian even if non-selfadjoint, works for any quantum collisions or motions, in the case of a continuum energy spectrum, in non-relativistic quantum mechanics and in one-dimensional quantum electrodynamics. The uniqueness of the (maximal) hermitian time operator (1) directly follows from the uniqueness of the Fourier-transformations from the time to the energy representation. The time operator (1) has been fruitfully used in the case, for instance, of tunnelling times (see Refs. (Olkhovsky & Recami, 1992; Olkhovsky et al., 1995; Olkhovsky & Agresti, 1997; Olkhovsky et al., 2004; Olkhovsky & Recami, 2007)), and of nuclear reactions and decays (see Refs. (Olkhovsky, 1984; 1990; 1992; 1998) and also Ref. (Olkhovsky et al., 2006)). We have discussed the advantages of such an approach with respect to POVM's, which is not applicable for three-dimensional particle collisions, within a wide class of Hamiltonians.

The mathematical properties of the present Time operator have actually demonstrated — without introducing any new physical postulates — that *time* can be regarded as a quantum-mechanical observable, at the same degree of other physical quantities (energy, momentum, spatial coordinates,...).

The commutation relations (eqs. (8), (22), (31)) here analyzed, and the uncertainty relations (9), result to be analogous to those known for other pairs of canonically conjugate observables (as for coordinate  $\hat{x}$  and momentum  $\hat{p}_x$ , in the case of eq. (9)). Of course, our new relations do not replace, but merely extend the meaning of the classic time and energy uncertainties, given e.g. in Ref. (Olkhovsky & Recami, 2008). In subsection 2.6, we have studied the properties of Time, as an observable, for quantum-mechanical systems with *discrete* energy spectra.

2. Let us recall that the Time operator (1), and relations (2), (3), (4), (15), (16), have been used for the temporal analysis of nuclear reactions and decays in Refs. (Olkhovsky, 1984; 1990; 1992; 1998); as well as of new phenomena, about time delays-advances in nuclear physics and about time resonances or explosions of highly excited compound nuclei, in Refs. (D'Arrigo et al., 1992; 1993; Olkhovskiy & Doroshko, 1992; Olkhovskiy et al., 2006). Let us also recall that, besides the time operator, other quantities, to which (maximal) hermitian operators correspond, can be analogously regarded as quantum-physical observables: For example, von Neumann himself (Recami, 1976; 1977; Von Neumann, 1955)) considered the case of the momentum operator  $-i\partial/\partial x$  in a semi-space with a rigid wall orthogonal to the  $x$ -axis at  $x = 0$ , or of the radial momentum  $-i\partial/\partial r$ , even if both act on packets defined only over the positive  $x$  or  $r$  axis, respectively.

Subsection 2.5 has been devoted to a new “hamiltonian approach”: namely, to the introduction of the analogue of the “Hamiltonian” for the case of the Time operator.

3. In Section 3, we have proposed a suitable generalization for the Time operator (or, rather, for a Space-Time operator) in relativistic quantum mechanics. For instance, for the Klein-Gordon case, we have shown that the hermitian part of the three-position operator  $\hat{x}$  is nothing but the Newton-Wigner operator, and corresponds to a point-like position; while its anti-hermitian part can be regarded as yielding the sizes of an extended-type (ellipsoidal) localization. When dealing with a 4-position operator, one finds that the Time operator is selfadjoint for unbounded energy spectra, while it is a (maximal) hermitian operator when the kinetic energy, and the total energy, are bounded from below, as for a free particle. We have extensively made recourse, in the latter case, to *bilinear* forms, which dispense with the necessity of eliminating the lower point — corresponding to zero velocity — of the spectra. It would be interesting to proceed to further generalizations of the 3- and 4-position operator for other relativistic cases, and analyze the localization problems associated with Dirac particles, or in 2D and 3D quantum electrodynamics, etc. Work is in progress on time analyses in 2D quantum electrodynamic, for application, e.g., to frustrated (almost total) internal reflections. Further work has still to be done also about the joint consideration of particles and antiparticles.

4. Section 4 has been devoted to the association of unstable states (decaying “resonances”) with the eigenvectors of quasi-hermitian (Agodi et al., 1973; Olkhovskiy et al., 2006; Recami et al., 1983) Hamiltonians.

5. Non-hermitian Hamiltonians, and non-unitary time-evolution operators, can play an important role also in microscopic quantum dissipation (Bonifacio, 1983; Bonifacio & Caldirola, 1983; Caldirola, 1941; Caldirola et al., 1978; Caldirola, 1979; Caldirola & Montaldi, 1979; Caldirola & Lugiato, 1982; Caldirola, 1983; Casagrande & Montaldi, 1977; Farias & Recami, 2007; Ghirardi & Weber, 1984; Janussis et al., 1980; 1981a,b; 1982a,b,c; 1984; 1991; 1995; Mignani, 1983; Recami & Farias, 2002): namely, in getting decoherence through interaction

with the environment (Farias & Recami, 2007; Recami & Farias, 2002). This topic is touched in Section 5; together with questions related with collisions in absorbing media. In particular, in Sec.5 we mention also the case of the optical model in nuclear physics; without forgetting that non-hermitian operators show up even in the case of tunnelling — e. g., in fission phenomena — with quantum dissipation, and of quantum friction. As to the former topic of microscopic quantum dissipation, among the many approaches to quantum irreversibility we have discussed in Sec. 5.2 a possible solution of the quantum measurement problem (via interaction with the environment) by the introduction of finite-difference equations (e. g., in terms of a “chronon”).

6. Let us eventually observe that the “dual equations” (26) and (27) seem to be promising also for the study the initial stage of our cosmos, when tunnellings can take place through the barriers which appear in quantum gravity in the limiting case of quasi-Schrödinger equations.

## 8. Acknowledgements

Part of this paper is based on work performed by one of us in cooperation with P. Smrz, and with A. Agodi, M. Baldo and A. Pennisi di Floristella. Thanks are moreover due for stimulating discussions, or kind collaboration, to Y. Aharonov, S. V. Belchikov, G. Battistoni, R. H. A. Farias, L. Fraietta, A. S. Holevo, V. L. Lyuboshitz, R. Mignani, V. Petrillo, P. Pizzochero, G. Salesi, B. N. Zakhariev, and M. Zamboni-Rached.

## 9. Appendix

### Time-dependent Schrödinger equation with dissipative terms

#### 9.1 Introduction

Let's consider the time-dependent Schrödinger equation:

$$i \frac{\partial}{\partial t} \Psi(x, t) = \left( -\frac{\partial^2}{\partial x^2} + V(x, t) \right) \Psi(x, t), \quad (107)$$

where we put  $\hbar = 1$ . Let us rewrite the time-dependent wave function (WF),  $\Psi(x, t)$  (which can be considered as a wave-packet (WP)), in the form of a Fourier integral:

$$\Psi(x, t) = \int_0^{E_0} g(E) e^{-iEt} \varphi(E, x) dE, \quad (108)$$

where  $\varphi(E, x)$  is the WF component independent of time, and  $g(E)$  is a weight factor. One can choose the function  $g(E)$  to be, e.g., a *Gaussian*:

$$g(E) = A e^{-a^2(k-\bar{k})^2}. \quad (109)$$

Here,  $A$  and  $a$  are constants, and  $\bar{k}$  is the selected value for the impulse, constituting the center of the WP. We substitute the Fourier-expansion (108) of WF into eq. (107). Thus, the l.h.s. of this equation transforms into

$$i \frac{\partial}{\partial t} \Psi(x, t) = \int_0^{E_0} g(E) e^{-iEt} \varphi(E, x) E dE. \quad (110)$$

Afterwards, the r.h.s. of eq. (107) gets transformed into

$$\left(-\frac{\partial^2}{\partial x^2} + V(x, t)\right) \Psi(x, t) = -\int_0^{E_0} g(E) e^{-iEt} \frac{\partial^2 \varphi(E, x)}{\partial x^2} dE + \int_0^{E_0} g(E) V(x, \bar{E}, t) e^{-iEt} \varphi(E, x) dE. \quad (111)$$

Therefore, the whole equation (107) has been transformed into

$$\int_0^{E_0} g(E) e^{-iEt} \varphi(E, x) E dE = -\int_0^{E_0} g(E) e^{-iEt} \frac{\partial^2 \varphi(E, x)}{\partial x^2} dE + \int_0^{E_0} g(E) V(x, \bar{E}, t) e^{-iEt} \varphi(E, x) dE. \quad (112)$$

Let us now apply the inverse Fourier-transformation to this equation. Its left part becomes

$$\begin{aligned} \frac{1}{2\pi} \int dt e^{iE't} \int_0^{E_0} g(E) e^{-iEt} \varphi(E, x) E dE &= \frac{1}{2\pi} \int_0^{E_0} dE E g(E) \varphi(E, x) \int e^{i(E'-E)t} dt = \\ &= \int_0^{E_0} g(E) \varphi(E, x) \delta(E' - E) E dE = g(E') E' \varphi(E', x); \end{aligned} \quad (113)$$

while its right part becomes

$$\begin{aligned} -\frac{1}{2\pi} \int dt e^{iE't} \int_0^{E_0} g(E) e^{-iEt} \frac{\partial^2 \varphi(E, x)}{\partial x^2} dE + \frac{1}{2\pi} \int dt e^{iE't} \int_0^{E_0} g(E) V(x, \bar{E}, t) e^{-iEt} \varphi(E, x) dE &= \\ = -\frac{1}{2\pi} \int_0^{E_0} dE g(E) \frac{\partial^2 \varphi(E, x)}{\partial x^2} \int e^{i(E'-E)t} dt + \frac{1}{2\pi} \int_0^{E_0} dE g(E) \varphi(E, x) \int V(x, \bar{E}, t) e^{i(E'-E)t} dt &= \\ = -g(E') \frac{\partial^2 \varphi(E', x)}{\partial x^2} + \frac{1}{2\pi} \int_0^{E_0} dE g(E) \varphi(E, x) \int V(x, \bar{E}, t) e^{i(E'-E)t} dt. \end{aligned} \quad (114)$$

As a result, we obtain eq. (112) in the form

$$g(E') E' \varphi(E', x) = -g(E') \frac{\partial^2 \varphi(E', x)}{\partial x^2} + \frac{1}{2\pi} \int_0^{E_0} dE g(E) \varphi(E, x) \int V(x, \bar{E}, t) e^{i(E'-E)t} dt. \quad (115)$$

## 9.2 The case of the simple Albreht's potential

Just as an example of a possible potential  $V(x, t)$ , let us choose

$$V(x, t) = V_0(x) + \gamma W_A(x). \quad (116)$$

where  $W_A(x)$  is the simple Albreht's dissipation term. Here,  $\gamma$  is a constant,  $V_0(x)$  is the usual stationary component of  $V(x)$ , and the dissipative component of  $V(x)$  has the form

$$W_A(x) = \langle p \rangle (x - \langle x \rangle), \quad (117)$$

where the averages are fulfilled by integrating over  $x$  by means of the functions  $\Psi^*(x, t)$  and  $\Psi(x, t)$ . For the right part of eq. (117) one gets

$$\begin{aligned} \langle p \rangle &= -i \int dx \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 g(E_1) g(E_2) e^{i(E_1-E_2)t} \varphi^*(E_1, x) \frac{\partial \varphi(E_2, x)}{\partial x}, \\ \langle x \rangle &= \int dx \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 g(E_3) g(E_4) e^{i(E_3-E_4)t} x \varphi^*(E_3, x) \varphi(E_4, x); \end{aligned} \quad (118)$$

so that the total potential  $V(x, t)$  becomes

$$\begin{aligned} V(x, t) &= V_0(x) - i\gamma \int dx_1 \int dx_2 \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 g(E_1) g(E_2) g(E_3) g(E_4) \\ &\times e^{i(E_1-E_2+E_3-E_4)t} (x - x_2) \varphi^*(E_1, x_1) \frac{\partial \varphi(E_2, x_1)}{\partial x_1} \varphi^*(E_3, x_2) \varphi(E_4, x_2). \end{aligned} \quad (119)$$

Taking into account this, we find the second term, in the r.h.s. of eq. (115), to be:

$$\begin{aligned} &\frac{1}{2\pi} \int_0^{E_0} dE g(E) \varphi(E, x) \int V(x, \bar{E}, t) e^{i(E'-E)t} dt = \\ &= \int_0^{E_0} dE g(E) \varphi(E, x) V_0(x) \delta(E' - E) - \\ &- i\gamma \int_0^{E_0} dE g(E) \varphi(E, x) \int dx_1 \int dx_2 \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 g(E_1) g(E_2) g(E_3) g(E_4) \\ &\times (x - x_2) \varphi^*(E_1, x_1) \frac{\partial \varphi(E_2, x_1)}{\partial x_1} \varphi^*(E_3, x_2) \varphi(E_4, x_2) \delta(E' - E + E_1 - E_2 + E_3 - E_4) = \\ &= g(E') \varphi(E', x) V_0(x) - \\ &- i\gamma \int dx_1 \int dx_2 \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 g(E_1) g(E_2) g(E_3) g(E_4) g(E'') \\ &\times (x - x_2) \varphi^*(E_1, x_1) \frac{\partial \varphi(E_2, x_1)}{\partial x_1} \varphi^*(E_3, x_2) \varphi(E_4, x_2) \varphi(E'', x), \end{aligned} \quad (120)$$

where

$$E'' = E' + E_1 - E_2 + E_3 - E_4. \quad (121)$$

As a consequence, the whole eq. (115) gets transformed into

$$\begin{aligned} &g(E') E' \varphi(E', x) = -g(E') \frac{\partial^2 \varphi(E', x)}{\partial x^2} + g(E') \varphi(E', x) V_0(x) \\ &- i\gamma \int dx_1 \int dx_2 \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 g(E_1) g(E_2) g(E_3) g(E_4) g(E'') \\ &\times (x - x_2) \varphi^*(E_1, x_1) \frac{\partial \varphi(E_2, x_1)}{\partial x_1} \varphi^*(E_3, x_2) \varphi(E_4, x_2) \varphi(E'', x) \end{aligned}$$

or (with the change of variables  $E' \rightarrow E$ )

$$\begin{aligned}
 & \left( -\frac{\partial^2}{\partial x^2} + V_0(x) - E \right) \varphi(E, x) = \\
 & = i\gamma \int dx_1 \int dx_2 \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 \frac{g(E_1)g(E_2)g(E_3)g(E_4)g(E'')}{g(E)} \\
 & \quad \times (x - x_2) \varphi^*(E_1, x_1) \frac{\partial \varphi(E_2, x_1)}{\partial x_1} \varphi^*(E_3, x_2) \varphi(E_4, x_2) \varphi(E'', x). \tag{122}
 \end{aligned}$$

We have thus obtained for this case the time-independent Schrödinger equation, by taking however into account dissipation via the parameter  $\gamma$ . Of course, when  $\gamma$  tends to zero, one goes back to the stationary Schrödinger equation.

### 9.3 Method of the successive approximations

Assuming the coefficient  $\gamma$  to be small, one can find the unknown function  $\varphi(x)$  in the simplified form

$$\varphi(x) = \varphi_0(x) + \gamma \varphi_1(x), \tag{123}$$

where as function  $\varphi_0(x)$  it has been used the standard WF of the time-independent Schrödinger equation with potential  $V_0(x)$  and energy  $E_0$ :

$$\left( -\frac{\partial^2}{\partial x^2} + V_0(x) \right) \varphi_0(x) = E_0 \varphi_0(x). \tag{124}$$

Substituting solution (123) into eq. (124), we obtain a new equation containing all the powers  $n$  of  $\gamma$ , namely, the  $\gamma^n$ . Let us confine ourselves, however, to write down this equation with accuracy up to  $\gamma^1$  only:

$$\begin{aligned}
 & \left( -\frac{\partial^2}{\partial x^2} + V_0(x) - E \right) \left( \varphi_0(E, x) + \gamma \varphi_1(E, x) \right) = \\
 & = i\gamma \int dx_1 \int dx_2 \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 \frac{g(E_1)g(E_2)g(E_3)g(E_4)g(E'')}{g(E)} \\
 & \quad \times (x - x_2) \varphi_0^*(E_1, x_1) \frac{\partial \varphi_0(E_2, x_1)}{\partial x_1} \varphi_0^*(E_3, x_2) \varphi_0(E_4, x_2) \varphi_0(E'', x), \tag{125}
 \end{aligned}$$

where the unknown  $\varphi_1(x)$  does not appear any longer, of course, into the r.h.s. of this equation. Taking  $E = E_0$ , we can rewrite in eq. (125), separately, the various terms with different powers of  $\gamma$ . When limiting ourselves to  $n = 0, 1$ , we obtain

$$\begin{aligned}
 \gamma^0: & \quad \left( -\frac{\partial^2}{\partial x^2} + V_0(x) - E_0 \right) \varphi_0(E_0, x) = 0, \\
 \gamma^1: & \quad \left( -\frac{\partial^2}{\partial x^2} + V_0(x) - E_0 \right) \varphi_1(E_0, x) = \\
 & = i\gamma \int dx_1 \int dx_2 \int_0^{E_0} dE_1 \int_0^{E_0} dE_2 \int_0^{E_0} dE_3 \int_0^{E_0} dE_4 \frac{g(E_1)g(E_2)g(E_3)g(E_4)g(E'')}{g(E_0)} \\
 & \quad \times (x - x_2) \varphi_0^*(E_1, x_1) \frac{\partial \varphi_0(E_2, x_1)}{\partial x_1} \varphi_0^*(E_3, x_2) \varphi_0(E_4, x_2) \varphi_0(E'', x), \tag{126}
 \end{aligned}$$

where

$$E'' = E_0 + E_1 - E_2 + E_3 - E_4. \quad (127)$$

The first equation holds when dissipation is absent. The second equation determines the unknown function  $\varphi_1$  in terms of the given  $\varphi_0$ : It results to be an ordinary differential equation of the second order, that can be solved by the ordinary numerical methods.

## 10. References

- Aberg, S., Semmes, P. B. & Nazarewicz, W. (1997). *Phys. Rev.* Vol. C56: 1762–1773.
- Agodi, A. & Eberle, E. (1960). *Nuovo Cimento* Vol. 18: 718.
- Agodi, A., Catara, F. & Di Toro, M. (1968). *Annals of Physics* Vol. 49: 445.
- Agodi, A. (1969). *Theory of Nuclear Structures (Trieste Lectures)*: p. 879.
- Agodi, A., Baldo, M. & Recami, E. (1973). *Annals of Physics* Vol. 77: 157.
- Aharonov, Y. & Bohm, D. (1961). *Phys. Rev.* Vol. A122: 1649.
- Aharonov, Y., Oppenhem, J., Popescu, S., Reznik, B. & Unruh, W. (1998). *Phys. Rev.* Vol. A57: 4130.
- Aharonov, Y., Erez, N. & Resnik, B. (2002). *Phys. Rev.* Vol. A65: 052124.
- Akhiezer, N. I. & Glazman, I. M. (1981). *The Theory of Linear Operators in Hilbert Space* (Boston: Pitman, Mass.).
- Akhiezer, A. I. & Berestezky, V. B. (1959). *Quantum Electrodynamics* [in Russian] (Moscow: Fizmatgiz).
- Albrecht, K. (1975). *Phys. Lett.* Vol. B56: 127.
- Anderson, A. (1989). Multiple scattering approach to one-dimensional potential problems, *Am. Journ. Phys.* Vol. Vol. 57 (No. 3): 230–235.
- Angelopoulou, P., et al. (1995). *Int. J. Mod. Phys.* Vol. B9: 2083.
- Atmanspacher, H. & Amann, A. (1998). *Internat. J. Theor. Phys.* Vol. 37: 629.
- Baldo, M. & Recami, E. (1969). *Lett. Nuovo Cim.* Vol. 2: 613.
- Blanchard, P. & Jadczyk, A. (1996). *Helv. Phys. Acta* Vol. 69: 613.
- Bonifacio, R. (1983). *Lett. N. Cim.* Vol. 37: 481.
- Bonifacio, R., & Caldirola, P., (1983). *Lett. N. Cim.* Vol. 38: 615.
- Buck, B., Merchant, A. C. & Perez, S. M. (1993). Half-lives of favored alpha decays from nuclear ground states, *At. Dat. Nucl. Dat. Tabl.* Vol. Vol. 54 (No. 1): 53–74.
- Busch, P., Grabowski, M. & Lahti, P. J. (1994). *Phys. Lett.* Vol. A191: 357.
- Caldeira, A. & Leggett, A. (1983). *Annals of Physics* Vol. 149: 374.
- Caldirola, P. (1941). *Nuovo Cimento* Vol. 18: 393.
- Caldirola, P. (1976). *Lett. N. Cim.* Vol. 16: 151.
- Caldirola, P. (1976). *Lett. N. Cim.* Vol. 17: 461.
- Caldirola, P. (1977). *Lett. N. Cim.* Vol. 18: 465.
- Caldirola, P., Casati, G. & Prosperetti, A. (1978). *Nuovo Cimento* Vol. A43: 127.
- Caldirola, P. (1979). *Rivista N. Cim.* Vol. 2: issue no. 13.
- Caldirola, P. & Montaldi, E. (1979). *Nuovo Cimento* Vol. B53: 291.
- Caldirola, P. & Lugiato, L. (1982). *Physica* Vol. A116: 248.
- Caldirola, P. (1983). Dissipation in quantum theory (40 years of research), *Hadronic Journal* Vol. 6: 1400–1433.
- Cardone, F., Maidanyuk, S. P., Mignani, R. & Olkhovsky, V. S. (2006). Multiple internal reflections during particle and photon tunneling, *Found. Phys. Lett.* Vol. Vol. 19 (No. 5): 441–457.
- Carleman, T. (1923). *Sur les Equations Integrales Anoyau R el et Sym etrique* (Uppsala).

- Casagrande, F. & Montaldi, E. (1977). *Nuovo Cimento* Vol. A40: 369.
- D'Arrigo, A., Doroshko, N. L., Eremin, N. V., Olkhovsky, V. S. et al. (1992). *Nucl. Phys.* Vol. A549: 375–386.
- D'Arrigo, A., Doroshko, N. L., Eremin, N. V., Olkhovsky, V. S. et al. (1993). *Nucl. Phys.* Vol. A564: 217–226.
- D ter Haar (1971). *Elements of Hamiltonian Mechanics* (Oxford).
- Delgado, V. (1999). *Phys. Rev.* Vol. A59: 1010.
- Denisov, V. Yu. & Ikezoe, H. (2005). Alpha-nucleus potential for alpha-decay and sub-barrier fusion, *Phys. Rev.* Vol. C72: 064613. URL: nucl-th/0510082.
- Egusquiza, I. L. & Muga, J. G. (1999). *Phys. Rev.* Vol. A61: 012104.
- Exner, P. (1983). *J. Math. Phys.* Vol. 24: 1129.
- Farias, A. R. H. & Recami, E. (2007). Introduction of a quantum of time (“chronon”) & its consequences for quantum mechanics, URL: quant-ph/97060509v3.
- Fermor, J. H. (1966). Quantum-mechanical tunneling, *Am. Journ. Phys.* Vol. 34: 1168–1170.
- Feshbach, H., Porter, C. E. & Weisskopf, V. F. (1954). *Phys. Rev.* Vol. 96: 448.
- Galitsky, V. M. & Migdal, A. B. (1958). *Sov. Phys. JETP* Vol. 34: 96.
- Gallardo, J. A., Ka'lnay, A. J., Stec, B. A. & Toledo, B. P. (1967). *Nuovo Cimento* Vol. A48: 1008.
- Gallardo, J. A., Ka'lnay, A. J., Stec, B. A. & Toledo, B. P. (1967). *Nuovo Cimento* Vol. A49: 393.
- Gallardo, J. A., Ka'lnay, A. J. & Risenberg, S. H. (1967). *Phys. Rev.* Vol. 158: 1484.
- Ghirardi, G. C. & Weber, T. (1984). *Lett. N. Cim.* Vol. 39: 157.
- Giannitrapani, R. (1997). *Int. J. Theor. Phys.* Vol. 36: 1575.
- Gisin, N. (1982). *Physica* Vol. A111: 364.
- Gózdź, A. & Dębicki, M. (2007). *Phys. At. Nuclei* Vol. 70: 529.
- Grot, N., Rovelli, C. & Tate, R. S. (1996). *Phys. Rev.* Vol. A54: 4676.
- Hasse, R. (1975). *J. Math. Phys.* Vol. 16: 2005.
- Hodgson, P. I. (1963). *The Optical Model of Elastic Scattering* (Clarendon Press; Oxford, UK).
- Holevo, A. S. (1978). *Rep. Math. Phys.* Vol. 13: 379.
- Holevo, A. S. (1982). *Probabilistic & Statistical Aspects of Quantum Theory* (Amsterdam).
- Janussis, A., et al. (1980). *Lett. N. Cim.* Vol. 29: 259.
- Janussis, A., et al. (1981). *Lett. N. Cim.* Vol. 30: 289.
- Janussis, A., et al. (1981). *Lett. N. Cim.* Vol. 31: 533.
- Janussis, A., et al. (1982). *Lett. N. Cim.* Vol. 34: 571.
- Janussis, A., et al. (1982). *Lett. N. Cim.* Vol. 35: 485.
- Janussis, A., et al. (1982). *Nuovo Cimento* Vol. B67: 161.
- Janussis, A., et al. (1984). *Lett. N. Cim.* Vol. 39: 75.
- Janussis, A., Brodimas, G. & Mignani, R. (1991). *J. Phys. A: Math. Gen.* Vol. 24: L775.
- Janussis, A., Leodaris, A. & Mignani, R. (1995). *Phys. Lett.* Vol. A197: 187.
- Ka'lnay, A. J. (1966). *Boletín del IMAF (Co'rdoba)* Vol. 2: 11.
- Ka'lnay, A. J. & Toledo, B. P. (1967). *Nuovo Cimento* Vol. A48: 997.
- Kanai, E. (1948). *Progr. Theor. Phys.* Vol. 3: 440.
- Kijowski, J. (1999). *Phys. Rev.* Vol. A59: 897.
- Kobe, D. H. & Aguilera-Navarro, V. C. (1994). *Phys. Rev.* Vol. A50: 933.
- Kocha'nski, P. & Wo'dkievicz, K. (1999). *Phys. Rev.* Vol. A60: 2689.
- Koning, A. G. & Delaroche, J. P. (2003). *Nucl. Phys.* Vol. A713: 231.
- Kostin, M. (1972). *J. Chem. Phys.* Vol. 57: 358.
- Kunieda, S., Chiba, S., Shilata, K., Ichihara, A. & Suchovitski, E. Sh. (2007). *J. Nucl. Sc. & Techn.* Vol. 44: 838.

- Landau, L. D. & Lifshitz, E. M. *Quantum mechanics, course of Theoretical Physics*, Vol. 3 (Nauka, Mockva, 1989), p. 768 — [in Russian; eng. variant: Oxford, Uk, Pergamon, 1982].
- Leo'n, J. (1997). *J. Phys.* Vol. A30: 4791.
- Maydanyuk, S. P., Olkhovsky, V. S. & Zaichenko, A. K. (2002). The method of multiple internal reflections in description of tunneling evolution of nonrelativistic particles and photons, *Journ. Phys. Stud.* Vol. 6 (No. 1): 1–16. URL: nucl-th/0407108.
- Maydanyuk, S. P., Olkhovsky, V. S. & Belchikov, S. V. (2002). The method of multiple internal reflections in description of nuclear decay, *Probl. At. Sci. Tech. (Voprosi atomnoi nauki i tehniki, RFNC-VNIIEF, Sarov, Russia)* Vol. 1: 16–19. URL: nucl-th/0409037.
- Maydanyuk, S. P. (2003). Time analysis of tunneling processes in nuclear collisions and decays, *Ph. D. dissertation* (Supervisor: Prof. V. S. Olkhovsky, Kiev), p. 147 [in Ukrainian].
- Maydanyuk, S. P. & Belchikov, S. V. (2011). Problem of nuclear decay by proton emission in fully quantum consideration: Calculations of penetrability and role of boundary conditions, *Journ. Mod. Phys.* Vol. 2: 572–585. URL: <http://www.SciRP.org/journal/jmp>.
- McVoy, K. W., Heller, L. & Bolsterli, M. (1967). Optical analysis of potential well resonances, *Rev. Mod. Phys.* Vol. 39 (No. 1): 245–258.
- Mignani, R. (1983). *Lett. N. Cim.* Vol. 38: 169.
- Moldauer, P. A. (1963). *Nucl. Phys.* Vol. 47: 65.
- Muga, J., Palao, J. & Leavens, C. (1999). *Phys. Lett.* Vol. A253: 21.
- Naimark, M. A. (1940). *Izvestiya Akademii Nauk SSSR, seriya matematicheskaya* [in Russian, partially in English] Vol. 4: 277.
- Naimark, M. A. (1943). *Izvestiya Akademii Nauk SSSR, seriya Matematika* Vol. 7: 237–244.
- Nikolaiev, M. V. & Olkhovsky, V. S. (1977). *Theor. Mathem. Phys.* Vol. 31: 418.
- Nimtz, G., Spieker, H. & Brodowsky, M. (1994). *J. Phys.-I (France)* Vol. 4: 1379.
- Gallardo, J. A., Ka'lnay, A. J., Stec, B. A. & Toledo, B. P. (1967). *Nuovo Cimento* Vol. A48: 1008.
- Olkhovsky, V. S. & Recami, E. (1968). *Nuovo Cimento* Vol. A53: 610.
- Olkhovsky, V. S. & Recami, E. (1969). *Nuovo Cimento* Vol. A63: 814.
- Olkhovsky, V. S. & Recami, E. (1970). *Lett. Nuovo Cim. (First Series)* Vol. 4: 1165.
- Olkhovsky, V. S. (1973). *Ukrainskiy Fiz. Zhurnal* [in Ukrainian & Russian] Vol. 18: 1910.
- Olkhovsky, V. S., Recami, E. & Gerasimchuk, A. (1974). *Nuovo Cimento* Vol. A22: 263.
- Olkhovsky, V. S. (1975). *Theor. Mathem. Phys.* Vol. 20: 774.
- Olkhovsky, V. S. & Zaichenko, A. K. (1981). *Nuovo Cimento* Vol. A63: 155.
- Olkhovsky, V. S. (1984). *Sov. J. Part. Nucl.* Vol. 15: 130.
- Olkhovsky, V. S. (1990). *Nukleonika* Vol. 35: 99.
- Olkhovsky, V. S. (1992). *Atti Accademia Peloritana dei Pericolanti, Sci., Mat. e Nat.* Vol. 70: 21 & 135.
- Olkhovsky, V. S. & Recami, E. (1992). *Phys. Rep.* Vol. 214: 339.
- Olkhovsky, V. S. & Doroshko, N. L. (1992). *Europhys. Lett.* Vol. 18 (No. 6): 483–486.
- Olkhovsky, V. S., Recami, E., Raciti, F. & Zaichenko, A. K. (1995). *J. de Phys. (France) I* Vol. 5: 1351.
- Olkhovsky, V. S. & Agresti, A. (1997). *Proc. of the Adriatico Research. Conf. on Tunneling & its Implications* (World Sci.): 327–355.
- Olkhovsky, V. S. (1998). *Mysteries, Puzzles & Paradoxes in Quantum Mechanics* (ed. by R. Bonifaccio, AIP): 272–276.
- Olkhovsky, V. S. & Maydanyuk, S. P. (2000). Method of multiple internal reflections in description of tunneling evolution through barriers, *Ukr. Phys. Journ.* Vol. 45 (No. 10): 1262–1269. URL: nucl-th/0406035.

- Olkhovsky, V. S., Recami, E. & Salesi, G. (2002). Tunneling through two successive barriers & the Hartman (Superluminal) effect, *Europhysics Letters* Vol. 57: 879–884. URL: [quant-ph/0002022](http://quant-ph/0002022).
- Olkhovsky, V. S., Recami, E. & Jakiel, J. (2004). *Phys. Rep.* Vol. 398: 133.
- Olkhovsky, V. S., Recami, E. & Zaichenko, A. K. (2005). Resonant & non-resonant tunneling through a double barrier, *Europhysics Letters* Vol. 70: 712–718. URL: [quant-th/0410128](http://quant-th/0410128).
- Olkhovsky, V. S., Dolinska, M. E. & Omelchenko, S. A. (2006). *Central European Journal of Physics* Vol. 4 (No. 2): 1–18.
- Olkhovsky, V. S. & Recami, E. (2007). *Intern. J. Mod. Phys* Vol. A22: 5063.
- Olkhovsky, V. S. & Recami, E. (2008). *Int. J. Mod. Phys.* Vol. B22: 1877.
- Olkhovsky, V. S. (2009). Time as a quantum observable, Canonically conjugated to energy, and foundations of self-consistent Time Analysis of Quantum Processes, *Advances in Math. Phys.* Vol. 2009 83 p.
- Olkhovsky, V. S. (2011). On time as a quantum observable canonically conjugate to energy, *Physics - Uspekhi* Vol. 54 (No. 8): 829–835.
- Pauli, W. (1926). *Handbuch der Physik* vol. 5/1 (Berlin: Ed. by S. Fluegge): p. 60.
- Pauli, W. (1980). *General Principles of Quantum Theory* (Berlin: Springer).
- Raciti, F. & Salesi, G. (1994). *J. Phys.-I (France)* Vol. 4: 1783.
- Recami, E. (1970). *Atti Accad. Naz. Lincei (Roma)* Vol. 49: 77.
- Recami, E. (1976). *An operator for the observable time in Recent Developments in Relativistic Q.F.T. & Its Application (Proc. of the XIII Karpacz Winter School on Theor. Phys.)*, vol.II, ed. by W.Karwowski (Wroclaw Univ.Press; Wroclaw), pp.251-265.
- Recami, E. (1977). *A time operator & the time-energy uncertainty relation in The Uncertainty Principle & Foundation of Quantum Mechanics*, ed. by C.Price & S. Chissik (London: J. Wiley) Chap.4, pp.21–28.
- Recami, E., Rodrigues, W. A. & Smrz, P. (1983). *Hadronic Journal* Vol. 6: 1773–1789.
- Recami, E. & Farias, A. R. H. (2002). A simple quantum equation for decoherence and dissipation, Report NSF-ITP-02-62 (KIPT, UCSB; Santa Barbara, CA). URL: [quant-ph/97060509v3](http://quant-ph/97060509v3).
- Recami, E. (2004). Superluminal tunneling through successive barriers. Does QM predict infinite group-velocities? *Journal of Modern Optics* Vol. 51: 913–923.
- Rosenbaum, D. M. (1969). *J. Math. Phys.* Vol. 10: 1127.
- Santilli, R. M. (1979). *Hadronic Journal* Vol. 2: 1460.
- Santilli, R. M. (1983). *Foundations of Theoretical Mechanics, Vol.II: Birkhoffian Generalization of Hamiltonian Mechanics* (Springer; Berlin).
- Schweber, S. (1961). *An Introduction to Relativistic Quantum Field Theory* chapter 5.3 (Row, Peterson and Co).
- Schwinger, J. S. (1960). *Annals of Physics* Vol. 9: 169.
- Sobiczewski, A. & Pomorski, K. (2007). Description of structure and properties of superheavy nuclei, *Prog. Part. Nucl. Phys.* Vol. 58: 292–349.
- Srinivas, M. D. & Vijayalakshmi, R. (1981). *Pramana J. Phys.* Vol. 16: 173.
- Stone, M. H. (1930). *Proc. Nat. Acad. Sci. USA* Vol. 16: issue no. 1.
- Toller, M. (1999). *Phys. Rev.* Vol. A59: 960.
- Von Neumann, J. (1955). *Mathematical foundations of quantum mechanics* (Princeton Univ. Press, Princeton, N.J.).
- Wang, Zhi-Yong & Xiong, Cai-Dong (2007). *Annals of Physics* Vol. 322: 2304.

# Order of Time Derivatives in Quantum-Mechanical Equations

Jan Jerzy Sławianowski

*Institute of Fundamental Technological Research, Polish Academy of Sciences  
5<sup>B</sup>, Pawińskiego str., 02-106 Warsaw, Poland*

## 1. Introduction

The problem is rather old, almost as quantum mechanics itself. The first, primary idea of Schrödinger was the relativistic one, with the d'Alembert operator on the left-hand side of quantum-mechanical equation, so, with the second-order time derivatives. Unfortunately, it turned out that the following results were in a rather clear contradiction with experimental data, although some kind of compatibility did exist. Schrödinger felt disappointed and at least temporarily he rejected his primary equation. Later on, basic on the idea of Lagrange-Hamilton optical-mechanical analogy and on certain de Broglie ideas, he in a sense derived his famous equation which seemed to remain in a beautiful agreement with spectroscopic data and was approximately compatible with the Bohr-Sommerfeld quantization rules. Nevertheless, it was of course drastically incompatible with the relativistic idea of Poincare symmetry. But, after the fall of the primary substantial interpretation by Schrödinger, it was compatible with the Born statistical interpretation of his formalism, and with the corresponding continuity equation for the probabilistic density (Veltman, 2003).

Later on history was rather complicated. Dirac formulated his relativistic quantum theory of electrons based on first-order space-time derivatives of multicomponent wave functions. The multicomponent character of waves had to do obviously with the particle spin. It was also understood that the relativistic velocity-dependence of the electron mass and the spin phenomena act in opposite directions, and because of this non-relativistic Schrödinger equation seemed to be better than his second order equation, rediscovered later on by Klein and Gordon. The formalism of quantum field theory rehabilitated the Klein-Gordon equation, i.e., primaevally the relativistic Schrödinger equation, as one describing some physics. And, let us also mention that the field-theoretic approach based on the Pauli exclusion principle removed certain problems with the quantum-mechanical Dirac equation for a single electron. And certain inconsistencies on one-particle relativistic theory were resolved. The only, and fundamental inadequacy which remained, was one connected with the essential non-linearity of the quantum field-theoretic equations for the field operators and the resulting interpretation difficulties. Nevertheless, they were in a sense solvable on the basis of renormalization procedure.

But in spite of everything said, the problem is still alive. There are certain not completely clear facts within the framework of field theory based on the Dirac-Clifford paradigm of first-order differential equations of quantum mechanics with  $\hbar/2$ -spin. One can show that they become more clear when we assume that in a sense some second-order equations are primary and the first-order ones are some approximations valid for slowly-varying fields. There are also

some arguments from geometrodynamics and gauge theories, one of fundamental methods in modern fields theory. It is known that there are certain disadvantages in geometrodynamical gauge models based on the Poincare group as a gauge group, in spite of certain correct results following from that approach. It seems that the main reason is the fact that Poincare group is not semi-simple. The best way out seems to take the simplest semisimple extension of this group, namely the conformal group of Minkowski space. It must be stressed that this group does not act in the space-time manifold, which in geometrodynamics is a general non-flat manifold of a dynamical structure. Instead, it acts as a purely internal group operating with internal degrees of freedom of our matter fields. To be more precise, instead its universal conformal group  $CO(1,3)$ , one should use its universal covering group  $SU(2,2)$  of pseudo-unitary mappings with the signature  $(+, +, -, -)$ , acting in the target spaces of matter fields and on the gauge connection components. The primary field equations are differential ones of the second-order in matter fields. After the careful rewriting in terms of the basic elements of Lie algebra  $SU(2,2)'$ , the internal group rules both matter and geometry (gravitation). The use of conformal group is interesting in itself. It is the smallest semi-simple group containing Poincare group. It is also the largest group which in the geometrically Minkowskian formulation preserves the family of relativistic uniformly accelerated motions (described by the flat time-like hyperboles). It turns out that there are interesting aspects of this approach, having some correspondence with the usual gravitation theory and with generally-relativistic spinor fields. In the specially-relativistic limit, the theory seems to predict the existence of pairs of fundamental quarks and leptons, just as it is really in Nature. We mean here the quark pairs  $(u, d)$ ,  $(c, s)$ ,  $(t, b)$  and those of leptons  $(\nu_e, e)$ ,  $(\nu_\mu, \mu)$ ,  $(\nu_\tau, \tau)$ . It is interesting that in this limit the fermion fields are described by the Klein-Gordon-Dirac equation combining the Klein-Gordon and Dirac operators, and that in this limit the Dirac behaviour of fields seems to be more remarkable. There are certain interesting facts concerning the spin-statistics problem. It seems that on the very fundamental level some fermion-boson mixing may appear, or that the two possibilities will be unified by some quite new approach. This framework seems to be related, in a rather unexpected way, to another aspect of the problem of the order of time derivatives in quantum mechanics. Namely, certain quite interesting aspects of the quantum-mechanical and quantum field-theoretic problems appears, when one temporarily forgets about the quantum nature of equations, and considers them simply as some Hamiltonian systems of mathematical physics. Certain primary ideas concerning this problem were formulated in our papers a few years ago (Sławianowski & Kovalchuk, 2002; 2008; 2010; Sławianowski et al., 2004; 2005). There are some arguments which seem to show that there is some so-to-speak inadequacy in the first-order Schrödinger equation. In any case, the second-order corrections seem to be just admissible if not desirable. There are also certain indications for that from the theory of stochastic processes. This approach has certain common points with the former gauge-theoretic one. Namely, once using the language of Hamiltonian dynamics (perhaps infinite-dimensional one) we do not feel any longer the usual reluctance of quantum people to the idea of non linearity. In particular, it turns out that the dynamical scalar product, i.e., one non-constant, but satisfying a closed system of equations with the wave amplitudes, is a natural constituent of the approach. The theory becomes then essentially nonlinear. Essentially, i.e., in such a way that nonlinearity is not an accidental term imposed onto some basic linear background. Everything is then nonlinear in the zeroth-order approximation. Nonlinearity is an essential feature, similar to one used in non-Abelian gauge theories and may be perhaps responsible for the decoherence and measurement paradoxes.

We have formulated some arguments in favour of  $SU(2,2)$  as a fundamental gauge group. Let us mention, incidentally, that this simply provokes the next question: Why the subgroup  $SU(2,2) \subset GL(4, \mathbb{C})$  but not just the whole  $GL(4, \mathbb{C})$ ? The latter group appears in a natural way as the structure group of the principal fibre bundle of the complexification of the usual bundle of frames over the four-dimensional space-time manifold. Obviously, it preserves the signature of sesquilinear forms, nevertheless changing them otherwise. Therefore, the bispinor sesquilinear Hermitian form  $G$  of signature  $(+, +, -, -)$  becomes an a priori free Hermitian form, the signature however being in a sense an integral of motion.

## 2. On the track of the scalar Klein-Gordon-Dirac formalism

The generally-relativistic Lagrangian of the Dirac field is given by the expression:

$$L = \frac{i}{2} e^\mu{}_A \gamma^{Ar}{}_s \left( \tilde{\Psi}_r D_\mu \Psi^s - D_\mu \tilde{\Psi}_r \Psi^s \right) \sqrt{|g|} - m \tilde{\Psi}_r \Psi^r \sqrt{|g|} \quad (1)$$

with the following meaning of symbols:

(a)  $\tilde{\Psi}$  denote the Dirac-conjugation of  $\Psi$ ,

$$\tilde{\Psi}_r = \overline{\Psi^s} G_{sr}, \quad (2)$$

where  $G$  denotes the Dirac-conjugation form of mass, i.e., sesquilinear Hermitian form of the natural signature  $(+, +, -, -)$ . If the Finkelstein-Penrose-Weizsäcker-van der Waerden point of view on the two-component spinors is accepted, then  $G$  is intrinsic, because the  $\mathbb{C}^4$ -space is then expressed as the Cartesian product of two mutually antidual copies of  $\mathbb{C}^2$ . Without this point of view, analytically  $\Psi$  is  $\mathbb{C}^4$ -valued.

(b) Dirac matrices  $\gamma^A$  satisfy the following anticommutation rules:

$$\{ \gamma^A, \gamma^B \} = \gamma^A \gamma^B + \gamma^B \gamma^A = 2\eta^{AB} I_4, \quad (3)$$

$$[ \gamma^{AB} ] = \text{diag} (1, -1, -1, -1). \quad (4)$$

Besides,  $\gamma^A$  are Hermitian with respect to  $\Gamma$ :

$$\Gamma^A{}_{\bar{r}s} = \overline{\Gamma^A{}_{s\bar{r}}} = \overline{\Gamma^A{}_{s\bar{r}}} = G_{\bar{r}z} \gamma^{Az}{}_s. \quad (5)$$

(c) The quantities  $e^\mu{}_A$  are components of the tetrad field. Its dual cotetrad  $e^A{}_\mu$  is analytically given by the reciprocal expression

$$e^A{}_\mu e^\mu{}_B = \delta^A{}_B. \quad (6)$$

(d) The metric tensor  $g_{\mu\nu}$  is built of  $e^A{}_\mu$  in a quadratic way:

$$g_{\mu\nu} = \eta_{AB} e^A{}_\mu e^B{}_\nu, \quad [\eta_{AB}] = \text{diag} (1, -1, -1, -1). \quad (7)$$

The Greek and capital Latin indices are shifted with the help of  $g_{\mu\nu}$  and  $\eta_{AB}$ .

(e) The operation  $D_\mu$  symbolizes the covariant differentiation of bispinors. It is given by the following sequence of expressions:

$$\Gamma^\alpha{}_{\beta\mu} = e^\alpha{}_A \Gamma^A{}_{B\mu} e^B{}_\beta + e^\alpha{}_A e^A{}_{\beta,\mu}, \quad (8)$$

$$\Gamma^A{}_{B\mu} = \frac{1}{2} \text{Tr} \left( \gamma^A \omega_\mu \gamma_B \right), \quad (9)$$

$$\omega_\mu = \frac{1}{2} \Gamma_{LK\mu} \Sigma^{LK} = \frac{1}{2} \eta_{LM} \Gamma^M_{K\mu} \Sigma^{LK}, \quad (10)$$

$$\Sigma^{LK} = \frac{1}{4} \left( \gamma^L \gamma^K - \gamma^K \gamma^L \right), \quad (11)$$

where, obviously, the following identities hold:

$$\eta_{AC} \Gamma^C_{B\mu} + \eta_{BC} \Gamma^C_{A\mu} = 0, \quad \nabla^\Gamma_\mu \delta_{\alpha\beta} = 0. \quad (12)$$

This means that the following is satisfied:

$$\Gamma^\alpha_{\beta\mu} = \left\{ \begin{matrix} \alpha \\ \beta\mu \end{matrix} \right\} + S^\alpha_{\beta\mu} + S_{\beta\mu}{}^\alpha - S_\mu{}^\alpha{}_\beta.$$

This means that  $\left\{ \begin{matrix} \alpha \\ \beta\mu \end{matrix} \right\}$  are coefficients of the Levi-Civita connection built of the metric  $g$ , and  $S^\alpha_{\beta\mu} = \Gamma^\alpha_{[\beta\mu]}$  is the torsion tensor of  $\Gamma^\alpha_{\beta\mu}$ .

In a sense this is a gauge theory. There are, however, certain objections against it, although from some point of view the theory works in a satisfactory way as a gauge frame. Let us quote mentioned objections:

(a) The use of tetrads seems to be essential here. This is important. In the standard gauge theories, e.g., in Salam-Weinberg model, or in quantum chromodynamics, the field of frames does not occur explicitly in the formalism. The idea of the local gauge Minkowski translations leads us far beyond the ground of the theory.

(b) Another non-pleasant and strange feature of the theory is the doubtful meaning of the gauge invariance under the group  $SU(2,2)$ , just as its quotient  $SL(2, \mathbb{C})$ . Namely, for the massless particles it is so that field equations are  $SU(2,2)/$ conformally invariant. But Lagrangian is not so, and because of this the Noether theorem is not either. And at the same time the mass form  $G_{\bar{r}s}$  just underlies the inertial properties of those Lagrangians. So, it is natural to expect that it should be essential for the invariance of Lagrangian. Therefore, perhaps Lagrangian should be transformed exactly up to the purely internal rule under  $SU(2,2)$ , just like it is done under the subgroup  $SL(2, \mathbb{C})$  in Einstein-Cartan theory. Namely, that for any  $A \in SU(2,2)$  it should be modified as follows under  $A \in SL(2, \mathbb{C})$  in (1):

$$(A\Psi)^r(x) = A^r_s \Psi^s(x). \quad (13)$$

This is a purely internal rule, like those for  $SL(2, \mathbb{C})$  in (1), without any external correction. In this way, the number of geometric degrees of freedom is increased, but it should be so if  $SU(2,2)$  are to be Hamiltonian symmetries.

(c) Finally, there is a strange feature of Lagrangian (1), namely, the one that it is essentially based on the covariant vector density

$$J^r_{s\mu} := \left( D_\mu \tilde{\Psi}_s \Psi^r - \tilde{\Psi}_s D_\mu \Psi^r \right) \sqrt{|g|}, \quad (14)$$

or, to be more honest, on its contravariant upper-index version

$$J^r_s{}^\mu := g^{\mu\nu} \left( D_\nu \tilde{\Psi}_s \Psi^r - \tilde{\Psi}_s D_\nu \Psi^r \right) \sqrt{|g|}. \quad (15)$$

The idea is that  $J^r_s{}^\mu$  looks as a typical bosonic current. What is the symmetry group responsible for it? The algebraic prescription for  $J^r_s{}^\mu$  does suggest that it is the group  $U(2,2)$

and that the corresponding Lagrangian for  $\Psi$  should be just the complex-four-dimensional Lagrangian for the field  $\Psi$ , this time invariant under the total  $U(2,2)$ , no longer by  $SL(2, \mathbb{C})$ . Nevertheless, some fundamental question remains, namely one concerning the relationship between Klein-Gordon equation of order two and first-order Dirac equation. It is though clear that the structural properties of differential equations are in very malicious way sensitive to the removing highest-order derivative term.

It is interesting to begin the analysis from some rather academic example of the scalar complex field interacting in a minimal way with the gauge field  $e_\mu$ , i.e., dynamically ruled by the unitary group  $U(1)$ . Namely, let us assume the primeval globally invariant by  $U(1)$  Lagrangian for  $\Psi : M \rightarrow \mathbb{C}$ ,

$$L_m = \left( \frac{1}{2} g^{\mu\nu} \overline{\partial_\mu \Psi} \partial_\nu \Psi - \frac{c}{2} \overline{\Psi} \Psi \right) \sqrt{|g|}. \quad (16)$$

Now, as usually we introduce the covector gauge field  $e_\mu$  and the covariant derivative of  $\Psi$ ,

$$D_\mu \Psi := \partial_\mu \Psi - iq e_\mu \Psi. \quad (17)$$

Substituting  $D_\mu$  instead of  $\partial_\mu$  to (16) we obtain as usual the locally-invariant expression

$$L_m = \frac{1}{2} g^{\mu\nu} \overline{D_\mu \Psi} D_\nu \Psi \sqrt{|g|} - \frac{c}{2} \overline{\Psi} \Psi \sqrt{|g|}. \quad (18)$$

This is Lagrangian for  $\Psi$ . The corresponding term for

$$f_{\mu\nu} = \partial_\mu e_\nu - \partial_\nu e_\mu \quad (19)$$

is as usually given by

$$L_g = -\frac{1}{4} g^{\mu\alpha} g^{\nu\lambda} f_{\mu\nu} f_{\alpha\lambda} \sqrt{|g|}, \quad (20)$$

and the total Lagrangian for  $(\Psi, f)$  is given by the sum

$$L = L_m + L_g \quad (21)$$

(the subscripts  $m, g$  refer respectively to the matter and gauge field).

Let us rewrite the matter term in the following form:

$$\begin{aligned} L_m &= q g^{\mu\nu} e_\mu \frac{i}{2} \left( \overline{\Psi} \partial_\nu \Psi - \overline{\partial_\nu \Psi} \Psi \right) \sqrt{|g|} \\ &\quad - \left( \frac{c}{2} - \frac{q^2}{2} g^{\mu\nu} e_\mu e_\nu \right) \overline{\Psi} \Psi \sqrt{|g|} + \frac{1}{2} g^{\mu\nu} \overline{\partial_\mu \Psi} \partial_\nu \Psi \sqrt{|g|}. \end{aligned} \quad (22)$$

The first term, built of the first derivatives and of the algebraic expressions of fields, leads to first-order differential equations with respect to  $\Psi$ . The last, Klein-Gordon term leads through variational principle to the second-order equations in  $\Psi$ . The rigorous field equations have the following form:

$$qie^\mu \partial_\mu \Psi - \left( \frac{c}{2} - \frac{q^2}{2} e^\mu e_\mu - \frac{iq}{2} e^\mu{}_{;\mu} \right) - \frac{1}{2} g^{\mu\nu} \partial_\mu \partial_\nu \Psi = 0, \quad (23)$$

$$\partial_\nu f^{\mu\nu} = \frac{qi}{2} \left( \overline{\Psi} \partial_\mu \Psi - (\partial_\mu \overline{\Psi}) \Psi \right) + q^2 e^\mu \overline{\Psi} \Psi. \quad (24)$$

Obviously, the semicolon symbol in (23) denotes the  $g$ -metric Levi-Civita affine connection, or rather divergence. It is interesting that the vector field  $e^\mu$  plays a role similar to that of gravitational tetrad, in spite of all differences. And in general, the pair  $(\Psi, e^\mu)$  is formally analogous to the pair  $(\Psi^r, e^A{}_\mu)$ , or equivalently  $(\Psi^r, e^A{}_\mu)$ , i.e., bispinor and tetrad/cotetrad. But there is no rigorous Clifford analogy. It is interesting that on the right hand side of (24) there is a combination of two terms: Dirac-like current and Schrödinger current.

It is interesting that the system of equations (23), (24) may be simplified by assuming that the system of first-order derivatives of  $\Psi$  is smaller than the system of quantities built of  $\Psi$  in an algebraic way, and similarly, the system of second derivatives of  $\Psi$  is smaller than the first- and zeroth-order derivatives of  $\Psi$ . But this means that the system (23), (24) may be approximated by the following one:

$$ie^\mu \partial_\mu \Psi - \left( \frac{c}{2q} - \frac{q}{2} e^\mu e_\mu - \frac{i}{2} e^\mu{}_{;\mu} \right) \Psi = 0, \quad (25)$$

$$\partial_\nu f^{\mu\nu} = q^2 e^\mu \bar{\Psi} \Psi. \quad (26)$$

It is interesting that this system, except the Clifford analogy, is structurally similar to the Dirac system of equations. It is difficult to state a priori if the essentially nonlinear system (25), (26) may have anything to do with reality. Nevertheless, the point is that it is both nonlinear, and as a system imposed on the pair  $(\Psi, e_\mu)$  it shows certain similarity to the Dirac-Maxwell system. And the bosonic current  $i(\partial_\mu \bar{\Psi} \Psi - \bar{\Psi} \partial_\mu \Psi) \sqrt{|g|}$  is an obvious counterpart of the  $SU(2,2)$  current given by  $i(D_\mu \tilde{\Psi}_s \Psi^r - \tilde{\Psi}_s D_\mu \Psi^r) \sqrt{|g|}$ . Obviously, the model (25), (26) is a bit non-physical and crazy, especially with its separation of terms. Nevertheless, it seems to follow from it that the above demands and objections concerning the  $U(2,2)$ -invariance and the particular role of the bosonic currents and tetrads may be easily answered on the basis of the spinor counterpart of  $L_m$  (22) and its first-order limit (23), (24).

### 3. Second order Klein-Gordon equation

We need a few things, for instance, affine connection in space-time manifold, spinor connection,  $U(2,2)$ -gauge field, metric tensor, and in certain approaches some field of frames, e.g., generalization of the tetrad field. The space-time manifold  $M$  is assumed structure-less and nothing but the differential-geometric structure is assumed in it. Unlike this, in the target space  $\mathbb{C}^4$ , we assume some internal geometry based on the use of some sesquilinear Hermitian  $G$  form of signature  $(+, +, -, -)$ , as mentioned above. This form does belong to the internal structure of  $\mathbb{C}^4$ , and to be more rigorous, we can assume it to be a complex linear space of dimension four, endowed with the mentioned neutral signature. When this form is fixed, it distinguishes within the complex group  $GL(4, \mathbb{C})$ , the pseudounitary group consisting of transformations preserving  $G$ , so that the following holds:

$$G_{\bar{r}s} = G_{z\bar{t}} \bar{U}^{\bar{z}}{}_{\bar{r}} U^t{}_s. \quad (27)$$

The Lie algebra of this group consists of linear mappings  $u$  which satisfy:

$$G_{\bar{r}z} u^z{}_s + \overline{G_{s\bar{z}} u^z{}_r} = 0. \quad (28)$$

So, roughly speaking,  $U(2,2)'$  consists of matrices (linear mapping of the target space) which are  $G$ -anti-Hermitian. Let us mention that any particular choice of  $\Gamma$  is only a matter of

convenience. It is only its global signature that matters. As mentioned, in the Weyl, Penrose, Finkelstein and Weizsäcker procedure the typical choice is

$$[G_{\bar{r}s}] = \begin{bmatrix} 0 & \mathbf{I}_2 \\ \mathbf{I}_2 & 0 \end{bmatrix}. \quad (29)$$

In the Dirac procedure one prefers the choice:

$$[G_{\bar{r}s}] = \begin{bmatrix} \mathbf{I}_2 & 0 \\ 0 & -\mathbf{I}_2 \end{bmatrix}. \quad (30)$$

Transition between these representations is described by the matrix

$$\frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_2 & \mathbf{I}_2 \\ \mathbf{I}_2 & -\mathbf{I}_2 \end{bmatrix}. \quad (31)$$

Similarly, the Weyl, Penrose, Finkelstein and Weizsäcker procedure leads to the following expressions for the Dirac matrices:

$$\gamma^A = \eta^{AB} \gamma_B = \begin{bmatrix} 0 & \tilde{\sigma}^A \\ \sigma^A & 0 \end{bmatrix} = \begin{bmatrix} 0 & \eta^{AB} \sigma_B \\ \sigma^A & 0 \end{bmatrix}. \quad (32)$$

In the Dirac representation we have that

$$\gamma^0 = \begin{bmatrix} \mathbf{I}_2 & 0 \\ 0 & -\mathbf{I}_2 \end{bmatrix} = \begin{bmatrix} \sigma^0 & 0 \\ 0 & -\sigma^0 \end{bmatrix}, \quad \gamma^R = \begin{bmatrix} 0 & \sigma^R \\ -\sigma^R & 0 \end{bmatrix}, \quad R = 1, 2, 3. \quad (33)$$

Obviously, those are two particular choices, we quote them only as the two most important ones.

The globally  $U(2,2)$ -invariant second-order Klein-Gordon Lagrangian for the  $\mathbb{C}^4$ -valued scalar field on  $M$  is given by

$$L_m(\Psi; g) = \frac{b}{2} g^{\mu\nu} \partial_\mu \bar{\Psi}^{\bar{r}} \partial_\nu \Psi^s G_{\bar{r}s} \sqrt{|g|} - \frac{c}{2} G_{\bar{r}s} \bar{\Psi}^{\bar{r}} \Psi^s \sqrt{|g|}. \quad (34)$$

Making use of the Dirac-conjugate field,

$$\tilde{\Psi}_r := \bar{\Psi}^{\bar{s}} G_{\bar{s}r}, \quad (35)$$

we can rewrite (34) in the following form:

$$L_m(\Psi; g) = \frac{b}{2} g^{\mu\nu} \partial_\mu \tilde{\Psi} \partial_\nu \Psi \sqrt{|g|} - \frac{c}{2} \tilde{\Psi} \Psi \sqrt{|g|}. \quad (36)$$

First let us consider the problem of the local  $U(2,2) \simeq U(4, G)$ -invariance. To do that we must begin with introducing the connection form of the  $U(4, G)$ -connection. This is a  $u(4, G)$ -valued differential form

$$M \ni x \mapsto \theta_x \in L(T_x M, u(4, G)) \quad (37)$$

transforming under the local  $U(4, G)$ -valued local transformations  $U : M \rightarrow U(4, G)$  as follows:

$$(U\theta)_x = U(x)\theta_x U(x)^{-1} - dU_x U(x)^{-1}. \quad (38)$$

This connection form is controlled by the two real parameters corresponding to  $SU(2,2) \simeq SU(4, G)$  and to the one-parameter dilatation group. The corresponding covariant derivative of the four-component Klein-Gordon field has the following form:

$$\nabla_\mu \Psi = \partial_\mu \Psi + g \left( \vartheta_\mu - \frac{1}{4} \text{Tr} \vartheta_\mu \mathbf{I} \right) \Psi + \frac{q}{4} \text{Tr} \vartheta_\mu \Psi = \partial_\mu \Psi + g \vartheta_\mu \Psi + \frac{q-g}{4} \text{Tr} \vartheta_\mu \Psi. \quad (39)$$

Similarly for the Dirac-conjugate field we have the following dual formula:

$$\nabla_\mu \tilde{\Psi} = \partial_\mu \tilde{\Psi} - g \tilde{\Psi} \left( \vartheta_\mu - \frac{1}{4} \text{Tr} \vartheta_\mu \mathbf{I} \right) - \frac{q}{4} \tilde{\Psi} \text{Tr} \vartheta_\mu = \partial_\mu \tilde{\Psi} - g \tilde{\Psi} \vartheta_\mu - \frac{q-g}{4} \tilde{\Psi} \text{Tr} \vartheta_\mu. \quad (40)$$

The curvature form  $\Phi = D\vartheta$  is then expressed as follows:

$$\Phi_{\mu\nu} = d\vartheta_{\mu\nu} + g [\vartheta_\mu, \vartheta_\nu] = \partial_\mu \vartheta_\nu - \partial_\nu \vartheta_\mu + g [\vartheta_\mu, \vartheta_\nu]. \quad (41)$$

Let us observe that the conserved Noether current following from the Noether theorem applied to (34) is given by

$$j^r_{s\mu} = \frac{b}{2} \left( \Psi^r \partial_\mu \tilde{\Psi}_s - \partial_\mu \Psi^r \tilde{\Psi}_s \right) \sqrt{|g|}. \quad (42)$$

One can show that the gauge invariant Lagrangian for the  $\Psi$ -matter has the following form:

$$L_m(\Psi, \vartheta, g) = \frac{b}{2} g^{\mu\nu} \nabla_\mu \tilde{\Psi} \nabla_\nu \Psi \sqrt{|g|} - \frac{c}{2} \tilde{\Psi} \Psi \sqrt{|g|}. \quad (43)$$

The gauge-invariant current

$$J(\Psi, \vartheta, g)^r_{s\mu} = \frac{b}{2} \left( \Psi^r \nabla_\mu \tilde{\Psi}_s - \nabla_\mu \Psi^r \tilde{\Psi}_s \right) \sqrt{|g|} \quad (44)$$

may be obtained from the Lagrangian (43) by performing its differentiation with respect to the connection  $\vartheta$ ,

$$\frac{\partial L_m(\Psi, \vartheta, g)}{\partial \vartheta^r_{s\mu}} = g J^s_{r\mu} + \frac{q-g}{4} J^z_{z\mu} \delta^s_r. \quad (45)$$

This was about the matter Lagrangian. What concerns the gauge Lagrangian, the simplest possibility of the gauge-invariant model is the following one:

$$L_{YM}(\vartheta, g) = \frac{a}{4} \text{Tr} (\Phi_{\mu\nu} \Phi_{z\lambda}) g^{\mu z} g^{\nu\lambda} \sqrt{|g|} + \frac{a'}{4} \text{Tr} \Phi_{\mu\nu} \text{Tr} \Phi_{z\lambda} g^{\mu z} g^{\nu\lambda} \sqrt{|g|}, \quad (46)$$

where  $a, a'$  are constants. The first term, controlled by the parameter  $a$  is the main, Maxwell-like expressions. The second term is additional one, built of the traces of field strengths. It is an auxiliary expression, nevertheless it is geometrically admissible and it may be some reasonable, helpful correction to the first one. In any case, it is a merely supplementary expression, although it may be convenient and physically justified.

Let us observe that in spite of the non-homogeneous transformation rule (38), the curvature two-form (41) transforms according to the tensorial homogeneous rule under (37), (38):

$$(U\Phi)_x = U(x)\Phi(x)U(x)^{-1}. \quad (47)$$

Because of this the Yang-Mills Lagrangian (46) is invariant under the local  $U(2,2) \simeq U(H, G)$  transformations (38). And similarly, the matter Lagrangian (43) is invariant.

This is the main, gauge constituent of the theory. Let us now mention only about the relationship of  $SU(2,2)$ -matrices  $\Phi$  to the representations  $SL(2, \mathbb{C}) \ni A \mapsto U[A] \in U(2,2)$  corresponding to the Weyl-Penrose-Finkelstein-Weizsäcker and to the Dirac representation of  $G_{\overline{78}}$ . In the first group of representation (W-P-F-W) we have the following realizations of  $U[A]$ :

$$U[A] = \begin{bmatrix} A & 0 \\ 0 & A^{-1+} \end{bmatrix}, \quad u[a] = \begin{bmatrix} a & 0 \\ 0 & -a^+ \end{bmatrix} \quad (48)$$

respectively for  $SL(2, \mathbb{C})$  and its Lie algebra. In Dirac representation

$$U[A] = \frac{1}{2} \begin{bmatrix} A + A^{-1+} & A - A^{-1+} \\ A - A^{-1+} & A + A^{-1+} \end{bmatrix}, \quad u[a] = \frac{1}{2} \begin{bmatrix} a - a^+ & a + a^+ \\ a + a^+ & a - a^+ \end{bmatrix} \quad (49)$$

for the group and algebra. The right “plus” superscript denotes obviously the Hermitian matrix conjugate. Obviously, quite independently of the choice of any representation the following holds:

$$U[A]\gamma_K U[A]^{-1} = \gamma_L P[A]^L{}_K, \quad (50)$$

where  $P : SL(2, \mathbb{C}) \rightarrow SO(1,3)^\uparrow$  is the covering projection. The mappings  $U : SL(2, \mathbb{C}) \rightarrow U(2,2)$  and  $P : SL(2, \mathbb{C}) \rightarrow SO(1,3)^\uparrow$  generate the corresponding homomorphisms of Lie algebras,  $u : SL(2, \mathbb{C})' \rightarrow U(2,2)'$  and  $p : SL(2, \mathbb{C})' \rightarrow SO(1,3)'$ . They are synchronized by

$$[u[a], (\gamma)_K] = \gamma_L p[a]^L{}_K. \quad (51)$$

It is also worth to note the following expressions:

$$P[A]^L{}_K = \frac{1}{4} \text{Tr} \left( \gamma^L U[A] \gamma_K U[A]^{-1} \right), \quad (52)$$

$$p[a]^L{}_K = \frac{1}{2} \text{Tr} \left( \gamma^L u[a] \gamma_K \right), \quad (53)$$

and

$$u[a] = \frac{1}{2} p[a]^L{}_K \Sigma^L{}_K, \quad (54)$$

where after the shift of indices we have that

$$\Sigma^{LK} = \frac{1}{4} \left( \gamma^L \gamma^K - \gamma^K \gamma^L \right) = \frac{1}{4} \left[ \gamma^L, \gamma^K \right]. \quad (55)$$

#### 4. What about the metric tensor?

In the gauge Lagrangians above, the metric tensor in a sense played the parameter role. Our idea was to construct the  $U(2,2) \simeq U(H, G)$ -invariant theory of gravitation. The main constituents of the theory were the four-component complex Klein-Gordon field and the corresponding Maxwell-like gauge field. We will show that there are interesting and very important points for which this is important, perhaps even just exciting. But there is some weak point which was not yet completely explained. It is just the role and physical status of the metric tensor, which is present in the Klein-Gordon and gauge Lagrangian, however its geometric and physical sense is not yet full understood. It is clear that it must occur there if we are to be able to construct Lagrangians. But what is its meaning and how to identify properly

its physical role? As one of potentials of gravitation, or as some secondary variable? And if the second possibility is to be chosen, what are the primary variables the byproduct of which is the metric tensor? Situation in this respect was clear only in the standard Einstein General Relativity. There it was just the only gravitational potential (or perhaps a superpotential if the connection coefficients were interpreted as proper potentials). But within any gauge framework the metric tensor is a merely one of a few potentials. In this paper we concentrate on the theory aspects not very sensitive to this problem. Instead, we shall present a few possibilities.

First of all, let us notice that quite naively, one can assume the Hilbert-Einstein term for the metric tensor  $g$ ,

$$L_{HE}(g) = -dR(g)\sqrt{|g|} + l\sqrt{|g|}. \quad (56)$$

where  $d, l$  are real constants. The special case  $d = 0$  is not to be a priori rejected. Namely, if  $d = 0$  and perhaps  $l = 0$ , then variation of the action functional with respect to  $g_{\mu\nu}$  enables one to express  $g_{\mu\nu}$  through the other variables. But of course, the choice (56) looks rather naive. In any case, the total Lagrangian of the form

$$L(\Psi, \vartheta, g) := L_m(\Psi, \vartheta, g) + L_{YM}(\vartheta, g) + L_{HE}(g) \quad (57)$$

leads, after the variational procedure for the action, to the following system of equations:

$$g^{\mu\nu} \overset{g}{\nabla}_\mu \overset{g}{\nabla}_\nu \Psi + \frac{c}{b} \Psi = 0, \quad (58)$$

$$\chi^{\mu\nu}{}_{;\nu} + g[\vartheta_\nu, \chi^{\mu\nu}] = gJ^\mu + \frac{q-g}{4} \text{Tr} J_\mu \mathbf{I}, \quad (59)$$

$$d \left( R(g)^{\mu\nu} - \frac{1}{2} R(g) g^{\mu\nu} \right) = \frac{l}{2} g^{\mu\nu} + \frac{1}{2} T^{\mu\nu}, \quad (60)$$

with the meaning of symbols as below:

1. The semicolon “;” is the  $g$ -Levi-Civita covariant differentiation.
2. The symbol  $\overset{g}{\nabla}_\mu$  denotes the complete covariant differentiation. The Levi-Civita covariant differentiation of the space-time indices is joined there with the  $U(2, 2) \simeq U(H, G)$  covariant differentials of internal indices. Let us quote a typical example:

$$\overset{g}{\nabla}_\mu Y^r{}_\nu + g\vartheta^r{}_{s\mu} Y^s{}_\nu + \frac{q-g}{4} \vartheta^z{}_{z\mu} Y^r{}_\nu - \overset{g}{\left\{ \begin{matrix} \lambda \\ \mu\nu \end{matrix} \right\}} Y^r{}_\lambda, \quad (61)$$

and similarly, i.e., dually, or in the Leibniz-multiplication sense, for other quantities.

3.  $\chi$  is the field momentum conjugate to  $\vartheta$ , so

$$\begin{aligned} \chi^r{}_{s\mu\nu} &= \frac{\partial L_{YM}}{\partial \vartheta^s{}_{r\mu,\nu}} = -a\Phi^r{}_{s\alpha\beta} g^{\alpha\mu} g^{\beta\nu} \sqrt{|g|} \\ &\quad - a' \delta^r{}_s \Phi^z{}_{z\alpha\beta} g^{\alpha\mu} g^{\beta\nu} \sqrt{|g|}, \end{aligned} \quad (62)$$

or in the shortened form with the  $g$ -shifting of indices,

$$\chi^{\mu\nu} = -a\Phi^{\mu\nu} \sqrt{|g|} - a' \text{I} \text{Tr} \Phi^{\mu\nu} \sqrt{|g|}. \quad (63)$$

4.  $T^{\mu\nu}$  denotes the symmetric energy-momentum tensor of the fields  $\vartheta, \Psi$ , so we have that

$$T^{\mu\nu} = T_m^{\mu\nu} + T_{YM}^{\mu\nu}, \quad (64)$$

where, obviously,

$$T_m^{\mu\nu} = -\frac{2}{\sqrt{|g|}} \left( \frac{\partial L_m}{\partial g^{\mu\nu}} - \left( \frac{\partial L_m}{\partial g^{\mu\nu,\alpha}} \right)_{,\alpha} \right), \quad (65)$$

$$T_{YM}^{\mu\nu} = -\frac{2}{\sqrt{|g|}} \left( \frac{\partial L_{YM}}{\partial g^{\mu\nu}} - \left( \frac{\partial L_{YM}}{\partial g^{\mu\nu,\alpha}} \right)_{,\alpha} \right). \quad (66)$$

We do not quote the explicit formulae.

As mentioned, the Hilbert-Einstein term of Lagrangian in (57) looks rather naive, although perhaps it may be reasonable. Equations resulting from the version with vanishing coefficients (or vanishing "d" at least) also seem to be not bad, and in any case not to be a priori rejected. And, as mentioned, the field equations following from the first two terms of (57) seem promisable. But, as said above, the Hilbert-Einstein term seems to spoil the whole taste of the gauge approach. In Einstein-Cartan theory it was the tetrad field who saved the situation, nevertheless, also for some price (as mentioned, in no other gauge theory one explicitly uses the field of frames as a dynamical variable). What may be done in our formalism to replace in a reasonable way the role of tetrad? We would like to answer this question before the further development of our theory. There are a few, at least three natural ways. Certainly there is no possibility to build the metric tensor from the gauge field, in the sense:

$$g_{\mu\nu} := p\vartheta^r{}_{s\mu}\vartheta^s{}_{r\nu} + q\vartheta^r{}_{r\mu}\vartheta^s{}_{s\nu}, \quad (67)$$

what apparently might seem natural. The point is, however that (67) is only globally, but not locally  $U(2,2)$ -invariant. But one can do it in a local way, by introducing some fields more elementary than the metric itself.

1. We may assume that besides the connection form  $\vartheta$ , the geometrodynamical sector involves some additional  $\mathbb{C}^4$ -valued ( $H$ -valued, let us say) differential one-form  $W$ :

$$M \ni x \mapsto W_x \in L(T_x M, \mathbb{C}^4).$$

Analytically we describe this object as  $W^r{}_\mu$ . And we assume that it is homogeneously transformable under the locally acting  $U(2,2)$ ,

$$W_x \mapsto U(x)W_x, \quad \text{i.e.,} \quad {}^l W^r{}_\mu = U^r{}_s(x)W^s{}_\mu. \quad (68)$$

This form gives rise to the metric tensor field on  $M$  as follows:

$$g(W)_{\mu\nu} := \text{Re} \left( \tilde{W}^r{}_\mu W^r{}_\nu \right) = \text{Re} \left( \tilde{W}_\mu W_\nu \right). \quad (69)$$

Therefore, this expression is the symmetric, thus real part of the Hermitian tensor  $W_x^* G$ . The quantity is locally  $U(2,2)$ -invariant. The simplest gauge-invariant Lagrangian is given by

$$L(W, \vartheta) = a \nabla \tilde{W}_{\mu\nu} \nabla W_{\alpha\lambda} g^{\mu\alpha} g^{\nu\lambda} \sqrt{|g|} + b \sqrt{|g|}. \quad (70)$$

In this expression  $a, b$  are some real constants,  $g^{\mu\alpha}g_{\alpha\nu} = \delta^{\mu}_{\nu}$ , and  $\nabla W$  denotes the exterior covariant differential of  $W$ , so that

$$\nabla W_{\mu\nu} = dW_{\mu\nu} + g(\vartheta_{\mu}W_{\nu} - \vartheta_{\nu}W_{\mu}) + \frac{{}'q - g}{4}(\text{Tr } \vartheta_{\mu}W_{\nu} - \text{Tr } \vartheta_{\nu}W_{\mu}). \quad (71)$$

Here  $'q$ , the coupling constant, is the kind of electric charge of  $W$ . Let us stress, the Lagrangian (70) is locally invariant under  $U(2,2)$ . It is interesting that after the  $SL(2, \mathbb{C})$ -reduction  $W$  is a 3/2-spin particle. This resembles the super-symmetric idea of gravitino.

- Let us suppose that besides of  $\vartheta$ , the geometric sector contains also another  $U(2,2)'$ -valued differential form  $W$ ,  $M \ni x \mapsto W_x \in L(T_x M, U(2,2)')$ . Analytically it is represented by the system of quantities  $W^r{}_{s\mu}$ . But unlike the connection form  $\vartheta$ , just like in the previous idea, it suffers a homogeneous transformation rule under  $U(2,2)$ ,

$$W_x \mapsto U(x)W_x U(x)^{-1}, \quad {}'W^r{}_{s\mu} = U^r{}_z W^z{}_{t\mu} U^{-1t}{}_s. \quad (72)$$

The corresponding metric field  $g(W)$  on  $M$  is given by

$$g(W)_{\mu\nu} = a\text{Tr}(W_{\mu}W_{\nu}) + b\text{Tr } W_{\mu}\text{Tr } W_{\nu}, \quad (73)$$

where  $a, b$  are constants and obviously  $a \neq 0$ ; the  $a$ -term is dominant, whereas the  $b$ -term is a merely correction.

The exterior covariant differential of  $W$  is given by

$$\nabla W_{\mu\nu} = dW_{\mu\nu} + g[\vartheta_{\mu}, W_{\nu}] - g[W_{\nu}, \vartheta_{\mu}]. \quad (74)$$

The corresponding Maxwell Lagrangian for the form  $W$  is given by

$$\begin{aligned} L(W, \vartheta) &= a\text{Tr}(\nabla W_{\mu\nu}\nabla W_{\kappa\lambda})g^{\mu\kappa}g^{\nu\lambda}\sqrt{|g|} \\ &+ b\text{Tr}(\nabla W_{\mu\nu})\text{Tr}(\nabla W_{\kappa\lambda})g^{\mu\kappa}g^{\nu\lambda}\sqrt{|g|} + c\sqrt{|g|} \end{aligned} \quad (75)$$

with constant coefficients  $a, b, c$ .

- There is also a different model, maximally economic in the sense that its only dynamical variables are  $\vartheta$  and  $\Psi$ . There is neither  $g$  nor any other geometric quantity used as Lagrangian argument. Instead, we use the metric-like tensor built in a locally-invariant gauge way from the basic field quantities:

$$g(\Psi, \vartheta)_{\mu\nu} = a\text{Re}(\nabla_{\mu}\tilde{\Psi}\nabla_{\nu}\Psi) = a\text{Re}(G_{\tilde{r}s}\nabla_{\mu}\tilde{\Psi}^{\tilde{r}}\nabla_{\nu}\Psi^s). \quad (76)$$

Obviously, it would be meaningless to substitute this metric to the usual Klein-Gordon Lagrangian, because the result would be trivial. However, there are modified Born-Infeld type schemes, in a sense very interesting ones, as usual Born-Infeld schemes are. The typical Born-Infeld scheme for  $(\Psi, \vartheta)$  is the following one:

$$L(\Psi, \vartheta) = \sqrt{\left| \det \left[ \frac{b}{2}g_{\mu\nu} + \frac{a}{4}\text{Tr}(F_{\mu\kappa}F_{\nu\lambda})g^{\kappa\lambda} + \frac{a'}{4}\text{Tr } F_{\mu\kappa}\text{Tr } F_{\nu\lambda}g^{\kappa\lambda} \right] \right|}. \quad (77)$$

This is the most natural Born-Infeld scheme.

Let us mention, there are also similar, but modified expressions with some “potential” terms, e.g.,

$$L(\Psi, \vartheta) = \frac{a}{4} \text{Tr} (F_{\mu\nu} F_{\kappa\lambda}) g^{\mu\kappa} g^{\nu\lambda} \sqrt{|g|} + \frac{a'}{4} \text{Tr} F_{\mu\nu} \text{Tr} F_{\kappa\lambda} g^{\mu\kappa} g^{\nu\lambda} \sqrt{|g|} + b \sqrt{|g|}. \quad (78)$$

In both expressions (77), (78) the space-time metric is given by (76), and  $a, a', b$  are some real constants.

## 5. The main ideas of the Klein-Gordon $U(2,2)$ -ruled theory

Without judging the three presented models of the metric field, or rather the four of them if the Hilbert-Einstein possibility is admitted, we have nevertheless formulated them. And from the point of view of aesthetic criteria, they look quite reasonable. But now, let us discuss the main results of the very  $U(2,2)$ -invariant Klein-Gordon gauge model of gravity as ruled by the first two terms of (57), or even by the total (57).

The field equations (58), (59), (60) become qualitatively readable when some special basis is chosen in the Lie algebra of  $U(2,2)$ . This basis is somehow related to the twistor geometry, although literally it is something else than the conformal geometry in Minkowskian space. The basis elements are built algebraically of  $\gamma^A$ -matrices.

Let us introduce the following matrices built algebraically of  $\gamma^A$ -s:

$$\gamma^5 = -\gamma_5 = -\gamma^0 \gamma^1 \gamma^2 \gamma^3, \quad (79)$$

$${}^A\gamma = i\gamma^A \gamma^5 = -i\gamma^5 \gamma^A, \quad (80)$$

$$\Sigma^{AB} = \frac{1}{4} (\gamma^A \gamma^B - \gamma^B \gamma^A). \quad (81)$$

It is clear that  ${}^A\gamma$ -s satisfy the opposite-sign anticommutation rules

$$\{ {}^A\gamma, {}^B\gamma \} = -2\eta^{AB} I. \quad (82)$$

One can show that

$${}^A\gamma = {}^B\gamma \eta_{BA} = -\frac{i}{6} \varepsilon_{ABCD} \gamma^B \gamma^C \gamma^D, \quad (83)$$

where the convention  $\varepsilon_{0123} = 1$  is used.

The Lie algebra  $U(2,2)' = U(H, G)'$  may be spanned in the  $\mathbb{R}$ -sense on the matrices

$$i\gamma^A, \quad i{}^A\gamma, \quad \Sigma^{AB}, \quad i\gamma^5, \quad iI_4. \quad (84)$$

Removing from this system the imaginary matrix  $iI_4$ , we obtain the basis of  $SU(2,2)' = SU(H, G)'$ . The matrices  $i\gamma^A, i{}^A\gamma$  do not  $\mathbb{R}$ -span Lie algebras. But it is clear that their sum and difference are bases of Abelian Lie subalgebras:

$$\tau_A := \frac{1}{2} (\gamma_A + {}^A\gamma), \quad \chi^A := \frac{1}{2} (\gamma^A - {}^A\gamma), \quad (85)$$

$$[\tau_A, \tau_B] = 0, \quad [\chi^A, \chi^B] = 0. \quad (86)$$

In the twistor language the quantities  $\tau_A$  generate Minkowskian translations, while  $\chi^A$  are generators of the group of proper conformal mappings. Obviously, this interpretation is true only within the framework of Minkowskian-conformal geometry. The literal meaning of this interpretation is lost within the internal interpretation of those mappings. Nevertheless, the commutation rules of the conformal group are still valid. It becomes internal group just like the Lorentz group in Einstein-Cartan theory.

The connection from  $\vartheta$  may be expanded as follows:

$$\vartheta_\mu = \frac{1}{2}\check{\Omega}^{AB}{}_\mu \Sigma_{AB} + B_\mu \frac{1}{i}\gamma^5 + A_\mu iI + e^A{}_\mu i\tau_A + f_{A\mu} i\chi^A, \quad (87)$$

where, obviously,  $\check{\Omega}^{AB}{}_\mu = -\check{\Omega}^{BA}{}_\mu$ .

It may be convenient to introduce the object

$$\Omega^A{}_{B\mu} := \check{\Omega}^A{}_{B\mu} + 2B_\mu \delta^A{}_B, \quad (88)$$

where the following holds:

$$B_\mu = \frac{1}{8}\Omega^A{}_{A\mu}, \quad \check{\Omega}^A{}_{B\mu} = \Omega^A{}_{B\mu} - \frac{1}{4}\Omega^C{}_{C\mu} \delta^A{}_B. \quad (89)$$

Therefore,  $8B_\mu$  may be identified with the trace, and  $\check{\Omega}^A{}_{B\mu}$  — with the trace-less part of the object  $\Omega^A{}_{B\mu}$ . So, it may be natural to write  $\vartheta_\mu$  as follows:

$$\vartheta_\mu = \frac{1}{2}\Omega^{AB}{}_\mu \left( \Sigma_{AB} + \frac{1}{4}n_{AB} \frac{1}{i}\gamma^5 \right) + e^A{}_\mu i\tau_A + f_{A\mu} i\chi^A + A_\mu iI, \quad (90)$$

or alternatively

$$\vartheta_\mu = \frac{1}{2g}\check{\Gamma}^{AB}{}_\mu \Sigma_{AB} + \frac{1}{4g}Q_\mu \frac{1}{i}\gamma^5 + \frac{1}{g}\varepsilon^A{}_\mu i\tau_A + \frac{1}{g}\varphi_{A\mu} i\chi^A + A_\mu iI, \quad (91)$$

or just as

$$\vartheta_\mu = \frac{1}{2g}\Gamma^{AB}{}_\mu \left( \Sigma_{AB} + \frac{1}{4}n_{AB} \frac{1}{i}\gamma^5 \right) + \frac{1}{g}\varepsilon^A{}_\mu i\tau_A + \frac{1}{g}\varphi_{A\mu} i\chi^A + A_\mu iI, \quad (92)$$

where the following auxiliary gauge symbols are used:

$$\Gamma^A{}_{B\mu} = g\Omega^A{}_{B\mu}, \quad \check{\Gamma}^A{}_{B\mu} = \Gamma^A{}_{B\mu} - \frac{1}{4}\Gamma^C{}_{C\mu} \delta^A{}_B, \quad (93)$$

$$Q_\mu = 4gB_\mu = \frac{g}{2}\Omega^A{}_{A\mu} = \frac{1}{2}\Gamma^A{}_{A\mu}, \quad (94)$$

$$\varepsilon^A{}_\mu = g e^A{}_\mu, \quad \varphi_{A\mu} = g f_{A\mu}. \quad (95)$$

The systems of differential forms  $[\Omega^A{}_B = (1/g)\Gamma^A{}_B]$ ,  $[e^A]$ ,  $[f_A]$  are parts of the connection form  $\vartheta$ , and because of this, the action of  $x$ -dependent matrices  $U$  on them is inhomogeneous. But when we restrict ourselves to the  $U$ -injected group  $SL(2, \mathbb{C})$ , the transformation rule for  $[e^A]$ ,  $[f_A]$  becomes homogeneous. It is just the correspondence rule with the situation of Einstein-Cartan theory where  $[e^A]$  was a gravitational cotetrad. Let us remind that in  $GL(2, \mathbb{C})$ -invariant spinor theory,  $Q_\mu = (1/2)\Gamma^A{}_{A\mu}$  was the Weyl covector, and there was

necessity to use an additional version of the cotetrad, transforming under dilatations in the inverse way. If we used the world metric  $g_{\mu\nu}$ , then  $Q_\mu$  was the Killing covector, i.e.,

$$\nabla_\lambda g_{\mu\nu} = -Q_\lambda g_{\mu\nu}, \quad (96)$$

in the sense of the  $GL(2, \mathbb{C})$ -part of the  $U(2, 2)$ -connection.

Let us stress that in certain formulae it is still more convenient to use the  $\gamma^A$ ,  ${}^A\gamma$ -expansion than those based on  $\tau_A, \chi^A$ , namely,

$$\vartheta_\mu = \frac{1}{2}\Omega^{AB}{}_\mu\Sigma_{AB} + B_\mu \frac{1}{i}\gamma^5 + A_\mu iI + E_{A\mu} i\gamma^A + F_{A\mu} i^A\gamma, \quad (97)$$

where the following notation is used:

$$E^A = \frac{1}{2}(e^A + \eta^{AB}f_B) = \eta^{AB}E_B, \quad F^A = \frac{1}{2}(e^A - \eta^{AB}f_B) = \eta^{AB}F_B. \quad (98)$$

Let us now take the following expansion of the curvature vector-valued two-form:

$$\Phi = T(e)^A i\tau_A + T(f)^A i\chi_A + \frac{1}{2}\tilde{R}^{AB}\Sigma_{AB} + G \frac{1}{i}\gamma^5 + FiI, \quad (99)$$

where the following partially clear symbols are used:

$$T(e)^A = de^A + g\Omega^A{}_B \wedge e^B = de^A + \Gamma^A{}_B \wedge e^B, \quad (100)$$

$$T(f)_A = df_A + gf_B \wedge \Omega^B{}_A = df_A + f_B \wedge \Gamma^B{}_A, \quad (101)$$

$$\begin{aligned} \tilde{R}^A{}_B &= R(\Omega)^A{}_B - \frac{1}{4}R(\Omega)^C{}_C \delta^A{}_B - 2ge^A \wedge f_B \\ &\quad + 2g\eta^{AC}\eta_{BD}e^D \wedge f_C = \frac{1}{g} \left( R(\Gamma)^A{}_B - \frac{1}{4}R(\Gamma)^C{}_C \delta^A{}_B \right. \\ &\quad \left. - 2g^2e^A \wedge f_B + 2g^2\eta^{AC}\eta_{BD}e^D \wedge f_C \right), \end{aligned} \quad (102)$$

$$G = \frac{1}{4g}dQ - ge^A \wedge f_A = \frac{1}{g} \left( \frac{1}{8}R(\Gamma)^A{}_A - g^2e^A \wedge f_A \right), \quad (103)$$

$$F = dA, \quad (104)$$

where  $R(\Gamma), R(\Omega)$  denote the curvature two-form:

$$R(\Gamma)^A{}_B = d\Gamma^A{}_B + \Gamma^A{}_C \wedge \Gamma^C{}_B, \quad R(\Omega)^A{}_B = d\Omega^A{}_B + g\Omega^A{}_C \wedge \Omega^C{}_B. \quad (105)$$

Let us remember that the torsion of a linear connection may be interpreted as a contribution to affine connection. The corresponding space-time objects, i.e., connections, torsions and curvatures are given by

$$\Gamma(e)^k{}_{ij} = e^k{}_A \Gamma^A{}_{Bj} e^B{}_i + e^k{}_A e^A{}_{ij}, \quad (106)$$

$$\Gamma(f)^k{}_{ij} = -f_{Ai} \Gamma^A{}_{Bj} f^{kB} + f^{kA} f_{Ai,j}, \quad (107)$$

$$S(e)^k{}_{ij} = \Gamma(e)^k{}_{[ij]} = -\frac{1}{2}e^k{}_A T(e)^A{}_{ij}, \quad (108)$$

$$S(f)^k{}_{ij} = \Gamma(f)^k{}_{[ij]} = -\frac{1}{2}f^{kA} T(f)_{Aij}, \quad (109)$$

$$R(e)^m{}_{kij} = e^m{}_A e^B{}_k R^A{}_{Bij}, \quad (110)$$

$$R(f)^m{}_{kij} = -f_{Ak} f^{mB} R^A{}_{Bij}. \quad (111)$$

We do not quote the total explicit form of field equations. Being generally-covariant, they are over-determined, therefore, as usual suspected to be inconsistent. Nevertheless, substituting to the field equations the non-excited matter  $\Psi = 0$ , and the above equations (106)–(111), one can show that there are certain non-trivial solutions. Namely, let us take the following Einstein-Dirac metrics:

$$\begin{aligned} h(e, \eta)_{\mu\nu} &= \eta_{AB} e^A_{\mu} e^B_{\nu}, \\ h(f, \eta)_{\mu\nu} &= \eta^{AB} f_{A\mu} f_{B\nu}. \end{aligned} \quad (112)$$

And now, let us substitute the following conditions to the field equations (58), (59), (60):

$$\Psi = 0, \quad f_{A\mu} = k\eta_{AB} e^B_{\mu}, \quad g_{\mu\nu} = ph(e, \eta)_{\mu\nu}, \quad (113)$$

$$Q_{\mu} = 0, \quad A_{\mu} = 0, \quad S(e)^{\lambda}_{\mu\nu} = S(f)^{\lambda}_{\mu\nu} = 0. \quad (114)$$

It is simply marvellous that the very complicated system of equations following from (58), (59), (60) after substituting (106)–(114) is solvable, moreover, it is reducible to something very simple. Namely, the very complicated system of equations for the geometric fields reduces step by step to

$$R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu} = -12\frac{g^2k}{p}g^{\mu\nu}, \quad (115)$$

where  $R^{\mu\nu}$  denotes the twice contravariant Ricci tensor built of  $g_{\alpha\beta}$ , and  $R$  is the curvature scalar. Substituting there  $k = 1$ ,  $p = 1$ , one obtains simply the following equation

$$R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu} = -12g^2g^{\mu\nu}. \quad (116)$$

In any case one deals here with the Einstein equation with the cosmological constant. It is remarkable that this coupling constant is proportional to the gauge coupling constant. Such a coupling between microphysical model and macroscopic or even just cosmic scale physics is marvellous and philosophically fascinating. When the Einstein-Hilbert dynamics of  $g_{\mu\nu}$  is used, we obtain also the condition

$$T^{\mu\nu} = 0. \quad (117)$$

But there is no contradiction between (117) and (115)/(116). There exist some common solutions, namely, ones corresponding to the constant curvature spaces:

$$R_{\alpha\beta\mu\nu} = \frac{4g^2k}{p} (g_{\alpha\mu}g_{\beta\nu} - g_{\alpha\nu}g_{\beta\mu}). \quad (118)$$

It is again the fascinating idea that the conformal flatness of space-time, expressed by (118) is so nicely compatible with the assumption that the theory is invariant under the covering group of the conformal group.

The field sector of the model corresponds smoothly to the gauge Poincare gravitation. When the field Lagrangian is expressed in terms of the above quantities  $e$ ,  $f$ ,  $S(e)$ ,  $S(f)$ ,  $R(e)$ ,  $R(f)$ , one obtains expression quadratic in field variables just like in Poincare gauge models. However, the use of the semisimple  $SU(2, 2)$  results in well-defined, rigorous ratio of constant coefficients, not accidental one like in Poincare model.

What concerns material sector, the use of the second-order Klein-Gordon equation and the presence of second-order derivatives of the matter field, is a drastic difference in comparison

with the first-order Dirac equation. However, the situation is not very bad, on the contrary, it may seem promising and desirable. If we substitute to the matter equation the Dirac-Einstein assumption, e.g., with  $p = 1$ ,  $k = 1$ , then we obtain the following Dirac-Klein-Gordon equation:

$$e^\mu{}_A i\gamma^A (\nabla_\mu + S^{\nu\mu} I_4) \Psi - \frac{4bg^2 - c}{2bg} \Psi + \frac{1}{2g} \overset{g}{\nabla}_\mu \overset{g}{\nabla}_\nu \Psi = 0, \quad (119)$$

where  $\nabla_\mu$  is the  $SL(2, \mathbb{C})$ -part of the  $U(2, 2)$ -covariant derivative and  $\overset{g}{\nabla}_\mu$  joints that differentiation with  $\Gamma^A{}_B$ -differentiation of objects with the capital Lorentz indices and with the  $g$ -Levi-Civita differentiation. It is interesting that the first two terms of (119) correspond exactly with the Dirac theory in Einstein-Cartan space-time. However, there is quite a natural question if the third, second-order d'Alembert term does not destroy completely this similarity.

The simplest way to answer this question is to consider the specially-relativistic situation, when  $g_{\mu\nu} = \eta_{\mu\nu}$ ,  $\Gamma^A{}_{B\mu} = 0$ ,  $e^A{}_\mu = \delta^A{}_\mu$ . It is clear that under this substitution one obtains the Dirac-Klein-Gordon differential equation

$$i\gamma^\mu \partial_\mu \Psi - \frac{4bg^2 - c}{2bg} \Psi + \frac{1}{2g} \eta^{\mu\nu} \partial_\mu \partial_\nu \Psi = 0. \quad (120)$$

It is obvious that the general solution of this equation is a combination of two Dirac waves with two possible masses, namely,  $m_\pm$ , where

$$m_\pm^2 = \frac{c}{b} - 2g^2 \left( 1 \pm \sqrt{\frac{c}{bg^2} - 3} \right). \quad (121)$$

This result is obtained when the amplitudes  $U(p) \exp(ip_\mu x^\mu)$  are substituted to (120). This means that there is a range of "Dirac" behaviour,  $c/b > 3g^2$ . The primeval mass parameter must be sufficiently large for that. Then the solution of (120) will be a superposition of two Dirac fields. If  $c/b = 3g^2$ , then one obtains the exactly Dirac behaviour. This means that there is no splitting of mass and that  $m = |g|$ . Below this threshold we are dealing with tachyonic or decay phenomena. It is also very interesting that if  $c/b = 4g^2$ , then one of the partner states is massless, namely,  $m_- = 0$ ,  $m_+ = 2|g|$ . Obviously, the doubling of mass within the Dirac behaviour needs some explanation in terms of experimental data. Let us quote three possibilities (Sławianowski & Kovalchuk, 2008):

- If the energy gap  $m_+ - m_-$  is very small (i.e.,  $|g|$  is small enough), then perhaps it is below the present accuracy of our experimental abilities.
- If the energy gap  $m_+ - m_-$  is so large that perhaps it is too difficult to create/excite the state of higher mass.
- And finally, the most important and promising explanation. Perhaps the mass-state-doubling does exist and is just observed. This would be just the explanation of the mysterious relationship between fundamental quarks and fermions in the standard model of weak interactions. We mean their occurrence in pairs  $(u, d)$ ,  $(c, s)$ ,  $(t, b)$  (quarks) and  $(\nu_e, e)$ ,  $(\nu_\mu, \mu)$ ,  $(\nu_\tau, \tau)$  (leptons) (Veltman, 2003). For example, the situation  $c/b = 4g^2$  might be a naive explanation of the pairing between heavy leptons and their neutrinos.

## 6. Some additional interpretation problems

In one of our earlier papers we have discussed the idea of higher-order derivatives just from the point of view of the order of time derivatives (Sławianowski & Kovalchuk, 2002; 2010; Sławianowski et al., 2004; 2005), because the other continuous variables were absent. The Schrödinger equation was then interpreted as a Hamiltonian system of mathematical physics. And it was just then where the second-order time derivatives seemed not only admissible, but just necessary. In the field problems, this concerns, of course, the occurrence of all second-order space-time derivatives of the field quantities/wave functions. And, as shown above, those second-order space-time derivatives just seem desirable, not only admissible. The next important question is: why just  $U(2,2) \subset GL(4, \mathbb{C})$ , not the total  $GL(4, \mathbb{C})$ ? But it is seen that it is only signature  $(+, +, -, -)$  of Hermitian  $G$ -forms, not the group  $U(2,2)$  itself that matters. Namely, the mass form  $G_{\bar{r}s}$  must be present in Lagrangian, however not as a fixed constant Hermitian form but as a dynamical,  $x$ -dependent form of signature  $(+, +, -, -)$ . The corresponding Lagrangian term would be proportional to

$$G^{s\bar{z}} G^{t\bar{r}} \frac{\partial G_{\bar{r}s}}{\partial x^\mu} \frac{\partial G_{\bar{z}t}}{\partial x^\nu} g^{\mu\nu}, \quad (122)$$

or perhaps to the Born-Infeld term

$$\sqrt{\det [G^{s\bar{z}} G^{t\bar{r}} G_{\bar{r}s, \mu} G_{\bar{z}t, \nu}]} \quad (123)$$

independent on the mentioned choice of the space-time metric  $g_{\mu\nu}$ . This term is evidently  $GL(4, \mathbb{C})$ -invariant.

## 7. References

- Sławianowski, J.J. & Kovalchuk, V. (2002). Klein-Gordon-Dirac equation: physical justification and quantization attempts. *Reports on Mathematical Physics*, Vol. 49, No. 2/3, 249–257.
- Sławianowski, J.J. & Kovalchuk, V. (2008). Search for the geometrodynamical gauge group. Hypotheses and some results, In: *Geometry, Integrability and Quantization IX*, Ivailo M. Mladenov, (Ed.), 66–132, SOFTEX, Sofia.
- Sławianowski, J.J. & Kovalchuk, V. (2010). Schrödinger and related equations as Hamiltonian systems, manifolds of second-order tensors and new ideas of nonlinearity in quantum mechanics. *Reports on Mathematical Physics*, Vol. 65, No. 1, 29–76.
- Sławianowski, J.J.; Kovalchuk, V.; Sławianowska A.; Gołubowska B.; Martens A.; Rożko E.E. & Zawistowski Z.J. (2004). Affine symmetry in mechanics of collective and internal modes. Part I. Classical models. *Reports on Mathematical Physics*, Vol. 54, No. 3, 373–427.
- Sławianowski, J.J.; Kovalchuk, V.; Sławianowska A.; Gołubowska B.; Martens A.; Rożko E.E. & Zawistowski Z.J. (2005). Affine symmetry in mechanics of collective and internal modes. Part II. Quantum models. *Reports on Mathematical Physics*, Vol. 55, No. 1, 1–45.
- Veltman M. (2003). *Facts and mysteries in elementary particle physics*, World Scientific, New Jersey.

# Theory of “Weak Value” and Quantum Mechanical Measurements

Yutaka Shikano

*Department of Physics, Tokyo Institute of Technology, Tokyo  
Japan*

## 1. Introduction

Quantum mechanics provides us many perspectives and insights on Nature and our daily life. However, its mathematical axiom initiated by von Neumann (121) is not satisfied to describe nature phenomena. For example, it is impossible not to explain a non self-adjoint operator, i.e., the momentum operator on a half line (See, e.g., Ref. (154).), as the physical observable. On considering foundations of quantum mechanics, the simple and specific expression is needed. One of the candidates is the *weak value* initiated by Yakir Aharonov and his colleagues (4). It is remarked that the idea of their seminal work is written in ref. (3). Furthermore, this quantity has a potentiality to explain the counter-factual phenomena, in which there is the contradiction under the classical logic, e.g., the Hardy paradox (64). If so, it may be possible to quantitatively explain quantum mechanics in the particle picture. In this review based on the author thesis (152), we consider the theory of the weak value and construct a measurement model to extract the weak value. See the other reviews in Refs. (12; 14; 15; 20).

Let the weak value for an observable  $A$  be defined as

$${}_f\langle A \rangle_i^w := \frac{\langle f|A|i \rangle}{\langle f|i \rangle}, \quad (1)$$

where  $|i\rangle$  and  $|f\rangle$  are called a pre- and post-selected state, respectively. As the naming of the “weak value”, this quantity is experimentally accessible by the weak measurement as explained below. As seen in Fig. 1, the weak value can be measured as the shift of a meter of the probe after the weak interaction between the target and the probe with the specific post-selection of the target. Due to the weak interaction, the quantum state of the target is only slightly changed but the information of the desired observable  $A$  is encoded in the probe by the post-selection. While the previous studies of the weak value since the seminal paper (4), which will be reviewed in Sec. 3, are based on the measurement scheme, there are few works that the weak value is focused on and is independent of the measurement scheme. Furthermore, in these 20 years, we have not yet understood the mathematical properties of the weak value. In this chapter, we review the historical backgrounds of the weak value and the weak measurement and recent development on the measurement model to extract the weak value.

## 2. Review of quantum operation

The time evolution for the quantum state and the operation for the measurement are called a quantum operation. In this section, we review a general description of the quantum operation. Therefore, the quantum operation can describe the time evolution for the quantum state, the

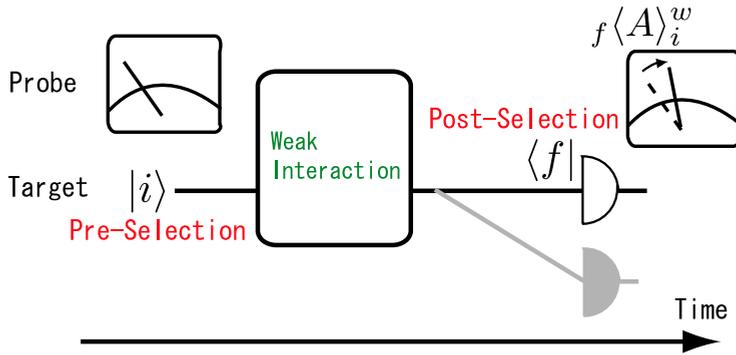


Fig. 1. Schematic figure of the weak measurement.

control of the quantum state, the quantum measurement, and the noisy quantum system in the same formulation.

## 2.1 Historical remarks

Within the mathematical postulates of quantum mechanics (121), the state change is subject to the Schrödinger equation. However, the state change on the measurement is not subject to this but is subject to another axiom, conventionally, von Neumann-Lüders projection postulate (105). See more details on quantum measurement theory in the books (31; 40; 194). Let us consider a state change from the initial state  $|\psi\rangle$  on the *projective measurement*<sup>1</sup> for the operator  $A = \sum_j a_j |a_j\rangle\langle a_j|$ . From the Born rule, the probability to obtain the measurement outcome, that is, the eigenvalue of the observable  $A$ , is given by

$$\Pr[A = a_m] = |\langle a_m | \psi \rangle|^2 = \text{Tr} [|\psi\rangle\langle\psi| \cdot |a_m\rangle\langle a_m|] = \text{Tr} \rho P_{a_m}, \quad (2)$$

where  $\rho := |\psi\rangle\langle\psi|$  and  $P_{a_m} = |a_m\rangle\langle a_m|$ . After the measurement with the measurement outcome  $a_m$ , the quantum state change is given by

$$|\psi\rangle \rightarrow |a_m\rangle, \quad (3)$$

which is often called the “collapse of wavefunction” or “state reduction”. This implies that it is necessary to consider the **non-unitary process** even in the isolated system. To understand the measuring process as quantum dynamics, we need consider the general theory of quantum operations.

## 2.2 Operator-sum representation

Let us recapitulate the general theory of quantum operations of a finite dimensional quantum system (122). All physically realizable quantum operations can be generally described by a completely positive (CP) map (127; 128), since the isolated system of a target system and an auxiliary system always undergoes the unitary evolution according to the axiom of quantum mechanics (121). Physically speaking, the operation of the target system should be described as a positive map, that is, the map from the positive operator to the positive operator, since the density operator is positive. Furthermore, if any auxiliary system is coupled to the target one,

<sup>1</sup> This measurement is often called the *von Neumann measurement* or the *strong measurement*.

the quantum dynamics in the compound system should be also described as the positive map since the compound system should be subject to quantum mechanics. Given the positive map, the positive map is called a CP map if and only if the positive map is also in the compound system coupled to any auxiliary system. One of the important aspects of the CP map is that all physically realizable quantum operations can be described only by operators defined in the target system. Furthermore, the auxiliary system can be environmental system, the probe system, and the controlled system. Regardless to the role of the auxiliary system, the CP map gives the same description for the target system. On the other hand, both quantum measurement and decoherence give the same role for the target system.

Let  $\mathcal{E}$  be a positive map from  $\mathcal{L}(\mathcal{H}_s)$ , a set of linear operations on the Hilbert space  $\mathcal{H}_s$ , to  $\mathcal{L}(\mathcal{H}_s)$ . If  $\mathcal{E}$  is completely positive, its trivial extension  $\mathcal{K}$  from  $\mathcal{L}(\mathcal{H}_s)$  to  $\mathcal{L}(\mathcal{H}_s \otimes \mathcal{H}_e)$  is also positive such that

$$\mathcal{K}(|\alpha\rangle) := (\mathcal{E} \otimes \mathbf{1})(|\alpha\rangle\langle\alpha|) > 0, \quad (4)$$

for an arbitrary state  $|\alpha\rangle \in \mathcal{H}_s \otimes \mathcal{H}_p$ , where  $\mathbf{1}$  is the identity operator. We assume without loss of generality  $\dim\mathcal{H}_s = \dim\mathcal{H}_e < \infty$ . Throughout this chapter, we concentrate on the case that the target state is pure though the generalization to mixed states is straightforward. From the complete positivity, we obtain the following theorem for quantum state changes.

**Theorem 2.1.** *Let  $\mathcal{E}$  be a CP map from  $\mathcal{H}_s$  to  $\mathcal{H}_s$ . For any quantum state  $|\psi\rangle_s \in \mathcal{H}_s$ , there exist a map  $\sigma$  and a pure state  $|\alpha\rangle \in \mathcal{H}_s \otimes \mathcal{H}_e$  such that*

$$\mathcal{E}(|\psi\rangle_s\langle\psi|) = {}_e\langle\tilde{\psi}|\mathcal{K}(|\alpha\rangle)|\tilde{\psi}\rangle_e, \quad (5)$$

where

$$|\psi\rangle_s = \sum_k \psi_k |k\rangle_s, \quad |\tilde{\psi}\rangle_e = \sum_k \psi_k^* |k\rangle_e, \quad (6)$$

which represents the state change for the density operator.

*Proof.* We can write in the Schmidt form as

$$|\alpha\rangle = \sum_m |m\rangle_s |m\rangle_e. \quad (7)$$

We rewrite the right hand sides of Eq. (5) as

$$\begin{aligned} \mathcal{K}(|\alpha\rangle) &= (\mathcal{E} \otimes \mathbf{1}) \left( \sum_{m,n} |m\rangle_s |m\rangle_e {}_s\langle n| {}_e\langle n| \right) \\ &= \sum_{m,n} |m\rangle_e {}_e\langle n| \mathcal{E}(|m\rangle_s {}_s\langle n|), \end{aligned} \quad (8)$$

to obtain

$${}_e\langle m|\mathcal{K}(|\alpha\rangle)|n\rangle_e = \mathcal{E}(|m\rangle_s {}_s\langle n|). \quad (9)$$

By linearity, the desired equation (5) can be derived.  $\square$

From the complete positivity,  $\mathcal{K}(|\alpha\rangle) > 0$  for all  $|\alpha\rangle \in \mathcal{H}_s \otimes \mathcal{H}_e$ , we can express  $\sigma(|\alpha\rangle)$  as

$$\mathcal{K}(|\alpha\rangle) = \sum_m s_m |\hat{s}_m\rangle\langle\hat{s}_m| = \sum_m |s_m\rangle\langle s_m|, \quad (10)$$

where  $s_m$ 's are positive and  $\{|\hat{s}_m\rangle\}$  is a complete orthonormal set with  $|s_m\rangle := \sqrt{s_m}|\hat{s}_m\rangle$ . We define the *Kraus operator*  $E_m$  (95) as

$$E_m|\psi\rangle_s := e\langle\tilde{\psi}|s_m\rangle. \quad (11)$$

Then, the quantum state change becomes the operator-sum representation,

$$\sum_m E_m|\psi\rangle_s\langle\psi|E_m^\dagger = \sum_m e\langle\tilde{\psi}|s_m\rangle\langle s_m|\tilde{\psi}\rangle_e = e\langle\tilde{\psi}|\mathcal{K}(|\alpha\rangle)|\tilde{\psi}\rangle_e = \mathcal{E}(|\psi\rangle_s\langle\psi|).$$

It is emphasized that the quantum state change is described solely in terms of the quantities of the target system.

### 2.3 Indirect quantum measurement

In the following, the operator-sum representation of the quantum state change is related to the indirect measurement model. Consider the observable  $A_s$  and  $B_p$  for the target and probe systems given by

$$A_s = \sum_j a_j|a_j\rangle_s\langle a_j|, \quad B_p = \sum_j b_j|b_j\rangle_p\langle b_j|, \quad (12)$$

respectively. We assume that the interaction Hamiltonian is given by

$$H_{int}(t) = g(A_s \otimes B_p) \delta(t - t_0), \quad (13)$$

where  $t_0$  is measurement time. Here, without loss of generality, the interaction is impulsive and the coupling constant  $g$  is scalar. The quantum dynamics for the compound system is given by

$$|s_m\rangle\langle s_m| = U(|\psi\rangle_s\langle\psi| \otimes |\phi\rangle_p\langle\phi|)U^\dagger, \quad (14)$$

where  $|\psi\rangle_s$  and  $|\phi\rangle_p$  are the initial quantum state on the target and probe systems, respectively. For the probe system, we perform the projective measurement for the observable  $B_p$ . The probability to obtain the measurement outcome  $b_m$  is given by

$$\begin{aligned} \Pr[B_p = b_m] &= \text{Tr}_s\langle b_m|U(|\psi\rangle_s\langle\psi| \otimes |\phi\rangle_p\langle\phi|)U^\dagger|b_m\rangle, \\ &= \text{Tr}_s E_m|\psi\rangle_s\langle\psi|E_m^\dagger = \text{Tr}_s |\psi\rangle_s\langle\psi|M_m, \end{aligned} \quad (15)$$

where the Kraus operator  $E_m$  is defined as

$$E_m := {}_p\langle b_m|U|\phi\rangle_p, \quad (16)$$

and  $M_m := E_m^\dagger E_m$  is called a *positive operator valued measure* (POVM) (45). The POVM has the same role of the spectrum of the operator  $A_s$  in the case of the projective measurement. To derive the projective measurement from the indirect measurement, we set the spectrum of the operator  $A_s$  as the POVM, that is,  $M_m = |a_m\rangle_s\langle a_m|$ . Since the sum of the probability distribution over the measurement outcome equals to one, we obtain

$$\begin{aligned} \sum_m \Pr[B_p = b_m] = 1 &\iff \sum_m \text{Tr} |\psi\rangle_s\langle\psi|M_m = \text{Tr} |\psi\rangle_s\langle\psi| \sum_m M_m = 1 \\ &\rightarrow \sum_m M_m = \mathbf{1}. \end{aligned} \quad (17)$$

Here, the last line uses the property of the density operator,  $\text{Tr} |\psi\rangle_s\langle\psi| = 1$  for any  $|\psi\rangle$ .

### 3. Review of weak value

In Secs. 2.1 and 2.3, the direct and indirect quantum measurement schemes, we only get the probability distribution. However, the probability distribution is not the only thing that is experimentally accessible in quantum mechanics. In quantum mechanics, the phase is also an essential ingredient and in particular the geometric phase is a notable example of an experimentally accessible quantity (150). The general experimentally accessible quantity which contains complete information of the probability and the phase seems to be the *weak value* advocated by Aharonov and his collaborators (4; 14). They proposed a model of weakly coupled system and probe, see Sec. 4.3, to obtain information to a physical quantity as a “weak value” only slightly disturbing the state. Here, we briefly review the formal aspects of the weak value.

For an observable  $A$ , the *weak value*  $\langle A \rangle_w$  is defined as

$$\langle A \rangle_w := \frac{\langle f | U(t_f, t) A U(t, t_i) | i \rangle}{\langle f | U(t_f, t_i) | i \rangle} \in \mathbb{C}, \quad (18)$$

where  $|i\rangle$  and  $\langle f|$  are normalized pre-selected ket and post-selected bra state vectors, respectively (4). Here,  $U(t_2, t_1)$  is an evolution operator from the time  $t_1$  to  $t_2$ . The weak value  $\langle A \rangle_w$  actually depends on the pre- and post-selected states  $|i\rangle$  and  $\langle f|$  but we omit them for notational simplicity in the case that we fix them. Otherwise, we write them explicitly as  ${}_f \langle A \rangle_i^w$  instead for  $\langle A \rangle_w$ . The denominator is assumed to be non-vanishing. This quantity is, in general, in the complex number  $\mathbb{C}$ . Historically, the terminology “weak value” comes from the *weak measurement*, where the coupling between the target system and the probe is weak, explained in the following section. Apart from their original concept of the weak value and the weak measurement, we emphasize that the concept of the weak value is independent of the weak measurement<sup>2</sup>. To take the weak value as *a priori* given quantity in quantum mechanics, we will construct the observable-independent probability space. In the conventional quantum measurement theory, the probability space, more precisely speaking, the probability measure, depends on the observable (151, Sec. 4.1)<sup>3</sup>.

Let us calculate the expectation value in quantum mechanics for the quantum state  $|\psi\rangle$  as

$$\begin{aligned} \text{Ex}[A] &= \langle \psi | A | \psi \rangle = \int d\phi \langle \psi | \phi \rangle \langle \phi | A | \psi \rangle = \int d\phi \langle \psi | \phi \rangle \cdot \langle \phi | \psi \rangle \frac{\langle \phi | A | \psi \rangle}{\langle \phi | \psi \rangle}, \\ &= \int d\phi |\langle \psi | \phi \rangle|^2 \phi \langle A \rangle_\psi^w, \end{aligned} \quad (19)$$

where  $h_A[|\phi\rangle] = \phi \langle A \rangle_\psi^w$  is complex random variable and  $dP := |\langle \phi | \psi \rangle|^2 d\phi$  is the probability measure and is independent of the observable  $A$ . Therefore, the event space  $\Omega = \{|\phi\rangle\}$  is taken as the set of the post-selected state. This formula means that the extended probability theory corresponds to the Born rule. From the conventional definition of the variance in

<sup>2</sup> This concept is shared in Refs. (1; 49; 51; 78; 81; 82; 117; 130).

<sup>3</sup> Due to this, the probability in quantum mechanics cannot be applied to the standard probability theory. As another approach to resolve this, there is the quantum probability theory (138).

quantum mechanics, we obtain the variance as

$$\begin{aligned}
\text{Var}[A] &= \int |h_A[|\phi\rangle]|^2 dP - \left( \int h_A[|\phi\rangle] dP \right)^2 \\
&= \int \left| \frac{\langle \phi|A|\psi\rangle}{\langle \phi|\psi\rangle} \right|^2 |\langle \phi|\psi\rangle|^2 d\phi - \left( \int \frac{\langle \phi|A|\psi\rangle}{\langle \phi|\psi\rangle} |\langle \phi|\psi\rangle|^2 d\phi \right)^2 \\
&= \int |\langle \phi|A|\psi\rangle|^2 d\phi - \left( \int \langle \psi|\phi\rangle \langle \phi|A|\psi\rangle d\phi \right)^2 \\
&= \int \langle \psi|A|\phi\rangle \langle \phi|A|\psi\rangle d\phi - (\langle \psi|A|\psi\rangle)^2 \\
&= \langle \psi|A^2|\psi\rangle - (\langle \psi|A|\psi\rangle)^2.
\end{aligned} \tag{20}$$

This means that the observable-independent probability space can be characterized by the weak value (155). From another viewpoint of the weak value, the statistical average of the weak value coincides with the expectation value in quantum mechanics (7). This can be interpreted as the probability while this allows the "negative probability"<sup>4</sup>. On this idea, the uncertainty relationship was analyzed on the Robertson inequality (58; 163) and on the Ozawa inequality (106), which the uncertainty relationships are reviewed in Ref. (151, Appendix A). Also, the joint probability for the compound system was analyzed in Refs. (27; 30). Furthermore, if the operator  $A$  is a projection operator  $A = |a\rangle\langle a|$ , the above identity becomes an analog of the Bayesian formula,

$$|\langle a|\psi\rangle|^2 = \int_{\phi} \langle a|\langle a|\rangle_{\psi}^w |\langle \phi|\psi\rangle|^2 d\phi. \tag{21}$$

The left hand side is the probability to obtain the state  $|a\rangle$  given the initial state  $|\psi\rangle$ . From this, one may get some intuition by interpreting the weak value  ${}_{\phi}\langle a|\langle a|\rangle_{\psi}^w$  as the complex conditional probability of obtaining the result  $|a\rangle$  under an initial condition  $|i\rangle$  and a final condition  $|f\rangle$  in the process  $|i\rangle \rightarrow |a\rangle \rightarrow |f\rangle$  (170; 171)<sup>5</sup>. Of course, we should not take the strange weak values too literally but the remarkable consistency of the framework of the weak values due to Eq. (21) and a consequence of the completeness relation,

$$\sum_a \langle a|\langle a|\rangle_w = 1, \tag{22}$$

may give a useful concept to further push theoretical consideration by intuition.

This interpretation of the weak values gives many possible examples of strange phenomena like a negative kinetic energy (11), a spin  $100\hbar$  for an electron (4; 23; 52; 60) and a superluminal propagation of light (142; 162) and neutrino (28; 176) motivated by the

<sup>4</sup> The concept of negative probability is not new, e.g., see Refs. (47; 57; 65; 66; 71). The weak value defined by Eq. (18) is normally called the transition amplitude from the state  $|\psi\rangle$  to  $\langle\phi|$  via the intermediate state  $|a\rangle$  for  $A = |a\rangle\langle a|$ , the absolute value squared of which is the probability for the process. But the three references quoted above seem to suggest that they might be interpreted as probabilities in the case that the process is counter-factual, i.e., the case that the intermediate state  $|a\rangle$  is not projectively measured. The description of intermediate state  $|a\rangle$  in the present work is counter-factual or virtual in the sense that the intermediate state would not be observed by projective measurements. Feynman's example is the counter-factual "probability" for an electron to have its spin up in the  $x$ -direction and also spin down in the  $z$ -direction (57).

<sup>5</sup> The interpretation of the weak value as a complex probability is suggested in the literature (118).

OPERA experiment (125). The framework of weak values has been theoretically applied to foundations of quantum physics, e.g., the derivation of the Born rule from the alternative assumption for *a priori* measured value (74), the relationship to the uncertainty relationship (72), the quantum stochastic process (190), the tunneling traverse time (135; 170; 171), arrival time and time operator (21; 39; 146; 147), the decay law (46; 187), the non-locality (32; 180; 181), especially, quantum non-locality, which is characterized by the modular variable, consistent history (87; 188), Bohmian quantum mechanics (98), semi-classical weak values on the tunneling (175), the quantum trajectory (192), and classical stochastic theory (177). Also, in quantum information science, the weak value was analyzed on quantum computation (35; 126), quantum communications (29; 36), quantum estimation, e.g., state tomography (67–69; 111; 158) and the parameter estimation (70; 73; 157), the entanglement concentration (113), the quasi-probability distribution (24; 61; 148; 183) and the cloning of the unknown quantum state with hint (161). Furthermore, this was applied to the cosmological situations in quantum-mechanical region, e.g., the causality (22), the inflation theory (42), backaction of the Hawking radiation from the black hole (34; 54; 55), and the new interpretation of the universe (9; 53; 62). However, the most important fact is that the weak value is experimentally accessible so that the intuitive argument based on the weak values can be either verified or falsified by experiments. There are many experimental proposals to obtain the weak value in the optical (2; 44; 88; 101; 112; 159; 197) and the solid-state (83; 84; 94; 115; 143; 144; 191; 200) systems. Recently, the unified viewpoint was found in the weak measurement (92).

On the realized experiments on the weak value, we can classify the three concepts: (i) testing the quantum theory, (ii) the amplification of the tiny effect in quantum mechanics, and (iii) the quantum phase.

- (i) Testing the quantum theory. The weak value can solve many quantum paradoxes seen in the book (14). The Hardy paradox (64), which there occurs in two Mach-Zehnder interferometers of the electron and the position, was resolved by the weak value (8) and was analyzed deeper (75). This paradoxical situation was experimentally demonstrated in the optical setup (107; 198). By the interference by the polarization (131) and shifting the optical axis (141), the spin beyond the eigenvalue is verified. By the latter technique, the three-box paradox (16; 188) was realized (139). Thereafter, the theoretical progresses are the contextuality on quantum mechanics (178), the generalized N-box paradox (99), and the relationship to the Kirkpatrick game (137). The weak value is used to show the violation of the Leggett-Garg inequality (110; 191). This experimental realizations were demonstrated in the system of the superconducting qubit (97), the optical systems (50; 134). Furthermore, since the weak value for the position observable  $|x\rangle\langle x|$  with the pre-selected state  $|\psi\rangle$  and the post-selection  $|p\rangle$  is given by

$$\langle x|x\rangle_w = \frac{\langle p|x\rangle\langle x|\psi\rangle}{\langle p|\psi\rangle} = \frac{e^{ixp}\psi(x)}{\phi(p)}, \quad (23)$$

we obtain the wavefunction  $\psi(x) := \langle x|\psi\rangle$  as the weak value with the multiplication factor  $1/\phi(0)$  with  $\phi(p) := \langle p|\psi\rangle$  in the case of  $p = 0$ . Using the photon transverse wavefunction, there are experimentally demonstrated by replacing the weak measurement for the position as the polarization measurement (109). This paper was theoretically criticized to compare the standard quantum state tomography for the phase space in Ref. (63) and was generalized to a conventionally unobservable (108). As other examples, there are the detection of the superluminal signal (37), the quantum non-locality (165), and the Bohmian trajectory (91; 149) on the base of the theoretical analysis (193).

- (ii) Amplification of the tiny effect in quantum mechanics. Since the weak value has the denominator, the weak value is very large when the pre- and post-selected states are almost orthogonal<sup>6</sup>. This is practical advantage to use the weak value. While the spin Hall effect of light (124) is too tiny effect to observe its shift in the conventional scheme, by almost orthogonal polarizations for the input and output, this effect was experimentally verified (76) to be theoretically analyzed from the viewpoint of the spin moments (96). Also, some interferometers were applied. The beam deflection on the Sagnac interferometer (48) was shown to be supported by the classical and quantum theoretical analyses (77)<sup>7</sup>. Thereafter, optimizing the signal-to-noise ratio (166; 184), the phase amplification (168; 169), and the precise frequency measurement (167) were demonstrated. As another example, there is shaping the laser pulse beyond the diffraction limit (136). According to Steinberg (172), in his group, the amplification on the single-photon nonlinearity has been progressed to be based on the theoretical proposal (56). While the charge sensing amplification was proposed in the solid-state system (200), there is no experimental demonstration on the amplification for the solid-state system. Furthermore, the upper bound of the amplification has not yet solved. Practically, this open problem is so important to understand the relationship to the weak measurement regime.
- (iii) Quantum phase. The argument of the weak value for the projection operator is the geometric phase as

$$\begin{aligned}
 \gamma &:= \arg \langle \psi_1 | \psi_2 \rangle \langle \psi_2 | \psi_3 \rangle \langle \psi_3 | \psi_1 \rangle \\
 &= \arg \frac{\langle \psi_1 | \psi_2 \rangle \langle \psi_2 | \psi_3 \rangle \langle \psi_3 | \psi_1 \rangle}{|\langle \psi_3 | \psi_1 \rangle|^2} = \arg \frac{\langle \psi_1 | \psi_2 \rangle \langle \psi_2 | \psi_3 \rangle}{\langle \psi_1 | \psi_3 \rangle} \\
 &= \arg_{\psi_1} \langle | \psi_2 \rangle \langle \psi_2 | \rangle_{\psi_3}^w. \tag{24}
 \end{aligned}$$

where the quantum states,  $|\psi_1\rangle$ ,  $|\psi_2\rangle$ , and  $|\psi_3\rangle$ , are the pure states (160). Here, the quantum states,  $|\psi_1\rangle$  and  $|\psi_3\rangle$ , are the post- and pre-selected states, respectively. Therefore, we can evaluate the weak value from the phase shift (174). Of course, vice versa (38). Tamate *et al.* proposal was demonstrated on the relationship to quantum eraser (90) and by the a three-pinhole interferometer (89). The phase shift from the zero mode to  $\pi$  mode was observed by using the interferometer with a Cs vapor (41) and the phase shift in the which-way path experiment was demonstrated (116). Furthermore, by the photonic crystal, phase singularity was demonstrated (164).

- (iv) Miscellaneous. The backaction of the weak measurement is experimentally realized in the optical system (79). Also, the parameter estimation using the weak value is demonstrated (73).

#### 4. Historical background – two-state vector formalism

In this section, we review the original concept of the two-state vector formalism. This theory is seen in the reviewed papers (15; 20).

<sup>6</sup> Unfortunately, the signal to noise ratio is not drastically changed under the assumption that the probe wavefunction is Gaussian on a one-dimensional parameter space.

<sup>7</sup> Unfortunately, the experimental data are mismatched to the theoretical prediction. While the authors claimed that this differences results from the stray of light, the full-order calculation even is not mismatched (93). However, this difference remains the open problem.

#### 4.1 Time symmetric quantum measurement

While the fundamental equations of the microscopic physics are time symmetric, for example, the Newton equation, the Maxwell equation, and the Schrödinger equation <sup>8</sup>, the quantum measurement is not time symmetric. This is because the quantum state after quantum measurement depends on the measurement outcome seen in Sec. 2. The fundamental equations of the microscopic physics can be solved to give the initial boundary condition. To construct the time symmetric quantum measurement, the two boundary conditions, which is called pre- and post-selected states, are needed. The concept of the pre- and post-selected states is called the two-state vector formalism (6). In the following, we review the original motivation to construct the time symmetric quantum measurement.

Let us consider the projective measurement for the observable  $A = \sum_i a_i |a_i\rangle\langle a_i|$  with the initial boundary condition denoted as  $|i\rangle$  at time  $t_i$ . To take quantum measurement at time  $t_0$ , the probability to obtain the measurement outcome  $a_j$  is given by

$$\Pr[A = a_j] = \|\langle a_j|U|i\rangle\|^2, \quad (25)$$

with the time evolution  $U := U(t_0, t_i)$ . After the projective measurement, the quantum state becomes  $|a_j\rangle$ . Thereafter, the quantum state at  $t_f$  is given by  $|\varphi_j\rangle := V|a_j\rangle$  with  $V = U(t_f, t_0)$ . the probability to obtain the measurement outcome  $a_j$  can be rewritten as

$$\Pr[A = a_j] = \frac{\|\langle \varphi_j|V|a_j\rangle\|^2 \|\langle a_j|U|i\rangle\|^2}{\sum_j \|\langle \varphi_j|V|a_j\rangle\|^2 \|\langle a_j|U|i\rangle\|^2}. \quad (26)$$

It is noted that  $\|\langle \varphi_j|V|a_j\rangle\|^2 = 1$ . Here, we consider the backward time evolution from the quantum state  $|\varphi_j\rangle$  at time  $t_f$ . We always obtain the quantum state  $|a_j\rangle$  after the projective measurement at time  $t_0$ . Therefore, the quantum state at time  $t_i$  is given by

$$|\tilde{i}\rangle := U^\dagger|a_j\rangle\langle a_j|V^\dagger|\varphi_j\rangle = U^\dagger|a_j\rangle. \quad (27)$$

In general,  $|\tilde{i}\rangle$  is different from  $|i\rangle$ . Therefore, projective measurement is time asymmetric. To construct the time-symmetric quantum measurement, we add the boundary condition at time  $t_f$ . Substituting the quantum state  $|\varphi_j\rangle$  to the *specific* one denoted as  $|f\rangle$ , which is called the post-selected state, the probability to obtain the measurement outcome  $a_j$ , Eq. (26), becomes

$$\Pr[A = a_j] = \frac{\|\langle f|V|a_j\rangle\|^2 \|\langle a_j|U|i\rangle\|^2}{\sum_j \|\langle f|V|a_j\rangle\|^2 \|\langle a_j|U|i\rangle\|^2}. \quad (28)$$

This is called the Aharonov-Bergmann-Lebowitz (ABL) formula (6). From the analogous discussion to the above, this measurement is time symmetric. Therefore, describing quantum mechanics by the pre- and post-selected states,  $|i\rangle$  and  $\langle f|$ , is called the “two-state vector formalism”.

#### 4.2 Protective measurement

In this subsection, we will see the noninvasive quantum measurement for the specific quantum state on the target system. Consider a system of consisting of a target and a probe defined in the Hilbert space  $\mathcal{H}_s \otimes \mathcal{H}_p$ . The interaction between the target and the probe is

<sup>8</sup> It is, of course, noted that thermodynamics does not have the time symmetric properties from the second law of thermodynamics.

given by

$$H_{int}(t) = g(t)(A \otimes \hat{P}), \quad (29)$$

where

$$\int_0^T g(t)dt =: g_0. \quad (30)$$

The total Hamiltonian is given by

$$H_{tot}(t) = H_s(t) + H_p(t) + H_{int}(t). \quad (31)$$

Here, we suppose that  $H_s(t)$  has discrete and non-degenerate eigenvalues denoted as  $E_i(t)$ . Its corresponding eigenstate is denoted as  $|E_i(t)\rangle$  for any time  $t$ . Furthermore, we consider the discretized time from the time interval  $[0, T]$ ;

$$t_n = \frac{n}{N}T \quad (n = 0, 1, 2, \dots, N), \quad (32)$$

where  $N$  is a sufficiently large number. We assume that the initial target state is the energy eigenvalue  $|E_i(t)\rangle$ <sup>9</sup> the initial probe state is denoted as  $|\xi(0)\rangle$ . Under the adiabatic condition, the compound state for the target and probe systems at time  $T$  is given by

$$\begin{aligned} |\Phi(T)\rangle := & |E_i(t_N)\rangle\langle E_i(t_N)|e^{-i\frac{T}{N}H_{tot}(t_N)}|E_i(t_{N-1})\rangle\langle E_i(t_{N-1})|e^{-i\frac{T}{N}H_{tot}(t_{N-1})} \dots \\ & \times |E_i(t_2)\rangle\langle E_i(t_2)|e^{-i\frac{T}{N}H_{tot}(t_2)}|E_i(t_1)\rangle\langle E_i(t_1)|e^{-i\frac{T}{N}H_{tot}(t_1)}|E_i(0)\rangle \otimes |\xi(0)\rangle. \end{aligned} \quad (33)$$

Applying the Trotter-Suzuki theorem (173; 182), one has

$$\begin{aligned} |\Phi(T)\rangle := & |E_i(t_N)\rangle\langle E_i(t_N)|e^{-i\frac{T}{N}H_{int}(t_N)}|E_i(t_N)\rangle\langle E_i(t_{N-1})|e^{-i\frac{T}{N}H_{int}(t_{N-1})} \dots \\ & \times |E_i(t_3)\rangle\langle E_i(t_2)|e^{-i\frac{T}{N}H_{int}(t_2)}|E_i(t_2)\rangle\langle E_i(t_1)|e^{-i\frac{T}{N}H_{int}(t_1)}|E_i(1)\rangle \otimes |\xi(T)\rangle. \end{aligned} \quad (34)$$

By the Taylor expansion with the respect to  $N$ , the expectation value is

$$\begin{aligned} \langle E_i(t_n)|e^{-i\frac{T}{N}g(t_n)A \otimes \hat{P}}|E_i(t_n)\rangle &= 1 - i\frac{T}{N}g(t_n)\text{Ex}[A(t_n)]\hat{P} - \frac{1}{2}\frac{T^2}{N^2}g^2(t_n)(\text{Ex}[A(t_n)])^2\hat{P}^2 \\ &\quad - \frac{1}{2}\frac{T^2}{N^2}g^2(t_n)\text{Var}[A(t_n)]\hat{P}^2 + O\left(\frac{1}{N^3}\right) \\ &\sim e^{-i\frac{T}{N}g(t_n)\text{Ex}[A(t_n)]\hat{P}} \left(1 - \frac{1}{2}\frac{T^2}{N^2}g^2(t_n)\text{Var}[A(t_n)]\hat{P}^2\right). \end{aligned} \quad (35)$$

<sup>9</sup> Due to this assumption, it is impossible to apply this to the arbitrary quantum state. Furthermore, while we seemingly need the projective measurement, that is, destructive measurement, for the target system to confirm whether the initial quantum state is in the eigenstates (145; 186), they did not apply this to the arbitrary state. For example, if the system is cooled down, we can pickup the ground state of the target Hamiltonian  $H_s(0)$ .

In the limit of  $N \rightarrow \infty$ , by quadrature by parts, we obtain

$$\begin{aligned} |\Phi(T)\rangle &\sim |E_i(T)\rangle \exp \left[ -i \left( \int_0^T g(t) \text{Ex}[A(t)] dt \right) \hat{P} \right] \\ &\quad \times \left[ 1 - \frac{T}{N} \left( \int_0^T g^2(t) \text{Var}[A(t)] dt \right) \hat{P}^2 \right] |\xi(T)\rangle + O\left(\frac{1}{N}\right) \\ &= |E_i(T)\rangle \exp \left[ -i \left( \int_0^T g(t) \text{Ex}[A(t)] dt \right) \hat{P} \right] |\xi(T)\rangle. \end{aligned} \quad (36)$$

Therefore, the shift of the expectation value for the position operator on the probe system is given by

$$\Delta[Q] = \int_0^T g(t) \text{Ex}[A(t)] dt. \quad (37)$$

It is emphasized that the quantum state on the target system remains to be the energy eigenstate of  $H_s$ . Therefore, this is called the *protective measurement* (5; 18). It is remarked that the generalized version of the protective measurement in Ref. (19) by the pre- and post-selected states and in Ref. (10) by the meta-stable state.

### 4.3 Weak measurement

From the above discussions, is it possible to combine the above two concepts, i.e., the time-symmetric quantum measurement without destroying the quantum state (189)? This answer is the *weak measurement* (4). Consider a target system and a probe defined in the Hilbert space  $\mathcal{H}_s \otimes \mathcal{H}_p$ . The interaction of the target system and the probe is assumed to be weak and instantaneous,

$$H_{int}(t) = g(A \otimes \hat{P})\delta(t - t_0), \quad (38)$$

where an observable  $A$  is defined in  $\mathcal{H}_s$ , while  $\hat{P}$  is the momentum operator of the probe. The time evolution operator becomes  $e^{-ig(A \otimes \hat{P})}$ . Suppose the probe initial state is  $|\xi\rangle$ . For the transition from the pre-selected state  $|i\rangle$  to the post-selected state  $|f\rangle$ , the probe wave function becomes  $|\xi'\rangle = \langle f|V e^{-ig(A \otimes \hat{P})} U|i\rangle|\xi\rangle$ , which is in the weak coupling case,

$$\begin{aligned} |\xi'\rangle &= \langle f|V e^{-ig(A \otimes \hat{P})} U|i\rangle|\xi\rangle = \langle f|V[\mathbf{1} - ig(A \otimes \hat{P})]U|i\rangle|\xi\rangle + O(g^2) \\ &= \langle f|VU|i\rangle - ig\langle f|VAU|i\rangle \otimes \hat{P}|\xi\rangle + O(g^2) = \langle f|VU|i\rangle (1 - ig\langle A \rangle_w \hat{P}) |\xi\rangle + O(g^2) \end{aligned} \quad (39)$$

where  $\langle f|VAU|i\rangle / \langle f|VU|i\rangle = \langle A \rangle_w$ . Here, the last equation uses the approximation that  $g\langle A \rangle_w \ll 1$ <sup>10</sup>. We obtain the shifts of the expectation values for the position and momentum operators on the probe as the following theorem:

**Theorem 4.1** (Jozsa (85)). *We obtain the shifts of the expectation values for the position and momentum operators on the probe after the weak measurement with the post-selection as*

$$\Delta[\hat{Q}] = g\text{Re}\langle A \rangle_w + mg\text{Im}\langle A \rangle_w \left. \frac{d \text{Var}[\hat{Q}]}{dt} \right|_{t=t_0}, \quad (40)$$

$$\Delta[\hat{P}] = 2g\text{Im}\langle A \rangle_w \text{Var}[\hat{P}], \quad (41)$$

<sup>10</sup> It is remarked that Wu and Li showed the second-order correction of the weak measurement (196). A further analysis was shown in Refs. (129; 132).

where

$$\Delta[\hat{Q}] := \frac{\langle \xi' | \hat{Q} | \xi' \rangle}{\langle \xi' | \xi' \rangle} - \langle \xi | \hat{Q} | \xi \rangle, \quad (42)$$

$$\Delta[\hat{P}] := \frac{\langle \xi' | \hat{P} | \xi' \rangle}{\langle \xi' | \xi' \rangle} - \langle \xi | \hat{P} | \xi \rangle, \quad (43)$$

$$\text{Var}[\hat{Q}] := \langle \xi | \hat{Q}^2 | \xi \rangle - (\langle \xi | \hat{Q} | \xi \rangle)^2, \quad (44)$$

$$\text{Var}[\hat{P}] := \langle \xi | \hat{P}^2 | \xi \rangle - (\langle \xi | \hat{P} | \xi \rangle)^2. \quad (45)$$

Here, the probe Hamiltonian is assumed as

$$\hat{H} = \frac{\hat{P}^2}{2m} + V(Q), \quad (46)$$

where  $V(Q)$  is the potential on the coordinate space.

*Proof.* For the probe observable  $\hat{M}$ , we obtain

$$\begin{aligned} \frac{\langle \xi' | \hat{M} | \xi' \rangle}{\langle \xi' | \xi' \rangle} &= \frac{\langle \xi | \hat{M} | \xi \rangle - ig \langle A \rangle_w \langle \xi | \hat{M} \hat{P} | \xi \rangle + ig \overline{\langle A \rangle_w} \langle \xi | \hat{P} \hat{M} | \xi \rangle}{\langle \xi | \xi \rangle - ig \langle A \rangle_w \langle \xi | \hat{P} | \xi \rangle + ig \overline{\langle A \rangle_w} \langle \xi | \hat{P} | \xi \rangle} \\ &= \frac{\langle \xi | \hat{M} | \xi \rangle + ig \text{Re} \langle A \rangle_w \langle \xi | [\hat{P}, \hat{M}] | \xi \rangle + g \text{Im} \langle A \rangle_w \langle \xi | \{\hat{P}, \hat{M}\} | \xi \rangle}{\langle \xi | \xi \rangle + 2g \text{Im} \langle A \rangle_w \langle \xi | \hat{P} | \xi \rangle} \\ &= (\langle \xi | \hat{M} | \xi \rangle + ig \text{Re} \langle A \rangle_w \langle \xi | [\hat{P}, \hat{M}] | \xi \rangle + g \text{Im} \langle A \rangle_w \langle \xi | \{\hat{P}, \hat{M}\} | \xi \rangle) \\ &\quad \times (1 - 2g \text{Im} \langle A \rangle_w \langle \xi | \hat{P} | \xi \rangle) + O(g^2) \\ &= \langle \xi | \hat{M} | \xi \rangle + ig \text{Re} \langle A \rangle_w \langle \xi | [\hat{P}, \hat{M}] | \xi \rangle \\ &\quad + g \text{Im} \langle A \rangle_w (\langle \xi | \{\hat{P}, \hat{M}\} | \xi \rangle - 2 \langle \xi | \hat{M} | \xi \rangle \langle \xi | \hat{P} | \xi \rangle) + O(g^2). \end{aligned} \quad (47)$$

If we set  $\hat{M} = \hat{P}$ , one has

$$\Delta[\hat{P}] = 2g \text{Im} \langle A \rangle_w \text{Var}[\hat{P}]. \quad (48)$$

If instead we set  $\hat{M} = \hat{Q}$ , one has

$$\Delta[\hat{Q}] = g \text{Re} \langle A \rangle_w + g \text{Im} \langle A \rangle_w (\langle \xi | \{\hat{P}, \hat{Q}\} | \xi \rangle - 2g \langle \xi | \hat{Q} | \xi \rangle \langle \xi | \hat{P} | \xi \rangle) \quad (49)$$

since  $[\hat{P}, \hat{Q}] = -i$ . From the Heisenberg equation with the probe Hamiltonian (46), we obtain the Ehrenfest theorem;

$$i \frac{d}{dt} \langle \xi | \hat{Q} | \xi \rangle = \langle \xi | [\hat{Q}, \hat{H}] | \xi \rangle = i \frac{\langle \xi | \hat{P} | \xi \rangle}{m} \quad (50)$$

$$i \frac{d}{dt} \langle \xi | \hat{Q}^2 | \xi \rangle = \langle \xi | [\hat{Q}^2, \hat{H}] | \xi \rangle = i \frac{\langle \xi | \{\hat{P}, \hat{Q}\} | \xi \rangle}{m}. \quad (51)$$

Substituting them into Eq. (49), we derive

$$\Delta[\hat{Q}] = g\text{Re}\langle A \rangle_w + mg\text{Im}\langle A \rangle_w \left. \frac{d \text{Var}[\hat{Q}]}{dt} \right|_{t=t_0} \quad (52)$$

since the interaction to the target system is taken at time  $t = t_0$ .  $\square$

Putting together, we can measure the weak value  $\langle A \rangle_w$  by observing the shift of the expectation value of the probe both in the coordinate and momentum representations. The shift of the probe position contains the future information up to the post-selected state.

**Corollary 4.2.** *When the probe wavefunction is real-valued in the coordinate representation, Eq. (40) can be reduced to*

$$\Delta[\hat{Q}] = g\text{Re}\langle A \rangle_w. \quad (53)$$

*Proof.* From the Schrödinger equation in the coordinate representation;

$$i \frac{\partial}{\partial t} \xi(Q) = \frac{1}{2m} \frac{\partial^2}{\partial Q^2} \xi(Q) + V(Q) \xi(Q), \quad (54)$$

where  $\xi(Q) \equiv \langle Q | \xi \rangle$ , putting  $\xi(Q) = R(Q) e^{iS(Q)}$ , we obtain the equation for the real part as

$$\frac{\partial}{\partial t} R(Q) + \frac{\partial}{\partial Q} \left( \frac{R(Q) \frac{\partial}{\partial Q} S(Q)}{m} \right) = 0. \quad (55)$$

Therefore, if the probe wavefunction is real-valued in the coordinate representation, one has  $\frac{\partial}{\partial Q} S(Q) = 0$  to obtain  $\frac{\partial}{\partial t} R = 0$ . Therefore, we obtain

$$\frac{d \text{Var}[\hat{Q}]}{dt} = 0 \quad (56)$$

for any time  $t$ . Vice versa. From this statement, we obtain the desired result from Eq. (40).  $\square$

It is noted that there are many analyses on the weak measurement, e.g., on the phase space (102), on the finite sample (179), on the counting statics (26; 104), on the non-local observable (32; 33), and on the complementary observable (197).

Summing up this section, the two-state vector formalism is called if the pre- and post-selected states are prepared and the weak or strong measurement is taken in the von-Neumann type Hamiltonian,  $H = gA\hat{P}\delta(t - t_0)$  between the pre- and post-selected states. In the case of the strong measurement, we obtain the expectation value  $\text{Ex}(A)$  in the probe. On the other hand, in the case of the weak measurement, we obtain the weak value  $\langle A \rangle_w$  in the probe.

## 5. Weak-value measurement for a qubit system

In this subsection, we consider the weak measurement in the case that the probe system is a qubit system (195). In general, the interaction Hamiltonian is given by

$$H_{int} = g[A \otimes (\vec{v} \cdot \vec{\sigma})] \delta(t - t_0), \quad (57)$$

where  $\vec{v}$  is a unit vector. Expanding the interaction Hamiltonian for the pre- and post-selected states,  $|\psi\rangle$  and  $|\phi\rangle$ , respectively up to the first order for  $g$ , we obtain the shift of the expectation

value for  $\vec{q} \cdot \vec{\sigma}$  as

$$\begin{aligned} \Delta[\vec{q} \cdot \vec{\sigma}] &= \frac{\langle \xi' | [\vec{q} \cdot \vec{\sigma}] | \xi' \rangle}{\langle \xi' | \xi' \rangle} - \langle \xi | [\vec{q} \cdot \vec{\sigma}] | \xi \rangle = g \langle \xi | i[\vec{v} \cdot \vec{\sigma}, \vec{q} \cdot \vec{\sigma}] | \xi \rangle \text{Re} \langle A \rangle_w \\ &\quad + g (\langle \xi | \{ \vec{v} \cdot \vec{\sigma}, \vec{q} \cdot \vec{\sigma} \} | \xi \rangle - 2 \langle \xi | \vec{v} \cdot \vec{\sigma} | \xi \rangle \langle \xi | \vec{q} \cdot \vec{\sigma} | \xi \rangle) \text{Im} \langle A \rangle_w + O(g^2) \\ &= 2g \{ (\vec{q} \times \vec{v}) \cdot \vec{m} \} \text{Re} \langle A \rangle_w + 2g \{ \vec{v} \cdot \vec{q} - (\vec{v} \cdot \vec{m})(\vec{q} \cdot \vec{m}) \} \text{Im} \langle A \rangle_w + O(g^2), \end{aligned} \quad (58)$$

where

$$|\xi'\rangle = \langle \phi | e^{-ig[A \otimes (\vec{v} \cdot \vec{\sigma})]} | \psi \rangle | \xi \rangle, \quad (59)$$

$$|\xi\rangle\langle\xi| =: \frac{1}{2}(\mathbf{1} + \vec{m} \cdot \vec{\sigma}). \quad (60)$$

Furthermore, the pre- and post-selected states are assumed to be

$$|\psi\rangle\langle\psi| =: \frac{1}{2}(\mathbf{1} + \vec{r}_i \cdot \vec{\sigma}), \quad (61)$$

$$|\phi\rangle\langle\phi| =: \frac{1}{2}(\mathbf{1} + \vec{r}_f \cdot \vec{\sigma}).$$

Since the weak value of the observable  $\vec{n} \cdot \vec{\sigma}$  is

$$\langle \vec{n} \cdot \vec{\sigma} \rangle_w = \frac{\langle \phi | \vec{n} \cdot \vec{\sigma} | \psi \rangle \langle \psi | \phi \rangle}{|\langle \phi | \psi \rangle|^2} = \vec{n} \cdot \frac{\vec{r}_i + \vec{r}_f + i(\vec{r}_i \times \vec{r}_f)}{1 + \vec{r}_i \cdot \vec{r}_f}, \quad (62)$$

we obtain

$$\Delta[\vec{q} \cdot \vec{\sigma}] = 2g \{ (\vec{q} \times \vec{v}) \cdot \vec{m} \} \frac{\vec{n} \cdot (\vec{r}_i + \vec{r}_f)}{1 + \vec{r}_i \cdot \vec{r}_f} + 2g \{ \vec{v} \cdot \vec{q} - (\vec{v} \cdot \vec{m})(\vec{q} \cdot \vec{m}) \} \frac{\vec{n} \cdot (\vec{r}_i \times \vec{r}_f)}{1 + \vec{r}_i \cdot \vec{r}_f} + O(g^2). \quad (63)$$

From Eq. (63), we can evaluate the real and imaginary parts of the weak value changing the parameter of the measurement direction  $\vec{q}$ . This calculation is used in the context of the Hamiltonian estimation (157).

Next, as mentioned before, we emphasize that the weak measurement is only one of the methods to obtain the weak value. There are many other approaches to obtain the weak value, e.g., on changing the probe state (59; 80; 103; 119), and on the entangled probe state (114). Here, we show another method to obtain the weak value in the case that the target and the probe systems are both qubit systems (133).

Let  $|\psi\rangle_s := \alpha|0\rangle_s + \beta|1\rangle_s$  be the pre-selected state for the target system. The initial probe state can be described as  $|\xi\rangle_p := \gamma|0\rangle_p + \eta|1\rangle_p$ . It is emphasized that the initial probe state is controllable. Here, the initial states are normalized, that is,  $|\alpha|^2 + |\beta|^2 = 1$  and  $|\gamma|^2 + |\eta|^2 = 1$ . Applying the Controlled-NOT (C-NOT) gate, we make a transform of the quantum state for the compound system to

$$|\psi\rangle_s \otimes |\xi\rangle_p \xrightarrow{\text{C-NOT}} |\Psi_c\rangle := (\alpha\gamma|0\rangle_s + \beta\eta|1\rangle_s)|0\rangle_p + (\alpha\eta|0\rangle_s + \beta\gamma|1\rangle_s)|1\rangle_p. \quad (64)$$

In the case of  $\gamma \sim 1$ , we obtain the compound state as

$$\alpha|0\rangle_s|0\rangle_p + \beta|1\rangle_s|1\rangle_p, \quad (65)$$

and similarly, in the case of  $\eta \sim 1$ , one has

$$\alpha|0\rangle_s|1\rangle_p + \beta|1\rangle_s|0\rangle_p. \quad (66)$$

Those cases can be taken as the standard von Neumann projective measurement. For the post-selected state  $|\phi\rangle$ , the probability to obtain the measurement outcome  $k$  on the probe is

$$\begin{aligned} \Pr[k] &:= \frac{\|({}_s\langle\phi| \otimes {}_p\langle k|) |\Psi_c\rangle\|^2}{\sum_{m \in \{0,1\}} \|({}_s\langle\phi| \otimes {}_p\langle m|) |\Psi_c\rangle\|^2} \\ &= \frac{|({}_s\langle\phi|0\rangle_s\langle 0|\psi\rangle_s\gamma + {}_s\langle\phi|1\rangle_s\langle 1|\psi\rangle_s\eta) \delta_{k,0} + ({}_s\langle\phi|0\rangle_s\langle 0|\psi\rangle_s\eta + {}_s\langle\phi|1\rangle_s\langle 1|\psi\rangle_s\gamma) \delta_{k,1}|^2}{\sum_{m \in \{0,1\}} \|({}_s\langle\phi| \otimes {}_p\langle m|) |\Psi_c\rangle\|^2} \\ &= \frac{|(\gamma - \eta) {}_s\langle\phi|k\rangle_s\langle k|\psi\rangle_s + \eta {}_s\langle\phi|\psi\rangle_s|^2}{|(\gamma - \eta) {}_s\langle\phi|0\rangle_s\langle 0|\psi\rangle_s + \eta {}_s\langle\phi|\psi\rangle_s|^2 + |(\gamma - \eta) {}_s\langle\phi|1\rangle_s\langle 1|\psi\rangle_s + \eta {}_s\langle\phi|\psi\rangle_s|^2} \\ &= \frac{|(\gamma - \eta) {}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w + \eta|^2}{1 - (\gamma - \eta)^2(1 - \sum_{m \in \{0,1\}} |{}_\phi\langle|m\rangle_s\langle m|\rangle_\psi^w|^2)}. \end{aligned} \quad (67)$$

Here, in the last line, the parameters  $\gamma$  and  $\eta$  are assumed to be real. Without the post-selection, the POVM to obtain the measurement outcome  $k$  is

$$E_k = (\gamma^2 - \eta^2)|k\rangle_s\langle k| + \eta^2. \quad (68)$$

Here, the coefficient of the first term means that the strength of measurement and the second term is always added. Therefore, we define the quantity to distinguish the probability for the measurement outcome  $k$  as

$$R[k] := \frac{\Pr[k] - \eta^2}{(\gamma^2 - \eta^2)}. \quad (69)$$

Putting together Eqs. (67) and (69), we obtain

$$R[k] = \frac{2\eta(\gamma - \eta)\text{Re} {}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w + (\gamma - \eta)^2[|{}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w|^2 + \eta^2(1 - |{}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w|^2)]}{(\gamma^2 - \eta^2)[1 - (\gamma - \eta)^2(1 - \sum_{m \in \{0,1\}} |{}_\phi\langle|m\rangle_s\langle m|\rangle_\psi^w|^2)]}. \quad (70)$$

Setting the parameters;

$$\gamma = \sqrt{\frac{1}{2} + \epsilon}, \quad \eta = \sqrt{\frac{1}{2} - \epsilon}, \quad (71)$$

one has

$$\begin{aligned} R[k] &= \frac{(1 - \epsilon)\text{Re} {}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w + \epsilon \left[ |{}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w|^2 + \left(\frac{1}{2} - \epsilon\right) (1 - |{}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w|^2) \right]}{2 \left[ 1 - \epsilon^2 \left( 1 - \sum_{m \in \{0,1\}} |{}_\phi\langle|m\rangle_s\langle m|\rangle_\psi^w|^2 \right) \right]} + O(\epsilon^2), \\ &= \frac{1}{2}\text{Re} {}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w - \frac{\epsilon}{2} \left( \text{Re} {}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w - \frac{1}{2}|{}_\phi\langle|k\rangle_s\langle k|\rangle_\psi^w|^2 \right) + O(\epsilon^2). \end{aligned} \quad (72)$$

From Eq. (72), it is possible to obtain the real part of the weak value from the first term and its imaginary part from the second term. Since the first order of the parameter  $\epsilon$  is the gradient on

changing the initial probe state from  $|\xi\rangle_p = \frac{1}{\sqrt{2}}(|0\rangle_p + |1\rangle_p)$ , realistically, we can evaluate the imaginary part of the weak value from the gradient of the readout. This method is also used in Ref. (198) on the joint weak value. It is emphasized that the weak value can be experimentally accessible by changing the initial probe state while the interaction is **not weak**<sup>11</sup>.

## 6. Weak values for arbitrary coupling quantum measurement

We just calculate an arbitrary coupling between the target and the probe systems (93; 120; 199). Throughout this section, we assume that the desired observable is the projection operator to be denoted as  $A^2 = A$  (153). In the case of the von-Neumann interaction motivated by the original work (4), when the pre- and post-selected states are  $|i\rangle$  and  $|f\rangle$ , respectively, and the probe state is  $|\xi\rangle$ , the probe state  $|\xi'\rangle$  after the interaction given by  $H_{int} = gA\hat{P}$  becomes

$$\begin{aligned} |\xi'\rangle &= \langle f|e^{-igA\hat{P}}|i\rangle|\xi\rangle = \langle f|\left(1 + \sum_{k=1}^{\infty} \frac{1}{k!}(-igA\hat{P})^k\right)|i\rangle|\xi\rangle = \langle f|\left(1 + A \sum_{k=1}^{\infty} \frac{1}{k!}(-ig\hat{P})^k\right)|i\rangle|\xi\rangle \\ &= \langle f|\left(1 - A + A \sum_{k=0}^{\infty} \frac{1}{k!}(-ig\hat{P})^k\right)|i\rangle|\xi\rangle = \langle f|\left(1 - A + Ae^{-ig\hat{P}}\right)|i\rangle|\xi\rangle \\ &= \langle f|i\rangle\left(1 - \langle A\rangle_w + \langle A\rangle_w e^{-ig\hat{P}}\right)|\xi\rangle. \end{aligned} \quad (73)$$

It is remarked that the desired observable  $B$ , which satisfies  $B^2 = 1$  (93; 120), corresponds to  $B = 2A - 1$ . Analogous to Theorem 4.1, we can derive the expectation values of the position and the momentum after the weak measurement. These quantities depends on the weak value  $\langle A\rangle_w$  and the generating function for the position and the momentum of the initial probe state  $|\xi\rangle$ .

## 7. Weak value with decoherence

The decoherence results from the coupled system to the environment and leads to the transition from the quantum to classical systems. The general framework of the decoherence was discussed in Sec. 2. In this section, we discuss the analytical expressions for the weak value.

While we directly discuss the weak value with decoherence, the weak value is defined as a complex number. To analogously discuss the density operator formalism, we need the operator associated with the weak value. Therefore, we define a *W operator*  $W(t)$  as

$$W(t) := U(t, t_i)|i\rangle\langle f|U(t_f, t). \quad (74)$$

To facilitate the formal development of the weak value, we introduce the ket state  $|\psi(t)\rangle$  and the bra state  $\langle\phi(t)|$  as

$$|\psi(t)\rangle = U(t, t_i)|i\rangle, \quad \langle\phi(t)| = \langle f|U(t_f, t), \quad (75)$$

so that the expression for the  $W$  operator simplifies to

$$W(t) = |\psi(t)\rangle\langle\phi(t)|. \quad (76)$$

<sup>11</sup> This point seems to be misunderstood. According to Ref. (134), the violation of the Leggett-Garg inequality (100) was shown, but the macroscopic realism cannot be denied since the noninvasive measurability is not realized.

By construction, the two states  $|\psi(t)\rangle$  and  $\langle\phi(t)|$  satisfy the Schrödinger equations with the same Hamiltonian with the initial and final conditions  $|\psi(t_i)\rangle = |i\rangle$  and  $\langle\phi(t_f)| = \langle f|$ . In a sense,  $|\psi(t)\rangle$  evolves forward in time while  $\langle\phi(t)|$  evolves backward in time. The time reverse of the  $W$  operator (76) is  $W^\dagger = |\phi(t)\rangle\langle\psi(t)|$ . Thus, we can say the  $W$  operator is based on the two-state vector formalism formally described in Refs. (16; 17). Even an apparently similar quantity to the  $W$  operator (76) was introduced by Reznik and Aharonov (140) in the name of "two-state" with the conceptually different meaning. This is because the  $W$  operator acts on a Hilbert space  $\mathcal{H}$  but the two-state vector acts on the Hilbert space  $\overrightarrow{\mathcal{H}}_1 \otimes \overleftarrow{\mathcal{H}}_2$ . Furthermore, while the generalized two-state, which is called a multiple-time state, was introduced (13), this is essentially reduced to the two-state vector formalism. The  $W$  operator gives the weak value of the observable  $A$ <sup>12</sup> as

$$\langle A \rangle_W = \frac{\text{Tr}(WA)}{\text{Tr } W}, \quad (77)$$

in parallel with the expectation value of the observable  $A$  by

$$\text{Ex}[A] = \frac{\text{Tr}(\rho A)}{\text{Tr } \rho} \quad (78)$$

from Born's rule. Furthermore, the  $W$  operator (74) can be regarded as a special case of a standard purification of the density operator (185). In our opinion, the  $W$  operator should be considered on the same footing of the density operator. For a closed system, both satisfy the Schrödinger equation. In a sense, the  $W$  operator  $W$  is the square root of the density operator since

$$W(t)W^\dagger(t) = |\psi(t)\rangle\langle\psi(t)| = U(t, t_i)|i\rangle\langle i|U^\dagger(t, t_i), \quad (79)$$

which describes a state evolving forward in time for a given initial state  $|\psi(t_i)\rangle\langle\psi(t_i)| = |i\rangle\langle i|$ , while

$$W^\dagger(t)W(t) = |\phi(t)\rangle\langle\phi(t)| = U(t_f, t)|f\rangle\langle f|U^\dagger(t_f, t), \quad (80)$$

which describes a state evolving backward in time for a given final state  $|\phi(t_f)\rangle\langle\phi(t_f)| = |f\rangle\langle f|$ . The  $W$  operator describes the entire history of the state from the past ( $t_i$ ) to the future ( $t_f$ ) and measurement performed at the time  $t_0$  as we shall see in Appendix 4.3. This description is conceptually different from the conventional one by the time evolution of the density operator. From the viewpoint of geometry, the  $W$  operator can be taken as the Hilbert-Schmidt bundle. The bundle projection is given by

$$\Pi : W(t) \rightarrow \rho_i(t) := W(t)W^\dagger(t). \quad (81)$$

When the dimension of the Hilbert space is  $N$ :  $\dim\mathcal{H} = N$ , the structure group of this bundle is  $U(N)$  (25, Sec. 9.3). Therefore, the  $W$  operator has richer information than the density operator formalism as we shall see a typical example of a geometric phase (155). Furthermore, we can express the probability to get the measurement outcome  $a_n \in A$  due to the ABL formula (28) using the  $W$  operator  $W$  as

$$\text{Pr}[A = a_n] = \frac{|\text{Tr } WP_{a_n}|^2}{\sum_n |\text{Tr } WP_{a_n}|^2}, \quad (82)$$

<sup>12</sup> While the original notation of the weak values is  $\langle A \rangle_w$  indicating the "w"eak value of an observable  $A$ , our notation is motivated by one of which the pre- and post-selected states are explicitly shown as  ${}_f\langle A \rangle_i^w$ .

where  $A = \sum_n a_n |a_n\rangle\langle a_n| =: \sum_n a_n P_{a_n}$ . This shows the usefulness of the  $W$  operator. Let us discuss a state change in terms of the  $W$  operator and define a map  $\mathcal{X}$  as

$$\mathcal{X}(|\alpha\rangle, |\beta\rangle) := (\mathcal{E} \otimes \mathbf{1})(|\alpha\rangle\langle\beta|), \quad (83)$$

for an arbitrary  $|\alpha\rangle, |\beta\rangle \in \mathcal{H}_s \otimes \mathcal{H}_e$ . Then, we obtain the following theorem on the change of the  $W$  operator such as Theorem 2.1.

**Theorem 7.1.** *For any  $W$  operator  $W = |\psi(t)\rangle_s\langle\phi(t)|$ , we expand*

$$|\psi(t)\rangle_s = \sum_m \psi_m |\alpha_m\rangle_s, \quad |\phi(t)\rangle_s = \sum_m \phi_m |\beta_m\rangle_s, \quad (84)$$

with fixed complete orthonormal sets  $\{|\alpha_m\rangle_s\}$  and  $\{|\beta_m\rangle_s\}$ . Then, a change of the  $W$  operator can be written as

$$\mathcal{E}(|\psi(t)\rangle_s\langle\phi(t)|) = {}_e\langle\tilde{\psi}(t)|\mathcal{X}(|\alpha\rangle, |\beta\rangle)|\tilde{\phi}(t)\rangle_e, \quad (85)$$

where

$$|\tilde{\psi}(t)\rangle_e = \sum_k \psi_k^* |\alpha_k\rangle_e, \quad |\tilde{\phi}(t)\rangle_e = \sum_k \phi_k^* |\beta_k\rangle_e, \quad (86)$$

and  $|\alpha\rangle$  and  $|\beta\rangle$  are maximally entangled states defined by

$$|\alpha\rangle := \sum_m |\alpha_m\rangle_s |\alpha_m\rangle_e, \quad |\beta\rangle := \sum_m |\beta_m\rangle_s |\beta_m\rangle_e. \quad (87)$$

Here,  $\{|\alpha_m\rangle_e\}$  and  $\{|\beta_m\rangle_e\}$  are complete orthonormal sets corresponding to  $\{|\alpha_m\rangle_s\}$  and  $\{|\beta_m\rangle_s\}$ , respectively.

The proof is completely parallel to that of Theorem 2.1.

**Theorem 7.2.** *For any  $W$  operator  $W = |\psi(t)\rangle_s\langle\phi(t)|$ , given the CP map  $\mathcal{E}$ , the operator-sum representation is written as*

$$\mathcal{E}(W) = \sum_m E_m W F_m^\dagger, \quad (88)$$

where  $E_m$  and  $F_m$  are the Kraus operators.

It is noted that, in general,  $\mathcal{E}(W)\mathcal{E}(W^\dagger) \neq \mathcal{E}(\rho)$  although  $\rho = WW^\dagger$ .

*Proof.* We take the polar decomposition of the map  $X$  to obtain

$$\mathcal{X} = \mathcal{K}u, \quad (89)$$

noting that

$$\mathcal{X}\mathcal{X}^\dagger = \mathcal{K}uu^\dagger\mathcal{K} = \mathcal{K}^2. \quad (90)$$

The unitary operator  $u$  is well-defined on  $\mathcal{H}_s \otimes \mathcal{H}_e$  because  $\mathcal{K}$  defined in Eq. (4) is positive. This is a crucial point to obtain this result (88), which is the operator-sum representation for the quantum operation of the  $W$  operator. From Eq. (10), we can rewrite  $\mathcal{X}$  as

$$\mathcal{X} = \sum_m |s_m\rangle\langle s_m|u = \sum_m |s_m\rangle\langle t_m|, \quad (91)$$

where

$$\langle t_m| = \langle s_m|u. \quad (92)$$

Similarly to the Kraus operator (16), we define the two operators,  $E_m$  and  $F_m^\dagger$ , as

$$E_m |\psi(t)\rangle_s := {}_e\langle \tilde{\psi}(t) | s_m \rangle, \quad {}_s\langle \phi(t) | F_m^\dagger := \langle t_m | \tilde{\phi}(t) \rangle_e, \quad (93)$$

where  $|\tilde{\psi}(t)\rangle_e$  and  $|\tilde{\phi}(t)\rangle_e$  are defined in Eq. (86). Therefore, we obtain the change of the  $W$  operator as

$$\begin{aligned} \sum_m E_m |\psi(t)\rangle_s \langle \phi(t) | F_m^\dagger &= \sum_m {}_e\langle \tilde{\psi}(t) | s_m \rangle \langle t_m | \tilde{\phi}(t) \rangle_e = {}_e\langle \tilde{\psi}(t) | \mathcal{X} | \tilde{\phi}(t) \rangle_e \\ &= \mathcal{E}(|\psi(t)\rangle_s \langle \phi(t)|), \end{aligned} \quad (94)$$

using Theorem 7.1 in the last line. By linearity, we obtain the desired result.  $\square$

Summing up, we have introduced the  $W$  operator (74) and obtained the general form of the quantum operation of the  $W$  operator (88) in an analogous way to the quantum operation of the density operator assuming the complete positivity of the physical operation. This can be also described from information-theoretical approach (43) to solve the open problem listed in Ref. (13, Sec. XII). However, this geometrical meaning has still been an open problem.

It is well established that the trace preservation,  $\text{Tr}(\mathcal{E}(\rho)) = \text{Tr} \rho = 1$  for all  $\rho$ , implies that  $\sum_m E_m^\dagger E_m = 1$ . As discussed in Eq. (17), the proof goes through as

$$1 = \text{Tr}(\mathcal{E}(\rho)) = \text{Tr} \left( \sum_m E_m \rho E_m^\dagger \right) = \text{Tr} \left( \sum_m E_m^\dagger E_m \rho \right) \quad (\forall \rho). \quad (95)$$

This argument for the density operator  $\rho = WW^\dagger$  applies also for  $W^\dagger W$  to obtain  $\sum_m F_m^\dagger F_m = 1$  because this is the density operator in the time reversed world in the two-state vector formulation as reviewed in Sec. 4. Therefore, we can express the Kraus operators,

$$E_m = {}_e\langle e_m | U | e_i \rangle_e, \quad F_m^\dagger = {}_e\langle e_f | V | e_m \rangle_e, \quad (96)$$

where

$$U = U(t, t_i), \quad V = U(t_f, t), \quad (97)$$

are the evolution operators, which act on  $\mathcal{H}_s \otimes \mathcal{H}_e$ .  $|e_i\rangle$  and  $|e_f\rangle$  are some basis vectors and  $|e_m\rangle$  is a complete set of basis vectors with  $\sum_m |e_m\rangle\langle e_m| = 1$ . We can compute

$$\sum_m F_m^\dagger E_m = \sum_m {}_e\langle e_f | V | e_m \rangle_e \langle e_m | U | e_i \rangle_e = {}_e\langle e_f | V U | e_i \rangle_e. \quad (98)$$

The above equality (98) may be interpreted as a decomposition of the history in analogy to the decomposition of unity because

$${}_e\langle e_f | V U | e_i \rangle_e = {}_e\langle e_f | S | e_i \rangle_e = S_{fi} \quad (99)$$

is the  $S$ -matrix element. On this idea, Ojima and Englert have developed the formulation on the  $S$ -matrix in the context of the algebraic quantum field theory (123) and the backaction of the Hawking radiation (55), respectively.

## 8. Weak measurement with environment

Let us consider a target system coupled with an environment and a general weak measurement for the compound of the target system and the environment. We assume that

there is no interaction between the probe and the environment and the same interaction between the target and probe systems (38). The Hamiltonian for the target system and the environment is given by

$$H = H_0 \otimes \mathbf{1}_e + H_1, \quad (100)$$

where  $H_0$  acts on the target system  $\mathcal{H}_s$  and the identity operator  $\mathbf{1}_e$  is for the environment  $\mathcal{H}_e$ , while  $H_1$  acts on  $\mathcal{H}_s \otimes \mathcal{H}_e$ . The evolution operators  $U := U(t, t_i)$  and  $V := U(t_f, t)$  as defined in Eq. (97) can be expressed by

$$U = U_0 K(t_0, t_i), \quad V = K(t_f, t_0) V_0, \quad (101)$$

where  $U_0$  and  $V_0$  are the evolution operators forward in time and backward in time, respectively, by the target Hamiltonian  $H_0$ .  $K$ 's are the evolution operators in the interaction picture,

$$K(t_0, t_i) = \mathcal{T} e^{-i \int_{t_i}^{t_0} dt U_0^\dagger H_1 U_0}, \quad K(t_f, t_0) = \overline{\mathcal{T}} e^{-i \int_{t_0}^{t_f} dt V_0 H_1 V_0^\dagger}, \quad (102)$$

where  $\mathcal{T}$  and  $\overline{\mathcal{T}}$  stand for the time-ordering and anti time-ordering products.

Let the initial and final environmental states be  $|e_i\rangle$  and  $|e_f\rangle$ , respectively. The probe state now becomes

$$|\zeta'\rangle = \langle f | \langle e_f | V U | e_i \rangle | i \rangle \left( \mathbf{1} - g \frac{\langle f | \langle e_f | V A U | e_i \rangle | i \rangle}{\langle f | \langle e_f | V U | e_i \rangle | i \rangle} \hat{P} + O(g^2) \right) |\zeta\rangle. \quad (103)$$

Plugging the expressions for  $U$  and  $V$  into the above, we obtain the probe state as

$$|\zeta'\rangle = N \zeta \left( \mathbf{1} - g \frac{\langle f | \langle e_f | K(t_f, t_0) V_0 A U_0 K(t_0, t_i) | e_i \rangle | i \rangle}{N} \hat{P} \right) |\zeta\rangle + O(g^2), \quad (104)$$

where  $N = \langle f | \langle e_f | K(t_f, t_0) V_0 U_0 K(t_0, t_i) | e_i \rangle | i \rangle$  is the normalization factor. We define the dual quantum operation as

$$\mathcal{E}^*(A) := \langle e_f | K(t_f, t_0) V_0 A U_0 K(t_0, t_i) | e_i \rangle = \sum_m V_0 F_m^\dagger A E_m U_0, \quad (105)$$

where

$$F_m^\dagger := V_0^\dagger \langle e_f | K(t_f, t_0) | e_m \rangle V_0, \quad (106)$$

$$E_m := U_0 \langle e_m | K(t_0, t_i) | e_i \rangle U_0^\dagger \quad (107)$$

are the Kraus operators. Here, we have inserted the completeness relation  $\sum_m |e_m\rangle \langle e_m| = \mathbf{1}$  with  $|e_m\rangle$  being not necessarily orthogonal. The basis  $|e_i\rangle$  and  $|e_f\rangle$  are the initial and final environmental states, respectively. Thus, we obtain the wave function of the probe as

$$\begin{aligned} |\zeta'\rangle &= N \left( \mathbf{1} - g \frac{\langle f | \mathcal{E}^*(A) | i \rangle}{N} \hat{P} \right) |\zeta\rangle + O(g^2) = N \left( \mathbf{1} - g \frac{\sum_m \langle f | V_0 F_m^\dagger A E_m U_0 | i \rangle}{\sum_m \langle f | V_0 F_m^\dagger E_m U_0 | i \rangle} \hat{P} \right) |\zeta\rangle + O(g^2) \\ &= N \left( \mathbf{1} - g \frac{\text{Tr} [A \sum_m E_m U_0 | i \rangle \langle f | V_0 F_m^\dagger]}{\text{Tr} [\sum_m E_m U_0 | i \rangle \langle f | V_0 F_m^\dagger]} \hat{P} \right) |\zeta\rangle + O(g^2) \\ &= N \left( \mathbf{1} - g \frac{\text{Tr} [\mathcal{E}(W) A]}{\text{Tr} [\mathcal{E}(W)]} \hat{P} \right) |\zeta\rangle + O(g^2) = N (\mathbf{1} - g \langle A \rangle_{\mathcal{E}(W)} \hat{P}) |\zeta\rangle + O(g^2), \end{aligned} \quad (108)$$

Analogous to Theorem 4.1, the shift of the expectation value of the position operator on the probe is

$$\Delta[Q] = g \cdot \text{Re}[\langle A \rangle_{\mathcal{E}(W)}] + mg \cdot \text{Im}[\langle A \rangle_{\mathcal{E}(W)}] \left. \frac{d \text{Var}[Q]}{dt} \right|_{t=t_0}. \quad (109)$$

From an analogous discussion, we obtain the shift of the expectation value of the momentum operator on the probe as

$$\Delta[P] = 2g \cdot \text{Var}[P] \cdot \text{Im}[\langle A \rangle_{\mathcal{E}(W)}]. \quad (110)$$

Thus, we have shown that the probe shift in the weak measurement is exactly given by the weak value defined by the quantum operation of the  $W$  operator due to the environment.

## 9. Summary

We have reviewed that the weak value is defined independent of the weak measurement in the original idea (4) and have explained its properties. Furthermore, to extract the weak value, we have constructed some measurement model to extract the weak value. I hope that the weak value becomes the fundamental quantity to describe quantum mechanics and quantum field theory and has practical advantage in the quantum-mechanical world.

## 10. Acknowledgment

The author acknowledges useful collaborations and discussion with Akio Hosoya, Yuki Susa, and Shu Tanaka. The author thanks Yakir Aharonov, Richard Jozsa, Sandu Popescu, Aephraim Steinberg, and Jeff Tollaksen for useful discussion. The author would like to thank the use of the utilities of Tokyo Institute of Technology and Massachusetts Institute of Technology and many technical and secretary supports. The author is grateful to the financial supports from JSPS Research Fellowships for Young Scientists (No. 21008624), JSPS Excellent Young Researcher Overseas Visit Program, Global Center of Excellence Program "Nanoscience and Quantum Physics" at Tokyo Institute of Technology during his Ph.D study.

## 11. References

- [1] J. Åberg and G. Mitchison, *J. Math. Phys.* 50, 042103 (2009).
- [2] G. S. Agarwal and P. K. Pathak, *Phys. Rev. A* 75, 032108 (2007).
- [3] Y. Aharonov, D. Z. Albert, A. Casher, and L. Vaidman, *Phys. Lett. A* 124, 199 (1987).
- [4] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. Lett.* 60, 1351 (1988).
- [5] Y. Aharonov, J. Anandan, and L. Vaidman, *Phys. Rev. A* 47, 4616 (1993).
- [6] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, *Phys. Rev.* 134, B1410 (1964).
- [7] Y. Aharonov and A. Botero, *Phys. Rev. A* 72, 052111 (2005).
- [8] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, *Phys. Lett. A* 301, 130 (2002).
- [9] Y. Aharonov and E. Y. Gruss, arXiv:quant-ph/0507269.
- [10] Y. Aharonov, S. Massar, S. Popescu, J. Tollaksen, and L. Vaidman, *Phys. Rev. Lett.* 77, 983 (1996).
- [11] Y. Aharonov, S. Popescu, D. Rohrlich, and L. Vaidman, *Phys. Rev. A* 48, 4084 (1993).
- [12] Y. Aharonov, S. Popescu, and J. Tollaksen, *Physics Today* 63, 27 (2010).
- [13] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, *Phys. Rev. A* 79, 052110 (2009).
- [14] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes* (Wiley-VCH, Weinheim, 2005).

- [15] Y. Aharonov and J. Tollaksen, in *Visions of Discovery: New Light on Physics Cosmology and Consciousness*, edited by R. Y. Chiao, M. L. Cohen, A. J. Leggett, W. D. Phillips, and C. L. Harper, Jr. (Cambridge University Press, Cambridge, 2011), p. 105.
- [16] Y. Aharonov and L. Vaidman, *Phys. Rev. A* 41, 11 (1990).
- [17] Y. Aharonov and L. Vaidman, *J. Phys. A* 24, 2315 (1991).
- [18] Y. Aharonov and L. Vaidman, *Phys. Lett. A* 178, 38 (1993).
- [19] Y. Aharonov and L. Vaidman, *Ann. N.Y. Acad. Sci.* 755, 361 (1995).
- [20] Y. Aharonov and L. Vaidman, in *Time in Quantum Mechanics*, Vol. 1, edited by J. G. Muga, R. Sala Mayato, and I. L. Egusquiza (Springer, Berlin Heidelberg, 2008) p. 399.
- [21] S. E. Ahnert and M. C. Payne, *Phys. Rev. A* 69, 042103 (2004).
- [22] J. Anandan, *Found. Phys. Lett.* 15, 415 (2002).
- [23] S. Ashhab and F. Nori, arXiv:0907.4823.
- [24] A. Bednorz and W. Belzig, *Phys. Rev. Lett.* 105, 106803 (2010).
- [25] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, 2006).
- [26] M. V. Berry and P. Shukla, *J. Phys. A* 43, 354024 (2010).
- [27] M. V. Berry, M. R. Dennis, B. McRoberts, and P. Shukla, *J. Phys. A* 44, 205301 (2011).
- [28] M. V. Berry, N. Brunner, S. Popescu, and P. Shukla, *J. Phys. A* 44, 205301 (2011).
- [29] A. Botero and B. Reznik, *Phys. Rev. A* 61, 050301 (2000)(R).
- [30] A. Botero, "Entanglement, Postselection and Precise Inferences in Joint Measurements of Incompatible Observables", in International Conference on Quantum Information, (Optical Society of America, 2008), paper JMB68.
- [31] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, 1992).
- [32] A. Brodtsch, *Master Thesis* at Tel-Aviv University (2008), arXiv:0811.1706.
- [33] A. Brodtsch and L. Vaidman, *J. Phys.: Conf. Ser.* 174, 012004 (2009).
- [34] R. Brout, S. Massar, R. Parentani, and Ph. Spindel, *Phys. Rep.* 260, 329 (1995).
- [35] T. Brun, L. Diósi, and W. T. Strunz, *Phys. Rev. A* 77, 032101 (2008).
- [36] N. Brunner, A. Acín, D. Collins, N. Gisin, and V. Scarani, *Phys. Rev. Lett.* 91, 180402 (2003).
- [37] N. Brunner, V. Scarani, M. Wegmüller, M. Legré, and N. Gisin, *Phys. Rev. Lett.* 93, 203902 (2004).
- [38] N. Brunner and C. Simon, *Phys. Rev. Lett.* 105, 010405 (2010).
- [39] P. Busch, *Found. Phys.* 20, 33 (1990).
- [40] P. Busch, P. Mittelstaedt, and P. J. Lahti, *Quantum Theory of Measurement* (Springer-Verlag, Berlin, 1991).
- [41] R. M. Camacho, P. B. Dixon, R. T. Glasser, A. N. Jordan, and J. C. Howell, *Phys. Rev. Lett.* 102, 013902 (2009).
- [42] D. Campo and R. Parentani, *Phys. Rev. D* 70, 105020 (2004).
- [43] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Phys. Rev. A* 80, 022339 (2009).
- [44] Y.-W. Cho, H.-T. Lim, Y.-S. Ra, and Y.-H. Kim, *New. J. Phys.* 12, 023036 (2010).
- [45] E. B. Davies and J. T. Lewis, *Commun. Math. Phys.* 17, 239 (1970).
- [46] P. C. W. Davies, *Phys. Rev. A* 79, 032103 (2009).
- [47] P. A. M. Dirac, *Proc. Roy. Soc. Lond. A* 180, 1 (1942).
- [48] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, *Phys. Rev. Lett.* 102, 173601 (2009).
- [49] J. Dressel, S. Agarwal, and A. N. Jordan, *Phys. Rev. Lett.* 104, 240401 (2010).
- [50] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan, *Phys. Rev. Lett.* 106, 040402 (2011).

- [51] J. Dressel and A. N. Jordan, arXiv:1110.0418.
- [52] I. M. Duck, P. M. Stevenson, E. C. G. Sudarshan, Phys. Rev. D 40, 2112 (1989).
- [53] G. F. R. Ellis and T. Rothman, Int. J. Theor. Phys. 49, 998 (2010).
- [54] F. Englert, arXiv:gr-qc/9502039.
- [55] F. Englert and Ph. Spindel, JHEP 12, 065 (2010).
- [56] A. Feizpour, X. Xing, and A. Steinberg, Phys. Rev. Lett. 107, 133603 (2011).
- [57] R. P. Feynman, Int. J. Theor. Phys. 21, 467 (1982).
- [58] J. L. Garretson, H. M. Wiseman, D. T. Pope, and D. T. Pegg, J. Opt. B 6, S506 (2004).
- [59] T. Geszti, Phys. Rev. A 81, 044102 (2010).
- [60] R. Golub and R. Gähler, Phys. Lett. A 136, 178 (1989).
- [61] M. A. de Gosson and S. M. de Gosson, arXiv:1109.3665.
- [62] E. Y. Gruss, *A Suggestion for a Teleological Interpretation of Quantum Mechanics*, as Master Thesis at The Hebrew University of Jerusalem (2000), arXiv:quant-ph/0006070.
- [63] E. Haapasalo, P. Lahti, and J. Schultz, Phys. Rev. A 84, 052107 (2011).
- [64] L. Hardy, Phys. Rev. Lett. 68, 2981 (1992).
- [65] H. F. Hofmann, Phys. Rev. A 62, 022103 (2000).
- [66] H. F. Hofmann, J. Phys. A 42, 275304 (2009).
- [67] H. F. Hofmann, Phys. Rev. A 81, 012103 (2010).
- [68] H. F. Hofmann, Phys. Rev. A 83, 022106 (2011).
- [69] H. F. Hofmann, New J. Phys. 13, 103009 (2011).
- [70] H. F. Hofmann, AIP Conf. Proc. 1363, 125 (2011).
- [71] H. F. Hofmann, arXiv:0911.0071.
- [72] H. F. Hofmann, arXiv:1005.0654.
- [73] H. F. Hofmann, M. E. Goggin, M. P. Almeida, and M. Barbieri, arXiv:1107.4735.
- [74] A. Hosoya and M. Koga, J. Phys. A 44, 415303 (2011).
- [75] A. Hosoya and Y. Shikano, J. Phys. A 43, 385307 (2010).
- [76] O. Hosten and P. Kwiat, Science 319, 787 (2008).
- [77] J. C. Howell, D. J. Starling, P. B. Dixon, P. K. Vudyasetu, and A. N. Jordan, Phys. Rev. A 81, 033813 (2010).
- [78] W. Hu, Found. Phys. 20, 447 (1990).
- [79] M. Iinuma, Y. Suzuki, G. Taguchi, Y. Kadoya, and H. F. Hofmann, New J. Phys. 13, 033041 (2011).
- [80] L. M. Johansen, Phys. Rev. Lett. 93, 120402 (2004).
- [81] L. M. Johansen, Phys. Lett. A 322, 298 (2004).
- [82] L. M. Johansen, Phys. Lett. A 366, 374 (2007).
- [83] A. N. Jordan and A. N. Korotkov, Contemp. Phys. 51, 125 (2010).
- [84] A. N. Jordan, B. Trauzettel, and G. Burkard, Phys. Rev. B 76, 155324 (2007).
- [85] R. Jozsa, Phys. Rev. A 76, 044103 (2007).
- [86] S. Kagami, Y. Shikano, and K. Asahi, Physica E 43, 761 (2011).
- [87] R. E. Kastner, Stud. Hist. Philos. M. P. 35, 57 (2004).
- [88] J. M. Knight and L. Vaidman, Phys. Lett. A 143, 357 (1990).
- [89] H. Kobayashi, S. Tamate, T. Nakanishi, K. Sugiyama, and M. Kitano, Phys. Rev. A 81, 012104 (2010).
- [90] H. Kobayashi, S. Tamate, T. Nakanishi, K. Sugiyama, and M. Kitano, J. Phys. Soc. Jpn. 80, 034401 (2011).
- [91] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg, Science 332, 1170 (2011).
- [92] A. G. Kofman, S. Ashhab, and F. Nori, arXiv:1109.6315.
- [93] T. Koike and S. Tanaka, Phys. Rev. A 84, 062106 (2011).

- [94] A. N. Korotkov and A. N. Jordan, *Phys. Rev. Lett.* 97, 166805 (2006).
- [95] K. Kraus, *Ann. Phys.* 64, 311 (1971).
- [96] C. M. Krowne, *Phys. Lett. A* 373, 466 (2009).
- [97] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve, and A. Korotkov, *Nat. Phys.* 6, 442 (2010).
- [98] C. R. Leavens, *Found. Phys.* 35, 469 (2005).
- [99] C. R. Leavens, I. P. Gimenez, D. Alonso, and R. Sala Mayato, *Phys. Lett. A* 359, 416 (2006).
- [100] A. J. Leggett and A. Garg, *Phys. Rev. Lett.* 54, 857 (1985).
- [101] C.-F. Li, X.-Y. Xu, J.-S. Tang, J.-S. Xu, and G.-C. Guo, *Phys. Rev. A* 83, 044102 (2011).
- [102] A. C. Lobo and C. A. Ribeiro, *Phys. Rev. A* 80, 012112 (2009).
- [103] A. Di Lorenzo and J. C. Egues, *Phys. Rev. A* 77, 042108 (2008).
- [104] A. Di Lorenzo, arXiv:1109.4183.
- [105] G. Lüders, *Ann. Physik* 8, 322 (1951).
- [106] A. P. Lund and H. M. Wiseman, *New J. Phys.* 12, 093011 (2010).
- [107] J. S. Lundeen and A. M. Steinberg, *Phys. Rev. Lett.* 102, 020404 (2009).
- [108] J. S. Lundeen and C. Bamber, arXiv:1110.0727.
- [109] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, *Nature* 474, 188 (2011).
- [110] S. Marcovitch and B. Reznik, arXiv:1103.2557.
- [111] S. Massar and S. Popescu, *Phys. Rev. A* 84, 052106 (2011).
- [112] D. Menzies and S. Croke, arXiv:0903.4181.
- [113] D. Menzies and N. Korolkova, *Phys. Rev. A* 76, 062310 (2007).
- [114] D. Menzies and N. Korolkova, *Phys. Rev. A* 77, 062105 (2008).
- [115] D. J. Miller, in *Wagga 2009: 33rd Annual Condensed Matter and Material Meeting*, paper WP10.
- [116] R. Mir, J. S. Lundeen, H. M. Mitchell, A. M. Steinberg, J. L. Garretson, and H. M. Wiseman, *New J. Phys.* 9, 287 (2007).
- [117] G. Mitchison, *Phys. Rev. A* 77, 052102 (2008).
- [118] G. Mitchison, R. Jozsa, and S. Popescu, *Phys. Rev. A* 76, 062105 (2007).
- [119] K. Nakamura, *J. Phys. Soc. Jpn.* 79, 125003 (2010).
- [120] K. Nakamura, A. Nishizawa, M.-K. Fujimoto, arXiv:1108.2114.
- [121] J. von Neumann, *Mathematische Grundlagen der Quantumechanik* (Springer, Berlin, 1932), [ Eng. trans. by R. T. Beyer, *Mathematical foundations of quantum mechanics* (Princeton University Press, Princeton, 1955). ]
- [122] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [123] I. Ojima, in *Quantum Bio-Informatics III, From Quantum Information and Bio-Informatics*, edited by L. Accardi, W. Freudenberg, and M. Ohya (World Scientific, Singapore, 2010) p. 277.
- [124] M. Onoda, S. Murakami, and N. Nagaosa, *Phys. Rev. Lett.* 93, 083901 (2004).
- [125] OPERA Collaboration, T. Adam *et al.*, arXiv:1109.4897.
- [126] O. Oreshukov and T. A. Brun, *Phys. Rev. Lett.* 95, 110409 (2005).
- [127] M. Ozawa, *J. Math. Phys.* 25, 79 (1984).
- [128] M. Ozawa, in *Squeezed and Nonclassical Light*, edited by P. Tombesi and E. R. Pike, (Plenum, New York, 1989), p. 263.
- [129] A. K. Pan and A. Matzkin, arXiv:1109.6824.
- [130] A. D. Parks, *J. Phys. A* 43, 035305 (2010).
- [131] A. D. Parks, D. W. Cullin, and D. C. Stoudt, *Proc. R. Soc. Lond. A* 454, 2997 (1998).

- [132] A. D. Parks and J. E. Gray, *Phys. Rev. A* **84**, 012116 (2011).
- [133] G. J. Pryde, J. L. O’Brien, A. G. White, T. C. Ralph, and H. M. Wiseman, *Phys. Rev. Lett.* **94**, 220405 (2005).
- [134] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O’Brien, A. G. White, and G. J. Pryde, *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1256 (2011).
- [135] A. Ranfagni, P. Fabeni, G. P. Pazzi, and D. Mugnai, *Phys. Rev. E* **48**, 1453 (1993).
- [136] A. Ranfagni, D. Mugnai, and R. Ruggeri, *J. Appl. Phys.* **95**, 2217 (2004).
- [137] T. Ravon and L. Vaidman, *J. Phys. A* **40**, 2873 (2007).
- [138] M. Rédei and S. J. Summers, *Stud. Hist. Philos. Sci. B* **38**, 390 (2007).
- [139] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, *Phys. Lett. A* **324**, 125 (2004).
- [140] B. Reznik and Y. Aharonov, *Phys. Rev. A* **52**, 2538 (1995).
- [141] N. W. M. Ritchie, J. G. Story, and R. G. Hulet, *Phys. Rev. Lett.* **66**, 1107 (1991).
- [142] D. Rohrlich and Y. Aharonov, *Phys. Rev. A* **66**, 042102 (2002).
- [143] A. Romito and Y. Gefen, *Physica E* **42**, 343 (2010).
- [144] A. Romito, Y. Gefen, and Y. M. Blanter, *Phys. Rev. Lett.* **100**, 056801 (2008).
- [145] C. Rovelli, *Phys. Rev. A* **50**, 2788 (1994).
- [146] J. Ruseckas and B. Kaulakys, *Phys. Rev. A* **66**, 052106 (2002).
- [147] J. Ruseckas and B. Kaulakys, *Lith. J. Phys.* **44**, 161 (2004).
- [148] T. Sagawa, arXiv:0901.4212.
- [149] L. K. Shalm, S. Kocsis, S. Ravets, B. Braverman, M. J. Stevens, R. P. Mirin, and A. M. Steinberg, “Observation of Bohmian Trajectories of a Single Photon Using Weak Measurements,” in *Quantum Electronics and Laser Science Conference, OSA Technical Digest (CD)* (Optical Society of America, 2010), paper QThK7.
- [150] A. Shapere and F. Wilczek, *Geometric Phases in Physics* (World Scientific, Singapore, 1988).
- [151] Y. Shikano, *Weak Values in Quantum Measurement Theory — Concepts and Applications —*, as Master Thesis at Tokyo Institute of Technology (2009).
- [152] Y. Shikano, *Time in Weak Values and Discrete Time Quantum Walk* as Ph. D. Dissertation at Tokyo Institute of Technology (2011).
- [153] Y. Shikano, unpublished (2010).
- [154] Y. Shikano and A. Hosoya, *J. Math. Phys.* **49**, 052104 (2008).
- [155] Y. Shikano and A. Hosoya, *J. Phys. A* **42**, 025304 (2010).
- [156] Y. Shikano, S. Kagami, S. Tanaka, and A. Hosoya, *AIP Conf. Proc.* **1363**, 177 (2011).
- [157] Y. Shikano and S. Tanaka, *Europhys. Lett.* **96**, 40002 (2011).
- [158] V. Shpitalnik, Y. Gefen, and A. Romito, *Phys. Rev. Lett.* **101**, 226802 (2008).
- [159] C. Simon and E. S. Polzik, *Phys. Rev. A* **83**, 040101 (2011)(R).
- [160] E. Sjöqvist, *Phys. Lett. A* **359**, 187 (2006).
- [161] E. Sjöqvist and J. Åberg, *Phys. Lett. A* **354**, 396 (2006).
- [162] D. Sokolovski, A. Z. Msezane, and V. R. Shaginyan, *Phys. Rev. A* **71**, 064103 (2005).
- [163] D. Sokolovski, *Phys. Rev. A* **76**, 042125 (2007).
- [164] D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann, *Phys. Rev. Lett.* **92**, 043601 (2004).
- [165] S. E. Spence and A. D. Parks, to appear in *Found. Phys.*, arXiv:1010.3289.
- [166] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, *Phys. Rev. A* **80**, 041803 (2009).
- [167] D. J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, *Phys. Rev. A* **82**, 063822 (2010).
- [168] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, *Phys. Rev. A* **82**, 011802 (2010).

- [169] D. J. Starling, P. B. Dixon, N. S. Williams, A. N. Jordan, and J. C. Howell, "Near Quantum Limited Optical Phase Measurements on a Dark Fringe," in *Frontiers in Optics, OSA Technical Digest (CD) (Optical Society of America, 2010)*, paper FThD4.
- [170] A. M. Steinberg, *Phys. Rev. Lett.* 74, 2405 (1995).
- [171] A. M. Steinberg, *Phys. Rev. A* 52, 32 (1995).
- [172] A. M. Steinberg, private communication (2011).
- [173] M. Suzuki, *Prog. Theor. Phys.* 56, 1454 (1976).
- [174] S. Tamate, H. Kobayashi, T. Nakanishi, K. Sugiyama, and M. Kitano, *New J. Phys.* 11, 093025 (2009).
- [175] A. Tanaka, *Phys. Lett. A* 297 307 (2002).
- [176] S. Tanimura, arXiv:1110.1790.
- [177] H. Tomita, arXiv:1109.4302.
- [178] J. Tollaksen, *J. Phys.: Conf. Ser.* 70, 012014 (2007).
- [179] J. Tollaksen, *J. Phys.: Conf. Ser.* 70, 012015 (2007).
- [180] J. Tollaksen, *J. Phys.: Conf. Ser.* 70, 012016 (2007).
- [181] J. Tollaksen, Y. Aharonov, A. Casher, T. Kaufherr, and S. Nussinov, *New J. Phys.* 12, 013023 (2010).
- [182] H. F. Trotter, *Proc. Amer. Math. Soc.* 10, 545 (1959).
- [183] M. Tsang, *Phys. Rev. A* 81, 013824 (2010).
- [184] M. D. Turner, C. A. Hagedorn, S. Schlamminger, and J. H. Gundlach, *Opt. Lett.* 36, 1479 (2011).
- [185] A. Uhlmann, *Rep. Math. Phys.* 24, 229 (1986).
- [186] W. G. Unruh, *Phys. Rev. A* 50, 882 (1994).
- [187] K. Urbanowski, *Cent. Eur. J. Phys.* 7, 696 (2009).
- [188] L. Vaidman, *Found. Phys.* 26, 895 (1996).
- [189] L. Vaidman, arXiv:quant-ph/9607023.
- [190] M. S. Wang, *Phys. Rev. Lett.* 79, 3319 (1997).
- [191] N. S. Williams and A. N. Jordan, *Phys. Rev. Lett.* 100, 026804 (2008).
- [192] H. M. Wiseman, *Phys. Rev. A* 65, 032111 (2002).
- [193] H. M. Wiseman, *New J. Phys.* 9, 165 (2007).
- [194] H. M. Wiseman and G. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2010).
- [195] S. Wu and K. Kølmer, *Phys. Lett. A* 374, 34 (2009).
- [196] S. Wu and Y. Li, *Phys. Rev. A* 83, 052106 (2011).
- [197] S. Wu and M. Życzkowski, arXiv:1106.4607.
- [198] K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, *New J. Phys.* 11, 033011 (2009).
- [199] X. Zhu, Y. Zhang, S. Pang, C. Qiao, Q. Liu, and S. Wu, *Phys. Rev. A* 84, 052111 (2011).
- [200] O. Zilberberg, A. Romito, and Y. Gefen, *Phys. Rev. Lett.* 106, 080405 (2011).

# Generalized Non-Relativistic Supersymmetric Quantum Mechanics

Thomas L. Markovich, Mason T. Biamonte, Eric R. Bittner  
and Donald J. Kouri  
*University of Houston*  
USA

## 1. Introduction

Symmetry has long been recognized as a powerful formal and computational tool in quantum mechanics, beginning with the seminal work of Wigner Wigner & Fano (1960) and Weyl Weyl (1950). Indeed, it is well understood that so-called “accidental degeneracies” were, in fact, not accidents at all but rather the result of “hidden symmetry” (e.g. the  $2l+1$  degeneracy of the hydrogen atom energy states). Because of this fundamental role in quantum mechanics, the discovery of new symmetries (and their possible “breaking” by interactions) is of enormous interest Griffiths (1987). In the latter half of the 20<sup>th</sup> century, a new hidden symmetry was discovered that led to much speculation in relativistic quantum field theory as applied to elementary particles. The essence of this so-called “supersymmetry” (SUSY) is that for every boson, there is also a fermion of the same mass (energy) and vice versa. Of course, this has not been observed in nature, leading to speculation that there exists some interaction in nature that “breaks” the symmetry. On the other hand, there is also substantial opinion held by many physicists that SUSY has no connection to physical reality.

This chapter is not aimed at addressing such issues. Rather, it was observed by many Junker (1996) that one did not have to deal with quantum field theory to encounter SUSY. Indeed, SUSY is also an intrinsic feature of ordinary, non-relativistic quantum mechanics (SUSY-QM). In this case, attention has not been focused on whether the symmetry exists in nature. Instead, it has been used primarily as a pedagogical tool. The reason for this lies in the intimate connection of SUSY-QM and the ladder operator approach to the harmonic oscillator, angular momentum, and the hydrogen atom Dirac (1958).

The essence of SUSY-QM is the factorization of the Hamiltonian for a one dimensional system in analogy with the harmonic oscillator. For most bound-state quantum systems, it is possible to define operators analogous to the lowering ( $\hat{a}$ ) and raising ( $\hat{a}^\dagger$ ) operators that factor the harmonic oscillator Hamiltonian. However, in the general case, these operators do not possess all the properties of the harmonic oscillator  $\hat{a}$  and  $\hat{a}^\dagger$ , but rather they behave as so-called “charge operators”. As such, the SUSY charge operators not only allow factorization of the one dimensional Hamiltonian, they form a Lie algebraic structure. This structure results in the generation of isospectral “sector Hamiltonians”. Unfortunately, almost all previous research concentrated on exactly soluble, one dimensional model systems. We became interested in the possibility of taking computational advantage of SUSY. Our idea was that symmetry in QM has long been known to lead to significant computational simplifications and advantages. We

therefore asked whether this could be the case for SUSY. It turns out that SUSY does, in fact, lead to significant computational advantages Kouri et al. (2010a); Kouri, Markovich, Maxwell & Bittner (2009); Kouri, Markovich, Maxwell & Bodman (2009). In particular, the structure of the degeneracies between sector Hamiltonians makes it possible to achieve significant progress in more accurate calculations of excited state energies and wave functions. Below we outline how the theory can be used as a new computational tool, first for one dimension and later for higher dimensional systems. In addition, we also introduce an entirely new class of system dependent coherent states.

There have been a number of suggested generalizations of SUSY-QM to treat more than one dimensional systems Andrianov et al. (1985); Andrianov, Borisov & Ioffe (1984a;b;c); Andrianov et al. (1986); Andrianov, Borisov, Ioffe & Eides (1984); Andrianov & Ioffe (1988); Andrianov et al. (2002); Cannata et al. (2002); Das & Pernice (1996). For the most part, these have involved the introduction of new “spin-like” variables. One early study instead introduced tensorial operators Stedman (1985), but at the cost of seriously affecting the nature of the energy level degeneracies. In addition, in the tensorial operator approach, he did not consider the application to any system in detail other than writing down the equations for the hydrogen atom without exploring their solutions. In the following sections, we present our generalization of SUSY-QM to allow the treatment of multi-particle, multi-dimensional systems. These include clusters of distinguishable particles and the electronic structure of atoms.

## 2. Introduction to supersymmetric quantum mechanics in one dimension

The general starting point is to define the so-called “superpotential”, usually denoted as  $W$ . In the theory,  $W$  is related to the ground state wave function through the well-known Riccati substitution Jafarpour & Afshar (2002):

$$\psi_0^1(x) = N e^{-\int_0^x W(x') dx'} \quad (2.1)$$

The relationship between the superpotential  $W$  and the physical interaction  $V(x)$  results from assuming that Equation (2.1) solves the standard Schrödinger equation with energy zero. This does not impose any restriction since the energy can be changed by adding any constant to the Hamiltonian. Thus,

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi_0^1}{dx^2} + V_1 \psi_0^1 = 0 \quad (2.2)$$

If we solve for  $W_1$  in Equation (2.1), we find that

$$W_1 = -\frac{d\psi_0^1}{dx} \frac{1}{\psi_0^1} = -\frac{d}{dx} \ln \psi_0^1 \quad (2.3)$$

and, if  $W_1$  is known,  $V_1$  is given by

$$V_1(x) = \frac{\hbar^2}{2m} \left( W_1^2(x) - \frac{dW_1}{dx} \right) \quad (2.4)$$

It is then evident that

$$-\frac{d^2 \psi_0^1}{dx^2} + \left( W_1^2(x) - \frac{dW_1}{dx} \right) \psi_0^1 = 0 \quad (2.5)$$

The Hamiltonian operator now can be factored in the form

$$-\frac{d^2}{dx^2} + W_1^2(x) - \frac{dW_1}{dx} = \left[ -\frac{d}{dx} + W_1(x) \right] \left[ \frac{d}{dx} + W_1(x) \right] \quad (2.6)$$

We define the “charge” operator and its adjoint by (assuming  $W_1$  is hermitian; *i.e.*,  $\psi_0^1$  is real)

$$Q_1 = \frac{d}{dx} + W_1, \quad Q_1^\dagger = -\frac{d}{dx} + W_1 \quad (2.7)$$

Then the “first sector” Hamiltonian is defined as

$$\mathcal{H}_1 = Q_1^\dagger Q_1 \quad (2.8)$$

Then it follows that for  $n > 0$ , (since for  $n = 0$ ,  $E_0 = 0$ ),

$$Q_1^\dagger Q_1 \psi_n^1 = E_n^1 \psi_n^1 \quad (2.9)$$

We then apply  $Q_1$  to the equation, to obtain

$$Q_1 Q_1^\dagger (Q_1 \psi_n^1) = E_n^1 Q_1 \psi_n^1 \quad (2.10)$$

Thus,  $Q_1 \psi_n^1$  is an eigenstate of  $\mathcal{H}_2$  with the same energy,  $E_n^1$ , as the state  $\psi_n^1$ . Similarly, consider the eigenstates of  $\mathcal{H}_2$ :

$$\mathcal{H}_2 \psi_n^2 = Q_1 Q_1^\dagger \psi_n^2 = E_n^2 \psi_n^2. \quad (2.11)$$

Application of  $Q_1^\dagger$ , then implies that  $Q_1^\dagger \psi_n^2$  is an eigenstate of  $\mathcal{H}_1$ :

$$(Q_1^\dagger Q_1)(Q_1^\dagger \psi_n^2) = E_n^2 Q_1^\dagger \psi_n^2 \quad (2.12)$$

It follows that the Hamiltonians  $\mathcal{H}_1$  and  $\mathcal{H}_2$  have identical spectra with the exception of the ground state, since the  $E_0^1 = 0$  wave function is unique. In the case of the ground state  $\psi_0^1$ , we recall that

$$Q_1 \psi_0^1 = 0 \quad (2.13)$$

which shows that the quantity  $Q_1 \psi_0^1$  cannot be used to generate the ground state of the second sector. Indeed, Equation (2.13) indicates that such a  $\psi_0^2$  would vanish identically.

Because of the uniqueness of the  $E_0^1 = 0$  state, the indexing of the first and second sector levels must be modified. Consider

$$Q_1 Q_1^\dagger \psi_n^2 = E_n^2 \psi_n^2 \quad (2.14)$$

Then

$$Q_1^\dagger Q_1 (Q_1^\dagger \psi_{n+1}^2) = E_{n+1}^2 (Q_1^\dagger \psi_{n+1}^2) \quad (2.15)$$

since  $Q_1 \psi_0^1 \equiv 0$ . So

$$E_n^2 = E_{n+1}^1 \quad (2.16)$$

and we conclude that

$$\psi_n^2 = \frac{Q_1 \psi_{n+1}^1}{\sqrt{E_{n+1}^1}} \quad \text{and} \quad \psi_{n+1}^1 = \frac{Q_1^\dagger \psi_n^2}{\sqrt{E_n^2}} \quad (2.17)$$

The next step in building a hierarchy of isospectral Hamiltonians is to define a second superpotential,  $W_2$ , according to

$$W_2 = -\frac{d}{dx} \ln \psi_0^2 \quad (2.18)$$

in much the same way as we did before. It is then clear that we can define an alternate form for  $\mathcal{H}_2$ , given by

$$\mathcal{H}_2 = Q_2^\dagger Q_2 + E_0^2, \quad (2.19)$$

where

$$Q_2 = \frac{d}{dx} + W_2 \quad (2.20)$$

We observe that  $\psi_0^2$  is automatically an eigenstate of this form for  $\mathcal{H}_2$

$$\mathcal{H}_2 \psi_0^2 = E_0^2 \psi_0^2 \quad (2.21)$$

Next consider the first excited state eigenvalue equation for the second sector:

$$\mathcal{H}_2 \psi_1^2 = E_1^2 \psi_1^2 \quad (2.22)$$

We apply  $Q_2$  to Equation (2.22) to find

$$\left( Q_2 Q_2^\dagger + E_0^2 \right) Q_2 \psi_1^2 = E_n^2 Q_2 \psi_1^2 \quad (2.23)$$

Then, by similar reasoning, we deduce that

$$Q_2 \psi_1^2 = \sqrt{E_1^2 - E_0^2} \psi_0^3. \quad (2.24)$$

Using the new charge operators  $Q_2$  and  $Q_2^\dagger$ , we then define the third sector Hamiltonian,

$$\mathcal{H}_3 = Q_2 Q_2^\dagger + E_0^2, \quad (2.25)$$

with ground state equation

$$\mathcal{H}_3 \psi_0^3 = E_0^3 \psi_0^3 \quad (2.26)$$

It follows that

$$Q_2^\dagger \psi_0^3 = \sqrt{E_1^2 - E_0^2} \psi_1^2 \quad (2.27)$$

and,

$$Q_2 Q_2^\dagger \psi_0^3 = \left( E_1^2 - E_0^2 \right) \psi_0^3 = E_0^3 \psi_0^3 \quad (2.28)$$

Thus, we conclude that

$$E_0^3 = E_1^2 - E_0^2 \quad (2.29)$$

It is clear that this procedure can be continued until one exhausts the number of bound excited states supported by  $\mathcal{H}_1$ . We also see that determining the excited state energies and wave

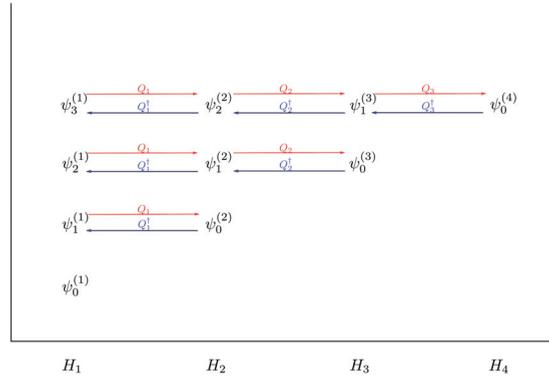


Fig. 1. Above, we display graphically how this hierarchical degeneracy is realized.

functions for  $\mathcal{H}_1$  involves solving for the ground state energies and wave functions for each sector Hamiltonian,  $\mathcal{H}_j, j > 1$ .

For a more concrete example, we present a general family of anharmonic oscillator. The ubiquity of this example in chemistry and physics is best exemplified by the fact that all nuclear vibrations in molecules are anharmonic, with the effect increasing as the vibrational energy gets closer to the dissociation limit Wilson et al. (1955). Additionally, anharmonicity is also present when studying the effects of rotation, through the centrifugal potential. For the one dimensional case, we consider an oscillator on the domain  $-\infty < x < \infty$ . In order to have potentials that are guaranteed to possess bound states, we shall postulate a superpotential

$$W(x) = \sum_{j=0}^J d_j x^{2j+1} \tag{2.30}$$

Then the corresponding sector one potential,  $V_1(x)$ , is

$$V_1(x) = \sum_{j=0}^J \sum_{j'=0}^J d_j d_{j'} x^{2(j+j'+1)} - \sum_{j=0}^J d_j (2j+1) x^{2j} \tag{2.31}$$

The “charge” operators are given by

$$Q_1 = \frac{d}{dx} + W_1 \text{ and } Q_1^\dagger = -\frac{d}{dx} + W_1. \tag{2.32}$$

Then the first sector ground state for a general member of this family is

$$\psi_0^+(x) = N e^{-\sum_{j=0}^J \frac{d_j x^{2j+2}}{(2j+2)}} \tag{2.33}$$

We stress that contrary to the periodic case Kouri, Markovich, Maxwell & Bodmann (2009), the solution of the sector two equation,

$$Q^\dagger \psi_0^2(x) = 0 \tag{2.34}$$

is not allowed because it is not normalizable. Thus, the ground state for the second sector satisfies

$$Q_1 Q_1^\dagger \psi_0^2(x) = E_0^2 \psi_0^2(x) = E_1^1 \psi_0^2(x) \quad (2.35)$$

where

$$E_1^1 \neq 0 \quad (2.36)$$

However, once  $\psi_0^2(x)$  is known, one can generate the first excited state  $\psi_1^1(x)$  according to

$$Q_1^\dagger \psi_0^2(x) = \sqrt{E_0^2} \psi_1^1(x) \quad (2.37)$$

The energy,  $E_1^1$ , of  $\psi_1^1(x)$  is, of course, equal to  $E_0^2$ .

We remark that the ground state,  $\psi_0^1(x)$  is equal to the product of the ground states for each separate term in  $W$ . This is to say that,

$$\psi_0^1(x) = N \prod_{j=0}^J e^{\frac{d_j x^{2j+2}}{(2j+2)}}, \quad (2.38)$$

where  $N$  is the normalization constant. This is true even though  $V_1(x)$  contains cross terms of the form

$$d_j d_{j'} x^{2(j+j'+1)}, \quad j \neq j' \quad (2.39)$$

In fact, even more general anharmonic oscillators can be dealt with. Thus, any function,  $g(x)$  can be added to  $W(x)$  in Equation (2.30), provided only that  $e^{-\int_0^x g(x') dx'}$  is  $\mathcal{L}_2$ . Thus, not only polynomial anharmonic potentials can be treated but many others.

### 3. Computational examples

In the following section we will explore the computational aspects of our SUSY-QM approach using two example anharmonic oscillator systems. To illustrate this approach to polynomial anharmonic oscillation we define  $W(x)$  to be

$$W_1(x) = x^3 + 2x, \quad (3.1)$$

which obviously yields a potential for the first sector of

$$V_1(x) = x^6 + 4x^4 + x^2 - 2. \quad (3.2)$$

Where  $x$  is defined on the domain  $-\infty < x < \infty$ . We can thus define  $\mathcal{H}_1$  as:

$$\mathcal{H}_1 = \left[ -\frac{d}{dx} + W_1(x) \right] \left[ \frac{d}{dx} + W_1(x) \right] \quad (3.3)$$

which satisfies the equation

$$\mathcal{H}_1 \psi(x)_0^{(1)} = 0, \quad (3.4)$$

and possesses an analytic ground state wave function of

$$\psi_0^{(1)} = N e^{-(\frac{x^4}{4} + x^2)}. \quad (3.5)$$

To get the second Hamiltonian in the hierarchy we next define  $\mathcal{H}_2$  as

$$\mathcal{H}_2 = \left[ \frac{d}{dx} + W_1(x) \right] \left[ -\frac{d}{dx} + W_1(x) \right] \quad (3.6)$$

so

$$V_2 = x^6 + 4x^4 + 7x^2 + 2 \quad (3.7)$$

and we must solve the equation

$$\mathcal{H}_2 \psi_0^{(2)} = E_0^{(2)} \psi_0^{(2)}. \quad (3.8)$$

Possessing  $\psi_0^{(2)}$ , we may develop the next Hamiltonian in the hierarchy. To do so, we begin by expressing  $\mathcal{H}_2$  in the following form

$$\mathcal{H}_2 = \left[ -\frac{d}{dx} + W_2(x) \right] \left[ \frac{d}{dx} + W_2(x) \right] + E_0^{(2)} \quad (3.9)$$

where

$$W_2(x) = -\frac{d}{dx} \ln \psi_0^{(2)}. \quad (3.10)$$

with

$$\psi_0^{(3)} = \frac{Q_2 \psi_1^{(2)}}{\sqrt{E_1^{(2)} - E_0^{(2)}}}. \quad (3.11)$$

From this point, one can obviously generate as many Hamiltonians as needed. It should also be noted that the excited state wave functions can be obtained by using the charge operators we have previously defined. We now turn to the proof of principle for this approach as a computational scheme to obtain improved excited state energies and wave functions in the Rayleigh-Ritz variational method. We should note that these results can be generalized to any system where a hierarchy of hamiltonians can be generated because of the nature of the Rayleigh-Ritz scheme. In the standard approach one calculates the energies and wave functions variationally, relying on the Hylleraas-Undheim theorem for convergence Hylleraas & Undheim (1930). This, however, is unattractive for higher energy states because they require a much larger basis to converge to the same error. We stress that this is true regardless of the specific basis set used. Of course, some bases will be more efficient than others but it is generally true that for a given basis, the Rayleigh-Ritz result is less accurate for excited states. We address this situation by always solving for ground states in the variational part of the problem.

To demonstrate our computational scheme, we investigate the first example system from the previous section. For this potential Equation (3.2), exact solutions are known for all states of  $\mathcal{H}_1$ . We use the exact results to assess the accuracy of the variational calculations. For our first variational calculations, we use the harmonic oscillator basis functions where:

$$\phi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} \mathcal{H}_n(x) e^{-\frac{x^2}{2}}, \quad (3.12)$$

with each matrix element determined using

$$\langle \phi_{n'} | \mathcal{H}_i | \phi_n \rangle \quad (3.13)$$

Using the hierarchy of hamiltonians, we present the converged eigenvalues in Table 1. In Table 1, all energies were obtained for each of the Hamiltonians,  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , by standard variational calculations using basis set sizes to achieve an accuracy of  $10^{-6}$ . It is easily seen that the ground state of  $\mathcal{H}_2$  is degenerate with the first excited state of  $\mathcal{H}_1$ . More interesting is the behavior of the excited state wave functions. Using the Cauchy criterion to measure

convergence, we show the basis set size (N) needed in a standard variational approach to obtain various eigenstates to the accuracy shown in Table 2. Clearly, excited state wave functions require substantially larger basis sets to achieve a high degree of accuracy. In Table 3 we show the results obtained for the same excited state wave functions obtained by applying the charge operator to the ground state wave function for  $\mathcal{H}_2$ . Again, N denotes the basis set required, and  $\Delta N$  is the reduction of basis set achieved by use of the charge operators.

$n$	$N\psi_0^n$	$\mathcal{H}_1$	$N\psi_1^n$	$\mathcal{H}_2$	$\Delta N$
0	$44\psi_0^1$	6.9441187e-07	$34\psi_0^1$	5.024450	18
1	$52\psi_1^1$	5.024449	$42\psi_0^1$	11.696825	20
2	$62\psi_2^1$	11.696820	$50\psi_0^1$	19.497666	-

Table 1. Energies for the Anharmonic Polynomic Oscillator using Hierarchy of Hamiltonians.

$n$	$N\psi_0^n$	$\mathcal{L}_2$	$\mathcal{L}_\infty$
0	$56\psi_0^1$	5.835283e-07	1.110223-16
1	$78\psi_1^1$	1.975656e-07	4.019723e-16
0	$68\psi_0^2$	2.303928e-07	2.220446e-16

Table 2. Wave function errors for the Anharmonic Polynomic Oscillator using the standard variational method for each hierarchy Hamiltonian. Each value has six significant figures.

$n$	$\Delta N$	$N\psi_0^n$	$\mathcal{L}_2$	$\mathcal{L}_\infty$
1	4	$74\psi_1^1$	4.083823e-07	2.086041e-16

Table 3. Wave function errors for the Anharmonic Polynomic Oscillator using Charge Operators to find excited states. Each value has six significant figures.

To find the solutions we used LAPACK routines to find these eigenvalues and vectors and GSL routines for numerical integration. Clearly, the use of the hierarchy of hamiltonians and charge operators provides more rapid convergence, which provides us with better methods to calculate the excited states.

The second example results from taking

$$W_1(x) = x^3 + x + e^x. \quad (3.14)$$

In this case,

$$V_1(x) = x^6 + 2x^4 + 2x^3e^x + 2xe^x + x^2 + e^{2x} - 3x^2 - e^x - 1 \quad (3.15)$$

Then

$$\mathcal{H}_1 = \left[ -\frac{d}{dx} + x^3 + x + e^x \right] \left[ \frac{d}{dx} + x^3 + x + e^x \right] \quad (3.16)$$

with

$$\mathcal{H}_1 \psi_0^1 = 0 \quad (3.17)$$

and the analytical ground state is

$$\psi_0^{(1)} = N e^{-\left(\frac{x^4}{4} + \frac{x^2}{2} + e^x\right)}. \quad (3.18)$$

Then the second sector Hamiltonian is

$$\mathcal{H}_2 = \left[ \frac{d}{dx} + x^3 + x + e^x \right] \left[ -\frac{d}{dx} + x^3 + x + e^x \right] \quad (3.19)$$

The ground state satisfies

$$\mathcal{H}_2 \psi_0^2 = E_0^2 \psi_0^2 \quad (3.20)$$

which must be solved numerically.

We performed the Rayleigh-Ritz calculations and found similar results for the second system described by  $W = x^3 + x + e^x$ . Because the excited states of this oscillator are not known analytically, we use the Cauchy convergence criterion

$$\int_{-\infty}^{\infty} |N \psi_n - N_{-1} \psi_n|^2 dx, \quad (3.21)$$

where  $N$  is the basis size. In Table 4, we give the converged energy levels (to 5 significant figures) obtained by standard variational calculations applied to  $\mathcal{H}_1$  and  $\mathcal{H}_2$ . In Table 5, we show the basis set sizes needed in standard variational calculations to converge the wave functions for  $\mathcal{H}_1$  and  $\mathcal{H}_2$  (again, the Cauchy criterion of convergence was used.) Finally, in Table 6, we show the results for excited states obtained using the charge operators applied to the ground state wave functions of  $\mathcal{H}_2$ . Again,  $\Delta N$  shows the reduction in the basis size gained by the charge operator approach.

Finally, we compared the numerical accuracy of the first excitation energy of the anharmonic oscillator described by Equation (3.2), but now using a  $n$ -point discrete variable representation (DVR) based upon the Tchebychev polynomials to compute the eigenspectra of the first and second sectors. In Figure 2 we show the numerical error in the first excitation energy by comparing  $E_1^1(n)$  (the first excited state energy from the standard variational calculation with  $n$ -DVR basis functions) and  $E_0^2(n)$  (the ground state of the sector two Hamiltonian computed with  $n$ -DVR basis functions) from an  $n$  point DVR to the numerically "exact" value corresponding to a 100 DVR points,

$$\epsilon_1^1(n) = \log_{10} |E_1^1(n) - E_1^1(\text{exact})|.$$

Likewise,

$$\epsilon_0^2(n) = \log_{10} |E_0^2(n) - E_1^1(\text{exact})|.$$

For any given basis size,  $\epsilon_0^2 < \epsilon_1^1$ . Moreover, over a range of  $15 < n < 40$  points, the excitation energy computed using the second sector's ground state is between 10 and 100 times more accurate than  $E_1^1(n)$ . This effectively reiterates our point that by using the SUSY hierarchy, one can systematically improve upon the accuracy of a given variational calculation. It also illustrates that our conclusion does not depend on the basis set used.

$n$	$N\psi_0^n$	$\mathcal{H}_1$	$N\psi_1^n$	$\mathcal{H}_2$
0	$50\psi_0^1$	2.703955e-06	$44\psi_0^1$	5.263075
1	$60\psi_1^1$	5.263075	$56\psi_0^1$	12.109712
2	$64\psi_2^1$	12.109717	$66\psi_0^1$	20.186019

Table 4. Energies for the Anharmonic Non-Polynomial Oscillator using Hierarchy of Hamiltonians, determined variationally.

$n$	$N\psi_0^n$	$\mathcal{L}_2$	$\mathcal{L}_\infty$
0	$70\psi_0^1$	3.7158761e-07	2.220446e-16
1	$88\psi_1^1$	6.477328e-08	1.221245e-15
0	$76\psi_0^2$	5.659010e-07	2.109424e-15

Table 5. Errors for the Anharmonic Non-Polynomial Oscillator wave functions using Hierarchy of Hamiltonians all determined variationally.

$n$	$\Delta N$	$N\psi_0^n$	$\mathcal{L}_2$	$\mathcal{L}_\infty$
1	14	$74\psi_1^1$	9.750546e-07	3.181791e-16

Table 6. Errors for the Anharmonic Non-Polynomial Oscillator using wave functions Charge Operators to find excited states by applying the correct charge operator to the appropriate ground state

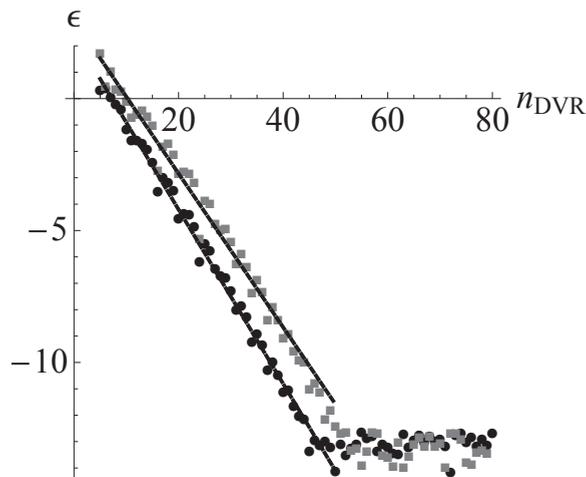


Fig. 2. Convergence of first excitation energy  $E_1^1$  for model potential  $V_1 = x^6 + 4x^4 + x^2 - 2$  using a  $n$ -point discrete variable representation (DVR). Gray squares:  $\epsilon = \log_{10} |E_1^1(n) - E_1^1(exact)|$ , Black squares:  $\epsilon = \log_{10} |E_0^2(n) - E_0^2(exact)|$ . Dashed lines are linear fits.

Another way to approximate excited state energies and wave functions of bound quantum systems is to take advantage of the Ricatti substitution for the purpose of constructing dynamically-adapted, system-specific coherent states. Perhaps the simplest procedure for creating an overcomplete set of such coherent states is to follow the work of Kouri, et al. Klauder & Skagerstam (1985); Kouri et al. (2003). In their approach, it was observed that the "fiducial" function,  $\langle x|\psi(0)\rangle = \psi_0^1(x)$ , could be translated an amount  $x_0$  by application of the shift operator  $e^{-x_0 \frac{d}{dx}}$

$$\begin{aligned} \psi(\alpha|x) &= N e^{ik_0(x-x_0)} e^{-x_0 \frac{d}{dx}} \psi(0|x) \\ &= N e^{ik_0(x-x_0)} e^{-\int_0^{x-x_0} W(x') dx'} \end{aligned} \tag{3.22}$$

where  $\alpha = x_0 + ik_0$  is a point in the phase space Klauder & Skagerstam (1985) which completely describes the coherent state and the de Broglie relation tells us that  $\langle \hat{p}_x \rangle = k_0$ . Using the set of coherent states defined above, we can select a finite subset which remains overcomplete by discretizing the otherwise continuous label  $\alpha = q + ik$  and setting up a von Neumann lattice in phase space with an appropriate density,  $D$ . We define an overcomplete basis of coherent states

$$\{\phi(\alpha_i|x) = N e^{ik_i(x-q_i)} e^{-\int_0^{x-q_i} W(x') dx'} : 1 \leq i \leq M, M \in \mathbb{N}\}, \tag{3.23}$$

where  $M$  is the number of basis functions and the phase space grid points are given by

$$\{(q_i, k_i)\} = \left\{ \left( m\Delta x \sqrt{\frac{2\pi}{D}}, \frac{n}{\Delta x} \sqrt{\frac{2\pi}{D}} \right) \right\}, m, n \in \mathbb{Z}, \tag{3.24}$$

and  $i$  being a joint index consisting of  $m$  and  $n$  Andersson (2001).

Due to the fact that the ground state of the system of interest solves the time-independent Schrödinger equation for the corresponding Hamiltonian, the coherent states defined above build in the dynamics of the system under investigation. This property leads to the expectation that these system-specific coherent states will prove more rapidly convergent in the approximation of excited state energies of bound quantum systems using variational methods.

For  $W(x) = x^3$ , and thus  $V(x) = x^6 - 3x^2$ , we carried out a variational calculation using the system-specific coherent states defined above and compared the accuracy in the approximation of the first three excited state energy eigenvalues with that achieved using the standard harmonic oscillator basis and the harmonic oscillator coherent states. To evaluate the accuracy of each method, we compare the results with a Chebyshev polynomial DVR (Discrete Variable Representation) calculation using 1000 points Littlejohn (2002). The number of decimal places reported in Tables 7-10 correspond to the number of decimal places of agreement with the DVR plus an additional significant figure which is either rounded up or down.

$E_0$	$E_1$	$E_2$	$E_3$
0.000000000	1.935482104	6.298495901	11.680970886

Table 7. DVR Comparison.

$M$	$E_0$	$E_1$	$E_2$	$E_3$
9	0.000000000	1.9355	6.3	11.69
15	0.000000000	1.93548218	6.2985	11.681

Table 8. System-Specific Coherent States.

$M$	$E_0$	$E_1$	$E_2$	$E_3$
9	0.2	2.0	15.0	43.0
15	0.2	2.0	9.0	43.0

Table 9. Harmonic Oscillator Coherent States.

$M$	$E_0$	$E_1$	$E_2$	$E_3$
9	0.07	2.0	6.9	23.0
15	0.01	1.99	6.5	12.0

Table 10. Harmonic Oscillator Basis.

The results indicate that, in fact, the system-specific coherent states provide more accurate approximations of the excited state energies for the anharmonic oscillator given by  $V(x) = x^6 - 3x^2$  when compared with other bases. Namely, they give seven decimal places of agreement with the DVR when 15 basis functions are used. The same number of harmonic oscillator basis functions only provides one decimal point of agreement, and 15 harmonic oscillator coherent states fails to give agreement in the ones place with the DVR. Despite the accuracy achieved using a small number of system-specific coherent states, the non-orthogonality and complex-valued nature of the basis necessitates the calculation of a complex overlap matrix, whose elements must be computed by numerical integration, which is computationally demanding. In order to eliminate numerical integration from the calculation, one can expand the ground state wave function used in the construction of the coherent states into an incomplete set of scaled Gaussians centered about the  $\{q_i: 1 \leq i \leq M\}$ . This expansion would allow us to compute the overlap matrix elements analytically, replacing numerical integration with function evaluation. In particular, we propose to use the Levenberg-Marquardt least-squares curve-fitting algorithm to build an arbitrarily-accurate approximation of each system-specific coherent state in the following manner:

$$\phi(\alpha_i|x) \approx N e^{ik_i(x-q_i)} \sum_{j=1}^{\tilde{M}} c_j e^{-(x-q_i)^2/\sigma_j}. \quad (3.25)$$

#### 4. Generalization to multi-dimensions

In our generalization, we make use of a vectorial approach that simultaneously treats more than one dimension and any number of distinguishable particles. We consider, therefore, a system of  $n$ -particles in three-dimensional space. We denote the coordinates of particle  $i$  by  $(x_i, y_i, z_i)$ . We then define an orthogonal hyperspace of dimension  $3n$ . We take the Hamiltonian for this system to be given by

$$\mathcal{H}_1 = -\nabla^2 + V_1 \quad (4.1)$$

where

$$\vec{\nabla} = \sum_j \vec{\epsilon}_j \frac{\partial}{\partial u_j} \quad (4.2)$$

and  $\vec{\epsilon}_j \cdot \vec{\epsilon}_k = \delta_{jk}$ . The subscript “1” indicates this is the “sector one” Hamiltonian. For simplicity we take the masses of the particles to be equal and use units such that  $\hbar^2/2m = 1$ . For the development here we assume a Cartesian coordinate space, but have provided an extension to more general curvilinear coordinates in a previous publication Kouri et al. (2010b).

As per usual in quantum mechanics, the ground-state wave function is a solution of the Schrödinger equation,

$$\mathcal{H}_1 \psi_0^{(1)} = E_0^{(1)} \psi_0^{(1)}. \quad (4.3)$$

We also emphasize that the lowest energy state,  $\psi_0^{(1)}$ , is nodeless.

We now define a vector superpotential,  $\vec{W}$ , as

$$\vec{W} = -\vec{\nabla} \ln \psi_0^{(1)}, \quad (4.4)$$

which is to say

$$\vec{W} = \sum_{j=1}^{3n} \vec{\epsilon}_j W_j = -\sum_{j=1}^{3n} \vec{\epsilon}_j \frac{\partial}{\partial u_j} \ln \psi_0^{(1)}. \quad (4.5)$$

It is straightforward to see that one can write  $\mathcal{H}_1$  in terms of  $\vec{W}$  as

$$\begin{aligned} (\mathcal{H}_1 - E_0^{(1)}) &= (-\vec{\nabla} + \vec{W}) \cdot (+\vec{\nabla} + \vec{W}) \\ &= (-\partial_i + W_i)(\partial_i + W_i) \\ &= Q_i^\dagger \cdot Q_i, \end{aligned} \quad (4.6)$$

where, according to the Einstein convention, we sum over repeated indices. Since  $(\vec{\nabla} + \vec{W})\psi_0^{(1)} \equiv \vec{Q}\psi_0^{(1)} \equiv \vec{0}$ , it is clear that  $(\mathcal{H}_1 - E_0^{(1)})\psi_0^{(1)} = 0$  as required. We can now define the sector two Hamiltonian such that, above the ground-state ( $E_0^{(1)}$ ), it is isospectral with  $\mathcal{H}_1$ . We do this as follows: for the first excited state in sector one we can write

$$Q_i^\dagger \cdot Q_i \psi_1^{(1)} = (E_1^{(1)} - E_0^{(1)}) \psi_1^{(1)}. \quad (4.7)$$

We then form the tensor product by operating on the left with  $\vec{Q}$  so that

$$(\vec{Q}\vec{Q}^\dagger) \cdot \vec{Q}\psi_1^{(1)} = (E_1^{(1)} - E_0^{(1)}) \vec{Q}\psi_1^{(1)}. \quad (4.8)$$

That is to say, using Einstein notation,

$$(Q_i Q_j^\dagger) Q_j \psi_1^{(1)} = (E_1^{(1)} - E_0^{(1)}) Q_i \psi_1^{(1)}. \quad (4.9)$$

It then follows that  $\vec{Q}\psi_1^{(1)}$  is an eigenstate of the tensor Hamiltonian  $\overleftrightarrow{\mathcal{H}}_2 = (\vec{Q}\vec{Q}^\dagger)$  with energy  $E_0^{(2)} = E_1^{(1)} - E_0^{(1)}$ . Since we are free to set the energy origin, taking  $E_0^{(1)} = 0$  gives  $E_0^{(2)} = E_1^{(1)}$ . It is also clear that  $\vec{Q}\psi_0^{(1)}$  cannot generate a lower energy eigenstate of  $\overleftrightarrow{\mathcal{H}}_2$  since  $\vec{Q}\psi_0^{(1)} = \vec{0}$ , so

that  $\vec{Q}\psi_1^{(1)}$  is indeed proportional to the ground state of  $\overleftarrow{\mathcal{H}}_2$ . The precise relation for obtaining the sector one state from a sector two state is given by

$$\psi_{n+1}^{(1)} = \frac{1}{\sqrt{E_{n+1}^{(1)} - E_0^{(1)}}} \vec{Q}^\dagger \cdot \vec{\psi}_n^{(2)}. \quad (4.10)$$

It is very instructive to illustrate this by considering a simple two dimensional separable harmonic oscillator model problem. This is because we can learn something of how our SUSY formalism works with an exactly soluble problem. We therefore consider a system described by the Hamiltonian

$$\mathcal{H} = -\frac{\partial^2}{\partial u_1^2} - \frac{\partial^2}{\partial u_2^2} + u_1^2 + u_2^2 \quad (4.11)$$

where again, we set  $\hbar^2/2m_1 = \hbar^2/2m_2 = 1$ . The solution of the Schrödinger equation is well known to be the product of one dimensional harmonic oscillator states,

$$\psi_{(n_1, n_2)}^{(1)} = N_{n_1, n_2} \mathcal{H}_{n_1}(u_1) \mathcal{H}_{n_2}(u_2) e^{-(u_1^2 + u_2^2)/2} \quad (4.12)$$

where  $N_{n_1, n_2}$  is the normalization constant and  $\mathcal{H}_n$  denotes the  $n^{\text{th}}$  Hermite polynomial. The ground state is

$$\psi_{(0,0)}^{(1)} = N_{0,0} e^{-(u_1^2 + u_2^2)/2} \quad (4.13)$$

with the zero point energy in this case equals to 2. We next generate the vector superpotential,  $\vec{W}_1$ , as

$$\vec{W}_1 = -\vec{\nabla} \ln \psi_{(0,0)}^{(1)} = u_1 \hat{e}_1 + u_2 \hat{e}_2 \quad (4.14)$$

where

$$\vec{\nabla} = \hat{e}_1 \frac{\partial}{\partial u_1} + \hat{e}_2 \frac{\partial}{\partial u_2}. \quad (4.15)$$

We consider

$$(-\vec{\nabla} + \vec{W}_1) \cdot (\vec{\nabla} + \vec{W}_1) = -\nabla^2 + \vec{W}_1 \cdot \vec{W}_1 - \vec{\nabla} \cdot \vec{W}_1 \quad (4.16)$$

we see that  $\vec{W}_1 \cdot \vec{W}_1 - \vec{\nabla} \cdot \vec{W}_1 = u_1^2 + u_2^2 - 2 = V - 2$ , so that

$$\mathcal{H}_1 = \vec{Q}_1^\dagger \cdot \vec{Q}_1 + 2 \quad (4.17)$$

It is easily verified that

$$\mathcal{H}_1 \psi_{(0,0)}^{(1)} = 2\psi_{(0,0)}^{(1)}, \quad (4.18)$$

as required. The first excited states of  $\mathcal{H}_1$  are doubly degenerate with energy  $E_{(1,0)}^{(1)} = E_{(0,1)}^{(1)} = 3$  and denoted by  $\psi_{(1,0)}^{(1)}$  and  $\psi_{(0,1)}^{(1)}$ . The next excited state,  $\psi_{(1,1)}^{(1)}$  is degenerate with  $\psi_{(0,2)}^{(1)}$  and  $\psi_{(2,0)}^{(1)}$  with energy  $E_{(1,1)}^{(1)} = E_{(2,0)}^{(1)} = E_{(0,2)}^{(1)} = 4$ .

We construct the rigorous sector two Hamiltonian as

$$\overleftarrow{\mathcal{H}}_2 = \overleftarrow{Q}_1 \overleftarrow{Q}_1^\dagger + 2 \overleftarrow{1}, \quad \overleftarrow{1} = \hat{\epsilon}_1 \hat{\epsilon}_1 + \hat{\epsilon}_2 \hat{\epsilon}_2, \quad (4.19)$$

which is a second rank tensor in this case. The Hamiltonian  $\overleftarrow{H}_2$  is then given by

$$\begin{aligned} \overleftarrow{\mathcal{H}}_2 = & \hat{\epsilon}_1 \hat{\epsilon}_1 \left[ -\frac{\partial^2}{\partial u_1^2} + u_1^2 + 3 \right] + \hat{\epsilon}_1 \hat{\epsilon}_2 \left[ -\frac{\partial^2}{\partial u_1 \partial u_2} + u_1 u_2 - u_1 \frac{\partial}{\partial u_2} + u_2 \frac{\partial}{\partial u_1} \right] \\ & + \hat{\epsilon}_2 \hat{\epsilon}_1 \left[ -\frac{\partial^2}{\partial u_1 \partial u_2} + u_1 u_2 - u_2 \frac{\partial}{\partial u_1} + u_1 \frac{\partial}{\partial u_2} \right] + \hat{\epsilon}_2 \hat{\epsilon}_2 \left[ -\frac{\partial^2}{\partial u_2^2} + u_2^2 + 3 \right]. \end{aligned} \quad (4.20)$$

The eigenvalue equation is

$$\overleftarrow{\mathcal{H}}_2 \cdot \vec{\psi}_{(n)}^{(2)} = E_{(n)}^{(2)} \vec{\psi}_{(n)}^{(2)} \quad (4.21)$$

with eigenstates

$$\vec{\psi}_{(n)}^{(2)} = \hat{\epsilon}_1 \psi_{(n)1}^{(2)} + \hat{\epsilon}_2 \psi_{(n)2}^{(2)}. \quad (4.22)$$

It is not difficult to show that there are two degenerate ground state solutions given by

$$\vec{\psi}_{(0)1}^{(2)} = \hat{\epsilon}_1 e^{-(u_1^2 + u_2^2)/2} \quad (4.23)$$

and

$$\vec{\psi}_{(0)2}^{(2)} = \hat{\epsilon}_2 e^{-(u_1^2 + u_2^2)/2}, \quad (4.24)$$

respectively. This is extremely interesting and in contrast to the usual situation in quantum mechanics. For most systems (excluding spin effects) the ground state is unique, i.e., non-degenerate.

We shall see that the degenerate states, Equation (4.23)-(4.24), are exactly what is required for the charge operator,  $\overleftarrow{Q}_1^\dagger$  to produce the doubly degenerate states  $\psi_{(1,0)}^{(1)}$  and  $\psi_{(0,1)}^{(1)}$ . Thus, recall that

$$\overleftarrow{Q}_1^\dagger = \hat{\epsilon}_1 \left( -\frac{\partial}{\partial u_1} + u_1 \right) + \hat{\epsilon}_2 \left( -\frac{\partial}{\partial u_2} + u_2 \right) \quad (4.25)$$

Then

$$\overleftarrow{Q}_1^\dagger \cdot \vec{\psi}_{(0)1}^{(2)} = 2u_1 e^{-(u_1^2 + u_2^2)/2} \propto \psi_{(1,0)}^{(1)} \quad \text{and} \quad \overleftarrow{Q}_1^\dagger \cdot \vec{\psi}_{(0)2}^{(2)} = 2u_2 e^{-(u_1^2 + u_2^2)/2} \propto \psi_{(0,1)}^{(1)} \quad (4.26)$$

Our results Equation (4.23)-(4.24) possess a remarkable property. Only one component is nonzero! We shall see that this is indicative of an extremely interesting property that we observe in the non-separable examples that we consider next. Indeed, we recall that in relativistic quantum mechanics, one obtains a tensor Hamiltonian and the solutions are characterized by large and small components. In the present case, the small component is exactly zero. In the degenerate pair of solutions, which component is zero changes. We stress, however, that any linear combination of the two degenerate solutions is also a solution of the same energy.

With a view toward the next section, where we consider a two dimensional, nonseparable anharmonic oscillator(or equivalently a pair of one dimensional coupled oscillators), we form the equivalent degenerate solutions

$$\vec{\phi}_{(0)1}^{(2)} = N e^{-(u_1^2 + u_2^2)/2} [\hat{e}_1 + \hat{e}_2] \quad \text{and} \quad e^{-(u_1^2 + u_2^2)/2} [\hat{e}_1 - \hat{e}_2] \quad (4.27)$$

In this case, both components of the 2-degenerate solutions are non-zero, of the same magnitude and of definite sign. In dealing with the two dimensional separable harmonic oscillator, the most convenient form is given by Equation (4.26).

A major concern is whether our approach satisfies the supersymmetric algebra which we will consider here. It is clear that we can define our Hamiltonian operator by

$$\mathcal{H} = \begin{pmatrix} \vec{Q}^\dagger \cdot \vec{Q} & 0 \\ 0 & \vec{Q} \vec{Q}^\dagger \end{pmatrix} = \begin{pmatrix} \mathcal{H}_1 & 0 \\ 0 & \overleftarrow{\mathcal{H}}_2 \end{pmatrix} \quad (4.28)$$

where the zero in the upper right is a row vector and the zero in the lower left is a column vector. This Hamiltonian will act on the state

$$\vec{\psi} = \begin{pmatrix} \psi_n^{(1)} \\ \vec{\psi}_{n-1}^{(2)} \end{pmatrix}. \quad (4.29)$$

Then, we can define a “super-charge” operator as

$$\mathcal{Q} = \begin{pmatrix} 0 & 0 \\ \vec{Q} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ Q_1 & 0 & 0 \\ Q_2 & 0 & 0 \end{pmatrix} \quad (4.30)$$

with the adjoint being

$$\mathcal{Q}^\dagger = \begin{pmatrix} 0 & \vec{Q}^\dagger \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & Q_1^\dagger & Q_2^\dagger \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.31)$$

If we take the product of  $\mathcal{Q}^\dagger$  and  $\mathcal{Q}$ , we find that

$$\overleftarrow{\mathcal{H}}_1 = \begin{pmatrix} \vec{Q}^\dagger \cdot \vec{Q} & 0 \\ 0 & 0 \end{pmatrix} \quad (4.32)$$

and similarly, if we take the product of  $\mathcal{Q}$  and  $\mathcal{Q}^\dagger$ , we find that

$$\overleftarrow{\mathcal{H}}_2 = \begin{pmatrix} 0 & 0 \\ 0 & \vec{Q} \vec{Q}^\dagger \end{pmatrix}. \quad (4.33)$$

It's straightforward to show that

$$[\mathcal{Q}, \mathcal{H}] = 0, \quad (4.34)$$

$$[\mathcal{Q}^\dagger \mathcal{Q}, \mathcal{Q} \mathcal{Q}^\dagger] = 0, \quad (4.35)$$

$$\mathcal{Q} \mathcal{Q} = \mathcal{Q}^\dagger \mathcal{Q}^\dagger = 0, \quad (4.36)$$

and

$$\overleftarrow{\mathcal{H}} = \{\mathcal{Q}, \mathcal{Q}^\dagger\} \quad (4.37)$$

where the braces indicate the anticommutator bracket. According to Wess and Bagger's text on supersymmetry, these are the necessary conditions for a superalgebra Wess & Bagger (1992).

We now consider in more detail the degeneracy between the two sectors,  $\vec{Q}_1$  and  $\vec{Q}_1^\dagger$  and the sector Hamiltonians  $\mathcal{H}_1$  and  $\overleftarrow{\mathcal{H}}_2$ . An important consequence of the algebra is the existence of "inter-twining" relations. These are of fundamental importance because they underlie the isospectral property and in addition, they can be used to establish the unique correspondence between the eigenstates of sectors 1 and 2. Indeed, they are responsible for establishing the completeness of the eigenvectors  $\{\overleftarrow{\psi}_n^{(2)}\}$  of  $\overleftarrow{\mathcal{H}}_2$ . It is of interest to note that inter-twining relations are essential to the fact that in ordinary quantum scattering, the continua of the full Hamiltonian,  $\mathcal{H}$ , and the unperturbed Hamiltonian  $\mathcal{H}_0$  (where  $\mathcal{H} = \mathcal{H}_0 + V$ , with  $V$  the perturbation responsible for scattering) coincide. In that case, the inter-twining relation is

$$\Omega^+ e^{i\mathcal{H}_0 t/\hbar} = e^{i\mathcal{H}t/\hbar} \Omega^+. \quad (4.38)$$

It is useful to derive the SUSY inter-twining relations explicitly. We have

$$\mathcal{H}_1 \psi_n^{(1)} = E_n^{(1)} \psi_n^{(1)}, \quad (4.39)$$

where  $\mathcal{H}_1$  is a standard Schrödinger operator (comprised of a Laplacian for the kinetic energy and a Hermitian potential  $V_1$ ). One result of this fact is that the ground state of  $\mathcal{H}_1$  is nodeless. In addition,  $\mathcal{H}_1$  is Hermitian and a well known postulate of quantum mechanics asserts that its eigenstates are complete. Essentially from a physical standpoint (as opposed to pure mathematics)  $\mathcal{H}_1$  is required to be Hermitian because (a) it represents an observable, implying only real eigenvalues (b) quantum mechanics further asserts that these eigenvalues are the only values that can be obtained when measuring the energy for the physical system represented by  $\mathcal{H}_1$ . This implies that any physically realizable state,  $\psi$ , of the system must be a superposition (in general) of these and only these eigenvectors. This then implies that the set  $\{\psi_n^{(1)}\}$  is complete on the physically allowed space of state vectors.

To derive the inter-twining relation, we again recall that the charge operator (which exactly factors  $\mathcal{H}_1$ ) is such that

$$\mathcal{H}_1 = \vec{Q}_1^\dagger \cdot \vec{Q}_1 + E_0^{(1)} \quad (4.40)$$

where

$$\vec{Q}_1 \psi_0^{(1)} \equiv \vec{0} \quad (4.41)$$

The general sector 1 Schrödinger equation is

$$(\vec{Q}_1^\dagger \cdot \vec{Q}_1 + E_0^{(1)}) \psi_n^{(1)} = E_n^{(1)} \psi_n^{(1)} \quad (4.42)$$

where we now assume that  $n > 0$ ; i.e.,  $\psi_n^{(1)}$  is an excited state of  $\mathcal{H}_1$ . We apply  $\vec{Q}_1$  to Equation (4.42) to find

$$(\vec{Q}_1 \vec{Q}_1^\dagger \cdot \vec{Q}_1 + E_0^{(1)}) \psi_n^{(1)} = E_n^{(1)} \vec{Q}_1 \psi_n^{(1)} \quad (4.43)$$

We define  $\overleftarrow{\mathcal{H}}_2$  as

$$\overleftarrow{\mathcal{H}}_2 \equiv \vec{Q}_1 \vec{Q}_1^\dagger + E_0^{(1)} \quad (4.44)$$

so that Equation (4.43) yields

$$\vec{Q}_1 \mathcal{H}_1 = \overleftarrow{\mathcal{H}}_2 \cdot \vec{Q}_1. \quad (4.45)$$

*This is the inter-twining relation.* Let us explore inter-twining consequences further. Suppose we consider an eigenstate,  $\psi_n^{(1)}$  of  $\mathcal{H}_1$ . It follows from Equation (4.45) that there is also a unique eigenstate of  $\overleftarrow{\mathcal{H}}_2$ ,  $\vec{Q}_1 \psi_n^{(1)}$  with the same energy. Next, assume that  $\overleftarrow{\mathcal{H}}_2$  possesses an eigenvalue  $E_\lambda^{(2)}$  that differs from all of the  $E_n^{(1)}$ . Then we have

$$\overleftarrow{\mathcal{H}}_2 \cdot \vec{\psi}_\lambda^{(2)} = E_\lambda^{(2)} \vec{\psi}_\lambda^{(2)} \quad (4.46)$$

Now  $\overleftarrow{\mathcal{H}}_2$  and  $\mathcal{H}_1$  are manifestly Hermitian. Taking the adjoint of Equation (4.45) yields

$$\mathcal{H}_1 \vec{Q}_1^\dagger = \vec{Q}_1^\dagger \cdot \overleftarrow{\mathcal{H}}_2 \quad (4.47)$$

which is again an inter-twining relation. We then take the scalar product of Equation (4.46) with  $\vec{Q}_1^\dagger$

$$\vec{Q}_1^\dagger \cdot \overleftarrow{\mathcal{H}}_2 \cdot \vec{\psi}_\lambda^{(2)} = E_\lambda^{(2)} \vec{Q}_1^\dagger \cdot \vec{\psi}_\lambda^{(2)} \quad (4.48)$$

But by the adjoint inter-twining relation, we have

$$\mathcal{H}_1 \vec{Q}_1^\dagger \cdot \vec{\psi}_\lambda^{(2)} = E_\lambda^{(2)} \vec{Q}_1^\dagger \cdot \vec{\psi}_\lambda^{(2)} \quad (4.49)$$

Thus, we find that  $\mathcal{H}_1$  also has the scalar eigenvector

$$\psi_\lambda^{(1)} \propto \vec{Q}_1^\dagger \cdot \vec{\psi}_\lambda^{(2)} \quad (4.50)$$

and its eigenvalue is equal to  $E_\lambda^{(2)}$ . This violates our initial assertion that  $\mathcal{H}_1$  did not have the eigenvalue  $E_\lambda^{(2)}$ . We conclude that for eigenvectors  $\vec{\psi}_n^{(2)}$ , there corresponds a unique eigenvector  $\psi_{n'}^{(1)}$ , where  $n' \equiv n + 1$ . That is,  $\vec{\psi}_0^{(2)}$  must have the same energy as the first excited state  $\psi_1^{(1)}$ . It can not be lower than  $E_1^{(1)}$  because it is the lowest eigenvalue of  $\overleftarrow{\mathcal{H}}_2$  and it cannot equal  $E_0^{(1)}$ . In fact, the inter-twining relation is sufficient to establish that  $\overleftarrow{\mathcal{H}}_2$  is a Schrödinger operator and as such, its eigenvectors must be complete on the space  $\vec{\psi}$ . Note that we are not saying that the  $\vec{\psi}_n^{(2)}$  span the the space generated by  $\mathcal{H}_1$ . They are completely separate vector spaces arising from two distinct Hermitian Hamiltonians. All of the above can be made mathematically rigorous but our purpose here is to supply a physically reasonable argument for the properties of the tensor sector. Finally, at no point in this discussion have we imposed a condition that the spectra of  $\mathcal{H}_1$  (and  $\overleftarrow{\mathcal{H}}_2$ ) are strictly discrete. The inter-twining relations hold for systems with a mixed discrete and continuous spectra and even for systems with a purely continuous spectrum.

## 5. Clusters of distinguishable particles

### 5.1 Degenerate case

We next consider a model non-separable two dimensional anharmonic oscillator system for sector one for which the ground state energy is zero and the ground state wave function is exactly given by

$$\psi_{(0)}^{(1)}(u_1, u_2) = N \exp(-u_1^2 u_2^2 - u_1^2 - u_2^2) \quad (5.1)$$

We can generate the superpotential corresponding to this ground state as

$$\vec{W}_1 = -\vec{\nabla} \ln \psi_{(0)}^{(1)}(u_1, u_2) \quad (5.2)$$

having the components  $W_{11} = (2u_1 u_2^2 + 2u_1)$  and  $W_{12} = (2u_1^2 u_2 + 2u_2)$ , respectively. Now using these components we can generate the model potential for sector one. Thus we get the Hamiltonian for sector one of the following form

$$\begin{aligned} \mathcal{H}_1 = -\nabla^2 + V_1(u_1, u_2) &= -\frac{\partial^2}{\partial u_1^2} - \frac{\partial^2}{\partial u_2^2} + (2u_1 u_2^2 + 2u_1)^2 + (2u_1^2 u_2 + 2u_2)^2 \\ &\quad - 2(u_1^2 + 1) - 2(u_2^2 + 1) \end{aligned} \quad (5.3)$$

In this case, the exact ground state energy is  $E_0^{(1)} = 0$ . The sector two tensor Hamiltonian can be generated with  $\vec{\nabla}$  and  $\vec{W}_1$ . The calculation for sector one and sector two eigenvalues and eigenfunctions is done variationally by diagonalizing each sector Hamiltonian in an approximate truncated basis. We choose to employ a basis of the direct product of the eigenstates of a harmonic oscillator in each dimension, each with frequency  $\omega = 2\sqrt{2}$ . The trial wave function for sector one is

$$\psi_{(trial)}^{(1)}(u_1, u_2) = \sum_{m,n} C_{m,n}^{(1)} \phi_m(\alpha, u_1) \phi_n(\alpha, u_2) \quad (5.4)$$

where  $\alpha = \sqrt{m\omega/\hbar}$ . Similarly for the sector two the trial wave functions for each component are

$$\begin{aligned} \psi_{(trial)1}^{(2)}(u_1, u_2) &= \sum_{m,n} C_{1,m,n}^{(2)} \phi_m(\alpha, u_1) \phi_n(\alpha, u_2) \\ \psi_{(trial)2}^{(2)}(u_1, u_2) &= \sum_{m,n} C_{2,m,n}^{(2)} \phi_m(\alpha, u_1) \phi_n(\alpha, u_2) \end{aligned} \quad (5.5)$$

Using these trial wave functions and treating the  $C_{m,n}$  as a variational parameters, we arrive at the Hermitian eigenvalue equation for both sectors. For sector one the form is

$$\mathcal{H}_1 C^{(1)} = E C^{(1)} \quad (5.6)$$

and that for sector two is

$$\begin{pmatrix} \mathcal{H}_{11}^{(2)} & \mathcal{H}_{12}^{(2)} \\ \mathcal{H}_{21}^{(2)} & \mathcal{H}_{22}^{(2)} \end{pmatrix} \begin{pmatrix} C_1^{(2)} \\ C_2^{(2)} \end{pmatrix} = E \begin{pmatrix} C_1^{(2)} \\ C_2^{(2)} \end{pmatrix} \quad (5.7)$$

Each term of the Hamiltonian matrix can be calculated analytically in the harmonic oscillator basis.

We have calculated energies and wave functions of the Hamiltonian in Equation (5.3) for sectors one and two using the variational approach we just described. In all calculations, we use the exact  $\psi_0^{(1)}$  to generate  $\vec{W}_1$  and  $\overleftrightarrow{\mathcal{H}}_2$  exactly. In Table 11 we compare sector one and sector two energies for different harmonic oscillator basis set sizes. The notation  $N_{u_1}$ ,  $N_{u_2}$  gives the number of basis functions for the variable  $u_1$ ,  $u_2$ , respectively. The first row gives the approximate results for  $E_0^{(1)}$ . The next is the doubly degenerate first excited state energy,  $E_1^{(1)}$  followed by the sector two ground state energy,  $E_0^{(2)}$ . The third row contains  $E_2^{(1)}$  and  $E_1^{(2)}$ , for different basis sets. It is easily seen that the doubly degenerate ground state of sector two is also isospectral with the doubly degenerate first excited state of sector one. This correspondence is clearly in accordance with the general SUSY prediction about the eigenstates for the two supersymmetric partner potentials. For the higher excited states this precise correspondence between the two sectors breaks down when we use a small number of basis functions (*i.e.*, there appear some "spurious" solutions) but it is gradually restored by increasing the basis size. We attribute this apparent breakdown of the SUSY-correspondence for higher states to the error that arises in the calculation due to the truncation of an infinite basis to a finite one. Essentially, some "spurious" eigenvalues appear in the SUSY-QM sector two spectrum, but they disappear as the basis size is increased. This may raise a question regarding the precise nature of the Hylleraas-Undheim theorem for the SUSY sector two tensor Hamiltonian.

The accuracy of the variational results are known for the ground state of sector one, since we know the exact energy is  $E_0^{(1)} = 0$ . Thus, the (10, 10) basis gives an error of  $9 \times 10^{-3}$  while the (60, 60) basis gives an error of  $4.9 \times 10^{-9}$ . In the case of the first excited state of sector one, the error for the (10, 10) basis (computed relative to (60, 60) basis result) is 0.0634. By contrast, the error in the (10, 10) basis result for the sector two ground state (again, relative to the (60, 60) basis result) is  $2.2 \times 10^{-4}$ . Consequently, the use of the sector two Hamiltonian for a ground state calculation enables us to obtain much improved accuracy for the first excitation energy of sector one. Basically, we estimate an increase in accuracy (defined as the ratio of the accuracy of the sector one result to that of the sector two result) to be a factor of 280. Our exploratory calculation thus clearly reveals that for the calculation of excited state energies, the SUSY-variational method requires a smaller number of basis functions to achieve the same order of accuracy. Of course, this level of accuracy resulted in part because we have used the exact  $\overleftrightarrow{\mathcal{H}}_2$ .

As this model problem has no analytical solution for the excited states, we have taken the results of the (60, 60) basis set calculation as the reference result for both sectors in order to check the convergence in wave functions. In Tables 12 and 13 we compare the  $\mathcal{L}_\infty$  and  $\mathcal{L}_2$  error of the first excited states of the sector one that we have obtained by the SUSY-variational calculation and the simple variational calculation. The  $\mathcal{L}_\infty$  error is defined as the absolute maximum difference between the solution computed with an infinite basis set ( $\psi^{(1)}(\infty)$ ) which we approximate with the (60,60) basis, and a smaller finite ( $n, n$ ) basis set ( $\psi^{(1)}(n)$ )

$$\mathcal{L}_\infty = \text{Max}\{|\psi^{(1)}(\infty) - \psi^{(1)}(n)|\}.$$

The  $\mathcal{L}_2$  error is defined by

$$\mathcal{L}_2 = \int_{-\infty}^{\infty} du_1 \int_{-\infty}^{\infty} du_2 |\psi^{(1)}(\infty) - \psi^{(1)}(n)|^2.$$

In the first column of Table 12 we show the difference in the number of basis states used (in each degree of freedom) and the maximum,  $N_{u_1} = N_{u_2} = 60$ , used for the reference result. Since  $\mathcal{L}_2$  and  $\mathcal{L}_\infty$  are computed relative to the  $N_{u_1}, N_{u_2} = 60, 60$  basis, they measure the degree of convergence of the calculations. It is clear from Tables 13 and 13 that the state obtained from the SUSY relation  $\psi_{(1,0)}^{(1)} = \vec{Q}_1^+ \cdot \vec{\psi}_{(0)}^{(2)}$  converges more rapidly than the result obtained directly from the variational solution for sector one. We note that the same level of convergence is obtained for both of the degenerate wave functions. Since the analytical solution for the ground state wave function of the sector one is known, we also have calculated the  $\mathcal{L}_2$  and  $\mathcal{L}_\infty$  error for this wave function, comparing the analytical and variational wave function of sector one for different numbers of basis states to determine a basis size which gives a satisfactory convergence. The results are given in Table 14. It is again clear that the variational results for the sector one ground state wave function are very well converged.

sector one $N_{u_1}, N_{u_2} = 10$ (in a.u.)	sector two $N_{u_1}, N_{u_2} = 10$ (in a.u.)	sector one $N_{u_1}, N_{u_2} = 40$ (in a.u.)	sector two $N_{u_1}, N_{u_2} = 40$ (in a.u.)	sector one $N_{u_1}, N_{u_2} = 60$ (in a.u.)	sector two $N_{u_1}, N_{u_2} = 60$ (in a.u.)
$9.0 \times 10^{-3}$	-	$4.0 \times 10^{-7}$	-	$5.0 \times 10^{-9}$	-
4.6	4.5849	4.58473	4.5847275	4.5847275	4.58472742
8.3	8.005	8.00007	8.0000005	8.000001	8.00000005

Table 11. Comparison of energy eigenvalues of sector one and sector two for different number of basis functions ( $N_{u_1}, N_{u_2}$ ).

$\Delta N = N_{ref} - n$	Error	$\psi_{(1,0)}^{(1)} = \vec{Q}_1^+ \cdot \vec{\psi}_{(0)}^{(2)}$	$\psi_{(1,0)}^{(1)}$
40 = 60 - 20	$\mathcal{L}_\infty$	$1.1 \times 10^{-4}$	$4.9 \times 10^{-4}$
30 = 60 - 30	$\mathcal{L}_\infty$	$2.2 \times 10^{-5}$	$8.5 \times 10^{-5}$
20 = 60 - 40	$\mathcal{L}_\infty$	$5.3 \times 10^{-6}$	$1.9 \times 10^{-5}$
10 = 60 - 50	$\mathcal{L}_\infty$	$1.6 \times 10^{-6}$	$3.9 \times 10^{-6}$

Table 12. Comparison between wave-function  $\mathcal{L}_\infty$  -Error for the doubly-degenerate sector one excited state, (1, 0) generated by standard variational calculation of the sector one and variational SUSY calculation for sector two ground state, followed by application of the SUSY Charge Operator for different number of basis ( $n = N_{u_1}, N_{u_2}$ ). ( $N_{ref} = (N_{u_1} = 60, N_{u_2} = 60)$ ).

In Figures 3(a-b) we show the two components of one of the degenerate sector two ground state wave functions and in Figures 3(c-d) we show the two components for the other degenerate sector two ground state wave function. It may seem problematic that the components  $\psi_{(0)1}^{(2)}$  and  $\psi_{(0)2}^{(2)}$  for the pair of sector two ground state wave functions have nodes. We shall see below that these nodes can be eliminated in a very simple manner. However, we stress that for each of the degenerate sector two ground state wave functions, there is a large and small component. Unlike the two dimensional separable harmonic oscillator case,

$\Delta N = N_{ref} - n$	Error	$\psi_{(1,0)}^{(1)} = \vec{Q}_1^\dagger \cdot \vec{\psi}_{(0)}^{(2)}$	$\psi_{(1,0)}^{(1)}$
40 = 60 - 20	$\mathcal{L}_2$	$3.0 \times 10^{-6}$	$1.8 \times 10^{-5}$
30 = 60 - 30	$\mathcal{L}_2$	$1.1 \times 10^{-7}$	$5.7 \times 10^{-7}$
20 = 60 - 40	$\mathcal{L}_2$	$6.2 \times 10^{-9}$	$3.0 \times 10^{-8}$
10 = 60 - 50	$\mathcal{L}_2$	$3.5 \times 10^{-10}$	$1.5 \times 10^{-9}$

Table 13. Comparison between wave-function  $\mathcal{L}_2$ -Error for the doubly-degenerate sector one Excited state, (1, 0) generated by standard variational calculation of the sector one and variational SUSY calculation for the sector two ground state, followed by application of the SUSY charge operator for different size basis ( $n = N_{u_1}, N_{u_2}$ ). ( $N_{ref} = (N_{u_1} = 60, N_{u_2} = 60)$ ).

$N_{u_1}, N_{u_2}$	$\mathcal{L}_\infty$	$\mathcal{L}_2$
20, 20	$9.3 \times 10^{-5}$	$1.5 \times 10^{-6}$
30, 30	$1.5 \times 10^{-5}$	$3.9 \times 10^{-8}$
40, 40	$3.4 \times 10^{-6}$	$2.0 \times 10^{-9}$
50, 50	$9.1 \times 10^{-7}$	$1.4 \times 10^{-10}$
60, 60	$2.8 \times 10^{-7}$	$1.3 \times 10^{-11}$

Table 14. Comparison between wave-function  $\mathcal{L}_2$  and  $L_\infty$  -Error for the 1<sup>st</sup> sector exact ground state wave function  $\psi_{(0,0)}^{(1)}(\infty)$  and variationally calculated ground state wave function  $\psi_{(0,0)}^{(1)}(n)$  for different number of basis states ( $n = N_{u_1}, N_{u_2}$ ).

the small component is not only non-zero but it has nodes. It is roughly ten times smaller in magnitude than the large component. The two degenerate states are  $90^\circ$  out of phase so far as their signs. In Figures 4(a-b) we show the first excited states  $\psi_{(0,1)}^{(1)}$  and  $\psi_{(1,0)}^{(1)}$  of the sector one that we have obtained after applying the SUSY charge operator to the sector two ground states and Figures 4(c-d) present the same states that were found variationally from the sector one Hamiltonian. The similarity of Figure 4a to 4c and 4b to 4d clearly reflects the correctness of our method. To eliminate the nodes in the components of  $\vec{\psi}_{(0)}^{(2)}$  and  $\vec{\psi}_{(0')}^{(2)}$ , we note that since they are degenerate, any linear combination of them is also a valid wave function. Accordingly, in analogy to the separable two dimensional harmonic oscillator considered previously we can define  $\vec{\phi}_{(0)}^{(2)}$  and  $\vec{\phi}_{(0')}^{(2)}$  by combining the components of  $\vec{\psi}_{(0)}^{(2)}$  and  $\vec{\psi}_{(0')}^{(2)}$  according to

$$\phi_{(0)1}^{(2)} = \psi_{(0)1}^{(2)} + \psi_{(0')1}^{(2)} \quad (5.8)$$

$$\phi_{(0)2}^{(2)} = \psi_{(0)2}^{(2)} + \psi_{(0')2}^{(2)} \quad (5.9)$$

$$\phi_{(0')1}^{(2)} = \psi_{(0)1}^{(2)} - \psi_{(0')1}^{(2)} \quad (5.10)$$

$$\phi_{(0')2}^{(2)} = \psi_{(0)2}^{(2)} - \psi_{(0')2}^{(2)}. \quad (5.11)$$

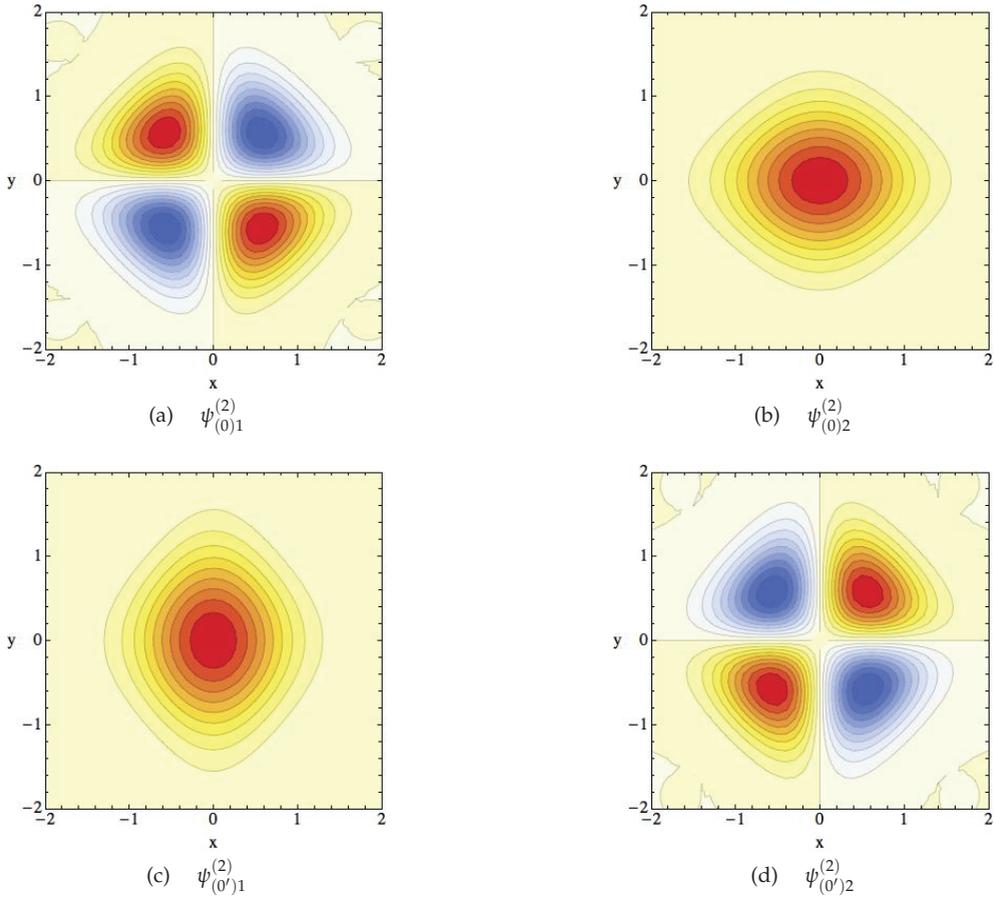


Fig. 3. (a-b) represent the two components of one of the degenerate sector two ground states. (c-d) represent the same for the other sector two ground state. Contour shading is such that red indicates positive amplitude and blue indicates negative amplitude. The prime on the quantum number "0" denotes the second of the 2 degenerate ground states.

In Figures 5(a-d) we show the components of  $\bar{\phi}_{(0)}^{(2)}$  and  $\bar{\phi}_{(0')}^{(2)}$ . These combinations are nodeless and have definite symmetry. We stress that the forms of the above  $\bar{\phi}_{(0)}^{(2)}$  and  $\bar{\phi}_{(0')}^{(2)}$  are analogous to the results in Equation (4.27), obtained for the degenerate separable two dimensional harmonic oscillator.

For sector two, in the case of a doubly degenerate first excited state of  $\mathcal{H}_1$ , we obtain a doubly degenerate ground state and the energies obtained by the Rayleigh-Ritz method are consistently lower for all the excited states of the sector one Hamiltonian, for the same basis size. In addition, the SUSY-QM sector two result for the first excited state energy is always several orders of magnitude more accurate than the Rayleigh-Ritz result for sector one for any given basis set size. Assessing the accuracy of the excited state wave functions is more

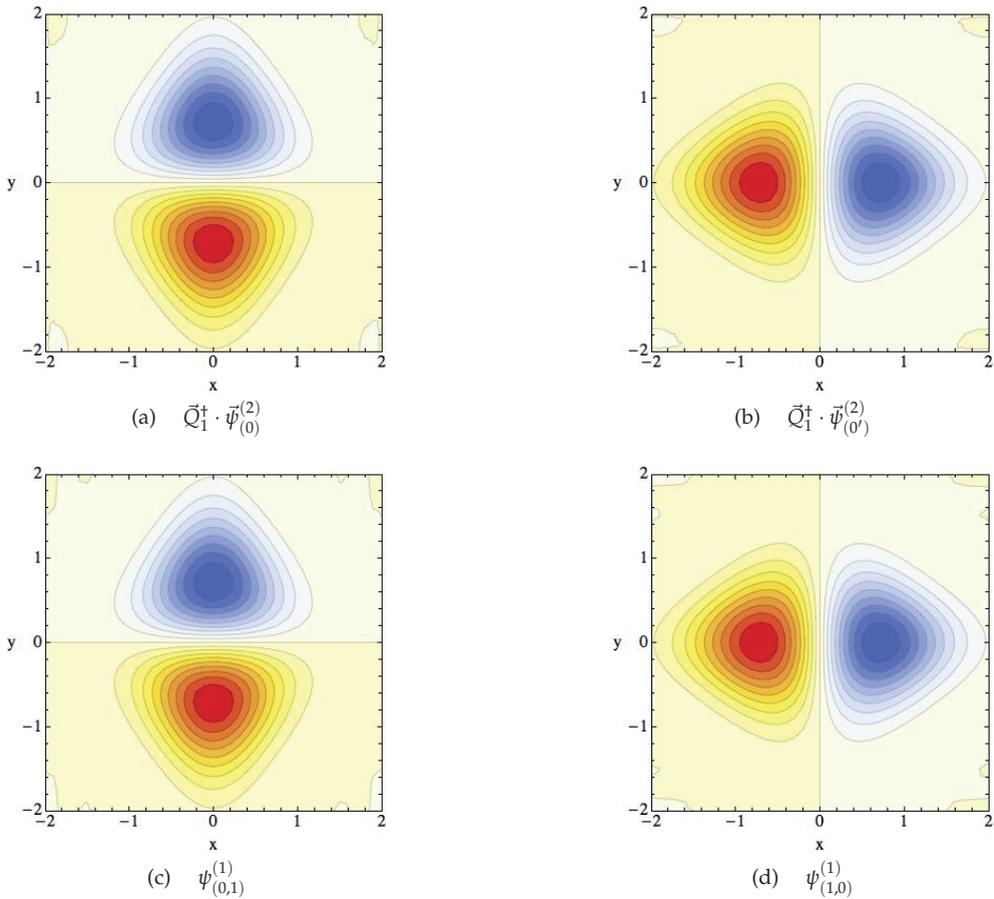


Fig. 4. (a-b) show the doubly degenerate  $1^{st}$  excited states of the sector one. The two states are generated by the SUSY -variational method. (c-d) the corresponding states of the sector one Hamiltonian, which are generated variationally.

difficult. We chose to do this in terms of convergence of the wave functions relative to the largest basis set results. However, we are able to assess the accuracy of our variational results quantitatively in the case of the sector one ground state, since it is exactly known. We report our results in terms of  $\mathcal{L}_2$  and  $\mathcal{L}_\infty$  measures, as is typical for assessing convergence and accuracy of functions in a Hilbert space. We find that the  $\mathcal{L}_2$  and  $\mathcal{L}_\infty$  accuracies of the SUSY-QM results are consistently better than the Rayleigh-Ritz results for excited states of sector one. As a further proof of this, we also consider the accuracy for the ground state of sector one (where we have the exact wave function), with the variational result. In fact, we find that the convergence of the sector two ground state is consistently better than the convergence of the variationally obtained ground state wave function for sector one.

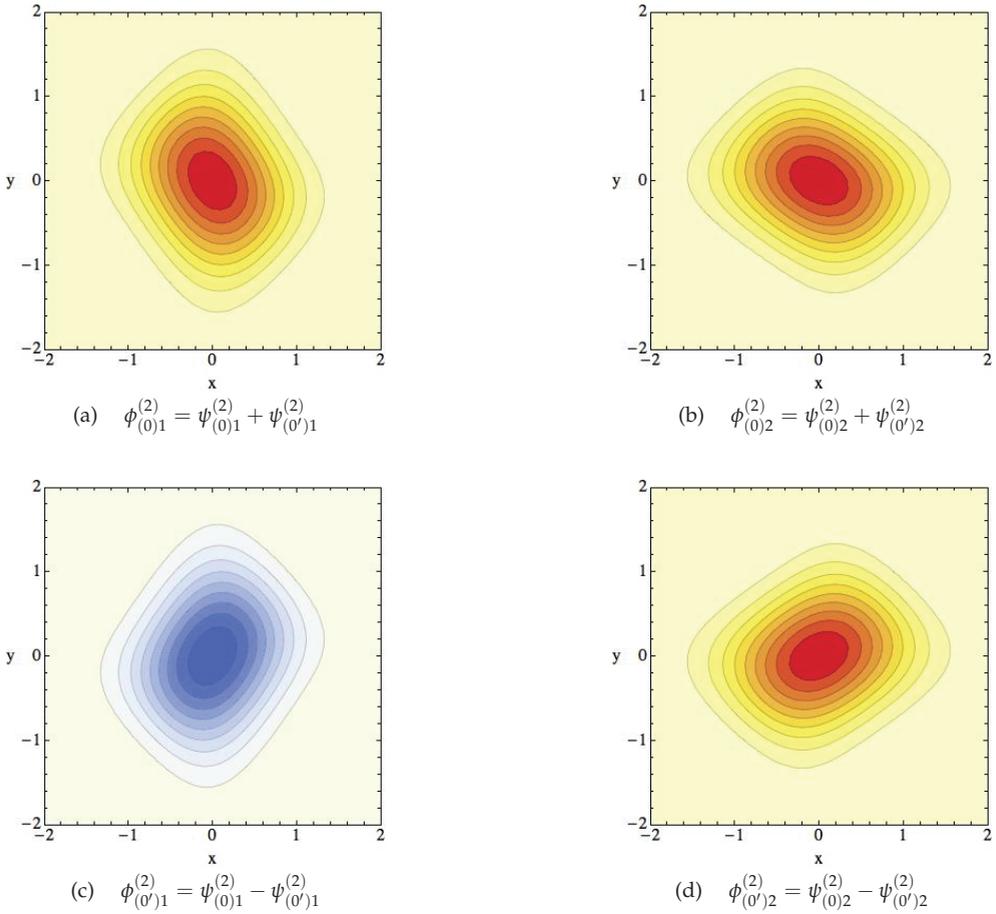


Fig. 5. Components of linear combinations of the two degenerate ground states of the sector two.

**5.2 Non-degenerate case**

For completeness of our presentation, we also consider a non-degenerate, two dimensional anharmonic oscillator model. To generate such a Hamiltonian, we modify the ground state in (5.1) to the form

$$\psi_{(0)}^{(1)}(u_1, u_2) = N \exp(-2u_1^2 u_2^2 - u_1^2 - \sqrt{2}u_2^2) \tag{5.12}$$

Then the exact Hamiltonians  $\mathcal{H}_1$  and  $\overleftrightarrow{\mathcal{H}}_2$  are readily generated. However, we also shall generate an approximate  $\overleftrightarrow{\mathcal{H}}_2$  using the variationally determined, approximate ground state wave function. The formal structure of the equations is the same as that above. In the case

where  $\psi_{0,approx}^{(1)}$  is used to develop  $\vec{W}_1$  and  $\overleftrightarrow{H}_2$ , we shall see that its accuracy is an extremely important consideration.

In the non-degenerate example, we have performed two distinct calculations. First we used the exact  $\vec{W}_1$  to construct the exact  $\overleftrightarrow{H}_2$ . The results are given in Tables 15 - 17. In Table 15, we see basically the same behavior as was obtained in the non-separable degenerate two dimensional example. The errors in the ground state energy are of similar size for both the degenerate and non-degenerate cases, with the same variation with basis set size. This behavior extends also to the first and second excited state energies. We conclude that the presence or absence of degeneracy does not affect the performance of our SUSY approach when the exact  $\vec{W}_1$  is used.

In the case of the  $\mathcal{L}_\infty$  and  $\mathcal{L}_2$  errors obtained when using the exact  $\vec{W}_1$ , we again see the same basic behavior with regard to the convergence of the wave functions.

However, the situation is more interesting when we use the variationally obtained approximate ground state,  $\psi_{0,approx}^{(1)}$  to generate  $\vec{W}_1^{approx}$  and thereby  $\overleftrightarrow{H}_{2,approx}$ . These results are shown in Table 18 and are compared to the exact  $\vec{W}_1$  results (the columns labeled  $E_1^{(1)}$  and  $E_0^{(2)}$ ). Results are shown for three different basis set sizes, (10,10), (20,20), and (30,30). Now because we are using an approximate  $\psi_0^{(1)}$  to generate  $\vec{W}_1^{approx}$ , it is important to note how the accuracy depends not only on basis size (which affects the accuracy of  $\psi_0^{(1)}$ ) but also how the accuracy of  $\vec{W}_1^{approx}$  is affected by errors in  $\psi_0^{(1)}$ . We have found that  $\vec{W}_1^{approx}$  is most sensitive to errors in region where  $\psi_0^{(1)}$  is small in magnitude. This is reasonable since  $\vec{W}_1^{approx} = -\vec{\nabla} \ln \psi_{0,approx}^{(1)}$  and we expect that  $(\partial \psi_0^{(1)} / \partial u_j) / \psi_0^{(1)}$  to be most sensitive to errors in regions where  $\psi_0^{(1)}$  is smallest in magnitude. In view of this, we have introduced  $\psi_{cutoff}$  levels at which we cease calculating  $\vec{W}_1$ . These correspond to cutoff values of  $\psi_{0,cutoff}^{(1)} = 10^{-10}, 10^{-5}, 10^{-3}$  and  $10^{-2}$ . Those results are in columns 4 - 7 in Table 18. It is clear that the SUSY result is always better than the sector one variational result, although this is only marginally the case with the very small cutoff values (i.e.,  $\psi_{0,cutoff}^{(1)} \leq 10^{-5}$ ). The best results are obtained with the  $10^{-2}$  cutoff value. While obviously, this is a single computational example, it is encouraging. However, additional careful studies are underway.

sector one $N_{u_1}, N_{u_2} = 10$ (in a.u.)	sector two $N_{u_1}, N_{u_2} = 10$ (in a.u.)	sector one $N_{u_1}, N_{u_2} = 40$ (in a.u.)	sector two $N_{u_1}, N_{u_2} = 40$ (in a.u.)	sector one $N_{u_1}, N_{u_2} = 60$ (in a.u.)	sector two $N_{u_1}, N_{u_2} = 60$ (in a.u.)
$6.4 \times 10^{-3}$	-	$3.4 \times 10^{-7}$	-	$3.6 \times 10^{-9}$	-
4.80	4.752	4.75181	4.75180771	4.75180778	4.75180770
6.70	6.65	6.64636	6.64634938	6.6463495	6.64634937

Table 15. Comparison of energy eigenvalues of sector one and sector two for different number of basis functions ( $N_{u_1}, N_{u_2}$ ). Exact sector one and sector two Hamiltonians are used.

Finally, in Figures 6(a-b), we give the two components of the (non-degenerate) ground state,  $\psi_{(0)1}^{(2)}$  and  $\psi_{(0)2}^{(2)}$ . In Figures 6(c-d), we display the components of the first excited state,  $\psi_{(1)1}^{(2)}$  and  $\psi_{(1)2}^{(2)}$ . We note that they are qualitatively similar to the results obtained for the degenerate

$\Delta N = N_{ref} - n$	Error	$\psi_{(1)}^{(1)} = \vec{Q}_1^+ \vec{\psi}_{(0)}^{(2)}$	$\psi_{(1)}^{(1)}$
40 = 60 - 20	$\mathcal{L}_\infty$	$9.5 \times 10^{-5}$	$4.5 \times 10^{-4}$
30 = 60 - 30	$\mathcal{L}_\infty$	$1.9 \times 10^{-5}$	$7.5 \times 10^{-5}$
20 = 60 - 40	$\mathcal{L}_\infty$	$4.5 \times 10^{-6}$	$1.7 \times 10^{-5}$
10 = 60 - 50	$\mathcal{L}_\infty$	$9.9 \times 10^{-7}$	$4.3 \times 10^{-6}$

Table 16. Comparison between wave-function  $\mathcal{L}_\infty$  -Error for the sector one excited state, generated by standard variational calculation of the sector one and variational SUSY calculation for sector two ground state, followed by application of the SUSY Charge Operator for different number of basis ( $n = N_{u_1}, N_{u_2}$ ). ( $N_{ref} = (N_{u_1} = 60, N_{u_2} = 60)$ ). Exact sector one and sector two Hamiltonians are used.

$\Delta N = N_{ref} - n$	Error	$\psi_{(1)}^{(1)} = \vec{Q}_1^+ \vec{\psi}_{(0)}^{(2)}$	$\psi_{(1)}^{(1)}$
40 = 60 - 20	$\mathcal{L}_2$	$1.5 \times 10^{-6}$	$8.9 \times 10^{-6}$
30 = 60 - 30	$\mathcal{L}_2$	$5.8 \times 10^{-8}$	$2.7 \times 10^{-7}$
20 = 60 - 40	$\mathcal{L}_2$	$3.2 \times 10^{-9}$	$1.5 \times 10^{-8}$
10 = 60 - 50	$\mathcal{L}_2$	$1.8 \times 10^{-10}$	$1.1 \times 10^{-9}$

Table 17. Comparison between wave-function  $\mathcal{L}_2$  -Error for the sector one excited state, generated by standard variational calculation of the sector one and variational SUSY calculation for sector two ground state, followed by application of the SUSY Charge Operator for different number of basis ( $n = N_{u_1}, N_{u_2}$ ). ( $N_{ref} = (N_{u_1} = 60, N_{u_2} = 60)$ ). Exact sector one and sector two Hamiltonians are used.

$n$	$E_1^{(1)}$	$E_0^{(2)}$	$\psi_{(0),cutoff}^{(1)}$ = $10^{-10}$	$\psi_{(0),cutoff}^{(1)}$ = $10^{-5}$	$\psi_{(0),cutoff}^{(1)}$ = $10^{-3}$	$\psi_{(0),cutoff}^{(1)}$ = $10^{-2}$
10, 10	4.80	4.752	4.794	4.791	4.78	4.756
20, 20	4.7532	4.75181	4.75317	4.75313	4.7526	4.75181
30, 30	4.752	4.751808	4.75187	4.75187	4.75186	4.751808

Table 18. Comparison among 1<sup>st</sup> excited state energy of sector one (calculated using analytical  $\vec{W}_1$ ), ground state energy of sector two (calculated using analytical  $\vec{W}_1$ ) and different sector two ground state energy that we obtained using  $\vec{W}_1^{approx}$  with different degrees of approximation for  $\psi_{0,approx}^{(1)}$  for different basis size ( $n = N_{u_1}, N_{u_2}$ ).

case ( Figures 3(a-b) and (c-d)! This suggests to us that the non-degenerate and degenerate cases are very similar so far as the wave functions are concerned. Again, in all cases the large component is nodeless and the small component has nodes. In Figures 7(a-b) we show the first ( $\psi_1^{(1)}$ ) and second excited state ( $\psi_2^{(1)}$ ), of the sector one that we have obtained after applying the SUSY charge operator to the sector two ground ( $\vec{\psi}_0^{(2)}$ ) and first excited state ( $\vec{\psi}_1^{(2)}$ ) and the Figures 7(c-d) presents the same states that were found variationally from the sector one Hamiltonian. These results also reflect the similarity between degenerate and non-degenerate case.

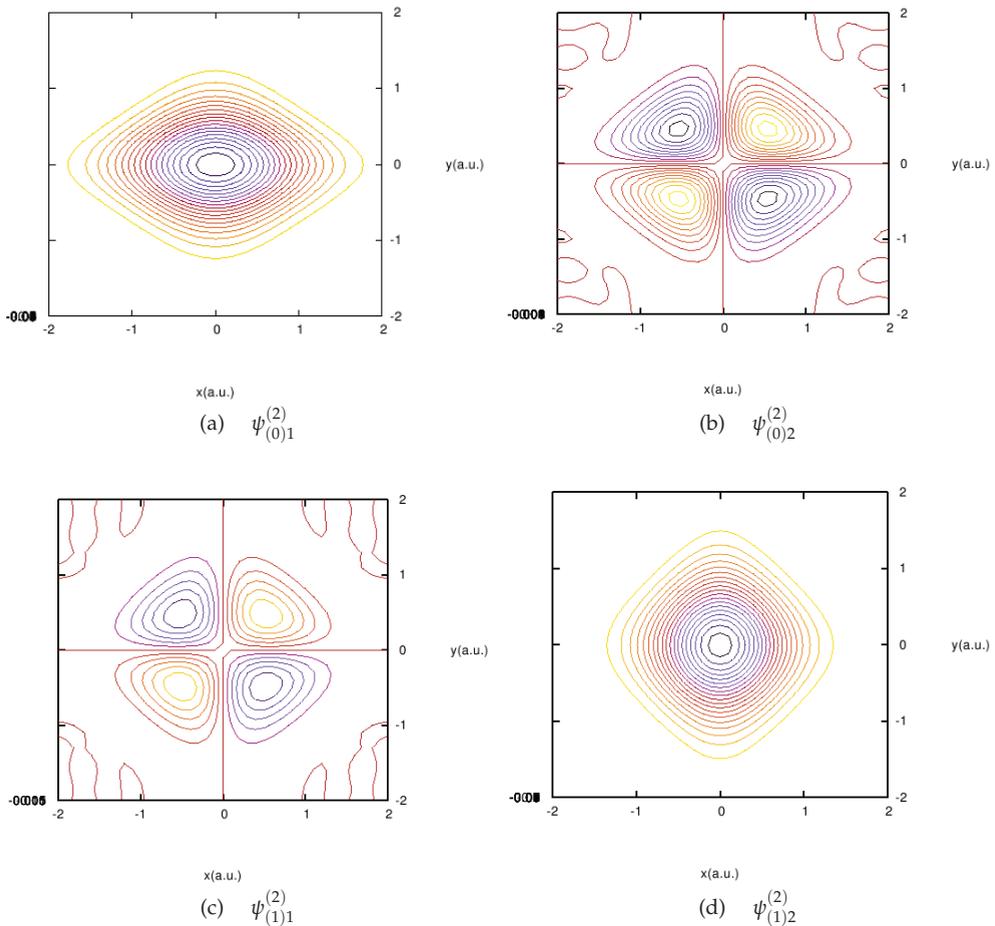


Fig. 6. (a-b) display the two components of the sector two ground state. (c-d) represent the two components for the sector two first excited state. Contour shading is such that red indicates positive amplitude and blue indicates negative amplitude.

For comparison, we display in Figures 8(a-b) the sum of  $\vec{\psi}_0^{(2)}$  and  $\vec{\psi}_1^{(2)}$  and in Figures 8(c-d) the difference of  $\vec{\psi}_0^{(2)}$  and  $\vec{\psi}_1^{(2)}$ . The results are qualitatively the same as those in the two dimensional separable and two dimensional non-separable degenerate cases. That is, both linear combinations are nodeless and of definite sign.

The two dimensional non-separable, non-degenerate case is interesting in that it appears that there is a similar relationship between  $\vec{\psi}_0^{(2)}$  and  $\vec{\psi}_1^{(2)}$  to that which was seen for the degenerate states  $\vec{\psi}_{(0)}^{(2)}$  and  $\vec{\psi}_{(0')}^{(2)}$ . That is, one component is nodeless and large and the second component has nodes and is smaller in magnitude. As in the degenerate case, sums and differences yield states with both components being nodeless. In this case, however, the non degenerate character of the states precludes simply using two different nodeless, orthogonal  $\vec{\phi}_{trial}^{(2)}$  states

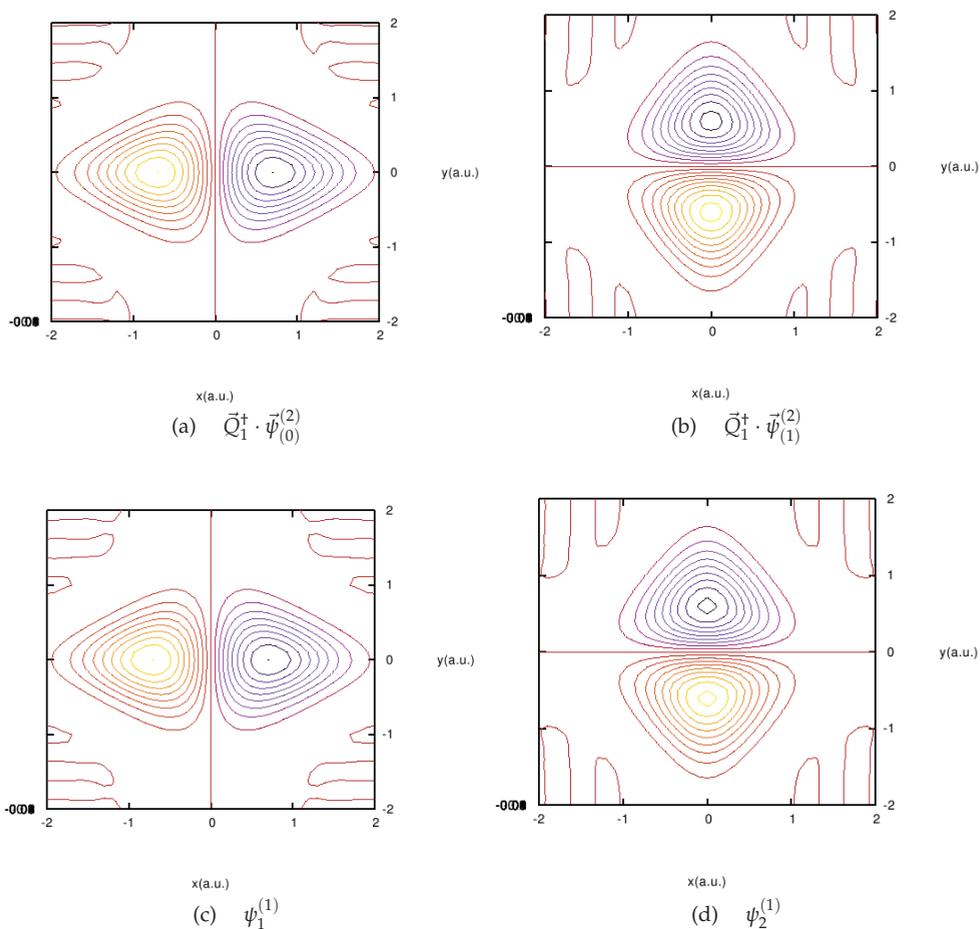


Fig. 7. (a-b) show the 1<sup>st</sup> and 2<sup>nd</sup> excited states of the sector one. The two states are generated by the SUSY-variational method. (c-d) the corresponding states of the sector one Hamiltonian, which are generated variationally. The exact  $\vec{W}_1$  is used.

for the quantum Monte Carlo method. Thus, the implementation of the quantum Monte Carlo method in the non-degenerate case appears to require further consideration. We are currently exploring this aspect of our multi-dimensional SUSY approach.

It is important to stress that our basic strategy is to use *only* the ground state results of the higher SUSY sectors. We believe that this will allow us to obtain the best results for both excited state energies and wave functions of the sector one Hamiltonian, while requiring the least computational effort.

Our upcoming computational studies will be to apply the present approach to more interesting, non-separable higher dimensional systems such as rare-gas atomic clusters where the structure and thermodynamics seem to require a fully quantum many-body treatment.

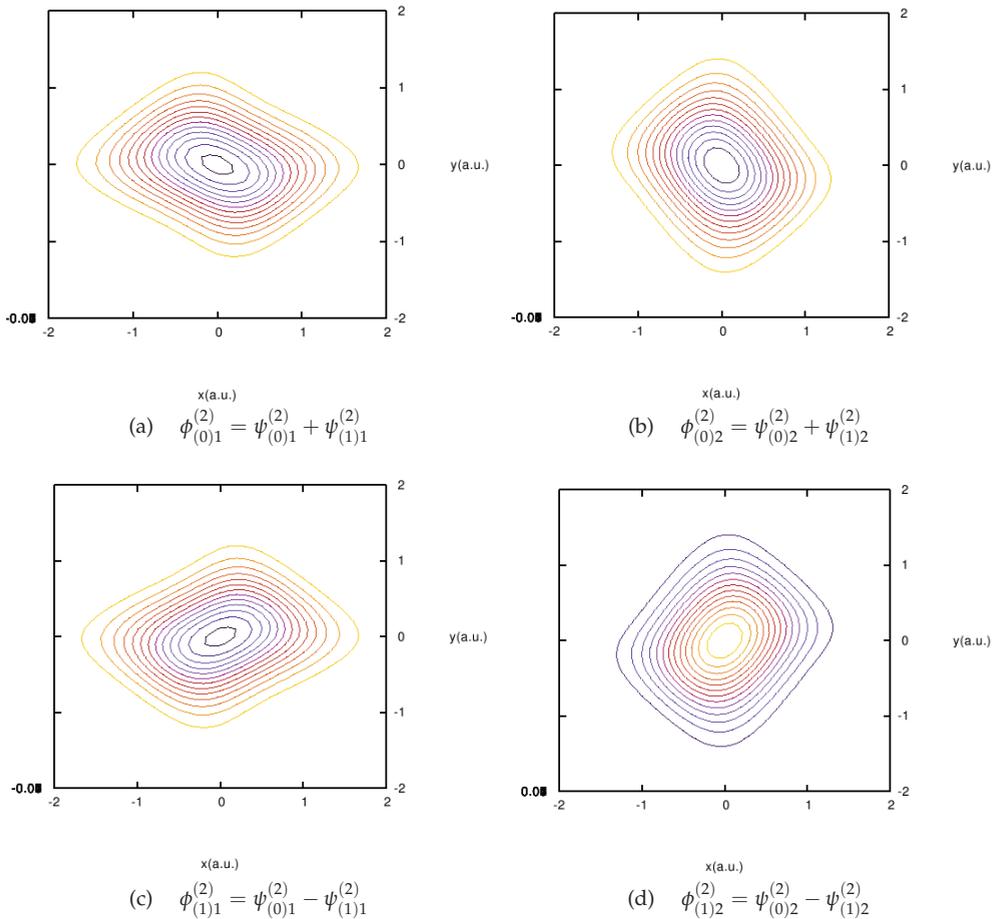


Fig. 8. Components of linear combinations of non-degenerate ground and first excited state.

Chakravarty (1995a;b); Derrickson & Bittner (2006; 2007); Franke et al. (1993); Lynden-Bell & Wales (1994); Rick et al. (1991); Schmidt et al. (2001); Wales & Doye (1997) For systems composed of a single type atom or molecule, we expect to encounter degeneracies. Thus, we expect the situation to mirror the present two dimensional non-separable degenerate case.

In dealing with such systems, we anticipate that as the number of particles is increased, we will find that a Monte Carlo based approach may be preferred.

Finally, we stress that this is the first formulation of a general SUSY approach for multi-dimensional and/or multi-particle systems. There remain many formal and computational questions, which we are continuing to explore. Our main conclusion is that there is sufficient promise that such studies are justified.

## 6. Electronic structure of atoms: Hydrogen and helium

In the previous sections, we have provided a generalization of SUSY-QM to treat any number of dimensions or particles with a focus on its usefulness as a computational tool for calculating accurate excited state energies and wave functions.

Because of the significant analytical and computational ramifications, and motivated by the future study of more complex electronic systems, we here apply our multi-dimensional generalization of SUSY-QM to the hydrogen atom in full three-dimensional detail. This is of interest because, until now, the standard application of SUSY-QM to the hydrogen atom required that we first separate out the angular degrees of freedom – effectively reducing the problem to a one-dimensional treatment [Kirchberg et al. (2003); Lahiri et al. (1987); Tangerman & Tjon (1993)]. With our vector superpotential approach, one can deal with the full three-dimensional nature of the hydrogen atom.

Our approach provides, for the first time, a SUSY-QM framework that can be employed to treat non-hydrogenic atoms. For example, the standard SUSY-QM treatment of the hydrogen atom cannot be extended readily to the helium atom because it is impossible reduce it to a one-dimensional system. In addition, the form of the three-dimensional vector superpotential for the hydrogen atom is of interest in its own right. It is quite different from the radial superpotential obtained in earlier SUSY-QM studies of the hydrogen atom. The present study thus lays the groundwork for a systematic SUSY-QM study of excited state energies and wave functions of atoms.

### 6.1 SUSY-QM for the three-dimensional hydrogen atom

We now consider the hydrogen atom. We begin by noting that the ground state is exactly given by

$$\psi_{1,0,0} = \frac{e^{-r}}{\sqrt{\pi}}, \quad (6.1)$$

where we have set the Bohr radius equal to 1. The Hamiltonian (in atomic units) is simply

$$\mathcal{H} = -\frac{1}{2}\nabla^2 - \frac{1}{r}. \quad (6.2)$$

Then, the vector superpotential is given by

$$\vec{W} = -\nabla \ln \psi_{1,0,0} = \hat{r}, \quad (6.3)$$

where  $\hat{r}$  is a unit vector in the direction of  $\vec{r}$ . This is an extremely interesting result. First, we see that the superpotential for the Coulomb interaction is, itself, non-singular. Second, in the standard approach, because the angular degrees of freedom have already been separated out, the superpotential is a scalar and it depends on the angular momentum squared (*i.e.* on  $l(l+1)$ ). The precise form for the ground state ( $l=0$ ) is

$$W_{radial} = 1 \quad (6.4)$$

In three dimensions, we have

$$\vec{W} = \vec{\epsilon}_x \frac{x}{r} + \vec{\epsilon}_y \frac{y}{r} + \vec{\epsilon}_z \frac{z}{r} = \hat{r}. \quad (6.5)$$

The magnitude of  $\vec{W}$  is equal to the radial superpotential, as one expects, but the individual components are radically different. Note that these components can also be written solely

in terms of angular functions (the direction cosines of  $\vec{r}$ ). To obtain the atomic potential for hydrogen, we form

$$\vec{W} \cdot \vec{W} - \nabla \cdot \vec{W} = 1 - \left( \frac{3}{r} - \frac{x^2 + y^2 + z^2}{r^3} \right) \quad (6.6)$$

$$= 1 - \frac{2}{r} = -2E_0 - \frac{2}{r}. \quad (6.7)$$

Now we recall that

$$H\psi_{m_l} = E_n\psi_{m_l} \quad (6.8)$$

and

$$-\frac{1}{2}\nabla^2\psi_{m_l} = \left[ E_n + \frac{1}{r} \right] \psi_{m_l} \quad (6.9)$$

yields

$$\nabla^2\psi_{1,0,0} = - \left[ 2E_0 + \frac{2}{r} \right] \psi_{1,0,0}. \quad (6.10)$$

Since the ground state energy of hydrogen in atomic units is  $-1/2$ , we find that Equations (6.7) and (6.10) are consistent and we have obtained the correct vector superpotential. Of great interest is the wave equation for the sector two problem. This Hamiltonian is given by

$$\overleftarrow{\mathcal{H}}_2 = -\frac{1}{2}\nabla\nabla + \frac{1}{2} [\vec{W}\vec{W} + \nabla\vec{W}]. \quad (6.11)$$

In the case of the hydrogen atom, because we have exact analytical expressions for the excited states of  $\mathcal{H}_1$ , it is a simple matter to generate analytical expressions for all the states of the sector two Hamiltonian. It is convenient to label the sector two states with an index indicating the  $n^{\text{th}}$  energy state (i.e., we use the principle quantum number  $n = 1, 2, \dots$ ) along with the quantum numbers of the sector one excited state from which they are obtained. Thus, the four degenerate ground states of  $\overleftarrow{\mathcal{H}}_2$  will be denoted by  $\overleftarrow{\psi}_{1,2p_x}^{(2)}$ ,  $\overleftarrow{\psi}_{1,2p_y}^{(2)}$ ,  $\overleftarrow{\psi}_{1,2p_z}^{(2)}$ ,  $\overleftarrow{\psi}_{1,2s}^{(2)}$ . We choose here to use the real states rather than those labeled by  $m_l = \pm 1$  and  $m_l = 0$  values. We find that these solutions are given by

$$\overleftarrow{\psi}_{1,2p_x}^{(2)} = N \left[ \hat{i}e^{-r/2} + \frac{x\hat{r}}{2}e^{-r/2} \right], \quad (6.12)$$

$$\overleftarrow{\psi}_{1,2p_y}^{(2)} = N \left[ \hat{j}e^{-r/2} + \frac{y\hat{r}}{2}e^{-r/2} \right], \quad (6.13)$$

$$\overleftarrow{\psi}_{1,2p_z}^{(2)} = N \left[ \hat{k}e^{-r/2} + \frac{z\hat{r}}{2}e^{-r/2} \right], \quad (6.14)$$

$$\overleftarrow{\psi}_{1,2s}^{(2)} = -N\frac{\vec{r}}{2}e^{-r/2}. \quad (6.15)$$

These equations can be verified by simply applying  $\vec{Q}$  to the first excited state wave functions of sector one. It is also easily verified that  $\vec{Q}^\dagger$  acting on these states regenerates the  $\psi_{2p}^{(1)}$  and  $\psi_{2s}^{(1)}$  states. Furthermore, in Figures 9 and 10, we provide plots of the  $\overleftarrow{\psi}_{1,2s}^{(2)}$  and  $\overleftarrow{\psi}_{1,2p_x}^{(2)}$ . It is straight forward to see that  $\overleftarrow{\psi}_{1,2p_y}^{(2)}$  and  $\overleftarrow{\psi}_{1,2p_z}^{(2)}$  are both similar to  $\overleftarrow{\psi}_{1,2p_x}^{(2)}$ .

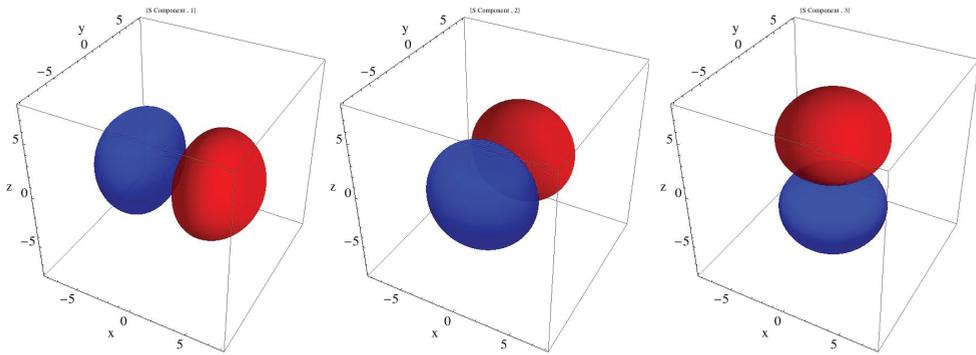


Fig. 9. The three components of the wave function for  $\bar{\psi}_{1,S}^{(2)}$ . Here, blue corresponds to positive values and red to negative.

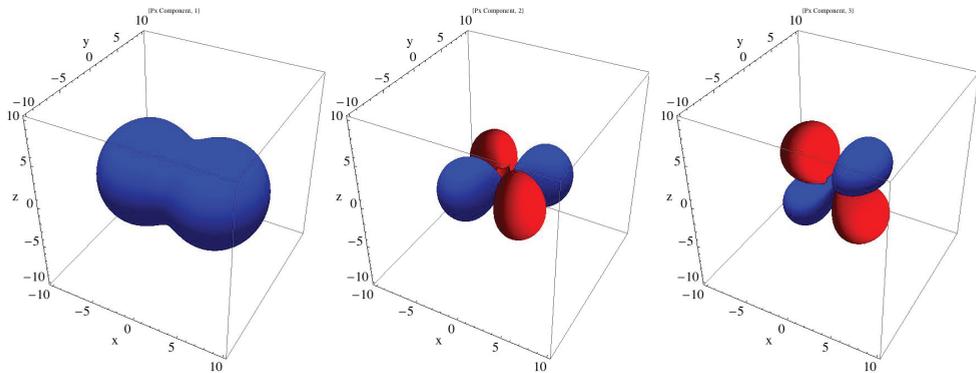


Fig. 10. The three components of the wave function for  $\bar{\psi}_{1,p_x}^{(2)}$ . Here, blue corresponds to positive values and red to negative.

**6.2 An approximate superpotential for the helium atom**

It is of interest to begin exploring how our approach to multidimensional SUSY-QM would deal with a two electron atom. It is clear that the usual radial (one-dimensional) hydrogen atom SUSY-QM treatment is not readily generalizable to deal with helium. We have carried out a Quantum Monte Carlo study of the sector one ground state of helium using the Padé Jastrow trial wave function Umrigar & Wilson (1988):

$$\psi_{T,\alpha}^{(1)} = e^{-2r_1} e^{-2r_2} e^{\frac{r_{12}}{2(1+\alpha r_{12})}}, \tag{6.16}$$

with the optimum  $\alpha$  given by  $\alpha = 0.353$ . This yields an energy of  $E_1^{(1)} \approx 2.878$ , which is in error by about 1%. This error is reasonable for a simple treatment neglecting relativistic interactions. The approximate  $\vec{W}$  is generated from

$$\vec{W}(\vec{r}_1, \vec{r}_2) = -\vec{\nabla} \ln \psi_{T,\alpha}^{(1)} \tag{6.17}$$

$$= -\vec{\nabla} \left[ -2r_1 - 2r_2 + \frac{r_{12}}{2(1 + \alpha r_{12})} \right]. \quad (6.18)$$

Here,

$$\vec{\nabla} = \epsilon_{1x} \frac{\partial}{\partial x_1} + \epsilon_{1y} \frac{\partial}{\partial y_1} + \epsilon_{1z} \frac{\partial}{\partial z_1} + \epsilon_{2x} \frac{\partial}{\partial x_2} + \epsilon_{2y} \frac{\partial}{\partial y_2} + \epsilon_{2z} \frac{\partial}{\partial z_2}, \quad (6.19)$$

where the  $\{\hat{\epsilon}_{ij}\}$  are orthonormal vectors. The resulting vector superpotential for the Padè-Jastrow trial function is readily found to be

$$\vec{W}_{He(PJ)} = 2\hat{r}_1 + 2\hat{r}_2 - \hat{r}_{12} \left[ 1 - \frac{\alpha}{(1 + \alpha r_{12})} \right]. \quad (6.20)$$

Thus, the structure of  $\vec{W}_{He(PJ)}$  is analogous to  $\vec{W}_H$  in that Coulomb interactions generate vector superpotentials that involve unit vectors anti-parallel to the direction of the forces. This is true in general for Coulombic interactions. This emphasizes the important distinction between our three-dimensional SUSY-QM treatment of an atom and the standard hydrogen atom one-dimensional radial SUSY-QM.

## 7. Aufbau approach for excited states

For multielectron atoms, it becomes necessary to consider how the aufbau principle acts in the second sector to permit efficient calculations of sector one excited states. This is because we can use this principle to design reasonable trial wave functions for a variational approach to the sector two ground state. In this section, we consider a simple aufbau description of the sector one helium excited states in order to design an approximate sector two ground state of helium. We assume that in the first excited state of sector one, we have one electron in the 1S orbital, given by wave function  $\alpha$ , and one electron in the 2S orbital, given by wave function  $\beta$  where

$$\alpha(r) = \frac{e^{-2r}}{\sqrt{\pi}} \quad \text{and} \quad \beta(r) = \frac{e^{-r}}{4\sqrt{2\pi}}(1-r). \quad (7.1)$$

Then, it is of interest to take the product of these states such that we have  $\alpha(r_1)\beta(r_2)$ , to which we can apply our  $\vec{Q}$  to find

$$\vec{Q}(\alpha(r_1)\beta(r_2)) = -e^{-2r_1-r_2} [2\hat{r}_1(1-r_2) + \hat{r}_2] \equiv \vec{\phi}_1^{(2)}. \quad (7.2)$$

The first excited sector one state of Helium is a triplet so the wave function is anti-symmetric under spatial electron exchange. It is clear that we require the second sector ground state also be anti-symmetric when we interchange labels 1 and 2. To obtain this, we apply  $P_{12}$  to Equation (7.2) to get

$$\vec{Q}(\alpha(r_2)\beta(r_1)) = -e^{-2r_2-r_1} [2\hat{r}_2(1-r_1) + \hat{r}_1] \equiv P_{12}\vec{\phi}_1^{(2)}, \quad (7.3)$$

where  $P_{12}$  exchanges the electron labels. Then, we can use Equations (7.2) and (7.3) as a ‘‘building blocks’’ to construct our ground state in the second sector by subtracting the first building block from the second. This gives us a second sector result of

$$\begin{aligned} \vec{\psi}_{1,triplet}^{(2)} &= -e^{-2r_1-r_2} [2\hat{r}_1(1-r_2) + \hat{r}_2] + \\ &\quad e^{-2r_2-r_1} [2\hat{r}_2(1-r_1) + \hat{r}_1]. \end{aligned} \quad (7.4)$$

And similarly, we can find the second sector first excited singlet state by simply adding the two building blocks as given below.

$$\vec{\psi}_{1,singlet}^{(2)} = -e^{-2r_1-r_2} [2\hat{r}_1(1-r_2) + \hat{r}_2] - e^{-2r_2-r_1} [2\hat{r}_2(1-r_1) + \hat{r}_1]. \quad (7.5)$$

Indeed, by taking the scalar product with  $\vec{Q}^\dagger$ , we can verify that  $\vec{\psi}_{1,triplet}^{(2)}$  and  $\vec{\psi}_{1,singlet}^{(2)}$  give the appropriate spatial wave functions. This is to say that, to within a multiplicative constant, we get that

$$\vec{Q}^\dagger \cdot \vec{\psi}_{1,triplet}^{(2)} = \psi_1^{(1)} = \alpha(r_1)\beta(r_2) - \alpha(r_2)\beta(r_1) \quad (7.6)$$

and

$$\vec{Q}^\dagger \cdot \vec{\psi}_{1,singlet}^{(2)} = \psi_2^{(1)} = \alpha(r_1)\beta(r_2) + \alpha(r_2)\beta(r_1). \quad (7.7)$$

From this, we observe that the aufbau principle in the second sector is remarkably simple for Helium. We merely need to take the building block  $\vec{\phi}_1^{(2)}$  and antisymmetrize or symmetrize appropriately. Of course our "building block",  $\phi_1^{(2)}$ , is neither symmetric nor antisymmetric under particle exchange.

However, this basis doesn't include the correlation. To do this, we can multiply our antisymmetrized second sector wave function by a correlation function, given by the Padé-Jastrow function which only depends on  $r_{12}$ . It is clear, then, that because our correlation function is only a function of  $r_{12}$ , its symmetry will not be affected by the application of  $\vec{Q}$  and, thus, we can simply multiply it by our second sector state of interest (where the minus corresponds to the triplet and the plus to the singlet):

$$\vec{\psi}_{1,triplet}^{(2)} = e^{\frac{r_{12}}{2(1+\delta r_{12})}} \left( -e^{-2r_1-r_2} [2\hat{r}_1(1-r_2) + \hat{r}_2] \mp e^{-2r_2-r_1} [2\hat{r}_2(1-r_1) + \hat{r}_1] \right). \quad (7.8)$$

Higher accuracy will result if we insert additional variational parameters (*e.g.*, effective charges, *etc.*) when doing computations.

## 8. Conclusions and perspectives

In this chapter, we began by presenting our computational approach to one dimensional systems. We showed with the anharmonic oscillator that we were able to achieve significant computational gains in a robust fashion, permitting more exact numerical solutions of one dimensional problems. Although anharmonic oscillator models are useful for a wide variety of problems in both chemistry and physics, it should be clear that other systems should show similar behaviour. The SUSY-QM approach enabled us to develop a hierarchy of isospectral Hamiltonians. This also led to the introduction of charge operators that transform wave functions between the various sectors, and the energies are always determined in a ground state setting. Because these are most easily and accurately obtained by the variational method, we realize a significant reduction in the basis size needed to yield accurate excited state wave functions. We then considered 2 specific examples of anharmonic oscillators. We concluded that using the SUSY hierarchy of hamiltonians and charge operators, provided faster convergence to the same level of accuracy and thus, provides a better method than the

standard variational approach. In most cases, only half as many basis functions were needed to generate the ground state wave function as were required for the first excited state of the same sector. As a result, the computational time for molecular models using anharmonic potentials will be significantly reduced, without sacrificing accuracy.

We also stress that our results do not depend on precisely what basis set is used for the calculations. Rather we are capitalizing on the general behavior of the Rayleigh-Ritz variational method with regard to accuracy and convergence rate for ground versus excited states of a given Hamiltonian.

We then presented our approach to generalizing SUSY-QM to deal with more than one dimension and more than one (distinguishable) particle. In general, previous attempts to do this have typically introduced Pauli spin matrices and so far as we are aware, none of these has been proved useful for the general case. Andrianov et al. (1985); Andrianov, Borisov & Ioffe (1984a;b;c); Andrianov et al. (1986); Andrianov, Borisov, Ioffe & Eides (1984); Andrianov & Ioffe (1988); Andrianov et al. (2002); Cannata et al. (2002); Das & Pernice (1996) One principle difficulty is that while the coordinates of different particles are independent variables, they are not defined relative to orthogonal axes. That is, there are only 3 independent, physical axes along which all particle positions are measured. Our approach introduces a higher dimensional vector space in which there is an orthonormal basis vector associated with each independent particle coordinate. This is analogous to the relativistic situation where each particle has its own coordinate system (and, of course, in the relativistic case, its own "proper time"). Here, however the device is a mathematical convenience (so far as we are currently aware) and it is, of course, non-relativistic. That is, we assume Galilean transformations. The most striking consequence similar to relativistic quantum mechanics is that our second sector Hamiltonian becomes a tensor in the expanded space. This does *not* increase the number of independent variables (i.e., the wave function for the second sector is a vector in the new hyperspace). It is shown that this tensor character is then absent from the 3<sup>rd</sup> sector Hamiltonian (which is once again a scalar operator). One in general obtains an alternating series of scalar and tensor Hamiltonians. *The occurrence of a tensor sector Hamiltonian is, of course, an added computational cost to the approach.* This is mitigated, to some degree, by the fact that we *never* must calculate an accurate wave function and energy except for *ground states*. It is this feature that makes the SUSY-QM approach attractive, since ground state energies and wave functions are the least computationally demanding of all and typically are obtained with the highest accuracy. Thus the computational effort of obtaining the second excited state energy and wave function again involves solving an equation comparable to that generated by the original  $H_1$ . A complication, however, arises due to the observed fact that for the ground state sector 2 wave function in both the degenerate and non-degenerate cases, one of the two components possesses nodes while the other component is nodeless. This is mitigated (in terms of accuracy of the sector 2 ground state calculation) by the fact that the component containing the node is (in the present computational examples) an order of magnitude smaller than the nodeless component. This appears to enable the variational evaluation of the ground state of the sector 2 to yield better accuracy for the first excited state energy and wave function than a comparable basis set calculation applied to  $H_1$ .

An extremely important question is, however, raised by the fact that the small component of  $\vec{\psi}_0^{(2)}$  has nodes. This is whether the presence of nodes will prevent us from applying a simple variational quantum Monte-Carlo method to obtain  $\vec{\psi}_0^{(2)}$ . We are currently exploring this question. However, in the present context, it does not appear to create difficulties for the variational approach.

Finally, we have shown how our multi-dimensional generalization of SUSY-QM can be applied to the hydrogen atom. Previously, most detailed attempts to treat the hydrogen atom first separated the angular degrees of freedom, leaving a one-dimensional radial wave equation. It was then possible to obtain the SUSY-QM factorization, yielding a scalar superpotential that, for the  $l = 0$  states, is simply  $W = 1$ . While these results are interesting, the one-dimensional radial SUSY-QM approach is not readily generalizable to treat even the helium atom.

In our approach, the full three-dimensional character of the hydrogen atom is considered, with the result being a vector-valued superpotential,  $\vec{W}$ , which for the hydrogen atom, is  $\vec{W} = \hat{r}$ . That is, the vector superpotential points in the opposite direction of the attractive Coulomb force between the electron and the nucleus. This is interesting also because, although the Coulomb potential is singular, its vector superpotential is not. It is important to note that such a superpotential was also obtained earlier by Stedman Stedman (1985). However, his sector two Hamiltonian differs from ours and produces “extra” states that are not degenerate with sector one.

The fact that  $\vec{W}$  for the three-dimensional hydrogen atom is a vector does not, in any way, modify the sector one dynamical equation. However, the sector two situation is radically affected! In the one-dimensional SUSY-QM case, there is no significant change in the basic mathematical structure of the sector two partner Hamiltonian. In the multi-dimensional case, the sector two Hamiltonian is a tensor. However, we have shown in previous studies, that many of the standard computational techniques remain valid. Of particular interest is the Dirac-Frankel-McLachlan Variational Method, since this is known to deal better with higher-dimensional systems Raab (2000).

In the case of the hydrogen atom, it is straight forward to generate all the sector two eigenstates. This is a consequence of the fact that exact analytical eigenstates of the three-dimensional hydrogen atom are known. It is then easy to apply the charge operator,  $\vec{Q}$ , to the excited hydrogen atom states and obtain sector two eigenstates. (We note that because of the four fold degeneracy for the sector two ground state, the resulting eigenstates can be super-posed in any manner convenient for the study at hand). It is of considerable interest to begin exploring how our multi-dimensional SUSY-QM treatment can be applied to the helium atom. In this case, the exact sector one ground state is, of course, unavailable. In our previous one and two-dimensional studies we have considered other systems for which an exact  $\vec{W}$  was not possible. In the case of helium, we chose to examine an accurate Padè-Jastrow approximation to the sector one ground state. In this case, it is easy to obtain an analytical (albeit approximate)  $\vec{W}$  that displays very reasonable intuitive character. In direct analogy with the exact hydrogen atom  $\vec{W}$ , we find that the  $\vec{W}_{He(PJ)}$  vector superpotential consists of a combination of unit vectors that again, are anti-parallel to the Coulomb forces associated with the helium atom potential energy. The next step in our study will consist of computations of a sector two ground state, which will allow us to obtain an approximate helium atom sector one first excited state energy and wave function.

Future studies will explore extending the approach to more than two electron atoms. There, the issue will be taking account of the electrons’ spin degrees of freedom. Our current plan is to employ the “spin-free” techniques of Matsen Matsen (1964; 1966; 1970); Matsen & Cantu (1968; 1969); Matsen et al. (1966); Matsen & Ellzey (1969); Matsen & Junker (1971); Matsen & Klein (1969; 1971); Matsen et al. (1971) and others Pauncz (1995).

We have also generalized the aufbau principle to work in the second sector Hamiltonian, demonstrating that we are able to produce reasonable forms of excited states by simply using

hydrogenic orbitals. The equations have a reasonable structure but variational computations are necessary. We shall report these results later.

## 9. References

- Andersson, L. (2001). Quantum dynamics using a discretized coherent state representation: An adaptive phase space method, *J. Chem. Phys.* 115: 1158–1165.
- Andrianov, A. A., Borisov, N. V., Eides, M. I. & Ioffe, M. V. (1985). Supersymmetric Origin Of Equivalent Quantum Systems, *Phys. Lett. A* 109: 143–148.
- Andrianov, A. A., Borisov, N. V. & Ioffe, M. V. (1984a). *JETP Lett.* 39: 93–97.
- Andrianov, A. A., Borisov, N. V. & Ioffe, M. V. (1984b). Factorization method and Darboux transformation for multidimensional Hamiltonians, *Theoretical and Mathematical Physics* 61: 1078–1089.
- Andrianov, A. A., Borisov, N. V. & Ioffe, M. V. (1984c). The Factorization method and quantum systems with equivalent energy spectra, *Phys. Lett. A* 105: 19–22.
- Andrianov, A. A., Borisov, N. V. & Ioffe, M. V. (1986). Nonstationary Approach To Scattering Theory For Supersymmetric Hamiltonians In Quantum Mechanics And Supersymmetry Of Nuclear Interactions, *Phys. Lett. B* 181: 141.
- Andrianov, A. A., Borisov, N. V., Ioffe, M. V. & Eides, M. I. (1984). Supersymmetric Mechanics: A New Look At The Equivalence Of Quantum Systems, *Theoretical and Mathematical Physics* 61: 965–972.
- Andrianov, A. A. & Ioffe, M. V. (1988). Pauli Fermions As Components Of N=2 Supersymmetrical Quantum Mechanics, *Phys. Lett. B* 205: 507–510.
- Andrianov, A. A., Ioffe, M. V. & Nishnianidze, D. N. (2002). Polynomial SUSY in quantum mechanics and second derivative Darboux transformations, *Phys. Lett. A* 201: 103–110.  
URL: <http://www.iop.org/EJ/ref/-target=inspec/0305-4470/35/6/305/20>
- Cannata, F., Ioffe, M. V. & Nishnianidze, D. N. (2002). New methods for the two-dimensional Schrödinger equation: SUSY-separation of variables and shape invariance, *Journal of Physics A: Mathematical and General* 35(6): 1389–1404.  
URL: <http://stacks.iop.org/0305-4470/35/1389>
- Chakravarty, C. (1995a). *The Journal of Chemical Physics* 102: 956–962.
- Chakravarty, C. (1995b). *The Journal of Chemical Physics* 103: 10663–10668.
- Das, A. & Pernice, S. (1996). *arXiv:hep-th/9612125v1* 1: 1.
- Derrickson, S. & Bittner, E. (2006). *Journal Of Physical Chemistry A* 110: 5333–5341.
- Derrickson, S. W. & Bittner, E. R. (2007). *Journal Of Physical Chemistry A* 111: 10345–10352.
- Dirac, P. (1958). *Lectures on Quantum Mechanics*, Oxford Press, Oxford University Press, Ely House, London W.1.
- Franke, G., Hilf, E. R. & Borrmann, P. (1993). *The Journal of Chemical Physics* 98: 3496.
- Griffiths, D. (1987). *Introduction to Elementary Particles*, John Wiley & Sons, New York, USA.
- Hylleraas, E. & Undheim, B. (1930). *Z. Physik* 65: 759.
- Jafarpour, M. & Afshar, D. (2002). *J. Phys. A* 35: 87–92.
- Junker, G. (1996). *Supersymmetric Methods in Quantum and Statistical Mechanics*, Springer-Verlag, New York.
- Kirchberg, A., Länge, J., Pisani, P. & Wipf, A. (2003). Algebraic solution of the supersymmetric hydrogen atom in d dimensions, *Annals of Physics* 303(2): 359–388.
- Klauder, J. & Skagerstam, B. (1985). *Coherent States*, World Scientific Publishing Co Pte Ltd.

- Kouri, D. J., Maji, K., Markovich, T. & Bittner, E. R. (2010a). New Generalization of Supersymmetric Quantum Mechanics to Arbitrary Dimensionality or Number of Distinguishable Particles, *The Journal of Physical Chemistry A* 114(32): 8202–8216.  
URL: <http://pubs.acs.org/doi/abs/10.1021/jp103309p>
- Kouri, D. J., Maji, K., Markovich, T. & Bittner, E. R. (2010b). New Generalization of Supersymmetric Quantum Mechanics to Arbitrary Dimensionality or Number of Distinguishable Particles, *Journal Of Physical Chemistry A* 114(32): 8202–8216.
- Kouri, D. J., Markovich, T., Maxwell, N. & Bittner, E. R. (2009). Supersymmetric Quantum Mechanics, Excited State Energies and Wave Functions, and the Rayleigh-Ritz Variational Principle: A Proof of Principle Study, *Journal Of Physical Chemistry A* 113(52): 15257–15264.
- Kouri, D. J., Markovich, T., Maxwell, N. & Bodman, B. G. (2009). The Heisenberg-Weyl Algebra on the Circle and a Related Quantum Mechanical Model for Hindered Rotation, *Journal Of Physical Chemistry A* 113: 7698–7705.
- Kouri, D. J., Markovich, T., Maxwell, N. & Bodmann, B. G. (2009). The Heisenberg-Weyl Algebra on the Circle and a Related Quantum Mechanical Model for Hindered Rotation, *Journal Of Physical Chemistry A* 113(26): 7698–7705.
- Kouri, D., Papadakis, M., Kakadiaris, I. & Hoffman, D. (2003). Properties of minimum uncertainty wavelets and their relations to the harmonic oscillator and the coherent states, *J. Phys. Chem. A* 107: 7318–7327.
- Lahiri, A., Roy, P. & Bagchi, B. (1987). Supersymmetry in atomic physics and the radial problem, *Journal of Physics A: Mathematical and General* 20: 3825.
- Littlejohn, R. G., C. M. C. J. T. M. K. A. P. B. (2002). A general framework for discrete variable representation basis sets, *J. Chem. Phys.* 116: 8691–8703.
- Lynden-Bell, R. M. & Wales, D. J. (1994). *The Journal of Chemical Physics* 101: 1460–1476.
- Matsen, F. A. (1964). Spin-Free Quantum Chemistry. II. Three-Electron Systems, *The Journal of Physical Chemistry* 68(11): 3282–3296.
- Matsen, F. A. (1966). Spin-Free Quantum Chemistry. IV. The pn Electron Configuration, *The Journal of Physical Chemistry* 70(5): 1568–1574.
- Matsen, F. A. (1970). Chemistry without spin, *Journal of the American Chemical Society* 92(12): 3525–3538.
- Matsen, F. A. & Cantu, A. A. (1968). Spin-free quantum chemistry. V. Spin density, *The Journal of Physical Chemistry* 72(1): 21–32.
- Matsen, F. A. & Cantu, A. A. (1969). Spin-free quantum chemistry. VII. The Slater determinant, *The Journal of Physical Chemistry* 73(8): 2488–2494.
- Matsen, F. A., Cantu, A. A. & Poshusta, R. D. (1966). Spin-Free Quantum Chemistry. 1a III. Bond Functions and the Pauling Rules, *The Journal of Physical Chemistry* 70(5): 1558–1567.
- Matsen, F. A. & Ellzey, M. L. (1969). Spin-free quantum chemistry. VIII. The crystal field problem, *The Journal of Physical Chemistry* 73(8): 2495–2502.
- Matsen, F. A. & Junker, B. R. (1971). Spin-free quantum chemistry. XI. Perturbation theory for interaction energies, *The Journal of Physical Chemistry* 75(12): 1878–1881.
- Matsen, F. A. & Klein, D. J. (1969). Spin-free quantum chemistry. VI. Spin conservation, *The Journal of Physical Chemistry* 73(8): 2477–2487.
- Matsen, F. A. & Klein, D. J. (1971). Spin-free quantum chemistry. IX. Aggregate theory of polyelectronic systems, *The Journal of Physical Chemistry* 75(12): 1860–1866.

- Matsen, F. A., Klein, D. J. & Foyt, D. C. (1971). Spin-free quantum chemistry. X. Effective spin Hamiltonian, *The Journal of Physical Chemistry* 75(12): 1866–1873.
- Pauncz, P. (1995). *The Symmetric Group in Quantum Chemistry*, CRC-Oress, CRC Press, Inc. 2000 Corporate Blvd., N.W., Boca Raton, Florida 33431.
- Raab, A. (2000). On the Dirac-Frenkel/McLachlan variational principle, *Chemical physics letters* 319(5-6): 674–678.  
URL: <http://www.sciencedirect.com/science/article/B6TFN-3YXJXVM-1B/2/1564c568595cdb01db9cfd5c68425cdb>
- Rick, S. W., Leitner, D. L., Doll, J. D., Freeman, D. L. & Frantz, D. D. (1991). *The Journal of Chemical Physics* 95: 6658–6667.
- Schmidt, M., Kusche, R., Hippler, T., Donges, J., Kronm uller, W., von Issendorff, B. & Haberland, H. (2001). *Physical Review Letters* 86: 1191–1194.
- Stedman, G. (1985). Simple supersymmetry: II. Factorisation method in quantum mechanics, *European Journal of Physics* 6(4): 225–231.  
URL: <http://iopscience.iop.org/0143-0807/6/4/002>
- Tangerman, R. & Tjon, J. (1993). Exact supersymmetry in the nonrelativistic hydrogen atom, *Arxiv preprint hep-th/9304097* .
- Umrigar, C. & Wilson, K. (1988). Optimized trial wave functions for quantum Monte Carlo calculations, *Physical Review Letters* .  
URL: <http://link.aps.org/doi/10.1103/PhysRevLett.60.1719>
- Wales, D. J. & Doye, J. (1997). *Journal Of Physical Chemistry A* 101: 5111–5116.
- Wess, J. & Bagger, J. (1992). *Supersymmetry and supergravity*, Princeton series in physics, Princeton University Press.  
URL: [http://books.google.com/books?id=4QrQZ\\_Rjq4UC](http://books.google.com/books?id=4QrQZ_Rjq4UC)
- Weyl, H. (1950). *The theory of groups and quantum mechanics*, Dover books on advanced mathematics, Dover Publications.  
URL: <http://books.google.com/books?id=jQbEcDDqGb8C>
- Wigner, E. P. & Fano, U. (1960). Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra, 28(4): 408–409.  
URL: <http://dx.doi.org/doi/10.1119/1.1935822>
- Wilson, E., Decius, J. & Cross, P.C. (1955). *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*, McGraw-Hill, New York.

# A Statistical Derivation of Non-Relativistic Quantum Theory

Ulf Klein

*University of Linz, Institute for Theoretical Physics  
Austria*

## 1. Introduction

Quantum theory (QT) may either be defined by a set of axioms or otherwise be 'derived' from classical physics by using certain assumptions. Today, QT is frequently identified with a set of axioms defining a Hilbert space structure. This mathematical structure has been created (by von Neumann) by abstraction from the linear solution space of the central equation of QT, the Schrödinger equation. Thus, deriving Schrödinger's equation is basically the same as deriving QT. To derive the most general version of the time-dependent Schrödinger equation, describing  $N$  particles with spin in an external gauge field, means to derive essentially the whole of non-relativistic QT.

The second way of proceeding is sometimes called 'quantization'. In the standard (canonical) quantization method one starts from a classical Hamiltonian whose basic variables are then 'transformed', by means of well-known correspondence rules,

$$p \rightarrow \frac{\hbar}{i} \frac{d}{dx}, \quad E \rightarrow -\frac{\hbar}{i} \frac{d}{dt}, \quad (1)$$

into operators. Then, all relevant classical observables may be rewritten as operators acting on states of a Hilbert space etc; the details of the 'derivation' of Schrödinger's equation along this lines may be found in many textbooks. There are formal problems with this approach which have been identified many years ago, and can be expressed e.g. in terms of Groenewold's theorem, see Groenewold (1946), Gotay (1999). Even more seriously, there is no satisfactory *explanation* for this 'metamorphosis' of observables into operators. This quantization method (as well as several other mathematically more sophisticated versions of it) is just a *recipe* or, depending on one's taste, "black magic", Hall (2005). Note that the enormous success of this recipe in various contexts - including field quantization - is no substitute for an explanation.

The choice of a particular quantization procedure will be strongly influenced by the preferred interpretation of the quantum theoretical formalism. If QT is interpreted as a theory describing individual events, then the Hamiltonian of classical mechanics becomes a natural starting point. This 'individuality assumption' is an essential part of the dominating 'conventional', or 'Copenhagen', interpretation (CI) of QT. It is well-known, that QT becomes a source of mysteries and paradoxes<sup>1</sup> whenever it is interpreted in the sense of CI, as a (complete) theory for individual events. Thus, the canonical quantization method and the CI are similar in

---

<sup>1</sup> I cannot report here a list, all the less a description, of all the quantum mechanical paradoxes found in the last eighty years.

two respects: both rely heavily on the concept of individual particles and both are rather mysterious.

This situation confronts us with a fundamental alternative. Should we accept the mysteries and paradoxes as inherent attributes of reality or should we not, instead, critically reconsider our assumptions, in particular the 'individuality assumption'. As a matter of fact, the dynamical numerical output of quantum mechanics consists of *probabilities*. A probability is a "deterministic" prediction which can be verified in a statistical sense only, i.e. by performing experiments on a large number of identically prepared individual systems, see Belinfante (1978), Margenau (1963). Therefore, the very structure of QT tells us that it is a theory about statistical ensembles only, see Ballentine (1970). If dogmatic or philosophical reasons 'force' us to interpret QT as a theory about individual events, we have to create complicated intellectual constructs, which are not part of the physical formalism, and lead to unsolved problems and contradictions.

The present author believes, like several other physicists [see e.g. Ali (2009); Ballentine (1970); Belinfante (1975); Blokhintsev (1964); Einstein (1936); Kemble (1929); Krüger (2004); Margenau (1935); Newman (1980); Pippard (1986); Ross-Bonney (1975); Toyozawa (1992); Tschudi (1987); Young (1980)] that QT is a purely statistical theory whose predictions can only be used to describe the behavior of statistical ensembles and not of individual particles. This statistical interpretation (SI) of QT eliminates all mysteries and paradoxes - and this shows that the mysteries and paradoxes are not part of QT itself but rather the result of a particular (mis)interpretation of QT. In view of the similarity discussed above, we adopt the statistical point of view, not only for the interpretation of QT itself, but also in our search for *quantization conditions*. The general strategy is to find a new set of (as) simple (as possible) statistical assumptions which can be understood in physical terms and imply QT. Such an approach would also provide an explanation for the correspondence rules (1).

The present paper belongs to a series of works aimed at such an explanation. Quite generally, the present work continues a long tradition of attempts, see Frieden (1989; 2004); Hall & Reginatto (2002a); Lee & Zhu (1999); Motz (1962); Rosen (1964); Schiller (1962a); Schrödinger (1926), to characterize QT by mathematical relations which can be understood in physical terms<sup>2</sup> (in contrast to the axiomatic approach). More specifically, it continues previous attempts to derive Schrödinger's equation with the help of statistical concepts, see Hall & Reginatto (2002b), Reginatto (1998a); Syska (2007), Klein (2009). These works, being quite different in detail, share the common feature that a statistical ensemble and not a particle Hamiltonian is used as a starting point for quantization. Finally, in a previous work, Klein (2011), of the present author an attempt has been undertaken to construct a complete statistical approach to QT with the help of a small number of very simple (statistical) assumptions. This work will be referred to as I. The present paper is a continuation and extension of I.

The quantization method reported in I is based on the following general ideas:

- QT should be a probabilistic theory in configuration space (not in phase space).
- QT should fulfill abstract versions of (i) a conservation law for probability (continuity equation) , and (ii) Ehrenfest's theorem. Such relations hold in all statistical theories no matter whether quantum or classical.
- There are no laws for particle trajectories in QT anymore. This arbitrariness, which represents a crucial difference between QT and classical statistics, should be handled by a statistical principle analogous to the principle of maximal entropy in classical statistics.

<sup>2</sup> The listing given here is far from complete.

These general ideas lead to the mathematical assumptions which represent the basis for the treatment reported in I. This work was restricted to a one-dimensional configuration space (a single particle ensemble with a single spatial degree of freedom). The present work generalizes the treatment of I to a  $3N$ -dimensional configuration space (ensembles representing an arbitrary number  $N$  of particles allowed to move in three-dimensional space), gauge-coupling, and spin. In a first step the generalization to three spatial dimensions is performed; the properly generalized basic relations are reported in section 2. This section contains also a review of the fundamental ideas.

In section 3 we make use of a mathematical freedom, which is already contained in our basic assumptions, namely the multi-valuedness of the variable  $S$ . This leads to the appearance of potentials in statistical relations replacing the local forces of single-event (mechanical) theories. The mysterious non-local action of the vector potential (in effects of the Aharonov-Bohm type) is explained as a consequence of the statistical nature of QT. In section 4 we discuss a related question: Which constraints on admissible forces exist for the present class of statistical theories? The answer is that only macroscopic (elementary) forces of the form of the Lorentz force can occur in nature, because only these survive the transition to QT. These forces are statistically represented by potentials, i.e. by the familiar gauge coupling terms in matter field equations. The present statistical approach provides a natural explanation for the long-standing question why potentials play an indispensable role in the field equations of physics.

In section 5 it is shown that among all statistical theories only the time-dependent Schrödinger equation follows the logical requirement of maximal disorder or minimal Fisher information. Spin one-half is introduced, in section 6, as the property of a statistical ensemble to respond to an external gauge field in two different ways. A generalized calculation, reported in sections 6 and 7, leads to Pauli's (single-particle) equation. In section 8 an alternative derivation, following Arunsalam (1970), and Gould (1995) is reported, which is particularly convenient for the generalization to arbitrary  $N$ . The latter is performed in section 9, which completes our statistical derivation of non-relativistic QT.

In section 10 the classical limit of QT is studied and it is stressed that the classical limit of QT is not classical mechanics but a classical statistical theory. In section 11 various questions related to the present approach, including the role of potentials and the interpretation of QT, are discussed. The final section 12 contains a short summary and mentions a possible direction for future research.

## 2. Basic equations for a class of statistical theories

In I three different types of theories have been defined which differ from each other with regard to the role of probability. We give a short review of the defining properties and supply some additional comments characterizing these theories.

The dogma underlying *theories of type 1* is determinism with regard to single events; probability does not play any role. If nature behaves according to this dogma, then measurements on identically prepared individual systems yield identical results. Classical mechanics is obviously such a deterministic type 1 theory. We shall use below (as a 'template' for the dynamics of our statistical theories) the following version of Newton's law, where the particle momentum  $p_k(t)$  plays the role of a second dynamic variable besides the spatial coordinate  $x_k(t)$ :

$$\frac{d}{dt}x_k(t) = \frac{p_k(t)}{m}, \quad \frac{d}{dt}p_k(t) = F_k(x, p, t). \quad (2)$$

In classical mechanics there is no restriction as regards the admissible forces. Thus,  $F_k$  is an arbitrary function of  $x_k, p_k, t$ ; it is, in particular, not required that it be derivable from a potential. Note that Eqs. (2) do *not* hold in the present theory; these relations are just used to establish a correspondence between classical mechanics and associated statistical theories.

Experimental data from atomic systems, recorded since the beginning of the last century, indicate that nature does not behave according to this single-event deterministic dogma. A simple but somewhat unfamiliar idea is, to construct a theory which is deterministic only in a statistical sense. This means that measurements on identically prepared individual systems do not yield identical results (no determinism with regard to single events) but repeated measurements on ensembles [consisting each time of a large (infinite) number of measurements on individual systems] yield identical results. In this case we have 'determinism' with regard to ensembles (expectation values, or probabilities).

Note that such a theory is far from chaotic even if our macroscopic anticipation of (single-event) determinism is not satisfied. Note also that there is no reason to assume that such a statistical theory for microscopic events is incompatible with macroscopic determinism. It is a frequently observed (but not always completely understood) phenomenon in nature that systems with many (microscopic) degrees of freedom can be described by a much smaller number of variables. During this process of elimination of variables the details of the corresponding microscopic theory for the individual constituents are generally lost. In other words, there is no reason to assume that a fundamental statistical law for individual atoms and a deterministic law for a piece of matter consisting of, say,  $10^{23}$  atoms should not be compatible with each other. This way of characterizing the relation between two physical theories is completely different from the common reductionistic point of view. Convincing arguments in favor of the former may, however, be found in Anderson (1972), Laughlin (2005). As discussed in I two types (referred to as type 2 and type 3) of indeterministic theories may be identified. In *type 2 theories* laws for individual particles exist (roughly speaking the individuality of particles remains intact) but the initial values are unknown and are described by probabilities only. An example for such a (classical-statistical) type 2 theory is statistical thermodynamics. On the other hand, in *type 3 theories* the amount of uncertainty is still greater, insofar as no dynamic laws for individual particles exist any more. A possible candidate for this 'extreme' type of indeterministic theory is quantum mechanics.

The method used in I to construct statistical theories was based on the following three assumptions,

- A local conservation law of probability with a particular form of the probability current.
- Two differential equations which are similar in structure to the canonical equations (2) but with observables replaced by expectation values.
- A differential version (minimal Fisher information) of the statistical principle of maximal disorder.

These (properly generalized) assumptions represent also the formal basis of the present work. The first and second of these cover type 2 as well as type 3 theories, while it will be shown that the third - the requirement of maximal disorder - does only hold for a single type 3 theory, namely quantum mechanics. In this sense quantum mechanics may be considered as the *most reasonable* theory among all statistical theories defined by the first two assumptions. There is obviously an analogy between quantum mechanics and the principle of minimal Fisher information on the one hand and classical statistical mechanics and the principle of maximal entropy on the other hand; both theories are realizations of the principle of maximal disorder.

Let us now generalize the basic equations of I (see section 3 of I) with respect to the number of spatial dimensions and with respect to gauge freedom. The continuity equation takes the form

$$\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial}{\partial x_k} \frac{\rho(x, t)}{m} \frac{\partial \tilde{S}(x, t)}{\partial x_k} = 0. \quad (3)$$

We use the summation convention, indices  $i, k, \dots$  run from 1 to 3 and are omitted if the corresponding variable occurs in the argument of a function. The existence of a local conservation law for the probability density  $\rho(x, t)$  is a necessity for a probabilistic theory. The same is true for the fact that the probability current takes the form  $j_k(x, t) = \rho(x, t) \tilde{p}_k(x, t)/m$ , where  $\tilde{p}_k(x, t)$  is the  $k$ -th component of the momentum probability density. The only non-trivial assumption contained in (3), is the fact that  $\tilde{p}_k(x, t)$  takes the form of the gradient,

$$\tilde{p}_k(x, t) = \frac{\partial \tilde{S}(x, t)}{\partial x_k}, \quad (4)$$

of a function  $\tilde{S}(x, t)$ . In order to gain a feeling for the physical meaning of (4) we could refer to the fact that a similar relation may be found in the Hamilton-Jacobi formulation of classical mechanics Sygne (1960); alternatively we could also refer to the fact that this condition characterizes 'irrotational flow' in fluid mechanics. Relation (4) could also be justified by means of the principle of simplicity; a gradient is the simplest way to represent a vector field, because it can be derived from a single scalar function.

In contrast to I we allow now for *multi-valued* functions  $\tilde{S}(x, t)$ . At first sight this seems strange since a multi-valued quantity cannot be an observable and should, consequently, not appear in equations bearing a physical meaning. However, only derivatives of  $\tilde{S}(x, t)$  occur in our basic equations. Thus, this freedom is possible without any additional postulate; we just have to require that

$$\tilde{S}(x, t) \text{ multi-valued, } \frac{\partial \tilde{S}}{\partial t}, \frac{\partial \tilde{S}}{\partial x_k} \text{ single-valued.} \quad (5)$$

(the quantity  $\tilde{p}$  defined in (4) is not multi-valued; this notation is used to indicate that this quantity has been defined with the help of a multi-valued  $\tilde{S}$ ). As discussed in more detail in section 3 this new 'degree of freedom' is intimately related to the existence of gauge fields. In contrast to  $\tilde{S}$ , the second dynamic variable  $\rho$  is a physical observable (in the statistical sense) and is treated as a single-valued function.

The necessary and sufficient condition for single-valuedness of a function  $\tilde{S}(x, t)$  (in a subspace  $\mathcal{G} \subseteq \mathcal{R}^4$ ) is that all second order derivatives of  $\tilde{S}(x, t)$  with respect to  $x_k$  and  $t$  commute with each other (in  $\mathcal{G}$ ) [see e.g. Kaempfer (1965)]. As a consequence, the order of two derivatives of  $\tilde{S}$  with respect to any one of the variables  $x_k, t$  must not be changed. We introduce the (single-valued) quantities

$$\tilde{S}_{[j,k]} = \left[ \frac{\partial^2 \tilde{S}}{\partial x_j \partial x_k} - \frac{\partial^2 \tilde{S}}{\partial x_k \partial x_j} \right], \quad \tilde{S}_{[0,k]} = \left[ \frac{\partial^2 \tilde{S}}{\partial t \partial x_k} - \frac{\partial^2 \tilde{S}}{\partial x_k \partial t} \right] \quad (6)$$

in order to describe the non-commuting derivatives in the following calculations.

The second of the assumptions listed above has been referred to in I as 'statistical conditions'. For the present three-dimensional theory these are obtained in the same way as in I,

by replacing the observables  $x_k(t)$ ,  $p_k(t)$  and the force field  $F_k(x(t), p(t), t)$  of the type 1 theory (2) by averages  $\overline{x_k}$ ,  $\overline{p_k}$  and  $\overline{F_k}$ . This leads to the relations

$$\frac{d}{dt} \overline{x_k} = \frac{\overline{p_k}}{m} \quad (7)$$

$$\frac{d}{dt} \overline{p_k} = \overline{F_k(x, p, t)}, \quad (8)$$

where the averages are given by the following integrals over the random variables  $x_k$ ,  $p_k$  (which should be clearly distinguished from the type I observables  $x_k(t)$ ,  $p_k(t)$  which will not be used any more):

$$\overline{x_k} = \int_{-\infty}^{\infty} d^3x \rho(x, t) x_k \quad (9)$$

$$\overline{p_k} = \int_{-\infty}^{\infty} d^3p w(p, t) p_k \quad (10)$$

$$\overline{F_k(x, p, t)} = \int_{-\infty}^{\infty} d^3x d^3p W(x, p, t) F_k(x, p, t). \quad (11)$$

The time-dependent probability densities  $W$ ,  $\rho$ ,  $w$  should be positive semidefinite and normalized to unity, i.e. they should fulfill the conditions

$$\int_{-\infty}^{\infty} d^3x \rho(x, t) = \int_{-\infty}^{\infty} d^3p w(p, t) = \int_{-\infty}^{\infty} d^3x d^3p W(x, p, t) = 1 \quad (12)$$

The densities  $\rho$  and  $w$  may be derived from the fundamental probability density  $W$  by means of the relations

$$\rho(x, t) = \int_{-\infty}^{\infty} d^3p W(x, p, t); \quad w(p, t) = \int_{-\infty}^{\infty} d^3x W(x, p, t). \quad (13)$$

The present construction of the statistical conditions (7) and (8) from the type 1 theory (2) shows two differences as compared to the treatment in I. The first is that we allow now for a  $p$ -dependent external force. This leads to a more complicated probability density  $W(x, p, t)$  as compared to the two decoupled densities  $\rho(x, t)$  and  $w(p, t)$  of I. The second difference, which is in fact related to the first, is the use of a multi-valued  $\tilde{S}(x, t)$ .

Note, that the  $p$ -dependent probability densities  $w(p, t)$  and  $W(x, p, t)$  have been introduced in the above relations in a purely formal way. We defined an expectation value  $\overline{p_k}$  [via Eq. (7)] and assumed [in Eq. (10)] that a random variable  $p_k$  and a corresponding probability density  $w(p, t)$  exist. But the validity of this assumption is not guaranteed. There is no compelling conceptual basis for the existence of these quantities in a pure configuration-space theory. If they exist, they must be defined with the help of additional considerations (see section 6 of I). The deeper reason for this problem is that the concept of measurement of momentum (which is proportional to the time derivative of position) is ill-defined in a theory whose observables are defined in terms of a large number of experiments at *one and the same* instant of time (measurement of a derivative requires measurements at different times). Fortunately, these considerations, which have been discussed in more detail in I, play not a prominent role [apart from the choice of  $W(x, p, t)$  discussed in section 4], for the derivation of Schrödinger's equation reported in the present paper<sup>3</sup>.

<sup>3</sup> These considerations seem relevant for attempts to define phase-space densities, e.g. of the Wigner type, in QT

Using the continuity equation (3) and the statistical conditions (7) and (8) the present generalization of the integral equation Eq. (24) of I may be derived. The steps leading to this result are very similar to the corresponding steps in I and may be skipped. The essential difference to the one-dimensional treatment is - apart from the number of space dimensions - the non-commutativity of the second order derivatives of  $\tilde{S}(x, t)$  leading to non-vanishing quantities  $\tilde{S}_{[j,k]}$ ,  $\tilde{S}_{[0,k]}$  defined in Eq. (6). The result takes the form

$$\begin{aligned}
& - \int_{-\infty}^{\infty} d^3x \frac{\partial \rho}{\partial x_k} \left[ \frac{\partial \tilde{S}}{\partial t} + \frac{1}{2m} \sum_j \left( \frac{\partial \tilde{S}}{\partial x_j} \right)^2 + V \right] \\
& + \int_{-\infty}^{\infty} d^3x \rho \left[ \frac{1}{m} \frac{\partial \tilde{S}}{\partial x_j} \tilde{S}_{[j,k]} + \tilde{S}_{[0,k]} \right] = \overline{F_k^{(e)}}(x, p, t)
\end{aligned} \tag{14}$$

In the course of the calculation leading to (14) it has been assumed that the macroscopic force  $F_k(x, p, t)$  entering the second statistical condition (8) may be written as a sum of two contributions,  $F_k^{(m)}(x, t)$  and  $F_k^{(e)}(x, p, t)$ ,

$$F_k(x, p, t) = F_k^{(m)}(x, t) + F_k^{(e)}(x, p, t), \tag{15}$$

where  $F_k^{(m)}(x, t)$  takes the form of a negative gradient of a scalar function  $V(x, t)$  (mechanical potential) and  $F_k^{(e)}(x, p, t)$  is the remaining  $p$ -dependent part.

Comparing Eq. (14) with the corresponding formula obtained in I [see Eq. (24) of I] we see that two new terms appear now, the expectation value of the  $p$ -dependent force on the r.h.s., and the second term on the l.h.s. of Eq. (14). The latter is a direct consequence of our assumption of a multi-valued variable  $\tilde{S}$ . In section 4 it will be shown that for vanishing multi-valuedness Eq. (14) has to agree with the three-dimensional generalization of the corresponding result [Eq. (24) of I] obtained in I. This means that the  $p$ -dependent term on the r.h.s. has to vanish too in this limit and indicates a relation between multi-valuedness of  $\tilde{S}$  and  $p$ -dependence of the external force.

### 3. Gauge coupling as a consequence of a multi-valued phase

In this section we study the consequences of the multi-valuedness [London (1927), Weyl (1929), Dirac (1931)] of the quantity  $\tilde{S}(x, t)$  in the present theory. We assume that  $\tilde{S}(x, t)$  may be written as a sum of a single-valued part  $S(x, t)$  and a multi-valued part  $\tilde{N}$ . Then, given that (5) holds, the derivatives of  $\tilde{S}(x, t)$  may be written in the form

$$\frac{\partial \tilde{S}}{\partial t} = \frac{\partial S}{\partial t} + e\Phi, \quad \frac{\partial \tilde{S}}{\partial x_k} = \frac{\partial S}{\partial x_k} - \frac{e}{c} A_k, \tag{16}$$

where the four functions  $\Phi$  and  $A_k$  are proportional to the derivatives of  $\tilde{N}$  with respect to  $t$  and  $x_k$  respectively (Note the change in sign of  $\Phi$  and  $A_k$  in comparison to Klein (2009); this is due to the fact that the multi-valued phase is now denoted by  $\tilde{S}$ ). The physical motivations for introducing the pre-factors  $e$  and  $c$  in Eq. (16) have been extensively discussed elsewhere, see Kaempfer (1965), Klein (2009), in an electro-dynamical context. In agreement with Eq. (16),  $\tilde{S}$  may be written in the form [Kaempfer (1965), Klein (2009)]

$$\tilde{S}(x, t; C) = S(x, t) - \frac{e}{c} \int_{x_0, t_0; C}^{x, t} [dx'_k A_k(x', t') - cd t' \phi(x', t')], \tag{17}$$

as a path-integral performed along an arbitrary path  $\mathcal{C}$  in four-dimensional space; the multi-valuedness of  $\tilde{S}$  simply means that it depends not only on  $x, t$  but also on the path  $\mathcal{C}$  connecting the points  $x_0, t_0$  and  $x, t$ .

The quantity  $\tilde{S}$  cannot be a physical observable because of its multi-valuedness. The fundamental physical quantities to be determined by our (future) theory are the four derivatives of  $\tilde{S}$  which will be rewritten here as two observable fields  $-\tilde{E}(x, t), \tilde{p}_k(x, t)$ ,

$$-\tilde{E}(x, t) = \frac{\partial S(x, t)}{\partial t} + e\Phi(x, t), \quad (18)$$

$$\tilde{p}_k(x, t) = \frac{\partial S(x, t)}{\partial x_k} - \frac{e}{c}A_k(x, t), \quad (19)$$

with dimensions of energy and momentum respectively.

We encounter a somewhat unusual situation in Eqs. (18), (19): On the one hand the left hand sides are observables of our theory, on the other hand we cannot solve our (future) differential equations for these quantities because of the peculiar multi-valued structure of  $\tilde{S}$ . We have to use instead the decompositions as given by the right hand sides of (18) and (19). The latter eight terms (the four derivatives of  $S$  and the four scalar functions  $\Phi$  and  $A_k$ ) are single-valued (in the mathematical sense) but need not be unique because only the left hand sides are uniquely determined by the physical situation. We tentatively assume that the fields  $\Phi$  and  $A_k$  are 'given' quantities in the sense that they represent an external influence (of 'external forces') on the considered statistical situation. An actual calculation has to be performed in such a way that fixed fields  $\Phi$  and  $A_k$  are chosen and then the differential equations are solved for  $S$  (and  $\rho$ ). However, as mentioned already, what is actually uniquely determined by the physical situation is the *sum* of the two terms on the right hand sides of (18) and (19). Consequently, a different set of fixed fields  $\Phi'$  and  $A'_k$  may lead to a physically equivalent, but mathematically different, solution  $S'$  in such a way that the sum of the new terms [on the right hand sides of (18) and (19)] is the *same* as the sum of the old terms. We assume here, that the formalism restores the values of the physically relevant terms. This implies that the relation between the old and new terms is given by

$$S'(x, t) = S(x, t) + \varphi(x, t) \quad (20)$$

$$\Phi'(x, t) = \Phi(x, t) - \frac{1}{e} \frac{\partial \varphi(x, t)}{\partial t} \quad (21)$$

$$A'_k(x, t) = A_k(x, t) + \frac{c}{e} \frac{\partial \varphi(x, t)}{\partial x_k}, \quad (22)$$

where  $\varphi(x, t)$  is an arbitrary, single-valued function of  $x_k, t$ . Consequently, all 'theories' (differential equations for  $S$  and  $\rho$  defined by the assumptions listed in section 2) should be form-invariant under the transformations (20)-(22). These invariance transformations, predicted here from general considerations, are (using an arbitrary function  $\chi = c\varphi/e$  instead of  $\varphi$ ) denoted as 'gauge transformations of the second kind'.

The fields  $\Phi(x, t)$  and  $A_k(x, t)$  describe an external influence but their numerical value is undefined; their value at  $x, t$  may be changed according to (21) and (22) without changing their physical effect. Thus, these fields *cannot play a local role* in space and time like forces and fields in classical mechanics and electrodynamics. What, then, is the physical meaning of these fields? An explanation which seems obvious in the present context is the following: They describe the *statistical effect* of an external influence on the considered system (ensemble of

identically prepared individual particles). The statistical effect of a force field on an ensemble may obviously *differ* from the local effect of the same force field on individual particles; thus the very existence of fields  $\Phi$  and  $A_k$  different from  $\vec{E}$  and  $\vec{B}$  is no surprise. The second common problem with the interpretation of the 'potentials'  $\Phi$  and  $A_k$  is their non-uniqueness. It is hard to understand that a quantity ruling the behavior of individual particles should not be uniquely defined. In contrast, this non-uniqueness is much easier to accept if  $\Phi$  and  $A_k$  rule the behavior of ensembles instead of individual particles. We have no problem to accept the fact that a function that represents a global (integral) effect may have many different local realizations.

It seems that this interpretation of the potentials  $\Phi$  and  $A_k$  is highly relevant for the interpretation of the effect found by Aharonov & Bohm (1959). If QT is interpreted as a theory about individual particles, the Aharonov-Bohm effects imply that a charged particle may be influenced in a nonlocal way by electromagnetic fields in inaccessible regions. This paradoxical prediction, which is however in strict agreement with QT, led even to a discussion about the reality of these effects, see Bocchieri & Loinger (1978), Roy (1980), Klein (1981), Peshkin & Tonomura (1989). A statistical interpretation of the potentials has apparently never been suggested, neither in the vast literature about the Aharonov-Bohm effect nor in papers promoting the statistical interpretation of QT; most physicists discuss this nonlocal 'paradox' from the point of view of 'the wave function of a single electron'. Further comments on this point may be found in section 11.

The expectation value  $F_k^{(e)}(x, p, t)$  on the right hand side of (14) is to be calculated using local, macroscopic forces whose functional form is still unknown. Both the potentials and these local forces represent an external influence, and it is reasonable to assume that the (nonlocal) potentials are the statistical representatives of the local forces on the r.h.s. of Eq. (14). The latter have to be determined by the potentials but must be uniquely defined at each space-time point. The gauge-invariant fields

$$E_k = -\frac{1}{c} \frac{\partial A_k}{\partial t} - \frac{\partial \Phi}{\partial x_k}, \quad B_k = \epsilon_{kij} \frac{\partial A_j}{\partial x_i}, \quad (23)$$

fulfill these requirements. As a consequence of the defining relations (23) they obey automatically the homogeneous Maxwell equations.

In a next step we rewrite the second term on the l.h.s. of Eq. (14). The commutator terms (6) take the form

$$\tilde{S}_{[0,k]} = -e \left( \frac{1}{c} \frac{\partial A_k}{\partial t} + \frac{\partial \Phi}{\partial x_k} \right), \quad \tilde{S}_{[j,k]} = \frac{e}{c} \left( \frac{\partial A_j}{\partial x_k} - \frac{\partial A_k}{\partial x_j} \right). \quad (24)$$

As a consequence, they may be expressed in terms of the local fields (23), which have been introduced above for reasons of gauge-invariance. Using (24), (23) and the relation (19) for the momentum field, Eq. (14) takes the form

$$\begin{aligned} & - \int_{-\infty}^{\infty} d^3x \frac{\partial \rho}{\partial x_k} \left[ \frac{\partial \tilde{S}}{\partial t} + \frac{1}{2m} \sum_j \left( \frac{\partial \tilde{S}}{\partial x_j} \right)^2 + V \right], \\ & + \int_{-\infty}^{\infty} d^3x \rho \left[ \frac{e}{c} \epsilon_{kij} \tilde{v}_i B_j + e E_k \right] = \overline{F_k^{(e)}}(x, p, t) \end{aligned} \quad (25)$$

with a velocity field defined by  $\tilde{v}_i = \tilde{p}_i / m$ . Thus, the new terms on the l.h.s. of (25) - stemming from the multi-valuedness of  $\tilde{S}$  - take the form of an expectation value (with  $\mathcal{R}^3$  as sample

space) of the Lorentz force field

$$\vec{F}_L(x, t) = e\vec{E}(x, t) + \frac{e}{c}\vec{v}(x, t) \times \vec{B}(x, t), \quad (26)$$

if the particle velocity is identified with the velocity field  $\vec{v}(x, t)$ .

The above steps imply a relation between potentials and local fields. From the present statistical (nonlocal) point of view the potentials are more fundamental than the local fields. In contrast, considered from the point of view of macroscopic physics, the local fields are the physical quantities of primary importance and the potentials may (or may not) be introduced for mathematical convenience.

#### 4. A constraint for forces in statistical theories

Let us discuss now the nature of the macroscopic forces  $F_k^{(e)}(x, p, t)$  entering the expectation value on the r.h.s. of Eq. (25). In our type I parent theory, classical mechanics, there are no constraints for the possible functional form of  $F_k^{(e)}(x, p, t)$ . However, this need not be true in the present statistical framework. As a matter of fact, the way the mechanical potential  $V(x, t)$  entered the differential equation for  $S$  (in the previous work I) indicates already that such constraints do actually exist. Let us recall that in I we tacitly restricted the class of forces to those derivable from a potential  $V(x, t)$ . If we eliminate this restriction and admit arbitrary forces, with components  $F_k(x, t)$ , we obtain instead of the above relation (25) the simpler relation [Eq. (24) of I, generalized to three dimensions and arbitrary forces of the form  $F_k(x, t)$ ]

$$-\int_{-\infty}^{\infty} d^3x \frac{\partial \rho}{\partial x_k} \left[ \frac{1}{2m} \sum_j \left( \frac{\partial S}{\partial x_j} \right)^2 + \frac{\partial S}{\partial t} \right] = \int_{-\infty}^{\infty} dx \rho F_k(x, t). \quad (27)$$

This is a rather complicated integro-differential equation for our variables  $\rho(x, t)$  and  $S(x, t)$ . We assume now, using mathematical simplicity as a guideline, that Eq. (27) can be written in the common form of a local differential equation. This assumption is of course not evident; in principle the laws of physics could be integro-differential equations or delay differential equations or take an even more complicated mathematical form. Nevertheless, this assumption seems rather weak considering the fact that all fundamental laws of physics take this 'simple' form. Thus, we postulate that Eq. (27) is equivalent to a differential equation

$$\frac{1}{2m} \sum_j \left( \frac{\partial S}{\partial x_j} \right)^2 + \frac{\partial S}{\partial t} + T = 0, \quad (28)$$

where the unknown term  $T$  describes the influence of the force  $F_k$  but may also contain other contributions. Let us write

$$T = -L_0 + V, \quad (29)$$

where  $L_0$  does not depend on  $F_k$ , while  $V$  depends on it and vanishes for  $F_k \rightarrow 0$ . Inserting (28) and (29) in (27) yields

$$\int d^3x \frac{\partial \rho}{\partial x_k} (-L_0 + V) = \int d^3x \rho F_k(x, t). \quad (30)$$

For  $F_k \rightarrow 0$  Eq. (30) leads to the relation

$$\int d^3x \frac{\partial \rho}{\partial x_k} L_0 = 0, \quad (31)$$

which remains true for finite forces because  $L_0$  does not depend on  $F_k$ . Finally, performing a partial integration, we see that a relation

$$F_k = -\frac{\partial V}{\partial x_k} + s_k, \quad \int_{-\infty}^{\infty} d^3x \rho s_k = 0, \quad (32)$$

exists between  $F_k$  and  $V$ , with a vanishing expectation value of the (statistically irrelevant) functions  $s_k$ . This example shows that the restriction to gradient fields, made above and in I, is actually not necessary. We may *admit* force fields which are arbitrary functions of  $x$  and  $t$ ; the statistical conditions (which play now the role of a 'statistical constraint') eliminate automatically all forces that cannot be written after statistical averaging as gradient fields.

This is very interesting and indicates the possibility that the present statistical assumptions leading to Schrödinger's equation may also be responsible, at least partly, for the structure of the real existing (gauge) interactions of nature.

Does this statistical constraint also work in the present  $p$ -dependent case? We assume that the force in (25) is a standard random variable with the configuration space as sample space (see the discussion in section 4 of I) and that the variable  $p$  in  $F_k^{(e)}(x, p, t)$  may consequently be replaced by the field  $\tilde{p}(x, t)$  [see (19)]. Then, the expectation value on the r.h.s. of (25) takes the form

$$\overline{F_k^{(e)}(x, p, t)} = \int_{-\infty}^{\infty} d^3x \rho(x, t) H_k(x, \frac{\partial \tilde{S}(x, t)}{\partial x}, t). \quad (33)$$

The second term on the l.h.s. of (25) has the *same* form. Therefore, the latter may be eliminated by writing

$$H_k(x, \frac{\partial \tilde{S}}{\partial x}, t) = \frac{e}{c} \epsilon_{kij} \frac{1}{m} \frac{\partial \tilde{S}}{\partial x_i} B_j + eE_k + h_k(x, \frac{\partial \tilde{S}}{\partial x}, t), \quad (34)$$

with  $h_k(x, p, t)$  as our new unknown functions. They obey the simpler relations

$$-\int_{-\infty}^{\infty} d^3x \frac{\partial \rho}{\partial x_k} \left[ \frac{\partial \tilde{S}}{\partial t} + \frac{1}{2m} \sum_j \left( \frac{\partial \tilde{S}}{\partial x_j} \right)^2 + V \right] = \int_{-\infty}^{\infty} d^3x \rho h_k(x, \frac{\partial \tilde{S}}{\partial x}, t). \quad (35)$$

On a first look this condition for the allowed forces looks similar to the  $p$ -independent case [see (27)]. But the dependence of  $h_k$  on  $x, t$  cannot be considered as 'given' (externally controlled), as in the  $p$ -independent case, because it contains now the unknown  $x, t$ -dependence of the derivatives of  $\tilde{S}$ . We may nevertheless try to incorporate the r.h.s by adding a term  $\tilde{T}$  to the bracket which depends on the derivatives of the multivalued quantity  $\tilde{S}$ . This leads to the condition

$$h_k(x, \frac{\partial \tilde{S}}{\partial x}, t) = -\frac{\partial \tilde{T}(x, \frac{\partial \tilde{S}}{\partial x}, t)}{\partial x_k} + s_k, \quad \int_{-\infty}^{\infty} d^3x \rho s_k = 0. \quad (36)$$

But this relation cannot be fulfilled for nontrivial  $h_k, \tilde{T}$  because the derivatives of  $\tilde{S}$  cannot be subject to further constraints beyond those given by the differential equation; on top of that the derivatives with regard to  $x$  on the r.h.s. create higher order derivatives of  $\tilde{S}$  which are

not present at the l.h.s. of Eq. (36). The only possibility to fulfill this relation is for constant  $\frac{\partial \tilde{S}}{\partial x}$ , a special case which has in fact already been taken into account by adding the mechanical potential  $V$ . We conclude that the statistical constraint leads to  $h_k = \tilde{T} = 0$  and that the statistical condition (35) takes the form

$$-\int d^3x \frac{\partial \rho}{\partial x_k} \left[ \frac{\partial \tilde{S}}{\partial t} + \frac{1}{2m} \sum_j \left( \frac{\partial \tilde{S}}{\partial x_j} \right)^2 + V \right] = 0. \quad (37)$$

Thus, only a mechanical potential and the four electrodynamic potentials are compatible with the statistical constraint and will consequently - assuming that the present statistical approach reflects a fundamental principle of nature - be realized in nature. As is well known all existing interactions follow (sometimes in a generalized form) the gauge coupling scheme derived above. The statistical conditions imply not only Schrödinger's equation but also the form of the (gauge) coupling to external influences and the form of the corresponding local force, the Lorentz force,

$$\vec{F}_L = e\vec{E} + \frac{e}{c}\vec{v} \times \vec{B}, \quad (38)$$

if the particle velocity  $\vec{v}$  is identified with the velocity field  $\vec{v}(x, t)$ .

In the present derivation the usual order of proceeding is just inverted. In the conventional deterministic treatment the form of the local forces (Lorentz force), as taken from experiment, is used as a starting point. The potentials are introduced afterwards, in the course of a transition to a different formal framework (Lagrange formalism). In the present approach the fundamental assumptions are the statistical conditions. Then, taking into account an existing mathematical freedom (multi-valuedness of a variable) leads to the introduction of potentials. From these, the shape of the macroscopic (Lorentz) force can be derived, using the validity of the statistical conditions as a constraint.

## 5. Fisher information as the hallmark of quantum theory

The remaining nontrivial task is the derivation of a local differential equation for  $S$  and  $\rho$  from the integral equation (37). As our essential constraint we will use, besides general principles of simplicity (like homogeneity and isotropy of space) the principle of maximal disorder, as realized by the requirement of minimal Fisher information. Using the abbreviation

$$\bar{L}(x, t) = \frac{\partial \tilde{S}}{\partial t} + \frac{1}{2m} \left( \frac{\partial \tilde{S}(x, t)}{\partial x} \right)^2 + V(x, t), \quad (39)$$

the general solution of (37) may be written in the form

$$\frac{\partial \rho}{\partial x_k} \bar{L}(x, t) = G_k(x, t), \quad (40)$$

where the three functions  $G_k(x, t)$  have to vanish upon integration over  $\mathcal{R}^3$  and are otherwise arbitrary. If we restrict ourselves to an isotropic law, we may write

$$G_k(x, t) = \frac{\partial \rho}{\partial x_k} L_0. \quad (41)$$

Then, our problem is to find a function  $L_0$  which fulfills the differential equation

$$\bar{L}(x, t) - L_0 = 0, \quad (42)$$

and condition (31). The method used in I for a one-dimensional situation, to determine  $L_0$  from the requirement of minimal Fisher information, remains essentially unchanged in the present three-dimensional case. The reader is referred to the detailed explanations reported in I.

In I it has been shown that this principle of maximal disorder leads to an anomalous variational problem and to the following conditions for our unknown function  $L_0$ :

$$\bar{L}(x, t) - L_0 \left( \rho, \frac{\partial \rho}{\partial x}, \frac{\partial^2 \rho}{\partial x \partial x} \right) = 0 \quad (43)$$

$$\delta \int d^3x \rho \left[ \bar{L}(x, t) - L_0 \left( \rho, \frac{\partial \rho}{\partial x}, \frac{\partial^2 \rho}{\partial x \partial x} \right) \right] = 0, \quad (44)$$

where  $L_0$  contains only derivatives of  $\rho$  up to second order and does not explicitly depend on  $x, t$ . If Eq. (43) is taken into account, the Euler-Lagrange equations of the variational problem (44) lead to the following differential equation

$$-\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_i} \frac{\partial \beta}{\partial \left( \frac{\partial^2 \rho}{\partial x_k \partial x_i} \right)} + \frac{\partial}{\partial x_k} \frac{\partial \beta}{\partial \left( \frac{\partial \rho}{\partial x_k} \right)} - \frac{\partial \beta}{\partial \rho} + \frac{\beta}{\rho} = 0 \quad (45)$$

for the variable  $\beta = \rho L_0$ . Eq. (45) is a straightforward generalization of the corresponding one-dimensional relation [equation (68) of I] to three spatial dimensions.

Besides (45) a further (consistency) condition exists, which leads to a simplification of the problem. The function  $L_0$  may depend on second order derivatives of  $\rho$  but this dependence must be of a special form not leading to any terms in the Euler-Lagrange equations [according to (43) our final differential equation for  $S$  and  $\rho$  must not contain higher than second order derivatives of  $\rho$ ]. Consequently, the first term in Eq. (45) (as well as the sum of the remaining terms) has to vanish separately and (45) can be replaced by the two equations

$$\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_i} \frac{\partial \beta}{\partial \left( \frac{\partial^2 \rho}{\partial x_k \partial x_i} \right)} = 0 \quad (46)$$

$$\frac{\partial}{\partial x_k} \frac{\partial \beta}{\partial \left( \frac{\partial \rho}{\partial x_k} \right)} - \frac{\partial \beta}{\partial \rho} + \frac{\beta}{\rho} = 0. \quad (47)$$

In I a new derivation of Fisher's functional has been obtained, using the general solution of the one-dimensional version of (45), as well as the so-called composition law. In the present three-dimensional situation we set ourselves a less ambitious aim. We know that Fisher's functional describes the maximal amount of disorder. If we are able to find a solution of (46), (47) that agrees with this functional (besides 'null-terms' giving no contribution to the Euler-Lagrange equations) then we will accept it as our correct solution. It is easy to see that this solution is given by

$$L_0 = B_0 \left[ -\frac{1}{2\rho^2} \sum_j \left( \frac{\partial \rho}{\partial x_j} \right)^2 + \frac{1}{\rho} \sum_j \frac{\partial^2 \rho}{\partial x_j^2} \right], \quad (48)$$

where  $B_0$  is an arbitrary constant. Eq. (48) presents again the three-dimensional (and isotropic) generalization of the one-dimensional result obtained in I. By means of the identity

$$\frac{\partial}{\partial x_i} \frac{\partial \sqrt{\rho}}{\partial x_i} \frac{\partial \sqrt{\rho}}{\partial x_k} = \frac{\partial \sqrt{\rho}}{\partial x_k} \frac{\partial}{\partial x_i} \frac{\partial \sqrt{\rho}}{\partial x_i} + \frac{1}{2} \frac{\partial}{\partial x_k} \frac{\partial \sqrt{\rho}}{\partial x_i} \frac{\partial \sqrt{\rho}}{\partial x_i}, \quad (49)$$

it is easily verified that the solution (48) obeys also condition (31). Using the decomposition (16) and renaming  $B$  according to  $B = \hbar^2/4m$ , the continuity equation (3) and the second differential equation (43) respectively, take the form

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_k} \frac{\rho}{m} \left( \frac{\partial S}{\partial x_k} - \frac{e}{c} A_k \right) = 0, \quad (50)$$

$$\frac{\partial S}{\partial t} + e\phi + \frac{1}{2m} \sum_k \left( \frac{\partial S}{\partial x_k} - \frac{e}{c} A_k \right)^2 + V - \frac{\hbar^2}{2m} \frac{\Delta \sqrt{\rho}}{\sqrt{\rho}} = 0. \quad (51)$$

The function  $S$  occurring in (50), (51) is single-valued but not unique (not gauge-invariant). If now the complex-valued variable

$$\psi = \sqrt{\rho} e^{i\frac{S}{\hbar}}, \quad (52)$$

is introduced, the two equations (50), (51) may be written in compact form as real and imaginary parts of the linear differential equation

$$\left( \frac{\hbar}{i} \frac{\partial}{\partial t} + e\phi \right) \psi + \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \vec{x}} - \frac{e}{c} \vec{A} \right)^2 \psi + V\psi = 0, \quad (53)$$

which completes our derivation of Schrödinger's equation in the presence of a gauge field.

Eq. (53) is in manifest gauge-invariant form. The gauge-invariant derivatives of  $\tilde{S}$  with respect to  $t$  and  $\vec{x}$  correspond to the two brackets in (53). In particular, the canonical momentum  $\partial S/\partial \vec{x}$  corresponds to the momentum operator proportional to  $\partial/\partial \vec{x}$ . Very frequently, Eq. (53) is written in the form

$$-\frac{\hbar}{i} \frac{\partial}{\partial t} \psi = H\psi, \quad (54)$$

with the Hamilton operator

$$H = \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \vec{x}} - \frac{e}{c} \vec{A} \right)^2 + V + e\phi, \quad (55)$$

Our final result, Eqs. (54), (55), agrees with the result of the conventional quantization procedure. In its simplest form, the latter starts from the classical relation  $H(x, p) = E$ , where  $H(x, p)$  is the Hamiltonian of a classical particle in a conservative force field, and  $E$  is its energy. To perform a "canonical quantization" means to replace  $p$  and  $E$  by differential expressions according to (1) and let then act both sides of the equation  $H(x, p) = E$  on states  $\psi$  of a function space. The 'black magic' involved in this process has been eliminated, or at least dramatically reduced, in the present approach, where Eqs. (54), (55) have been derived from a set of assumptions which can all be interpreted in physical terms.

The Hamiltonian (55) depends on the potentials  $\Phi$  and  $\vec{A}$  and is consequently a non-unique (not gauge-invariant) mathematical object. The same is true for the time-development operator  $U(H)$  which is an operator function of  $H$ , see e.g. Kobe & Yang (1985). This non-uniqueness is a problem if  $U(H)$  is interpreted as a quantity ruling the time-evolution of a single particle. It is no problem from the point of view of the SI where  $H$  and  $U(H)$  are primarily convenient mathematical objects which occur in a natural way if the time-dependence of statistically relevant (uniquely defined) quantities, like expectation values and transition probabilities, is to be calculated.

## 6. Spin as a statistical degree of freedom

Spin is generally believed to be a phenomenon of quantum-theoretic origin. For a long period of time, following Dirac's derivation of his relativistic equation, it was also believed to be essentially of relativistic origin. This has changed since the work of Schiller (1962b), Levy-Leblond (1967), Arunsalam (1970), Gould (1995), Reginatto (1998b) and others, who showed that spin may be derived entirely in the framework of non-relativistic QT without using any relativistic concepts. Thus, a new derivation of non-relativistic QT like the present one should also include a derivation of the phenomenon of spin. This will be done in this and the next two sections.

A simple idea to extend the present theory is to assume that sometimes - under certain external conditions to be identified later - a situation occurs where the behavior of our statistical ensemble of particles cannot longer be described by  $\rho$ ,  $S$  alone but requires, e.g., the double number of field variables; let us denote these by  $\rho_1$ ,  $S_1$ ,  $\rho_2$ ,  $S_2$  (we restrict ourselves here to spin one-half). The relations defining this generalized theory should be formulated in such a way that the previous relations are obtained in the appropriate limits. One could say that we undertake an attempt to introduce a new (discrete) degree of freedom for the ensemble. If we are able to derive a non-trivial set of differential equations - with coupling between  $\rho_1$ ,  $S_1$  and  $\rho_2$ ,  $S_2$  - then such a degree of freedom could exist in nature.

Using these guidelines, the basic equations of the generalized theory can be easily formulated. The probability density and probability current take the form  $\rho = \rho_1 + \rho_2$  and  $\vec{j} = \vec{j}_1 + \vec{j}_2$ , with  $\vec{j}_i$  ( $i = 1, 2$ ) defined in terms of  $\rho_i$ ,  $S_i$  exactly as before (see section 2). Then, the continuity equation is given by

$$\frac{\partial(\rho_1 + \rho_2)}{\partial t} + \frac{\partial}{\partial x_l} \left( \frac{\rho_1}{m} \frac{\partial \tilde{S}_1}{\partial x_l} + \frac{\rho_2}{m} \frac{\partial \tilde{S}_2}{\partial x_l} \right) = 0, \quad (56)$$

where we took the possibility of multi-valuedness of the "phases" already into account, as indicated by the notation  $\tilde{S}_i$ . The statistical conditions are given by the two relations

$$\frac{d}{dt} \overline{x_k} = \frac{\overline{p_k}}{m} \quad (57)$$

$$\frac{d}{dt} \overline{p_k} = \overline{F_k^{(T)}(x, p, t)}, \quad (58)$$

which are similar to the relations used previously (in section 2 and in I), and by an additional equation

$$\frac{d}{dt} \overline{s_k} = \overline{F_k^{(R)}(x, p, t)}, \quad (59)$$

which is required as a consequence of our larger number of dynamic variables. Eq. (59) is best explained later; it is written down here for completeness. The forces  $F_k^{(T)}(x, p, t)$  and  $F_k^{(R)}(x, p, t)$  on the r.h.s. of (58) and (59) are again subject to the "statistical constraint", which has been defined in section 3. The expectation values are defined as in (9)-(11).

Performing mathematical manipulations similar to the ones reported in section 2, the l.h.s. of Eq. (58) takes the form

$$\begin{aligned} \frac{d}{dt} \overline{p_k} = \int d^3x \left[ \frac{\partial \rho_1}{\partial t} \frac{\partial \tilde{S}_1}{\partial x_k} + \frac{\partial \rho_2}{\partial t} \frac{\partial \tilde{S}_2}{\partial x_k} \right. \\ \left. - \frac{\partial \rho_1}{\partial x_k} \frac{\partial \tilde{S}_1}{\partial t} - \frac{\partial \rho_2}{\partial x_k} \frac{\partial \tilde{S}_2}{\partial t} + \rho_1 \tilde{S}_{[0,k]}^{(1)} + \rho_2 \tilde{S}_{[0,k]}^{(2)} \right], \end{aligned} \quad (60)$$

where the quantities  $\tilde{S}_{[j,k]}^{(i)}$ ,  $i = 1, 2$  are defined as above [see Eq. (6)] but with  $\tilde{S}$  replaced by  $\tilde{S}_i$ . Let us write now  $\tilde{S}$  in analogy to section 2 in the form  $\tilde{S}_i = S_i + \tilde{N}_i$ , as a sum of a single-valued part  $S_i$  and a multi-valued part  $\tilde{N}_i$ . If  $\tilde{N}_1$  and  $\tilde{N}_2$  are to represent an external influence, they must be identical and a single multi-valued part  $\tilde{N} = \tilde{N}_1 = \tilde{N}_2$  may be used instead. The derivatives of  $\tilde{N}$  with respect to  $t$  and  $x_k$  must be single-valued and we may write

$$\frac{\partial \tilde{S}_i}{\partial t} = \frac{\partial S_i}{\partial t} + e\Phi, \quad \frac{\partial \tilde{S}_i}{\partial x_k} = \frac{\partial S_i}{\partial x_k} - \frac{e}{c} A_k, \quad (61)$$

using the same familiar electrodynamic notation as in section 2. In this way we arrive at eight single-valued functions to describe the external conditions and the dynamical state of our system, namely  $\Phi$ ,  $A_k$  and  $\rho_i$ ,  $S_i$ .

In a next step we replace  $\rho_i$ ,  $S_i$  by new dynamic variables  $\rho$ ,  $S$ ,  $\vartheta$ ,  $\varphi$  defined by

$$\begin{aligned} \rho_1 &= \rho \cos^2 \frac{\vartheta}{2}, & S_1 &= S + \frac{\hbar}{2} \varphi, \\ \rho_2 &= \rho \sin^2 \frac{\vartheta}{2}, & S_2 &= S - \frac{\hbar}{2} \varphi. \end{aligned} \quad (62)$$

A transformation similar to Eq. (62) has been introduced by Takabayasi (1955) in his reformulation of Pauli's equation. Obviously, the variables  $S$ ,  $\rho$  describe 'center of mass' properties (which are common to both states 1 and 2) while  $\vartheta$ ,  $\varphi$  describe relative (internal) properties of the system.

The dynamical variables  $S$ ,  $\rho$  and  $\vartheta$ ,  $\varphi$  are not decoupled from each other. It turns out (see below) that the influence of  $\vartheta$ ,  $\varphi$  on  $S$ ,  $\rho$  can be described in a (formally) similar way as the influence of an external electromagnetic field if a 'vector potential'  $\vec{A}^{(s)}$  and a 'scalar potential'  $\phi^{(s)}$ , defined by

$$A_l^{(s)} = -\frac{\hbar c}{2e} \cos \vartheta \frac{\partial \varphi}{\partial x_l}, \quad \phi^{(s)} = \frac{\hbar}{2e} \cos \vartheta \frac{\partial \varphi}{\partial t}, \quad (63)$$

are introduced. Denoting these fields as 'potentials', we should bear in mind that they are not externally controlled but defined in terms of the internal dynamical variables. Using the abbreviations

$$\hat{A}_l = A_l + A_l^{(s)}, \quad \hat{\phi} = \phi + \phi^{(s)}, \quad (64)$$

the second statistical condition (58) can be written in the following compact form

$$\begin{aligned} & - \int d^3x \frac{\partial \rho}{\partial x_l} \left[ \left( \frac{\partial S}{\partial t} + e\hat{\phi} \right) + \frac{1}{2m} \sum_j \left( \frac{\partial S}{\partial x_j} - \frac{e}{c} \hat{A}_j \right)^2 \right] \\ & + \int d^3x \rho \left[ -\frac{e}{c} v_j \left( \frac{\partial \hat{A}_l}{\partial x_j} - \frac{\partial \hat{A}_j}{\partial x_l} \right) - \frac{e}{c} \frac{\partial \hat{A}_l}{\partial t} - e \frac{\partial \hat{\phi}}{\partial x_l} \right] \\ & = \overline{F_l^{(T)}(x, p, t)} = \int d^3x \rho F_l^{(T)}(x, p, t), \end{aligned} \quad (65)$$

which shows a formal similarity to the spinless case [see (14) and (24)]. The components of the velocity field in (65) are given by

$$v_j = \frac{1}{m} \left( \frac{\partial S}{\partial x_j} - \frac{e}{c} \hat{A}_j \right). \quad (66)$$

If now fields  $E_l$ ,  $B_l$  and  $E_l^{(s)}$ ,  $B_l^{(s)}$  are introduced by relations analogous to (23), the second line of (65) may be written in the form

$$\int d^3x \rho \left[ (e\vec{E} + \frac{e}{c}\vec{v} \times \vec{B})_l + (eE^{(s)} + \frac{e}{c}\vec{v} \times B^{(s)})_l \right], \quad (67)$$

which shows that both types of fields, the external fields as well as the internal fields due to  $\vartheta$ ,  $\varphi$ , enter the theory in the same way, namely in the form of a Lorentz force.

The first, externally controlled Lorentz force in (67) may be eliminated in exactly the same manner as in section 3 by writing

$$\overline{F_l^{(T)}(x, p, t)} = \int d^3x \rho (e\vec{E} + \frac{e}{c}\vec{v} \times \vec{B})_l + \int d^3x \rho F_l^{(I)}(x, p, t). \quad (68)$$

This means that one of the forces acting on the system as a whole is again given by a Lorentz force; there may be other nontrivial forces  $F^{(I)}$  which are still to be determined. The second 'internal' Lorentz force in (67) can, of course, not be eliminated in this way. In order to proceed, the third statistical condition (59) must be implemented. To do that it is useful to rewrite Eq. (65) in the form

$$\begin{aligned} & - \int d^3x \frac{\partial \rho}{\partial x_l} \left[ \left( \frac{\partial S}{\partial t} + e\hat{\phi} \right) + \frac{1}{2m} \sum_j \left( \frac{\partial S}{\partial x_j} - \frac{e}{c} \hat{A}_j \right)^2 \right] \\ & + \int d^3x \frac{\hbar}{2} \rho \sin \vartheta \left( \frac{\partial \vartheta}{\partial x_l} \left[ \frac{\partial \varphi}{\partial t} + v_j \frac{\partial \varphi}{\partial x_j} \right] - \frac{\partial \varphi}{\partial x_l} \left[ \frac{\partial \vartheta}{\partial t} + v_j \frac{\partial \vartheta}{\partial x_j} \right] \right) \\ & = \overline{F_l^{(I)}(x, p, t)} = \int d^3x \rho F_l^{(I)}(x, p, t), \end{aligned} \quad (69)$$

using (67), (68) and the definition (63) of the fields  $A_l^{(s)}$  and  $\phi^{(s)}$ .

We interpret the fields  $\varphi$  and  $\vartheta$  as angles (with  $\varphi$  measured from the  $y$ -axis of our coordinate system) determining the direction of a vector

$$\vec{s} = \frac{\hbar}{2} (\sin \vartheta \sin \varphi \vec{e}_x + \sin \vartheta \cos \varphi \vec{e}_y + \cos \vartheta \vec{e}_z), \quad (70)$$

of constant length  $\frac{\hbar}{2}$ . As a consequence,  $\vec{s}$  and  $\vec{s}$  are perpendicular to each other and the classical force  $\vec{F}^{(R)}$  in Eq. (59) should be of the form  $\vec{D} \times \vec{s}$ , where  $\vec{D}$  is an unknown field. In contrast to the 'external force', we are unable to determine the complete form of this 'internal' force from the statistical constraint [an alternative treatment will be reported in section 8] and set

$$\vec{F}^{(R)} = -\frac{e}{mc} \vec{B} \times \vec{s}, \quad (71)$$

where  $\vec{B}$  is the external 'magnetic field', as defined by Eq. (23), and the factor in front of  $\vec{B}$  has been chosen to yield the correct  $g$ -factor of the electron.

The differential equation

$$\frac{d}{dt} \vec{s} = -\frac{e}{mc} \vec{B} \times \vec{s} \quad (72)$$

for particle variables  $\vartheta(t)$ ,  $\varphi(t)$  describes the rotational state of a classical magnetic dipole in a magnetic field, see Schiller (1962b). Recall that we do *not* require that (72) is fulfilled in the

present theory. The present variables are the fields  $\vartheta(x, t)$ ,  $\varphi(x, t)$  which may be thought of as describing a kind of 'rotational state' of the statistical ensemble as a whole, and have to fulfill the 'averaged version' (59) of (72).

Performing steps similar to the ones described in I (see also section 2), the third statistical condition (59) implies the following differential relations,

$$\begin{aligned} \dot{\varphi} + v_j \frac{\partial \varphi}{\partial x_j} &= \frac{e}{mc} \frac{1}{\sin \vartheta} (B_z \sin \vartheta - B_y \cos \vartheta \cos \varphi - B_x \cos \vartheta \sin \varphi) \\ &+ \frac{\cos \varphi}{\sin \vartheta} G_1 - \frac{\sin \varphi}{\sin \vartheta} G_2, \end{aligned} \quad (73)$$

$$\dot{\vartheta} + v_j \frac{\partial \vartheta}{\partial x_j} = \frac{e}{mc} (B_x \cos \varphi - B_y \sin \varphi) - \frac{G_3}{\sin \vartheta}, \quad (74)$$

for the dynamic variables  $\vartheta$  and  $\varphi$ . These equations contain three fields  $G_i(x, t)$ ,  $i = 1, 2, 3$  which have to obey the conditions

$$\int d^3x \rho G_i = 0, \quad \vec{G} \vec{s} = 0, \quad (75)$$

and are otherwise arbitrary. The 'total derivatives' of  $\varphi$  and  $\vartheta$  in (69) may now be eliminated with the help of (73),(74) and the second line of Eq. (69) takes the form

$$\begin{aligned} &\int d^3x \frac{\partial \rho}{\partial x_l} \frac{e}{mc} s_j B_j + \int d^3x \rho \frac{e}{mc} s_j \frac{\partial}{\partial x_l} B_j \\ &+ \int d^3x \rho \frac{\hbar}{2} \left( \cos \varphi \frac{\partial \vartheta}{\partial x_l} G_1 - \sin \varphi \frac{\partial \vartheta}{\partial x_l} G_2 + \frac{\partial \varphi}{\partial x_l} G_3 \right). \end{aligned} \quad (76)$$

The second term in (76) presents an external macroscopic force. It may be eliminated from (69) by writing

$$\overline{F_l^{(I)}}(x, p, t) = \int d^3x \rho \left( -\mu_j \frac{\partial}{\partial x_l} B_j \right) + \overline{F_l^{(V)}}(x, p, t), \quad (77)$$

where the magnetic moment of the electron  $\mu_i = -(e/mc)s_i$  has been introduced. The first term on the r.h.s. of (77) is the expectation value of the well-known electro-dynamical force exerted by an inhomogeneous magnetic field on the translational motion of a magnetic dipole; this classical force plays an important role in the standard interpretation of the quantum-mechanical Stern-Gerlach effect. It is satisfying that both translational forces, the Lorentz force as well as this dipole force, can be derived in the present approach. The remaining unknown force  $\vec{F}^{(V)}$  in (77) leads (in the same way as in section 3) to a mechanical potential  $V$ , which will be omitted for brevity.

The integrand of the first term in (76) is linear in the derivative of  $\rho$  with respect to  $x_l$ . It may consequently be added to the first line of (69) which has the same structure. Therefore, it represents (see below) a contribution to the generalized Hamilton-Jacobi differential equation. The third term in (76) has the mathematical structure of a force term, but does not contain any externally controlled fields. Thus, it must also represent a contribution to the generalized Hamilton-Jacobi equation. This implies that this third term can be written as

$$\int d^3x \rho \frac{\hbar}{2} \left( \cos \varphi \frac{\partial \vartheta}{\partial x_l} G_1 - \sin \varphi \frac{\partial \vartheta}{\partial x_l} G_2 + \frac{\partial \varphi}{\partial x_l} G_3 \right) = \int d^3x \frac{\partial \rho}{\partial x_l} L'_0, \quad (78)$$

where  $L'_0$  is an unknown field depending on  $G_1, G_2, G_3$ .

Collecting terms and restricting ourselves, as in section 5, to an isotropic law, the statistical condition (69) takes the form of a generalized Hamilton-Jacobi equation:

$$\bar{L} := \left( \frac{\partial S}{\partial t} + e\hat{\phi} \right) + \frac{1}{2m} \sum_j \left( \frac{\partial S}{\partial x_j} - \frac{e}{c} \hat{A}_j \right)^2 + \mu_i B_i = L_0. \quad (79)$$

The unknown function  $L_0$  must contain  $L'_0$  but may also contain other terms, let us write  $L_0 = L'_0 + \Delta L_0$ .

## 7. 'Missing' quantum spin terms from Fisher information

Let us summarize at this point what has been achieved so far. We have four coupled differential equations for our dynamic field variables  $\rho, S, \vartheta, \varphi$ . The first of these is the continuity equation (56), which is given, in terms of the present variables, by

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_l} \left[ \frac{\rho}{m} \left( \frac{\partial S}{\partial x_j} - \frac{e}{c} \hat{A}_j \right) \right] = 0. \quad (80)$$

The three other differential equations, the evolution equations (73), (74) and the generalized Hamilton-Jacobi equation (79), do not yet possess a definite mathematical form. They contain four unknown functions  $G_i, L_0$  which are constrained, but not determined, by (75), (78).

The simplest choice, from a formal point of view, is  $G_i = L_0 = 0$ . In this limit the present theory agrees with Schiller's field-theoretic (Hamilton-Jacobi) version, see Schiller (1962b), of the equations of motion of a classical dipole. This is a classical (statistical) theory despite the fact that it contains [see (63)] a number  $\hbar$ . But this classical theory is not realized in nature; at least not in the microscopic domain. The reason is that the simplest choice from a formal point of view is not the simplest choice from a physical point of view. The postulate of maximal simplicity (Ockham's razor) implies equal probabilities and the principle of maximal entropy in classical statistical physics. A similar principle which is able to 'explain' the nonexistence of classical physics (in the microscopic domain) is the principle of minimal Fisher information Frieden (2004). The relation between the two (classical and quantum-mechanical) principles has been discussed in detail in I.

The mathematical formulation of the principle of minimal Fisher information for the present problem requires a generalization, as compared to I, because we have now several fields with coupled time-evolution equations. As a consequence, the spatial integral (spatial average) over  $\rho(\bar{L} - L_0)$  in the variational problem (44) should be replaced by a space-time integral, and the variation should be performed with respect to all four variables. The problem can be written in the form

$$\delta \int dt \int d^3x \rho (\bar{L} - L_0) = 0 \quad (81)$$

$$E_a = 0, \quad a = S, \rho, \vartheta, \varphi, \quad (82)$$

where  $E_a = 0$  is a shorthand notation for the equations (80), (79), (74) (73). Eqs. (81), (82) require that the four Euler-Lagrange equations of the variational problem (81) agree with the differential equations (82). This imposes conditions for the unknown functions  $L_0, G_i$ . If the solutions of (81), (82) for  $L_0, G_i$  are inserted in the variational problem (81), the four

relations (82) become redundant and  $\rho(\bar{L} - L_0)$  becomes the Lagrangian density of our problem. Thus, Eqs. (81) and (82) represent a method to construct a Lagrangian.

We assume a functional form  $L_0(\chi_\alpha, \partial_k \chi_\alpha, \partial_k \partial_l \chi_\alpha)$ , where  $\chi_\alpha = \rho, \theta, \varphi$ . This means  $L_0$  does not possess an explicit  $x, t$ -dependence and does not depend on  $S$  (this would lead to a modification of the continuity equation). We further assume that  $L_0$  does not depend on time-derivatives of  $\chi_\alpha$  (the basic structure of the time-evolution equations should not be affected) and on spatial derivatives higher than second order. These second order derivatives must be taken into account but should not give contributions to the variational equations (a more detailed discussion of the last point has been given in I).

The variation with respect to  $S$  reproduces the continuity equation which is unimportant for the determination of  $L_0, G_i$ . Performing the variation with respect to  $\rho, \theta, \varphi$  and taking the corresponding conditions (79), (74) (73) into account leads to the following differential equations for  $L_0, G_1 \cos \varphi - G_2 \sin \varphi$  and  $G_3$ ,

$$-\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_i} \frac{\partial \rho L_0}{\partial \frac{\partial^2 \rho}{\partial x_k \partial x_i}} + \frac{\partial}{\partial x_k} \frac{\partial \rho L_0}{\partial \frac{\partial \rho}{\partial x_k}} - \rho \frac{\partial L_0}{\partial \rho} = 0 \quad (83)$$

$$-\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_i} \frac{\partial \rho L_0}{\partial \frac{\partial^2 \theta}{\partial x_k \partial x_i}} + \frac{\partial}{\partial x_k} \frac{\partial \rho L_0}{\partial \frac{\partial \theta}{\partial x_k}} - \frac{\partial \rho L_0}{\partial \theta} - \frac{\hbar \rho}{2} (G_1 \cos \varphi - G_2 \sin \varphi) = 0 \quad (84)$$

$$-\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_i} \frac{\partial \rho L_0}{\partial \frac{\partial^2 \varphi}{\partial x_k \partial x_i}} + \frac{\partial}{\partial x_k} \frac{\partial \rho L_0}{\partial \frac{\partial \varphi}{\partial x_k}} - \frac{\partial \rho L_0}{\partial \varphi} - \frac{\hbar}{2} \rho G_3 = 0. \quad (85)$$

The variable  $S$  does not occur in (83)-(85) in agreement with our assumptions about the form of  $L_0$ . It is easy to see that a proper solution (with vanishing variational contributions from the second order derivatives) of (83)-(85) is given by

$$L_0 = \frac{\hbar^2}{2m} \left[ \frac{1}{\sqrt{\rho}} \frac{\partial}{\partial \bar{x}} \frac{\partial}{\partial \bar{x}} \sqrt{\rho} - \frac{1}{4} \sin^2 \theta \left( \frac{\partial \varphi}{\partial \bar{x}} \right)^2 - \frac{1}{4} \left( \frac{\partial \theta}{\partial \bar{x}} \right)^2 \right] \quad (86)$$

$$\hbar G_1 \cos \varphi - \hbar G_2 \sin \varphi = \frac{\hbar^2}{2m} \left[ \frac{1}{2} \sin 2\theta \left( \frac{\partial \varphi}{\partial \bar{x}} \right)^2 - \frac{1}{\rho} \frac{\partial}{\partial \bar{x}} \rho \frac{\partial \theta}{\partial \bar{x}} \right] \quad (87)$$

$$\hbar G_3 = -\frac{\hbar^2}{2m} \frac{1}{\rho} \frac{\partial}{\partial \bar{x}} \left( \rho \sin \theta^2 \frac{\partial \varphi}{\partial \bar{x}} \right). \quad (88)$$

A new adjustable parameter appears on the r.h.s of (86)- (88) which has been identified with  $\hbar^2/2m$ , where  $\hbar$  is again Planck's constant. This second  $\hbar$  is related to the quantum-mechanical principle of maximal disorder. It is in the present approach not related in any obvious way to the previous "classical"  $\hbar$  which denotes the amplitude of a rotation; compare, however, the alternative derivation of spin in section 8.

The solutions for  $G_1, G_2$  may be obtained with the help of the second condition ( $\vec{G}\vec{s} = 0$ ) listed in Eq. (75). The result may be written in the form

$$\begin{aligned} G_1 &= \frac{\hbar}{2m} \frac{1}{\rho} \frac{\partial}{\partial \bar{x}} \rho \left( \frac{1}{2} \sin 2\theta \sin \varphi \frac{\partial \varphi}{\partial \bar{x}} - \cos \varphi \frac{\partial \theta}{\partial \bar{x}} \right) \\ G_2 &= \frac{\hbar}{2m} \frac{1}{\rho} \frac{\partial}{\partial \bar{x}} \rho \left( \frac{1}{2} \sin 2\theta \cos \varphi \frac{\partial \varphi}{\partial \bar{x}} + \sin \varphi \frac{\partial \theta}{\partial \bar{x}} \right). \end{aligned} \quad (89)$$

Eqs. (88) and (89) show that the first condition listed in (75) is also satisfied. The last condition is also fulfilled:  $L_0$  can be written as  $L'_0 + \Delta L_0$ , where

$$L'_0 = -\frac{\hbar^2}{8m} \left[ \sin^2 \vartheta \left( \frac{\partial \varphi}{\partial \vec{x}} \right)^2 - \left( \frac{\partial \vartheta}{\partial \vec{x}} \right)^2 \right], \quad \Delta L_0 = \frac{\hbar^2}{2m} \frac{1}{\sqrt{\rho}} \frac{\partial}{\partial \vec{x}} \frac{\partial}{\partial \vec{x}} \sqrt{\rho}, \quad (90)$$

and  $L'_0$  fulfills (78). We see that  $L'_0$  is a quantum-mechanical contribution to the rotational motion while  $\Delta L_0$  is related to the probability density of the ensemble (as could have been guessed considering the mathematical form of these terms). The last term is the same as in the spinless case [see (51)].

The remaining task is to show that the above solution for  $L_0$  does indeed lead to a (appropriately generalized) Fisher functional. This can be done in several ways. The simplest is to use the following result due to Reginatto (1998b):

$$\int d^3x (-\rho L_0) = \frac{\hbar^2}{8m} \sum_{j=1}^3 \int d^3x \sum_{k=1}^3 \frac{1}{\rho^{(j)}} \left( \frac{\partial \rho^{(j)}}{\partial x_k} \right)^2, \quad (91)$$

$$\rho^{(1)} := \rho \sin^2 \frac{\vartheta}{2} \cos^2 \frac{\varphi}{2}, \quad \rho^{(2)} := \rho \sin^2 \frac{\vartheta}{2} \sin^2 \frac{\varphi}{2}, \quad \rho^{(3)} := \rho \cos^2 \frac{\vartheta}{2}. \quad (92)$$

The functions  $\rho^{(j)}$  represent the probability that a particle is at space-time point  $x, t$  and  $\vec{s}$  points into direction  $j$ . Inserting (86) the validity of (91) may easily be verified. The r.h.s. of Eq. (91) shows that the averaged value of  $L_0$  represents indeed a Fisher functional, which completes our calculation of the 'quantum terms'  $L_0, G_i$ .

Summarizing, our assumption, that under certain external conditions four state variables instead of two may be required, led to a nontrivial result, namely the four coupled differential equations (80), (79), (74), (73) with  $L_0, G_i$  given by (86), (89), (88). The external condition which stimulates this splitting is given by a gauge field; the most important case is a magnetic field  $\vec{B}$  but other possibilities do exist (see below). These four differential equations are equivalent to the much simpler differential equation

$$\left( \frac{\hbar}{i} \frac{\partial}{\partial t} + e\phi \right) \hat{\psi} + \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \vec{x}} - \frac{e}{c} \vec{A} \right)^2 \hat{\psi} + \mu_B \vec{\sigma} \vec{B} \hat{\psi} = 0, \quad (93)$$

which is linear in the complex-valued two-component state variable  $\hat{\psi}$  and is referred to as Pauli equation (the components of the vector  $\vec{\sigma}$  are the three Pauli matrices and  $\mu_B = -e\hbar/2mc$ ). To see the equivalence one writes, see Takabayasi (1955), Holland (1995),

$$\hat{\psi} = \sqrt{\rho} e^{iS} \begin{pmatrix} \cos \frac{\vartheta}{2} e^{i\frac{\varphi}{2}} \\ \sin \frac{\vartheta}{2} e^{-i\frac{\varphi}{2}} \end{pmatrix}, \quad (94)$$

and evaluates the real and imaginary parts of the two scalar equations (93). This leads to the four differential equations (80), (79), (74) (73) and completes the present spin theory.

In terms of the real-valued functions  $\rho, S, \vartheta, \varphi$  the quantum-mechanical solutions (86), (88), (89) for  $L_0, G_i$  look complicated in comparison to the classical solutions  $L_0 = 0, G_i = 0$ . In terms of the variable  $\hat{\psi}$  the situation changes to the contrary: The quantum-mechanical equation becomes simple (linear) and the classical equation, which has been derived by Schiller (1962b), becomes complicated (nonlinear). The simplicity of the

underlying physical principle (principle of maximal disorder) leads to a simple mathematical representation of the final basic equation (if a complex-valued state function is introduced). One may also say that the linearity of the equations is a consequence of this principle of maximal disorder. This is the deeper reason why it has been possible, see Klein (2009), to derive Schrödinger's equation from a set of assumptions including linearity.

Besides the Pauli equation we found, as a second important result of our spin calculation, that the following local force is compatible with the statistical constraint:

$$\vec{F}^L + \vec{F}^I = e \left( \vec{E} + \frac{1}{c} \vec{v} \times \vec{B} \right) - \vec{\mu} \cdot \frac{\partial}{\partial \vec{x}} \vec{B}. \quad (95)$$

Here, the velocity field  $\vec{v}(x, t)$  and the magnetic moment field  $\vec{\mu}(x, t) = -(e/mc)\vec{s}(x, t)$  have been replaced by corresponding particle quantities  $\vec{v}(t)$  and  $\vec{\mu}(t)$ ; the dot denotes the inner product between  $\vec{\mu}$  and  $\vec{B}$ . The first force in (95), the Lorentz force, has been derived here from first principles without any additional assumptions. The same cannot be said about the second force which takes this particular form as a consequence of some additional assumptions concerning the form of the 'internal force'  $\vec{F}^R$  [see (71)]. In particular, the field appearing in  $\vec{F}^R$  was arbitrary as well as the proportionality constant (g-factor of the electron) and had to be adjusted by hand. It is well-known that in a relativistic treatment the spin term appears automatically if the potentials are introduced. Interestingly, this unity is not restricted to the relativistic regime. Following Arunsalam (1970) and Gould (1995) we report in the next section an alternative (non-relativistic) derivation of spin, which does not contain any arbitrary fields or constants - but is unable to yield the expression (95) for the macroscopic electromagnetic forces.

In the present treatment spin has been introduced as a property of an ensemble and not of individual particles. Similar views may be found in the literature, see Ohanian (1986). Of course, it is difficult to imagine the properties of an ensemble as being completely independent from the properties of the particles it is made from. The question whether or not a property 'spin' can be ascribed to single particles is a subtle one. Formally, we could assign a probability of being in a state  $i$  ( $i = 1, 2$ ) to a particle just as we assign a probability for being at a position  $\vec{x} \in R^3$ . But contrary to position, no classical meaning - and no classical measuring device - can be associated with the discrete degree of freedom  $i$ . Experimentally, the measurement of the 'spin of a single electron' is - in contrast to the measurement of its position - a notoriously difficult task. Such experiments, and a number of other interesting questions related to spin, have been discussed by Morrison (2007).

## 8. Spin as a consequence of a multi-valued phase

As shown by Arunsalam (1970), Gould (1995), and others, spin in non-relativistic QT may be introduced in exactly the same manner as the electrodynamic potentials. In this section we shall apply a slightly modified version of their method and try to derive spin in an alternative way - which avoids the shortcoming mentioned in the last section.

Arunsalam (1970) and Gould (1995) introduce the potentials by applying the well-known minimal-coupling rule to the free Hamiltonian. In the present treatment this is achieved by making the quantity  $S$  multi-valued. The latter approach seems intuitively preferable considering the physical meaning of the corresponding classical quantity. Let us first review the essential steps [see Klein (2009) for more details] in the process of creating potentials in the *scalar* Schrödinger equation:

- Chose a free Schrödinger equation with single-valued state function.
- 'Turn on' the interaction by making the state function multi-valued (multiply it with a multi-valued phase factor)
- Shift the multi-valued phase factor to the left of all differential operators, creating new terms (potentials) in the differential equation.
- Skip the multi-valued phase. The final state function is again single-valued.

Let us adapt this method for the derivation of spin (considering spin one-half only). The first and most important step is the identification of the free Pauli equation. An obvious choice is

$$\left[ \frac{\hbar}{i} \frac{\partial}{\partial t} + \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \vec{x}} \right)^2 + V \right] \bar{\psi} = 0, \quad (96)$$

where  $\bar{\psi}$  is a single-valued *two-component* state function; (96) is essentially a duplicate of Schrödinger's equation. We may of course add arbitrary vanishing terms to the expression in brackets. This seems trivial, but some of these terms may vanish *only* if applied to a single-valued  $\bar{\psi}$  and may lead to non-vanishing contributions if applied later (in the second of the above steps) to a multi-valued state function  $\bar{\psi}^{multi}$ .

In order to investigate this possibility, let us rewrite Eq. (96) in the form

$$\left[ \hat{p}_0 + \frac{1}{2m} \vec{\hat{p}} \vec{\hat{p}} + V \right] \bar{\psi} = 0, \quad (97)$$

where  $\hat{p}_0$  is an abbreviation for the first term of (96) and the spatial derivatives are given by

$$\vec{\hat{p}} = \hat{p}_k \vec{e}_k, \quad \hat{p}_k = \frac{\hbar}{i} \frac{\partial}{\partial x_k}. \quad (98)$$

All terms in the bracket in (97) are to be multiplied with a 2x2 unit-matrix  $E$  which has not be written down. Replace now the derivatives in (97) according to

$$\hat{p}_0 \Rightarrow \hat{p}_0 M_0, \quad \vec{\hat{p}} \Rightarrow \vec{\hat{p}}_k M_k, \quad (99)$$

where  $M_0, M_k$  are hermitian 2x2 matrices with constant coefficients, which should be constructed in such a way that the new equation agrees with (97) for single-valued  $\bar{\psi}$ , i.e. assuming the validity of the condition

$$(\hat{p}_i \hat{p}_k - \hat{p}_k \hat{p}_i) \bar{\psi} = 0. \quad (100)$$

This leads to the condition

$$M_0^{-1} M_i M_k = E \delta_{ik} + T_{ik}, \quad (101)$$

where  $T_{ik}$  is a 2x2 matrix with two cartesian indices  $i, k$ , which obeys  $T_{ik} = -T_{ki}$ . A solution of (101) is given by  $M_0 = \sigma_0, M_i = \sigma_i$ , where  $\sigma_0, \sigma_i$  are the four Pauli matrices. In terms of this solution, Eq. (101) takes the form

$$\sigma_i \sigma_k = \sigma_0 \delta_{ik} + i \epsilon_{ikl} \sigma_l. \quad (102)$$

Thus, an alternative free Pauli-equation, besides (96) is given by

$$\left[ \frac{\hbar}{i} \frac{\partial}{\partial t} + \frac{1}{2m} \left( \frac{\hbar}{i} \right)^2 \sigma_i \frac{\partial}{\partial x_i} \sigma_k \frac{\partial}{\partial x_k} + V \right] \bar{\psi} = 0. \quad (103)$$

The quantity in the bracket is the generalized Hamiltonian constructed by Arunsalam (1970) and Gould (1995). In the present approach gauge fields are introduced by means of a multi-valued phase. This leads to the same formal consequences as the minimal coupling rule but allows us to conclude that the second free Pauli equation (103) is *more appropriate* than the first, Eq. (96), because it is more general with regard to the consequences of multi-valuedness. This greater generality is due to the presence of the second term on the r.h.s. of (102). The second step is to turn on the multi-valuedness in Eq. (103),  $\bar{\psi} \Rightarrow \bar{\psi}^{multi}$ , by multiplying  $\bar{\psi}$  with a multi-valued two-by-two matrix. This matrix must be chosen in such a way that the remaining steps listed above lead to Pauli's equation (93) in presence of an gauge field. Since in our case the final result (93) is known, this matrix may be found by performing the inverse process, i.e. performing a singular gauge transformation  $\hat{\psi} = \Gamma \bar{\psi}^{multi}$  of Pauli's equation (93) from  $\hat{\psi}$  to  $\bar{\psi}^{multi}$ , which *removes* all electrodynamic terms from (93) and creates Eq. (103). The final result for the matrix  $\Gamma$  is given by

$$\Gamma = E \exp \left\{ i \frac{e}{\hbar c} \int^{x,t} [dx'_k A_k(x', t') - cd t' \phi(x', t')] \right\}, \quad (104)$$

and agrees, apart from the unit matrix  $E$ , with the multi-valued factor introduced previously [see (17) and (52)] leading to the electrodynamic potentials. The inverse transition from (103) to (93), i.e. the creation of the potentials and the Zeeman term, can be performed by using the inverse of (104).

The Hamiltonian (103) derived by Arunsalam (1970) and Gould (1995) shows that spin can be described by means of the same abelian gauge theory that leads to the standard quantum mechanical gauge coupling terms; no new adjustable fields or parameters appear. The only requirement is that the appropriate free Pauli equation (103) is chosen as starting point. The theory of Dartora & Cabrera (2008), on the other hand, started from the alternative (from the present point of view inappropriate) free Pauli equation (96) and leads to the conclusion that spin must be described by a non-abelian gauge theory.

As far as our derivation of non-relativistic QT is concerned we have now two alternative, and in a sense complementary, possibilities to introduce spin. The essential step in the second (Arunsalam-Gould) method is the transition from (96) to the equivalent free Pauli equation (103). This step is a remarkable short-cut for the complicated calculations, performed in the last section, leading to the various terms required by the principle of minimal Fisher information. The Arunsalam-Gould method is unable to provide the shape (95) of the corresponding macroscopic forces but is very powerful insofar as no adjustable quantities are required. It will be used in the next section to perform the transition to an arbitrary number of particles.

## 9. Transition to N particles as final step to non-relativistic quantum theory

In this section the present derivation of non-relativistic QT is completed by deriving Schrödinger's equation for an arbitrary number  $N$  of particles or, more precisely, for statistical ensembles of identically prepared experimental arrangements involving  $N$  particles.

In order to generalize the results of sections 2 and 5, a convenient set of  $n = 3N$  coordinates  $q_1, \dots, q_n$  and masses  $m_1, \dots, m_n$  is defined by

$$\begin{aligned} (q_1, q_2, q_3, \dots, q_{n-2}, q_{n-1}, q_n) &= (x_1, y_1, z_1, \dots, x_N, y_N, z_N), \\ (m_1, m_2, m_3, \dots, m_{n-2}, m_{n-1}, m_n) &= (m_1, m_1, m_1, \dots, m_N, m_N, m_N). \end{aligned} \quad (105)$$

The index  $I = 1, \dots, N$  is used to distinguish particles, while indices  $i, k, \dots$  are used here to distinguish the  $3N$  coordinates  $q_1, \dots, q_n$ . No new symbol has been introduced in (105) to distinguish the masses  $m_I$  and  $m_i$  since there is no danger of confusion in any one of the formulas below. However, the indices of masses will be frequently written in the form  $m_{(i)}$  in order to avoid ambiguities with regard to the summation convention. The symbol  $Q$  in arguments denotes dependence on all  $q_1, \dots, q_n$ . In order to generalize the results of section 8 a notation  $x_{I,k}$ ,  $\vec{x}_I$ , and  $m_I$  (with  $I = 1, \dots, N$  and  $k = 1, 2, 3$ ) for coordinates, positions, and masses will be more convenient.

The basic relations of section 2, generalized in an obvious way to  $N$  particles, take the form

$$\frac{\partial \rho(Q, t)}{\partial t} + \frac{\partial}{\partial q_k} \frac{\rho(Q, t)}{m_{(k)}} \frac{\partial S(Q, t)}{\partial q_k} = 0 \quad (106)$$

$$\frac{d}{dt} \bar{q}_k = \frac{1}{m_{(k)}} \bar{p}_k \quad (107)$$

$$\frac{d}{dt} \bar{p}_k = \overline{F_k(Q, t)} \quad (108)$$

$$\bar{q}_k = \int dQ \rho(Q, t) q_k \quad (109)$$

Here,  $S$  is a single-valued variable; the multi-valuedness will be added later, following the method of section 8.

The following calculations may be performed in complete analogy to the corresponding steps of section 2. For the present  $N$ -dimensional problem, the vanishing of the surface integrals, occurring in the course of various partial integrations, requires that  $\rho$  vanishes exponentially in arbitrary directions of the configuration space. The final conclusion to be drawn from Eqs. (106)- (109) takes the form

$$\sum_{j=1}^n \frac{1}{2m_{(j)}} \left( \frac{\partial S}{\partial q_j} \right)^2 + \frac{\partial S}{\partial t} + V = L_0, \quad \int dQ \frac{\partial \rho}{\partial q_k} L_0 = 0, \quad (110)$$

The remaining problem is the determination of the unknown function  $L_0$ , whose form is constrained by the condition defined in Eq. (110).

$L_0$  can be determined using again the principle of minimal Fisher information, see I for details. Its implementation in the present framework takes the form

$$\delta \int dt \int dQ \rho (L - L_0) = 0 \quad (111)$$

$$E_a = 0, \quad a = S, \rho, \quad (112)$$

where  $E_S = 0$ ,  $E_\rho = 0$  are shorthand notations for the two basic equations (110) and (106). As before, Eqs. (111), (112) represent a method to construct a Lagrangian. After determination of  $L_0$  the three relations listed in (111), (112) become redundant and (112) become the fundamental equations of the  $N$ -particle theory.

The following calculation can be performed in complete analogy to the case  $N = 1$  reported in section 5. All relations remain valid if the upper summation limit 3 is replaced by  $3N$ . This is also true for the final result, which takes the form

$$L_0 = \frac{\hbar^2}{4\rho} \left[ -\frac{1}{2\rho} \frac{1}{m_{(j)}} \frac{\partial \rho}{\partial q_j} \frac{\partial \rho}{\partial q_j} + \frac{1}{m_{(j)}} \frac{\partial^2 \rho}{\partial q_j \partial q_j} \right]. \quad (113)$$

If a complex-valued variable  $\psi$ , defined as in (52), is introduced, the two basic relations  $E_a = 0$  may be condensed into the single differential equation,

$$\left[ \frac{\hbar}{i} \frac{\partial}{\partial t} + \sum_{I=1}^N \frac{1}{2m_{(I)}} \left( \frac{\hbar}{i} \frac{\partial}{\partial x_{I,k}} \right) \left( \frac{\hbar}{i} \frac{\partial}{\partial x_{I,k}} \right) + V \right] \psi = 0, \quad (114)$$

which is referred to as  $N$ -particle Schrödinger equation, rewritten here in the more familiar form using particle indices. As is well-known, only approximate solutions of this partial differential equation of order  $3N + 1$  exist for realistic systems. The inaccessible complexity of quantum-mechanical solutions for large  $N$  is not reflected in the abstract Hilbert space structure (which is sometimes believed to characterize the whole of QT) but plays probably a decisive role for a proper description of the mysterious relation between QT and the macroscopic world.

Let us now generalize the Arunsalam-Gould method, discussed in section 8, to an arbitrary number of particles. We assume, that the considered  $N$ -particle statistical ensemble responds in  $2^N$  ways to the external electromagnetic field. This means we restrict ourselves again, like in section 6, 7 to spin one-half. Then, the state function may be written as  $\psi(x_1, s_1; \dots; x_I, s_I; \dots; x_N, s_N)$  where  $s_I = 1, 2$ . In the first of the steps listed at the beginning of section 8, a differential equation, which is equivalent to Eq. (114) for single-valued  $\psi$  but may give non-vanishing contributions for multi-valued  $\psi$ , has to be constructed. The proper generalization of Eq. (103) to arbitrary  $N$  takes the form

$$\left[ \frac{\hbar}{i} \frac{\partial}{\partial t} + \sum_{I=1}^N \frac{1}{2m_{(I)}} \left( \frac{\hbar}{i} \sigma_k^{(I)} \frac{\partial}{\partial x_{I,k}} \right) \left( \frac{\hbar}{i} \sigma_l^{(I)} \frac{\partial}{\partial x_{I,l}} \right) + V \right] \psi = 0, \quad (115)$$

where the Pauli matrices  $\sigma_k^{(I)}$  operate by definition only on the two-dimensional subspace spanned by the variable  $s_I$ . In the second step we perform the replacement

$$\psi \Rightarrow \exp \left\{ -\frac{i}{\hbar} \sum_{I=1}^N \frac{e_I}{c} \sum_{k=1}^3 \int^{\vec{x}_I, t} \left[ dx'_{I,k} A_k(x'_I, t') - c dt' \phi(x'_I, t') \right] \right\} \psi, \quad (116)$$

using a multi-valued phase factor, which is an obvious generalization of Eq. (104). The remaining steps, in the listing of section 8, lead in a straightforward way to the final result

$$\left[ \frac{\hbar}{i} \frac{\partial}{\partial t} + \sum_{I=1}^N e_I \phi(x_I, t) + \sum_{I=1}^N \sum_{k=1}^3 \frac{1}{2m_{(I)}} \left( \frac{\hbar}{i} \frac{\partial}{\partial x_{I,k}} - \frac{e_I}{c} A_k(x_I, t) \right)^2 + \sum_{I=1}^N \mu_B^{(I)} \sigma_k^{(I)} B_k(x_I, t) + V(x_1, \dots, x_N, t) \right] \psi = 0 \quad (117)$$

where  $\mu_B^{(I)} = -\hbar e_I / 2m_I c$  and  $\vec{B} = \text{rot} \vec{A}$ . The mechanical potential  $V(x_1, \dots, x_N, t)$  describes a general many-body force but contains, of course, the usual sum of two-body potentials as a special case. Eq. (117) is the  $N$ -body version of Pauli's equation and completes - in the sense discussed at the very beginning of this paper - the present derivation of non-relativistic QT.

## 10. The classical limit of quantum theory is a statistical theory

The classical limit of Schrödinger's equation plays an important role for two topics discussed in the next section, namely the interpretation of QT and the particular significance of potentials in QT; to study these questions it is sufficient to consider a single-particle ensemble described by a single state function. This 'classical limit theory' is given by the two differential equations

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_k} \frac{\rho}{m} \left( \frac{\partial S}{\partial x_k} - \frac{e}{c} A_k \right) = 0, \quad (118)$$

$$\frac{\partial S}{\partial t} + e\phi + \frac{1}{2m} \sum_k \left( \frac{\partial S}{\partial x_k} - \frac{e}{c} A_k \right)^2 + V = 0, \quad (119)$$

which are obtained from Eqs. (50) and (51) by performing the limit  $\hbar \rightarrow 0$ . The quantum mechanical theory (50) and (51) and the classical theory (118) and (119) show fundamentally the same mathematical structure; both are initial value problems for the variables  $S$  and  $\rho$  obeying two partial differential equations. The difference is the absence of the last term on the l.h.s. of (51) in the corresponding classical equation (119). This leads to a *decoupling* of  $S$  and  $\rho$  in (119); the identity of the classical object described by  $S$  is no longer affected by statistical aspects described by  $\rho$ .

The field theory (118), (119) for the two 'not decoupled' fields  $S$  and  $\rho$  is obviously very different from classical mechanics which is formulated in terms of trajectories. The fact that one of these equations, namely (119), agrees with the Hamilton-Jacobi equation, does not change the situation since the presence of the continuity equation (118) cannot be neglected. On top of that, even if it could be neglected, Eq. (119) would still be totally different from classical mechanics: In order to construct particle trajectories from the partial differential equation (119) for the field  $S(x, t)$ , a number of clearly defined mathematical manipulations, which are part of the classical theory of canonical transformations, see Greiner (1989), must be performed. The crucial point is that the latter theory is *not* part of QT and cannot be added 'by hand' in the limit  $\hbar \rightarrow 0$ . Thus, (118), (119) is, like QT, an *indeterministic* theory predicting not values of single event observables but only probabilities, which must be verified by ensemble measurements.

Given that we found a solution  $S(x, t)$ ,  $\rho(x, t)$  of (118), (119) for given initial values, we may ask which experimental predictions can be made with the help of these quantities. Using the fields  $\vec{p}(x, t)$ ,  $\vec{E}(x, t)$  defined by Eqs. (19), (18), the Hamilton-Jacobi equation (119) takes the form

$$\frac{\vec{p}^2(x, t)}{2m} + V(x, t) = \vec{E}(x, t), \quad (120)$$

The l.h.s. of (120) depends on the field  $\vec{p}$  in the same way as a classical particle Hamiltonian on the (gauge-invariant) kinetic momentum  $\vec{p}$ . We conclude that the field  $\vec{p}(x, t)$  describes a mapping from space-time points to particle momenta: If a particle (in an external electromagnetic field) is found at time  $t$  at the point  $x$ , then its kinetic momentum is given by  $\vec{p}(x, t)$ . This is not a deterministic prediction since we can not predict if a single particle will be or will not be at time  $t$  at point  $x$ ; the present theory gives only a probability  $\rho(x, t)$  for such an event. Combining our findings about  $\vec{p}(x, t)$  and  $x$  we conclude that the experimental prediction which can be made with the help of  $S(x, t)$ ,  $\rho(x, t)$  is given by the following phase

space probability density:

$$w(x, p, t) = \rho(x, t) \delta^{(3)}\left(p - \frac{\partial \tilde{S}(x, t)}{\partial x}\right). \quad (121)$$

Eq. (121) confirms our claim that the classical limit theory is a statistical theory. The one-dimensional version of (121) has been obtained before by means of a slightly different method in I. The deterministic element [realized by the delta-function shaped probability in (121)] contained in the classical statistical theory (118), (119) is *absent* in QT, see I.

Eqs. (118), (119) constitute the mathematically well-defined limit  $\hbar \rightarrow 0$  of Schrödinger's equation. Insofar as there is general agreement with regard to two points, namely that (i) 'non-classicality' (whatever this may mean precisely) is expressed by a nonzero  $\hbar$ , and that (ii) Schrödinger's equation is the most important relation of quantum theory, one would also expect general agreement with regard to a further point, namely that Eqs. (118), (119) present essentially (for a three-dimensional configuration space) the *classical limit* of quantum mechanics. But this is, strangely enough, not the case. With a few exceptions, see Van Vleck (1928), Schiller (1962a), Ballentine (1994), Shirai (1998), Klein (2009), most works (too many to be quoted) take it for granted that the classical limit of quantum theory is classical mechanics. The objective of papers like Rowe (1991), Werner & Wolff (1995), Landau (1996), Allori & Zanghi (2009) devoted to "...the classical limit of quantum mechanics.." is very often not the problem: "*what is the classical limit of quantum mechanics?*" but rather: "*how to bridge the gap between quantum mechanics and classical mechanics?*". Thus, the fact that classical mechanics is the classical limit of quantum mechanics is considered as *evident* and any facts not compatible with it - like Eqs. (118), (119) - are denied.

What, then, is the reason for this widespread denial of reality? One of the main reasons is the principle of reductionism which still rules the thinking of most physicists today. The reductionistic ideal is a hierarchy of physical theories; better theories have an enlarged domain of validity and contain 'inferior' theories as special cases. This principle which has been extremely successful in the past *dictates* that classical mechanics is a special case of quantum theory. Successful as this idea might have been during a long period of time it is not necessarily universally true; quantum mechanics and classical mechanics describe different domains of reality, both may be true in their own domains of validity. Many phenomena in nature indicate that the principle of reductionism (alone) is insufficient to describe reality, see Laughlin & Pines (2000). Releasing ourselves from the metaphysical principle of reductionism, we accept that the classical limit of quantum mechanics for a three-dimensional configuration space is the statistical theory defined by Eqs. (118), (119). It is clear that this theory is not realized in nature (with the same physical meaning of the variables) because  $\hbar$  is different from zero. But this is a different question and does not affect the conclusion.

## 11. Extended discussion

In this paper it has been shown, continuing the work of I, that the basic differential equation of non-relativistic QT may be derived from a number of clearly defined assumptions of a statistical nature. Although this does not exclude the possibility of other derivations, we consider this success as a strong argument in favor of the statistical interpretation of QT.

This result explains also, at least partly, the success of the canonical quantization rules (1). Strictly speaking, these rules have only be derived for a particular (though very important) special case, the Hamiltonian. However, one can expect that (1) can be verified for all

meaningful physical observables<sup>4</sup>. On the other hand, it cannot be expected that the rules (1) hold for arbitrary functions of  $x, p$ ; each case has to be investigated separately. Thus, the breakdown of (1), as expressed by Groenewold's theorem, is no surprise.

The fundamental Ehrenfest-like relations of the present theory establish [like the formal rules (1)] a *correspondence* between particle mechanics and QT. Today, philosophical questions concerning, in particular, the 'reality' of particles play an important role in the thinking of some physicists. So: 'What is this theory about.. ?' While the present author is no expert in this field, the concept of *indeterminism*, as advocated by the philosopher Popper (1982), seems to provide an appropriate philosophical basis for the present work.

The present method to introduce gauge fields by means of a multi-valued dynamic variable ('phase function') has been invented many years ago but leads, in the context of the present statistical theory, nevertheless to several new results. In particular, it has been shown in section 3, that only the Lorentz force can exist as fundamental macroscopic force if the statistical assumptions of section 2 are valid. It is the only force (in the absence of spin effects, see the remarks below) that can be incorporated in a 'standard' differential equation for the dynamical variables  $\rho, S$ . The corresponding terms in the statistical field equations, *representing* the Lorentz force, are given by the familiar gauge (minimal) coupling terms containing the potentials. The important fact that all forces in nature follow this 'principle of minimal coupling' is commonly explained as a consequence of local gauge invariance. The present treatment offers an alternative explanation.

Let us use the following symbolic notation to represent the relation between the local force and the terms representing its action in a statistical context:

$$\Phi, \vec{A} \Rightarrow e\vec{E} + \frac{e}{c}\vec{v} \times \vec{B}. \quad (122)$$

The fields  $\vec{E}$  and  $\vec{B}$  are uniquely defined in terms of the potentials  $\phi$  and  $\vec{A}$  [see (23)] while the inverse is not true. Roughly speaking, the local fields are 'derivatives' of the potentials - and the potentials are 'integrals' of the local field; this mathematical relation reflects the physical role of the potentials  $\phi$  and  $\vec{A}$  as statistical representatives of the the local fields  $\vec{E}$  and  $\vec{B}$ , as well as their non-uniqueness. It might seem that the logical chain displayed in (122) is already realized in the classical treatment of a particle-field system, where potentials have to be introduced in order to construct a Lagrangian, see e.g. Landau & Lifshitz (1967). However, in this case, the form of the local force is not derived but postulated. The present treatment 'explains' the form of the Lagrangian - as a consequence of the basic assumptions listed in section 2.

The generalization of the present theory to spin, reported in sections 6 and 7, leads to a correspondence similar to Eq. (122), namely

$$\vec{\mu}\vec{B} \rightarrow \vec{\mu} \cdot \frac{\partial}{\partial \vec{x}} \vec{B}. \quad (123)$$

The term linear in  $\vec{B}$ , on the l.h.s. of (123), plays the role of a 'potential' for the local force on the r.h.s. The points discussed after Eq. (122) apply here as well [As a matter of fact we consider  $\vec{B}$  as a unique physical quantity; it would not be unique if it would be defined in terms of the tensor on the r.h.s. of (123)]. We see here a certain analogy between gauge and spin interaction terms. Unfortunately, the derivation of the spin force on the r.h.s. of (123) requires - in contrast to the Lorentz force - additional assumptions (see the remarks in sections 7, 8).

<sup>4</sup> As indicated by preliminary calculations of the angular momentum relation corresponding to Eq. (8)

Our notation for potentials  $\phi$ ,  $\vec{A}$ , fields  $\vec{E}$ ,  $\vec{B}$ , and parameters  $e$ ,  $c$  suggests that these quantities are electro-dynamical in nature. However, this is not necessarily true. By definition, the fields  $\vec{E}$ ,  $\vec{B}$  obey four equations (the homogeneous Maxwell equations), which means that additional conditions are required in order to determine these six fields. The most familiar possibility is, of course, the second pair of Maxwell's equations. A second possible realization for the fields  $\vec{E}$ ,  $\vec{B}$  is given by the *inertial* forces acting on a mass  $m$  in an arbitrarily accelerated reference frame, see Hughes (1992). The inertial gauge field may also lead to a spin response of the ensemble; such experiments have been proposed by Mashhoon & Kaiser (2006). It is remarkable that the present theory establishes a (admittedly somewhat vague) link between the two extremely separated physical fields of inertia and QT.

It is generally assumed that the electrodynamic potentials have a particular significance in QT which they do not have in classical physics. Let us analyze this statement in detail. The first part of the statement, concerning the significance of the potentials, is of course true. The second part, asserting that in classical physics all external influences *can* be described solely in terms of field strengths, is wrong. More precisely, it is true for classical mechanics but not for classical physics in general. A counterexample - a theory belonging to classical physics but with potentials playing an indispensable role - is provided by the classical limit (118),(119) of Schrödinger's equation. In this field theory the potentials play an indispensable role because (in contrast to particle theories, like the canonical equations) no further derivatives of the Hamiltonian, which could restore the fields, are to be performed. This means that the significance of the potentials is not restricted to quantum theory but rather holds for the whole class of *statistical* theories discussed above, which contains both quantum theory and its classical limit theory as special cases. This result is in agreement with the statistical interpretation of potentials proposed in section 3.

The precise characterization of the role of the potentials is of particular importance for the interpretation of the Aharonov-Bohm effect. The 'typical quantum-mechanical features' observed in these phase shift experiments should be identified by comparing the quantum mechanical results not with classical mechanics but with the predictions of the classical statistical theory (118), (119). The predictions of two statistical theories, both of which use potentials to describe the influence of the external field, have to be compared.

The limiting behavior of Schrödinger's equation as  $\hbar \rightarrow 0$ , discussed in section 10, is very important for the proper interpretation of QT. The erroneous belief (wish) that this limit can (must) be identified with classical mechanics is closely related to the erroneous belief that QT is able to describe the dynamics of individual particles. In this respect QT is obviously an *incomplete* theory, as has been pointed out many times before, during the last eighty years, see e.g. Einstein (1949), Margenau (1935), Ballentine (1970), Held (2008). Unfortunately, this erroneous opinion is historically grown and firmly established in our thinking as shown by the ubiquitous use of phrases like 'the wave function of the electron'. But it is clear that an erroneous identification of the domain of validity of a physical theory will automatically create all kinds of mysteries and unsolvable problems - and this is exactly what happens. Above, we have identified one of the (more subtle) problems of this kind, concerning the role of potentials in QT, but many more could be found. Generalizing the above argumentation concerning potentials, we claim that characteristic features of QT cannot be identified by comparison with classical mechanics. Instead, quantum theory should be compared with its classical limit, which is in the present 3D-case given by (118), (119) - we note in this context that several 'typical' quantum phenomena have been explained by Kirkpatrick (2003) in terms of classical probability theory. One has to compare the solutions of the classical, nonlinear

equations (118), (119) with those of the quantum mechanical, linear equations, (50), (51), in order to find out which 'typical quantum-mechanical features' are already given by statistical (nonlocal) correlations of the classical limit theory and which features are really quantum-theoretical in nature - related to the nonzero value of  $\hbar$ .

## 12. Summary

In the present paper it has been shown that the method reported in I, for the derivation of Schrödinger's equation, can be generalized in such a way that essentially all aspects of non-relativistic QT are taken into account. The success of this derivation from statistical origins is interpreted as an argument in favor of the SI. The treatment of gauge fields and spin in the present statistical framework led to several remarkable new insights. We understand now why potentials (and not local fields) occur in the field equations of QT. The non-uniqueness of the potentials and the related concept of gauge invariance is not a mystery any more. Spin is derived as a kind of two-valuedness of a statistical ensemble. The local forces associated with the gauge potentials, the Lorentz force and the force experienced by a particle with magnetic moment, can also be derived. Apart from some open questions in the area of non-relativistic physics, a major problem for future research is a relativistic generalization of the present theory.

## 13. References

- Aharonov, Y. & Bohm, D. (1959). Significance of electromagnetic potentials in quantum theory, *Phys. Rev.* 115(3): 485.
- Ali, A. H. (2009). The ensemble quantum state of a single particle, *Int. J. Theor. Phys.* 48: 194–212.
- Allori, V. & Zanghi, N. (2009). On the classical limit of quantum mechanics, *Foundations of Physics* 39: 20–32.
- Anderson, P. W. (1972). More is different, *Science* 177(4047): 393–396.
- Arunsalam, V. (1970). Hamiltonians and wave equations for particles of spin 0 and spin  $\frac{1}{2}$  with nonzero mass, *Am. J. Phys.* 38: 1010–1022.
- Ballentine, L. E. (1970). The statistical interpretation of quantum mechanics, *Reviews of Modern Physics* 42: 358–381.
- Ballentine, L. E. (1994). Inadequacy of Ehrenfest's theorem to characterize the classical regime, *Physical Review A* 50: 2854–2859.
- Belinfante, F. J. (1975). *Measurements and Time Reversal in Objective Quantum Theory*, Pergamon Press, Oxford.
- Belinfante, F. J. (1978). Can individual elementary particles have individual properties?, *Am. J. Phys.* 46(4): 329–336.
- Blokhintsev, D. I. (1964). *Quantum Mechanics*, Reidel, Dordrecht.
- Bocchieri, P. & Loinger, A. (1978). Nonexistence of the Aharonov-Bohm effect, *Nuovo Cimento* 47A: 475–482.
- Dartora, C. A. & Cabrera, G. G. (2008). Magnetization, spin current, and spin-transfer torque from SU(2) local gauge invariance of the nonrelativistic Pauli-Schrödinger theory, *Physical Review B* 78: 012403.
- Dirac, P. A. M. (1931). Quantised singularities in the electromagnetic field, *Proc. R. Soc. London, Ser. A* 133: 60–72.
- Einstein, A. (1936). Physics and reality, *J. Franklin Inst.* 221: 349.

- Einstein, A. (1949). *Reply to Criticism*, Harper and Row, New York, p. 665.
- Frieden, B. R. (1989). Fisher information as the basis for the Schrödinger wave equation, *Am. J. Phys.* 57(11): 1004–1008.
- Frieden, B. R. (2004). *Science from Fisher Information, a Unification*, Cambridge University Press, Cambridge.
- Gotay, M. J. (1999). On the Groenewold-Van Hove problem for  $R^{2n}$ , *J. Math. Phys.* 40: 2107–2116.
- Gould, R. J. (1995). The intrinsic magnetic moment of elementary particles, *Am. J. Phys.* 64: 597–601.
- Greiner, W. (1989). *Classical Mechanics, Systems of particles and Hamiltonian dynamics*, Springer, New York.
- Groenewold, H. J. (1946). On the principles of elementary quantum mechanics, *Physica* 12: 405–460.
- Hall, M. J. & Reginatto, M. (2002a). Quantum mechanics from a Heisenberg-type equality, *Fortschr. Phys.* 50: 5–7.
- Hall, M. J. & Reginatto, M. (2002b). Schrödinger equation from an exact uncertainty principle, *J. Phys. A* 35: 3289–3303.
- Hall, M. J. W. (2005). Exact uncertainty approach in quantum mechanics and quantum gravity, *Gen. Relativ. Grav.* 37: 1505–1515.
- Held, C. (2008). Axiomatic quantum mechanics and completeness, *Foundations of Physics* 38: 707–732.
- Holland, P. R. (1995). *The quantum theory of motion*, Cambridge University Press, Cambridge, U.K.
- Hughes, R. J. (1992). On Feynman's proof of the Maxwell equations, *Am. J. Phys.* 60: 301–306.
- Kaempfer, F. A. (1965). *Concepts in Quantum Mechanics*, Academic Press, New York.
- Kemble, E. C. (1929). The general principles of quantum mechanics. part I, *Rev. Mod. Phys.* 1: 157–215.
- Kirkpatrick, K. A. (2003). 'Quantal' behavior in classical probability, *Foundations of Physics Letters* 16: 199–224.
- Klein, U. (1981). Comment on 'Condition for nonexistence of Aharonov-Bohm effect', *Physical Review D* 23: 1463–1465.
- Klein, U. (2009). Schrödinger's equation with gauge coupling derived from a continuity equation, *Foundations of physics* 39: 964.
- Klein, U. (2011). The statistical origins of quantum mechanics, *Physics Research International* Article ID 808424.  
URL: <http://downloads.hindawi.com/journals/phys/2010/808424.pdf>
- Kobe, D. H. & Yang, K. (1985). Gauge transformation of the time-evolution operator, *Phys. Rev. A* 32: 952–958.
- Krüger, T. (2004). An attempt to close the Einstein-Podolsky-Rosen debate, *Can. J. Phys.* 82: 53–65.
- Landau, L. D. & Lifshitz, E. M. (1967). *Classical theory of fields*, Vol. II of *Course of theoretical physics*, 5 edn, Pergamon Press, Oxford. Translation from Russian, Nauka, Moscow, 1973.
- Landau, L. J. (1996). Macroscopic observation of a quantum particle in a slowly varying potential - on the classical limit of quantum theory, *Annals of Physics* 246: 190–227.
- Laughlin, R. B. (2005). *A Different Universe*, Basic Books, Cambridge.

- Laughlin, R. B. & Pines, D. (2000). The theory of everything, *Proc. Natl. Acad. Sci. USA* 97: 28–31.
- Lee, Y. C. & Zhu, W. (1999). The principle of minimal quantum fluctuations for the time-dependent Schrödinger equation, *J. Phys. A* 32: 3127–3131.
- Levy-Leblond, J. M. (1967). Nonrelativistic particles and wave equations, *Comm. Math. Physics* 6: 286–311.
- London, F. (1927). Quantenmechanische Deutung der Theorie von Weyl, *Z. Phys.* 42: 375.
- Margenau, H. (1935). Quantum-mechanical description, *Physical Review* 49: 24–242.
- Margenau, H. (1963). Measurements in quantum mechanics, *Annals of Physics* 23: 469–485.
- Mashhoon, B. & Kaiser, H. (2006). Inertia of intrinsic spin, *Physica B* 385-386: 1381–1383.
- Morrison, M. (2007). Spin: All is not what it seems, *Studies in history and philosophy of modern physics* 38: 529–557.
- Motz, L. (1962). Quantization and the classical Hamilton-Jacobi equation, *Phys. Rev.* 126: 378–382.
- Newman, R. G. (1980). Probability interpretation of quantum mechanics, *Am. J. Phys.* 48: 1029–1034.
- Ohanian, H. C. (1986). What is spin ?, *Am. J. Phys.* 54: 500–505.
- Peshkin, M. & Tonomura, A. (1989). *The Aharonov-Bohm Effect*, Lecture Notes in Physics, Springer Verlag, Berlin.
- Pippard, A. B. (1986). The interpretation of quantum mechanics, *Eur. J. Phys.* 7: 43–48.
- Popper, K. R. (1982). *The Open Universe - an Argument for Indeterminism*, Rowman and Littlefield, Totowa.
- Reginatto, M. (1998a). Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information, *Phys. Rev. A* 58: 1775–1778.
- Reginatto, M. (1998b). Derivation of the Pauli equation using the principle of minimum Fisher information, *Physics Letters A* 249: 355–357.
- Rosen, N. (1964). The relation between classical and quantum mechanics, *Am. J. Phys.* 32: 597–600.
- Ross-Bonney, A. A. (1975). Does god play dice ? - a discussion of some interpretations of quantum mechanics, *Nuovo Cimento* 30 B: 55.
- Rowe, E. G. P. (1991). Classical limit of quantum mechanics (electron in a magnetic field), *Am. J. Phys.* 59: 1111–1117.
- Roy, S. M. (1980). Condition for nonexistence of Aharonov-Bohm effect, *Physical Review Letters* 44(3): 111–114.
- Schiller, R. (1962a). Quasi-classical theory of the nonspinning electron, *Phys. Rev.* 125(3): 1100–1108.
- Schiller, R. (1962b). Quasi-Classical Theory of the Spinning Electron, *Phys. Rev.* 125(3): 1116–1123.
- Schrödinger, E. (1926). Quantisierung als Eigenwertproblem, Erste Mitteilung, *Annalen der Physik* 79: 361.
- Shirai, H. (1998). Reinterpretation of quantum mechanics based on the statistical interpretation, *Foundations of Physics* 28: 1633–1662.
- Synge, J. L. (1960). *Classical Dynamics*, in *Encyclopedia of Physics: Principles of Classical Mechanics and Field theory*, Springer, Berlin, pp. 1–223.
- Syska, J. (2007). Fisher information and quantum-classical field theory: classical statistics similarity, *phys. stat. sol.(b)* 244: 2531–2537.

- Takabayasi, T. (1955). The vector representation of spinning particles in the quantum theory I, *Progress in Theoretical Physics*. 14: 283–302.
- Toyozawa, Y. (1992). Theory of measurement, *Progress of Theoretical Physics* 87: 293–305.
- Tschudi, H. R. (1987). On the statistical interpretation of quantum mechanics, *Helvetica Physica Acta* 60: 363–383.
- Van Vleck, J. H. (1928). The correspondence principle in the statistical interpretation of quantum mechanics, *Proc. Natl. Acad. Sci. U.S.* 14: 178–188.
- Werner, R. F. & Wolff, M. P. H. (1995). Classical mechanics as quantum mechanics with infinitesimal  $\hbar$ , *Physics Letters A* 202: 155–159.
- Weyl, H. (1929). Elektron und Gravitation I, *Z. Phys.* 56: 330–352.
- Young, R. H. (1980). Quantum mechanics based on position, *Foundations of Physics* 10: 33–56.

# The Quantum Completeness Problem

Carsten Held

*Philosophisches Seminar, Universität Erfurt, Erfurt  
Germany*

## 1. Introduction

Quantum mechanics (QM) is complete in a precise sense. It cannot be supplemented by more informative descriptions of physical systems given certain reasonable assumptions; this is what the no-hidden-variables proofs show. Theorems of Kochen-Specker-type (KS theorems) crucially employ an assumption of context-independence of the observables considered while those of Bell-type theorems use an assumption of locality. Since locality amounts to context-independence of local observables the former theorems can be considered as more general than the latter. What exactly do these theorems show in terms of physics? In which sense do they prove QM to be complete?

The standard answer to these questions is indeed older than the theorems themselves. It was given by the inventors of QM, notably von Neumann (who himself devised a no-hidden-variables argument) and Dirac. It says that the mathematical entity representing the maximal QM information we can have about a system  $S$  represents all the physical properties that  $S$  has. This is an informative physical interpretation of the theorems. Since it embodies the idea that the QM information about  $S$  is the complete representation of its physical state it may be called the *completeness assumption*.

In this paper, I will try to show that the completeness assumption is not the correct interpretation of the KS theorems for an ultimately simple reason: it cannot be squared with QM itself. The argument proceeds as follows. Initially, I explicate four properties of QM – properties not mentioned but represented in standard axiomatisations – that will drive the argument, and moreover cast the completeness assumption into a precise form: COMP (basically a weakened version of the well-known eigenstate-eigenvalue link) (sec.2). Then I consider the central equation of the QM statistical algorithm: the trace formula. From the QM properties I conclude that the formula must be explicated in one of two ways. Neither option, however, can be harmonised with both QM and a general probability principle (sec.3). I discuss whether the argument is just an involved form of the measurement problem or else unduly neglects the notion of QM measurement – both with a negative result (sec.4). I consider a fundamental objection that indeed circumvents the argument and show that it violates one of the four QM properties (sec.5). Finally, I briefly consider what the KS theorems tell us about QM given that the argument against COMP is correct (sec.6).

## 2. Preliminaries

### 2.1 Properties of quantum mechanics

Initially, we must identify the features of QM needed for the argument. The theory can be introduced in various ways, hence comes in different versions but the four properties we

need are shared by all of them. I choose a standard formulation of QM, using the Schrödinger picture and density operators, and illustrate everything from there but the argument in no way hinges on this choice. Suppose that the standard formulation comes as a set of axioms. The number of the axioms and their individual details will differ but the axioms will rule three important things. First, there will be an axiom fixing the mathematical entity representing the state of a QM system  $S$  for a fixed time  $t_0$ : a density operator  $W(t_0)$  on an appropriate Hilbert space. Second, there will be an axiom ruling the time evolution of state  $W(t_0)$ : an equation  $W(t) = U(t) W(t_0) U(t)^{-1}$  where  $U(t) = \exp[-iHt]$  is a function of time, and  $H$ , possibly time-dependent ( $H = H(t)$ ), is a unitary operator representing  $S$ 's total energy. Finally, there will be an axiom decreeing how QM generates predictions. It will either be or directly imply a statistical algorithm. Assuming a nondegenerate and discrete observable  $A$  on  $S$  with eigenvalues  $a_i$  (with  $i = 1, 2, \dots$ ) and  $a_k$  an arbitrary one of the  $a_i$ , one form of the algorithm will include an equivalent of the following equation:  $p(a_k) = \text{Tr}(W P_{a_k})$ , called the *trace formula*. Note that  $p(a_k)$ , on the left abbreviates  $p([A] = a_k)$  where  $[A] = a_k'$  means 'the value of  $A$  on  $S$  equals  $a_k'$ ' or equivalently ' $S$  has value  $a_k$  of  $A$ '.

This little sketch of one form of QM illustrates four crucial properties that it shares with all other forms. These properties are independent of the axiomatisation details. In particular, they are independent of: a) which specific form the statistical algorithm (often called the Born Rule) takes, wherein the trace formula ultimately resides; b) whether or not we have projection in QM. The properties are dependent on the Schrödinger picture but have counterparts in the Heisenberg picture such that an argument parallel to the one given here can be constructed. (I will, however, not try to show this and leave the Heisenberg picture mostly out of discussion.) The first property is this:

**Time Parameter Uniqueness:** QM contains a unique time-parameter  $t$ .

The time-parameter in QM has a clear role to play. Given an initial value  $t_0$  of  $t$ , variable values of  $t$  relate states to  $W(t_0)$  by  $W(t) = U(t) W(t_0) U(t)^{-1}$ . The variable  $t$  can be assumed to take values ad libitum, but there is no second time *parameter* in QM. In particular, QM leaves no conceptual room for re-interpreting two occurrences of  $t$  in any QM equation as two independent parameters ( $t$  and  $t'$ , say) capable of taking different values. (The evolution equation  $W(t) = U(t) W(t_0) U(t)^{-1}$  where  $t$ , if it takes a value, must take the same in all three places is one trivial example of this property.)

The second property is a trivial consequence of the Schrödinger picture:

**Time Reference:** If QM ascribes to  $S$  a state  $W$ , then  $W = W(t)$ , i.e.  $W$  is time-dependent throughout.

Given  $W = W(t_0)$ , Time Reference is satisfied. Up to a measurement, all other states are produced from  $W(t_0)$  via time evolution, so are time-dependent by construction. The situation will be different upon measurement – but only if we have a form of QM where unitary evolution is interrupted: projection. As I said, the argument below will be independent of whether or not QM contains a projection postulate because it employs just time-evolution up to a “first” measurement. But note that Time Reference is also respected, given projection, in the situation during or after a measurement. If we accept projection we assume a “collapse” of  $S$ 's state upon measurement of  $W(t_1)$ . There will be an immediate problem about specifying the time-reference for collapsed state  $W'$ . But collapse interpretations (like the GRW theory) usually assume that a collapsed state  $W'$  can be an input for a new unitary evolution. For this to be possible, the collapsed state must also be time-dependent:  $W' = W'(t_2)$  (where  $t_1, t_2$  are two different values of the same parameter  $t$ ).

**State Uniqueness:** If QM attributes a state  $W(t)$  to  $S$ , then for all  $W'(t)$  attributed to  $S$ ,  $W(t) = W'(t)$ .

This property again is a trivial consequence of our using the Schrödinger picture. QM would break down if the state attributed to  $S$  per time were not unique. This would simply result in conflicting inputs for the trace formula, hence lead to conflicting predictions.

The last property relates to the trace formula  $p(a_k) = \text{Tr}(W P_{a_k})$  and may be called Generality.

**Generality:** Given an arbitrary state  $W$  of  $S$  and an arbitrary observable  $A$ , it is the case that  $p(a_k) = \text{Tr}(W P_{a_k})$ .

The emphasis here is on *arbitrary*  $W$  and  $A$ . When representing a physical system in QM, we pick a Hilbert space whose dimension depends on the observables we are interested in. Given that such a choice has been made, Generality simply says that the trace formula holds good for *all* states and observables on that space.

These four properties suffice to create a conflict with the completeness assumption that I call the *completeness problem*. Plainly, the properties are shared by all versions of QM using the Schrödinger picture and density operators. Moreover, they have counterparts in versions making other choices. I thus assume that QM as such leads to the completeness problem.

## 2.2 Completeness

KS theorems show the impossibility of assigning values to QM observables given essentially two constraints. These theorems translate the idea of assigning properties (values of observables) to  $S$  into the task of assigning values to operators forming a certain algebra. Because the algebra is non-Boolean it is possible to find finite sets of operators such that not all of them can consistently be assigned the real numbers that are the allowed values of the associated observables. One condition of the arguments is that the algebra is based on a Hilbert space  $\mathcal{H}$  with  $\dim \mathcal{H} > 2$ . This condition is unquestionable because most QM systems require representation on a Hilbert space  $\mathcal{H}$  with at least  $\dim \mathcal{H} = 3$ . But there are two substantial assumptions going into the theorems that can be questioned: (1) Algebraic relations among operators representing QM observables are mirrored exactly in the associated values; this is a condition known in the literature as Functional Composition. (2) There is a one-one correspondence between observables and the operators representing them; since this can be interpreted as an assumption of independence of any single observable from the observables considered in conjunction with it this condition is generally called Noncontextuality.<sup>1</sup> Given these two assumptions, relations among operators translate straightforwardly into relations among values of observables that cannot simultaneously be satisfied.<sup>2</sup> Trivially, all the sets of operators are such that they cannot all have one eigenstate in common.

Generally, the completeness assumption is seen as being embodied in the so-called eigenstate-eigenvalue link (EE)<sup>3</sup> stating that  $S$  has  $a_k$  if and only if it is in the pertaining eigenstate  $P_{a_k}$ . Following this policy we can, as a first stab, rule this equivalence:

<sup>1</sup> See Redhead (1987, ch. 5) for a (by now classic) discussion.

<sup>2</sup> See Kochen and Specker (1967). For a simple and unified form of the Bell and Kochen-Specker theorems for two and three particles, see Mermin (1990); see also Held (2009). For the simplest Kochen-Specker type arguments in three and more dimensions see the references in Bub (1999, 118-9).

<sup>3</sup> For the expression and a clear formulation see Fine (1973), 20. The link was first proposed, however, by Dirac (1930/1958, 46-7) and von Neumann (1932/1955, 253); see Bub (1999, 29) for a discussion.

$$[A] = a_k \text{ if and only if } S \text{ is in } \mathbf{P}_{ak}. \quad (*)$$

This is insufficient, for a formal and a physical reason. In the Schrödinger picture due to Time Reference, the state is time-dependent:  $\mathbf{P}_{ak} = \mathbf{P}_{ak}(t)$ . So in (\*) there is an implicit time reference on the right. Now, while it is logically possible to have no matching time reference on the left this would clearly leave (\*) incomplete. Moreover, the statement that  $[A] = a_k$  is physically void if not referring to time concretely. In particular, we will want to tell stories such that at certain times it is true of S that  $[A] = a_k$ , but not at certain others. The only natural way to make (\*) meaningful formally and physically is explication on both sides:

$$[A] = a_k \text{ at } t \text{ if and only if } S \text{ is in } \mathbf{P}_{ak}(t). \quad (\text{EE})$$

This is the full-fledged EE.<sup>4</sup> One of its directions will suffice to make sense of the KS impossibility results. The 'only if'-direction is:

$$\text{If } S \text{ is not in } \mathbf{P}_{ak}(t), \text{ then } \neg [A] = a_k \text{ at } t.$$

Now, we know (by State Uniqueness) that the state  $W(t)$  of S is unique, i.e. that if S is in a state  $W(t) \neq \mathbf{P}_{ak}$ , then S is not in  $\mathbf{P}_{ak}(t)$ . Hence, from the 'only if' direction of EE:

$$\text{If } S \text{ is in a state } W(t) \neq \mathbf{P}_{ak}, \text{ then } \neg [A] = a_k \text{ at } t.$$

Let's weaken this one step further - to see what is minimally needed to make sense of the results.<sup>5</sup> Assume only that when S is in a *pure* state other than  $\mathbf{P}_{ak}(t)$ , then it does not have  $a_k$  at t. Identify this as our completeness assumption (COMP):

$$\text{If } S \text{ is in a pure state } W(t) \neq \mathbf{P}_{ak}, \text{ then } \neg [A] = a_k \text{ at } t. \quad (\text{COMP})$$

COMP makes eminent physical sense of the KS impossibility results. Every KS set includes observables that do not all have one eigenstate in common. Whatever S's state - if it's a pure state then the set contains observables for which it is not an eigenstate. These observables do not have values while the others may have. The contradiction that would arise from attributing values to all observables is avoided.

The crucial role of time in this reasoning should be made as explicit as possible. The starting point is that certain relations among values  $[A]$ ,  $[B]$ ,  $[C]$  ... of certain observables  $A$ ,  $B$ ,  $C$  ... (that do not all have one eigenstate in common) cannot *jointly* be satisfied. We want to concretize this in terms of physics in the sense that  $A$ ,  $B$ ,  $C$  ... cannot all *simultaneously* have values. Suppose that we want to express that  $A$  has a value while  $B$  and  $C$  have none, *at the same time*  $t$ :  $[A] = a_k$  at  $t$ ,  $\neg [A] = a_i$  at  $t$  (for all  $i \neq k$ ),  $\neg [B] = b_i$  at  $t$  (for all  $i$ ),  $\neg [C] = c_i$  at  $t$  (for all  $i$ ). An interpretive principle can imply this set of propositions only if it provides time references for the value ascriptions. COMP does just this in the only conceivable way. It uses the time reference attached to state  $W(t)$  and transfers it to negations of value ascriptions. It is hence crucial to relate these negations to time  $t$  in order for COMP to explicate the interpretation of the KS results we intend.

<sup>4</sup> While I say that (\*) *must* be explicated as EE to make formal and physical sense, a competing intuition has it that (\*) is perfectly o.k. as an expression of the eigenstate-eigenvalue link and simply *implies* EE. The net result is the same, of course. We are entitled to proceed with EE as the legitimate explication of the link.

<sup>5</sup> This intention can also be put another way. The full EE is not needed to make sense of the KS results. Modal interpretations reject EE (assume that if S's state is a reduction mixture, then it is possible that  $[A] = a_k$ ), but accept COMP. The ensuing argument creates a problem for these interpretations, too.

### 3. Main argument

#### 3.1 Explicating the trace formula

From Time Reference, we know that the trace formula inevitably carries a time-reference on the right:  $p(a_k) = \text{Tr}(W(t) P_{ak})$ . This creates a problem that has been little discussed in the foundational literature. The equation has a time-reference on the right, but it is unclear whether and eventually where there is a matching one on the left. Does  $p(a_k)$  also have a time-reference? If so it deserves to be explicated. Recall that  $p(a_k)$  is the probability for the event referred to by proposition  $'[A] = a_k'$ . So we interpret  $p(a_i)$  as a probability *function* (for a fixed pair of operators) from a set of propositions into  $[0, 1]$ . So we can cast the problem in these questions: Is  $p(a_k)$  a function of time? And if so, then in which sense is it such a function?

The first question must be answered positively because of the composition of functions. The trace formula *identifies* a magnitude (a probability) that is a function of propositions (attributing to S the different eigenvalues of A) with another magnitude (the trace) that is a function of operators (state times measured observable) one of which in turn is a function of a certain parameter (time), hence the probability function is itself a function of the parameter (time).

Now in which sense is  $p(a_k)$  time-dependent? Since formally it is a function there are basically two possibilities. Either this function depends on  $t$  *implicitly* - ' $p(a_k(t))$ ' - or *explicitly* - ' $p(a_k, t)$ '. In the first case, it depends on time because its argument, the proposition  $'[A] = a_k'$ , does. The argument then, in fact is the proposition  $'[A] = a_k$  at  $t'$ . In the second case, it depends on two arguments directly, on the proposition  $'[A] = a_k'$  and on time  $t$ . For a clearer visual distinction I rewrite this second option as ' $p(\underline{t})(a_k)$ '. Schematically then the explication must be either ' $p(a_k(\underline{t}))$ ' or ' $p(\underline{t})(a_k)$ '. Of course, the index could also be in both places at once. I ignore this third possibility here but consider it below (in sec. 3.3). We have just seen that the index must turn up somewhere and these are the two minimal possibilities. For obvious reasons the first reading can be called the *internal interpretation*:

$$p(a_k(t)) = \text{Tr}(W(t) P_{ak}).$$

and the second the *external interpretation*:

$$p(t)(a_k) = \text{Tr}(W(t) P_{ak}).$$

These two are the only apparent choices. Without arguing for it further, I assume that there are no others.

#### 3.2 The internal interpretation

I imagine that QM can be condensed into a number of theorems and, if applied to a certain system, a number of state specifications as necessary. I then can assume QM (applied to a certain system) to be a conjunction of theorems and state specifications. I can also speak of QM as a conjunct in a conjunction and henceforth will speak of the conjunction of QM and COMP (applied to system S with observable A) as complete QM.

I begin with a fundamental idea about probability that I hope finds universal acceptance: Probability is quantified possibility. More precisely: If a physical theory assigns an event a non-zero probability, then, given the theory's truth, this event is possible. The weakest form

of possibility is logical possibility. Using it we can implement the idea in the following *probability principle*: If a theory  $T$  for a proposition  $F$  (describing an event) yields another proposition  $p(F) > 0$  then it must be compatible with  $F$ , i.e. it is not the case that the set  $\{T, F\}$  makes derivable a contradiction (in first-order logic, I shall assume).

Now suppose that  $S$  is in a pure state  $W(t_1) \neq P_{a_k}$ , for some value  $t_1$  of  $t$ , such that for the probability calculated from the trace formula it is the case that  $1 > p(a_k) > 0$ . (This is a state specification such that  $S$  is in a pure non-eigenstate of  $A$  at  $t_1$ .) From COMP, we know that in state  $W(t_1) \neq P_{a_k}$   $S$  does not have value  $a_k$  at  $t_1$ , i.e.  $\neg[A] = a_k$  at  $t_1$ . But from the internal interpretation we know that  $p(a_k) = p(a_k(t_1))$  which is the probability that  $[A] = a_k$  at  $t_1$ . This probability is strictly positive, hence the theory producing it should be compatible with  $[A] = a_k$  at  $t_1$  – which complete QM evidently is not (given the state specification in question). Contradiction. It would not be correct to say that the internal interpretation contradicts COMP. In fact, the correct thing to say is that complete QM (including a trivially admissible state specification) plus the internal interpretation contradicts the probability principle. I trust that the reader finds this principle as fundamental as I do. So, the internal interpretation is untenable in complete QM.<sup>6</sup>

### 3.3 The external interpretation

There is only one tangible alternative and hence only one conclusion to be drawn. Since the internal interpretation is in conflict with complete QM we must employ the external interpretation to make sense of it.

We must consider (briefly) the following idea. Assuming that the time-parameter, in the external interpretation, refers to the probability for an event (not the event itself) we might still be free to assume that the parameter *also* refers to the event. Explicitly, the probability calculated from the trace formula, given a state  $W(t_1)$ , might be  $p(t_1)(a_k(t_1))$ . Obviously, this is not a helpful idea because we now have a strictly positive probability at  $t_1$  for  $S$  having  $a_k$  at  $t_1$ , i.e. for  $[A] = a_k$  at  $t_1$ . But we also still have, from  $W(t_1) \neq P_{a_k}(t_1)$  and COMP, that  $S$  does *not* have value  $a_k$  at  $t_1$ , i.e.  $\neg[A] = a_k$  at  $t_1$ . So, there still is a contradiction with our probability principle. The only way out is to assume that the time-parameter (here its value  $t_1$ ), which, given the external interpretation, the probability inherits from the state  $W(t_1)$ , does *not* also refer to the event for whose realization the probability is given. What it does refer to is *only* the probability itself.

However, we know from Time Parameter Uniqueness that QM contains just one parameter  $t$ . If it is not this parameter whose values are time-references for the predicted events in the trace formula, then these events can carry no time-references, at all. Hence, the expression ' $a_k$ ' in the formula ' $p(t)(a_k) = \text{Tr}(W(t)P_{a_k})$ ' is necessarily timeless in complete QM. Note also the following fact. The trace formula can be suitably generalised to encompass discrete and continuous observables. The relevant counterpart of proposition ' $[A] = a_k$ ' in that generalised form must of course also be timeless to match the discrete case. But then this must also hold for the continuous case, i.e. for a continuous observable  $A$ , a value  $a$  and an arbitrary Borel subset  $B$  of the real line it must be the case that:

$$p(t)(a \in B) = \text{Tr}(W(t)P_A(B)) \quad (3.1)$$

where the proposition ' $a \in B$ ' is necessarily timeless.

<sup>6</sup> See Held (2008, sec.s 2 and 3) for an earlier version of this argument.

Serious problems ensue. Familiar interpretations of QM notions cannot be maintained and important applications collapse. E.g., the so-called conventional interpretation of the wavefunction turns out to be false in complete QM. However, modern approaches to QM gravity (reasonably) construct probabilities for spacetime events that have this interpretation as a limiting case, so these approaches are also endangered. Similarly, QM transition probabilities and propagators cannot be defined in their usual textbook ways.<sup>7</sup> These problems should convince us right away that the external interpretation is not the one we use in real-life QM. Of course, the factual practice of QM does not constitute a knock-down objection in matters foundational but indeed such an objection is comparatively easy to come by. We can, with standard QM techniques, construct probabilities that carry an explicit *internal* time-reference – despite the fact that complete QM now forbids it.

The idea is as follows. In eq. (3.1) set  $A = X$  (the observable position),  $a = x$  and  $B = \Delta x$  (a value and value range of  $X$ ). Proposition ' $x \in \Delta x$ ' cannot carry the index  $t$ . And as in the discrete case this entails that the proposition cannot carry an index at all. Instead, the probability itself is time-indexed, i.e. we have another instance of the external interpretation:

$$p(x \in \Delta x) = p(t)(x \in \Delta x).$$

Consider now an observable that is a function of position, e.g.  $S$ 's potential energy  $V(x)$  and assume it to be explicitly time-dependent, i.e.  $V(x, t)$ . (Since the Hamiltonian  $H$  is a function of  $V$ , it is here that we use the possibility that  $H = H(t)$ .) Since  $V$  is a function of position (we can measure it via position measurements), we can exploit the Statistical Functional Composition Principle (a direct consequence of the trace formula; see Redhead 1987, 18) to construct probabilities for values  $V$  of  $V(x, t)$  and equate them with those for position:

$$p(V \in \Delta V(t)) = p(x \in \Delta x)$$

(where  $\Delta V(t)$  is time-dependent because  $V(x, t)$  is). The values  $V$  of  $V(x, t)$  on the left are manifestly time-dependent.<sup>8</sup> We thus have constructed a probability that explicitly has an internal time-reference and identified it with one that, given the external interpretation tailored to complete QM, cannot have such an internal time-reference. Contradiction, again.

So the external interpretation of the trace formula cannot be correct, either. We *need* time-references for the events predicted by the trace formula to do sensible QM. But the internal interpretation was so easily brought into a conflict with complete QM that it seems implausible to maintain. So in the presence of COMP, neither interpretation is tenable and there is no obvious third possibility. The argument, at this point, is inescapably turning against COMP. But COMP is our best (at any rate, most generally accepted) shot at how to interpret the KS no-hidden variables results. So we should very thoroughly look at possible objections. In the next section, I will consider whether the argument is a variant of the infamous measurement problem or otherwise trades on a neglect of the notion of QM measurement. Neither will turn out to be the case. I will then look at the only clear objection to the argument.

<sup>7</sup> Both these claims are substantiated in a companion paper (Held, in preparation).

<sup>8</sup> Details of the construction are again given in Held (in preparation).

## 4. Measurement

### 4.1 The measurement problem

An obvious intuition about the previous argument is that it constitutes just a peculiar form of the measurement problem and it is old wisdom that any interpretation of QM endorsing COMP must try to deal with *that* difficulty. But only the second half of this intuition is correct, not the first. Indeed it is true that a formulation of the measurement problem, in general, *presupposes* COMP. Let me quickly recall this.

For simplicity, assume that observable  $A$  on  $S$  has just two distinct eigenvalues  $a_1$  and  $a_2$ . Let  $S$  be measured by an  $A$ -measurement apparatus  $M$  and let  $[t_1, t_3]$  be an interval such that  $t_1$  is the time of measurement onset and  $t_3$  some later time. Assume that  $S$  is in a superposition of  $A$ -eigenstates, e.g. in  $\mathbf{P}_b(t_1) = |b\rangle\langle b|$  where  $|b\rangle = 1/\sqrt{2}(|a_1\rangle + |a_2\rangle)$ . First, assume that from our observation of a suitable pointer observable  $M$  on  $M$  we learn that  $[M] = m_1$  at  $t_3$ . From this we want to conclude something about  $S$ 's value of  $A$ , e.g.  $[A] = a_1$  at  $t_2$  - where  $t_2$  is an arbitrary time from  $[t_1, t_3]$ , including the boundaries. All choices of  $t_2$  are perfectly possible. In particular, it is consistent to assume jointly that  $S$  is in  $\mathbf{P}_b(t_1)$  (a projector onto a non-eigenstate of  $A$ ) and  $[A] = a_1$  at  $t_1$ . Similarly, if  $S$  is in  $\mathbf{P}_b(t_2)$ , where  $t_2$  is an arbitrary time with  $t_2 \in ]t_1, t_3]$ , ( $S$  stays in the superposition due to unitary time evolution) we can consistently assume  $[A] = a_1$  at  $t_2$ . Now, assume also EE plus State Uniqueness (from sec.2.1) or alternatively COMP. Then on pain of contradiction, we can assume jointly neither  $\mathbf{P}_b(t_1)$  and  $[A] = a_1$  at  $t_1$  nor  $\mathbf{P}_b(t_2)$  and  $[A] = a_1$  at  $t_2$ . We can no longer draw any conclusions from our assumed observation  $[M] = m_1$  at  $t_3$  to  $S$  - which is what generally a measurement of  $S$  by means of  $M$  is intended to deliver.

In its more familiar form, the measurement problem refers not to conclusions from observed pointer values to  $S$  values but instead to the QM state of the apparatus  $M$ . Let an  $A$ -measurement on  $S$  by  $M$  be modelled in the combined system  $S + M$  by initial states  $\mathbf{P}_{R_i}(t_1) = |R_i\rangle\langle R_i|$ , where  $|R_i\rangle = |a_i\rangle|m_{ready}\rangle$  and final states  $\mathbf{P}_{A_i}(t_2) = |A_i\rangle\langle A_i|$ , where  $|A_i\rangle = |a_i\rangle|m_i\rangle$ , for  $i = 1, 2$ , and where  $t_2 \in ]t_1, t_3]$ . Then, if  $S + M$  is in state  $\mathbf{P}_R(t_1) = |R\rangle\langle R|$ , where  $|R\rangle = |b\rangle|m_{ready}\rangle$  and  $|b\rangle$  as before, it follows that  $S + M$  is in state  $\mathbf{P}_B(t_2) = |B\rangle\langle B|$  with  $|B\rangle = 1/\sqrt{2}(|a_1\rangle|m_1\rangle + |a_2\rangle|m_2\rangle)$ . Again, it is perfectly consistent to assume that  $S + M$  is in  $\mathbf{P}_B(t_2)$  and  $[M] = m_1$  at  $t_2$  (or alternatively  $[M] = m_1$  at  $t_2$ ). Now, assume EE in the form:  $[M] = m_i$  at  $t$  if and only if  $S$  is in  $\mathbf{P}_{m_i}(t)$  (for  $i = 1, 2$ ) plus again State Uniqueness. Then, again on pain of contradiction, we can assume jointly neither  $\mathbf{P}_B(t_2)$  and  $[M] = m_1$  at  $t_2$  nor  $\mathbf{P}_B(t_2)$  and  $[M] = m_2$  at  $t_2$  for any  $t_2 \in ]t_1, t_3]$ . The apparatus has been pushed into a QM state such that, given EE, it is impossible to assume it as showing one of the pointer values - in contradiction with what we really observe. (As is well-known, COMP here is too weak to generate the contradiction because  $M$ 's state at  $t_2$  is a reduction mixture. As is also well-known, moving from EE to COMP alone does not suffice to avoid the measurement problem since the  $M$  reduction mixture cannot be given an ignorance interpretation.)

So, the measurement problem confronts QM states and hypothetical observations by means of EE plus State Uniqueness or COMP - which means that it *presupposes* EE or COMP. The argument of sec.3 is not a form of the measurement problem simply because it does not mention any observations at all that could be confronted with QM states. On the other hand, the measurement problem requires COMP or the stronger EE for its formulation. And the argument of sec.3 shows that COMP is incompatible with QM, even without that confrontation. If sound the argument provides grounds to *reject* the measurement problem

because it rests on a combination of COMP (or EE) and QM, which is in a deeper sense incoherent, i.e. leads to a consistency problem without invoking observations (though using other assumptions).

#### 4.2 The value assumption

Now, while the previous argument cannot be said to be a version of the measurement problem one may still think that it unduly neglects the crucial role of measurement in QM. I will try to dispel this impression. Evidently, the argument has not mentioned measurement. This would be inevitable if the two readings of the trace formula were integrated into a full axiom system for QM. Such a system would contain a statistical algorithm and that algorithm would be (or directly entail) one for generating probabilities via the trace formula. In formulating the latter algorithm – the Born Rule – we would have to take a stand on the notion of measurement. Generally, measurements and outcomes are mentioned explicitly and many think that this is an unavoidable feature of the Born Rule.<sup>9</sup> But considering the appropriate wording of the Born Rule here would just needlessly complicate the argument – which requires nothing but the trace formula itself.<sup>10</sup>

Let me however explicate an assumption that has tacitly been made in the argument (and will be tentatively questioned below, in sec.5). Recall that in the trace formula I took  $p(a_k)$  to be the probability for  $[A] = a_k$ , i.e. explicitly for ‘S has value  $a_k$  of  $A$ ’. The argument in sec.3.2 would not get off the ground if  $p(a_k)$  were *not* this probability for  $[A] = a_k$ , because the conflict arises between  $p(a_k) > 0$  (from the trace formula and an admissible state specification) and  $\neg [A] = a_k$  (from COMP). The conflict becomes sharp, if both statements are nailed to the same time and we assume that  $p(F) > 0$  entails that  $F$  can consistently be added as an assumption (our probability principle). But even then a contradiction will only arise if in ‘ $p(a_k)$ ’ the “inner” expression ‘ $a_k$ ’ really means ‘S has value  $a_k$  of  $A$ ’. If that is not the case then even fixing the same time for both propositions will not generate any problem. So the assumption at work in the previous argument is that in the trace formula:

$$p(a_k) = p([A] = a_k),$$

i.e.  $p(a_k)$  is the probability that  $S$  has value  $a_k$  of  $A$ .

This very elementary assumption may be called the *value assumption* because it amounts to the claim that the events concerning  $S$ , for which QM encodes probabilities, are simply identical to  $S$  having a value of an observable. Note carefully that this assumption does not involve a claim about *when*  $S$  eventually has value  $a_k$  of  $A$  and hence is not per se in a conflict with certain widespread intuitions about QM probabilities. Instead it helps to disentangle these intuitions – because we can separate the question of what the QM events (“inside” the probability of the trace formula) are from the one when they eventually occur.

I will consider three intuitions about measurement. First, there is the intuition that QM probabilities concern events that are *conditional upon measurement*. We can implement this in QM in different ways. First of all, we can write the following version of the Born Rule: If  $S$  is in state  $W(t)$  and  $A$  is an observable on  $S$  and *if an  $A$ -measurement on  $S$  is made* then:  $p(a_k) = \text{Tr}(W(t) P_{a_k})$ . Here measurement is mentioned as one of several conditions for the

<sup>9</sup> Among many possible references see Redhead (1987, 8), Hughes (1989, 70), van Fraassen (1991, 142-143), Albert (1992, 35), Bub (1999, 31), Nielsen & Chuang (2000, 84-85), Landsman (2009, 66).

<sup>10</sup> In this sense, the version of the argument of sec.3.2 given in Held (2008) is inferior to the present one.

trace formula which itself remains formally untouched. In particular, no reason emerges for assuming that the event for which a probability (given the conditions) is encoded should *not* be just that  $S$  has value  $a_k$  of  $A$ . Alternatively, we could insist that this special condition be directly written into the trace formula itself. Now, it is well-known that if conditional probabilities are introduced via the familiar Kolmogorov definition, then a conditional probability  $p(a_k | M_A)$ , where  $M_A$  is the “event” of an  $A$ -measurement being initiated, is not generally well-defined.<sup>11</sup> Suppose nevertheless, for a moment, that QM probabilities were simple conditional probabilities. Again, the value assumption could stay essentially unaffected. Probability  $p(a_k)$  would now equal  $p(a_k | M_A)$  but in that latter expression the component ‘ $a_k$ ’ could keep its meaning. If, on independent grounds we can manage to nail the ‘ $a_k$ ’ in  $p(a_k | M_A)$  to the same time  $t$ , for which an  $A$ -value is prohibited via COMP, the conflict is just the same as before. Finally, consider  $p(a_k \text{ given } M_A)$ , where ‘given’ is a conditional connective awaiting further analysis.<sup>12</sup> The situation will still be the same. Appreciating that QM probabilities are conditional on measurement will not in itself contradict the assumption that the conditioned events just consist in  $S$  having a value of an observable. All in all, conditioning on measurement is not in itself, in conflict with the value assumption.

Second, there is a strong intuition that, because QM probabilities are conditional on measurement, the events they concern are measurement *outcomes*. What is the consequence of this? Again, nothing in itself prohibits that a measurement outcome be reported in the statement ‘ $S$  has value  $a_k$  of  $A$ ’. Indeed, it is not entirely implausible to assume that a measurement outcome just is a value possessed (at some time) conditional on a measurement (at an earlier time). So, we can call the events in question by the name of outcomes and still respect the value assumption. Admittedly, the notion of an outcome is ambiguous enough to be interpreted, alternatively, as *not* meaning a possessed value but something else. This ambiguity, however, should tell against a use of “outcome” in the foundations of QM without clarifying whether or not an outcome is a (perhaps conditionally) possessed value of  $S$  or not.

Finally, it is sometimes said that QM systems upon measurement *display* or *show* values of observables<sup>13</sup> and hence QM probabilities should be interpreted as concerning such displaying or showing events. Since displaying or showing something is representing something, this is just a terribly misplaced metaphor. But even if we take QM systems as representing something upon measurement, there still are only two options. Does  $S$  have or possess what it represents (or displays or shows) upon measurement? If so the value assumption is met, if no it isn’t. Again, the idea in itself is ambiguous enough to be interpreted in a sense that is not in conflict with that assumption.

Before moving on I describe a simple non-QM example where all three intuitions plus the value assumption are respected. Our present picture of QM measurement may be described by using the following analogy. Such measurement is not like opening a box and determining the colour of the ball inside. Instead, it is like rolling a die and recording the result. Consider this latter picture. Assume that we have a fair die and want to define probabilities for the experiment of rolling it. We understand that the type of event, for which we define probabilities, is that the die shows a certain one of its six faces at the end of the

<sup>11</sup> See van Fraassen and Hooker (1976), 229.

<sup>12</sup> See Halpin (1991), Butterfield (1993), Held (2008, sec.5) for more on the appropriate analysis.

<sup>13</sup> See Earman’s portrayal of a ‘non-realist’ interpretation of QM measurement in Earman (1986), 219.

experiment. The probability  $p(6) = 1/6$  is directly understood as being the probability at the start of the experiment and as being conditional upon actually rolling the die. If measuring what the die shows is indeed doing the experiment of rolling it then, if the die in the end shows 6, that event is indeed *conditional upon measurement*. That event is also naturally described as the experiment's *outcome* and the die can meaningfully be said to *display* or *show* the number 6. Yet, there can be no doubt that the die in the end also *has* the value 6 (of the observable 'which face is up'). It would be confused here to assume that outcomes are not possessed values. Yes, they are but they are not values possessed initially - in contrast with what we assume about the ball in the box. The whole burden of distinguishing the die "measurement" from the "classical" colour measurement in the ball is carried by implicit time-references. In particular, in the case of the die  $p(6) = 1/6$  is the measurement outcome probability *at the beginning* of the experiment and the eventual outcome 6 is an event *at its end*. Accordingly, I now turn to a school of thought about QM measurement where probability and QM event refer to different times but the value assumption is respected. I will not try to be exhaustive but only consider two prominent authors representing that school, one a founding father of QM (von Neumann), the other a leading contemporary interpreter (van Fraassen).

### 4.3 Von Neumann's account

As has been pointed out many times before, Dirac and von Neumann can be credited with having first assumed EE - even if only implicitly.<sup>14</sup> In von Neumann's formulation of the Born Rule the key expression is that, given a state  $W$  and an observable  $A$ ,  $p([A] = a_k)$  (or equivalently  $p([P_{ak}] = 1)$  is the probability that  $A$  (or  $P_{ak}$ ) 'takes on' a value  $a_k$  (or 1) (see von Neumann 1932/1955, 200, 206, 211). There is only one sentence, however, where he explicitly says that the Born Rule 'does not tell us what values [the observables] *have* in the state  $W$ , but only with what probability they *take on* all possible values' (ibid., 206; my italics and notation). (In passing, I note that von Neumann does not explicitly claim this taking-on to be conditional on measurement.) Only here we explicitly learn that 'taking on' a value is opposed to 'having' it, i.e. that 'taking on' is a dynamical process.<sup>15</sup> By a dynamical process I mean a transition in time from  $S$  having no value of  $A$  to having one of the  $a_i$  - which (since von Neumann accepts EE) simultaneously is a transition from  $W \neq P_{ai}$  (for any  $i$ ) to  $P_{ai}$  (for some  $i$ ): the famous "collapse of the wave-function". As that latter expression suggests, the transition is construed best as taking place instantaneously; at any rate, this is the standard understanding.

What are we to say about this instantaneous transition? Suppose, quite unreasonably, that it is instantaneous in the sense of taking up no interval of time, at all. Because of Time Reference (also explicitly endorsed by von Neumann; see 1932/1955, 198)  $W = W(t)$ . Now, are we to think that the transition is from  $W(t) \neq P_{ai}$  (for any  $i$ ) to  $P_{ai}$  (for some  $i$ ), where  $P_{ai} = P_{ai}(t)$ ? Given State Uniqueness, this is contradictory. So this interpretation is implausible. Suppose, more reasonably, that the transition is instantaneous in the sense of

<sup>14</sup> See note 3 above.

<sup>15</sup> The other two passages are ambiguous. Suppose that for a function  $F(t)$  of parameter  $t$ , a real number  $f$  and value  $t_1$  of  $t$ ,  $F(t_1) = f$ . One might describe this by saying that  $F$  takes on value  $f$  at  $t_1$ . But of course  $F$  can simultaneously be said to *have* value  $f$  at  $t_1$ . Accordingly, the taking-on here is not a dynamical process. This interpretation, though implausible in a physics context, is at least possible for the passages on p. 200 and p. 211.

taking up an infinitesimal interval  $dt$  – where, as usual, “infinitesimal” means arbitrarily small but finite. Then the instantaneous transition can be viewed as being one from  $W(t)$  to  $P_{ai}(t+dt)$  (for some  $i$ ) and from  $S$  having no value of  $A$  at  $t$  to having one of the  $a_i$  at  $t+dt$  – a perfectly consistent interpretation of von Neumann’s idea. But notice that  $[t, t+dt]$  though arbitrarily small is finite and cannot go to zero (on pain of reproducing the previous contradiction).

Now, should von Neumann be interpreted as proposing a version of the Born Rule for transitions or for the values possessed at the end of such transitions? Naively, it seems legitimate to equate a probability for  $S$  taking on a value of  $A$  in the interval  $[t, t+dt]$  as the probability for  $S$  having that value of  $A$  at  $t+dt$  but is this interpretation admissible here? Indeed, it is not only legitimate but a necessity. Clearly, if  $W(t) = P_{ak}$ , then  $p([P_{ak}] = 1 \text{ at } t+dt)$  cannot be the probability for  $S$  taking on  $a_k$  within  $[t, t+dt]$  for that would entail that at  $t$   $S$  does not have  $a_k$  at  $t$  – which contradicts EE. But Generality requires that we give the trace formula a uniform interpretation (either transitions or possessed values) for all states. Hence, it must be possessed values. One *must* interpret von Neumann’s Born probabilities as being for the events of  $S$  having that value of  $A$  at  $t+dt$ .<sup>16</sup> This is to say that von Neumann *must* endorse the value assumption.<sup>17</sup> Moreover, he must endorse the following version of the trace formula:

$$p([P_{ak}] = 1 \text{ at } t+dt) = \text{Tr}(W(t) P_{ak}) \quad (4.1)$$

(where  $p([P_{ak}] = 1 \text{ at } t+dt) = p([A] = a_k \text{ at } t+dt)$ . But why should an interpreter propose this? For an answer, I can only refer back to sec. 3.2. Von Neumann endorses COMP (because he assumes EE and certainly endorses State Uniqueness). In the presence of COMP,  $p([P_{ak}] = 1 \text{ at } t) = \text{Tr}(W(t) P_{ak})$  looks implausible (for  $W(t) \neq P_{ai}$ , for any  $i$ ) and indeed, given the probability principle, does yield a contradiction. It is just natural to avoid the threat of that contradiction by forcing apart time references for state and possible value.

However, against this line of thought the second part of the above argument (sec. 3.3) now is seen to have bite. If (4.1) is the proposed explication of the trace formula for a state  $W(t)$  and observable  $P_{ak}$  then this proposal is in trouble. Obviously, a value of the time-reference  $t+dt$  on the left cannot come from  $W(t)$  because  $t$  and  $t+dt$  are supposed to take different values while the size of  $dt$  is nowhere determined. (Recall that the infinitesimal size of  $dt$

<sup>16</sup> Explicitly, the argument is this. Assume that  $p(T(a_i)) = \text{Tr}(W(t) P_{ai})$  where  $T(a_i)$  is the transition from  $S$  not having  $a_i$  at  $t$  to  $S$  having  $a_i$  at  $t+dt$  (for all  $i$ ). Presupposing Generality, assume  $W(t) = P_{ak}(t)$ . Hence,  $S$  is in the eigenstate pertaining to  $a_k$  at  $t$ . Then by EE,  $S$  has  $a_k$  at  $t$ . But the  $p(T(a_i))$  have just been defined as probabilities for transitions starting from  $S$  not having  $a_i$  (for all  $i$ ). Hence,  $S$  does not have any of the values  $a_i$  at  $t$  and a fortiori not  $a_k$ . Contradiction. In words: von Neumann’s Born probabilities cannot concern transitions from no value of  $A$  to value of  $A$ , in general, on pain of violating EE, when  $W$  projects onto an eigenstate of  $A$ .

<sup>17</sup> In fact, at a much later point von Neumann in a sense *denies* the value assumption when he writes: ‘Indeed experience only makes statements of this type: an observer has made a certain (subjective) observation; and never any like this: a physical quantity has a certain value.’ (von Neumann 1932/1955, 420). This idea is a denial of the value assumption if we assume that QM probabilities concern what we experience rather than what is the case. Note however that the present context is different because it speaks not of experience but of what is the case: von Neumann’s decree that the trace formula fixes ‘with what probability [observables] take on ... possible values’ (206) implies that such taking-on of a value can happen. If it does then, as a result, the observable *has* a value – regardless of whether an observer has a mere ‘(subjective) observation’ of that or more.

does absolutely no work here. The interval cannot be set to zero without reproducing the previous contradiction, so the limit of making  $dt$  ever smaller simply does not exist. But which time value then is  $t + dt$  ?) So there is nothing at all in QM to provide a value of  $t + dt$  on the left because (due to Time Parameter Uniqueness) there is no second time-parameter in QM. Hence (4.1) cannot be a formula of QM. Instead the outcome  $[P_{ak}] = 1$  in (4.1) *must* lack a time-reference. And, as we saw in sec.3.3, this in turn can be shown to be incompatible with QM itself.

#### 4.4 Van Fraassen's account

A detailed discussion of QM measurement is presented in van Fraassen (1991). While this discussion is critical of von Neumann's classical account, it is in one sense similar to the latter. Both are part of a school of thought that distributes state plus probability and eventual outcome to different times in order to make sense of QM measurement in light of COMP. Moreover, van Fraassen also respects the value assumption. I quote a longer passage, where he characterizes QM measurement *before* differentiating collapse and no-collapse (modal) interpretations (one version of which he advocates). Van Fraassen writes:

Born's Rule says [... that the] probability that a measurement of yes-no observable  $P$  will yield value 1, if made in state  $W$ , equals  $\text{Tr}(W P)$ . This statement gives us the probability of the occurrence of an event, conditional on the occurrence of a certain measurement process in a given state. [...] But what is this event? During the measurement process we see *an initial situation change into a final situation*. The *initial situation* can be analysed into two parts: the system of interest is in state  $W$ , and the environment (presence of and interaction with the measuring apparatus in its groundstate) has a certain character – call it  $IN$ . In *the final situation*, this system's state *will have evolved into*, say,  $W'$ , and the environment will have a new character – call it  $OUT$  – which includes e.g. that the pointer on the apparatus now sits at the number 1. So we have the transition:  $IN, W \rightarrow W', OUT$ . This transition could be indeterministic in two ways. Recall that we have a probability for  $OUT$ . Another possible transition – one that could have occurred but did not – may be described as  $IN, W \rightarrow W'', OUT'$  and it too had a probability. (1991, 244-5; my italics)

The two ways of indeterminism then differentiate two types of interpretation. Either, in a no-collapse interpretation, given 'the initial situation ( $IN, W$ ), the final state  $W'$  is completely determined (so  $W'' = W'$ ), and the probability we are given is really the probability of character  $OUT$ , given that the system is then in state  $W'$ .' Or, in a collapse interpretation, given 'the final state, the outcome character is completely determined ( $OUT \neq OUT'$  then  $W' \neq W''$ ), and the probability we are given is the probability of the transition from state  $W$  to  $W'$ , when the initial situation has character  $IN$ .' (1991, 245).

From this passage, we learn two things about QM measurement. First, although van Fraassen does not explicate time-references for  $W$  and  $W'$  (or  $W''$ ) it is entirely obvious from (the italicised portions of) the text that  $W'$  (or  $W''$ ) is the system's state at a time *later* than the one of  $W$ . Second, it is also clear from the text that the Born Rule probability is one *for* an event at the *later* time:  $W'$  and character  $OUT$  are both part of the 'final situation' and the original probability is either one for  $OUT$  only or for  $OUT$  plus  $W'$ . All in all, van Fraassen straightforwardly explains that in the trace formula  $p([P] = 1) = \text{Tr}(W P)$  the elements ' $W'$  (state) and ' $[P] = 1$ ' (value) refer to different times. In a collapse interpretation, this

probability simultaneously is one for  $W' = \mathbf{P}$  such that we can rewrite the formula as  $p(\mathbf{P}) = \text{Tr}(W \mathbf{P})$ . This of course implies the weaker formula. Moreover, since by assumption  $W'$  is a state later than  $W$ , in this rewritten version the two time references are even more obvious.

Explicitly, van Fraassen starts with  $W$ , the initial state of  $S$  ('the system of interest') as an input for the trace formula. Hence, he proposes a version of formula  $p([\mathbf{P}_{ak}] = 1) = \text{Tr}(W \mathbf{P}_{ak})$ , where ' $W$ ' (initial state of  $S$ ) and ' $[\mathbf{P}_{ak}] = 1$ ' (value of  $S$ ) refer to *different times*. This is a literal interpretation of the passage and I emphasize right away that it is doubtful whether it really catches van Fraassen's intention but leave that worry aside for a moment. What van Fraassen clearly does endorse is the value assumption. When expounding his own interpretation he takes the Born Rule to deliver probabilities for 'value states' or better for the truth of 'value attributing propositions' where such a proposition 'states that observable  $A$  actually *has* a value  $a_k$ ' (van Fraassen 1991, 275; my italics and notation).

Assume now that in van Fraassen's equation  $p([\mathbf{P}_{ak}] = 1) = \text{Tr}(W \mathbf{P}_{ak})$  the state  $W$  equals  $W(t_1)$  where  $t_1$  is the time of the onset of interaction between  $S$  and  $M$ . This is clearly suggested by his characterizing  $W$  as the 'initial state'. Assume that the second half of my interpretation is also correct, i.e. that ' $[\mathbf{P}_{ak}] = 1$ ' is an equivalent of his value attributing proposition which, if true, is true at a later time. Van Fraassen then is effectively putting forward this version of the trace formula:

$$p([\mathbf{P}_{ak}] = 1 \text{ at } t') = \text{Tr}(W(t) \mathbf{P}_{ak}) \quad (4.2)$$

where  $t' > t$ . But of course, the problem with (4.2) is the same as with von Neumann's (4.1). QM contains no theory giving us  $t'$  for a given  $t$ , so (4.2) cannot be a formula of QM. This means that also in van Fraassen's version of the trace formula the internal time-reference must be cancelled as a matter of principle - which in sec. 3.3 was shown to be in conflict with QM.

I have recorded doubts whether van Fraassen should be saddled with the interpretation leading to (4.2).<sup>18</sup> Indeed, it is very clear that his account of measurement aims at another state. To quote him again, in a no-collapse interpretation the trace formula provides 'the probability of character OUT, *given that the system is then in state  $W'$* .' We may read this as saying that really  $W'(t_2)$  of  $S$  is the appropriate input for the formula. This reading is corroborated by looking at van Fraassen's own version of a no-collapse interpretation. Referring back to my discussion of a superposition measurement in sec.4.1, the interpretation holds that the state of composite system  $S + M$  at  $t_2$  equals  $\mathbf{P}_B(t_2) = |B'\rangle\langle B'|$  with  $|B'\rangle = 1/\sqrt{2}(|a_1\rangle|m_1\rangle + |a_2\rangle|m_2\rangle)$ . At  $t_2$ ,  $S$  and  $M$  have become correlated, their reduced states are mixtures and denial of EE allows us to ascribe values to

<sup>18</sup> I here refer to van Fraassen (1991). In an earlier exposition of his modal interpretation van Fraassen formulates an equivalent of (4.2) explicitly. He gives this version of the Born Rule: 'When  $A$  is an observable pertaining to  $S$  and the process  $w(t_1) \rightarrow w(t_2)$  is a measurement of  $A$  on  $S$  by  $M$ , with pointer observable  $M$ , then  $\alpha_1(t_1) \langle A, E \rangle$  is the probability that  $w(t_2)$  is in  $\langle A, E \rangle$  and equally the probability that it is in  $\langle M, E' \rangle$ ' (van Fraassen 1981, 305; notation adapted). Van Fraassen's ' $S$  is in  $\langle A, E' \rangle$ ' is equivalent to ' $[A] \in E'$ ' and for discrete  $A$  may be simplified to ' $[A] = a_k$ ' or equivalently ' $[\mathbf{P}_{ak}] = 1$ '.  $\alpha$  is a map from the set of value attributions into  $[0, 1]$  given by  $\alpha \langle A, E \rangle = \text{Tr}(\alpha \mathbf{P}_{M, E})$ , where  $\alpha$  is a density operator, i.e. in the present notation we can introduce a similar map  $W$  with:  $W([\mathbf{P}_{ak}] = 1) = \text{Tr}(W \mathbf{P}_{ak})$ . Putting in time references from the quotation, we get  $p([\mathbf{P}_{ak}] = 1 \text{ at } t_2) = \text{Tr}(W(t_1) \mathbf{P}_{ak})$  - which is an instance of (4.2).

observables  $A$  and  $M$  (on  $S$  and  $M$ ). This again seems to imply that it is a state  $W'(t_2)$  (e.g. a reduced state of  $S$ ) that should go into the trace formula – at least in a no-collapse interpretation. (Note that for a collapse interpretation using  $W'(t_2)$ , a collapsed state as an input for the formula would not, generally make sense.)

Van Fraassen's exposition only suggests this alternative interpretation but another interpreter, Vermaas (1999, 195), explicates it – for a certain type of state  $W'(t_2)$ .<sup>19</sup> Vermaas also endorses a no-collapse interpretation and is motivated by a similar account of measurement. He uses  $W'(t_2)$ , where this is the reduced state of the apparatus  $M$ , in order to calculate probabilities for values of the pointer observable  $M$  or explicitly:  $p([P] = 1 \text{ at } t_2) = \text{Tr}(W'(t_2) P)$  (where  $P$  now is a projector onto an eigenspace of  $M$ ).<sup>20</sup> Apparently, it seems unobjectionable *in this particular situation* to write the trace formula like this. But of course the formula cannot plausibly be restricted to this situation. We cannot help ourselves to a version here which we are not obliged to in general. When we consider possible interpretations of *the* trace formula any restriction to a certain type of state (apparatus mixture) is illegitimate. Neither the formula nor any standard version of the Born Rule (where it ultimately resides) nor the algorithm for expectation values know of any such restriction. And it is this fact that has led me to explicitly claim Generality of the trace formula as a property of QM. Now, if we read Vermaas' version as a general proposal for interpreting the trace formula then clearly we are back to square one. We must allow  $W(t_1)$  as a possible input for the formula, where  $W(t_1)$  is  $S$ 's state at an arbitrary time (it might be measurement onset or not). According to the proposal Vermaas makes (using  $W'(t_2)$ !), we then have, for (the equally admissible)  $W(t_1)$ , the formula  $p([P_{ak}] = 1 \text{ at } t_1) = \text{Tr}(W(t_1) P_{ak})$ . But this just is the internal interpretation of the formula from sec.3.2. If  $W(t_1)$  is a superposition of  $A$ -eigenstates this interpretation is straightforwardly brought into conflict with COMP.

One may question whether I am doing justice to the proposal. What is intended is a version of the trace formula:  $p([P] = 1 \text{ at } t_2) = \text{Tr}(W'(t_2) P)$ , where  $W'(t_2)$  is an arbitrary state in the sense of pertaining either to  $S$ ,  $M$  or  $S + M$ , but a non-arbitrary one in the sense of referring to a time  $t_2$  *after measurement onset*. By contrast,  $W(t_1)$  does not unambiguously refer to such a time, indeed can be taken to be the state at measurement onset. So the idea is to formulate a version of the trace formula roughly like this:

$$p([P_{ak}] = 1 \text{ at } t_2) = \text{Tr}(W'(t_2) P_{ak}) \quad (4.3)$$

where  $W'(t_2)$  explicitly is considered as a state *after* measurement onset. Taking  $t_1$  to be measurement onset,  $W(t_1)$  is not a legitimate input for (4.2). I understand the idea but still cannot see how it respects the QM property of Generality. At an arbitrary time  $t$ , the system's state is represented by an arbitrary member  $W(t)$  of the set of all possible density operators on the appropriate space and the trace formula processes *that* state. So to respect Generality,  $W(t_1)$  has to be an admissible input.

One may propose that Generality should be relinquished. I have two replies. First of all, I insist that Generality is a feature of ordinary QM. No version of the trace formula known to me makes any restrictions on the allowed input states and observables.<sup>21</sup> KS type theorems

<sup>19</sup> Vermaas here characterizes a common property of three modal interpretations that are all stronger than van Fraassen's (see Vermaas 1999, ch.4), so we may take his proposal as an interpretation of the trace formula in modal interpretations, in general.

<sup>20</sup> This explicitly is the content of Vermaas' proposal (1999, 195).

<sup>21</sup> In particular, this is true for the references in note 9 above.

concern observables in ordinary QM, COMP was intended as an interpretation of these theorems referring to ordinary QM, and it is ordinary QM that has been shown to be in a conflict with COMP. So the proposal to relinquish Generality amounts to a change of the subject because it does not talk about ordinary QM.

Second, I do not see how (4.3) really could work. Since Generality is a formal property of QM, to abandon it means to substantially revise the formalism and it is not a priori clear whether the revision will give us a sensible new QM. Consider this. Regardless of whether we interpret  $W'(t_2)$  in (4.3) as a state of  $S$ ,  $M$  or  $S + M$ , it is a state at some time during (or at the end of) an  $A$ -measurement, hence is not independent of a choice of observable  $A$ . Accordingly, the trace formula is no longer general in yet another sense, i.e. it no longer processes arbitrary observables. Suppose that we want to calculate  $W'(t_2)$  via the QM time-evolution directly from the preparation  $W(t_0)$  – insisting with (4.2) that an  $A$ -measurement must start at some time  $t_1$  in between, but *not* specifying a state  $W(t_1)$  – then  $W(t_0)$  has to include some state of the  $A$ -measurement apparatus leading to the appropriate ‘ready’-state, i.e. we would no longer be able to *refer* to states of QM systems apart from devices measuring specific observables on them. Suppose alternatively that we want to recover something like Generality, given (4.3). Assume that, for observables  $A$  and  $B$  with  $[A, B] \neq 0$ , we want the trace formula to yield probabilities for both choices. This inevitably will force that we *do* explicate a state  $W(t_1)$  of  $S$  alone (calculated from  $W(t_0)$  and with  $t_0 < t_1 < t_2$ ) that is neutral between those choices. So the algorithm containing the new trace formula (4.3) (a new version of the Born Rule) will have to make reference to two different states  $W(t_1)$  and  $W'(t_2)$ . And this will have to involve an exact theory about how the two states are related, i.e. given  $W(t_1)$ , where  $t_1$  may be measurement onset, we must name sufficient conditions for how to pick time  $t_2$  such that  $W'(t_2)$  can be put into the trace formula. The usual choice – that  $t_2$  is a time when  $S$  and  $M$  have become perfectly correlated (see van Fraassen 1991, 200) – is unavailable because that correlation presupposes an apparatus measuring a specific observable, mention of which we just wanted to avoid.

In sum, (4.3) simply does not respect one of the properties of QM, i.e. Generality. Giving up this property forces a formidable revision of QM that does not yet exist. I don’t think that modal interpreters like van Fraassen or Vermaas really intend such a revision. But the only alternative to (4.3) is the literal interpretation of van Fraassen’s general account of measurement – in effect (4.2). The argument of sec.3.2 can be viewed as explaining why interpreters put forward proposals like this. (Van Fraassen, though exceptionally clear on the matter, certainly is not the only interpreter doing so; vide von Neumann.) And sec.3.3 above presents an argument why (4.2), like (4.1) before, cannot be a component of QM. All in all, both van Fraassen and von Neumann exemplify an important tradition of thought about QM measurement that as a whole is challenged by the above argument. Our consideration of this tradition, rather than showing how to avoid or mitigate the original completeness problem, exposes it even more starkly.

## 5. Objection

The mentioned school of thought about QM measurement tries to accommodate COMP by accepting the value assumption and suggesting two time references, either by separating the time references of state and eventual value ((4.1 and (4.2)) or by tacitly referring to a state at a time after measurement onset ((4.3)). Both roads, I have argued, are blocked. So we should now consider the remaining alternative. Could the value assumption be dropped?

This *is* a fundamental objection and it makes the whole argument of sec.3.2 collapse immediately. To repeat: Given the value assumption, in the trace formula  $p(a_k) = p([A] = a_k)$  where  $[A] = a_k$  means that S has value  $a_k$  of A. Now, choose  $W(t)$  such that  $p(a_k) > 0$  and, by COMP,  $\neg[A] = a_k$  at t. Setting  $p(a_k) = p(a_k(t))$  we get an implausibility and, assuming the probability principle, a contradiction. It all depends on the first step: the value assumption. To drop it means to assume that QM probabilities do *not* concern events that can be identified with S having a value (of an observable) at a specific time. Instead they concern events of another type – but events that can be assumed to occur at time t, where t is the parameter from  $W(t)$ .

We can very broadly delimit what these events are. The purpose of giving up the value assumption is to find a replacement for  $[A] = a_k$  at  $t'$  in  $p([A] = a_k$  at t) such that the latter does not (via the probability principle) get into conflict with  $\neg[A] = a_k$  at t, from COMP (given a pure non-eigenstate of A). A positive suggestion for the events is that they are situations which are *as if*  $[A] = a_k$  at t. Let's write  $p("[A] = a_k$  at  $t'$ ") – using double quotes – for the probability that the situation at t is as if S had  $a_k$  at t. In fact, this is the broadest construal of an alternative to  $p([A] = a_k$  at t) – no double quotes – that I can think of. Its appeal, I believe, comes from the following intuition. A physical situation can have features that are as if some fact F were the case while in fact it must, for whatever reason, *remain open* whether really F is the case.<sup>22</sup> Think, e.g., of a concrete reading of a value of pointer observable M from which we draw conclusions to S values (perhaps at the same or perhaps at an earlier time). Isn't the whole situation at the end of a measurement as if S had  $a_k$  (at some time, leave that open for the moment)? And isn't it plausible that one can remain agnostic whether S really had  $a_k$  (at that time)?

This interpretation of the 'as if'-proposal may be appealing but here it is not available. Recall once more the structure of the argument in sec.3.2 and plug in the 'as if'-proposal. Choose  $W(t)$  such that  $p("[A] = a_k$  at  $t'$ ")  $> 0$  and, by COMP,  $\neg[A] = a_k$  at t. Since the probability is positive the probability principle applies, i.e. the proposition "within" the probability must be consistent with  $\neg[A] = a_k$  at t. For a check, resolve the double quotes of that proposition in the manner suggested: 'It is as if  $[A] = a_k$  at t but in fact it *remains open* whether  $[A] = a_k$  at  $t'$ '. Clearly this sentence does *not* safely avoid a contradiction with  $\neg[A] = a_k$  at  $t'$ . So the appealing version of the 'as if'-interpretation is not the one that can salvage the consistency of complete QM. Instead what we need is that propositions of type " $[A] = a_k$  at  $t'$ " (with double quotes) are interpreted as follows: 'It is as if  $[A] = a_k$  at t but in fact *it is not the case* that  $[A] = a_k$  at  $t'$ '. This statement is the only explication safely consistent with  $\neg[A] = a_k$  at  $t'$  and hence is the only available reading of the 'as if'-interpretation. But now recall that the proposal must be taken as an alternative reading of all probabilities within the trace formula. So, if we adopt the 'as if'-proposal we automatically propose what QM probabilities are, in general. Such probabilities now concern the *mere* display of values. (For contrast, recall the display of number 6 in the die in sec.4.2). Writing the display of a value

<sup>22</sup> This also is a plausible interpretation of von Neumann's idea that in a QM context experience only consists in statements of '(subjective) observation', not statements ascertaining what is the case (see note 17). I have argued above that in the context of his 'taking on'-idea probabilities must concern what is the case, hence the value assumption must be respected. Here, I explore the idea that probabilities do concern only what we experience, not what is the case, i.e. that the value assumption is rejected. I do not think that von Neumann can coherently maintain both his ideas but I will not try to show this as it would embroil me in questions of text exegesis.

in French quotes, we have that " $a_i$ "  $\leftrightarrow$  " $\langle a_i \rangle$ " &  $\neg a_i$ . We have that an 'as if'-probability  $p$  (" $a_i$ ") equals  $p$  (" $\langle a_i \rangle$ " &  $\neg a_i$ ) (for all  $i$ ).

As an interpretation of a fundamental physical theory this sounds wildly implausible and it can indeed be shown to be so. Consider what can be the only motivation for the proposal: to harmonize QM and completeness in the form of either EE or COMP. So the proposal had better be compatible with these two assumptions. But it cannot be so if it also respects Generality, i.e. is a state-independent version of the trace formula. Assume that the formula is now to be read as:

$$p("[A] = a_k \text{ at } t") = \text{Tr}(W(t) P_{ak}) \quad (5.1)$$

for arbitrary  $W(t)$  and  $A$ . Now, suppose that  $W(t) = P_{ak}$ . Then, from (5.1)  $p("[a_k(t)]") = 1$ . But this equals the following probability:  $p("\langle a_k(t) \rangle$ " &  $\neg a_k(t)$ ). Since the probability of this conjunction equals 1 we may assume that both conjuncts are true. In particular it is true that  $\neg[A] = a_k$  at  $t$ . We have deduced: If  $S$  is in  $W(t) = P_{ak}$  then  $\neg[A] = a_k$  at  $t$ . This contradicts EE, so a defender of full EE cannot maintain (5.1) for all states.

Given COMP, there is no contradiction but still we have deduced that if  $S$  is in  $W(t) = P_{ak}$  then  $\neg[A] = a_k$  at  $t$ . Since  $A$  is arbitrary and  $k$  is an arbitrary one of the indices  $i$ , we have derived that if  $S$  is in an arbitrary  $A$ -eigenstate then it does not have a value of  $A$ . But recalling COMP, when  $S$  is in an  $A$ -eigenstate then it does not have a value of  $B$ , for any  $B$  with  $[A, B] \neq 0$ . So when  $S$  is in an arbitrary pure state, it does not have a value of *any* observable. This is an implausible result. We should recall why COMP was endorsed in the first place, namely to accommodate the classical no hidden variables proofs (of KS-type). These proofs show that (given two assumptions) observables on  $S$  forming a KS set cannot *all* be assumed to simultaneously have values. Now, we learn that given  $S$ 's state is pure, none of the observables in an *arbitrary* set has a value anyway (at a given time). So, given a pure state the claims of all KS-type proofs are vacuously true. Now, if  $S$ 's state is an ignorance mixture then  $S$  really is in one of the pure states we happen to be ignorant about, so assuming  $S$ 's state to be such a mixture will not make a difference. So the defender of (5.1) and COMP must claim that only if  $S$  is in a mixture not interpretable as an expression of our ignorance, i.e. a reduction mixture, it can be assumed to have a value of *some* observable and the KS arguments are non-vacuously true. However, it is perfectly admissible to assume  $S$  – say, Kochen and Specker's original spin-1 system – to be the only QM system in the universe. In this case, there is no larger system such that  $S$ 's reduced state could be interpreted as a non-trivial reduction mixture – which again means that the KS arguments are only vacuously true. Hence, this line of reasoning is implausible. Accordingly, the 'as if'-interpretation cannot deliver a version of the trace formula that is plausible and respects Generality.

As in the case of proposal (4.3), one may seriously consider giving up Generality. Again, this means to give up something that is a property of ordinary QM. Abandoning it may be a possible reaction to the argument (which concerns ordinary QM) but certainly is no valid objection. And again the price of abandoning Generality is high. It is a formal property of QM, hence without it we must revise the formalism. Neither has such a revision been tried anywhere nor is it obvious how a reasonable new QM could arise. Suppose that we bite the bullet and dismiss the project of one general trace formula. We would have two trace formulas instead:

$$p (" [A] = a_k \text{ at } t'') = \text{Tr} (W (t) P_{ak}), \text{ if } W (t) \neq P_{ai}, \text{ for any } i; W (t) \text{ pure} \tag{5.2a}$$

$$p ([A] = a_k \text{ at } t) = \text{Tr} (W (t) P_{ak}) \text{ if } W (t) = P_{ai} \text{ for any } i \tag{5.2b}$$

According to (5.2b), for the eigenstates of  $A$  we have  $p (a_k (t)) = \delta_{ik}$  for  $W (t) = P_{ai}$ . And by (5.2a),  $p (" a_k (t) ")$  is the appropriate probability for a pure non-eigenstate of  $A$ . How to treat mixed states? Assume first that they are lumped together with the  $P_{ai}$  in (5.2b), i.e. the trace formula produces probabilities for the truly possessed values for these mixtures (rather than the 'as if'-probabilities of (5.2a)). It seems plausible that for an observable  $A \otimes B$  on a compound system the 'as if'-interpretation (5.2a) is true with respect to  $A \otimes B$  if and only if it is true for  $A$  and  $B$ . Now, let  $S + M$  be in state  $P_B = |B\rangle\langle B|$  with  $|B\rangle = \frac{1}{2}(\sqrt{3} |a_1\rangle + |m_1\rangle + \frac{1}{2} |a_2\rangle + |m_2\rangle)$ , where the  $|a_i\rangle$  and  $|m_i\rangle$  (for  $i = 1, 2$ ) are eigenstates of  $A$  and  $M$ , respectively. Then the reduced states  $\# P_B$  of  $S$  and  $\# P_B$  of  $M$  are mixtures of the respective eigenstates and the probabilities are calculated via (5.2b), according to our assumption. But  $S + M$  is in a *pure* non-eigenstate of  $A + M$ , hence the 'as if'-version (5.2a) applies. Since e.g.  $\text{Tr} (P_{m1} (\# P_B)) = \text{Tr} ((P_{a1} \otimes P_{m1}) P_B)$  (where  $P_{a1}$  and  $P_{m1}$  project onto rays containing  $|a_1\rangle$  and  $|m_1\rangle$ , respectively, in the respective subspaces) we have that  $p (m_1) = p (" a_1 " \& " m_1 ")$  and since  $P_B$  is a perfectly correlated state also  $p (m_1) = p (" m_1 ")$ . Now, events  $\langle m_i \rangle$  &  $\neg m_i$  are also perfectly correlated (for otherwise we blur the distinction between (5.2a and 5.2b)). Hence,  $p (m_1) = p (" m_1 ") = p (\langle m_1 \rangle \& \neg m_1) = p (\neg m_1)$ . This means that  $p (m_1) = p (\neg m_1) = \frac{3}{4}$ . Hence, we cannot plausibly treat mixtures via (5.2a).

Suppose alternatively that we treat mixtures like pure non-eigenstates of  $A$ , i.e. apply an 'as if'-interpretation as in (5.2a). Assume that the system in question is the measurement apparatus  $M$ , that there is an observable  $M$  on  $M$  and that the state of  $S + M$  is a superposition such that  $M$ 's reduced state  $\# P (t_1) = \sum_i a_i P_{mi}$ , where the  $P_{mi}$  project onto the pointer eigenstates. Then for this state (of  $M$ !) the conditions for (5.2a) are in place. Hence, we have probabilities  $p (" m_i ") = p (\langle m_i \rangle \& \neg m_i)$  (for all  $i$ ), i.e. each probability concerns the event that the  $M$ -measurement apparatus  $M$  is in a situation that is as if  $M$  had  $m_i$  at  $t_1$  and  $M$  does not have  $m_i$  at  $t_1$  (for all  $i$ ). But  $M$  is an observable accessible to the naked eye. So QM plus (5.2a, b) proposes that direct observation of a macroscopic object is regularly mistaken and universally so, if we have reason to believe the observed object  $M$  is in a non-eigenstate of the pointer observable.

We need to hear nothing further, I believe. An empirical theory that must assume systematic error in the direct observations eventually testing it is absurd. So the 'as if'-proposal fails not only as a general proposal for the trace formula, but also in terms of its content.<sup>23</sup> Hence, the only tangible proposal not obeying the value assumption fails. The argument of sec.3 remains intact.

---

<sup>23</sup> Note also that the measurement problem, in the second version above, is trivially dissolved. Nothing prohibits that we calculate probabilities for the values of pointer observables (of a macroscopic apparatus) from the trace formula. But these must now, of course, also be probabilities for 'as if'-events. Consider the probability  $p (" [M] = m_i \text{ at } t_3 ")$  deduced from a state  $W (t_3)$  of  $S + M$  and  $t_3$  is the end of the  $S$ - $M$  interaction. We have a probability for a situation that is as if  $M$  showed  $m_i$  at  $t_3$  while  $M$  does not in fact show  $m_i$  at  $t_3$ . If we take to heart this interpretation then there can of course be no measurement problem. This problem was that we see, e.g.,  $[M] = m_i$  at  $t_3$  while the superposition affecting  $M$  prohibits (via EE) that  $[M] = m_i$  at  $t_3$ . But given the 'as if'-interpretation our vision is deceptive anyway. It is not the case that  $[M] = m_i$  at  $t_3$ . If our observations are by assumption fallacious they can be in conflict with ascribed QM states but the conflict is harmless just because they are fallacious. Cancelling one of two contradictory statements always removes a contradiction.

## 6. Consequences

The argument here developed shows that EE, the standard interpretation of QM completeness originating with Dirac and von Neumann, is implausible. Given State Uniqueness, EE entails COMP, which cannot be true under plausible assumptions. This may seem but a minor result. While many theoreticians not concerned with the foundations of QM still reproduce EE many others concerned with these foundations reject it right away. But the above result is of considerable strength. First of all, adopting or rejecting EE in the past often was a matter of general interpretive preferences, so any argument relating it to the structure of QM itself must be welcome. Moreover, in the presence of State Uniqueness, which is a property of QM anyway, COMP is weaker than EE, hence the result is stronger. Finally, the assumptions made are mostly very general and so the argument should guide our future thought about time-references in the formalism of QM.

Of course, if the result is any good then its trivial consequence is that, given the assumptions, the negation of COMP is true. But the discussion hopefully has highlighted what exactly this negation is: If  $S$  is in some state  $W(t)$  such that for an arbitrary observable  $P_{ak}$  the trace formula yields:  $1 > p(a_k) = \text{Tr}(W(t) P_{ak}) > 0$  (i.e.  $W(t)$  is a non-eigenstate of  $A$ ), then this must be explicated by the internal interpretation, hence must explicitly run:  $p(a_k(t)) > 0$  and it must be legitimate to assume that  $S$  has value  $a_k$  of  $A$  at time  $t$  - *even when  $W(t)$  is a pure non-eigenstate of  $A$* . It should be obvious that this conclusion, if inescapable, has radical implications. Let me nevertheless illustrate them by means of a well-known and elegant KS variant due to Mermin (1990). Let  $S$  be a two-particle spin  $\frac{1}{2}$  system (hence we must have a space  $\mathcal{H}$  with  $\dim \mathcal{H} = 4$ ) and consider (for an orthogonal triple  $x, y, z$  of vectors in  $\mathbb{R}^3$ ) these nine spin observables:  $\sigma^1_x \sigma^2_x \sigma^1_y \sigma^2_y \sigma^1_x \sigma^2_x \sigma^1_y \sigma^2_y \sigma^1_x \sigma^2_y \sigma^1_y \sigma^2_x \sigma^1_z \sigma^2_z$ . Writing again  $[A]$  for the value of an observable  $A$ , one can derive:

$$\begin{aligned}
 [\sigma^1_x] [\sigma^2_x] [\sigma^1_x \sigma^2_x] &= 1 \\
 [\sigma^1_x] [\sigma^2_y] [\sigma^1_x \sigma^2_y] &= 1 \\
 [\sigma^1_y] [\sigma^2_x] [\sigma^1_y \sigma^2_x] &= 1 \\
 [\sigma^1_y] [\sigma^2_y] [\sigma^1_y \sigma^2_y] &= 1 \\
 [\sigma^1_x \sigma^2_y] [\sigma^1_y \sigma^2_x] [\sigma^1_z \sigma^2_z] &= 1 \\
 [\sigma^1_x \sigma^2_x] [\sigma^1_y \sigma^2_y] [\sigma^1_z \sigma^2_z] &= -1
 \end{aligned}$$

The derivation uses the constraints mentioned in sec.2.2: Functional Composition (here the Product Rule:  $[AB] = [A] [B]$ ) and Noncontextuality (every operator represents a unique observable such that two occurrences of some  $[A]$  have the same value). The result is so elegant because the impossibility is so easy to see. Since every value appears twice the product of all the left sides must be positive, while the one of all the right sides is negative. Now let  $S$  be in the singlet state. Then it is straightforward to show that all nine observables get strictly positive probabilities. Hence, given  $\neg$  COMP, it must be possible to assume that all nine observables have values at time  $t$ . Since this is impossible given the two constraints at least one of them must be false.

The problem outlined should direct further research in the foundations of QM. As indicated the measurement problem over which so much ink has been spilled is a *consequence* of yoking COMP and QM together despite the hidden inconsistency I have tried to uncover here. The only conclusion I can see is that attempts to resolve the measurement problem are misguided and that we must again investigate hidden-variables interpretations more

seriously. It should give us pause to see that elementary textbooks of QM directed at undergraduates and ignorant of interpretational subtleties rectify, without further ado, equivalents of  $p(a_k) = \text{Tr}(W(t) P_{a_k})$  into  $p(a_k(\underline{t})) = \text{Tr}(W(\underline{t}) P_{a_k})$ , that theoretical physicists defining transition probabilities and propagators do so as well as those trying to wed quantum theories and general relativity while orthodox interpretations either ignore the problem or, via the measurement-outcome idea, smuggle in references to times other than  $t$  that simply are not provided by  $W(t)$ . To my mind, all this recommends a sincere look at how far away we are from understanding QM.

The next logical step is to try and answer which of the two conditions – Functional Composition and Noncontextuality – QM systems violate. If we look at the two possibilities of contextuality discussed in the (not so recent) literature then measurement-induced Contextuality is not an option because, as we saw, QM systems possess the properties for which we make QM predictions at measurement onset. The alternative, called Ontological Contextuality (see Redhead 1987, 137), is little more than a label and so far it is nothing that can guide an understanding of the non-Boolean lattice structure of QM observables. But here, to my mind, should be the starting point for future research. Finally, it should be pointed out that the argument relates predictions of properties of QM systems to the times of measurement onset in measurements of these systems. This *does* entail that the properties measured in a QM system exist in it at measurement onset, but it *does not* entail that measurement is faithful in the sense of revealing pre-existing properties. QM properties might very well be relational in the sense of being relations between system and apparatus or between system and observer. It is just that these properties, if they are relations, come into existence, at the latest, at the time system and apparatus meet. To make sense of QM under these conditions is certainly a fundamental task for present philosophy of physics.

## 7. References

- Albert, D. (1992). *Quantum mechanics and experience*. Cambridge, Massachusetts: Harvard University Press.
- Bub, J. (1999). *Interpreting the quantum world*. Cambridge: Cambridge University Press. Revised paperback edition of the 1997 edition.
- Butterfield, J. (1993). Forms for probability ascriptions. *International Journal of Theoretical Physics*, 32, 2271-2286.
- Butterfield, J. (1995). Quantum theory and the mind I: worlds, minds and quanta. *Proceedings of the Aristotelian Society Supplementary Volume LXIX*.
- Dirac, P.A.M. (1930/1958). *The principles of quantum mechanics*. Oxford: Clarendon Press, 1930. 4<sup>th</sup> edition, 1958.
- Earman, J. (1986). *A primer on determinism*. Dordrecht: Reidel.
- Fine, A. (1973). Probability and the interpretation of quantum mechanics. *British Journal for the Philosophy of Science*, 24, 1-37.
- Greenberger, D., Hentschel, K., and Weinert, F. (eds.) *Compendium of quantum physics*. Berlin: Springer, 64-70.
- Halpin, J.F. (1991). What is the logical form of probability assignment in quantum mechanics? *Philosophy of Science*, 58, 36-60.
- Held, C. (2008). Axiomatic quantum mechanics and completeness. *Foundations of Physics*, 38, 707-732.
- Held, C. (2009). The Kochen-Specker theorem. In Greenberger et al. (2009), 331-35.

- Held, C. (in preparation). Incompatibility of standard quantum completeness and quantum mechanics.
- Hughes, R.I.G. (1989). *The structure and interpretation of quantum mechanics*. Cambridge, Mass.: Harvard University Press.
- Kochen, S. & Specker, E. (1967). The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17, 59-87.
- Landsman, K. (2009). The Born rule and its interpretation. In Greenberger et al. (2009), 64-70.
- Mermin, N. D. (1990). Simple unified form of the major no-hidden variables theorems. *Physical Review Letters*, 65, 3373-3376.
- Nielsen, M.A. & Chuang, I.L. (2000). *Quantum computation and quantum information*. Cambridge: Cambridge University Press.
- Redhead, M. (1987). *Incompleteness, nonlocality, and realism: a prolegomenon to the philosophy of quantum mechanics*. Oxford: Clarendon Press
- Van Fraassen, B. C. (1981). A modal interpretation of quantum mechanics. In E. Beltrametti and B. C. van Fraassen (eds.) *Current issues in quantum logic* (Singapore: World Scientific): 229-58. Reprinted in L. Sklar (ed.). *The philosophy of physics* (New York: Garland, 2000): 279-308.
- Van Fraassen, B.C. (1991). *Quantum mechanics: an empiricist view*. Oxford: Clarendon Press.
- van Fraassen, B.C. and Hooker, C.A. (1976). A semantic analysis of Niels Bohr's philosophy of quantum theory. In W.L. Harper and C.A. Hooker (eds.): *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science, vol. III: Foundations and Philosophy of Statistical Theories in the Physical Sciences*, (Dordrecht, Reidel) 221-241.
- Von Neumann, J. (1932/1955). *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer, 1932; english translation: *Mathematical foundations of quantum mechanics*. Princeton: Princeton University Press, 1955.

# Quantum Correlations in Successive Spin Measurements

Ali Ahanj

*Department of Physics, Khayyam Higher Education Institute (KHEI), Mashhad, and  
School of Physics, Institute for Research in Fundamental Science (IPM),  
Tehran,  
Iran*

## 1. Introduction

Quantum Mechanics (QM) represents one of the pillars of modern physics: so far a huge amount of theoretical predictions deriving from this theory have been confirmed by very accurate experimental data. No doubts can be raised on the validity of this theory. Nevertheless, even after one century since its birth, many problems related to the interpretation of this theory persist: non-local effects of entangled states, wave function reduction and the concept of measurement in QM, the transition from a microscopic probabilistic world to a macroscopic deterministic world perfectly described by classical mechanics and so on. A possible way out from these problems would be if QM represents a statistical approximation of an unknown deterministic theory, where all observables have well defined values fixed by unknown variables, the so called Hidden Variable Theories (HVT). Therefore, the debate whether QM is a complete theory and probabilities have a non-epistemic character (i.e. nature is intrinsically probabilistic) or whether it is a statistical approximation of a deterministic theory and probabilities are due to our ignorance of some parameters (i.e. they are epistemic) dates to the beginning of the theory itself.

The fundamental paper where this problem clearly emerged appeared in 1935 when Einstein, Podolsky and Rosen asked this question by considering an explicit example (Einstein et al., 1935). For this purpose, they introduced the concept of element of reality according to the following definition: if, without disturbing a system in any way, one can predict with certainty the value of a physical quantity, then there is an element of physical reality corresponding to this quantity. They formulated also the reasonable hypothesis (consistent with special relativity) that every non-local action was forbidden. A theory is complete when it describes every element of reality. They concluded that either one or more of their premises was wrong or Quantum Mechanics was not a complete theory, in the sense that not every element of physical reality had a counterpart in the theory.

This problem led to the search of a "complete theory" by adding hidden variables to the wave function in order to implement realism. For a long time, there was a general belief among quantum physicists that quantum mechanics can not be replaced by some complete theory (HVT) due to Von Neumann's impossibility proof (who imposed an unwarranted constraint on HVT). But in sixties we got two theorems due to J. S. Bell (Bell, 1964) (Bell, 1966) and Kochen and Specker (Kochen & Specker, 1967). These theorems showed that quantum mechanics can

not be replaced by some classes of HVT, namely local and non-contextual HVT. The most celebrated of this kind of HVT was presented by Bohm in 1952 (Bohm, 1952). Bohm, just prior to developing his HV interpretation, introduced a simplified scenario involving two spin-half particles with correlated spins, rather than two particles with correlated positions and momenta as used by EPR. The EPR-Bohm scenario has the advantage of being experimentally accessible.

In 1964 John Bell (Bell, 1964) derived an inequality (which is a statistical result, and is called Bell's inequality BI) using locality and reality assumptions of EPR-Bohm, and showed that the singlet state of two spin-1/2 particles violates this inequality, and hence the contradiction with quantum mechanics.

Contemporary versions of the argument are based on the Clauser, Horne, Shimony and Holt (CHSH) inequality (Clauser et al., 1969), rather than the original inequality used by Bell. There is a very good reason for that. While Bell's argument applied only to the singlet state, the CHSH inequality is violated by all pure entangled states (Gisin & Peres, 1992). Early versions of CHSH inequalities involved only two observers, each one having a choice of two (mutually incompatible) experiments. The various outcomes of each experiment were lumped into two sets, arbitrarily called +1 and -1. Possible generalizations involve more than two observers, or more than two alternative experiments for each observer, or more than two distinct outcomes for each experiment. We may consider  $n$ -partite systems, each subject to a choice of  $m$   $v$ -valued measurements. This gives a total of  $(mv)^n$  experimentally accessible probabilities. The set of Bell inequalities is then the set of inequalities that bounds this region of probabilities to those accessible with a local hidden variable model. Thus for each value of  $n$ ,  $m$  and  $v$  the set of local realistic theories is a polytope bounded by a finite set of linear Bell inequalities. The CHSH inequalities apply to a situation  $(n, m, v) = (2, 2, 2)$ . Gisin et al (Gisin & Bechmann-Pasquinucci, 1998) have found a family of Bell inequalities for the case with the number of measurements is arbitrary, i.e.  $(n, m, v) = (2, m, 2)$ . Collins et al (Collins, Gisin, Linden, Massar & Popescu., 2002) and Kaszlikowski et al (Kaszlikowski et al., n.d.) have produced inequalities for arbitrarily high dimensional systems, i.e.  $(n, m, v) = (2, 2, v)$ . The most complete study of Bell inequalities is for the case  $(n, m, v) = (n, 2, 2)$ .  $n$ -particle generalizations of the CHSH inequality were first proposed by Mermin (Mermin, 1990), and Belinskii and Klyshko (Belinskii & Klyshko, 1993), and have been extended by Werner and Wolf (Werner & Wolf, 2000), and Zukowski and Brukner (Zukowski & Brukner, 2002) to give the complete set for two dichotomic observables per site.

On the theoretical side, "violation of Bell's inequalities" had become synonymous with "non-classical correlation", i.e., entanglement. One of the first papers in which finer distinctions were made was the construction of states with the property that they satisfy all the usual assumptions leading to the Bell inequalities, but can still not be generated by a purely classical mechanism (are not "separable" in modern terminology) (Werner, 1989). This example pointed out a gap between the obviously entangled states (violating a Bell inequality) and the obviously non-entangled ones, which are merely classical correlated (separable). In 1995 Popescu (Popescu, 1995) (and later (Bennett et al., 1996)) narrowed this gap considerably by showing that after local operations and classical communication one could "distill" entanglement, leading once again to violations, even from states not violating any Bell inequality initially. To summarize this phase: it became clear that violations of Bell inequalities, while still a good indicator for the presence of non-classical correlations by no means capture all kinds of "entanglement".

Bell inequalities are statistical predictions about measurements made on two particles,

typically photons or particles with spin  $\frac{1}{2}$ . So some people were trying to show a direct contradiction (which is not a statistical one) of quantum mechanics with local realism. Greenberger, Horne and Zeilinger (GHZ)(Greenberger et al., 1990) found a way to show more immediately, without inequalities, that results of quantum mechanics are inconsistent with the assumptions of EPR. It focuses on just one event, not the statistics of many events. Their proof relies on *eight* dimensional Hilbert space, unlike the case of Bell's theorem, which is valid in *four* dimensions. Heywood and Redhead (Heywood & Redhead, 1983) have provided a direct contradiction (without inequalities) of quantum mechanics with local realism for a particular state of two spin-1 particles. Finally, Hardy (Hardy, 1992) gave a proof of non locality for two particles with spin  $\frac{1}{2}$  that only requires a total of *four* dimensions in Hilbert space like Bell's proof but does not require inequalities. This was accomplished by considering a particular experimental setup consisting of two over-lapping Mach-Zehnder interferometers, one for positrons one for electrons, arranged so that if the electron and positron each take a particular path then they will meet and annihilate one another with probability equal to 1. This arrangement is required to produce assymmetric entangled state which only exhibits non locality without any use of inequality. The argument has been generalized to two spin  $s$  particles by Clifton and Niemann (Clifton & Niemann, 1992) and to  $N$  spin half particles by Pagonis and Clinton (Pagonis & Clifton, 1992).

Later, Hardy showed that this kind of non locality argument can be made for almost all entangled states of two spin- $\frac{1}{2}$  particles except for maximally entangled one (Hardy, 1993). This proof was again simplified by Goldstein (Goldstein, 1994) and extended it to the case of bipartite systems whose constituents belong to Hilbert spaces of arbitrary dimensions.

Conceptually, as well as mathematically, space and time are differently described in quantum mechanics. While time enters as an external parameter in the dynamical evolution of a system, spatial coordinates are regarded as quantum mechanical observables. Moreover, spatially separated quantum systems are associated with the tensor product structure of the Hilbert state-space of the composite system. This allows a composite quantum system to be in a state that is not separable regardless of the spatial separation of its components. We speak about *entanglement in space*. On the other hand, time in quantum mechanics is normally regarded as lacking such a structure. Because of different roles time and space play in quantum theory one could be tempted to assume that the notion of "*entanglement in time*" cannot be introduced in quantum physics.

In this chapter we propose and analyze a particular scenario to account for the deviations of QM from 'realism' (defined below), which involves correlations in the outputs of successive measurements of noncommuting operators on a spin- $s$  state. The correlations for successive measurements have been used previously by Popescu (Popescu, 1995) in the context of nonlocal quantum correlations, in order to analyze a class of Werner states which are entangled but do not break (bipartite) Bell-type inequality. Although local HVT can simulate the quantum correlations between the outputs of single ideal measurement on each part of the system, it fails to simulate the correlations of the second measurements on each part. Leggett and Garg have used consecutive measurements to challenge the applicability of QM to macroscopic phenomena (Leggett & Garg, 1985). While the temporal Bell inequalities, considered in refs. (Leggett & Garg, 1985) (see also (Paz & Mahler, 1993)), are for histories, we deal here with Bell-type inequalities with predetermined measurement values at different times. The temporal Bell inequalities deal with measurement of the same observable at different times, whereas we deal here with different measurements at different times. Finally there is a large literature on the problem of information of a quantum state that can be

obtained by measuring the same operator successively on a single system. The research in this area is elegantly summarized in (Alter & Yamamoto, n.d.). Bell-type inequality with successive measurements was first considered by (Brukner et al., 2004). They have derived CHSH-type inequality (Clauser et al., 1969) for two successive measurements on an arbitrary state of a single qubit and have shown that every such state would violate that inequality for proper choice of the measurement settings. They have shown that the quantum mechanical correlation for three successive measurements, for any single qubit input state is the product of two consecutive correlations each of which is the correlation of two consecutive measurements – a scenario quite uncommon for spatial correlations. As an application of their approach, they have used the correlations in two successive measurements to overcome the limitations in RAM of a computer to calculate a Boolean function whose input bits are supplied sequentially in time.

We consider and analyze the correlations between the outputs of successive measurements for a general spin  $S$  state as against the general qubit state. We show that, for  $S > \frac{1}{2}$ , the quantum mechanical correlation for three successive measurements is not a product of two successive correlations, that is, the correlations in two successive measurements. We show that for  $S = \frac{1}{2}$ , the correlation between the outputs of measurements from  $n - k$  to  $n$  (last  $k$  out of  $n$  successive measurements)  $k = 0, 1, \dots, n - 1$ , depend on the measurement prior to  $n - k$ , when  $k$  is even, while for odd  $k$ , these correlations are independent of the outputs of measurements prior to  $n - k$ . Further, we show that all qubit states break the Bell type inequalities corresponding to  $n$  successive measurements, where  $n$  is any finite number. Finally, we study Hardy's nonlocality arguments for the correlations between the outputs of  $n$  successive measurements for all  $s$ -spin measurements. We show that the maximum probability of success of Hardy's argument in the successive measurement is much higher than the spatial ones in a certain sense.

The chapter is organized as follows. In Section 2 we describe the basic scenario in detail. Section 3 formulates the implications of hidden variable theory (HVT) for this scenario in terms of Bell-type inequalities. Section 4 evaluates these inequalities for mixed input states of single spin- $s$  system for two and three successive measurements (considering various values of  $s$ ). Section 5 deals with  $n$  successive measurements on spin-1/2 system. Section 6 explains the logical structure of Hardy's argument on time locality and, we show that no time-local stochastic HVT (SHVT) can simultaneously satisfy Hardy's argument. Finally we conclude with summary and comments in Section 7. Mathematical details are relegated to Appendices A and B.

## 2. Basic scenario

Consider the following sequence of measurements. A quantum particle with spin  $s$ , prepared in the initial state  $\rho_0$ , is sent through a cascade of Stern-Gerlach (SG) measurements for the spin components along the directions given by the unit vectors  $\hat{a}_1, \hat{a}_2, \hat{a}_3, \dots, \hat{a}_n$  (*i.e.*, measurement of observables of the form  $\vec{S} \cdot \hat{a}_i$ , where  $\vec{S} = (S_x, S_y, S_z)$  is the vector of spin angular momentum operators  $S_x, S_y, S_z$  and  $\hat{a}$  is a unit vector from  $\mathbb{R}^3$ ). Each measurement has  $2s + 1$  possible outcomes. For the  $i$ -th measurement, we denote these outcomes (eigenvalues) by  $\alpha_i \in \{s, s - 1, \dots, -s\}$ . We denote by  $\langle \alpha_i \rangle$  the quantum mechanical (ensemble) average  $\langle \vec{S} \cdot \hat{a}_i \rangle$ , by  $\langle \alpha_i \alpha_j \rangle$  the average  $\langle (\vec{S} \cdot \hat{a}_i)(\vec{S} \cdot \hat{a}_j) \rangle$  etc.

Each of the  $(2s + 1)^n$  possible outcomes, which one gets after performing  $n$  consecutive measurements, corresponds to a particular combination of the results of the measurements at previous  $n - 1$  steps and the result of the measurement at the  $n$ -th step. The probability

of each of these  $(2s + 1)^n$  outcomes is the joint probability for such combinations. Note that even though the spin observables  $\vec{S} \cdot \hat{a}_1, \vec{S} \cdot \hat{a}_2, \dots, \vec{S} \cdot \hat{a}_n$ , whose measurements are being performed at times  $t_1, t_2, \dots, t_n$  respectively (with  $t_1 < t_2 < \dots < t_n$ ) do not commute, above-mentioned joint probabilities for the outcomes are well defined because each of these spin observables act on different states (Fine, 1982) (Anderson et al., 2005) (Ballentine, 1990). We emphasize that this is the joint probability for the results of  $n$  actual measurements and not a joint probability distribution for hypothetical simultaneous values of  $n$  noncommuting observables. Moreover, various sub-beams (*i.e.*, wave functions) emerging from every Stern-Gerlach apparatus (corresponding to  $(2s + 1)$  outcomes) in every stage of measurement are separated without any overlap or recombination between them. In other words, the eigen wave packet  $\psi_{s-j, t_i, \hat{a}_i}(x)$ , corresponding to the eigen value  $s - j$  of the observable  $\vec{S} \cdot \hat{a}_i$ , measured at time  $t_i$ , will not have any part in the regions where the SG setups, for measurement of the observables  $\vec{S} \cdot \hat{a}_{i+1, s}, \vec{S} \cdot \hat{a}_{i+1, s-1}, \dots, \vec{S} \cdot \hat{a}_{i+1, s-j+1}, \vec{S} \cdot \hat{a}_{i+1, s-j-1}, \dots, \vec{S} \cdot \hat{a}_{i+1, -s}$ , are situated. We further assume that, between two successive measurements, the spin state does not change with time *i.e.*  $\vec{S}$  commutes with the interaction Hamiltonian, if any. Also, throughout the string of measurements, no component (*i.e.*, sub-beam) is blocked. It is to be mentioned here that the time of each of the measurements are measured by a common clock.

### 3. Implications of HVT

HVT assumes that in every possible state of the system, all observables have well defined (sharp) values (Redhead, 1987). On the measurement of an observable in a given state, the value possessed by the observable in that state (and no other value) results. To gain compatibility with QM and the experiments, a set of 'hidden' variables is introduced which is denoted collectively by  $\lambda$ . For given  $\lambda$ , the values of all observables are specified as the values of appropriate real valued functions defined over the domain  $\Lambda$  of possible values of hidden variables. For the spin observable  $\vec{S} \cdot \hat{a}$ , we denote the value of  $\vec{S} \cdot \hat{a}$  in the QM (spin) state  $|\psi\rangle$  by  $\alpha$ . Considered as a function,  $\alpha : \Lambda \rightarrow \mathbf{R}$ , we represent the value of  $\vec{S} \cdot \hat{a}$  when the hidden variables have the value  $\lambda$  by  $\alpha(\lambda)$ . More generally, we may require that a value of  $\lambda$  gives the probability density  $p(\alpha|\lambda)$  over the values of  $\alpha$  rather than specifying the value of  $\alpha$  (stochastic HVT). We denote the probability density function for the hidden variables in the state  $|\psi\rangle$  by  $\rho_\psi$  ( $\rho_\psi(\lambda)d\lambda$  measures the probability that the collective hidden variable lies in the range  $\lambda$  to  $\lambda + d\lambda$ ). Then the average value of  $\vec{S} \cdot \hat{a}$  in the state  $|\psi\rangle$  is

$$\langle \alpha \rangle = \int_{\Lambda} \alpha(\lambda) \rho_\psi(\lambda) d\lambda, \quad (1)$$

where the integration is over  $\Lambda$  defined above. In the general case (SHVT)

$$\langle \alpha \rangle = \int_{\Lambda} \alpha p(\alpha|\lambda) \rho_\psi(\lambda) d\lambda. \quad (2)$$

We now analyze the consequences of SHVT for our scenario. In general, the outputs of  $k$ th and  $l$ th experiments may be correlated so that,

$$p(\alpha_i, \hat{a}_k \& \alpha_j, \hat{a}_l) \neq p(\alpha_i; \hat{a}_k) p(\alpha_j; \hat{a}_l). \quad (3)$$

However, in SHVT we suppose that these correlations have a common cause represented by a stochastic hidden variable  $\lambda$  so that

$$p(\alpha_i, \hat{a}_k \& \alpha_j, \hat{a}_\ell | \lambda) = p(\alpha_i, \hat{a}_k | \lambda) p(\alpha_j, \hat{a}_\ell | \lambda). \tag{4}$$

As a consequence of equation (4), the probability  $p(\alpha_i, \hat{a}_k | \lambda)$  obtained in a measurement ( $\vec{S} \cdot \hat{a}_k$  say) performed at time  $t_k$  is independent of any other measurement ( $\vec{S} \cdot \hat{a}_l$  say) made at some earlier or later time  $t_l$ . This is called locality in time (Leggett & Garg, 1985) (Brukner et al., 2004).

One should note that for a two dimensional quantum mechanical system, one can always assign values (deterministically or probabilistically) to the observables with the help of a HVT. Once the measurement is done, the system will be prepared in an output state (namely, an eigenstate of the observable), and the earlier HVT may or may not work to reproduce the values of the observables to be measured on that output state (prepared after the first measurement). We have considered possibility of existence of a HVT for every input qubit state which can reproduce the measurement outcomes of  $n$  successive measurements.

Equation (4) is the crucial equation expressing the fundamental implication of SHVT to the successive measurement scenario. We now obtain the Bell type inequalities from equation (4) which can be compared with QM. Here we assume that in HVT all probabilities corresponding to outputs of measurements account for the possible changes in the values of the observable being measured, (due to the interaction of the measuring device and the system), occurring in the previous measurements.

Now  $\langle \alpha_i \alpha_j \rangle$  is the expectation value of obtaining the outcome  $\alpha_i$  in the measurement of the observable  $\vec{S} \cdot \hat{a}_i$  at time  $t_i$  as well as the outcome  $\alpha_j$  in the measurement of the observable  $\vec{S} \cdot \hat{a}_j$  at later time  $t_j$ . Due to the HVT, we must have (dropping  $\hat{a}_k, \hat{a}_\ell$ )

$$\langle \alpha_i \alpha_j \rangle = \int \rho(\lambda) E(\alpha_i, \alpha_j, \lambda) d\lambda, \tag{5}$$

where

$$\begin{aligned} E(\alpha_i, \alpha_j, \lambda) &= \sum_{\alpha_i, \alpha_j} \alpha_i \alpha_j p(\alpha_i, \alpha_j | \lambda) = \sum_{\alpha_i} \alpha_i p(\alpha_i | \lambda) \sum_{\alpha_j} \alpha_j p(\alpha_j | \lambda) \\ &= E(\alpha_i, \lambda) E(\alpha_j, \lambda) \end{aligned} \tag{6}$$

by equation (4). Now let us consider the case of two successive measurements, with options  $\hat{a}_1, \hat{a}'_1$  and  $\hat{a}_2, \hat{a}'_2$  respectively for measuring spin components. In each run of the experiment, a random choice between  $\{\hat{a}_1, \hat{a}'_1\}$  and  $\{\hat{a}_2, \hat{a}'_2\}$  is made. Define  $\theta_i$  ( $i = 1, 1'$ ) to be the angle between  $\hat{a}_i$  and the positive z-axis,  $\theta_{ij}$  ( $i = 1, 1'$  and  $j = 2, 2'$ ) is the angle between  $\hat{a}_j$  and  $\hat{a}_i$ . Using condition (6) and the result (Shimony, n.d.) (Jarrett, 1984)

$$-2s^2 \leq xy + xy' + x'y - x'y' \leq 2s^2, \quad x, y, x', y' \in \{-s, -s+1, \dots, s-1, s\},$$

we obtain

$$-2s^2 \leq E(\alpha_1, \alpha_2, \lambda) + E(\alpha_1, \alpha'_2, \lambda) + E(\alpha'_1, \alpha_2, \lambda) - E(\alpha'_1, \alpha'_2, \lambda) \leq 2s^2. \tag{7}$$

Multiplying by  $\rho(\lambda)d\lambda$  and integrating over  $\Lambda$ , we get the CHSH-type inequality (Clauser et al., 1969) (involving the hidden variable  $\lambda$ ) corresponding to performing two successive measurements of spin-s observables on a spin-s initial state:

$$|\langle BI \rangle| = \frac{1}{2} |\langle \alpha_1 \alpha_2 \rangle + \langle \alpha_1 \alpha'_2 \rangle + \langle \alpha'_1 \alpha_2 \rangle - \langle \alpha'_1 \alpha'_2 \rangle| \leq s^2. \tag{8}$$

Similarly, using the algebraic fact

$$-2s^3 \leq xyz' + xy'z + x'y'z - x'y'z' \leq 2s^3,$$

where

$$x, y, z, x', y', z' \in \{-s, -s+1, \dots, s-1, s\}$$

and<sup>1</sup>

$$E(\alpha_i, \alpha_j, \alpha_k, \lambda) = E(\alpha_i, \lambda)E(\alpha_j, \lambda)E(\alpha_k, \lambda),$$

we can prove Mermin-Klyshko Inequality (MKI) (Mermin, 1990), (Belinskii & Klyshko, 1993) for three successive measurements,

$$|\langle MKI \rangle| = \frac{1}{2} |\langle \alpha_1 \alpha_2 \alpha'_3 \rangle + \langle \alpha_1 \alpha'_2 \alpha_3 \rangle + \langle \alpha'_1 \alpha_2 \alpha_3 \rangle - \langle \alpha'_1 \alpha'_2 \alpha'_3 \rangle| \leq s^3. \quad (9)$$

Let  $|\langle MKI' \rangle| \leq s^3$ , where  $|\langle MKI' \rangle|$  is obtained from equation (9) by interchanging primes with non-primes in MKI. It is easily shown that

$$|\langle SI \rangle| = |\langle MKI \rangle + \langle MKI' \rangle| \leq |\langle MKI \rangle| + |\langle MKI' \rangle| \leq 2s^3. \quad (10)$$

This is the Svetlichny inequality (SI) (Svetlichny, 1987), (Seevinck & Svetlichny, 2002), (Collins, Gisin, Popescu, Roberts & Scarani, 2002).

For  $n$  successive measurements on spin  $s$  system, we define the MK polynomials recursively as follows:

$$M_1 = \alpha_1, \quad M'_1 = \alpha'_1, \quad (11)$$

$$M_n = \frac{1}{2} M_{n-1} (\alpha_n + \alpha'_n) + \frac{1}{2} M'_{n-1} (\alpha_n - \alpha'_n), \quad (12)$$

where  $M'_n$  are obtained from  $M_n$  by interchanging all primed and non-primed  $\alpha$ 's. The recursive relation (12) gives, for all  $1 \leq k \leq n-1$  (Collins, Gisin, Popescu, Roberts & Scarani, 2002), (Cabello, 2002a):

$$M_n = \frac{1}{2} M_{n-k} (M_k + M'_k) + \frac{1}{2} M'_{n-k} (M_k - M'_k). \quad (13)$$

In particular, we have

$$M_2 = BI = \frac{1}{2} (\alpha_1 \alpha_2 + \alpha'_1 \alpha_2 + \alpha_1 \alpha'_2 - \alpha'_1 \alpha'_2), \quad (14)$$

$$M_3 = MKI = \frac{1}{2} (\alpha_1 \alpha_2 \alpha'_3 + \alpha_1 \alpha'_2 \alpha_3 + \alpha'_1 \alpha_2 \alpha_3 - \alpha'_1 \alpha'_2 \alpha'_3). \quad (15)$$

We now show that in HVT,

$$|\langle M_n \rangle| \leq s^n. \quad (16)$$

<sup>1</sup> This is obtained by using equation (4) and the similar argument as has been used in deriving equation (6).

First note that (16) is true for  $n = 2, 3$  (equations (8), (9)). Suppose it is true for  $n = k$  i.e.  $\text{Max}|\langle M_k \rangle| = s^k$ . Now

$$|\langle M_{k+1} \rangle| = \frac{1}{2} |\langle M_k \alpha_{k+1} \rangle + \langle M_k \alpha'_{k+1} \rangle + \langle M'_k \alpha_{k+1} \rangle - \langle M'_k \alpha'_{k+1} \rangle|.$$

Since HVT applies here we can use (4) to get

$$|\langle M_{k+1} \rangle| = \frac{1}{2} |\langle M_k \rangle (\langle \alpha_{k+1} \rangle + \langle \alpha'_{k+1} \rangle) + \langle M'_k \rangle (\langle \alpha_{k+1} \rangle - \langle \alpha'_{k+1} \rangle)|.$$

This implies, by induction hypothesis (and using the fact that  $\max|\langle M_2 \rangle| = s^2$ ), that

$$\max |\langle M_{k+1} \rangle| = s \max |\langle M_k \rangle| = s^{k+1}.$$

This result is derived for  $n$  spin- $s$  particles by Cabello (Cabello, 2002a).

We now define a quantity, denoted by  $\eta_n$ , which will be required later on.  $\eta_n$ , is the ratio between maximum  $|\langle M_n \rangle|$  given by quantum correlation between  $n$  successive measurement's outputs and the maximal classical one,

$$\eta_n = \frac{\max |\langle M_n \rangle_{QM}|}{s^n}. \quad (17)$$

## 4. Mixed input state for arbitrary spin

### 4.1 Two successive measurements (BI)

We first deal with the case when input state is a mixed state whose eigenstates coincide with those of  $\vec{S} \cdot \hat{a}_0$  for some  $\hat{a}_0$  whose eigenvalues we denote by  $\alpha_0 \in \{-s, \dots, s\}$ . For spin 1/2 this is the most general mixed state because given any density operator  $\rho_0$  for spin 1/2 (corresponding to some point within the Bloch sphere), we can find an  $\hat{a}_0$  such that the eigenstates of  $\vec{S} \cdot \hat{a}_0$  and  $\rho_0$  coincide. However, for  $s > 1/2$ , our choice forms a restricted class of mixed states. We note that these are the only states accessible via SG experiments. Thus we have

$$\rho_0 = \sum_{\alpha_0} p_{\alpha_0} |\vec{S} \cdot \hat{a}_0, \alpha_0\rangle \langle \vec{S} \cdot \hat{a}_0, \alpha_0|; \quad \left( \sum_{\alpha_0} p_{\alpha_0} = 1 \right) \quad (18)$$

After the first measurement along  $\hat{a}_1$ , the resulting state of the system is

$$\rho_1 = \sum_{\alpha_1} M_{\alpha_1}^\dagger \rho_0 M_{\alpha_1}, \quad (19)$$

where

$$M_{\alpha_1}^\dagger = M_{\alpha_1} = |\vec{S} \cdot \hat{a}_1, \alpha_1\rangle \langle \vec{S} \cdot \hat{a}_1, \alpha_1|.$$

Now  $\langle \alpha_1 \alpha_2 \rangle_{QM}$  is the expectation value (according to QM) that given the initial state  $\rho_0$  (given in equation (18)), the 1st measurement along  $\hat{a}_1$  will give rise to any value  $\alpha_1 \in \{-s, -s+1, \dots, s-1, s\}$ , and then, on the after-measurement state  $\rho_1$  (given in equation (19)), if one

performs measurement along  $\hat{a}_2$ , one of the values  $\alpha_2 \in \{-s, -s + 1, \dots, s - 1, s\}$  will arise. So

$$\begin{aligned} \langle \alpha_1 \alpha_2 \rangle_{QM} &= \text{Tr}(\rho_1 \vec{S} \cdot \hat{a}_1 \vec{S} \cdot \hat{a}_2) = \\ &= \sum_{\alpha_0 \alpha_1 \alpha_2} p_{\alpha_0} \alpha_1 \alpha_2 |\langle \vec{S} \cdot \hat{a}_0, \alpha_0 | \vec{S} \cdot \hat{a}_1, \alpha_1 \rangle|^2 |\langle \vec{S} \cdot \hat{a}_1, \alpha_1 | \vec{S} \cdot \hat{a}_2, \alpha_2 \rangle|^2. \end{aligned} \tag{20}$$

Note that, since  $\vec{S} \cdot \hat{a}_i$  are complete observables, all of whose eigenvalues are non degenerate, the probabilities factorize like those of a Markov chain (Beck & Graudenz, 1992). Every factor in (20) corresponds to the transition amplitude between two successive measurements. By equation (A.12), we get

$$\langle \alpha_1 \alpha_2 \rangle = \frac{1}{2} \cos \theta_{12} [A \cos^2 \theta_1 + B], \tag{21}$$

where

$$A = 3\chi - s(s + 1), \quad B = s(s + 1) - \chi, \quad \chi = \sum_{\alpha_0=-s}^{+s} \alpha_0^2 p_{\alpha_0}.$$

It is to be noted that although  $A$  can have positive and negative values,  $B$  will always be positive. Moreover, for all  $\theta \in [0, 2\pi]$ , if  $A \geq 0$ ,  $A \cos^2 \theta_1 + B$  is always positive and if  $A < 0$ ,  $A \cos^2 \theta_1 + B \geq B + A = 2\chi \geq 0$ . We now have the following expression for the quantity  $BI$ , appeared in equation (8):

$$\begin{aligned} BI &= \frac{1}{4} \{ (A \cos^2 \theta_1 + B) (\cos \theta_{12} + \cos \theta'_{12}) + (A \cos^2 \theta'_1 + B) (\cos \theta''_{12} - \cos \theta'''_{12}) \} \\ &= \frac{3\chi - s(s + 1)}{4} \{ \cos^2 \theta_1 (\cos \theta_{12} + \cos \theta'_{12}) + \cos^2 \theta'_1 (\cos \theta''_{12} - \cos \theta'''_{12}) \} \\ &+ \frac{s(s + 1) - \chi}{4} \{ (\cos \theta_{12} + \cos \theta'_{12}) + (\cos \theta''_{12} - \cos \theta'''_{12}) \}, \end{aligned} \tag{22}$$

where (according to Appendix A)  $\theta_1$  is the angle between  $\hat{a}_0$  and  $\hat{a}_1$ ,  $\theta'_1$  is the angle between  $\hat{a}_0$  and  $\hat{a}'_1$ ,  $\theta_{12}$  is the angle between  $\hat{a}_1$  and  $\hat{a}_2$ ,  $\theta'_{12}$  is the angle between  $\hat{a}_1$  and  $\hat{a}'_2$ ,  $\theta''_{12}$  is the angle between  $\hat{a}'_1$  and  $\hat{a}_2$ ,  $\theta'''_{12}$  is the angle between  $\hat{a}'_1$  and  $\hat{a}'_2$ . We have used, in Eq (22), the expressions for  $A$  and  $B$  in terms of  $\chi$  and  $s$ . Note that the second term in equation (22) (*i.e.*, the term with the factor  $\frac{s(s+1)-\chi}{4}$ ) is similar to the expression for  $\langle \alpha_1 \alpha_2 \rangle + \langle \alpha_1 \alpha'_2 \rangle + \langle \alpha'_1 \alpha_2 \rangle - \langle \alpha'_1 \alpha'_2 \rangle$  corresponding to the CHSH inequality (Clauser et al., 1969). And hence, its maximum value will occur when we choose all the four vectors  $\hat{a}_1, \hat{a}'_1, \hat{a}_2, \hat{a}'_2$  on the same plane. But we also have to take care about maximization of the first term in equation (22) and that might require these four vectors to be on different planes. In order to resolve this issue, we now consider the spherical-polar co-ordinates  $(\theta_1, \phi_1), (\theta'_1, \phi'_1), (\theta_2, \phi_2), (\theta'_2, \phi'_2)$  of the vectors  $\hat{a}_1, \hat{a}'_1, \hat{a}_2, \hat{a}'_2$  respectively, where  $\theta_1, \theta'_1, \theta_2, \theta'_2 \in [0, \pi]$  and  $\phi_1, \phi'_1, \phi_2, \phi'_2 \in [0, 2\pi]$ . Then  $BI$  has the form

$$\begin{aligned} BI &= \frac{1}{4} (A \cos^2 \theta_1 + B) [\cos \theta_1 (\cos \theta_2 + \cos \theta'_2) + \sin \theta_1 \sin \theta_2 \cos(\phi_1 - \phi_2) \\ &+ \sin \theta_1 \sin \theta'_2 \cos(\phi_1 - \phi'_2)] + \frac{1}{4} (A \cos^2 \theta'_1 + B) [\cos \theta'_1 (\cos \theta_2 - \cos \theta'_2) \\ &+ \sin \theta'_1 \sin \theta_2 \cos(\phi'_1 - \phi_2) - \sin \theta'_1 \sin \theta'_2 \cos(\phi'_1 - \phi'_2)]. \end{aligned} \tag{23}$$

Here also the maximum value of  $|BI|$  will occur when all the vectors  $\hat{a}_1, \hat{a}'_1, \hat{a}_2, \hat{a}'_2$  lie on the same plane. This is obtained by:

$$\frac{\partial BI}{\partial \phi_1} = \frac{\partial BI}{\partial \phi_2} = \frac{\partial BI}{\partial \phi'_1} = \frac{\partial BI}{\partial \phi'_2} = 0 \Rightarrow \phi_1 = \phi'_1 = \phi_2 = \phi'_2.$$

In that case, the maximum value of  $|BI|$  will occur when  $\theta'_1 = \pi - \theta_1, \theta_2 = \pi/2, \theta'_2 = 0$  and (correspondingly) the quantity

$$\eta_2 = \frac{|BI|}{s^2} = \left( \frac{1}{2s^2} \right) |(\sin \theta_1 + \cos \theta_1)|(A \cos^2 \theta_1 + B) \quad (24)$$

is maximized over all possible values of  $\theta_1$  ( $A \cos^2 \theta_1 + B \geq 0$ ). If  $\eta_2 > 1$ , the correlations for two successive measurements violate the CHSH-type inequality (8), and hence a contradiction with the above-mentioned HVT. In fact  $\frac{\partial \eta_2}{\partial \theta_1} = 0$  implies that

$$B \tan^3 \theta_1 + (2A - B) \tan^2 \theta_1 + (3A + B) \tan \theta_1 - (A + B) = 0. \quad (25)$$

Real roots (for  $\tan \theta_1$ ) of this equation give values of  $\theta_1$  for which  $\eta_2$  is maximum. The maximum value of  $\eta_2$  is evaluated at these  $\theta_1$ 's.

We find that for  $s = \frac{1}{2}, \chi = 1/4$  for all  $\rho_0$ , and so  $A = 0, B = 1/4$ . So, from equation (24), we have  $\eta_2 = \sin \theta_1 + \cos \theta_1$ . Therefore equation (25) becomes  $\tan^3 \theta_1 - \tan^2 \theta_1 + \tan \theta_1 - 1 = 0$ , whose only one real solution is  $\tan \theta_1 = 1$ . So  $\theta_1 = \pi/4$  or  $5\pi/4$ .  $\theta_1 = \pi/4$  gives the maximum possible value  $\eta_2 = \sqrt{2} > 1$ . Thus all possible spin-1/2 states break BI for (proper choices of) two successive measurements. This can be compared with the measurement correlations corresponding to measurement of spin observables on space-like separated two particles scenario where only the entangled pure states break BI while not all entangled mixed states break it (Werner, 1989).

From now on, we will use the range of values of the quantity  $\xi \equiv \chi/s^2$  to identify the parametric region of the initial density matrix  $\rho_0$  where the inequality (8) will be violated. Thus we see that for all spin-1/2 input states  $\rho_0, \xi = 1$ .

For a spin-1 system, we first consider all input states  $\rho_0$  none of which have a contribution of  $S_z = 0$  eigenstate. In this case  $\chi = 1, A = B = 1$ . So  $\eta_2 = (1/2)(\sin \theta_1 + \cos \theta_1)(\cos^2 \theta_1 + 1)$  and equation (25) takes the form  $\tan^3 \theta_1 + \tan^2 \theta_1 + 4 \tan^3 \theta_1 - 2 = 0$ . The only real root of this equation is  $\tan \theta_1 \approx 0.433$ . Thus the maximum possible value of  $\eta_2$  is (using equation (24)) 1.2112 (approximately). Thus we see that all input spin-1 states  $\rho_0$ , none of which has a component along  $|S_z = 0\rangle$ , break BI (equation (8)) for proper choice of the observables.

Next, for  $s = 1$ , we consider the state  $\rho_0$ , for which  $p_{\alpha_0=0} = 1$ , i.e.,  $\rho_0 = |\vec{S} \cdot \hat{a}_0, 0\rangle \langle \vec{S} \cdot \hat{a}_0, 0|$ . In this case,  $\chi = 0$ , and so,  $A = -2, B = 2, \xi = 0$ . Then equation (25) takes the form  $2 \tan^3 \theta_1 - 6 \tan^2 \theta_1 - 4 \tan \theta_1 = 0$ . It has three real solutions, which corresponds to  $\theta_1 = 0$  (or  $\pi$ ),  $74.3165^\circ$  (approx.),  $150.6836^\circ$  (approx.). The maximum possible value of  $\eta_2$  occurs at  $\theta_1 = 74.3165^\circ$ , and the corresponding value is given by  $\eta_2 \approx 1.1428$ . Thus the state  $\rho_0 = |\vec{S} \cdot \hat{a}_0, 0\rangle \langle \vec{S} \cdot \hat{a}_0, 0|$  breaks the BI (equation (8)).

For  $s = 1$ , when  $0 < p_{\alpha_0=0} \equiv p_0$  (say)  $< 1$ , we have  $\chi = 1 - p_0 = \xi, A = 1 - 3p_0 = 3\xi - 2$  and  $B = 1 + p_0 = 2 - \xi$ . We then have  $\eta_2 = (1/2)|\sin \theta_1 + \cos \theta_1| \{ (3\xi - 2) \cos^2 \theta_1 + (2 - \xi) \}$  (by equation (24)), and equation (25) becomes  $(2 - \xi) \tan^3 \theta_1 + (7\xi - 6) \tan^2 \theta_1 + 4(2\xi - 1) \tan \theta_1 - 2\xi = 0$ . In this case, one can show numerically that the BI will break (i.e.,  $\eta_2 > 1$ ) if and only if either  $0 < \xi < 0.33$  or  $0.77 < \xi < 1$  (equivalently, either  $0.67 < p_0 < 1$  or  $0 < p_0 < 0.23$ ).

Thus we see that, when  $s = 1$ , only those input states  $\rho_0$  ( given in equation (18)) will break BI for each of which  $p_{\alpha_0=0} \in [0, 0.23] \cup (0.67, 1]$ . Next we consider the situations where  $s > 1$ . Note that, by definition (true for all  $s$ ),

$$\xi = (p_s + p_{-s}) + (p_{s-1} + p_{-s+1}) \left(1 - \frac{1}{s}\right)^2 + (p_{s-2} + p_{-s+2}) \left(1 - \frac{2}{s}\right)^2 + \dots,$$

where  $0 \leq p_s, p_{-s}, p_{s-1}, p_{-s+1}, p_{s-2}, p_{-s+2}, \dots \leq 1$  and  $\sum_{\alpha_0=-s}^s p_{\alpha_0} = 1$ . Therefore, we must have  $0 \leq \xi \leq 1$ . In this case, equations (24) and (25) respectively take the forms

$$\eta_2 = \frac{1}{2s^2} |\sin\theta_1 + \cos\theta_1| \left[ \{3\xi s^2 - s(s+1)\} \cos^2\theta_1 + \{s(s+1) - \xi s^2\} \right], \quad (26)$$

$$\begin{aligned} & \{s(s+1) - \xi s^2\} \tan^3\theta_1 + \{7\xi s^2 - 3s(s+1)\} \tan^2\theta_1 + \\ & \{8\xi s^2 - 2s(s+1)\} \tan\theta_1 - 2\xi s^2 = 0. \end{aligned} \quad (27)$$

Let us first consider the input states of the form

$$\rho_0 = p_s |\vec{S} \cdot \hat{a}_0, s\rangle \langle \vec{S} \cdot \hat{a}_0, s| + p_{-s} |\vec{S} \cdot \hat{a}_0, -s\rangle \langle \vec{S} \cdot \hat{a}_0, -s|, \quad (28)$$

with  $p_s + p_{-s} = 1$  and  $p_{s-1}, p_{-s+1}, p_{s-2}, p_{-s+2}, \dots = 0$ . Thus we see here that  $\xi = 1$ . Also equations (26) and (27) have respectively been turned into the forms

$$\eta_2 = (1/2s) |\sin\theta_1 + \cos\theta_1| \{(2s-1)\cos^2\theta_1 + 1\}, \quad (29)$$

$$\tan^3\theta_1 + (4s-3)\tan^2\theta_1 + (6s-2)\tan\theta_1 - 2s = 0. \quad (30)$$

As here  $s > 1$ , therefore the last equation will have only one positive root and the other two roots will be complex. The positive root will correspond to an angle  $\theta_1^{max}(s) \in (0, \pi/4)$  for which it can be shown that  $\eta_2^{max}(s) \equiv \eta_2(\theta_1^{max}(s)) > 1$  for all  $s > 1$ . Hence, in this case, BI is violated. If  $\rho_0$  has contribution from neither of the states corresponding to  $\alpha_0 = \pm s$  (i.e.,  $p_s = p_{-s} = 0$ ), we have

$$\xi = (p_{s-1} + p_{-s+1}) \left(1 - \frac{1}{s}\right)^2 + (p_{s-2} + p_{-s+2}) \left(1 - \frac{2}{s}\right)^2 + \dots \leq \left(1 - \frac{1}{s}\right)^2.$$

From equation (24) it follows that

$$\begin{aligned} \eta_2 &= \frac{1}{\sqrt{2s}} |\sin(\theta_1 + \pi/4)| \left[ (s+1-s\xi) + (3s\xi-s-1)\cos^2\theta_1 \right] \\ &< \frac{1}{\sqrt{2s}} \times 1 \times [(s+1-s\xi) + (3s\xi-s-1) \times 1] = \sqrt{2}\xi \leq \sqrt{2} \left(1 - \frac{1}{s}\right)^2. \end{aligned}$$

But the quantity  $\sqrt{2}(1-1/s)^2$  is less than 1 for all  $s = 1/2, 1, 3/2, \dots, 6$ . Therefore, for  $s > 1$ , if the initial state  $\rho_0$  has contribution from neither of the states corresponding to  $\alpha_0 = \pm s$ , BI will be satisfied for all  $s \leq 6$ . Thus we see that whenever  $s \in \{3/2, 2, 5/2, \dots, 6\}$ , in order that  $\rho_0$  violates BI, the associated quantity  $\xi$  must have values near 1. In table 1, we have given the ranges of values of  $\xi$  (obtained numerically) for which BI is violated, starting from  $s = 1/2$ . The case when  $s \rightarrow \infty$  has also been considered in table 1.

$s$	$\zeta$	$s$	$\zeta$	$s$	$\zeta$
$\frac{1}{2}$	$\zeta = 1$	$\frac{5}{2}$	$0.847 \leq \zeta \leq 1$	$\frac{9}{2}$	$0.858 \leq \zeta \leq 1$
1	$0 \leq \zeta \leq 0.33$ and $0.77 \leq \zeta \leq 1$	3	$0.851 \leq \zeta \leq 1$	5	$0.859 \leq \zeta \leq 1$
$\frac{3}{2}$	$0.824 \leq \zeta \leq 1$	$\frac{7}{2}$	$0.854 \leq \zeta \leq 1$	$\frac{11}{2}$	$0.860 \leq \zeta \leq 1$
2	$0.84 \leq \zeta \leq 1$	4	$0.856 \leq \zeta \leq 1$	6	$0.862 \leq \zeta \leq 1$
				$\infty$	$0.87 \leq \zeta \leq 1$

Table 1. The ranges of  $\zeta$ , for which BI is violated.

$s$	$\eta_2$	$s$	$\eta_2$	$s$	$\eta_2$
$\frac{1}{2}$	$\sqrt{2}$	$\frac{5}{2}$	1.1638	$\frac{9}{2}$	1.1538
1	1.2112	3	1.1599	5	1.1526
$\frac{3}{2}$	1.1817	$\frac{7}{2}$	1.1572	$\frac{11}{2}$	1.1517
2	1.17	4	1.1553	6	1.1509
				$\infty$	1.143

Table 2. The maximum violation of BI for different spin values for two successive measurements.

The maximum violation of Bell inequality, characterized by  $\eta_2$ , decreases monotonically with  $s$ . Table 2 summarizes ( obtained numerically) the maximum allowed value of  $\eta_2$  for each  $s$ . We see from this table that for all spin values  $s$ , BI is broken. Note that there is a sharp decrease in  $\eta_2$  from  $s = \frac{1}{2}$  to  $s = 1$ , while  $\eta_2$  decreases slowly as  $s$  increases from 1. A possible reason is that, for  $s = 1/2$ , all states break BI while for  $s \geq 1$ , only a fraction of spin states break it. We now consider a case where the initial state  $\rho_0^{max}$  (given in equation (28)) is contaminated by the maximally noisy state, resulting in the state

$$\rho(f) = (1 - f) \rho_0^{max} + \frac{f}{2s + 1} I, \tag{31}$$

where the positive parameter  $f (\leq 1)$  is the probability of the noise contamination of the state  $\rho_0^{max}$ . Proceeding as before ( see equation (21)), we get

$$\langle \alpha_1 \alpha_2 \rangle = \frac{1}{2} \cos \theta_{12} [A' \cos^2 \theta_1 + B'] \tag{32}$$

where

$$A' = (1 - f)(2s - 1)s; \quad B' = (1 - f)s + \frac{2}{3} f(s + 1)s,$$

which leads to

$$\eta_{noise} = \left( \frac{1}{2s^2} \right) (\sin \theta_1 + \cos \theta_1) (A' \cos^2 \theta_1 + B'). \tag{33}$$

Using the maximization procedure (i.e., taking  $\frac{\partial \eta_{noise}}{\partial \theta_1} = 0$ ),  $\tan \theta_1$  for maximum  $\eta_{noise}$  is given by a real root of

$$B' \tan^3 \theta_1 + (2A' - B') \tan^2 \theta_1 + (3A' + B') \tan \theta_1 - (A' + B') = 0. \tag{34}$$

The range of  $f$  for which  $\eta_{noise} > 1$  is tabulated in table 3. Note that for  $s = \frac{1}{2}$  the state corresponding to  $f = 1$  ( the random mixture) also breaks BI! Of course we have already

$s$	$f$	$s$	$f$	$s$	$f$
$\frac{1}{2}$	$0 \leq f \leq 1$	$\frac{5}{2}$	$f < 0.287$	$\frac{9}{2}$	$f < 0.239$
1	$f < 0.696$	3	$f < 0.267$	5	$f < 0.234$
$\frac{3}{2}$	$f < 0.395$	$\frac{7}{2}$	$f < 0.254$	$\frac{11}{2}$	$f < 0.230$
2	$f < 0.321$	4	$f < 0.245$	6	$f < 0.227$
				$\infty$	$f < 0.195$

Table 3. The range of the noise  $f$  over which BI is violated.

shown that for  $s = \frac{1}{2}$ , BI is broken for all states. This indicates that the notion of “classicality”, compatible with the usual local HVT, is different in nature from the notion of classicality that would arise from the non-violation of BI here.

Table 3 answers the question, “what is the maximum fraction of noise that can be added to  $\rho_0^{\max}$ , which maximally breaks BI, so that the state has stronger than “classical correlations<sup>2</sup>?” We see that the corresponding fraction of noise (*i.e.*, for which BI is violated) decreases monotonically with  $s$ , or with the dimension of the Hilbert space. This may be compared with the results of Collins and Popescu (Collins & Popescu, 2001) who found that the nonlocal character of the correlations between the outcomes of measurements performed on entangled systems separated in space is robust in the presence of noise. They showed that, for any fraction of noise, by taking the Hilbert space of large enough dimension, one can find bipartite entangled states giving nonlocal correlations. These results have been obtained by considering two successive measurements on each part of the system. On the other hand, in the present case of successive measurements on the single spin state, we see that the fraction of noise that can be added so that the quantum correlations continue to break Bell inequality, falls off monotonically with  $s$ , or the dimension of the Hilbert space. For  $s = \frac{1}{2}$  all fractions  $f \leq 1$  are allowed, while for large  $s$ ,  $f < 0.195$ .

### 4.2 Three successive measurements (MKI)

We again assume the input state to be given by equation (18). Using equation (A.19) :

$$\langle \alpha_1 \alpha_2 \alpha_3 \rangle = \frac{1}{16} \cos \theta_{23} \{ \cos \theta_1 [M \cos^2 \theta_{12} + N] + R [3 \cos^2 \theta_{12} - 1] \} \tag{35}$$

where

$$M = \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} \alpha_0 [9\alpha_0^2 + s(s+1) - 3],$$

$$N = \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} \alpha_0 [5s(s+1) - 3\alpha_0^2 + 1],$$

$$R = \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} \alpha_0 [5\alpha_0^2 - 3s(s+1) + 1],$$

$\theta_1$  is the angle between  $\hat{a}_0$  and  $\hat{a}_1$  (measured with respect to the right-handed system  $(\hat{a}_0, \hat{a}_1, (\hat{a}_0 \times \hat{a}_1)/|\hat{a}_0 \times \hat{a}_1|)$ ),  $\theta_{12}$  is the angle between  $\hat{a}_1, \hat{a}_2$  (measured with respect to the right-handed system  $(\hat{a}_1, \hat{a}_2, (\hat{a}_1 \times \hat{a}_2)/|\hat{a}_1 \times \hat{a}_2|)$ ), etc.

<sup>2</sup> *i.e.*, correlations obeying “realism” and “locality in time”, as described in section 3.

$s$	$\eta_3$	$s$	$\eta_3$	$s$	$\eta_3$
$\frac{1}{2}$	$\sqrt{2}$	$\frac{5}{2}$	1.1736	$\frac{9}{2}$	1.1634
1	1.2178	3	1.1698	5	1.1621
$\frac{3}{2}$	1.1907	$\frac{7}{2}$	1.1670	$\frac{11}{2}$	1.1610
2	1.1793	4	1.1650	6	1.1601
				$\infty$	1.1527

Table 4. The maximum violation of MKI for different spin values for three successive measurements.

We now consider the pure state  $|\vec{S} \cdot \hat{a}_0, s\rangle \langle \vec{S} \cdot \hat{a}_0, s|$  instead of considering the most general state  $\rho_0$ , given in equation (18). So here  $M = s(2s - 1)(5s + 3)$ ,  $N = s(2s^2 + 5s + 1)$ , and  $R = s(2s - 1)(s - 1)$ . Substituting the correlations like that in equation (35) in the MKI (given in equation (9)), using the above-mentioned values of  $M$ ,  $N$ ,  $R$ , and then finding out the conditions (numerically) for which  $\eta_3 \equiv |MKI|/s^3$  is maximized, we get the maximum possible  $\eta_3$ -values for different spins as summarized in table 4.

We see that  $\eta_3 > 1$  for all spins and  $\eta_3 > \eta_2$  except  $s = \frac{1}{2}$ , while  $\eta_3 = \eta_2 = \sqrt{2}$  for  $s = 1/2$ . Also  $\eta_3$ , like  $\eta_2$ , decreases monotonically with  $s$ . It is interesting, in the case of two and three successive measurements of spin  $s$  prepared in a pure state, the maximum violation of BI and MKI tends to a constant for arbitrary large  $s$ .

$$\eta_3(s \rightarrow \infty) = 1.153 \text{ (approx.)},$$

$$\eta_2(s \rightarrow \infty) = 1.143 \text{ (approx.)}.$$

It is thus seen that large quantum numbers do not guarantee “classical” (as defined in this chapter) behavior.

It is straightforward to check that, three successive measurements satisfy Svetlichny Inequality (SI) (equation (10)). The reason is that, for all  $s$ , the settings of the measurement directions which maximize  $MKI'$  are obtained from those which maximize  $MKI$  by interchanging primes on the corresponding unit vectors. Thus these two settings are incompatible so that we cannot get a single set of measurement directions, which maximize both  $MKI$  and  $MKI'$ . In fact, for all  $s$ , the measurement directions which maximize  $MKI$  ( $MKI'$ ) correspond to  $MKI' = 0$  ( $MKI = 0$ ).

We now consider the situation of three consecutive observations but two-fold correlations for two measurements  $\vec{S} \cdot \hat{a}_1$  and  $\vec{S} \cdot \hat{a}_3$  performed, say, at time  $t_1$  and  $t_3$ , but where an additional measurement ( $\vec{S} \cdot \hat{a}_2$ ) is performed at time  $t_2$  lying between  $t_1$  and  $t_3$  ( $t_1 < t_2 < t_3$ ). By substituting Eq (A.20) in Bell type inequality Eq(8) and simplifying, we obtain:

$$\begin{aligned}
 |BI| &= \frac{1}{2} [|\cos \theta_{32} + \cos \theta_{3'2}| |\langle \alpha_1 \alpha_2 \rangle| + |\cos \theta_{32} - \cos \theta_{3'2}| |\langle \alpha'_1 \alpha_2 \rangle|] \\
 &\leq \frac{1}{2} [|\cos \theta_{32} + \cos \theta_{3'2}| |\langle \alpha_1 \alpha_2 \rangle| + |\cos \theta_{32} - \cos \theta_{3'2}| |\langle \alpha'_1 \alpha_2 \rangle|] \\
 &\leq \cos \theta_{32} s^2 \leq s^2.
 \end{aligned} \tag{36}$$

We have used  $\max |\langle \alpha_1 \alpha_2 \rangle| = \max |\langle \alpha'_1 \alpha_2 \rangle| = s^2$ .

So, the correlation function (36) for a given measurement performed at  $t_2$  cannot violate the Bell type inequality for measurements at  $t_1$  and  $t_3$ . Therefore, any measurement performed at time  $t_2$  “disentangles” events at time  $t_1$  and  $t_3$  if  $t_1 < t_2 < t_3$  (Brukner et al., 2004).

## 5. $n$ successive measurements for Spin- $\frac{1}{2}$

### 5.1 Violation Mermin-Klyshko Inequality (MKI)

We consider now  $n$  successive measurements in direction  $\vec{S} \cdot \hat{a}_i$ , ( $i = 1, 2, 3, \dots, n$ ) on a spin  $s = \frac{1}{2}$  particle in a mixed state. For simplicity we take the eigenvalues to be  $\alpha_k = \pm 1$ , *i.e.*, the eigenvalues of  $\sigma_z$  are taken here as  $\pm 1$  instead of  $\pm(1/2)$ . We also write  $|\alpha_k\rangle$  for  $|\vec{S} \cdot \hat{a}_k, \alpha_k\rangle$ . The initial state is taken as

$$\rho_0 = p_{+1}|\alpha_0 = +1\rangle\langle\alpha_0 = +1| + p_{-1}|\alpha_0 = -1\rangle\langle\alpha_0 = -1|. \quad (37)$$

For a spin- $\frac{1}{2}$  system, we have

$$|\langle\alpha_{k-1}|\alpha_k\rangle|^2 = \frac{1}{2}(1 + \alpha_{k-1}\alpha_k \cos\theta_{k-1,k}) \quad (38)$$

where  $\cos\theta_{k-1,k} = \hat{a}_{k-1} \cdot \hat{a}_k$  for  $k = 1, 2, \dots, n$ . So, given the input state  $|\alpha_0\rangle$ , the (joint) probability that the measurement outcomes will be  $\alpha_1 \in \{+1, -1\}$  in the first measurement,  $\alpha_2 \in \{+1, -1\}$  in the second measurement,  $\dots$ ,  $\alpha_n \in \{+1, -1\}$  in the  $n$ -th measurement, will be given by

$$p(\alpha_1, \alpha_2, \dots, \alpha_n) = \frac{1}{2^n} \prod_{i=1}^n (1 + \alpha_{i-1}\alpha_i \cos\theta_{i-1,i}). \quad (39)$$

Thus we see that given the input state  $\rho_0 = \sum_{\alpha_0=\pm 1} p_{\alpha_0} |\alpha_0\rangle\langle\alpha_0|$ , the average output state after  $n$  successive measurements will be given by

$$\rho_n = \sum_{\alpha_0, \alpha_1, \dots, \alpha_n = \pm 1} p_{\alpha_0} p(\alpha_1, \alpha_2, \dots, \alpha_n) |\alpha_n\rangle\langle\alpha_n|.$$

Then, for  $n$  successive measurements on spin-1/2 system,

$$\begin{aligned} \langle\alpha_{n-1}\alpha_n\rangle_{QM} &= \sum_{\alpha_{n-1}, \alpha_n = \pm 1} \alpha_{n-1} \{\text{coeff. of } |\alpha_{n-1}\rangle\langle\alpha_{n-1}| \text{ in } \rho_{n-1}\} \alpha_n |\langle\alpha_{n-1}|\alpha_n\rangle|^2 \\ &= \sum_{\alpha_0 = \pm 1} p_{\alpha_0} \sum_{\alpha_1, \alpha_2, \dots, \alpha_n = \pm 1} \alpha_{n-1} \alpha_n p(\alpha_1, \alpha_2, \dots, \alpha_n) \\ &= \sum_{\alpha_0 = \pm 1} p_{\alpha_0} 2^{-n} \sum_{\alpha_1, \alpha_2, \dots, \alpha_n = \pm 1} \prod_{i=1}^n \alpha_{n-1} \alpha_n (1 + \alpha_{i-1} \alpha_i \cos\theta_{i-1,i}) \\ &= \cos\theta_{n-1,n} \end{aligned} \quad (40)$$

by equation (39). Further

$$\begin{aligned} \langle\alpha_n\rangle_{QM} &= \sum_{\alpha_n = \pm 1} \alpha_n \{\text{coeff. of } |\alpha_n\rangle\langle\alpha_n| \text{ in } \rho_n\} \\ &= \sum_{\alpha_0 = \pm 1} p_{\alpha_0} \sum_{\alpha_1, \alpha_2, \dots, \alpha_n = \pm 1} \alpha_n p(\alpha_1, \alpha_2, \dots, \alpha_n) \\ &= \sum_{\alpha_0 = \pm 1} p_{\alpha_0} 2^{-n} \sum_{\alpha_1, \alpha_2, \dots, \alpha_n = \pm 1} \prod_{i=1}^n \alpha_n (1 + \alpha_{i-1} \alpha_i \cos\theta_{i-1,i}) \\ &= (p_{+1} - p_{-1}) \cos\theta_1 \cos\theta_{12} \dots \cos\theta_{n-1,n} \end{aligned} \quad (41)$$

where  $\theta_1 \equiv \theta_{0,1}$ ,  $\theta_{12} \equiv \theta_{1,2}$ , etc. Now equations (40) and (41) give,

$$\langle \alpha_n \rangle_{QM} = \langle \alpha_1 \rangle_{QM} \langle \alpha_2 \alpha_3 \rangle_{QM} \cdots \langle \alpha_{n-1} \alpha_n \rangle_{QM}. \quad (42)$$

Further,

$$\begin{aligned} & \langle \alpha_{n-k} \cdots \alpha_n \rangle_{QM} = \\ & \sum_{\alpha_0} p_{\alpha_0} 2^{-n} \sum_{\alpha_1, \alpha_2, \dots, \alpha_n = \pm 1} \prod_{i=1}^n (\alpha_{n-k} \cdots \alpha_n) (1 + \alpha_{i-1} \alpha_i \cos \theta_{i-1, i}) = \\ & \begin{cases} \langle \alpha_1 \rangle_{QM} \langle \alpha_2 \alpha_3 \rangle_{QM} \cdots \langle \alpha_{n-1} \alpha_n \rangle_{QM} & k \text{ even} \\ \langle \alpha_{n-k} \alpha_{n-k+1} \rangle_{QM} \langle \alpha_{n-k+2} \alpha_{n-k+3} \rangle_{QM} \cdots \langle \alpha_{n-1} \alpha_n \rangle_{QM} & k \text{ odd} \end{cases} \end{aligned} \quad (43)$$

All of the above results are inherently quantum and are not compatible with HVT ( see the discussion in the next paragraph). The first two results ((41) and (42)) are the special cases of the last result (43) for  $k = 1$  and  $k = 0$  (with  $\alpha_0 = 1$ ). If the number of variables ( which are averaged) is odd (i.e.  $k$  is even) the average depends on the measurements prior to  $(n - k)$ , while in the other case the average does not depend on the measurements prior to  $(n - k)$ . For example, for two successive measurements ( taking  $n = 2$  and  $k = 1$ ), gives  $\langle \alpha_1 \alpha_2 \rangle = \cos \theta_{12}$ , which is independent of the initial state. On the other hand, for three successive measurements ( taking  $n = 3$  and  $k = 2$ ), we have  $\langle \alpha_1 \alpha_2 \alpha_3 \rangle = \langle \alpha_1 \rangle \langle \alpha_2 \alpha_3 \rangle$  – showing its dependence on the initial state (as  $\langle \alpha_1 \rangle = (p_{+1} - p_{-1}) \cos \theta_1$  depends upon the initial state  $\rho_0 = \sum_{\alpha_0 = \pm 1} p_{\alpha_0} |\alpha_0\rangle \langle \alpha_0|$ ). Moreover, the correlation  $\langle \alpha_1 \alpha_2 \alpha_3 \alpha_4 \rangle_{QM}$  for four successive measurements ( for example ) turns out to be dependent only on the two ‘disjoint’ correlations  $\langle \alpha_1 \alpha_2 \rangle_{QM}$  and  $\langle \alpha_3 \alpha_4 \rangle_{QM}$ . In general, we have:

$$\langle \alpha_1 \alpha_2, \dots, \alpha_{2p} \rangle = \langle \alpha_1 \alpha_2 \rangle \langle \alpha_3 \alpha_4 \rangle \dots \langle \alpha_{2p-1} \alpha_{2p} \rangle \quad (44)$$

and

$$\langle \alpha_1 \alpha_2, \dots, \alpha_{2p+1} \rangle = \langle \alpha_1 \rangle \langle \alpha_2 \alpha_3 \rangle \dots \langle \alpha_{2p} \alpha_{2p+1} \rangle. \quad (45)$$

Interestingly if  $\hat{a}_0 \perp \hat{a}_1$  so that  $\cos \theta_1 = 0$  ( and so,  $\langle \alpha_1 \rangle_{QM} = 0$ ) or, if the initial state is the random mixture  $(1/2) \sum_{\alpha_0 = \pm 1} |\alpha_0\rangle \langle \alpha_0|$  ( and so  $\langle \alpha_1 \rangle = 0$ ), then for all even  $k$ ,

$$\langle \alpha_{n-k} \cdots \alpha_n \rangle_{QM} = 0$$

and so

$$\langle \alpha_1 \alpha_2 \cdots \alpha_{n=2p+1} \rangle_{QM} = 0.$$

We shall now show that for  $n$  successive experiments ( with  $n > 1$ ), QM violates the inequality  $|\langle MKI \rangle| \leq s^3$  ( see equation (9)) up to  $\sqrt{2}$  for  $s = 1/2$  systems. We take the eigenvalues to be  $\alpha_k = \pm 1$  so  $|\langle M_k \rangle|_{HVT} \leq 1$ . We have already shown that for  $n = 2$  and  $n = 3$ , the corresponding MKI’s are violated (section 4). Now, we know that temporal two-fold correlations  $\langle \alpha_{k-1} \alpha'_k \rangle$ ,  $\langle \alpha'_{k-1} \alpha_k \rangle$ ,  $\langle \alpha_{k-1} \alpha_k \rangle$  and  $\langle \alpha'_{k-1} \alpha'_k \rangle$  are independent on the previous measurements. So by using equations (43) and (13) we find that

$$|\langle M_k \rangle| = \frac{1}{2} |\langle M_{k-2} \rangle [ \langle \alpha_{k-1} \alpha'_k \rangle + \langle \alpha'_{k-1} \alpha_k \rangle ] + \langle M'_{k-2} \rangle [ \langle \alpha_{k-1} \alpha_k \rangle - \langle \alpha'_{k-1} \alpha'_k \rangle ]|. \quad (46)$$

We now consider the spherical-polar co-ordinates  $(\theta_{k-1}, \phi_{k-1})$ ,  $(\theta'_{k-1}, \phi'_{k-1})$ ,  $(\theta_k, \phi_k)$ ,  $(\theta'_k, \phi'_k)$  of the vectors  $\hat{a}_{k-1}$ ,  $\hat{a}'_{k-1}$ ,  $\hat{a}_k$ ,  $\hat{a}'_k$  respectively, where all  $\theta \in [0, \pi]$  and all  $\phi \in [0, 2\pi]$ . Then  $|\langle M_k \rangle|$  has the form

$$\begin{aligned} |\langle M_k \rangle| &= \frac{1}{2} |\langle M_{k-2} \rangle| [\cos \theta_{k-1} \cos \theta'_k + \sin \theta_{k-1} \sin \theta'_k \cos(\phi_{k-1} - \phi'_k)] \\ &\quad + \cos \theta'_{k-1} \cos \theta_k + \sin \theta'_{k-1} \sin \theta_k \cos(\phi'_{k-1} - \phi_k)] \\ &\quad + \langle M'_{k-2} \rangle [\cos \theta_{k-1} \cos \theta_k + \sin \theta_{k-1} \sin \theta_k \cos(\phi_{k-1} - \phi_k) \\ &\quad - \cos \theta'_{k-1} \cos \theta'_k - \sin \theta'_{k-1} \sin \theta'_k \cos(\phi'_{k-1} - \phi'_k)]. \end{aligned} \quad (47)$$

We know from

$$\frac{\partial |\langle M_k \rangle|}{\partial \phi_{k-1}} = \frac{\partial |\langle M_k \rangle|}{\partial \phi_k} = \frac{\partial |\langle M_k \rangle|}{\partial \phi'_{k-1}} = \frac{\partial |\langle M_k \rangle|}{\partial \phi'_k} = 0 \Rightarrow \phi_{k-1} = \phi'_{k-1} = \phi_k = \phi'_k.$$

The maximum value of  $|\langle M_k \rangle|$  will occur when all the vectors  $\hat{a}_{k-1}$ ,  $\hat{a}'_{k-1}$ ,  $\hat{a}_k$ ,  $\hat{a}'_k$  lie on the same plane. We obtain:

$$\begin{aligned} |\langle M_k \rangle| &\leq \frac{1}{2} |\langle M_{k-2} \rangle| [\cos(\theta_{k-1} - \theta'_k) + \cos(\theta'_{k-1} - \theta_k)] \\ &\quad + \langle M'_{k-2} \rangle [\cos(\theta_{k-1} - \theta_k) - (\cos \theta'_{k-1} - \theta'_k)] \\ &\leq \frac{1}{2} |\langle M_{k-2} \rangle| [\cos(\theta_{k-1} - \theta'_k) + \cos(\theta'_{k-1} - \theta_k)] \\ &\quad + \frac{1}{2} |\langle M'_{k-2} \rangle| [\cos(\theta_{k-1} - \theta_k) - (\cos \theta'_{k-1} - \theta'_k)]. \end{aligned} \quad (48)$$

By substituting  $x = \theta_{k-1} - \theta'_k$ ,  $y = \theta'_{k-1} - \theta_k$ ,  $z = \theta_{k-1} - \theta_k$  and  $\theta'_{k-1} - \theta'_k = x + y - z$  in above-equation and by using

$$\frac{\partial |\langle M_k \rangle|}{\partial x} = \frac{\partial |\langle M_k \rangle|}{\partial y} = \frac{\partial |\langle M_k \rangle|}{\partial z} = 0,$$

we get  $x = y = -z = \pi/4$ . Finally, by using the fact  $|\langle M_k \rangle| + |\langle M'_k \rangle| \leq 2$ , we obtain:

$$|\langle M_k \rangle| \leq \frac{\sqrt{2}}{2} \{|\langle M_{k-2} \rangle| + |\langle M'_{k-2} \rangle|\} \leq \sqrt{2}. \quad (49)$$

One can get this result by induction hypothesis. Thus, we conclude that QM violates the MKI ( $|\langle M_n \rangle| \leq 1$ ) for  $n$  successive measurements upto  $\sqrt{2}$ , i.e.,

$$\eta_n = \sqrt{2}. \quad (50)$$

## 5.2 Violation Scarani-Gisin inequality (SCI)

Although in contrast to correlations in space there are no genuine multi-mode correlations in time, we will see that temporal correlations can be stronger than spatial ones in a certain sense. We denote by  $\max[B_{QM}^{space}(i, j)]$  the maximal value of the Bell expression for qubits  $i$  and  $j$  (Bell inequality is obtained by Eq(8)). Scarani and Gisin (Scarani & Gisin, 2001) found an interesting bound that holds for arbitrary state of three qubits:

$$\max[B_{QM}^{space}(1, 2)] + \max[B_{QM}^{space}(2, 3)] \leq 2. \quad (51)$$

Physically, this means that no two pairs of qubits of a three-qubit system can violate the CHSH inequalities simultaneously. This is because if two systems are highly entangled, they can not be entangled highly to another systems. Let us denote by  $\max[B_{QM}^{time}(i, j)]$  the maximal value of the Bell expression for two consecutive observations of a single qubit at times  $i$  and  $j$ . Since quantum correlations between two successive measurements do not depend on the initial state ( see Eq (40)), one can obtain:

$$\begin{aligned} & \left[ B_{QM}^{time}(k-1, k) + B_{QM}^{time}(k, k+1) \right] = \\ & \frac{1}{2} [\cos \theta_{k-1, k} + \cos \theta_{k-1, k'} + \cos \theta_{k-1', k} - \cos \theta_{k-1', k'}] + \\ & \frac{1}{2} [\cos \theta_{k, k+1} + \cos \theta_{k, k+1'} + \cos \theta_{k', k+1} - \cos \theta_{k', k+1'}]. \end{aligned} \quad (52)$$

By selecting,

$$\theta_{k-1, k} = \theta_{k-1, k'} = \theta_{k-1', k} = \theta_{k, k+1} = \theta_{k, k+1'} = \theta_{k', k+1} = \frac{\pi}{4}$$

and

$$\theta_{k-1', k'} = \theta_{k', k+1'} = \frac{3\pi}{4},$$

we obtain:

$$\max \left[ B_{QM}^{time}(k-1, k) + \max[B_{QM}^{time}(k, k+1) \right] = \sqrt{2} + \sqrt{2} = 2\sqrt{2} > 2. \quad (53)$$

Thus, although there are no genuine three-fold temporal correlations, a specific combination of two-fold correlations can have values that are not achievable with correlations in space for any three-qubit system. In fact, one would need two pairs of maximally entangled two-qubit states to achieve the bound in (53). Also note that the local realistic bound is 2, which is equal to the bound in (51). Similar conclusion can be obtained for the sum of  $n$  successive measurements.

$$\max \left[ B_{QM}^{time}(1, 2) \right] + \max \left[ B_{QM}^{time}(2, 3) \right] + \dots + \max \left[ B_{QM}^{time}(n-1, n) \right] = n\sqrt{2} > n. \quad (54)$$

### 5.3 Violation chained Bell inequalities (CHI)

Generalized CHSH inequalities may be obtained by providing more than two alternative experiments to each process. We consider two successive measurements on a spin- $\frac{1}{2}$  particle in a mixed state, such that the first experiment can measure spin component along one of the directions  $\hat{a}_1, \hat{a}_3, \dots, \hat{a}_{2n-1}$  and the second experiment along one of the directions  $\hat{b}_2, \hat{b}_4, \dots, \hat{b}_{2n}$ . The results of these measurements are called  $\alpha_r$  ( $r = 1, 3, \dots, 2n-1$ ) and  $\beta_s$  ( $s = 2, 4, \dots, 2n$ ), respectively, and their values are  $\pm 1$  (in unit if  $\hbar/2$ ). We have a generalized CHSH inequality (Braunstein & Caves, 1990),(Peres, 1993):

$$CBI = \frac{1}{2} |\langle \alpha_1 \beta_2 \rangle + \langle \beta_2 \alpha_3 \rangle + \langle \alpha_3 \beta_4 \rangle + \dots + \langle \alpha_{2n-1} \beta_{2n} \rangle - \langle \beta_{2n} \alpha_1 \rangle| \leq n-1 \quad (55)$$

This upper bound is violated by quantum correlations in two successive measurements, increasingly with larger  $n$ . In order to obtain the maximum value above-inequality, we consider the spherical-polar co-ordinates  $(\theta_k, \phi_k), (k = 1, 3, \dots, 2n-1)$  of the vectors

$\hat{a}_1, \hat{a}_3, \dots, \hat{a}_{2n-1}$  and ( $k = 2, 4, \dots, 2n$ ) for vectors  $\hat{b}_2, \hat{b}_4, \dots, \hat{b}_{2n}$ . The maximum value  $|CBI|$  will occur when all the vectors lie on the same plane. this is because of:

$$\frac{\partial(CBI)}{\partial\phi_k} = 0 \Rightarrow \phi_1 = \phi_2 = \dots = \phi_n.$$

After partial differential over all  $\theta_k$ , we get:

$$\frac{\partial(CBI)}{\partial\theta_k} = 0 \Rightarrow \theta_{12} = \theta_{23} = \dots = \theta_{2n-1,2n} = \theta.$$

Therefore, we obtain:

$$CBI = (2n - 1) \cos \theta - \cos(2n - 1)\theta. \quad (56)$$

So,

$$\frac{\partial(CBI)}{\partial\theta} = 0 \Rightarrow \theta = \pi/2n. \quad (57)$$

By substituting , we obtain:

$$CBI = (2n - 1) \cos \frac{\beta}{2n} - \cos \frac{(2n - 1)\beta}{2n} = 2n \cos \frac{\beta}{2n} \quad (58)$$

We know  $\cos(\frac{\pi}{2n})$  tends to  $(1 - \frac{\pi^2}{8n^2})$  for  $n \rightarrow \infty$ . Therefore the maximum CBI can be made arbitrarily close to  $2n$ .

#### 5.4 Violation Bell inequalities involving Tri and Bi-measurements correlations

It would be interesting to consider Bell inequalities involving both two and three successive measurement correlations. The simplest way of obtaining such an inequality would be by adding genuinely bipartite correlations to the tripartite correlations considered in Mermin's inequality. For instance, a straightforward calculation would allow us to prove that any local realistic theory must satisfy the following inequality (Cabello, 2002b):

$$\begin{aligned} -5 \leq & \langle \alpha_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha_1 \alpha'_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha'_2 \alpha_3 \rangle - \\ & \langle \alpha_1 \alpha'_2 \rangle - \langle \alpha_1 \alpha'_3 \rangle - \langle \alpha_2 \alpha_3 \rangle \leq 3 \end{aligned} \quad (59)$$

A numerical calculation shows that both the GHZ and W states give a same maximal violation of the inequality (59). However, if we assign a higher weight to the bipartite correlations appearing in the inequality, then we can reach a Bell inequality such as

$$\begin{aligned} -8 \leq & \langle \alpha_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha_1 \alpha'_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha'_2 \alpha_3 \rangle - \\ & 2\langle \alpha_1 \alpha'_2 \rangle - 2\langle \alpha_1 \alpha'_3 \rangle - 2\langle \alpha_2 \alpha_3 \rangle \leq 4, \end{aligned} \quad (60)$$

which is violated by the W state but not by GHZ state (Cabello, 2002b). It is not difficult to show that three successive measurements correlations for spin 1/2 break the hybrid Bell inequalities.

$$\begin{aligned} -5.34 \leq & \langle \alpha_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha_1 \alpha'_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha'_2 \alpha_3 \rangle - \\ & \langle \alpha_1 \alpha'_2 \rangle - \langle \alpha_1 \alpha'_3 \rangle - \langle \alpha_2 \alpha_3 \rangle \leq 3.8 \end{aligned} \quad (61)$$

and

$$\begin{aligned} -8.2 \leq & \langle \alpha_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha_1 \alpha'_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha_2 \alpha'_3 \rangle - \langle \alpha'_1 \alpha'_2 \alpha_3 \rangle - \\ & 2\langle \alpha_1 \alpha'_2 \rangle - 2\langle \alpha_1 \alpha'_3 \rangle - 2\langle \alpha_2 \alpha_3 \rangle \leq 4.8. \end{aligned} \quad (62)$$

So two successive measurements correlations are relevant to those of three successive measurements. This behavior is analogous to three particle W state (Cabello, 2002b).

## 6. Hardy's argument for $n$ successive measurements for all spin- $s$ measurements

Hardy's nonlocality argument is considered weaker than Bell inequalities in the bipartite case, as every maximally entangled state of two spin- $\frac{1}{2}$  particles violates Bell's inequality maximally but none of them satisfies Hardy-type nonlocality conditions. The scenario in successive spin measurements is quite different, however. We showed in previous sections that all  $n$  successive spin- $s$  measurements break Bell-type inequalities, in contrast to the bipartite case, where only the entangled states break it. In this section, we prove that all  $n$  successive spin- $s$  measurements satisfy Hardy-type argument conditions. Consider four yes/no-type events  $A, A', B$  and  $B'$ , where  $A$  and  $A'$  may happen at time  $t_1$ , and  $B$  and  $B'$  may happen at another time,  $t_2$  ( $t_2 > t_1$ ). The joint probability that, at the first time ( $t_1$ ),  $A$  and, at the second time ( $t_2$ ),  $B$  are "yes" is 0. The joint probability that, at the first time ( $t_1$ ),  $A$  is "no" and, at the second time ( $t_2$ ),  $B'$  is "yes" is 0. The joint probability that, at the first time ( $t_1$ ),  $A'$  is "yes" and, at the second time ( $t_2$ ),  $B$  is "no", is 0. The joint probability that both  $A'$  and  $B'$  are "yes" is nonzero. We can write this as follows:

$$\begin{aligned}
 p(A = +1, B = +1) &= 0, \\
 p(A = -1, B' = +1) &= 0, \\
 p(A' = +1, B = -1) &= 0, \\
 p(A' = +1, B' = +1) &= p \neq 0.
 \end{aligned} \tag{63}$$

We show that these four statements are not compatible with time-local realism. The nonzero probability appearing in the argument is the measure of violation of time-local realism. It is interesting that two successive  $s$ -spin measurements violate time-local realism. We deal with the case where the input state is a pure state whose eigenstates coincide with those of  $s \cdot \hat{a}_0$  for some  $\hat{a}_0$  whose eigenvalues we denote  $\alpha_0 = j$ . Hardy's argument for a system of  $n$  successive spin- $s$  measurements, in its minimal form (Parasuram & Ghosh, n.d.), is given by following conditions:

$$\begin{aligned}
 p(s \cdot \hat{a}_1 = j, s \cdot \hat{a}_2 = j, \dots, s \cdot \hat{a}_n = j) &= 0, \\
 p(s \cdot \hat{a}_1 = j - 1, s \cdot \hat{a}'_2 = j, \dots, s \cdot \hat{a}'_n = j) &= 0, \\
 p(s \cdot \hat{a}_1 = j - 2, s \cdot \hat{a}'_2 = j, \dots, s \cdot \hat{a}'_n = j) &= 0, \\
 &\vdots \\
 &\vdots \\
 p(s \cdot \hat{a}_1 = -j, s \cdot \hat{a}'_2 = j, \dots, s \cdot \hat{a}'_n = j) &= 0, \\
 &\vdots \\
 &\vdots \\
 &\vdots
 \end{aligned} \tag{65}$$

$$\begin{aligned}
 p(s \cdot \hat{a}'_1 = j, \dots, s \cdot \hat{a}_l = j - 1, \dots, s \cdot \hat{a}'_n = j) &= 0, \\
 p(s \cdot \hat{a}'_1 = j, \dots, s \cdot \hat{a}_l = j - 2, \dots, s \cdot \hat{a}'_n = j) &= 0,
 \end{aligned} \tag{66}$$

$$\begin{aligned} & \cdot \\ & \cdot \\ & p (s.\hat{a}'_1 = j, \dots, s.\hat{a}_l = -j, \dots, s.\hat{a}'_n = j) = 0, \end{aligned} \quad (67)$$

$$\begin{aligned} & \cdot \\ & \cdot \\ & \cdot \\ & p (s.\hat{a}'_1 = j, s.\hat{a}'_2 = j, \dots, s.\hat{a}_n = j - 1) = 0, \\ & p (s.\hat{a}'_1 = j, s.\hat{a}'_2 = j, \dots, s.\hat{a}_n = j - 2) = 0, \\ & \cdot \\ & \cdot \\ & p (s.\hat{a}'_1 = j, s.\hat{a}'_2 = j, \dots, s.\hat{a}_n = -j) = 0, \\ & p (s.\hat{a}'_1 = j, s.\hat{a}'_2 = j, \dots, s.\hat{a}'_n = j) = p. \end{aligned} \quad (68)$$

First, we prove here that all time-local SHVTs predict  $p = 0$ . Suppose that a time-local SHVT reproducing, in accordance with Eq.(4), the quantum predictions exist. Accordingly, if we consider, for example, Eq.(64), we must have

$$\begin{aligned} & p(s.\hat{a}'_1 = j, \dots, s.\hat{a}_l = j - 1, \dots, s.\hat{a}'_n = j) \\ & = \int_{\Lambda} d\lambda \rho(\lambda) p_{\lambda}(s.\hat{a}'_1 = j, \dots, s.\hat{a}_l = j - 1, \dots, s.\hat{a}'_n = j) \\ & = \int_{\Lambda} d\lambda \rho(\lambda) p_{\lambda}(s.\hat{a}'_1 = j) \dots p_{\lambda}(s.\hat{a}_l = j - 1) \dots p_{\lambda}(s.\hat{a}'_n = j) \\ & = 0, \end{aligned} \quad (69)$$

where the second equality is implied by the time-locality condition of Eq.(4). The last equality in Eq.(69) can be fulfilled if and only if the product  $p_{\lambda}(s.\hat{a}'_1 = j) \dots p_{\lambda}(s.\hat{a}'_n = j)$  vanishes every time within  $\Lambda$ . An equivalent result holds for Eqs.(64-69), leading to:

$$\begin{aligned} & p_{\lambda} (s.\hat{a}_1 = j) p_{\lambda} (s.\hat{a}_2 = j) \dots p_{\lambda} (s.\hat{a}_n = j) = 0, \end{aligned} \quad (70)$$

$$\begin{aligned} & p_{\lambda} (s.\hat{a}_1 = j - 1) p_{\lambda} (s.\hat{a}'_2 = j) \dots p_{\lambda} (s.\hat{a}'_n = j) = 0, \\ & p_{\lambda} (s.\hat{a}_1 = j - 2) p_{\lambda} (s.\hat{a}'_2 = j) \dots p_{\lambda} (s.\hat{a}'_n = j) = 0, \end{aligned}$$

$$\begin{aligned} & \cdot \\ & \cdot \\ & p_{\lambda} (s.\hat{a}_1 = -j) p_{\lambda} (s.\hat{a}'_2 = j) \dots p_{\lambda} (s.\hat{a}'_n = j) = 0, \end{aligned} \quad (71)$$

$$\begin{aligned} & \cdot \\ & \cdot \\ & \cdot \\ & p_{\lambda} (s.\hat{a}'_1 = j) \dots p_{\lambda} (s.\hat{a}_l = j - 1) \dots p_{\lambda} (s.\hat{a}'_n = j) = 0, \\ & p_{\lambda} (s.\hat{a}'_1 = j) \dots p_{\lambda} (s.\hat{a}_l = j - 2) \dots p_{\lambda} (s.\hat{a}'_n = j) = 0, \\ & \cdot \\ & \cdot \\ & p_{\lambda} (s.\hat{a}'_1 = j) \dots p_{\lambda} (s.\hat{a}_l = -j) \dots p_{\lambda} (s.\hat{a}'_n = j) = 0, \end{aligned} \quad (72)$$

$$\begin{aligned}
& \cdot \\
& \cdot \\
& \cdot \\
& p_\lambda (s.\hat{a}'_1 = j)p_\lambda(s.\hat{a}'_2 = j) \dots p_\lambda(s.\hat{a}_n = j - 1) = 0, \\
& p_\lambda (s.\hat{a}'_1 = j)p_\lambda(s.\hat{a}'_2 = j) \dots p_\lambda(s.\hat{a}_n = j - 2) = 0, \\
& \cdot \\
& \cdot \\
& p_\lambda (s.\hat{a}'_1 = j)p_\lambda(s.\hat{a}'_2 = j) \dots p_\lambda(s.\hat{a}_n = -j) = 0, \tag{73} \\
& p_\lambda (s.\hat{a}'_1 = j)p_\lambda(s.\hat{a}'_2 = j) \dots p_\lambda(s.\hat{a}'_n = j) = p \neq 0, \tag{74}
\end{aligned}$$

where the first  $2jn + 1$  equations are supposed to hold almost every time within  $\Lambda$ , while the last equation has to be satisfied in a subset of  $\Lambda$  whose measure according to the distribution  $\rho(\lambda)$  is nonzero. To prove the more general result that no conceivable time-local SHVT can simultaneously satisfy Eqs.(70)-(74), a manipulation of those equations is required. To this end, let us sum all equations in each set. We obtain

$$\begin{aligned}
(1 - p_\lambda(s.\hat{a}_1 = j)) [p_\lambda(s.\hat{a}'_2 = j) \dots p_\lambda(s.\hat{a}'_n = j)] &= 0, \\
\cdot \\
\cdot \\
(1 - p_\lambda(s.\hat{a}_l = j)) [p_\lambda(s.\hat{a}'_1 = j) \dots p_\lambda(s.\hat{a}'_n = j)] &= 0, \\
\cdot \\
\cdot \\
(1 - p_\lambda(s.\hat{a}_n = j)) [p_\lambda(s.\hat{a}'_1 = j) \dots p_\lambda(s.\hat{a}'_{n-1} = j)] &= 0. \tag{75}
\end{aligned}$$

Now let us partition the set of hidden variables  $\Lambda$  and define the following subsets  $A_1, A_2, \dots, A_n$ , and  $B$  as:

$$\begin{aligned}
A_1 &= \{\lambda \in \Lambda | p_\lambda(s.\hat{a}_1 = j) = 0\}, \\
& \cdot \\
& \cdot \\
A_l &= \{\lambda \in \Lambda | p_\lambda(s.\hat{a}_l = j) = 0\}, \tag{76} \\
& \cdot \\
& \cdot
\end{aligned}$$

$$\begin{aligned}
A_n &= \{\lambda \in \Lambda | p_\lambda(s.\hat{a}_n = j) = 0\}, \\
B &= \Lambda - \{A_1 \cup A_2 \cup \dots \cup A_n\}. \tag{77}
\end{aligned}$$

We have that, for all  $\lambda$  belonging to  $B$ ,  $p_\lambda(s.\hat{a}_1 = j)p_\lambda(s.\hat{a}_2 = j) \dots p_\lambda(s.\hat{a}_n = j) \neq 0$ . If set  $B$  had a nonzero measure according to the distribution  $\rho$ , that is, if  $\int_B d\lambda \rho(\lambda) \neq 0$ , there would be violation of Eq.(70) and, consequently, of Eq.(64). Therefore, to fulfill Eq.(70), the set  $A_1 \cup A_2 \cup \dots \cup A_n$  must coincide with  $\Lambda$  apart from a set of zero measure, and we are left only with hidden variables belonging to either  $A_1$  or  $A_2$  or ... or  $A_n$ . If  $\lambda$

belongs to  $A_l$ , then, by definition,  $p_\lambda(s.\hat{a}_l = j) = 0$ , so that Eq.(75) can be satisfied only if  $p_\lambda(s.\hat{a}_1 = j) \dots p_\lambda(s.\hat{a}_n = j) = 0$ . Hence, for any  $\lambda \in \{A_1 \cup A_2 \cup \dots \cup A_n\}$ , we obtain a result leading to a contradiction of Eq.(74), which requires that there is a set of nonzero  $\rho$  measure within  $\Lambda$  where both probabilities do not vanish. To summarize, we have shown that it is not possible to exhibit any time-local hidden-variable model, satisfying Hardy's logic for  $n$  successive measurements.

Now, we show that in quantum theory for the  $n$  successive spin measurement, sometimes  $p > 0$ . So, we consider  $n$  successive measurements in directions  $s.\hat{a}_i$  ( $i = 1, 2, \dots, n$ ) on spin- $s$  particles. For a spin- $s$  system, we have ( see Appendix-B):

$$|\langle \alpha_{k-1} | \alpha_k \rangle| = |\langle s.\hat{a}_{k-1} | s.\hat{a}_k \rangle| = d_{\alpha_{k-1}, \alpha_k}^{(s)}(\beta_k - \beta_{k-1}), \quad (78)$$

where  $\beta_k$  is the angle between the  $\hat{a}_k$  and the  $+z$  axes. So, given the input state  $|\alpha_0\rangle$ , the ( joint) probability that the measurement outcomes will be  $|\alpha_1\rangle \in \{+j, \dots, -j\}$  in the first measurement,  $|\alpha_2\rangle \in \{+j, \dots, -j\}$  in the second measurement, ...,  $|\alpha_n\rangle \in \{+j, \dots, -j\}$  in the  $n$ -th measurement, is given by

$$\begin{aligned} p(\alpha_1, \alpha_2, \dots, \alpha_n) &= \prod_{k=1}^n |\langle \alpha_{k-1} | \alpha_k \rangle| \\ &= \prod_{k=1}^n d_{\alpha_{k-1}, \alpha_k}^2(\beta_k - \beta_{k-1}). \end{aligned} \quad (79)$$

We deal with the case where the input state is a pure state whose eigenstates coincide with those of  $\vec{S}.\hat{a}_0$  for some  $\hat{a}_0$  whose eigenvalues we denote  $\alpha_0 = j$ . Now, by substituting Eq.(79) in the minimal form of Hardy's argument [ Eqs.(64)-(68)], we have

$$d_{jj}^2(\beta_1) d_{jj}^2(\beta_2 - \beta_1) \dots d_{jj}^2(\beta_n - \beta_{n-1}) = 0, \quad (80)$$

$$d_{j,j-1}^2(\beta_1) d_{j,j-1}^2(\beta'_2 - \beta_1) \dots d_{jj}^2(\beta'_n - \beta'_{n-1}) = 0,$$

$$d_{j,j-2}^2(\beta_1) d_{j,j-2}^2(\beta'_2 - \beta_1) \dots d_{jj}^2(\beta'_n - \beta'_{n-1}) = 0,$$

⋮

$$d_{j,-j}^2(\beta_1) d_{j,-j}^2(\beta'_2 - \beta_1) \dots d_{jj}^2(\beta'_n - \beta'_{n-1}) = 0,$$

⋮

$$d_{jj}^2(\beta'_1) \dots d_{j,j-1}^2(\beta_l - \beta'_{l-1}) d_{j,j-1}^2(\beta'_{l+1} - \beta_l) \dots$$

$$d_{jj}^2(\beta'_n - \beta'_{n-1}) = 0,$$

$$d_{jj}^2(\beta'_1) \dots d_{j,j-2}^2(\beta_l - \beta'_{l-1}) d_{j,j-2}^2(\beta'_{l+1} - \beta_l) \dots$$

$$d_{jj}^2(\beta'_n - \beta'_{n-1}) = 0,$$

⋮

⋮

$$d_{jj}^2(\beta'_1) \dots d_{j,-j}^2(\beta_l - \beta'_{l-1}) d_{j,-j}^2(\beta'_{l+1} - \beta_l) \dots$$

$$d_{jj}^2(\beta'_n - \beta'_{n-1}) = 0, \quad (81)$$

$$\begin{aligned}
& \cdot \\
& \cdot \\
& d_{jj}^2(\beta'_1) d_{jj}^2(\beta'_2 - \beta'_1) \dots d_{j,j-1}^2(\beta_n - \beta'_{n-1}) = 0, \\
& d_{jj}^2(\beta'_1) d_{jj}^2(\beta'_2 - \beta'_1) \dots d_{j,j-2}^2(\beta_n - \beta'_{n-1}) = 0, \\
& \cdot \\
& \cdot \\
& d_{jj}^2(\beta'_1) d_{jj}^2(\beta'_2 - \beta'_1) \dots d_{j,-j}^2(\beta_n - \beta'_{n-1}) = 0, \tag{82} \\
& d_{jj}^2(\beta'_1) d_{jj}^2(\beta'_2 - \beta'_1) \dots d_{j,j}^2(\beta'_n - \beta'_{n-1}) = p. \tag{83}
\end{aligned}$$

From Eq (80), at least one of the factors must be 0. So

$$\begin{aligned}
d_{jj}^2(\beta_1) = 0 &\implies \beta_1 = \pi \\
&\text{or} \\
d_{jj}^2(\beta_2 - \beta_1) &\implies |\beta_2 - \beta_1| = \pi \\
&\text{or} \\
&\cdot \\
&\cdot \\
&\cdot \\
d_{jj}^2(\beta_n - \beta_{n-1}) &\implies |\beta_n - \beta_{n-1}| = \pi
\end{aligned}$$

To satisfy all equations (81)-(82), we have the following conditions:

$$\begin{aligned}
& (\beta_1 = 0) \text{ or } (\beta'_2 = \beta_1) \\
& \text{and} \\
& (\beta_2 = \beta'_1) \text{ or } (\beta_2 = \beta'_3) \\
& \cdot \\
& \cdot \\
& \cdot \\
& (\beta_l = \beta'_{l-1}) \text{ or } (\beta_l = \beta'_{l+1}) \tag{84} \\
& \cdot \\
& \cdot \\
& \cdot \\
& \text{and} \\
& (\beta_n = \beta'_{n-1}).
\end{aligned}$$

Now, we can calculate the maximum value  $p$  by using these conditions. For example, if we select  $\beta_1 = \pi$ , so we must have  $\beta_2 = \beta_1 = \pi$ . In this case,

$$p = d_{jj}^2(\beta'_1) d_{jj}^2(\pi - \beta'_1) d_{jj}^2(\beta'_3 - \pi) \dots d_{j,j}^2(\beta'_n - \beta'_{n-1}). \tag{85}$$

By substituting  $d_{jj}^j(\beta) = \cos^{2j}(\beta/2)$ , we have

$$p = \cos^{4j}\left(\frac{\beta'_1}{2}\right) \cos^{4j}\left(\frac{\pi - \beta'_1}{2}\right) \cos^{4j}\left(\frac{\beta'_3 - \pi}{2}\right) \dots \cos^{4j}\left(\frac{\beta'_n - \beta'_{n-1}}{2}\right). \quad (86)$$

By selecting  $\beta'_n = \beta'_{n-1} = \dots = \beta'_3 = \pi$  and  $\beta'_1 = \pi$ ,

$$p \leq \left(\frac{1}{2}\right)^{4j}. \quad (87)$$

We can obtain this result in the general case. We choose  $|\beta_l - \beta_{l-1}| = \pi$ , where  $2 \leq l \leq n$ . Without loss of generality, we select  $\beta_l = \pi$  and  $\beta_{l-1} = 0$ . From the results obtained with Eq.(84),  $\beta'_{l+1} = \beta_l = \pi$  or  $\beta'_{l-1} = \beta_l = \pi$ . Exactly for (l-1)th in Eq.(84), we have  $\beta'_l = \beta_{l-1} = 0$  or  $\beta'_{l-2} = \beta_{l-1} = 0$ . So we have four cases: (i)  $\beta'_{l-1} = \pi$  and  $\beta'_{l-2} = 0$  (ii)  $\beta'_{l-1} = \pi$  and  $\beta'_l = 0$  (iii)  $\beta'_{l+1} = \pi$  and  $\beta'_l = 0$  and (iv)  $\beta'_{l+1} = \pi$  and  $\beta'_{l-2} = 0$ . It is easy see that for the first three cases, the maximum value of  $p$  is 0, but in the fourth case, by selecting  $\beta'_1 = \beta'_2 = \dots = \beta'_{l-1} = 0$  and  $\beta'_{l+2} = \beta'_{l+3} = \dots = \beta'_n = \pi$ , we get

$$p = \cos^{4j}\left(\frac{\beta'_l}{2}\right) \sin^{4j}\left(\frac{\beta'_l}{2}\right) \leq \left(\frac{1}{2}\right)^{4j}. \quad (88)$$

We see that  $p > 0$  for all spins, and also, the maximum probability of success of Hardy's non-time locality is independent of the number of successive measurements and decreases with  $s$ .

## 7. Summary and comments

*Entanglement in space* displays one of the most interesting features of quantum mechanics, often called quantum non locality. Locality in space and realism impose constraints -Bell's inequalities- on certain combinations of correlations for measurements of spatially separated systems, which are violated by quantum mechanics. Non locality is one of the strangest properties of quantum mechanics, and understanding this notion remains an important problem.

*Entanglement in time* is not introduced in quantum mechanics because of different roles time and space play in quantum theory. The meaning of locality in time is that the results of measurement at time  $t_2$  are independent of any measurement performed at some earlier time  $t_1$  or later time  $t_3$ . The temporal Bell's inequalities are derived from the realistic hidden variable theory.

In this chapter we have considered a hidden variable theory of successive measurements on a single spin- $s$  system. In all the previous scenarios comparing HVT and QM the principal hypothesis being tested was that, in a given state (having spatial correlation), HVT implies the existence of a joint probability distribution for all observables even if some of them are not compatible. QM is shown to contradict the consequence of this requirement as it does not assign joint probabilities to the values of incompatible observables. The particular implication that is tested is whether the marginal of the observable  $A$  in the joint distribution of the compatible observables  $A$  and  $B$  is the same as the marginal for  $A$  in joint distribution

for the observables  $A$  and  $C$  even if  $B$  and  $C$  are not compatible. In other words, HVT implies noncontextuality for which QM can be tested. The celebrated theorem of Bell and Kochen-Specker showed that QM is contextual (Bell, 1966),(Kochen & Specker, 1967). In our scenario, the set of measured observables have a well defined joint probability distribution as each of them acts on a different state. Note that the Bell-type inequalities we have derived follow from equation (4) which says that, for a given value of stochastic hidden variable  $\lambda$ , the joint probability for the outcomes of successive measurements must be statistically independent. In other words the hidden variable  $\lambda$  completely decides the probabilities of individual measurement outcomes independent of other measurements. We show that QM is not consistent with this requirement of HVT. A Bell-type inequality (for single particle), testing contextuality of QM was proposed by Basu et al. (Basu et al., 2001) and was shown that it could be empirically tested. However, the approach given in the present chapter furnishes a test for realistic nature of QM independent of contextuality. We have compared QM with HVT for different values of spin and for different number of successive measurements. The dependence of the deviation of QM from HVT on the spin value and on the number of successive measurements opens up new possibilities for comparison of these models, and may lead to a sharper understanding of QM.

In the following, I bring some of the key surprising results obtained in this chapter.

1- We obtained temporal Mermin-Klyshko inequality (MKI) and Svetlichny inequality (SI) for  $n$  successive measurements by using realism and non locality in time. We showed quantum correlations violate temporal MKI and satisfy temporal SI.

2- It was interesting that, for a spin- $s$  particle, maximum deviation of quantum mechanics from realism was obtained for all convex combinations of  $\alpha_0 = \pm 1$  states (the case when input state is a mixed state whose eigenstates coincide with those of  $\vec{S} \cdot \hat{a}_0$  for some  $\hat{a}_0$  whose eigenvalues we denote by  $\alpha_0 \in \{-s, \dots, s\}$ ). This is surprising as one would expect pure states to be more 'quantum' than the mixed ones thus breaking Bell inequalities by larger amount.

3- All spin 1/2 states maximally break Mermin-Klyshko inequalities for  $n$  successive measurements ( $\eta_n = \sqrt{2}$ ) as against only the entangled states break it in multipartite case. Interestingly that for  $s = \frac{1}{2}$  the random mixture ( maximum noisy state) also breaks BI. This indicates that the notion of "classicality", compatible with the usual local HVT, is different in nature from the notion of classicality that would arise from the non-violation of BI here.

4- We saw that for all spins, BI and MKI is violated in two and three successive measurements ( $\eta_2 > 1, \eta_3 > 1$ ) and the value of violation MKI in three successive measurements is a little more than the value of violation BI in two successive measurements  $\eta_3 > \eta_2$  except  $s = \frac{1}{2}$ , while  $\eta_3 = \eta_2 = \sqrt{2}$  for spin  $s = \frac{1}{2}$ . Also  $\eta_3$  and  $\eta_2$  decrease monotonically with increase in the value of  $s$ . It is interesting, in the case of two and three successive measurements of spin  $s$  prepared in a pure state, that the maximum violation of BI and MKI falls off as the spin of the particle increases, but tends to a constant for arbitrary large  $s$ ,  $\eta_2(s \rightarrow \infty) = 1.143$  and  $\eta_3(s \rightarrow \infty) = 1.153$ . It is thus seen that large quantum numbers do not guarantee classical behavior.

5- We showed that for  $s = \frac{1}{2}$ , the correlation between the outputs of measurements from last  $k$  out of  $n$  successive measurements ( $k < n$ ) depend on the measurement prior to  $(n - k)$ , when  $k$  is even, while for odd  $k$ , these correlations are independent of the outputs of measurements prior to  $n - k$ .

6- Interestingly if the initial state is the random mixture or first Stern-Gerlach measurements for the qubit component along the directions  $a_1$  perpendicular to initial state  $\hat{a}_0 \perp \hat{a}_1$  so that,  $\langle \alpha_1 \rangle_{QM} = 0$ , then always quantum averages for all odd number of successive measurements

are zero.

7- We proved that the correlation function between first and third measurement ( $t_1$  and  $t_3$ ) on spin- $s$  particle for a given measurement performed at  $t_2$  can not violate the temporal Bell inequality. Therefore, any measurement performed at time  $t_2$  disentangles events at time  $t_1$  and  $t_3$  if  $t_1 < t_2 < t_3$ .

8- Three successive measurements on spin- $s$  particles do not break Svetlinchi Inequality. But it is proved that three successive measurements on qubit violate Scarani-Gisin inequality. Thus, although there are no genuine three-fold temporal correlations, a specific combination of two-fold correlations can have values that are not achievable with correlations in space for any three-qubit system.

9- Also we showed that three successive measurements violate two types of Bell inequalities involving two and three successive measurements. So two successive measurement correlations are relevant to those of three successive measurements. This behavior is analogous to three particle W-state.

10- Quantum correlations between two successive measurements on a qubit violates chained Bell inequality which is obtained by providing more than two alternative experiments in every step.

11- Also, we have studied Hardy's argument for the correlations between the outputs of  $n$  successive measurements for all  $s$ -spin measurements. We have shown that the maximum probability of success of Hardy's argument for  $n$  successive measurements is  $(\frac{1}{2})^{4s}$ , which is independent of the number of successive measurements of spin ( $n$ ) and decreases with increase of  $s$ . This can be compared with the correlations corresponding to measurement of spin observables in a spacelike separated two-particles scenario where only the non-maximally entangled states of any spin- $s$  bipartite system respond to Hardy's nonlocality test.

## 8. Appendix A

We evaluate  $\langle \alpha_1 \rangle$ ,  $\langle \alpha_1 \alpha_2 \rangle$  and  $\langle \alpha_1 \alpha_2 \alpha_3 \rangle$  in the state  $\rho_0$  given in(4.1).

$$(|\vec{S} \cdot \hat{a}_0, \alpha_0\rangle \equiv |\hat{a}_0, \alpha_0\rangle)$$

$$\langle \alpha_1 \rangle = \sum_{\alpha_1=-s}^s \alpha_1 p(\alpha_1) = \langle \hat{a}_0, \alpha_0 | \vec{S} \cdot \hat{a}_1 | \hat{a}_0, \alpha_0 \rangle = \langle \hat{a}_1, \alpha_0 | e^{i\vec{S} \cdot \hat{n} \theta_1} (\vec{S} \cdot \hat{a}_1) e^{-i\vec{S} \cdot \hat{n} \theta_1} | \hat{a}_1, \alpha_0 \rangle \quad (\text{A.1})$$

where  $\theta_1$  is the angle between  $\hat{a}_0$  and  $\hat{a}_1$  and  $\hat{n}$  is the unit vector along the direction defined by  $\hat{n} = \hat{a}_0 \times \hat{a}_1$ . By using Baker- Hausdorff Lemma

$$e^{iG\lambda} A e^{-iG\lambda} = A + i\lambda[G, A] + \left(\frac{i^2 \lambda^2}{2!}\right) [G, [G, A]] + \dots \quad (\text{A.2})$$

we get,

$$\begin{aligned} \langle \alpha_1 \rangle &= \langle \hat{a}_1, \alpha_0 | \vec{S} \cdot \hat{a}_1 | \hat{a}_1, \alpha_0 \rangle + \frac{i\theta_1}{1!} \langle \hat{a}_1, \alpha_0 | [\vec{S} \cdot \hat{n}, \vec{S} \cdot \hat{a}_1] | \hat{a}_1, \alpha_0 \rangle \\ &+ \frac{i^2 \theta_1^2}{2!} \langle \hat{a}_1, \alpha_0 | [\vec{S} \cdot \hat{n}, [\vec{S} \cdot \hat{n}, \vec{S} \cdot \hat{a}_1]] | \hat{a}_1, \alpha_0 \rangle + \dots \end{aligned} \quad (\text{A.3})$$

By using:

$$\langle \hat{a}_1, \alpha_0 | \vec{S} \cdot \hat{a}_1 | \hat{a}_1, \alpha_0 \rangle = \alpha_0, \quad (\text{A.4})$$

$$\langle \hat{a}_1, \alpha_0 | [\vec{S} \cdot \hat{n}, \vec{S} \cdot \hat{a}_1] | \hat{a}_1, \alpha_0 \rangle = \langle \hat{a}_1, \alpha_0 | (i\vec{S} \cdot (\hat{n} \times \hat{a}_1)) | \hat{a}_1, \alpha_0 \rangle = 0, \quad (\text{A.5})$$

and

$$\langle \hat{a}_1, \alpha_0 | [\vec{S} \cdot \hat{n}, [\vec{S} \cdot \hat{n}, \vec{S} \cdot \hat{a}_1]] | \hat{a}_1, \alpha_0 \rangle = \langle \hat{a}_1, \alpha_0 | \vec{S} \cdot \hat{a}_1 | \hat{a}_1, \alpha_0 \rangle = \alpha_0. \quad (\text{A.6})$$

Terms with odd powers of  $\theta_1$  vanish

$$\langle \alpha_1 \rangle = \alpha_0 - \frac{\theta_1^2}{2!} \alpha_0 + \frac{\theta_1^4}{4!} \alpha_0 - \dots = \alpha_0 \cos \theta_1. \quad (\text{A.7})$$

If the initial state is mixed state(4.1):

$$\langle \alpha_1 \rangle = \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} \alpha_0 \cos \theta_1. \quad (\text{A.8})$$

Further we compute

$$\langle \alpha_1 \alpha_2 \rangle = \sum_{\alpha_1} \alpha_1 |\langle \hat{a}_0, \alpha_0 | \hat{a}_1, \alpha_1 \rangle|^2 \sum_{\alpha_2} \alpha_2 |\langle \hat{a}_1, \alpha_1 | \hat{a}_2, \alpha_2 \rangle|^2.$$

By using (A.7)

$$\begin{aligned} \langle \alpha_1 \alpha_2 \rangle &= \cos \theta_{12} \sum_{\alpha_1} \alpha_1^2 |\langle \hat{a}_0, \alpha_0 | \hat{a}_1, \alpha_1 \rangle|^2 = \cos \theta_{12} \langle \hat{a}_0, \alpha_0 | (\vec{S} \cdot \hat{a}_1)^2 | \hat{a}_0, \alpha_0 \rangle \\ &= \cos \theta_{12} \langle \hat{a}_1, \alpha_0 | e^{i\vec{S} \cdot \hat{n} \theta_1} (\vec{S} \cdot \hat{a}_1)^2 e^{-i\vec{S} \cdot \hat{n} \theta_1} | \hat{a}_1, \alpha_0 \rangle \end{aligned} \quad (\text{A.9})$$

Using the Baker-Hausdorff Lemma, and using

$$\begin{aligned} &\langle \hat{a}_1, \alpha_0 | [\vec{S} \cdot \hat{n}, [\vec{S} \cdot \hat{n}, [\vec{S} \cdot \hat{n}, \dots [\vec{S} \cdot \hat{n}, (\vec{S} \cdot \hat{a}_1)^2]] \dots]] | \hat{a}_1, \alpha_0 \rangle \\ &= \begin{cases} 0 & \text{if } \vec{S} \cdot \hat{n} \text{ occurs odd number of times} \\ 3\alpha_0^2 - s^2 - s & \text{if } \vec{S} \cdot \hat{n} \text{ occurs } 2p \text{ times} \end{cases} \end{aligned} \quad (\text{A.10})$$

we get,

$$\langle \alpha_1 \alpha_2 \rangle = \frac{1}{2} \cos \theta_{12} [(s^2 + s - \alpha_0^2) + (3\alpha_0^2 - s^2 - s) \cos^2 \theta_1]. \quad (\text{A.11})$$

If the initial state is mixed state (4.1),

$$\langle \alpha_1 \alpha_2 \rangle = \frac{1}{2} \cos \theta_{12} \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} [(s^2 + s - \alpha_0^2) + (3\alpha_0^2 - s^2 - s) \cos^2 \theta_1]. \quad (\text{A.12})$$

Next we calculate,

$$\langle \alpha_1 \alpha_2 \alpha_3 \rangle = \sum_{\alpha_1} \alpha_1 |\langle \hat{a}_0, \alpha_0 | \hat{a}_1, \alpha_1 \rangle|^2 \sum_{\alpha_2} \alpha_2 |\langle \hat{a}_1, \alpha_1 | \hat{a}_2, \alpha_2 \rangle|^2 \sum_{\alpha_3} \alpha_3 |\langle \hat{a}_2, \alpha_2 | \hat{a}_3, \alpha_3 \rangle|^2. \quad (\text{A.13})$$

By using (A.7) and (A.11) we get,

$$\langle \alpha_1 \alpha_2 \alpha_3 \rangle = \frac{1}{2} \alpha_0 \cos \theta_1 \cos \theta_{23} \sin^2 \theta_{12} s(s+1) + \frac{1}{2} \cos \theta_{23} (3 \cos^2 \theta_{12} - 1) A, \quad (\text{A.14})$$

where,

$$A = \sum_{\alpha_1} \alpha_1^3 |\langle \hat{a}_0, \alpha_0 | \hat{a}_1, \alpha_1 \rangle|^2 = \langle \hat{a}_1, \alpha_0 | e^{i\vec{S} \cdot \hat{n} \theta_1} (\vec{S} \cdot \hat{a}_1)^3 e^{-i\vec{S} \cdot \hat{n} \theta_1} | \hat{a}_1, \alpha_0 \rangle. \quad (\text{A.15})$$

Using Baker-Hausdorff lemma and

$$\begin{aligned} & \langle \hat{a}_1, \alpha_0 | [\vec{S} \cdot \hat{n}, [\vec{S} \cdot \hat{n}, [\vec{S} \cdot \hat{n}, \dots [\vec{S} \cdot \hat{n}, (\vec{S} \cdot \hat{a}_1)^3]] \dots]] | \hat{a}_1, \alpha_0 \rangle \\ &= \begin{cases} 0 & \text{if } \vec{S} \cdot \hat{n} \text{ occurs odd number of times} \\ Y(X - \alpha_0^3) + X & \text{if } \vec{S} \cdot \hat{n} \text{ occurs } 2p \text{ times} \end{cases} \end{aligned} \quad (\text{A.16})$$

where,

$$\begin{aligned} X &= 6\alpha_0^3 + \alpha_0(1 - 3s(s+1)) \\ Y &= 3^{2p-2} + 3^{2p-4} + \dots + 3^2 = \left(\frac{9}{8}\right)[9^{2p-2} - 1] \end{aligned} \quad (\text{A.17})$$

we get,

$$A = \frac{1}{8} \sum_{j=0}^{\infty} (-1)^j [(9^j - 1)X - (9^j - 9)\alpha_0^3] \left(\frac{\theta_{01}^2}{2j!}\right).$$

This gives,

$$A = \frac{1}{8} \alpha_0 \{ [3\alpha_0^2 + 3s(s+1) - 1] \cos \theta_{01} + [5\alpha_0^2 - 3s(s+1) + 1] \cos 3\theta_{01} \} \quad (\text{A.18})$$

After substituting (A.18) in (A.14) and simplifying,

$$\langle \alpha_1 \alpha_2 \alpha_3 \rangle = \frac{1}{16} \cos \theta_{23} \{ \cos \theta_{11} [M \cos^2 \theta_{12} + N] + R [3 \cos^2 \theta_{12} - 1] \} \quad (\text{A.19})$$

where,

$$\begin{aligned} M &= \alpha_0 [9\alpha_0^2 + s(s+1) - 3] \\ N &= \alpha_0 [5s(s+1) - 3\alpha_0^2 + 1] \\ R &= \alpha_0 [5\alpha_0^2 - 3s(s+1) + 1]. \end{aligned}$$

If the initial state is a mixed state (4.1),

$$\begin{aligned} M &= \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} \alpha_0 [9\alpha_0^2 + s(s+1) - 3] \\ N &= \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} \alpha_0 [5s(s+1) - 3\alpha_0^2 + 1] \\ R &= \sum_{\alpha_0=-s}^{+s} p_{\alpha_0} \alpha_0 [5\alpha_0^2 - 3s(s+1) + 1]. \end{aligned}$$

Also, by using equations (A.12) and (A.13), one obtains:

$$\begin{aligned}
 \langle \alpha_1 \alpha_3 \rangle &= \sum_{\alpha_1} \alpha_1 |\langle \hat{a}_0, \alpha_0 | \hat{a}_1, \alpha_1 \rangle|^2 \sum_{\alpha_2} |\langle \hat{a}_1, \alpha_1 | \hat{a}_2, \alpha_2 \rangle|^2 \sum_{\alpha_3} \alpha_3 |\langle \hat{a}_2, \alpha_2 | \hat{a}_3, \alpha_3 \rangle|^2 \\
 &= \cos \theta_{32} \langle \alpha_1 \alpha_2 \rangle \\
 &= \frac{1}{2} \cos \theta_{32} \cos \theta_{21} [(s^2 + s - \alpha_0^2) + (3\alpha_0^2 - s^2 - s) \cos^2 \theta_1], \tag{A.20}
 \end{aligned}$$

and we can obtain:

$$\begin{aligned}
 \langle \alpha_2 \alpha_3 \rangle &= \sum_{\alpha_1} |\langle \hat{a}_0, \alpha_0 | \hat{a}_1, \alpha_1 \rangle|^2 \sum_{\alpha_2} \alpha_2 |\langle \hat{a}_1, \alpha_1 | \hat{a}_2, \alpha_2 \rangle|^2 \sum_{\alpha_3} \alpha_3 |\langle \hat{a}_2, \alpha_2 | \hat{a}_3, \alpha_3 \rangle|^2 \\
 &= \cos \theta_{32} \langle \alpha_2^2 \rangle \\
 &= \frac{1}{2} s \cos \theta_{32} \{ (s+1) \sin^2 \theta_{12} + \frac{1}{2} (3 \cos^2 \theta_{12} - 1) [1 + (2s-1) \cos^2 \theta_1] \}. \tag{A.21}
 \end{aligned}$$

## 9. Appendix B

Let us consider a situation where an ensemble of systems prepared in state  $\rho = |s, \hat{a}_0 = \alpha_0 \rangle \langle s, \hat{a}_0 = \alpha_0|$  at time  $t = 0$ , is subjected to a measurement of the observable  $A(t_1) = s, \hat{a}_1$  at time  $t_1$  followed by a measurement of the observable  $B(t_2) = s, \hat{a}_2$  at time  $t_2$  ( $t_2 > t_1 > 0$ ), where we have adopted the Heisenberg picture of time evolution. Further, let us assume that both  $A(t_1)$  and  $B(t_2)$  have purely discrete spectra. Let  $\{\alpha_1\} = \{-s, -s+1, \dots, s\}$  and  $\{\alpha_2\} = \{-s, -s+1, \dots, s\}$  denote the eigenvalues and  $P^{A(t_1)}(\alpha_1) = |s, \hat{a}_1 = \alpha_1 \rangle \langle s, \hat{a}_1 = \alpha_1|$ ,  $P^{B(t_2)}(\alpha_2) = |s, \hat{a}_2 = \alpha_2 \rangle \langle s, \hat{a}_2 = \alpha_2|$  the corresponding eigenprojectors of  $A(t_1)$  and  $B(t_2)$  respectively. Then the joint probability that a measurement of  $A(t_1)$  yields the outcome  $\alpha_1$  and a measurement of  $B(t_2)$  yields the outcome  $\alpha_2$  is given by

$$\begin{aligned}
 Pr_{A(t_1), B(t_2)}^\rho(\alpha_1, \alpha_2) &= \text{Tr} \left[ P^{B(t_2)}(\alpha_2) P^{A(t_1)}(\alpha_1) \rho P^{A(t_1)}(\alpha_1) P^{B(t_2)}(\alpha_2) \right] \\
 &= |\langle s, \hat{a}_0 | s, \hat{a}_1 \rangle|^2 |\langle s, \hat{a}_1 | s, \hat{a}_2 \rangle|^2. \tag{B.1}
 \end{aligned}$$

We know that  $|s, a = \alpha \rangle = \sum_{m=-s}^{+s} d_{\alpha, m}^{(s)}(\beta) |m \rangle$  where  $d_{\alpha, m}^{(s)}(\beta) \equiv \langle s, \alpha | \exp\left(\frac{-iS_y(\beta)}{\hbar}\right) |s, m \rangle$  and it obtains Wigner's formula (Sakurai, n.d.) and  $\alpha_i \in -s, \dots, s$  and  $\beta_i$  is the angle between the  $\hat{a}_i$  and the z axes. In contrast,

$$\begin{aligned}
 |\langle s, \hat{a}_1 | s, \hat{a}_2 \rangle| &= \sum_m \langle m | d_{\alpha_1, m}^{*(s)}(\beta_1) \sum_{m'} d_{\alpha_2, m'}^{(s)}(\beta_2) |m' \rangle \\
 &= \sum_m d_{\alpha_1, m}^{*(s)}(\beta_1) d_{\alpha_2, m}^{(s)}(\beta_2) \\
 &= d_{\alpha_1 \alpha_2}^{(s)}(\beta_2 - \beta_1). \tag{B.2}
 \end{aligned}$$

So, we obtain

$$pr_{QM}(\alpha_1, \alpha_2) = |d_{\alpha_0 \alpha_1}^{(s)}(\beta_1 - \beta_0)|^2 |d_{\alpha_1 \alpha_2}^{(s)}(\beta_2 - \beta_1)|^2. \tag{B.3}$$

## 10. References

- Alter, O. & Yamamoto, Y. (n.d.). *Quantum Measurement of a Single System*, A Wiley-interscience Publication John Wiley and sons, INC.
- Anderson, E., Barnett, S. M. & Aspect, A. (2005). *Phys. Rev. A* 72: 042104.
- Ballentine, L. E. (1990). *Quantum Mechanics*, prentice Hall, Englewood Cliffs, NJ.
- Basu, S., Bandyopadhyay, S., Kar, G. & Home, D. (2001). *Phys. Lett. A* 279: 284.
- Beck, C. & Graudenz, D. (1992). *Phys. Rev. A* 46: 6265.
- Belinskii, A. V. & Klyshko, D. N. (1993). *Phys. Usp.* 36: 653.
- Bell, J. S. (1964). *Physics* 1: 195.
- Bell, J. S. (1966). *Rev Mod. Phys.* 38: 447.
- Bennett, C. H., Brassard, G., Popescu, S., Schumacher, B., Smolin, J. & Wootters, W. K. (1996). *Phys. Rev. Lett* 76: 722.
- Bohm, D. (1952). *Phys. Rev.* 85: 166.
- Braunstein, L. & Caves, C. M. (1990). *Ann. Phys. (NY)* 202: 22.
- Brukner, C., Taylor, S., Cheung, S. & Vedral, V. (2004). *The Seventh International Conference on Quantum Communication Measurement and Computing, Glasgow, United Kingdom*, pp. 25–29.
- Cabello, A. (2002a). *Phys. Rev. A* 65: 062105.
- Cabello, A. (2002b). *Phys. Rev. A* 65: 032108.
- Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. (1969). *Phys. Rev. Lett* 23: 880.
- Clifton, R. & Niemann, P. (1992). *Phys. Lett. A* 166: 177.
- Collins, D., Gisin, N., Linden, N., Massar, S. & Popescu, S. (2002). *Phys. Rev. Lett.* 88: 040404.
- Collins, D., Gisin, N., Popescu, S., Roberts, D. & Scarani, V. (2002). *Phys. Rev. Lett* 88: 170405.
- Collins, D. & Popescu, S. (2001). *J. Phys. A :Math. Gen* 34: 6821.
- Einstein, A., Podolski, B. & Rosen, N. (1935). *Phys Rev* 47.
- Fine, A. (1982). *Phys. Rev. Lett.* 48: 291.
- Gisin, N. & Bechmann-Pasquinucci, H. (1998). *Phys. Lett. A* 246.
- Gisin, N. & Peres, A. (1992). *Phys. Lett. A* 162.
- Goldstein, S. (1994). *Phys. Rev. Lett.* 72: 1951.
- Greenberger, D. M., Horne, M. A. & Zeilinger, A. (1990). *Am. J. Phys.* 58: 1131–1143.
- Hardy, L. (1992). *Phys. Rev. Lett* 68: 2981.
- Hardy, L. (1993). *Phys. Rev. Lett.* 71: 1665.
- Heywood, P. & Redhead, M. L. (1983). *Found. Phys.* 13: 481.
- Jarrett, J. P. (1984). On the physical significance of the locality conditions in the bell arguments, *Nous* 18: 569–89.
- Kaszlikowski, D., Gnacinski, P., Zukowski, M., Miklaszewski, W. & Zeilinger, A. (n.d.). *Phys.Rev. Lett.* 85.
- Kochen, K. & Specker, E. (1967). *J. Math.* 17: 59.
- Leggett, A. J. & Garg, A. (1985). *Phys. Rev. Lett.* 54: 857.
- Mermin, N. D. (1990). *Phys. Rev. Lett* 65: 1838.
- Pagonis, C. & Clifton, R. (1992). *Phys. Lett. A* 168: 100.
- Parasuram, K. & Ghosh, S. (n.d.). *arXiv: quant-ph/0903.3020* .
- Paz, J. P. & Mahler, G. (1993). *Phys. Rev. Lett.* 71: 3235.
- Peres, A. (1992). *Phys. Rev. A* 46: 4413.
- Peres, A. (1993). *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers.
- Popescu, S. (1995). *Phys. Rev. Lett.* 74: 2619.
- Redhead, M. (1987). *Incompleteness, Nonlocality, and Realism*, Clarendon Press, Oxford.

- Sakurai, J. J. (n.d.). *Modern Quantum Mechanics, (revised edition)*, Addison-Wesley, 1999.
- Scarani, V. & Gisin, N. (2001). *Phys. Rev. Lett.* 87: 117901.
- Seevinck, M. & Svetlichny, G. (2002). *Phys. Rev. Lett.* 89: 060401.
- Shimony, A. (n.d.). Historical philosophica, and physical inquiries into the foundations of quantum mechanics, *In Sixty-Two Years of Uncertainty*, Plenum, New York.
- Svetlichny, G. (1987). *Phys. Rev. D* 35: 3066.
- Werner, R. F. (1989). *Phys. Rev. A* 40: 4277.
- Werner, R. & Wolf, M. (2000). *Phys. Rev. A.* 61: 062102.
- Zuckowski, M. & Brukner, C. (2002). *Phys. Rev. Lett.* 88: 21041.

# The Gluon Emission Model for Vector Meson Decay\*

D. White

*Dept. of Biological, Chemical and Physical Sciences,  
Roosevelt University, Chicago, Illinois  
USA*

## 1. Introduction

Quantum Chromodynamics (QCD), the heart of the so-called Standard Model, was developed along the lines of the most successful theoretical structure in all of physics, namely, Quantum Electrodynamics (QED), which represents the interactions between the electron and the electromagnetic field, photons serving as the mediator between the two entities. QCD, therefore, contains many objects which are analogous to those within QED. There are the quarks, for example, which come in six “flavors” (up, down, strange, charm, bottom, top), serving as the analogous construct to the electron and its cousins, the muon and the tauon. Plexification, of course, always accompanies the proceeding by analogy to any theory, and so we find that quarks must come not only in flavors, but also in “colors” (three total), and they must be fractionally charged (one or two thirds of the electron charge in absolute value). What mediates the quark and the strong field, analogous to the photon in QED, is the gluon. In QED all quantum events are described by a coupling between the electron and the photon, called the fine structure constant of magnitude approximately 0.007. Regarding the coupling between quarks and the strong field responsible for hadronization ... the production of hadron pairs in a colliding beams experiment, for example ... the situation in QCD is quite different. Work continues at present in the field of high-energy physics to determine the precise nature of the quark-gluon coupling, but one overarching behavior pattern of such coupling, called the strong coupling, is that it is *not* a constant. Rather, it varies generally as the reciprocal of the natural logarithm of the energy wrapped up in the colliding beams.

In the work which follows we will have occasion to investigate the phenomenon of vector meson formation and decay in accord with a QCD model, called the Gluon Emission Model (GEM), first developed by F. E. Close in the 1970s. The GEM follows rigorously the precepts of QED proper, the *only* QCD quantity entering into the calculations being the strong coupling parameter, which replaces the fine structure constant in the relevant places. The GEM thus provides for a self-contained formalism that follows the constructs of QED essentially as closely as is possible at the present time. As we will see, even the precise form for the strong coupling parameter may be determined within the GEM, the valid range

---

\* Much work presented in this Chapter is taken from D. White, “GEM and the  $Y(1S)$ ”, *The Journal of Informatics and Mathematical Sciences*, Vol. 2, Nos. 2 & 3 (2010), pp. 71 - 93.

being the range of energy encompassed by the known vector mesons themselves. Let us begin by reviewing a central feature of QED:

In all quantum systems in which natural decay occurs between an excited level and the ground state, the absorption cross-section goes as<sup>1</sup>

$$\sigma(\omega) = K\alpha|V|^2 (1/m)^2(1/\omega)L(\omega) \quad (1)$$

where  $K$  is a constant,  $\omega$  represents photon frequency,  $|V|^2$  represents the square of the matrix element descriptive of the photon emission process, the system has mass  $m$ ,  $L(\omega)$  is a Lorentz Amplitude with a peak at  $\omega = \omega_0$  and with a width  $\Gamma$ , and  $\alpha = (1/137.036)$  represents the fine structure constant. Assuming "asymptotic freedom", i.e., that we may ignore the masses of the decay products (light hadron pairs) in relation to the total energy involved in the system under investigation, we may employ Eq. 1 to predict the width of vector mesons by making the following substitutions to take us from a general quantum electrodynamics (QED) to a specific quantum chromodynamics (QCD) process:

We substitute for the photon frequency  $\omega$  the gluon energy  $Q_0$ . We evaluate the right hand side of Eq. 1 at a specific vector meson mass,  $m_v$ , i.e.,  $Q_0 = m = m_v$ . (Hence, the associated Lorentz Amplitude equals unity.) We require  $|V|^2$  to be proportional to  $\sum_i(q_i)^4$ , where  $q_i$  = quark charge (in units of electron charge magnitude) associated with the quarks comprising the relevant vector meson. The above criterion is consistent with a spin-spin interaction<sup>2</sup> proportional to  $q_i^2$ , where  $i$  denotes quark flavor, giving rise to spin-flip transitions, and the sum is required only in the case of the  $\rho$ , as it comprises both the up quark ( $u$ ) of charge  $q_u = 2/3$  and the down quark ( $d$ ) of charge  $q_d = -1/3$ . We postulate  $|V|^2$  to be proportional to only  $\sum_i(q_i)^4$ , i.e., the precise form of the interaction is universal to all vector mesons in their ground states, except for quark charge differences.

We replace  $\alpha$  by  $\alpha_s$ , the strong coupling parameter, which has the well-known form from QCD gauge invariance theories of<sup>3</sup>:

$$\alpha_s = B[\ln(Q_0/\Lambda)]^{-1} \quad (2)$$

where  $B$  is a constant and  $\Lambda$  is a parameter, called the QCD scale factor, to be determined. Again, we emphasize that commensurate with the above replacements is that we must assume that the initial energy involved in the formation of a given vector meson is extremely high, i.e., in the "asymptotically free" region of energy space, where the masses of emerging hadron pairs as decay products can be neglected. Accordingly, then, we find in terms of the above ansatz (normalizing to the  $\rho$ )

$$\Gamma_v = A(m_\rho/m_v)^3(\sum_i(q_i)^4)[\ln(m_v/\Lambda)]^{-1} \quad (3)$$

where  $\Gamma_v$  represents the width of a given vector meson,  $v$ ,  $A$  is a constant to be determined, and  $\Lambda$ , the QCD scale factor, is to be determined, as well.

The constants,  $A$  and  $\Lambda$ , may be determined (see Section 2 below for the determination of the values of  $A$  and  $\Lambda$ ) by simultaneously fitting the width of the  $\rho$  and the width of the kaon branch of the  $\phi$  to the form of Eq. 3 above, and  $B$  may be determined by evaluating  $\alpha_s$  at the  $Y(1S)$  energy through the utilization of the experimentally determined partial width associated with the  $Y(1S) \rightarrow e^+e^-$  decay in conjunction with the GEM-theoretical hadronic width of the  $Y(1S)$  (see Section 2 below). In conventional terms we will find that the hadronic width of any vector meson may be expressed as the following:

$$\Gamma_v \approx (\alpha_s / 2\pi)(10,042)(2m_e)(m_\rho/m_v)^3(\sum_i(q_i)^4) \quad (4)$$

where  $m_e$  represents the electron mass of 0.511 Mev, so that  $2m_e = 1.022$  Mev.,  $\alpha_s$  represents the strong coupling parameter, given by  $\alpha_s = 1.2[\ln(m_v/50 \text{ Mev})]^{-1}$ ,  $m_\rho$  represents the mass of the  $\rho$ -meson,  $m_v$  represents the mass of the vector meson with designate "v", and  $q_i$  represents the charge of the relevant quark type(s) "i" to undergo the spin-flip to form the vector meson under consideration. As mentioned above, the  $q_i$  involved in  $\rho$  formation, are  $q_u = 2/3$  and  $q_d = -1/3$ , where "u" designates an "up quark" and "d" designates a "down quark". Only  $q_s = -1/3$ , where "s" designates a "strange quark", is involved in the formation of the kaon branch of the  $\phi$ , whereas  $q_u$ ,  $q_d$ , and  $q_s$  are all involved in the formation of the  $K^*(892)$ . In addition, as we will see below, the  $q_i$  mainly associated with the  $J(3097)$  is actually  $q_s$ , and that associated with the  $Y(1S)$  is actually  $q_c = 2/3$ , where "c" is the designate for the "charm quark".

## 2. The constants, A and $\Lambda$ , and the specific form of $\alpha_s$

Focusing now upon Eq. 3 above, A and  $\Lambda$  can be determined simultaneously by utilizing the appropriate  $q_i$  and  $m_v$  associated with the  $\rho$  and  $\phi$  mesons in conjunction with their published widths (see, for example, pdg.lbl.gov; "Meson Table" (2004)). The result<sup>4</sup> is that  $A \approx 1960$  Mev and  $\Lambda \approx 50$  Mev. What is most interesting about the above result is the extremely small value for  $\Lambda$  as per the GEM applied to the  $\rho$  and the  $\phi$ , as the accepted value<sup>5</sup> for  $\Lambda$  as of 1996 is around 290 Mev. Nevertheless, in QCD,  $\Lambda$  is considered to be an arbitrary parameter, so no "rules" laid forth within QCD itself are violated by such result. Hence, the GEM formulation of  $\Gamma_v$  becomes:

$$\Gamma_v \approx (1960 \text{ Mev}) (m_\rho/m_v)^3(\sum_i(q_i)^4) [\ln(m_v/50 \text{ Mev})]^{-1} \quad (5)$$

Now, in the asymptotically free regions of energy space we expect the ratio of a vector meson's electron/positron partial width ( $\Gamma_{ee}$ ) to its hadronic width to be approximately  $(\alpha/\alpha_s)$ , where  $\alpha$  is the fine structure constant =  $(1/137.036)$ , due to electromagnetic, rather than strong coupling at the pair vertex. At the  $Y(1S)$  energy ( $m_Y = 9460$  Mev according to the 2004 "Meson Table") from the 2004 "Meson Table" on p. 86,  $\Gamma_{ee} = 1.31$  Kev, while the *theoretical* hadronic width, as per the GEM, of the  $Y(1S)$  is, as we will see below, assuming single gluon emission, 41 Kev. (As we will in addition see below, the GEM requires that the resonant state at the  $Y(1S)$  energy be characterized by an essentially instantaneous transition from its root  $bb^*$  state (b represents the bottom quark, while  $b^*$  represents the bottom anti-quark) to an excited  $cc^*$  state (c represents the charm quark, while  $c^*$  represents the charm anti-quark) before its decay.)

Thus,

$$\alpha_s |_{Y \text{ energy}} \approx \alpha (41/1.31) = 0.2284 \quad (6)$$

Setting (from the general form for  $\alpha_s$  described above)  $0.2284 = B[\ln(9460/50)]^{-1}$ , we obtain  $B \approx 1.2$ ; hence, the GEM determines that

$$\alpha_s \approx 1.2[\ln(Q/50 \text{ Mev})]^{-1} \quad (7)$$

With the strong coupling parameter defined as above, it presumably valid over the entire range of energy from the  $\rho$  energy to that of the  $Y(1S)$ , the GEM width formula for vector mesons takes the form, then, of

$$\Gamma_v \approx (1633 \text{ Mev}) (m_p/m_v)^3 (\sum_i (q_i)^4) \alpha_s \tag{8}$$

Expressing Eq. 8 in more conventional form, we have:

$$\Gamma_v \approx (\alpha_s / 2\pi)(10,263 \text{ Mev})(m_p/m_v)^3 (\sum_i (q_i)^4) \tag{9}$$

Yet more formally, we finally obtain:

$$\Gamma_v \approx (\alpha_s / 2\pi)(10,042)(2m_e)(m_p/m_v)^3 (\sum_i (q_i)^4) \tag{10}$$

### 3. The Feynman Diagrams of the GEM

As per F. Close, the GEM treats the virtual photon and the gluon as, essentially, two aspects of the same entity, which we will call “the four-momentum propagator” designated as “ $\zeta$ ”. Thus, as stated in Section 2 above, the ratio of the partial width associated with a given decaying pair of quarks comprising a given vector meson associated with electron-positron decay to the hadronic width of same is simply  $(\alpha/\alpha_s)$ , where, again, “ $\alpha$ ” represents the fine structure constant =  $(1/137.036)$ . Hence, the general form for the partial width of a vector meson undergoing  $e^+e^-$  decay would be given by

$$\Gamma_{v-ee} \approx (\alpha/2\pi)(10,042)(2m_e)(m_p/m_v)^3 (\sum_i (q_i)^4) \tag{11}$$

A relevant Feynman Diagram will make the various aspects of the GEM easier to picture, so let us look now to Fig. 1 below, which represents the Feynman Diagram (FD) associated the formation and decay of vector meson “X” in its simplest possible form.

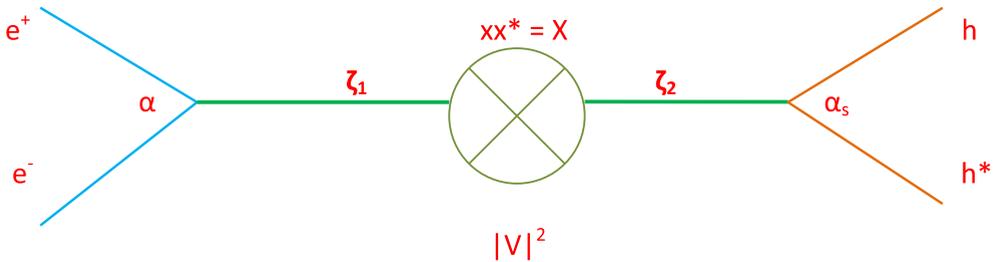


Fig. 1. Basic Feynman Diagram for Conventional Vector Meson Formation/Decay via the GEM

In Figure 1  $\zeta_1$  represents, in part, a virtual photon created at the  $e^+e^-$  annihilation vertex, coupling at said vertex represented as  $\alpha$ ; then, in Close’s terms, the virtual photon couples to a gluon with coupling strength “1”, which then couples to the  $xx^*$  ... a given quark - anti-quark pair, also with coupling strength “1”. In our notation  $\zeta_1$  simply represents a four-momentum propagator, created at the  $e^+e^-$  vertex and absorbed (as a gluon) at the  $xx^*$  node. The details of the absorption of  $\zeta_1$  are contained in the integrated absorption cross-section as exhibited in the Introduction, and  $|V|^2$ , proportional to  $q_x^4$ , describes the formation of the spin one resonance. From there  $\zeta_2$  (a gluon) is emitted, resulting in coupling to hadrons ( $h$ ;  $h^*$ ), the coupling at the latter vertex of magnitude  $\alpha_s$ . The calculation of the width of the  $xx^*$  state then, given the stated mechanism of a spin-flip of one of the “x quarks” due to a spin -

spin interaction proportional to  $q_x^2$ , proceeds straight along the dictates of standard QED, except for the replacement of  $\alpha$  by  $\alpha_s$  at the  $hh^*$  vertex.

For comparison, immediately below we present the FD associated with the same X meson, assumed to exist in the realm of asymptotic freedom, decaying into an electron-positron pair.

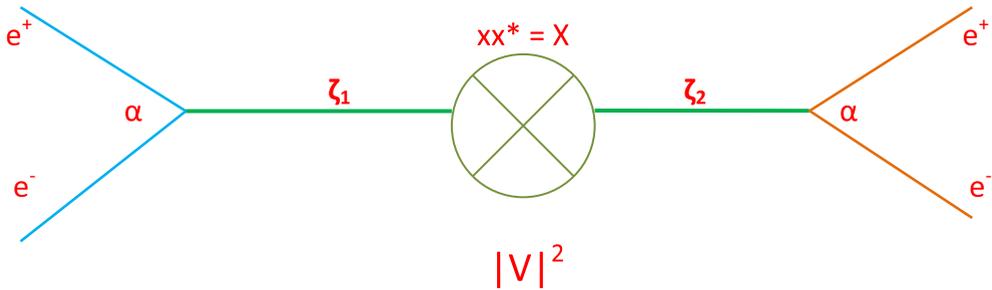


Fig. 2. Basic Feynman Diagram for Conventional Vector Meson Formation & Decay into an Electron/Positron Pair via the GEM

The only fundamental difference between Figure 1 and Figure 2 is that in Figure 2  $\zeta_2$  starts out as a gluon and ends up as a virtual photon at the right hand vertex, at which point the coupling, of course, is now  $\alpha$ . Hence, all in the width calculation associated with Figure 1 is the same in Figure 2, except that  $\alpha_s$  in Eq. 10 is replaced by  $\alpha$ . Of note, too, and we shall return to the point made here, Figure 2 represents rigorously a straight-forward calculation in QED, again, given the stated mechanism for the formation of the resonance state. However, it is also important to note that Figure 2 applies only to vector mesons existing in the realm of “asymptotic freedom”, i.e., to the J(3097), the Y(1S), and “Toponium”, or the “T” meson.

Immediately below we will view the detailed FDs required by the GEM to describe the widths of the  $\rho$ , the  $\phi$ , the  $K^*(892)$  ... a very interesting case, as the  $K^*(892)$  is not conventionally thought of as a vector meson per se, though it is of the spin one variety ... the J(3097), the Y(1S), and the “T”.

**The  $\rho$ -meson**

Although the width of the  $\rho$  (and the  $\phi$ ) as determined by the GEM is guaranteed to be a match to experiment by construction, the  $\rho$  is a good place to start with the elucidation of the application of the GEM to the various spin one mesons because of the simplicity involved. Let us begin by viewing Figure 3 below ... the FD associated with the formation and decay of the  $\rho$  meson.

In Figure 3  $\zeta_1$  represents a virtual photon created at the  $e^+e^-$  vertex which transmutes to a gluon, which, in turn, is absorbed by the  $[q_u q_u^* + q_d q_d^*]$  combination;  $\zeta_2$  represents the emitted gluon, which converts to pion pairs. The application of Eq. 10 results in the following for the hadronic width of the  $\rho$ :

$$\Gamma_\rho \approx (\alpha_s / 2\pi)(10,042)(2m_\rho)(\Sigma_i(q_i)^4) \approx (\alpha_s / 2\pi)(10,042)(2m_\rho)(17/81) \tag{12a}$$

where

$$\alpha_s = 1.2[\ln(776/50)]^{-1} = 0.4376 \tag{12b}$$

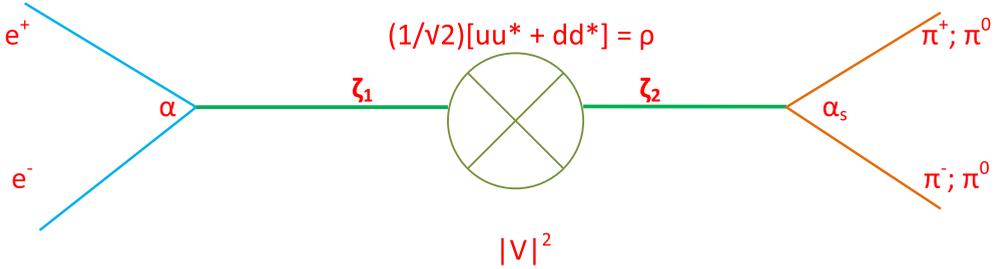


Fig. 3. Basic Feynman Diagram for Formation and Decay of the  $\rho$  meson via the GEM

Hence,

$$\Gamma_\rho \approx 150 \text{ Mev} \tag{12c}$$

Though adaptation of Figure 2 and Eq. 11 do not formally apply, as asymptotic freedom does not apply to the  $\rho$ , we note that in the event that it were to apply, we would obtain for the electron/positron partial width,  $\Gamma_{\rho-ee}$ , the following:

$$\Gamma_{\rho-ee} \approx (\alpha / 2\pi)(10,042)(2m_e)(17/81) \approx 2.50 \text{ Mev}$$

a figure<sup>7</sup> about 355 times too high, indicating that the transmutation coupling of the  $\zeta_2$  gluon to its virtual photon identity is only 0.0028, as opposed to 1 in the asymptotically free energy regime.

**The  $\phi$ -meson**

Application of the GEM to the kaon branch of the  $\phi$  meson ( $\phi_K$ ) follows similar lines as to the  $\rho$ . The FD associated with the formation and decay of the kaon branch of the  $\phi$  follows:

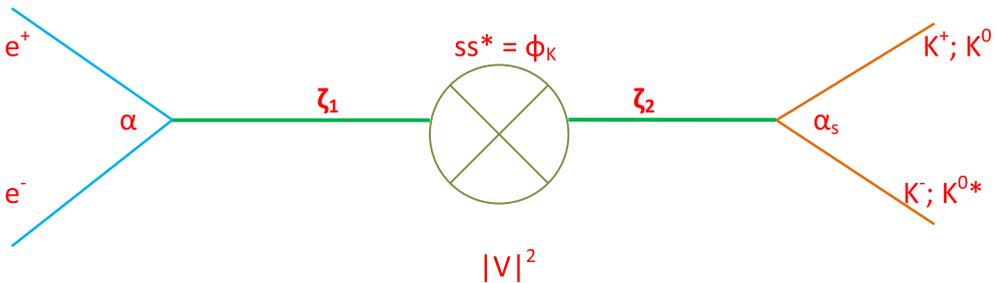


Fig. 4. Basic Feynman Diagram for Formation and Decay of the Kaon Branch of the  $\phi$  Meson via the GEM

For the hadronic width of the kaon branch of the  $\phi$  we obtain:

$$\Gamma_{\phi-K} \approx (\alpha_s / 2\pi)(10,042)(2m_e)(m_\rho / m_v)^3 (\sum_i (q_i)^4) \approx (\alpha_s / 2\pi)(10,042)(2m_e)(776 / 1019)^3 (1/81) \tag{13a}$$

where

$$\alpha_s = 1.2[\ln(1019/50)]^{-1} = 0.3981 \tag{13b}$$

Hence,

$$\Gamma_{\phi-K} \approx 3.55 \text{ Mev} \quad (13c)$$

Again applying Eq. 11 to the kaon branch of the  $\phi$ , we obtain for its  $e^+e^-$  partial width the following:

$$\Gamma_{\phi-K-ee} \approx (\alpha / 2\pi)(10,042)(2m_e)(776/1019)^3(1/81) \approx 0.0650 \text{ Mev}$$

a figure<sup>7</sup> still way too high as compared to experiment, but here about 52 times so, indicating that the  $\zeta_2$  gluon to virtual photon transmutation coupling has risen to 0.0194.

### The $K^*(892)$

The situation regarding the  $K^*(892)$  is highly interesting. Close had developed the GEM in the 1970s to describe two distinct processes: (1) the production of pion pairs associated with the  $\rho$  resonance and (2) the production of kaon pairs associated with the  $\phi$  resonance. In a sense, then, the GEM was first envisioned to be "route specific", i.e., the spin-flip process involving up and down quarks, which resonates at the  $\rho$  mass, was thought of as "the pion route" in thinking of the decay of quark - anti-quark structures, while the spin-flip process involving the strange quark, which resonates at the  $\phi$  mass, was thought of as the corresponding "kaon route". At that time no one had thought of applying the GEM to the  $K^*(892)$ , because, although energetically possible, the  $K^*(892)$  did not exhibit "a pion route" in its decay; rather, the  $K^*(892)$  decays almost exclusively into various  $\{\pi, K\}$  combinations, with equal probability of occurrence among the various allowed decay products. Such circumstance led to the invention of the "isospin" quantum number, a half integer value for which signifying a forbidden decay route that is energetically possible. However, since the spin associated with the  $K^*(892)$  is one, it is quite feasible that the GEM, appropriately mitigated to fit the situation pertaining to the  $K^*(892)$ 's isospin, may be applied to the  $K^*(892)$  resonance. In fact, the GEM has been applied to the  $K^*(892)$  quite successfully<sup>4, 8</sup>. The reasoning leading to the proper mitigation is as follows:

Since pions and kaons are the decay products of the  $K^*(892)$ , with the various types of pions combining with correspondingly allowed various types of kaons and all types showing up with equal probability, it is reasonable to assume that the  $K^*(892)$  ... for purposes of discussion here considered as a composite entity of mass, 894 Mev, i.e., no distinction as to charged mode versus neutral mode being made ... comprises a linear combination of  $\{uu^*, dd^*, \text{ and } ss^*\}$  in equal measure. Symbolically, we may represent the  $K^*(892)$ , therefore, as

$$K^*(892) = (1/\sqrt{3})[uu^* + dd^* + ss^*] \quad (14)$$

Now, the associated value of  $(\sum_i(q_i)^4)$  would be  $(18/81)$ , but the "pion route" does not occur, though it is energetically possible. So, segmenting the decay in terms of "routes", the  $\{\pi, K\}$  route, whose  $(\sum_i(q_i)^4) = (18/81)$  does occur, whereas the "pion route", whose  $(\sum_i(q_i)^4) = (17/81)$  does not occur. The allowed route is thus favored over the forbidden route by the factor  $(18/17)$ , therefore. Hence, we postulate that the isospin quantum number =  $(1/2)$  assigned to the  $K^*(892)$  signifies that of the energetically possible routes available to the  $K^*(892)$  resonance,  $(18/35)$  of them manifests in the decay process (the  $\{\pi, K\}$  route), whereas  $(17/35)$  of them fails to materialize (the pion route). We thus multiply the right hand side of Eq. 10 by  $(18/35)$  to obtain the width of the  $K^*(892)$ . First, let us view the associated FD: The GEM yields for the width of the  $K^*(892)$  the following:

$$\Gamma_{K^*} \approx (18/35)(\alpha_s / 2\pi)(10,042)(2m_e)(m_p/m_v)^3(\sum_i(q_i)^4)$$

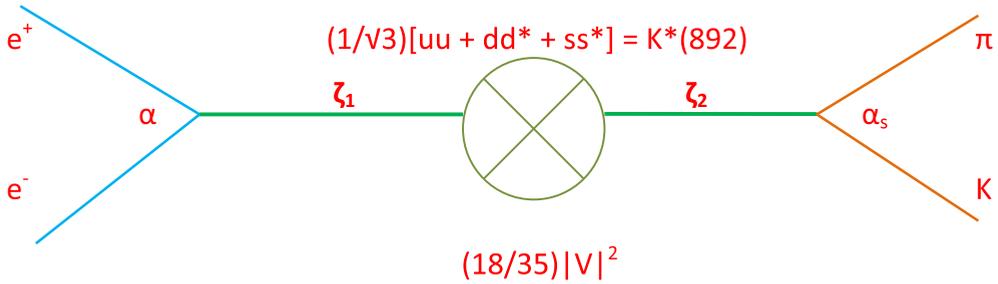


Fig. 5. Basic Feynman Diagram for Formation and Decay of the  $K^*(892)$  via the GEM

$$\approx (18/35)(\alpha_s / 2\pi)(10,042)(2m_e)(776/894)^3(18/81) \tag{15a}$$

where

$$\alpha_s = 1.2[\ln(894/50)]^{-1} = 0.4161 \tag{15b}$$

Hence,

$$\Gamma_{K^*} \approx 50.80 \text{ Mev} \tag{15c}$$

The average of the widths associated with the charged and neutral modes of the  $K^*(892)$  is stated<sup>9</sup> as  $\Gamma_{K^*}(\text{PDG}) = 50.75 \text{ Mev}$ . Hence, the GEM as applied to the  $K^*(892)$  provides for fabulous agreement with experiment. Moreover, the GEM demonstrates quite clearly that the  $K^*(892)$  is not a “strange meson” in the usual sense, i.e., it is seen not as a  $us^*$ ,  $su^*$ ,  $ds^*$ , or  $sd^*$  structure at all; rather it is seen, similar to the theoretical structures of the  $\rho$  and the  $\phi$ , as comprising a linear combination of more than one type of quark - anti-quark pair, its specific nature expressed via Eq. 14.

### The J(3097)

Application of the GEM in accord with Figure 1, with  $x = c$ , seems reasonably straightforward, but it turns out to be problematic. However, when one sees that the hadronic width of the J(3097), designated as simply the “J” henceforth, given by the application of Eq. 10 in accord with Figure 1 with  $x = c$ , is roughly sixteen times too large, as compared to experimental results, coupled with the fact that the hadronic width of the  $Y(1S)$  given by the application of Eq. 10 in accord with Figure 1 with  $x = b$  is roughly sixteen times too small, as compared with experimental results, it becomes obvious as to what physically must transpire as regards both the J and the  $Y(1S)$ . Restricting the discussion to the J for the time being, in what we call “the zeroth order approximation”, the basic  $cc^*$  structure of the J must make a point-like transition to an  $ss^*$  structure of equal mass, whereupon one of the s quarks undergoes a spin flip to form the associated resonance<sup>10</sup>, the point-like transition from  $cc^*$  to  $ss^*$  instantaneous, thus having no influence on the J’s width. Indeed, the resonance does not even form until an s (or  $s^*$ ) quark undergoes a spin-flip. That the  $cc^*$  to  $ss^*$  transition is necessary is quite understandable: The J is not massive enough for it to be able to decay into hadrons via emission of two c quarks; hence, it must transition to a quark pair of lesser bare mass each. The simplest possible assumption is that the  $cc^*$  transitions to the quark pair type characterized by the next smallest mass, viz., the s type. Nothing prevents the  $cc^*$  structure from decaying into leptons ( $e^+e^-$  and  $\mu^+\mu^-$ ), however. It is found<sup>10</sup> in fact, that in order for both the hadronic width of the J and the leptonic width of the J as determined via the GEM to match the results of experiment,  $(8/9)^{\text{ths}}$  of the  $cc^*$  structure must undergo a slightly “un-

point-like" transition to  $ss^*$ , described by a form factor,  $f < 1$ , which, in turn, decays into both hadrons and leptons as per Eq. 10 and Eq. 11, respectively, while  $(1/9)^{\text{th}}$  of the original  $cc^*$  structure remains to decay into leptons exclusively. We may picture the complete details of the J formation and decay via the following two arrays of FDs, the first such array descriptive of what we may now call "the first order approximation" to the width of the J, the second such array descriptive of what we call "the second order approximation".

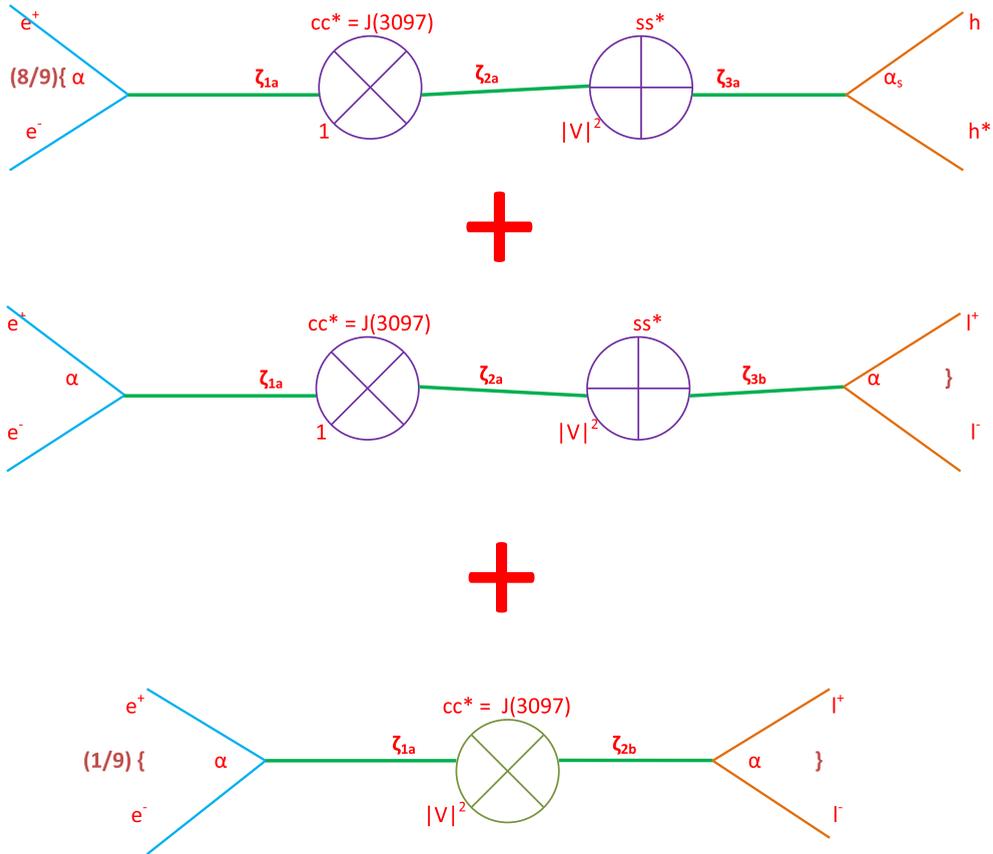


Fig. 6a. Feynman Diagram Array Characterizing the Formation and Decay of the J(3097) in First Order Approximation via the GEM

In Figure 6a above "1" represents a leptonic decay product,  $\zeta_{2a}$  represents the gluon involved in a point-like transition from  $cc^*$  to  $ss^*$ , and all other "zeta" designates should be understood from previous discussion. Transforming the schematic representation of Figure 6a into the calculation of the full (hadronic plus leptonic) width of the J in first order approximation, denoted as  $\Gamma_{J\text{-full-1}}$ , proceeds as follows (the factors of "2" in Eq. 16a, immediately in front of the factors " $(\alpha / 2\pi)$ " take into account muon pair production in accord with "e-u universality"):

$$\Gamma_{J\text{-full-1}} \approx (8/9)\{(\alpha_s / 2\pi)(10,042)(2m_e)(m_p/m_l)^3(q_s)^4 + 2(\alpha / 2\pi)(10,042)(2m_e)(m_p/m_l)^3(q_s)^4\}$$

$$+ (1/9)\{2(\alpha / 2\pi)(10,042)(2m_e)(m_p / m_j)^3(q_c)^4\} \quad (16a)$$

Thus,

$$\begin{aligned} \Gamma_{J\text{-full-1}} \approx & (8/9)\{(\alpha_s / 2\pi)(10,042)(2m_e)(776/3097)^3(1/81) + \\ & 2(\alpha / 2\pi)(10,042)(2m_e)(776/3097)^3(1/81)\} + \\ & (1/9)\{2(\alpha / 2\pi)(10,042)(2m_e)(776/3097)^3(16/81)\} \end{aligned}$$

The value of the strong coupling parameter at the J mass is given by

$$\alpha_s = 1.2[\ln(3097/50)]^{-1} = 0.2908 \quad (16b)$$

Therefore,

$$\Gamma_{J\text{-full-1}} \approx (8/9)\{92.2491 \text{ Kev} + 4.6298 \text{ Kev}\} + (1/9)\{74.0769 \text{ Kev}\} \approx 94.35 \text{ Kev} \quad (16c)$$

The value for  $\Gamma_{J\text{-full-1}}$  obtained via the first approximation of the GEM is a match to experiment, as according to PDG (2009), the full width of the J via experiment is  $(93.2 \pm 2.1)$  Kev . As well, the hadronic width alone via the first approximation of the GEM is a match to experiment (82.00 Kev via the GEM vs.  $(81.7 \pm 0.5)$  Kev via experiment (PDG (2009)); the leptonic width via the first approximation of the GEM is 12.35 Kev, which is about 11% more than that reported by the PDG currently  $(11.10 \pm 0.16)$  Kev (PDG (2009)).

The first approximation assumes that  $(8/9)^{\text{th}}$  of the original  $cc^*$  state undergo a point-like transition to an excited  $ss^*$  state, leaving  $(1/9)^{\text{th}}$  of the original  $cc^*$  state to decay into leptons. A point-like transition is instantaneous, so it has no effect on the width of the original construction (i.e., the J). In terms of a *form factor*,  $f$ , a point-like transition is consistent with  $f = 1$ . As it is difficult to see how *any fraction* of the original  $cc^*$  state could “know” to make an instantaneous transition, leaving a remnant to do other things, we believe a second order approximation is in order. Our reasoning is simply that, logically, we feel that there simply *must* be some type of communication between the  $cc^*$  and  $ss^*$  states *before* the  $cc^*$  to  $ss^*$  transition takes place in order for the proper remnant to consistently remain to decay into leptons. Hence, we reason that  $f < 1$  describes the  $cc^*$  to  $ss^*$  transition. Statistically,  $f = (1 - q_s^2) = (8/9)$  is necessary to describe the hadronic width of the J. Since  $f$  is not appreciably different than 1, the leptonic width of the J, relative to the first order approximation, will be mitigated slightly. The second order FD for the J follows:

In Figure 6b  $f = (8/9)$  multiplies the entire array. Denoting the full width of the J in second order approximation by  $\Gamma_{J\text{-full-2}}$ , we find in accord with Figure 6b:

$$\begin{aligned} \Gamma_{J\text{-full-2}} \approx & (8/9)[(\alpha_s / 2\pi)(10,042)(2m_e)(m_p / m_j)^3(q_s)^4 + 2(\alpha \\ & / 2\pi)(10,042)(2m_e)(m_p / m_j)^3(q_s)^4 \\ & + (1/9)\{2(\alpha / 2\pi)(10,042)(2m_e)(m_p / m_j)^3(q_c)^4\}] \end{aligned} \quad (16d)$$

Thus,

$$\begin{aligned} \Gamma_{J\text{-full-2}} \approx & (8/9)[(\alpha_s / 2\pi)(10,042)(2m_e)(776/3097)^3(1/81) + \\ & 2(\alpha / 2\pi)(10,042)(2m_e)(776/3097)^3(1/81) + \\ & (1/9)\{2(\alpha / 2\pi)(10,042)(2m_e)(776/3097)^3(16/81)\}] \end{aligned}$$

Again, the value of the strong coupling parameter at the J mass is given by Eq. 16b, viz.,

$$\alpha_s = 1.2[\ln(3097/50)]^{-1} = 0.2908$$

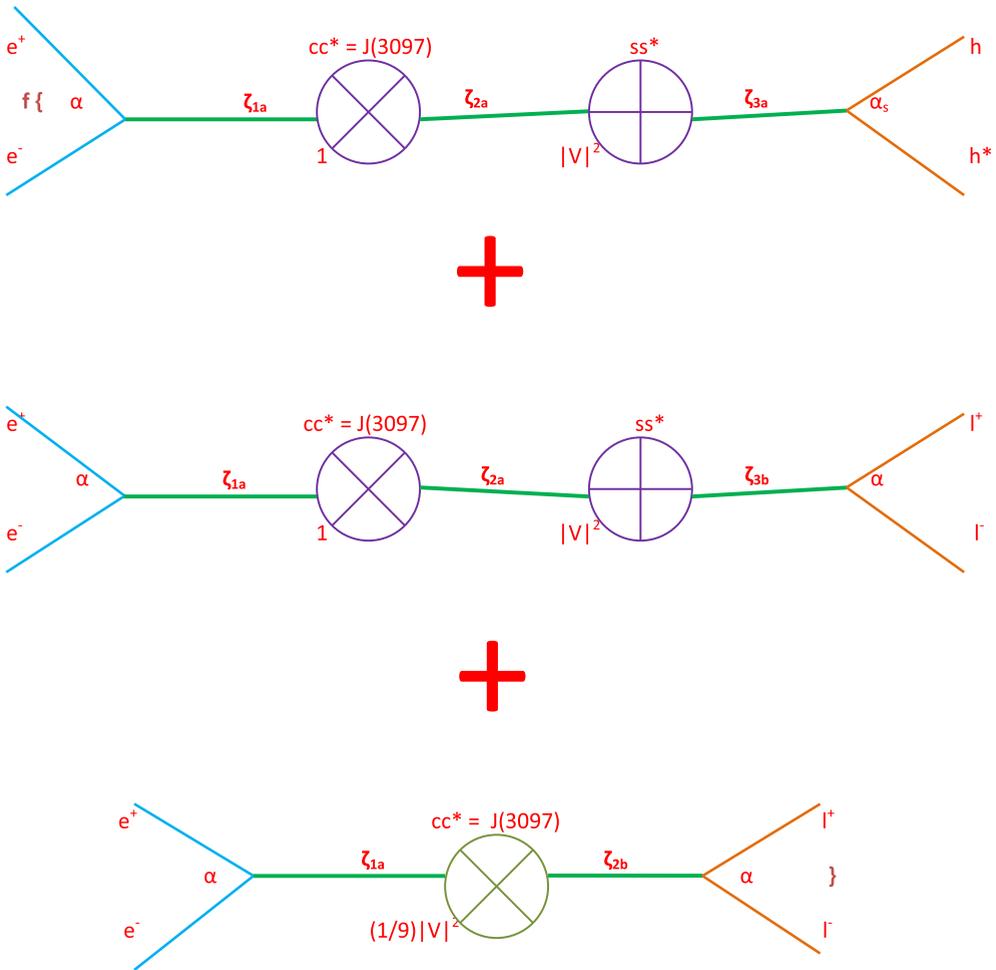


Fig. 6b. Feynman Diagram Array Characterizing the Formation and Decay of the J(3097) in Second Order Approximation via the GEM

Therefore,

$$\Gamma_{J\text{-full-2}} \approx (8/9)[92.2491 \text{ Kev} + 4.6298 \text{ Kev} + (1/9)\{74.0769 \text{ Kev}\}] \approx 93.43 \text{ Kev} \quad (16e)$$

The full width of the J under second order approximation is thus nearly an exact match to experiment (93.4 Kev via the GEM vs. 93.2 Kev from PDG (2009)). The hadronic width of the J is unchanged from first to second approximation; so, it remains a match with experiment (82.0 Kev via the GEM vs. 81.7 Kev from PDG (2009)). As well, the leptonic width of the J via the GEM (11.4 Kev) is now only 2.7% higher than that reported by the PDG ((11.1 ± 0.2) Kev).

**The Y(1S)**

Analogous to the J, the Y(1S), originally a bb\* construction, must transition to a cc\* excited state of the same mass as that of the bb\* state in order to decay into hadrons. Unlike the J, however, there is no reason to suspect that leptons emerge from the bb\* state. Hence, we assume that all types of Y(1S) decays ensue from the cc\* excited state. Corroborative evidence abounds in support of such assumption, as we shall see, so let us proceed with the viewing of the two FDs which depict the hadronic decay of the Y(1S) and the leptonic decay of the Y(1S), respectively:

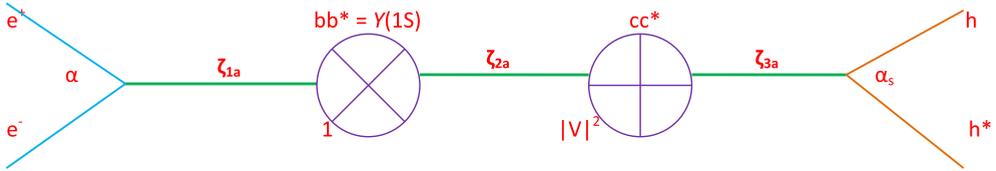


Fig. 7a. Basic Feynman Diagram for Y(1S) Formation and Decay into Hadrons via the GEM

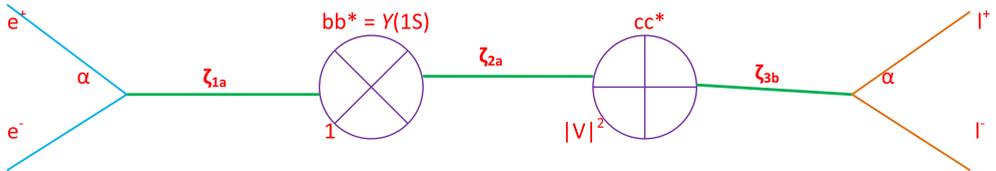


Fig. 7b. Basic Feynman Diagram for Y(1S) Formation and Decay into Leptons via the GEM

From Eq. 10 the hadronic width of the Y(1S), denoted by  $\Gamma_{Y-H}$ , via the GEM theoretical structure is given by:

$$\Gamma_{Y-H} \approx (\alpha_s / 2\pi)(10,042)(2m_e)(m_p / m_Y)^3(q_c)^4 \approx (\alpha_s / 2\pi)(10,042)(2m_e)(776 / 9460)^3(16 / 81) \quad (17a)$$

where

$$\alpha_s = 1.2[\ln(9460/50)]^{-1} = 0.2289 \quad (17b)$$

Hence,

$$\Gamma_{Y-H} \approx 40.76 \text{ Kev} \quad (17c)$$

The PDG in the 2008 Meson Table (PDG (2008), p.119) reports the corresponding figure as

$$\Gamma_{Y-H}(\text{PDG}) = 49.99 \text{ Kev} \quad (17d)$$

a figure 23% higher than the GEM-theoretical result.

However, if we look at the leptonic width of the Y(1S), denoted by  $\Gamma_{Y-L}$ , as derived via the GEM, we find from Eq. 11 (the right hand side of same multiplied by "3" to take into account muon and tauon pairs in accord with "e-μ-τ universality") that

$$\Gamma_{Y-L} \approx 3(\alpha / 2\pi)(10,042)(2m_e)(m_p / m_Y)^3(q_c)^4 \approx 3(\alpha / 2\pi)(10,042)(2m_e)(776 / 9460)^3(16 / 81) \quad (18a)$$

Hence,

$$\Gamma_{Y-L} \approx 3.90 \text{ Kev} \quad (18b)$$

which represents a match to the PDG's report from the same 2008 Meson Table of

$$\Gamma_{Y-L}(\text{PDG}) = (4.03 \pm 0.14) \text{ Kev} \quad (18c)$$

Specific to the  $e^+e^-$  partial width ( $\Gamma_{Y-ee}$ ), the GEM obviously determines  $\Gamma_{Y-ee} \approx 1.30 \text{ Kev}$ , while the PDG in the above-mentioned source (p. 119) states  $\Gamma_{Y-ee}(\text{PDG}_1) \approx 1.34 \text{ Kev}$  directly, but indirectly, in terms of its stated fractional branching ratio on p.119, a different value is inferred, viz.,  $\Gamma_{Y-ee}(\text{PDG}_2) \approx 1.29 \text{ Kev}$ . From the latter we infer that according to the PDG (2008), the experimentally determined value for the  $e^+e^-$  partial width of the  $Y(1S)$  is given by

$$\Gamma_{Y-ee}(\text{PDG}) = (1.31 \pm 0.03) \text{ Kev} \quad (18d)$$

a match to that of the GEM, i.e.,

$$\Gamma_{Y-ee} \approx 1.30 \text{ Kev} \quad (18e)$$

Herein (i.e., the match between Eq. 18d and Eq. 18e) lies the source of a paradox that the hadronic width as given by the GEM (i.e.,  $\sim 41 \text{ Kev}$ ) should be correct, though it is so seriously discrepant with that reported by the PDG (i.e.,  $50 \text{ Kev}$ ). The paradox unfolds as follows: In order to obtain the constant "B" in the general expression for  $\alpha_s$ , once  $\Lambda$  was determined, the assumption was made that, since the  $Y(1S)$  exists well into the realm of asymptotic freedom,

$$\alpha/\alpha_s = (e^+e^- \text{ partial width})/(\text{hadronic partial width})$$

as associated with the  $Y(1S)$ .

In Section 2 we inserted  $\Gamma_{Y-ee}(\text{PDG}) = 1.31 \text{ Kev}$  for the  $e^+e^-$  partial width, and for the hadronic partial width, we inserted the *GEM-theoretical width*, i.e.,  $\Gamma_{Y-H} \approx 41 \text{ Kev}$ . We then obtained the general relation,

$$\alpha_s = B[\ln(9460/50)]^{-1} = \alpha(41/1.31)$$

from which we solved for "B" to obtain,  $B = 1.2$

In turn, as "B" is a multiplier on the right hand sides of all width calculations via the GEM theory, and as all width calculations, as seen above, represent nearly exact matches with experiment in all cases except as to the hadronic width of the  $Y(1S)$ , it is difficult to fathom the source of the disparity between  $\Gamma_{Y-H} = 41 \text{ Kev}$  and  $\Gamma_{Y-H}(\text{PDG}) = 50 \text{ Kev}$ .

After a good number of years of pondering, it turns out that there is, actually, a very simple, and at the same time a very plausible solution to the paradox mentioned above, viz., we postulate an additional route for  $Y(1S)$  decay into hadrons, a route assumed not to have a high probability of occurrence for the J or the other vector mesons of mass less than that of the J. As the basis for the existence of the additional route available to the  $Y(1S)$ , we point to the fact that there is roughly three times the energy spectrum available to the  $Y(1S)$  in its decay (9460 Mev worth) as compared to the next lightest vector meson, i.e., the J (3097 Mev worth). With three times the energy spectrum (as compared to the J) available to the  $Y(1S)$ , we think it plausible that decays resulting in hadrons as products may be allowed to take place through the *bifurcation* of the gluon emitted from the resonance state (or more simply stated: via emission of two gluons), rather than what has heretofore been assumed in accord

with Figure 7 a, in which a single gluon,  $\zeta_{3a}$ , converts to hadrons to mark the final stage of the decay process. Specifically, we propose that, in addition to the route as described immediately above, a route exists in which  $\zeta_{3a}$  bifurcates into two gluons, each of which then converts to hadrons. The FD associated with the proposed additional route is seen immediately below.

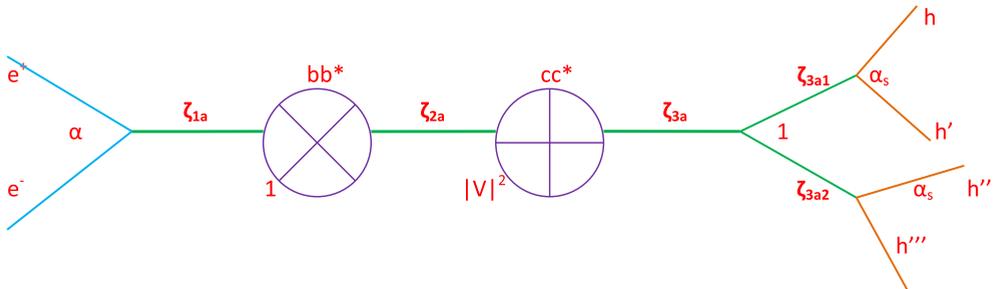


Fig. 8. Basic Feynman Diagram for Postulated Additional Route for Y(1S) Formation and Decay into Hadrons (h, h', h'', and h''') via the GEM

The additional route, which we denote as the “bifurcated gluon route for hadron decay” (BGRHD), effectively adds  $\alpha_s$  times  $\Gamma_{Y-H}$  or  $(0.2289)(40.76 \text{ Kev}) = 9.33 \text{ Kev}$  to the GEM-theoretical width of the Y(1S). The reformulated situation regarding the Y(1S) may be summarized, therefore, as follows:

Denoting the partial width due to the BGRHD as  $\Gamma_{Y-BGH}$ , we have

$$\Gamma_{Y-BGH} = 9.33 \text{ Kev} \tag{19}$$

From above we have

$$\Gamma_{Y-H} = 40.76 \text{ Kev}$$

Also from above we have

$$\Gamma_{Y-L} = 3.90 \text{ Kev}$$

The net hadronic width of the Y(1S) as per the GEM would now be given by

$$\Gamma_{Y-(H+BGH)}(\text{GEM 2010}) = 50.09 \text{ Kev} \tag{20}$$

which now represents a nearly perfect match to

$$\Gamma_{Y-H}(\text{PDG}) = 49.99 \text{ Kev}$$

In addition the full width of the Y(1S) as per the GEM would now be given by

$$\Gamma_{Y-\text{full}}(\text{GEM 2010}) = 53.99 \text{ Kev} \tag{21}$$

which also represents a nearly perfect match to  $\Gamma_{Y-\text{full}}(\text{PDG}) = (54.02 \pm 1.25) \text{ Kev}$

With the addition of the BGRHD the calculation of “B” in the expression for  $\alpha_s$  is uncompromised, while at the same time the major discrepancy between the hadronic width of the Y(1S) as determined via the GEM versus via the methods engaged by the PDG is completely removed. For that reason we believe the postulate as to the addition of the

BGRHD is a viable one. In fact, if we postulate that in addition to the BGRHD there is a companion route for leptons, i.e., a bifurcated gluon route for lepton decay (BGRLD), whose FD is identical to that of Figure 9, except that on the far right hand side of the diagram, each “ $\alpha_s$ ” is replaced by “ $\alpha$ ” and “h, h', h'', and h'''” are replaced by “ $l_i^+, l_i^-, l_j^+,$  and  $l_j^-$ ”, respectively, where “i” and “j” denote lepton types and  $i = j$  is allowed, (3.90 Kev / 137.036) = 0.03 Kev would be added to

$$\Gamma_{Y\text{-full}}(\text{GEM 2010})$$

above, thus bringing

$$\Gamma_{Y\text{-full}}(\text{GEM 2010}) \rightarrow 54.02 \text{ Kev} \tag{22}$$

i.e., the realization of an exact match to experiment.

### 4. Speculations based upon the GEM

#### The T-Meson

To address the T-meson, thought to be a  $tt^*$  (where “t” represents the top quark) state of mass approximately 340000 Mev, but never “discovered” to date, is quite speculative on our part, but we think it important to do so because the GEM provides a perfectly logical reason as to why the T has yet to be “found”, i.e., unequivocally shown to exist by experiment. Said reason is just the opposite of the prevailing view as to the “invisibility” of the T, which is: “the T doesn’t last long enough for it to be found.” In a sense such is true; after all, the  $bb^*$  of the  $Y(1S)$  transitions instantaneously to a  $cc^*$  state according to the GEM, but the mass of the original  $bb^*$  state is preserved in the resulting  $cc^*$  state, thus allowing for the “finding” of a resonance at the  $Y(1S)$  mass. Assuming the T to act in like manner to the  $Y(1S)$ , the following FD would apply as regards hadron production:

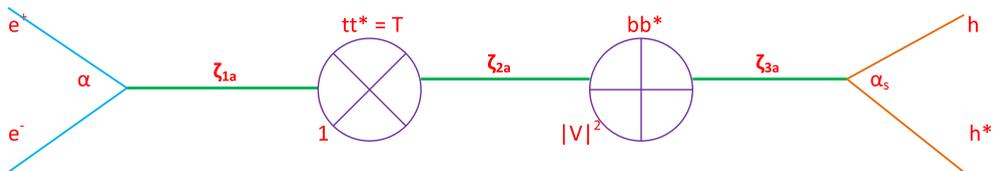


Fig. 9. Basic Feynman Diagram for T Formation and Decay into Hadrons via the GEM

The hadronic width of the T, from Eq. 4 would be:

$$\begin{aligned} \Gamma_T &\approx (\alpha_s / 2\pi)(10,042)(2m_e)(m_p / m_T)^3(q_b)^4 \\ &\approx (\alpha_s / 2\pi)(10,042)(2m_e)(776 / 340000)^3(1/81) \end{aligned} \tag{23a}$$

where

$$\alpha_s = 0.90[\ln(340000/50)]^{-1} = 0.1020 \tag{23b}$$

(In Eq. 23b the constant “1.2” in the expression for  $\alpha_s$  becomes<sup>11</sup> “0.90” beyond 100000 Mev, and in Eq. 23a  $q_b = -1/3$ .)

Hence,

$$\Gamma_T \approx 0.024 \text{ ev} \tag{23c}$$

Thus, we see that, contrary to the “convenient explanation” as to why the T has not so far been observed, the T lives for a *very long time* (about 6 ps)! It’s just that its width to mass ratio makes it impossible right now for the experimental apparatus to pick up such a narrow signal amongst the “noise” inherent in the energy background needed to produce the T.

**Color-by-Color Disengagement from Lepton Production**

We must go a bit beyond the scope of the material thus far presented in order to discuss the phenomenon of color-by-color disengagement from lepton production, as we must now make reference to the excited states of the  $\Psi$ -series and  $Y$ -series mesons, designated respectively as  $\Psi(NS)$  and  $Y(NS)$ , where  $N > 1$  designates an excited state. It turns out that as regards lepton decay, partial widths of the above objects, when  $N = 2$ , the associated form factors ( $f_i$ ) which provide for excellent agreement with experiment are as follows<sup>12</sup>:

For the  $\Psi(2S)$   $f_1 = 1 - q_s^2 = 8/9$  (as was found above for the  $J = \Psi(1S)$ ), and for the  $Y(2S)$   $f_2 = 1 - q_c^2 = 5/9$  (analogous to the  $\Psi(2S)$ ). We will designate the form factor for the  $Y(1S)$  (as per above) as  $f_3 = 1$ . What is highly interesting, as it turns out, is that in terms of the above form factors, the leptonic partial widths of many of the  $\Psi(NS)$  and  $Y(NS)$  states are excellently described by the GEM if the latitude exists to multiply the resulting GEM-width formulas involving the appropriate form factors by “ $(n/6)$ ”, where “ $n$ ” is an integer, the interpretation being that as  $N$  increases, the number of quark colors participating in lepton decay decreases ... effectively by “half colors” at a time. We illustrate the quark color disengagement phenomenon by reproducing the results from the reference<sup>12</sup> associated with footnote # 12 below:

Meson	Mass (Mev)	$\Gamma_{ee}(\text{GEM})$	$\Gamma_{ee}(\text{PDG})$	# of Colors Operative
$\Psi(1S)$	3097	5.72	$5.55 \pm 0.16$	3
$\Psi(2S)$	3686	2.26	$2.36 \pm 0.04$	2
$\Psi(3S)$	4039	0.86	$0.86 \pm 0.07$	1
$\Psi(4S)$	4153	0.79	$0.83 \pm 0.07$	1
$\Psi(5S)$	4421	0.65	$0.58 \pm 0.07$	1

Chart 1. Color Participation in Lepton Production in the  $\Psi$ -Series

In Chart 1 all electron/positron partial widths are expressed in Kev. Note that a near-match with experiment occurs for the  $J(3097)$ , as per Section C above, assuming three colors are operative in electron/positron decay ( $n = 6$ ). A near-match with experiment results for the  $\Psi(2S)$  assuming two colors are operative in said decay ( $n = 4$ ), and statistical matches with experiment are evident if we assume only one color participates in electron/positron decay ( $n = 2$ ) as associated with the  $\Psi(3S)$ ,  $\Psi(4S)$ , and  $\Psi(5S)$ . As to the  $\Psi(NS)$  objects, then, the GEM provides an excellent match to experiment (from the PDG’s 2009 “Meson Table”) if we assume that sequentially progressive disengagement ... color-by-color ... of quark colors manifests in lepton decay.

The situation is similar to, but slightly different than the above, as regards the  $Y(NS)$  series, as seen in Chart 2 below.

Again, all partial widths are expressed in Kev, and, again, we see excellent agreement with experiment if we assume  $n = 6$  associated with the  $Y(2S)$  (three colors operative),  $n = 5$  associated with the  $Y(3S)$  (either “ $2\frac{1}{2}$  colors” operative or an even mix of three colors and two colors taking part in the decay),  $n = 3$  associated with the  $Y(4S)$  (either “ $1\frac{1}{2}$  colors”

Meson	Mass (Mev)	$\Gamma_{ee}(\text{GEM})$	$\Gamma_{ee}(\text{PDG})$	# of Colors Operative
Y(2S)	10023	0.624	$0.612 \pm 0.11$	3
Y(3S)	10355	0.471	$0.443 \pm 0.008$	2½
Y(4S)	10579	0.266	$0.272 \pm 0.029$	1½
Y(5S)	10860	0.33	$0.31 \pm 0.07$	2
Y(6S)	11019	0.157	$0.13 \pm 0.03$	1

Chart 2. Color Participation in Lepton Production in the Y-Series

operative or an even mix of two colors and one color taking part in the decay),  $n = 4$  associated with the Y(5S) (two colors operative), and  $n = 2$  associated with the Y(6S) (one color operative). At the present juncture, perhaps the best guess as how to handle the results associated with the Y(3S) and Y(4S) would be to speculate that the GEM description should be due to a nearly even mix of two and three contributing colors in one case, and a nearly even mix of one and two contributing colors in the other case, rather than making the claim that the GEM has shown that the quark colors become fragmented beyond 10 000 Mev, especially when one notes that all results associated with integer color contributions in Chart 2 represent statistical matches with experiment. Note, as well, that the progressive color disengagement behavior with increasing mass is preserved if only integer color contributions are considered.

**Vector Mesons as Vacuum Excitations**

One basic reality demonstrated by the GEM is the prime role that the electromagnetic interaction plays in vector meson formation and decay. The spin-flip responsible for all spin one mesons takes place via the electromagnetic interaction. The transitioning of the four-momentum from  $cc^*$  states to  $ss^*$  in the  $\psi(\text{NS})$  decays, as evidenced via the mathematical form of the form factor,  $f_1$ , is seen to take place via the electromagnetic interaction, as is the case regarding the analogous transition from  $bb^*$  to  $cc^*$  in the Y(NS) decays. A second reality demonstrated by the GEM represents a radical departure from current assumptions about the structures of vector mesons in general ... but especially about the structure of the  $K^*$  ... viz., that all vector mesons are represented by

$$\chi_v = (1/\sqrt{n})[\sum_{i=1}^n (Q_i Q_i^*)] \tag{24}$$

where  $Q_i$  represents a quark of flavor, "i",  $Q_i^*$  represents the associated anti-quark, and "n" represents the number of flavors operative at the energy scale of the relevant meson's rest energy. Thus, for example, the neutral  $K^*$  does *not* comprise a  $ds^*$  or an  $sd^*$ , and the charged  $K^*$  does *not* comprise a  $us^*$  or an  $su^*$ . Rather, there is a general " $K^*$  construction" given by Eq. 14, the decay of which features a "favored energy" of 892 Mev resulting in a net charge of  $\pm 1$  amongst its decay products ... *and* another "favored energy" at 896 Mev resulting in no net charge amongst its decay products. In the literature similar considerations apply to the various  $D^*$  and  $B^*$  states<sup>13</sup>. In other words the GEM illustrates that vector mesons are not actually "unstable particles" which form at collision sites, but rather are manifestations of a "quark sea" as part of the construction of what we call "the vacuum" ... much analogous to Dirac's idea of the "electron sea" of old. Just as the electrons in the Dirac Sea were thought to be excited via the electromagnetic interaction, we see from the above that one may surely think of the formation of vector mesons as an electromagnetic excitation of relevant quarks in a "quark sea" ... where a given quark is promoted to a positive energy state by a virtual gluon, which, unlike in electron/positron production, where the electron flies away from its

“vacuum hole”, produces the circumstance that a tightly bound quark/anti-quark pair becomes virtually extant, in which vicinity the spin-flip of one of the quarks occurs, thus producing what we call “a vector meson”. Whether one thinks of the vacuum in accord with Brian Greene as “the fabric of the cosmos”, or, more conventionally, in accord with the pioneers of Quantum Electrodynamics (QED), as “a sea of leptons” (i.e., electrons, muons, taus, and six flavors of quarks), the vacuum is certainly a “something” as opposed to “nothing at all”. If one treats vector mesons as vacuum excitations, i.e., “an excitation of the sea of leptons” and follows the rules laid down by the pioneers of QED (i.e., the calculation of the relevant Feynman Diagrams) with the one exception of replacing  $\alpha$  by  $\alpha_s$  (it derived solely via the GEM) in the hadronic partial width calculations, one acquires fine agreement with experiment. Hence, we believe, vector mesons thought of as vacuum excitations makes perfect sense.

## 5. Summary

The Gluon Emission Model has been shown to serve very nicely as a basis for calculations of not only the widths of the  $\rho$  meson, the  $\phi$  meson, the  $K^*(892)$ , the  $J$  meson, and the  $Y$  meson, but also for the determination of the strong coupling parameter,  $\alpha_s$ , over the entire range of energy over which the above objects exist. We have seen that the GEM has built into its framework two precepts of prime importance for the carrying out of the above types of calculations: (1) the specification of a quark spin-flip matrix element as the central determinant of a vector meson resonance and (2) the virtual photon and the gluon as two aspects of the same entity, viz., the four-momentum propagator. The prime significance of (1) is that the square of the quark spin-flip matrix elements in vector meson width calculations are proportional to  $q_i^4$ , where  $q_i$  represents the magnitude of the charge of quark type “ $i$ ”. The significance of (2) is that the virtual photon and the gluon essentially obtain their identities from what the vertices of origin and termination are in the relevant Feynman Diagram. The ramifications of (1) are that, as  $(2/3)^4$  is 16 times  $(1/3)^4$ , it is quite easy to determine that the  $cc^*$  (charm – anti-charm) structure of the  $J(3097)$  must transmute to an  $ss^*$  (strange – anti-strange) in nearly a point-like manner, such that it is the  $ss^*$  structure that undergoes the spin-flip at the  $J(3097)$  resonance. Similarly, the  $Y(1S)$  must transmute in point-like manner from its original  $bb^*$  (bottom – anti-bottom) structure to a  $cc^*$  structure before decaying. The ramification of (2) is that the leptonic width to hadronic width ratio associated with the same basic decaying structure must be in the ratio of  $\alpha$  to  $\alpha_s$ . We saw that the GEM predicts the hadronic width of the  $Y(1S)$  to be  $\sim 41$  Kev, assuming that the  $Y(1S)$ , as lower energy mesons do, decays solely via the emission of a single gluon, whereas the figure for same as stated in the 2008 Meson Table from the Particle Data Group (PDG) is  $\sim 50$  Kev. The discrepancy noted above (23%) is seen to be extremely important, because, if we were to assume that the GEM was in error by such amount, it turns out that all other GEM calculations, currently essentially exactly on the mark as to the  $\rho$ , the  $\phi$ , the  $K^*(892)$ , the  $J$ , and  $\alpha_s$ , would have to be rendered as 23% too large by bringing the GEM’s determination of the  $Y(1S)$  in line with the PDG’s determination of same through adjustment of the GEM’s determination of  $\alpha_s$ . Hence, in order to make the GEM as currently constructed fit the PDG as to the hadronic width of the  $Y(1S)$ , all other GEM calculations would be discrepant by the same amount, i.e., 23%, at each diverse point of the energy spectrum where the GEM has been successfully applied. Clearly, then, what needed to be addressed are the details in the GEM’s determination of the width of the  $Y(1S)$ , with an eye

towards any reasonable modifications that might remove the above-mentioned disparity, with the satisfactory result being the assumption of an additional decay route for the  $Y(1S)$ , i.e., the bifurcated gluon route.

Unlike the theoretical structures prevalent in the literature that one encounters as to determining the width of the vector mesons, the GEM theory is about as simple as it gets: One fundamental process is posited for the formation and decay of any spin one meson, i.e., a quark spin-flip; the gluon absorption cross-section for said process is then integrated over energy, and from there, the Feynman Diagram resulting in hadron or lepton pairs is then calculated.

Form factors associated with the  $\psi(1S)$  and  $Y(2S)$ , calculated directly from relevant experimental data, are given by  $f_1 = (1-q_s^2) = (8/9)$  in the case of the  $\psi$ -series mesons and  $f_2 = (1-q_c^2) = (5/9)$  in the case of the  $Y$ -series mesons, where  $q_s = -1/3$  represents the charge of the strange quark and  $q_c = 2/3$  represents the charge of the charm quark. The form factors represent the fraction of the originally produced quark/anti-quark ( $QQ^*$ ) state which makes a transition to a  $QQ^*$  state of the next lowest mass ...  $ss^*$  in the case of the  $\psi$ -series mesons and  $cc^*$  in the case of the  $Y$ -series mesons ... and thus figure prominently into the calculation of the hadronic and leptonic widths of a given meson via the constructs of the Gluon Emission Model. We have seen that  $f_1 = (8/9)$  is representative of all  $\psi$ -states, if, and only if, it is assumed that one quark color (in the case of the  $\psi(2S)$ ) or two quark colors (in all other cases) become disengaged from lepton production. A similar set of circumstances is observed as to the  $Y$ -series mesons, such illustrating that all three quark colors are functional in lepton production in  $Y(2S)$  decay, fewer than three functional in  $Y(3S)$  and  $Y(4S)$  decay, with likely only one color functioning in  $Y(5S)$  and  $Y(6S)$  decay. For each meson series, then, lepton decay is characterized by the phenomenon of sequential disengagement of quark color from lepton production as a function of increasing mass.

Finally, we have seen that the GEM suggests, contrary to a rigid interpretation of the Standard Model, in which vector mesons are treated as unstable particles, that vector mesons are quite realizable as electromagnetic vacuum excitations of a constituent "quark sea", analogous to the "Dirac Sea" of electrons of old. Specifically, the GEM construct yields agreement with experiment only if it is assumed that vector mesons are represented as linear combinations of quark spin-flip excitation possibilities. The  $K^*(892)$  is a case in point, but, further, there appears to be no hope for reliable width calculations of any  $\psi$ -series mesons if such elements of said series are represented solely as  $cc^*$  objects. A like statement, of course, holds for the elements of the  $Y$ -series as immutably  $bb^*$  "particles".

## 6. References

- [1] F. E. Close (1979), *An Introduction to Quarks and Partons*, Academic Press.
- [2] R. H. Dalitz (1977), "Glossary for New Particles and New Quantum Numbers", *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences*, Vol. 355 (1983), pp. 601 & ff.
- [3] S. Gasiorowicz and J. L. Rosner, *American Journal of Physics* 49, pp. 954 & ff. (1981).
- [4] Brian Greene (2004), *The Fabric of the Cosmos*, Knopf.
- [5] E. Merzbacher (1970), *Quantum Mechanics*, Wiley.
- [6] M. Schmelling (1996), [arxiv.org/abs/hep-ex/9701002](http://arxiv.org/abs/hep-ex/9701002), "Status of the Strong Coupling Constant", *Plenary Talk given at the XXVIII International Conference on High Energy Physics*, Warsaw, Poland, July, 1996.

- [7] PDG (2004), pdg.lbl.gov, "Meson Table".
- [8] PDG (2008), pdg.lbl.gov, "Meson Table".
- [9] PDG (2009), pdg.lbl.gov, "Meson Table".
- [10] D. White (2008), "The Gluon Emission Model for Hadron Production Revisited", *Journal of Interdisciplinary Mathematics*, Vol. 11, No. 4, pp. 543 - 551.
- [11] D. White (2008), "GEM and the  $K^*(892)$ ", *Journal of Applied Global Research*, Vol. 1 (3), pp. 1 - 4.
- [12] D. White (2009), "GEM and the Leptonic Width of the  $J(3097)$ ", *Journal of Applied Global Research*, Vol. 2 (4), pp. 1 - 5.
- [13] D. White (2009), "Mathematical Modeling of the Strong Coupling Parameter Based upon the Gluon Emission Model for Hadron Production Associated with Vector Meson Decay", *Journal of Interdisciplinary Mathematics*, Vol. 12, No. 6, pp 825-838.
- [14] D. White (2010), " Evidence for Color-by-Color Disengagement from the Process of Lepton Production Associated with the  $\psi$ -Series and  $Y$ -Series Mesons", *Communications in Mathematics and Applications*, Vol. 1, No. 3, pp. 183 - 193.
- [15] D. White (2010), "GEM and the  $D^*$  &  $D_s^*$  Mesons", *Global Journal of Mathematical Sciences*, Vol. 2, No. 3, pp. 159 - 166.
- [16] D. White (2010), "GEM and the  $Y(1S)$ ", *The Journal of Informatics and Mathematical Sciences*, Vol. 2, Nos. 2 & 3, pp. 71 - 93.

# Vector Correlations in Collision of Atom and Diatomic Molecule

Xian-Fang Yue  
Jining University  
China

## 1. Introduction

The use of crossed molecule beam (CMB) and polarized laser techniques has allowed the precisely measurements of the vector correlations in atom and diatomic molecule collision reactions. There are a variety of experimental studies aiming at exploring the vector correlations underlying a collision reaction.[1-15] To name a few are those of Zare and co-workers, [1-3] those of Herschbach and co-workers [4-5], those of Fano and co-workers [6], as well as those of Han and co-workers [7-9]. These experimental investigations impelled the development of theoretical methods to simulate the measured profiles and to explore the unmeasured ones. Generally, theoretical explorations on vector correlations of collision reactions fall into two categories: one is the reagent  $k$ - $j$ ,  $j$ - $k$ - $k'$  vector correlations, and the other is the product  $k$ - $j'$ ,  $k$ - $k'$ - $j'$  vector correlations. Where  $k/k'$  and  $j/j'$  denotes the relative velocity of reagent/product and reagent/product rotational angular momentum in the centre of mass (CM) frame. As reported in many previous papers, both of the quantum scattering and the quasi-classical trajectory (QCT) calculations can be used to investigate the vector correlations. In this chapter, we report the product  $k$ - $j'$ ,  $k$ - $k'$ - $j'$  vector correlations of the  $\text{N}(^2\text{D}) + \text{H}_2(\nu, j) \rightarrow \text{NH}(\nu', j') + \text{H}$  collision reactions with the QCT method.

The  $\text{N}(^2\text{D}) + \text{H}_2 \rightarrow \text{NH} + \text{H}$  collision reaction plays an important role in the chemistry of nitrogen containing fuels and of nitrogen in the atmosphere [16]. It has attracted many investigations from not only experimental viewpoints [17-25], but also theoretical viewpoints [26-37]. Experimentally, Suzuki et al. [17] measured the rate constants for this reaction by employing a pulse radiolysis-resonance absorption technique at temperatures between 213 and 300 K. They found that the temperature dependence of the rate constants exhibits an Arrhenius behavior. Umemoto and coworkers [18-20] measured the vibrational and rotational state distributions of the nascent NH and ND molecules formed in the  $\text{N}(^2\text{D}) + \text{H}_2$  and  $\text{N}(^2\text{D}) + \text{D}_2$  reactions. In their experiments, the  $\text{N}(^2\text{D})$  atoms were generated by two-photon dissociation of NO, while the nascent NH and ND molecules were detected by laser-induced fluorescence (LIF) technique. They found that the nascent vibrational distributions have  $\text{NH}(\nu''=1)/\text{NH}(\nu''=0) = 0.8 \pm 0.1$  and  $\text{ND}(\nu''=1)/\text{ND}(\nu''=0) = 1.0 \pm 0.1$ , and that the rotational populations of these vibrational states are broad and hot. More recently, Casavecchia et al. [21-25] carried out a series of experimental and theoretical studies on the  $\text{N}(^2\text{D}) + \text{H}_2$  and  $\text{N}(^2\text{D}) + \text{D}_2$  reactions. They have measured the angular and velocity distributions of the nascent NH and ND products in the title reactions under CMB experiments with mass spectrometric detection technique.

Theoretically, Honvault and Launay [26] carried out the first accurate quantum mechanics (QM) calculations of the  $N(^2D) + H_2$  reaction on the Pederson et al. PES [27] with the hyperspherical method. They calculated the vibrationally and rotationally state resolved integral cross sections at collision energies of 1.6, 2.5, and 3.8 kcal/mol. A forward-backward symmetry was also found in their differential cross sections (DCS) calculations, which indicates that a complex of  $NH_2$  is formed in the reaction. In a recent work, Ho et al. [28] improved and refined the Pederson et al. PES [27], and developed a more accurate adiabatic ground  $1^2A''$  PES. The new PES is fitted by a set of 2715 ab initio points resulting from the multireference configuration interaction (MRCI) calculations and a better algorithm. Moreover, the new PES is free of spurious small scale features and is in much better agreement with the ab initio calculations, especially in key stationary point regions including the  $C_{2v}$  minimum, the  $C_{2v}$  transition state, and the N-H-H linear barrier. Employing this new PES, both QM [29-32] and QCT [28, 31-33] calculations have been performed for the title reactions.

Most of previous studies focused on the calculations of the reaction probabilities, integral cross sections and thermal rate constants. Very few investigations paid attention to the vector properties calculations for the  $N(^2D) + H_2$  reactions. In this chapter, we present a much detailed calculations on the vector correlations of this collision reaction.

## 2. Methodology and theory

### 2.1 Quasi-classical trajectory calculations

The QCT calculation method is the same as that in Refs. [38-42]. Six-order symplectic integration was used in the QCT calculations, in which the classical Hamilton's equations are integrated numerically for motion in three dimensions. In a given Hamiltonian system whose Hamiltonian can be partitioned, i.e., written as  $H = T(p) + U(q)$ , we define the derivative terms in the Hamiltonian movement equations as  $hq_i = \partial H / \partial p_i$  for the generalized momentum, and  $hp_i = \partial H / \partial q_i$  for the generalized coordinates; let  $dt$  be the full time step. The algorithm can be expressed as [42]

$$\begin{aligned}
 & \text{do } i = 0, 1, 2, \dots, n-2, n-1 \\
 & \quad p = p + dta(2i)hp \\
 & \quad q = q + dta(2i+1)hq \\
 & \quad \text{enddo} \\
 & p = p + dta(2n)hp. \tag{1}
 \end{aligned}$$

The last step can be concatenated with the first step in a continuing calculation. Then there are  $2n$  substeps in every step, where  $n$  is 8 for six-order symplectic integrator. Thus, the number of the calling to the potential derivatives is also  $n$ . The coefficients are defined by  $2n + 1$  values of  $a(i)$ , which is taken directly from Ref. [43-44] without revision.

The accurate  $1^2A''$  state PES constructed recently by Ho et al. [28] is employed for the  $N(^2D) + H_2 \rightarrow NH + H$  calculations. The full potential is written in terms of a many body expansion:

$$V(R_{\text{NH}_a}, R_{\text{NH}_b}, R_{\text{HH}}) = V^{(1)} + V_{\text{NH}}^{(2)}(R_{\text{NH}_a}) + V_{\text{NH}}^{(2)}(R_{\text{NH}_b}) + V_{\text{HH}}^{(2)}(R_{\text{HH}}) + V^{(3)}(R_{\text{NH}_a}, R_{\text{NH}_b}, R_{\text{HH}}), \quad (2)$$

Where the one-body term  $V^{(1)}$  defines the zero of energy as the three-atom limit and the two-body terms  $V_{\text{NH}}^{(2)}$  and  $V_{\text{HH}}^{(2)}$  represent the energy of H + NH and N + H<sub>2</sub>, respectively. The two-body functions are interpolated from separate calculations of the two-body potentials using the reciprocal power radial kernel. The three-body function  $V^{(3)}(R_{\text{NH}_a}, R_{\text{NH}_b}, R_{\text{HH}})$  in the coordinate system is interpolated from the differences between the ab initio data and the corresponding one- and two-body sums. During reactive encounter, the total angular momentum is conserved

$$L + j = L' + j', \quad (3)$$

where  $L$  and  $L'$  are the reagent and product orbital momenta, respectively. In the mass weighted coordinates, the rotational angular momentum (AM) of reagent molecule H<sub>2</sub> is

$$j = \mu_{\text{HH}} \mathbf{R}_{\text{HH}} \times \dot{\mathbf{R}}_{\text{HH}} = \sqrt{j(j+1)}\hbar, \quad (4)$$

and reagent orbital AM is

$$l = \mu_r \mathbf{R}_{\text{N-HH}} \times \dot{\mathbf{R}}_{\text{N-HH}} = \sqrt{l(l+1)}\hbar. \quad (5)$$

Accordingly, the product rotational AM is

$$j' = \mu_{\text{NH}} \mathbf{R}_{\text{NH}} \times \dot{\mathbf{R}}_{\text{NH}} = \sqrt{j'(j'+1)}\hbar, \quad (6)$$

and the product orbital AM is

$$l' = \mu_p \mathbf{R}_{\text{NH-H}} \times \dot{\mathbf{R}}_{\text{NH-H}} = \sqrt{l'(l'+1)}\hbar. \quad (7)$$

Where  $\mu_r$  and  $\mu_p$  are reagent and product reduced masses, respectively. The Hamilton's motion equation is solved by the symplectic integration method. The accuracy of the numerical integration is verified by checking the conservations of the total energy and the total angular momentum for every trajectory. In the calculation, batches of  $1 \times 10^5$  trajectories are run for each reaction and the integration step size is chosen to be 0.1 femtosecond (fs). The trajectories start at an initial distance of 15 Å between the N atom and the center of the mass of the H<sub>2</sub> molecules. The impact parameter  $b$  is optimized through the repeated computation with  $1 \times 10^5$  trajectories for all of the present studies. The maximum impact parameter  $b_{\text{max}}$  was obtained when there is no reactive trajectory any more, if a tiny larger value of  $b$  is done. All the optimized maximum impact parameters used for the present study are displayed in Table 1.

## 2.2 Polarization-dependent differential cross-sections (PDDCSs)

The CM frame was used as the reference frame in the present work, which is depicted in Figure 1. The reagent relative velocity vector  $k$  is parallel to the z-axis. The x-z plane is the

scattering plane which contains the initial and the final relative velocity vectors,  $k$  and  $k'$ .  $\theta_t$  is the angle between the reagent relative velocity and product relative velocity (so-called scattering angle).  $\theta_r$  and  $\phi_r$  are the polar and azimuthal angles of the final rotational angular momentum  $j'$ .

The distribution function  $P(\theta_r)$  describing the  $k$ - $j'$  correlation can be expanded in a series of Legendre polynomials as [38-41]

$$P(\theta_r) = \frac{1}{2} \sum_k [k] a_0^k P_k(\cos \theta_r), \quad (8)$$

where  $[k] = 2k + 1$ . The  $a_0^k$  coefficients are given by

$$a_0^k = \langle P_k(\cos \theta_r) \rangle. \quad (9)$$

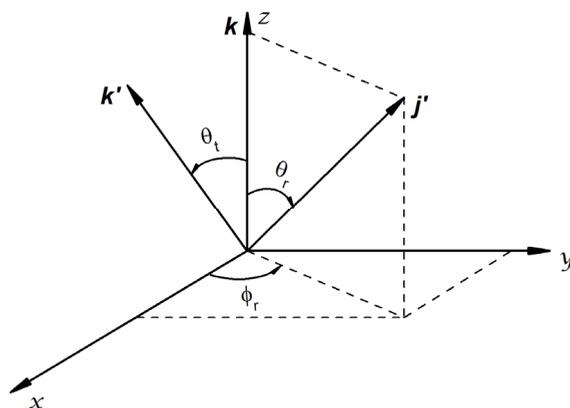


Fig. 1. Center-of-mass coordinate system used to describe the  $k$ ,  $k'$  and  $j'$  correlations.

The expanding coefficients  $a_0^k$  are called orientation ( $k$  is odd) and alignment ( $k$  is even) parameters.

The dihedral angle distribution function  $P(\phi_r)$  describing  $k$ - $k'$ - $j'$  correlation can be expanded in Fourier series as [38-41]

$$P(\phi_r) = \frac{1}{2\pi} \left( 1 + \sum_{\text{even}, n \geq 2} a_n \cos n\phi_r + \sum_{\text{odd}, n \geq 1} b_n \sin n\phi_r \right), \quad (10)$$

where

$$a_n = 2 \langle \cos n\phi_r \rangle, \quad (11)$$

$$b_n = 2 \langle \sin n\phi_r \rangle. \quad (12)$$

In this calculation,  $P(\phi_r)$  is expanded up to  $n = 24$ , which shows good convergence. The joint probability density function of angles  $\theta_r$  and  $\phi_r$ , which determine the direction of  $j'$ , can be written as

$$\begin{aligned}
 P(\theta_r, \phi_r) &= \frac{1}{4\pi} \sum_{kq} [k] a_q^k C_{kq}(\theta_r, \phi_r)^* \\
 &= \frac{1}{4\pi} \sum_k \sum_{q \geq 0} [a_q^k \cos q\phi_r - a_{q\mp}^k i \sin q\phi_r] C_{kq}(\theta_r, 0). \quad (13)
 \end{aligned}$$

this calculation, the polarization parameter is evaluated as

$$a_{q\pm}^k = 2 \langle C_{k|q|}(\theta_r, 0) \cos q\phi_r \rangle, \quad k \text{ is even}, \quad (14)$$

$$a_{q\pm}^k = 2i \langle C_{k|q|}(\theta_r, 0) \sin q\phi_r \rangle, \quad k \text{ is odd}. \quad (15)$$

In the calculation,  $P(\theta_r, \phi_r)$  is expanded up to  $k = 7$ , which is sufficient for good convergence.

The full three-dimensional angular distribution associated with  $k-k'-j'$  correlation can be represented by a set of generalized polarization-dependent differential cross-sections (PDDCSs) in the CM frame. The fully correlated CM angular distribution is written as [38-41]

$$P(\omega_t, \omega_r) = \sum_{kq} \frac{[k]}{4\pi} \frac{1}{\sigma} \frac{d\sigma_{kq}}{d\omega_t} C_{kq}(\theta_r, \phi_r)^*, \quad (16)$$

Where the angles  $\omega_t = \theta_t, \phi_t$  and  $\omega_r = \theta_r, \phi_r$ .  $\sigma$  is the integral cross section.  $C_{kq}(\theta_r, \phi_r)$  are

modified spherical harmonics.  $\frac{1}{\sigma} \frac{d\sigma_{kq}}{d\omega_t}$  is a generalized PDDCS.  $C_{kq}(\theta_r, \phi_r)$  are modified

spherical harmonics.  $\frac{1}{\sigma} \frac{d\sigma_{kq}}{d\omega_t}$  is a generalized PDDCS, and  $\frac{1}{\sigma} \frac{d\sigma_{kq}}{d\omega_t}$  yields

$$\frac{1}{\sigma} \frac{d\sigma_{k0}}{d\omega_t} = 0 \quad (k \text{ is odd}), \quad (17)$$

$$\frac{1}{\sigma} \frac{d\sigma_{kq+}}{d\omega_t} = \frac{1}{\sigma} \frac{d\sigma_{kq}}{d\omega_t} + \frac{1}{\sigma} \frac{d\sigma_{k-q}}{d\omega_t} = 0 \quad (k \text{ even, } q \text{ odd or } k \text{ odd, } q \text{ even}), \quad (18)$$

$$\frac{1}{\sigma} \frac{d\sigma_{kq-}}{d\omega_t} = \frac{1}{\sigma} \frac{d\sigma_{kq}}{d\omega_t} - \frac{1}{\sigma} \frac{d\sigma_{k-q}}{d\omega_t} = 0 \quad (k \text{ even, } q \text{ even or } k \text{ odd, } q \text{ odd}). \quad (19)$$

The PDDCS is written in the following form:

$$\frac{1}{\sigma} \frac{d\sigma_{kq\pm}}{d\omega_t} = \sum_{k_1} \frac{[k_1]}{4\pi} S_{kq\pm}^{k_1} C_{k_1-q}(\theta_t, 0), \quad (20)$$

where the  $S_{kq\pm}^{k_1}$  is evaluated by using the expected value expression to be

$$S_{kq\pm}^{k_1} = \langle C_{k_1q}(\theta_t, 0) C_{kq}(\theta_r, 0) [(-1)^q e^{iq\phi_r} \pm e^{-iq\phi_r}] \rangle, \quad (21)$$

where the angular brackets represent an average over all angles. The differential cross-section is given by

$$\frac{1}{\sigma} \frac{d\sigma_{00}}{d\omega_t} = P(\omega_t) = \frac{1}{4\pi} \sum [k_1] h_0^{k_1}(k_1, 0) P_{k_1}(\cos\theta_t). \quad (22)$$

The bipolar moments  $h_0^{k_1}(k_1, 0)$  are evaluated by using the expectation values of the Legendre moments of the differential cross-section and expressed as

$$h_0^{k_1}(k_1, 0) = \langle P_{k_1}(\cos\theta_t) \rangle. \quad (23)$$

The PDDCS with  $q = 0$  is presented by

$$\frac{1}{\sigma} \frac{d\sigma_{k0}}{d\omega_t} = \frac{1}{4\pi} \sum [k_1] S_{k0}^{k_1} P_{k_1}(\cos\theta_t), \quad (24)$$

where  $S_{k0}^{k_1}$  is evaluated by the expected value expression and given as

$$S_{k0}^{k_1} = \langle P_{k_1}(\cos\theta_t) P_k(\cos\theta_r) \rangle. \quad (25)$$

Many photon-initiated bimolecular reaction experiments are sensitive to only those polarization moments with  $k = 0$  and  $k = 2$ . In order to compare calculations with experiments, PDDCS<sub>00</sub> ( $\frac{2\pi}{\sigma} \frac{d\sigma_{00}}{d\omega_t}$ ), PDDCS<sub>20</sub> ( $\frac{2\pi}{\sigma} \frac{d\sigma_{20}}{d\omega_t}$ ), PDDCS<sub>22+}</sub> ( $\frac{2\pi}{\sigma} \frac{d\sigma_{22+}}{d\omega_t}$ ), and PDDCS<sub>21-</sub>

( $\frac{2\pi}{\sigma} \frac{d\sigma_{21-}}{d\omega_t}$ ) are calculated. In the above calculations, PDDCSs are expanded up to  $k_1 = 7$ , which is sufficient for good convergence.

### 3. Results and discussion

#### 3.1 The $k$ - $j'$ vector correlation

The product  $P(\theta_r)$  distribution describes the  $k$ - $j'$  vector correlation with  $k j' = \cos(\theta_r)$ . Figure 2 displays the calculated  $P(\theta_r)$  distributions of NH products from the  $N(^2D) + H_2 \rightarrow NH + H$  reaction at collision energies  $E_c = 2.0, 3.8, 5.1, 7.0, 9.0,$  and  $11.0$  kcal/mol. It is clear that the  $P(\theta_r)$  distribution peaks at  $\theta_r = 90^\circ$ , and exhibits a very good symmetry with respect to  $90^\circ$ . This indicates that the product rotational angular momentum vector ( $j'$ ) is aligned along the direction at right angle to the relative velocity direction ( $k$ ). With the collision energy increasing, the peak at  $\theta_r = 90^\circ$  becomes small, which means that high

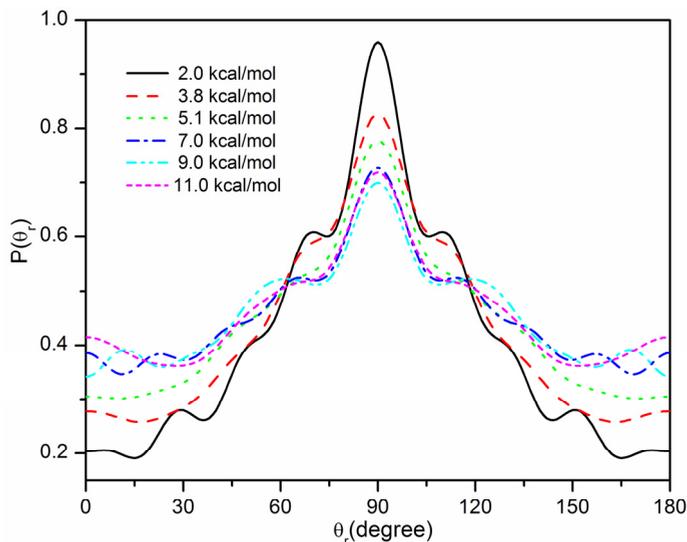


Fig. 2.  $P(\theta_r)$  distributions for the NH products in the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0) \rightarrow \text{NH} + \text{H}$  reaction at  $E_c = 2.0, 3.8, 5.1, 7.0, 9.0,$  and  $11.0$  kcal/mol collision energies.

collision energy weakens the product rotational alignment. However, there is a very tiny exception that the  $P(\theta_r)$  peak at  $\theta_r = 90^\circ$  at  $11.0$  kcal/mol is a little bit smaller than that at  $9.0$  kcal/mol. The location and height of the barrier of the PES is related with the entrance and exit channels, which may be the reason of this tiny disagreement. Figure 3 shows the

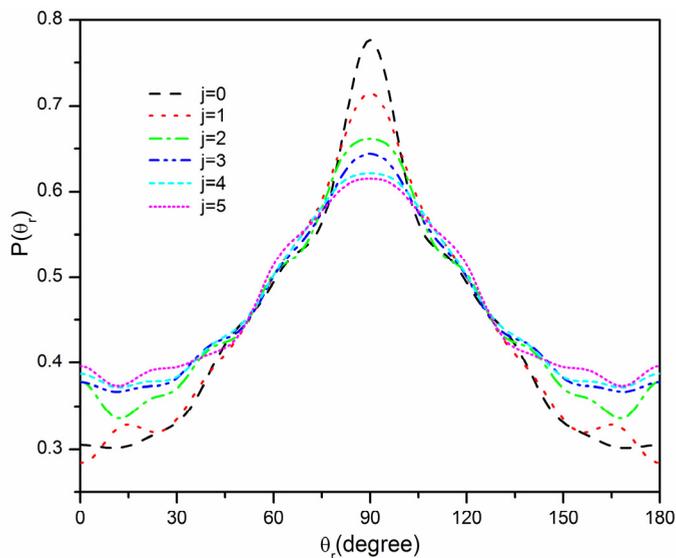


Fig. 3.  $P(\theta_r)$  distributions for the NH products in the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0, 1, 2, 3, 4, 5) \rightarrow \text{NH} + \text{H}$  reaction at  $E_c = 5.1$  kcal/mol collision energy.

effect of the reagent rotational excitation on the product  $P(\theta_r)$  distributions at collision energy of 5.1 kcal/mol. As clearly shown in Fig. 3, the higher the rotational excitation is, the smaller the  $P(\theta_r)$  peak at  $\theta_r = 90^\circ$ . This indicates that the reagent rotational excitation depresses the product rotational alignment.

### 3.2 The $k-k'-j'$ vector correlation

Under the  $\phi_r$  definition of the dihedral angle between the planes consisting of  $k-k'$  and  $k-j'$ , the  $P(\phi_r)$  distribution describes the  $k-k'-j'$  vector correlation and can provide both product alignment and product orientation information. Figure 4 presents the product  $P(\phi_r)$  distributions for the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0) \rightarrow \text{NH} + \text{H}$  reaction at collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol. As can be seen, in contrast to the symmetric feature of the  $P(\theta_r)$  distribution with respect to  $\theta_r = 90^\circ$ , the  $P(\phi_r)$  distribution is asymmetric with respect to the scattering plane where  $\phi_r = 180^\circ$ , but symmetric with respect to that where  $\phi_r = 270^\circ$ . The product  $P(\phi_r)$  distribution behaves a large peak at  $\phi_r = 270^\circ$  and a small peak at  $\phi_r = 90^\circ$  for the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0) \rightarrow \text{NH} + \text{H}$  reaction under each collision energy. Another two peaks emerge at about  $\phi_r = 0^\circ$  and  $\phi_r = 180^\circ$  when collision energy is 2.0 kcal/mol, which disappear in higher collision energies. The peaks at  $\phi_r = 270^\circ$  for the higher collision energies are larger than that for the collision energy  $E_c = 2.0$  kcal/mol, with the largest peak situated at the collision energy  $E_c = 5.1$  kcal/mol. However, the peaks at  $\phi_r = 90^\circ$  have no large change with the collision energy increasing. Therefore, we can conclude that the most NH product rotational angular momentum tend to align along the direction of  $y$  axis which is perpendicular to the scattering  $k-k'$  plane, and orientate along the negative directions of  $y$  axis. That is to say, the product molecules prefer a counterclockwise rotation (see from the negative direction of  $y$  axis) in the plane parallel to the scattering plane.

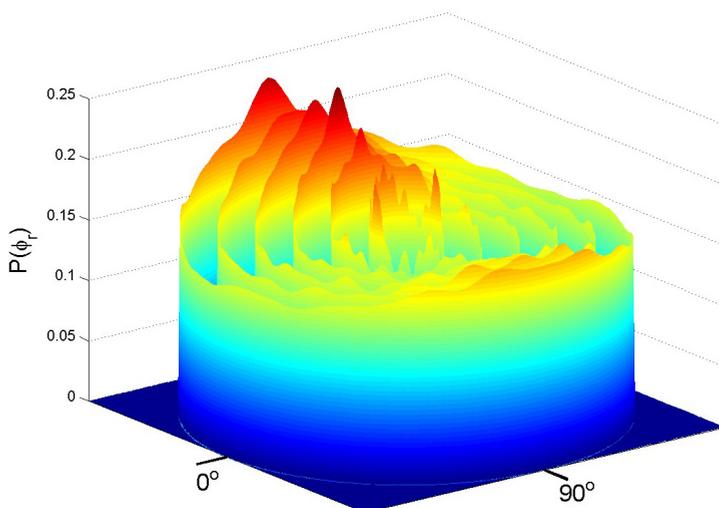


Fig. 4.  $P(\phi_r)$  distributions as a function of the dihedral angle  $\phi_r$  at collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol (from inner to outer) for the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0) \rightarrow \text{NH} + \text{H}$  reaction.

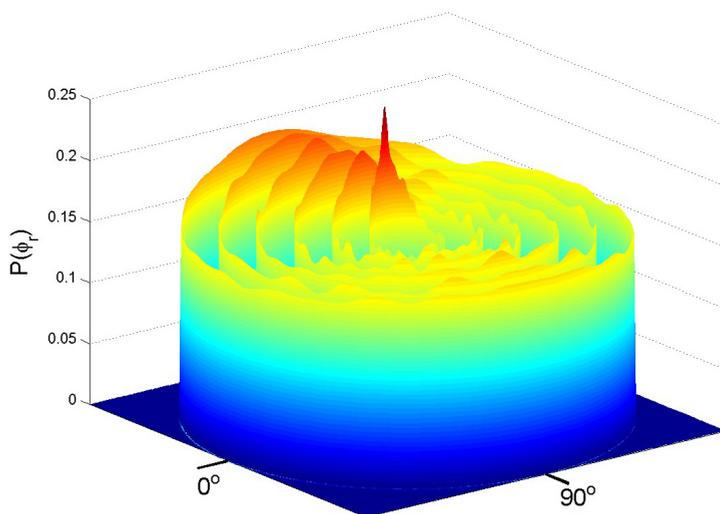
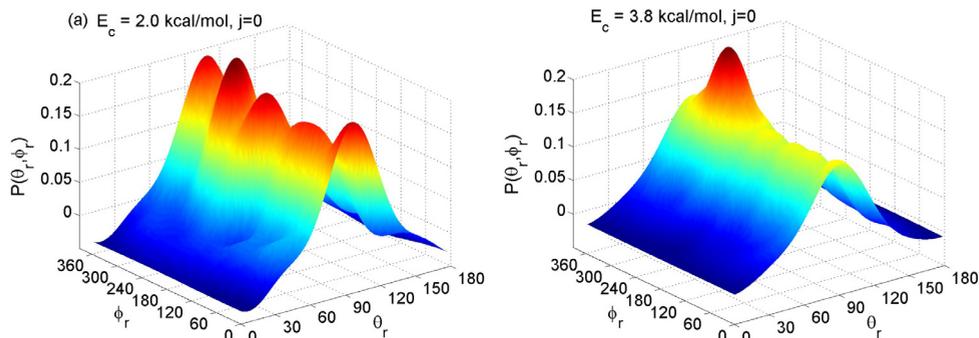
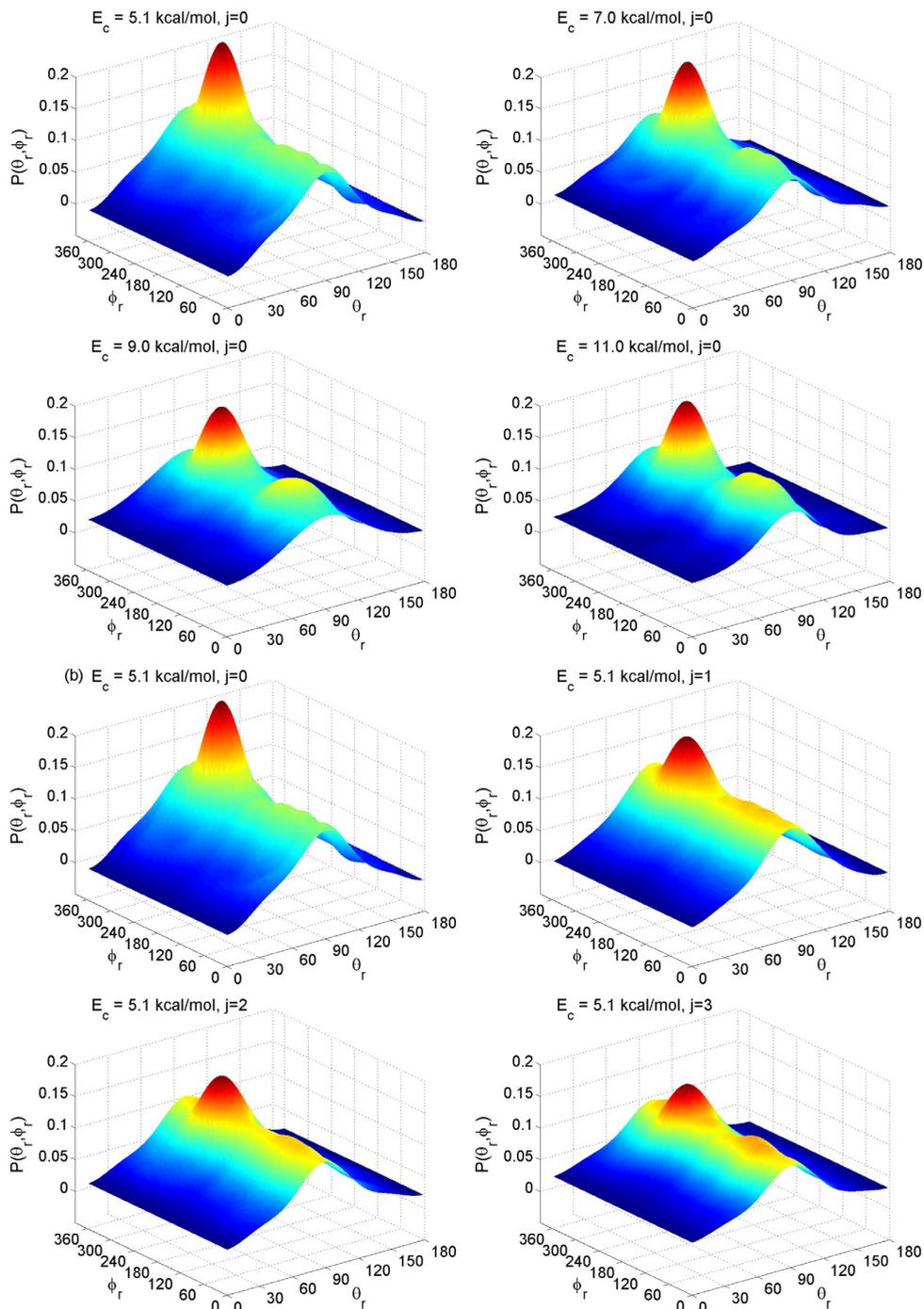


Fig. 5.  $P(\phi_r)$  distributions as a function of the dihedral angle  $\phi_r$  for the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0, 1, 2, 3, 4, 5) \rightarrow \text{NH} + \text{H}$  reactions (from inner to outer) at the collision energy of 5.1 kcal/mol.

Figure 5 reveals the rotational excitation effect on the product  $P(\phi_r)$  distributions for the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0, 1, 2, 3, 4, 5) \rightarrow \text{NH} + \text{H}$  reactions at 5.1 kcal/mol collision energy. Obviously, all of the  $P(\phi_r)$  distributions appear a large peak at about  $\phi_r = 270^\circ$  and a small or no peak at about  $\phi_r = 90^\circ$ , which means that the orientation of the product rotational angular momentum tends to point to the negative direction of  $y$  axis. The peaks at about  $\phi_r = 270^\circ$  for the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=1-5) \rightarrow \text{NH} + \text{H}$  reactions are smaller than that for the ground rotational state reaction, while there is very tiny change for the peaks at about  $\phi_r = 90^\circ$ . This indicates that the product rotational orientation becomes weaker when the reagent is excited to a higher rotational state.





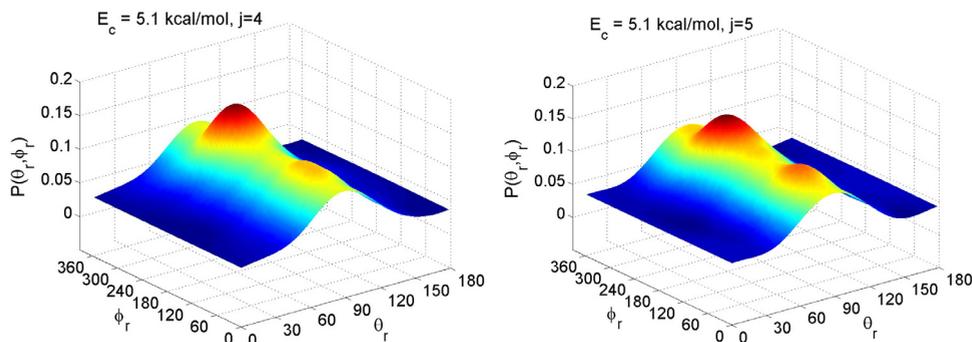


Fig. 6.  $P(\theta_r, \phi_r)$  distributions as a function of both the polar  $\theta_r$  and dihedral angle  $\phi_r$  for the (a)  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0) \rightarrow \text{NH} + \text{H}$  reaction at collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol, (b)  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0, 1, 2, 3, 4, 5) \rightarrow \text{NH} + \text{H}$  reactions at the 5.1 kcal/mol collision energy.

Further supporting information comes from the joint  $P(\theta_r, \phi_r)$  distributions. Figure 6 presents the product  $P(\theta_r, \phi_r)$  distributions for the  $\text{N}(^2\text{D}) + \text{H}_2 \rightarrow \text{NH} + \text{H}$  reaction under different collision energies (Fig. 6a) and different reagent rotational excitations (Fig. 6b). The tomographical features are in good consistence with the separate  $P(\theta_r)$  and  $P(\phi_r)$  distributions. For example, the  $P(\theta_r, \phi_r)$  distribution in the  $\text{N}(^2\text{D}) + \text{H}_2(v=0, j=0) \rightarrow \text{NH} + \text{H}$  reaction at 2.0 kcal/mol collision energy has one largest peak at about  $(\theta_r, \phi_r) = (90^\circ, 270^\circ)$ , one small peak at about  $(\theta_r, \phi_r) = (90^\circ, 90^\circ)$ , and two medium peaks at about  $(\theta_r, \phi_r) = (90^\circ, 0^\circ)$  and  $(\theta_r, \phi_r) = (90^\circ, 180^\circ)$ , respectively. These features reflect the characteristics of the  $P(\theta_r)$  (Fig. 2) and  $P(\phi_r)$  (Fig. 4) distributions for this concrete reaction. Similar things can be found for other specific reactions.

The atom (A) and diatomic molecule (BC) collision interaction can be described qualitatively by "Impulse model", [45] which applies the Newtonian mechanics and consider energy (e.g. endothermic translational, internal (rotational and vibrational), repulsive, and attractive energies) and momentum conservation. The QCT method simulates such interaction or reaction in the time domain step by step (time step, e.g. 0.1 fs) along each trajectory, where the energy and momentum conservations must be controlled from one step to the next step. Based on this impulsive model proposed for an  $\text{A} + \text{BC} \rightarrow \text{AB} + \text{C}$  collision system in previous works, in the present system, the product angular momentum vector  $j'$  can be written as  $j' = L \sin^2 \beta + j \cos^2 \beta + \mathbf{J}_1 m_B / m_{AB}$ . Here  $L$  is the reagent orbital angular momentum and  $j$  is the reagent rotational angular momentum.  $\mathbf{J}_1 = \sqrt{\mu_{BC} E_r} (\mathbf{r}_{AB} \times \mathbf{r}_{CB})$ , with  $\mathbf{r}_{AB}$  and  $\mathbf{r}_{CB}$  being the unit vectors and B pointing to

A and C, respectively,  $\mu_{BC}$  is the reduced mass of the BC molecule and  $E_r$  is the repulsive energy between B and C atoms.  $\cos^2 \beta$  is the mass factor. With respect to the scattering plane, the first and second terms are symmetric terms, and thus the preferred direction of the product angular momentum is determined only by the third term of  $\mathbf{J}_1 m_B / m_{AB}$ , which is eventually traced back to the impulsive energy  $E_r$  of the collision system. As a consequence, we can say that the impulsive energy  $E_r$  leads to the preferred direction of the

product angular momentum vector. At a low collision energy of 2.0 kcal/mol, it seems that this effect of impulsive energy is not so evident, but it becomes more effective as collision energy increases, thus leading to the more preference for the counterclockwise rotation in the planes parallel to the scattering plane (Fig. 4). In the case of reactions for the reagent rotational excitation at collision energy of 5.1 kcal/mol, the alignment and orientation becomes weaker with the higher reagent rotational excitation. It is possibly because the rotational excitation weakens the impulsive energy  $E_r$ . However, the effect of collision energy and rotational excitation on the product rotational alignment and orientation is not always monotonic. As discussed in previous vector correlation studies, the product rotational alignment and orientation is also largely affected by the PES. [46-47] Therefore, the non-monotonic variety of the product alignment and orientation may be resulted by the PES.

### 3.3 The PDDCSs distributions

The generalized PDDCSs describe the  $k$ - $k'$ - $j'$  correlation and the scattering direction of the product molecule. Figure 7 shows the calculated results of the PDDCSs for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction under collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol. The  $PDDCS_{00}$  is simply proportional to the differential cross-section (DCS), and only describes the  $k$ - $k'$  correlation or the product angular distributions. Figure 7a plots the  $PDDCS_{00}$  as a function of scattering angle  $\theta_t$  for the above reaction. Apparently, the  $PDDCS_{00}$  shows a large peak at about  $\theta_t = 180^\circ$  and a tiny peak at about  $\theta_t = 0^\circ$  for the collision energy  $E_c = 2.0$  kcal/mol, which means that the NH product angular distribution is almost the backward scattering. With the collision energy increasing, it shows a decreasing behavior for the peak at about  $\theta_t = 180^\circ$ , but a monotonously increasing behavior for the peak at  $\theta_t = 0^\circ$ . This indicates that the product angular distribution changes from the backward scattering to the both backward and forward scattering. Up to the collision energy  $E_c = 11.0$  kcal/mol, both of the peaks at about  $\theta_t = 180^\circ$  and  $0^\circ$  are nearly the same magnitude, which means that the NH product angular distribution is equally backward and forward scattering in the reaction. These results are in good agreement with recent exact QM results by Lin et al. [30]. The asymmetry in DCS was ascribed to the contributions of the fast insertion component of the reaction and the abstraction channel. The  $PDDCS_{20}$  is the expectation value of the second Legendre moment  $\langle P_2(\cos\theta_r) \rangle$  and contains the alignment information of  $j'$  with respect to  $k$ . As shown in Fig. 7b, the behaviour of the  $PDDCS_{20}$  distribution demonstrates an opposite trend to that of  $PDDCS_{00}$  and obviously depends on the scattering angle  $\theta_t$ . It can be clearly seen from Fig. 7b that the  $PDDCS_{20}$  values are negative for both backward and forward scatterings, but they are close to zero for sideways scattering. These results suggest that the  $j'$  polarizes preferentially along the direction perpendicular to  $k$  when the products are scattered forward and backward. This is consistent with the product alignment prediction from the  $P(\theta_r)$  distribution shown in Fig. 2. Figure 7c and d illustrate the PDDCSs distributions with  $q \neq 0$ . All of the PDDCSs with  $q \neq 0$  are equal to zero at the extremities of forward and backward scattering. The  $PDDCS_{22+}$  value is positive or negative, depending on the preference of  $j'$  alignment along the  $x$  axis or  $y$  axis. It can be seen from Fig. 7c that the  $PDDCS_{22+}$  values of the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction at collision energies of larger than 2.0 kcal/mol are negative for all

scattering angles, which indicates that the alignments of the NH products prefer to be along the  $y$  axis. However, for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction at 2.0 kcal/mol, the  $PDDCS_{22+}$  values are positive in the ranges of  $\theta_t = 30^\circ \sim 73^\circ$ ,  $83^\circ \sim 98^\circ$ , and  $140^\circ \sim 180^\circ$ , with two largest positive peaks at  $\theta_t = 55^\circ$  and  $155^\circ$ , respectively. These two positive peaks are corresponding with the  $P(\phi_r)$  distributions (Fig. 4) in which there have two medium peaks at about  $(\theta_r, \phi_r) = (90^\circ, 0^\circ)$  and  $(\theta_r, \phi_r) = (90^\circ, 180^\circ)$  for the 2.0 kcal/mol collision energy. This demonstrates that the product  $j'$  alignment is not only along the  $y$  axis, but also along the  $x$ -axis. The  $PDDCS_{21-}$  is related to  $\langle -\sin 2\theta_r \cos \phi_r \rangle$ , and its distribution is depicted in Fig. 7d. As shown in Fig. 7d, the  $PDDCS_{21-}$  distribution shows a strongest polarization at about  $\theta_t = 159^\circ$  for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction at 2.0 kcal/mol collision energy. Correspondingly, the  $PDDCS_{21-}$  distributions show a medium polarization for collision energies of 3.8 and 5.1 kcal/mol at respect  $\theta_t = 144^\circ$  and  $140^\circ$ , and a weak polarization for collision energies of 7.0, 9.0, and 11.0 kcal/mol corresponding with  $\theta_t = 83^\circ$ ,  $26^\circ$ , and  $18^\circ$ . These results indicate that the product angular distributions are anisotropic for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction under each collision energy. The polarization degree becomes weaker with the increasing collision energy.

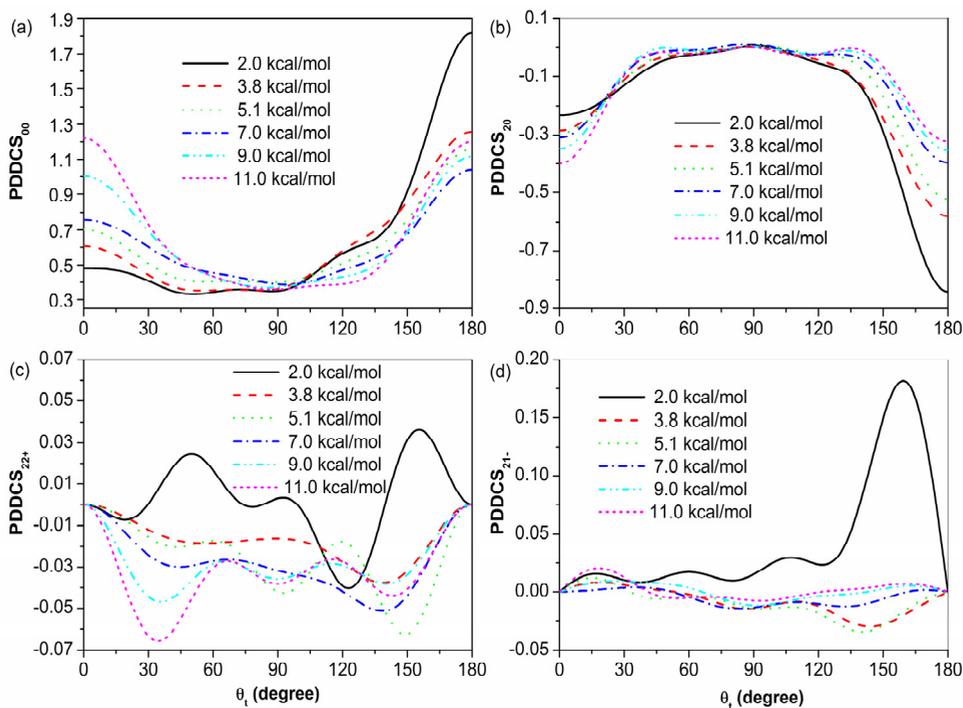


Fig. 7. (a)  $PDDCS_{00}$ , (b)  $PDDCS_{20}$ , (c)  $PDDCS_{22+}$  and (d)  $PDDCS_{21-}$  distributions as a function of scattering angle  $\theta_t$  for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction at collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol.

Figure 8 shows the rotational excitation effect on the PDDCSs distributions for the  $N(^2D) + H_2(v=0, j=0-5) \rightarrow NH + H$  reactions at the 5.1 kcal/mol collision energy. As depicted in Fig. 8a, with the increasing rotational excitation, the peak at about  $\theta_t = 180^\circ$  for the  $PDDCS_{00}$  distribution becomes weaker, but stronger for the peak at about  $\theta_t = 0^\circ$ . These results means that the rotational excitation leads to the product angular distribution changing from the backward scattering to the both backward and forward scattering. The  $PDDCS_{20}$  distributions in Fig. 8b show negative values in both of the backward and forward scatterings, but they are close to zero for the sideways scatterings. These results suggest that the  $j'$  polarizes preferentially along the direction perpendicular to  $k$ , which is consistent with the product  $P(\theta_r)$  distributions in Fig. 3. The  $PDDCS_{22+}$  values are negative for all scattering angles, meaning the alignments of the NH products prefer to be along the  $y$  axis. As can be seen from Fig. 8c, the rotational excitation diminishes the  $PDDCS_{22+}$  peak, implying that the rotational excitation weakened the product alignment. The  $PDDCS_{21-}$  distribution has two largest peaks for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction at  $\theta_t = 15^\circ$  and  $140^\circ$ . However, no distinct large peaks were found in the  $PDDCS_{21-}$  distribution for other rotational excited reactions. These characteristics illuminate that the rotational excitation reduce the anisotropy of the product angular distribution.

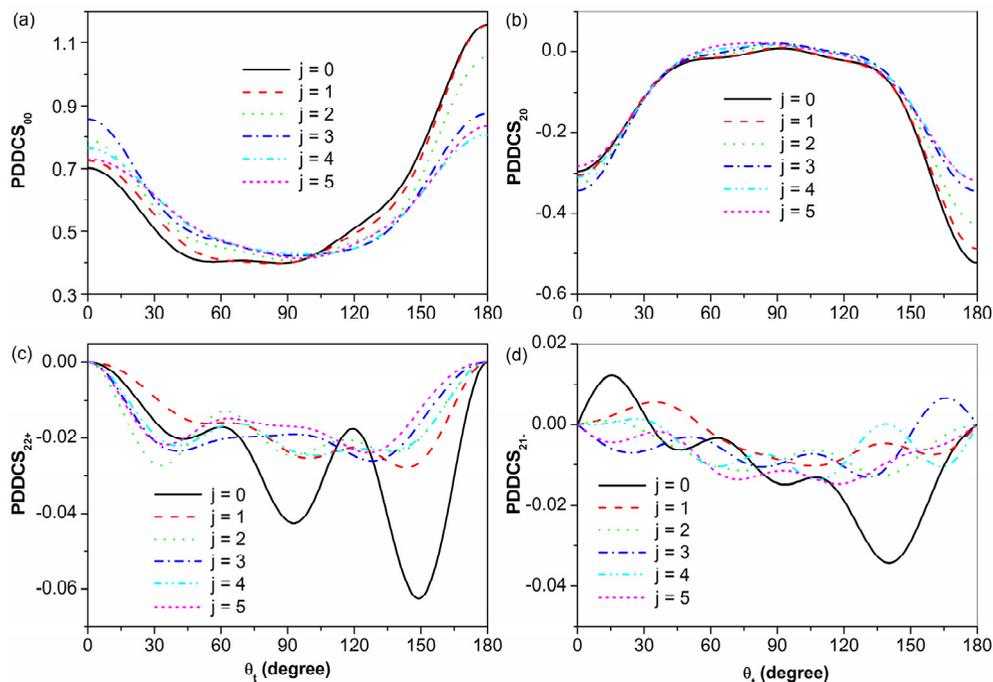


Fig. 8. (a)  $PDDCS_{00}$ , (b)  $PDDCS_{20}$ , (c)  $PDDCS_{22+}$  and (d)  $PDDCS_{21-}$  distributions as a function of scattering angle  $\theta_t$  for the  $N(^2D) + H_2(v=0, j=0, 1, 2, 3, 4, 5) \rightarrow NH + H$  reaction at the collision energy of 5.1 kcal/mol.

### 3.4 Rotational alignment parameter $A_0^{(2)}$

A simple way to express the degree of product rotational polarization of  $j'$  can be through the CM frame alignment parameter  $A_0^{(2)} = 2\langle P_2(\mathbf{j}' \cdot \mathbf{K}) \rangle = \langle 3\cos^2\theta_r - 1 \rangle$ , which is also a usual and common parameter measured in vector correlation experiments. [1] In this expression, the brackets denote an average over the distributions of  $j'$  with respect to  $\mathbf{K}$ . The product rotational alignment parameter  $A_0^{(2)}$  has also been calculated at the present work, which is shown in Table 1. It can be seen from the  $A_0^{(2)}$  expression that, for  $j'$  parallel or antiparallel to  $\mathbf{K}$ ,  $\langle P_2(\mathbf{j}' \cdot \mathbf{K}) \rangle = 1$  and then the alignment parameter takes the value  $A_0^{(2)} = 2$ , while for  $j'$  perpendicular to  $\mathbf{K}$ ,  $A_0^{(2)} = -1$ . The values of the alignment parameter discussed above are limiting cases that represent the maximum possible alignment. In general, these parameters take values of small magnitude, which indicates a distribution that tends toward one of the above limits. If all alignment parameters are zero, the  $j'$  distribution is isotropic about  $\mathbf{K}$ . All of the  $A_0^{(2)}$  values displayed in Table 1 are negative, which means that most of the product rotational angular momentum  $j'$  tend to be perpendicular to  $\mathbf{K}$ . For the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reactions at collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol, the  $A_0^{(2)}$  value becomes larger with the collision energy increasing. For the  $N(^2D) + H_2(v=0, j=0-5) \rightarrow NH + H$  reactions at the collision energy of 5.1 kcal/mol, the  $A_0^{(2)}$  value also becomes larger with the rotational excitation increasing. These trends imply that both of the increasing collision energy and reagent rotational excitation depress the product rotational alignment. This is in good consistent with the the  $P(\theta_r)$  distributions (Fig. 2 and 3).

$E_c$ (kcal/mol)	J	$b_{\max}$ (Å)	$A_0^{(2)}$
2.0	0	1.087	-0.39117
3.8		1.673	-0.22509
5.1		1.804	-0.21372
7.0		1.885	-0.21098
9.0		1.924	-0.21018
11.0		1.947	-0.21786
5.1	1	1.811	-0.22699
	2	1.811	-0.17626
	3	1.810	-0.14679
	4	1.817	-0.13044
	5	1.831	-0.13044

Table 1. Values of the impact parameters  $b_{\max}$  and product rotational alignment parameters  $A_0^{(2)}$  calculated with 100,000 trajectories for the  $N(^2D) + H_2(v=0, j=0-5) \rightarrow NH + H$  reactions under collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol.

#### 4. Conclusion

In this chapter, we presented a detailed vector correlation study on the  $N(^2D) + H_2(v=0, j=0-5) \rightarrow NH + H$  reaction at collision energies of 2.0, 3.8, 5.1, 7.0, 9.0, and 11.0 kcal/mol. Three angular distributions  $P(\theta_r)$ ,  $P(\phi_r)$ ,  $P(\theta_r, \phi_r)$ , and four PDDCSs were computed by the QCT method with a six-order symplectic integration. Under an accurate  $1^2A''$  state PES, batches of 100,000 trajectories were running for the investigated reactions. It was found that the  $P(\theta_r)$  distribution has a large peak at about  $\theta_r = 90^\circ$ , meaning that the product angular momentum  $j'$  is aligned perpendicular to  $k$ . With the collision energy increasing, for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reactions, the product rotational alignment becomes weaker due to the reduction of the  $P(\theta_r)$  peak at about  $\theta_r = 90^\circ$ . Similarly, the rotational excitation also demonstrated a reducing behaviour for the product rotational alignment in the reactions of  $N(^2D) + H_2(v=0, j=0-5) \rightarrow NH + H$  at 5.1 kcal/mol. However, the increasing collision energy enhanced the peak at about  $\phi_r = 270^\circ$ , demonstrating that the product rotational orientation was enhanced by the increasing collision energy. For the  $N(^2D) + H_2(v=0, j=0-5) \rightarrow NH + H$  reactions at 5.1 kcal/mol, the rotational excitation reduced the  $P(\phi_r)$  peak at about  $\phi_r = 270^\circ$ , and therefore depressed the product rotational orientation. The  $P(\theta_r, \phi_r)$  distributions were consistent with the  $P(\theta_r)$  and  $P(\phi_r)$  distributions. These  $P(\theta_r)$ ,  $P(\phi_r)$  and  $P(\theta_r, \phi_r)$  distributions had been interpreted by an "Impulsive model". At the low collision energy of 2.0 kcal/mol, the PDDCS<sub>00</sub> shows a large peak at about  $\theta_t = 180^\circ$  and a very tiny peak at about  $\theta_t = 0^\circ$  for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction, which indicates that the product angular distribution is the backward scattering. As the collision energy increasing, the product angular distribution changes from the backward scattering to the both backward and forward scatterings. Similar behavior was found for the rotational excitation which also leads to the product angular distribution changing from the backward scattering to the both backward and forward scatterings. The PDDCS<sub>20</sub> shows an opposite distribution to the PDDCS<sub>00</sub>. The PDDCS<sub>22+</sub> values in the investigated reactions are negative except for the  $N(^2D) + H_2(v=0, j=0) \rightarrow NH + H$  reaction at 2.0 kcal/mol, meaning that most of the product angular momentum  $j'$  prefer to align along the  $y$  axis. This alignment was reduced by both of the increasing collision energy and rotational excitation. For the PDDCS<sub>21-</sub> distribution, the collision energy increasing and reagent rotational excitation decreased the largest peaks, which indicate that the product rotational polarization is anisotropic. This anisotropic distribution was weakened by the rotational excitation and increasing collision energy. The calculated  $A_0^{(2)}$  values reflect that the product rotational alignment is perpendicular to  $k$ , which is consistent with the aforementioned  $P(\theta_r)$  distributions.

#### 5. Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 21003062) and the Young Funding of Jining University (No. 2009QNKJ02).

#### 6. References

Alexander, A. J. et al. (2000). Product rotational angular momentum polarization in the reaction  $O(^1D_2) + H_2 \rightarrow OH + H$ . *Physical Chemistry Chemical Physics*, Vol.2, No.4, (February 2000), pp. 571-580, ISSN 1463-9076

- Aoiz, F. J. et al. (2004). Classical Stereodynamics in Ar + NO Inelastic Collisions. *Physical Chemistry Chemical Physics*, Vol.6, No.18, (September 2004), pp. 4407-4415, ISSN 1463-9076
- Aquilanti, V. et al. (2005). Orienting and Aligning Molecules for Stereochemistry and Photodynamics. *Physical Chemistry Chemical Physics*, Vol.7, No.2, (January 2005), pp. 291-300, ISSN 1463-9076
- Balucani, N. et al. (2001). Dynamics of the N(<sup>2</sup>D) + D<sub>2</sub> Reaction from Crossed-Beam and Quasiclassical Trajectory Studies. *The Journal of Physical Chemistry A*, Vol.105, No.11, (February 2001), pp. 2414-2422, ISSN 1089-5639
- Balucani, N. et al. (2002). Quantum Effects in the Differential Cross Sections for the Insertion Reaction N(<sup>2</sup>D)+H<sub>2</sub>. *Physical Review Letters*, Vol.89, No.1, (July 2002), pp. 013201, ISSN 0031-9007
- Balucani, N. et al. (2006). Experimental and Theoretical Differential Cross Sections for the N(<sup>2</sup>D) + H<sub>2</sub> Reaction. *The Journal of Physical Chemistry A*, Vol.110, No.2, (January 2006), pp. 817-829, ISSN 1089-5639
- Bañares L. et al. (2005). Influence of Rotation and Isotope Effects on the Dynamics of the N(<sup>2</sup>D)+H<sub>2</sub> Reactive System and of Its Deuterated Variants. *The Journal of Chemical Physics*, Vol.123, No.22, (December 2005), pp. 224301, ISSN 3063-3070
- Barnwell, J. D. et al. (1983). Angular Correlations in Chemical Reactions. Statistical Theory for Four-vector Correlations. *The Journal of Physical Chemistry*, Vol.87, No.15, (July 1983), pp. 2781-2786, ISSN 1089-5639
- Brouard, M. et al. (2000). Product State Resolved Stereodynamics: Rotational Polarization of OH(<sup>2</sup>Π;*v',N',Ω,f*) Scattered from the Reaction, H+CO<sub>2</sub>→OH+CO. *The Journal of Chemical Physics*, Vol.113, No.8, (August 2000), pp. 3173-3180, ISSN 0021-9606
- Casavecchia, P. et al. (1999). Reactive Scattering of Oxygen and Nitrogen Atoms. *Accounts of Chemical Research*, Vol.32, No.6, (February 1999), pp. 503-511, ISSN 0001-4842.
- Castillo, J. F. (2007). Wave Packet and Quasiclassical Trajectory Calculations for the N(<sup>2</sup>D) + H<sub>2</sub> Reaction and Its Isotopic Variants. *Chemical Physics*, Vol.332, No.1, (January 2007), pp. 119-131, ISSN 0301-0104
- Chen, M. D. et al. (2002). Theoretical studies of product polarization and state distributions of the H + HCl reaction. *Chemical Physics*, Vol.283, No.3, (October 2002), pp. 463-472, ISSN 0301-0104
- Chen, M. D. et al. (2003). Theoretical Study of Stereodynamics for the Reactions Cl+H<sub>2</sub>/HD/D<sub>2</sub>. *The Journal of Chemical Physics*, Vol.118, No.10, (March 2003), pp. 4463-4470, ISSN 3063-3070
- Chu, T. S. (2009). Quantum Mechanics and Quasiclassical Study of the H/D+FO → OH/OD+F, HF/DF+O Reactions: Chemical Stereodynamics. *Journal of Computational Chemistry*, Vol.31, No.7, (October 2009), pp. 1385-1396, ISSN 0192-8651
- Chu, T. S. et al. (2006). A Quantum Wave Packet Dynamics Study of the N(<sup>2</sup>D) + H<sub>2</sub> Reaction. *The Journal of Physical Chemistry A*, Vol.110, No.4, (February 2006), pp. 1666-1671, ISSN 1089-5639
- Fano, U. & Macek, J. H. (1973). Impact Excitation and Polarization of the Emitted Light. *Reviews of Modern Physics*, Vol.45, No.4, (October 1973), pp. 553-573, ISSN 0034-6861
- Han, K. L. et al. (1991). Chemical Reaction Dynamics of Barium Atom with Alkyl Bromides. *Chemical Physics Letters*, Vol.181, No.5, (July 1991), pp. 474-478, ISSN 0009-2614

- Ho, T. S. et al. (2003). Implementation of a Fast Analytic Ground State Potential Energy Surface for the  $N(^2D)+H_2$  Reaction. *The Journal of Chemical Physics*, Vol.119, No.6, (August 2003), pp. 9091-9100, ISSN 3063-3070
- Honvault, P. & Launay, J. M. (1999). A Quantum-mechanical Study of the Dynamics of the  $N(^2D)+H_2 \rightarrow NH+H$  reaction. *The Journal of Chemical Physics*, Vol.111, No.15, (October 1999), pp. 6665-6667, ISSN 0021-9606
- Li, R. J. et al. (1994). Rotational Alignment of Product Molecules from the Reactions  $Sr+CH_3Br$ ,  $C_2H_5Br$ ,  $n-C_3H_7Br$ ,  $i-C_3H_7Br$  by Means of PLIF. *Chemical Physics Letters*, Vol.220, No.3-5, (April 1994), pp. 281-285, ISSN 0009-2614
- Lin, S. Y. & Guo, H. (2006). Exact Quantum Dynamics of  $N(^2D)+H_2 \rightarrow NH+H$  Reaction: Cross-sections, Rate constants, and Dependence on Reactant Rotation. *The Journal of Chemical Physics*, Vol.124, No.3, (January 2006), pp. 031101, ISSN 3063-3070
- Lin, S. Y. et al. (2007). Differential and Integral Cross Sections of the  $N(^2D) + H_2 \rightarrow NH + H$  Reaction from Exact Quantum and Quasi-Classical Trajectory Calculations. *The Journal of Physical Chemistry A*, Vol.111, No.12, (March 2007), pp. 2376-2384, ISSN 1089-5639
- M. Alagia, M. et al. (1999). Exploring the reaction dynamics of nitrogen atoms: A combined crossed beam and theoretical study of  $N(^2D)+D_2 \rightarrow ND+D$ . *The Journal of Chemical Physics*, Vol.110, No.18, (May 1999), pp. 8857-8860, ISSN 0021-9606
- Marinakis, S. et al. (2008). Rotational Angular Momentum Polarization: The Influence of Stray Magnetic Fields. *The Journal of Physical Chemistry*, Vol.128, No.2, (January 2008), pp. 021101, ISSN 1089-5639
- McClelland, G. M. & Herschbach, D. R. (1979). Symmetry Properties of Angular Correlations for Molecular Collision Complexes. *The Journal of Physical Chemistry*, Vol.83, No.11, (May 1979), pp. 1445-1454, ISSN 1089-5639
- Orr-Ewing, A. J. & Zare, R. N. (1994). Orientation and Alignment of Reaction Products. *Annual Review of Physical Chemistry*, Vol.45, (October 1994), pp. 315-366, ISSN 0066-426X
- Orr-Ewing, A. J. et al. (1997). Scattering-angle Resolved Product Rotational Alignment for the Reaction of Cl with Vibrationally Excited Methane. *The Journal of Chemical Physics*, Vol.106, No.14, (April 1997), pp. 5961-5971, ISSN 0021-9606
- Pederson, L. A. et al. (1999). Potential Energy Surface and Quasiclassical Trajectory Studies of the  $N(^2D)+H_2$  Reaction. *The Journal of Chemical Physics*, Vol.110, No.18, (May 1999), pp. 9091-9100, ISSN 0021-9606
- Perkins, B. G. & Nesbitt, D. J. (2008). Stereodynamics in State-resolved Scattering at the Gas-Liquid Interface. *Proceedings of the National Academy of Sciences of the United States of America*, Vol.105, No.35, (September 2008), pp. 12684-12689, ISSN 0027-8424
- Rao, B. J. & Mahapatra, S. (2007). Quantum Wave Packet Dynamics of  $N(^2D)+H_2$  Reaction. *The Journal of Chemical Physics*, Vol.127, No.24, (December 2007), pp. 244307, ISSN 3063-3070
- Schlier, C. & Seiter, A. (1998). Symplectic Integration of Classical Trajectories: A Case Study. *The Journal of Physical Chemistry A*, Vol.102, No.47, (July 1998), pp. 9399-9404, ISSN 1089-5639
- Schlier, C. & Seiter, A. (2000). High-order Symplectic Integration: An Assessment. *Computer Physics Communications*, Vol. 130, No.1-2, (July 2000), pp. 176-189. ISSN 0010-4655

- Shafer-Ray, N. E. et al. (1995). Beyond State-to-State Differential Cross Sections: Determination of Product Polarization in Photoinitiated Bimolecular Reactions. *The Journal of Physical Chemistry*, Vol.99, No.19, (May 1995), pp. 7591-7603, ISSN 1089-5639
- Stapelfeldt, H. & Seideman, T. (2003). Colloquium: Aligning Molecules with Strong Laser Pulses. *Reviews of Modern Physics*, Vol.75, No.2, (April 2003), pp. 543-557, ISSN 0034-6861
- Suzuki, T. et al. (1993). Reactions of N(2 <sup>2</sup>D) and N(2 <sup>2</sup>P) with H<sub>2</sub> and D<sub>2</sub>. *Journal of the Chemical Society, Faraday Transactions*, Vol.89, No.7, pp. 995-999, ISSN 0956-5000
- Umemoto, H. & Matsumoto, K. (1996). Production of NH(ND) Radicals in the Reactions of N(2<sup>2</sup>D) with H<sub>2</sub>(D<sub>2</sub>): Nascent Vibrational Distributions of NH(X <sup>3</sup>Σ<sup>-</sup>) and ND(X <sup>3</sup>Σ<sup>-</sup>). *The Journal of Chemical Physics*, Vol.104, No.23, (June 1996), pp. 9640-9643, ISSN 0021-9606
- Umemoto, H. et al. (1997). Nascent Rotational and Vibrational State Distributions of NH(X <sup>3</sup>Σ<sup>-</sup>) and ND(X <sup>3</sup>Σ<sup>-</sup>) Produced in the Reactions of N(2 <sup>2</sup>D) with H<sub>2</sub> and D<sub>2</sub>. *The Journal of Chemical Physics*, Vol.106, No.12, (March 1997), pp. 4985-4991, ISSN 0021-9606
- Umemoto, H. et al. (2000). Verification of the Insertion Mechanism of N(2 <sup>2</sup>D) into H-H Bonds by the Vibrational State Distribution Measurement of NH(X <sup>3</sup>Σ<sup>-</sup>, 0 ≤ v" ≤ 3). *The Journal of Chemical Physics*, Vol.112, No.13, (April 2000), pp. 5762-5766, ISSN 0021-9606
- Wang, M. L. et al. (1997). Reaction Dynamics of the Sr(3P<sub>J</sub>) + RI → SrI\* + R (R = CH<sub>3</sub>, CH<sub>3</sub>CH<sub>2</sub>) Systems: Rotational Alignment, Electronic State Branching Ratio and Vibrational State Population of Products. *Chemical Physics Letters*, Vol.278, No.4-6, (October 1997), pp. 307-312, ISSN 0009-2614
- Wang, M. L. et al. (1998). Product Rotational Polarization in Photo-initiated Bimolecular Reactions A+BC: Dependence on the Character of the Potential Energy Surface for Different Mass Combinations. *The Journal of Physical Chemistry A*, Vol.102, No.50, (November 1998), pp. 10204-10210, ISSN 1089-5639
- Wang, M. L. et al. (1998). Product Rotational Polarization in the Photoinitiated Bimolecular Reaction A+BC→AB+C on Attractive, Mixed and Repulsive Surfaces. *The Journal of Chemical Physics*, Vol.109, No.13, (October 1998), pp. 5446-5454, ISSN 3063-3070
- Wright, A.N. & Winkler C.A. (1968). *Active Nitrogen*. Academic, ISBN 0127651500, New York
- Yue, X. F. et al. (2010). Quasi-Classical Trajectory Study of the Reactions N(2D) with H<sub>2</sub>, D<sub>2</sub>, and HD. *Journal of Theoretical and Computational Chemistry*, Vol.9, No.5, (October 2010), pp. 919-924, ISSN 0219-6336
- Yue, X. F. et al. (2010). Quasi-classical Trajectory Study of the N(2D) + H<sub>2</sub> → NH + H and N(2D) + D<sub>2</sub> → ND + D reactions: Vector Correlation. *Journal of Molecular Structure: THEOCHEM*, Vol.955, No.1-3, (September 2010), pp. 61-65, ISSN 2210-271X
- Yue, X. F. et al. (2010). Stereodynamics Study of the Reactions N(2D) + HD → NH + D and ND + H. *Chinese Physics B*, Vol.19, No.4, (April 2010), pp. 043401, ISSN 1674-1056
- Zhan, J. P. et al. (1997). Rotational Alignment of Products from the NOCl + Ca Chemiluminescent Reaction. *The Journal of Physical Chemistry A*, Vol.101, No.41, (October 1997), pp. 7486-7489, ISSN 1089-5639

---

Zhang, X. & Han, K. L. (2006). High-order Symplectic Integration in Quasi-classical Trajectory Simulation: Case Study for  $O(^1D) + H_2$ . *International Journal of Quantum Chemistry*, Vol.106, No.8, (July 2006), pp. 1815-1819, ISSN 0020-7608

# Schrödinger Transform of Image: A New Tool for Image Analysis

Liantang Lou<sup>1</sup>, Hua Zeng<sup>1\*</sup>, Jipeng Xiong<sup>1</sup>, Lingling Li<sup>2</sup> and Wenliang Gao<sup>1</sup>

*<sup>1</sup>School of Science, Hubei Province Key Laboratory of Intelligent Robot,  
Wuhan Institute of Technology, Wuhan*

*<sup>2</sup>Department of Computer Science and Application, ZhengZhou Institute of  
Aeronautical Industry Management, Zhengzhou  
China*

## 1. Introduction

Image segmentation is the process of separating or grouping an image into different parts . These parts normally correspond to something that human beings can easily separate and view as individual objects. Computers have no means of intelligently for recognizing objects, and a large number of different methods have been developed in order to segment images, ranging from the simple thresholding method to advanced graph-cut methods. The segmentation process is based on various features found in the image. Those features might be histograms information, information about the pixels that indicate edges or boundaries or texture information and so on.

Approaches of Image processing and analysis based on partial differential equation, such as deformable models or snakes (Terzopoulos et al., 1987; Kass, et al., 1987), balloon models (Cohen, L. D., 1991; Cohen, L. D. & Cohen, I., 1993), geometric models (Caselles et al., 1993), discrete dynamic contour models (Lobergt & Viergever, 1995), geodetic active contours (Caselles et al., 1995) and topology adaptive deformable model (McInerney & Terzopoulos, 1999), whose physical background is principle of minimum action or force equilibrium in classical mechanics, are being extensively applied to image segmentation, image smooth, image inpainting, extraction of boundary and so on. Xu and Prince analyzed the reason why snake methods have poor convergence to boundaries with large curvatures and replaced the gradient field with the gradient vector field (GVF), which has a larger capture region and slowly changes away from the boundaries (Xu & Prince, 1998). Consequently, the dependence on initial positions is decreased but the field can attract the moving contour to the right position. Parametric deformable models have high computational efficiency and can easily incorporate a priori knowledge. However, these models cannot naturally handle topological changes and are sensitive to initial conditions. Geometric deformable models are based on the level set method (Osher & Sethian, 1988), which was initially proposed to handle topological changes during the curve evolution. Geometric deformable models have the advantage of naturally handling the topological

---

\* Corresponding Author

changes and are widely studied for medical segmentation (Malladi & Sethian, 1996). Another popular geometric model is proposed by Chan and Vese (Chan & Vese, 1999; 2002). Chan-Vese's model is a simplified version of the Mumford-Shah energy model. The algorithm extracts the desired object through simultaneously minimizing the intensity variations inside and outside the contour. In 1997, Cohen and Kimmel described a method for integrating object boundaries by searching the path of a minimal active deformable model's energy between two points (Cohen, L. D. & Kimmel, 1997). But they are easy to fall into the local minimum, sensitive to noise, do not have topology adaptive, poor convergence to concave boundaries. Lou and Ding used point tracking by estimating the maximum probability of a particle in quantum mechanics moving from one point to another, and did not impose any smoothness constraints to ensure the extraction of the details of a concave contour (Lou & Ding, 2007a). To overcome the main drawbacks of global minimal for active contour models that the contour was only extracted partially for low SNR images, maximal probability method of boundary extraction based on particle motion was proposed (Lou et al, 2007b). Schrödinger transform of image was first given by Lou, Zhan, Fu and Ding, and the probability  $P(b,a)$  that a particle moved from a point  $a$  to another point  $b$  was computed according to I-Type Schrödinger transform of image (Lou et al, 2008). In the chapter, a new tool for image, Schrödinger Transform of image, is investigated.

The remainder of the chapter was organized as follows. First, we gave physical explanation of boundary extraction from the point view of classical mechanics and quantum mechanics in Section 2. Next, we defined Schrödinger transform of image, discussed its properties and computation in Section 3. Then we investigated scale parameter and potential function of Schrödinger transform in Section 4. And then, we constructed high and low pass filter and carried through automatic contour extraction for multiple objects using I-type Schrödinger transform of image, segmented image using deformation of II-type Schrödinger transform of image in Section 5. Finally, we gave our conclusion in Section 6.

## 2. Physical explanation of boundary extraction

Boundary extraction is belong to field of image progressing while classical mechanics and quantum mechanics are pure physical concept. However, we can find their common ground from the point of view of particle motion. Boundary of object can be thought of as trajectory of moving particle while the law of motion of particle were investigated in classical mechanics and quantum mechanics using two determinacy and nondeterminacy method.

### 2.1 Physical explanation of boundary extraction from the point of view of classical mechanics

Deformable models are the elastic curves defined within an image domain that can move under the influence of internal forces arising from curve smoothness and external forces computed from the image data. The internal and external forces are so defined that the deformable contour has a minimum energy at the true object boundary. The following mapping can represent the deformable contour model:

$$\begin{aligned} \mathbf{x}(s) : [0, 1] &\rightarrow \mathbb{R}^2 \\ s &\mapsto (x(s), y(s))' \end{aligned} \quad (1)$$

where  $s \in [0, 1]$  is the parameterization variable of the object boundary and 0 and 1 correspond to the start and end points of the boundary. The deformable contour is a curve  $\mathbf{x}(s) = (x(s), y(s))$  that moves in the spatial domain of the image to minimize the energy:

$$E = \int_0^1 \frac{1}{2} (w_1 |\mathbf{x}'(s)|^2 + w_2 |\mathbf{x}''(s)|^2) + E_{ext}(\mathbf{x}(s)) ds \quad (2)$$

where  $w_1, w_2$  are the weighting parameters controlling the contour's tension and rigidity respectively.  $\mathbf{x}'(s)$  and  $\mathbf{x}''(s)$  are the first and second order derivatives of  $\mathbf{x}(s)$  with respect to  $s$ . The external energy  $E_{ext}$  is a function of the image data depending on the goal of application. For object boundary extraction, it is defined as the image gradient. Suppose  $\mathbf{x}(s)$  has a local minimum of  $E$  at  $s$ , the following Euler-Lagrange Equation has to be satisfied:

$$w_1 \mathbf{x}''(s) - w_2 \mathbf{x}''''(s) - \nabla E_{ext} = 0 \quad (3)$$

To solve Eq. (3), the boundary conditions of  $\mathbf{x}(s)$  should be given. If the boundary is closed, a periodic boundary condition  $\mathbf{x}(0) = \mathbf{x}(1)$  can be attached. Each term in Eq. (3) is considered as a force applied on the curve. The internal force  $w_1 \mathbf{x}''(s) - w_2 \mathbf{x}''''(s)$  resists stretching and bending while the external potential force  $-\nabla E_{ext}$  pulls the contour towards the boundary. Thus, the object boundary is obtained either when the force equilibrium of Eq. (3) or a minimum of the energy in Eq. (2) is reached.

## 2.2 Physical explanation of boundary extraction from the point of view of quantum mechanics

The deformable contour  $\mathbf{x}(s) = (x(s), y(s))$  can further be considered as the path of a moving particle in the image if the parameterization variable  $s$  is replaced by the time variable  $t$ . By referring the work of Feynman and Hibbs (Feynman & Hibbs, 1965), we explain boundary extraction from the point of view of quantum mechanics as follows:

Suppose a particle moves from the position  $a$  at the time  $t_a$  to the position  $b$  at the time  $t_b$ , e.g.,  $a = \mathbf{x}(t_a)$ ,  $b = \mathbf{x}(t_b)$ . According to the theory of quantum mechanics, the amplitude,  $K(b, a)$ , called kernel or propagator, contains the total contribution of all paths between  $a$  and  $b$ , which is different from Eq. (2) where only a specific path from  $a$  to  $b$  with the minimum energy is concerned. In order to distinguish these two types of contours, we refer to the contour determined by  $K(b, a)$  as the quantum contour, denoted as  $Q(t)$ , while the traditional deformable contour determined by minimizing the energy between  $a$  and  $b$  as the classical contour, denoted by  $X(t)$ . Obviously, the classical contour  $X(t)$  is considered as a specific case of the quantum contour  $Q(t)$  when a single path is concerned. In physics, the energy functional of a path,  $E$ , is defined by:

$$E = \int_{t_a}^{t_b} L(\mathbf{x}'(t), \mathbf{x}(t), t) dt \quad (4)$$

where  $L$  is the Lagrangian function of system. For a moving particle with the mass  $m$ , and potential  $V(\mathbf{x}, t)$ , the Lagrangian function is determined by:

$$L = \frac{m}{2} (\mathbf{x}'(t))^2 - V(\mathbf{x}, t) \quad (5)$$

Comparing Eq. (4) to Eq. (2), we find that the Lagrangian functional of a path in Eq. (4) is similar to the energy functional of a contour curve in Eq. (2). Thus, we refer to deformable model as the motion of a particle described by classical mechanics.

In quantum mechanics, the total contribution of all paths between  $a$  to  $b$  is calculated by

$$K(b,a) = \sum_{R(a,b)} \phi(\mathbf{x}(t)) \quad (6)$$

where  $R(a,b)$  is the set of all paths between  $a$  and  $b$ .  $\phi(\mathbf{x}(t))$  is the contribution of a path  $\mathbf{x}(t)$  with a phase proportional to its energy  $E(\mathbf{x}(t))$ , *i.e.*,

$$\phi(\mathbf{x}(t)) = C \cdot e^{(2\pi j/h)E(\mathbf{x}(t))} \quad (7)$$

where  $h$  is the Planck's Constant and  $C$  is a constant.

According to the theory of quantum mechanics, the probability<sup>1</sup> of a particle moving from the position  $a$  at time  $t_a$  to the position  $b$  at  $t_b$ , denoted by  $P(b,a)$ , is equal to the square amplitude of  $K(b,a)$ , *i.e.*,

$$P(b,a) = |K(b,a)|^2 \quad (8)$$

For a system with a simple Lagrangian function,  $K(b,a)$  can be calculated directly from the path integral ((Feynman & Hibbs, 1965)) while for a system with a complex Lagrangian function, it is difficult and time-consuming to estimate the value of  $P(b,a)$  from  $K(b,a)$ . In order to avoid such difficulty, we estimate the probability of a particle moving from point  $a$  to point  $b$  directly from specific particle models, *e.g.*, a free particle or a particle moving through a Gaussian slit, where their motions can be used to describe the boundary of an object of interest with a known probability density of appearing at a point.

There is a stronger motivation of adapting the quantum mechanics than the similarity between equations (2) and (4). Snakes, Deformable models, level set methods, etc, are all based on classical mechanics in a form of partial differential equation. Although classical physics is adequate to explain virtually all phenomena one will ever directly experience in one's life, certain phenomena cannot be explained by classical physics. In many respects, quantum mechanics presents the physics that underlies physical reality at its most basic level. Quantum theory can be thought of as the generalization of classical mechanics and many non-classical phenomena that do not have a classical analog are known in the quantum physics world. The relationship between classical and quantum mechanics is of central importance to the philosophy of physics. Classical mechanics extends the elementary Newtonian concepts to the Lagrangian and Hamiltonian formulations, to the least action principle, to the angle-action variables, etc, in ways that are the essential framework of quantum mechanics. However, there are significant distinctions between the two theories that arise not because of quantization, but rather from the nonessential tendency to describe macroscopic systems by instantaneous values for position, speed and acceleration, and microscopic systems by time-averaged position probability densities. Probability is a bridge between classical mechanics and quantum mechanics. The detail discussion of this relationship is out of this chapter's scope, which can be founded by hundreds of literatures in quantum mechanics field.

---

<sup>1</sup> Here the probability refers to a relative probability.

### 3. Schrödinger transform of image

For a complex system, however, it is difficult to calculate the value of  $P(b,a)$ . In 1995, Williams and Jacobs derived the probability that a particle with a random walk passes through a given position and orientation on a path joining two boundary fragments, which is obtained by the product of two vector-field convolutions (Williams & Jacobs, 1995).

Although some specific particle motions have been considered, a general analytic expression of the probability for complex system is still open. In order to calculate the value of  $P(b,a)$ , a numerical approximation of  $P(b,a)$  is needed. In this section, we will try to compute the probability  $P(b,a)$  by using Schrödinger transform of image.

#### 3.1 Definition of Schrödinger transform of image

The active contour model or Snake model had their profound physical background. If the parameter  $s$  in the deformable contour curve  $\mathbf{x}(s) = (x(s), y(s))$  could be understood as time  $t$ , object contour curve  $\mathbf{x}(t)$  could be considered as the path of the particle in plane motion.

Suppose a particle moves from the position  $a$  at the time  $t_a$  to the position at the time  $t_b$ , e.g.,  $a = x(t_a)$ ,  $b = x(t_b)$ . According to the theory of quantum mechanics, the probability of a particle moving from the position  $a$  to  $b$  at  $t_b$ , denoted by  $P(b,a)$ , is dependent on the kernel  $K(b,a)$ , which is the sum of all paths contribution between  $\mathbf{x}_a$  and  $\mathbf{x}_b$ , i.e.,

$$K(b,a) = \sum_{R(a,b)} \phi(\mathbf{x}(t)), \quad (9)$$

where  $R(a,b)$  is the set of all paths between  $\mathbf{x}_a$  and  $\mathbf{x}_b$ .  $\phi(\mathbf{x}(t))$  is the contribution of a path  $\mathbf{x}(t)$  with a phase proportional to its energy  $E(\mathbf{x}(t))$ , i.e.,

$$\phi(\mathbf{x}(t)) = Ce^{(2\pi i/h)E(\mathbf{x}(t))}, \quad (10)$$

where  $h$  is the Planck's Constant,  $C$  is a constant, and  $E = \int_{t_a}^{t_b} \left( \frac{m}{2} |\mathbf{x}'(t)|^2 - V(\mathbf{x}, t) \right) dt$  is the energy functional of path. The probability from point  $\mathbf{x}_a$  at time  $t_a$  to point  $\mathbf{x}_b$  at time  $t_b$  is the square of absolute value of kernel  $K(b,a)$  from  $\mathbf{x}_a$  to  $\mathbf{x}_b$ , that is,

$$p(a,b) = |K(b,a)|^2.$$

We must solve the problem of computing the kernel  $K(b,a)$  to introduce law of particle motion in quantum mechanics into image processing and analysis. For a system with a simple Lagrangian function,  $K(b,a)$  can be calculated directly from the path integral (see 0) while for a system with a complex Lagrangian function, it is difficult and time-consuming to estimate the value of  $P(b,a)$  from  $K(b,a)$ . Replacing the kernel  $K(b,a)$  with the wave function  $u(\mathbf{x}, t)$  in the position  $\mathbf{x}$  at the time  $t$ , then  $u(\mathbf{x}, t)$  satisfied the following Schrödinger equation:

$$\bar{h}i \cdot \frac{\partial u}{\partial t} = -\frac{\bar{h}^2}{2m} \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + V(\mathbf{x}, t)u(\mathbf{x}, t), \quad (11)$$

where  $\hbar = h / 2\pi = 1.054 \times 10^{-27} \text{ crg} \cdot \text{s}$ ,  $i$  is imaginary unit. So, the relation between the probability  $P(b,a)$  and image gradient  $G(\mathbf{x})$  can be given in Fig. 1.

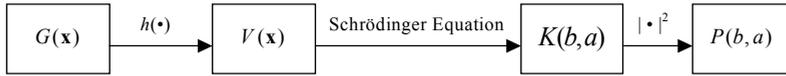


Fig. 1. the relation between the probability  $P(b,a)$  and image gradient  $G(\mathbf{x})$

We could rewrite Eq.(11) as the initial-value problem:

$$\begin{cases} i \cdot u_t + a\Delta u = v(\mathbf{x})u \\ u|_{t=0} = \varphi(\mathbf{x}) \end{cases} \tag{12}$$

Continuous Schrödinger transform of image  $\varphi(\mathbf{x})$  based on  $v(\mathbf{x})$  is defined as the solution of Eq.(12). And the transform is called I-type Schrödinger transform when  $v(\mathbf{x})=0$ , otherwise the transform is called II-type Schrödinger transform.

By applying Fourier transform to equation (12) and making use of the properties of Fourier transform, we have

$$\begin{cases} i \cdot \hat{u}_t - a|\mathbf{y}|^2 \hat{u} = \hat{v}(\mathbf{y}) * \hat{u} \\ \hat{u}|_{t=0} = \hat{\varphi}(\mathbf{y}) \end{cases}, \tag{13}$$

where the mark '\*' denotes convolution of two functions, '^ ^' denotes Fourier transform of function. When  $v(\mathbf{x})=0$ , both  $u(\mathbf{x},t)$  and  $\hat{u}(\mathbf{y},t)$  have the following analytic solutions (see 0):

$$\hat{u} = e^{-ait|\mathbf{y}|^2} \hat{\varphi}(\mathbf{y}) \tag{14}$$

$$u(\mathbf{x},t) = \frac{1}{4\pi ait} \iint e^{\frac{i}{4ait}|\mathbf{x}-\mathbf{y}|^2} \hat{\varphi}(\mathbf{y}) d\mathbf{y}, \quad \mathbf{x} \in R^2, t > 0, \tag{15}$$

When  $v(\mathbf{x}) \neq 0$ ,  $u(\mathbf{x},t)$  and  $\hat{u}(\mathbf{y},t)$  also have analytic solutions (L. C. Evans, 1998), but they are too complex to be used to compute their numerical solutions. We give the following definition of discrete Schrödinger transform of image because of Eq.(12) and Eq.(13):

Supposed both  $\varphi(\mathbf{x})$  and  $v(\mathbf{x})$  are  $m \times n$  images, then two-dimensional discrete Schrödinger transform of image  $\varphi(\mathbf{x})$  based on  $v(\mathbf{x})$  is expressed with the following differential equation which its Fourier transform satisfies:

$$\begin{cases} i \cdot \tilde{u}_t = (V + a|\mathbf{y}|^2) \tilde{u} \\ \tilde{u}|_{t=0} = \tilde{\varphi} \end{cases}, \tag{16}$$

where  $\tilde{u}_t$  is  $mn$ -dimensional column vector formed by concatenating all the rows of  $m \times n$  matrix  $\hat{u}_t$ .  $mn \times mn$  matrix  $|y|$  was diagonal matrix whose diagonal elements express distance.  $mn \times mn$  matrix  $V$  is a block cyclic matrix, i.e.,

$$V = \begin{pmatrix} V_0 & V_{m-1} & \cdots & V_1 \\ V_1 & V_0 & \cdots & V_2 \\ \vdots & \vdots & \ddots & \vdots \\ V_{m-1} & V_{m-2} & \cdots & V_0 \end{pmatrix}, \tag{17}$$

where  $V_i$  is a cyclic matrix,

$$V_i = \begin{pmatrix} \hat{v}(i,0) & \hat{v}(i,n-1) & \cdots & \hat{v}(i,1) \\ \hat{v}(i,1) & \hat{v}(i,0) & \cdots & \hat{v}(i,2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{v}(i,n-1) & \hat{v}(i,n-2) & \cdots & \hat{v}(i,0) \end{pmatrix}, \tag{18}$$

Discrete Schrödinger transform of image  $\varphi(x)$  based on  $v(x)$  is denoted by  $S_v(\varphi)$  while I-type discrete Schrödinger transform of image  $\varphi(x)$  is denoted by  $S(\varphi)$ . Obviously, the solution of Eq.(16) is

$$S_v(\varphi) = \hat{u}(y) = e^{-it(V+a|y|^2)} \hat{\varphi}(y). \tag{19}$$

If the matrix  $V + a|y|^2$  can be diagonalized, that is,  $V + a|y|^2 = P^{-1}DP$ , then

$$\hat{u}(y) = P^{-1}Diag(e^{-itd_1}, e^{-itd_2}, \dots, e^{-itd_{mn}})P\hat{\varphi}(y), \tag{20}$$

where  $D = Diag(d_1, d_2, \dots, d_{mn})$  was diagonal matrix. Eq.(19) degenerates into Eq.(14) when  $v(x) = 0$ .

### 3.2 Computation of Schrödinger transform

The transfer function of Schrödinger transform of image is  $e^{-it(V+a|y|^2)}$ . For a given  $m \times n$  image and constant  $at$ , I-type Schrödinger transform of image  $\varphi(x, y)$  can be computed as the following steps:

1. Suppose the low frequency component be in the center of image, we construct a  $m \times n$  distance matrix  $D = (d_{uv})$ , where

$$d_{st} = (u - m / 2)^2 + (v - n / 2)^2$$

2. Compute the transfer function of I-type Schrödinger transform,  $H = (h_{uv})$

$$h_{uv} = e^{-atid_{uv}}$$

3. Compute the Fourier transform of  $\varphi(x, y)$ ,  $\hat{\varphi}(u, v)$

4. Compute the Fourier transform of  $\mathbb{S}(\varphi)$ ,  $\hat{\mathbb{S}}(\varphi)$ , according to the following formula:

$$\hat{\mathbb{S}}(\varphi) = e^{-aitd_{uv}} \hat{\varphi}(u, v)$$

5. The I-type Schrödinger transform of image  $\varphi(x, y)$  is the modulus of the inverse Fourier transform of  $\hat{\mathbb{S}}(\varphi)$ .

It is difficult to directly compute II-type Schrödinger transform of image using Eq.(19). We can compute II-type Schrödinger transform using a two step method since the matrix  $V$  is Block Circulant Matrix and the matrix  $|\mathbf{y}|^2$  is diagonal matrix in Eq.(19). The block circulant matrix  $V$  is similar to a diagonal matrix, that is,

$$V = WD_VW^{-1} \tag{21}$$

Here,  $W = W_m \otimes W_n$ ,  $W_m$  is a  $m \times m$  matrix with  $W_m(j, k) = e^{-2\pi ijk/m}$ ,  $W^{-1} = W_m^{-1} \otimes W_n^{-1}$ ,  $W_m^{-1}(j, k) = \frac{1}{m} e^{2\pi ijk/m}$ , and  $D_V$  is a  $mn \times mn$  diagonal matrix.

We know that the exponential function satisfies  $e^{x+y} = e^x e^y$  for any real numbers (scalars)  $x$  and  $y$ . The same goes for commuting matrices: if the matrices  $X$  and  $Y$  commute (meaning that  $XY = YX$ ), then  $e^{X+Y} = e^X e^Y$ . It is usually necessary for  $A$  and  $B$  to commute for the law to still hold. However, in mathematics, the Lie product formula, named for Sophus Lie, holds for all matrices  $A$  and  $B$ , even ones which do not commute. That is, for arbitrary real or complex matrices  $A$  and  $B$ ,

$$e^{A+B} = \lim_{N \rightarrow \infty} \left( e^{A/N} e^{B/N} \right)^N \tag{22}$$

The formula has applications, for example, in the path integral formulation of quantum mechanics. It allows one to separate the Schrödinger evolution operator into alternating increments of kinetic and potential operators. Hence, we rewrite (19) as

$$\hat{\mathbb{S}}_v(\varphi) = e^{-it(V+a|\mathbf{y}|^2)} \hat{\varphi}(\mathbf{y}) \approx \left( e^{-itV/N} e^{-iat|\mathbf{y}|^2/N} \right)^N \hat{\varphi}(\mathbf{y}) \tag{23}$$

For a given  $N$ , denote,

$$\hat{\varphi}_k(\mathbf{y}) = \left( e^{-itV/N} e^{-iat|\mathbf{y}|^2/N} \right)^k \hat{\varphi}(\mathbf{y}) \tag{24}$$

So,

$$\hat{\varphi}_k(\mathbf{y}) = e^{-itV/N} e^{-iat|\mathbf{y}|^2/N} \hat{\varphi}_{k-1}(\mathbf{y}) \tag{25}$$

Where  $\hat{\varphi}_0(\mathbf{y}) = \hat{\varphi}(\mathbf{y})$ .

Replacing  $V$  by Eq.(21) in Eq.(25), we get that

$$\hat{\varphi}_k(\mathbf{y}) = W e^{-itD_v/N} W^{-1} \hat{\mathbb{S}}(\varphi_{k-1}) \tag{26}$$

That is,

$$W^{-1}\hat{\varphi}_k(\mathbf{y}) = We^{-itD_V/N}W^{-1}\hat{S}(\varphi_{k-1}) \tag{27}$$

Note that,  $\hat{S}(\varphi_{k-1})$  are Fourier transforms of I-type of image  $\varphi_{k-1}(\mathbf{x})$  and  $W^{-1}$  is inverse Fourier transform. Hence,

$$\varphi_k(\mathbf{x}) = e^{-itD_V/N}\mathbb{S}(\varphi_{k-1})(\mathbf{x}) \tag{28}$$

Here,  $k=1,2,\dots,N$ ,  $D_V$  is the inverse Fourier transform of  $V$  while  $V$  is the Fourier transform of  $v(\mathbf{x})$ , that is,  $D_V$  is a  $mn \times mn$  diagonal matrix with the diagonal elements  $v(\mathbf{x})$ . So, we obtain the following two step method of computing II-type Schrödinger transform:

- Step 1: Using the formual  $\hat{S}(\varphi_{k-1}) = e^{-itV/N} e^{-iat|\mathbf{y}|^2/N} \hat{\varphi}_{k-1}(\mathbf{y})$ , compute I-type Schrödinger transform of image  $\varphi(x,y)$  in the frequency domain of image (see step (1) to step (4) of computing I-type Schrödinger transform of image );
- Step 2: Compute II-type Schrödinger transform of image  $\varphi(x,y)$  in the spatial domain of image according to Eq.(28), that is,

$$\mathbb{S}_v(\varphi)(\mathbf{x}) \approx \varphi_N(\mathbf{x}) = e^{-itD_V/N}\mathbb{S}(\varphi_{N-1})(\mathbf{x}) \tag{29}$$

Eq.(29) means that discrete Schrödinger transform of image  $\varphi(\mathbf{x})$  based on  $v(\mathbf{x})$  is the approximative superposition of the following two equations

$$\begin{cases} i \cdot \tilde{u}_t^I = a|\mathbf{y}|^2 \tilde{u}^I \\ \tilde{u}^I|_{t=0} = \tilde{\varphi} \end{cases} \tag{30}$$

and

$$\begin{cases} i \cdot \vec{u}_t = V\vec{u} \\ \vec{u}|_{t=0} = \vec{u}^I \end{cases} \tag{31}$$

Obviously, for continous Schrödinger transform of image  $\varphi(\mathbf{x})$  based on  $v(\mathbf{x})$ , the conclusion mentioned above is not true proposition, that is, eq.(12) can not decompose the following two initial-value problem:

$$\begin{cases} i \cdot u_t^I + a\Delta u^I = 0 \\ u^I|_{t=0} = \varphi(\mathbf{x}) \end{cases} \tag{32}$$

and

$$\begin{cases} i \cdot u_t = v(\mathbf{x})u \\ u|_{t=0} = u^I(\mathbf{x}) \end{cases} \tag{33}$$

Hence, there are essential differences between continous Schrödinger transform continous and discrete Schrödinger transform.

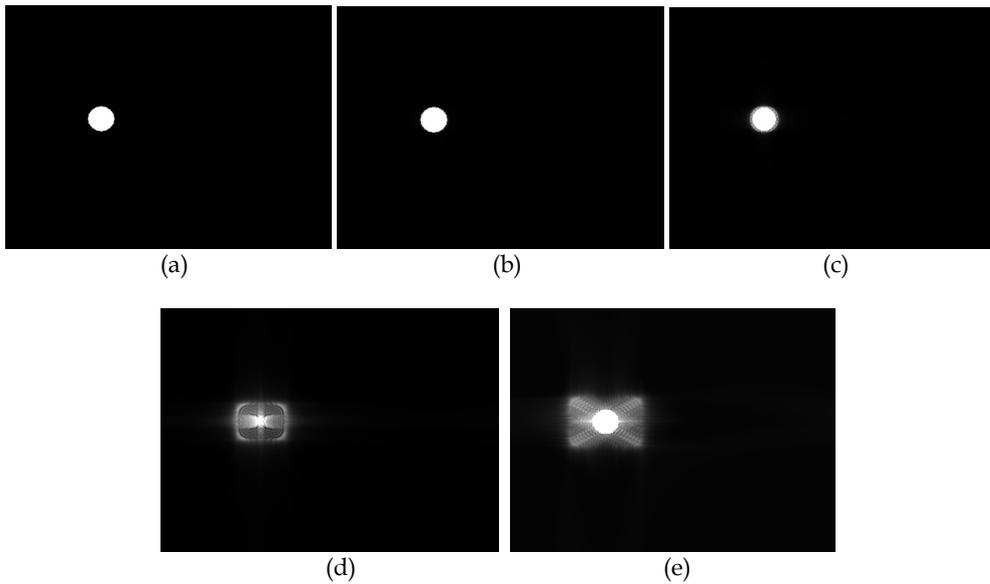
**3.3 Properties and meaning of Schrödinger transform**

There are a variety of properties associated with the Schrödinger transform of image. The following are some of the most relevant for I-type Schrödinger transform of image. Energy conservation property also exists for the Schrödinger transform of image like the Fourier transform.

Proposition 1. (Energy Conservation Theorem) Let  $u(\mathbf{x}, t)$  be Schrödinger transform of image  $\phi(\mathbf{x})$ , then

$$\int_{-\infty}^{+\infty} |u(\mathbf{x}, t)|^2 dx = \int_{-\infty}^{+\infty} |\phi(\mathbf{x})|^2 dx$$

The proposition can be proved according to equation (14) and energy conservation properties of 2D Fourier Transform of image. The energy conservation properties of Schrödinger transform show that energy will diffuse from high energy to low energy while total energy is invariable (Fig. 2).



(a) original image, (b), (c), (d) and (e) are Schrödinger transforms with parameters 0.00001, 0.00005, 0.0005 and 0.001, respectively.

Fig. 2. Schrödinger transforms of a circular disc image with different parameters  $at$

Proposition 2. Let filter transfer function be  $H(\mathbf{y}) = \frac{1}{2} (1 + e^{-at|\mathbf{y}|^2})$ , where  $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ,  $|\mathbf{y}|^2 = y_1^2 + y_2^2$ , If  $|\mathbf{y}|^2 \leq M$  (There exists highest frequency for digital image), then  $|H(\mathbf{y})|^2$  is a decrease function of  $|\mathbf{y}|^2$  when  $at \leq \frac{\pi}{2M}$ , that is,  $H(\mathbf{y})$  is low pass filter.

Proposition 3. Let filter transfer function be  $H(\mathbf{y}) = 1 - e^{-at\|\mathbf{y}\|^2}$ , where  $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ,  $\|\mathbf{y}\|^2 = y_1^2 + y_2^2$ , If  $\|\mathbf{y}\|^2 \leq M$  (There exists highest frequency for digital image), then  $|H(\mathbf{y})|^2$  is a increase function of  $\|\mathbf{y}\|^2$  when  $at \leq \frac{\pi}{2M}$ , that is,  $H(\mathbf{y})$  is high pass filter.

High and low pass filter can be obtained using Schrödinger transform of image according to the above two propositions. And Schrödinger transform of image can be applied to image processing and analysis, such as, boundary extraction, edge enhancement, etc.

The energy conservation properties of Schrödinger transform show that energy will diffuse from high energy to low energy while total energy is invariable (Fig. 2). The following experiments (see Fig. 3 and Fig. 4) show the meaning and function of Schrödinger transformation of image, that is, Schrödinger transformation of image can be seen as the result of primitive image shrinking inside and spreading outward at the center of object, like as interference wave. The bigger  $at$  is, the more obvious the interference is. On the other hand, contour curves of object in the transformed image are similar to contour curves of object in the original image, and they are too similar to draw them manually. Hence, I-type Schrödinger transform of image is isotropic.

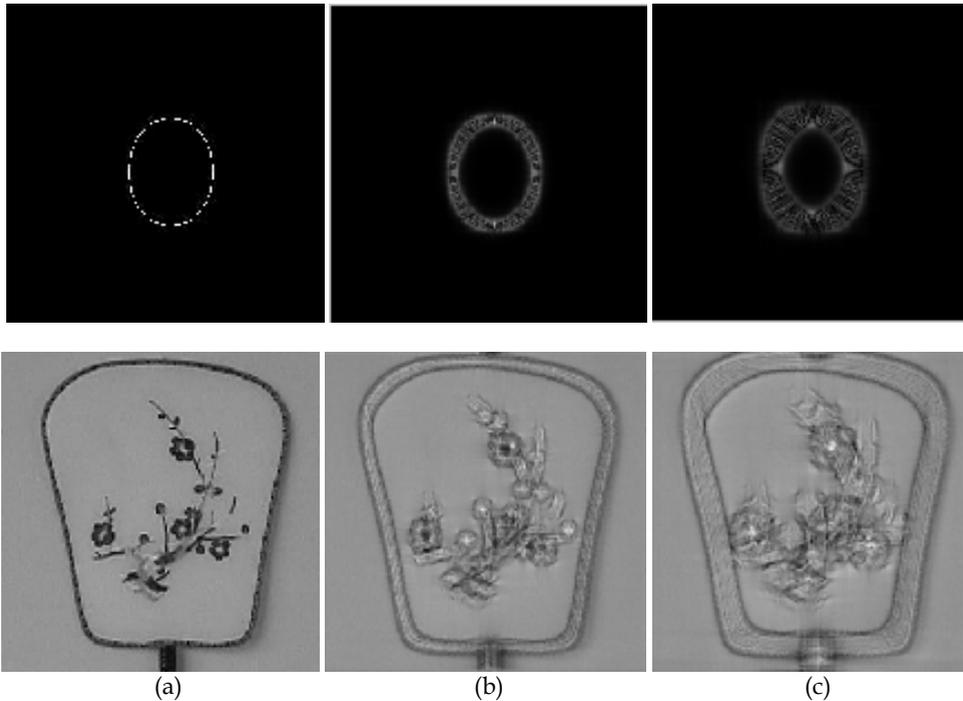
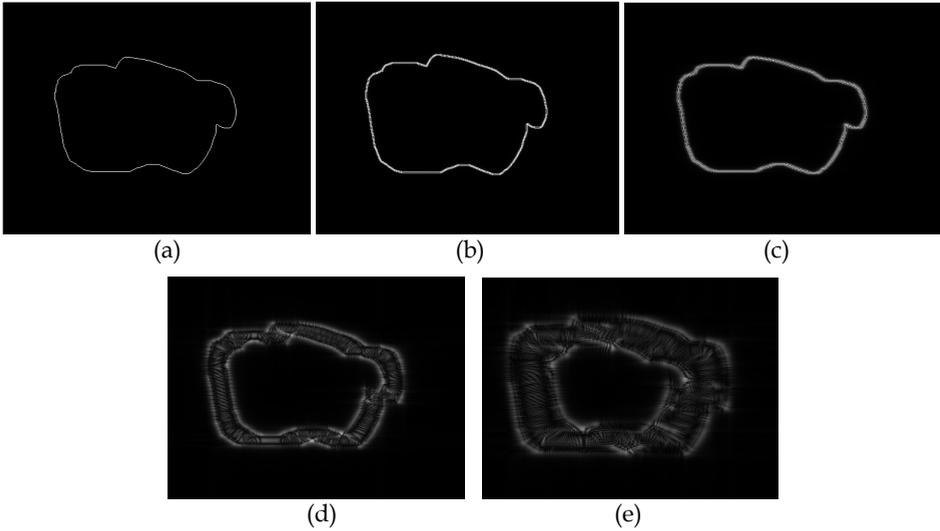


Fig. 3. I-type Schrödinger transform of image. (a) The original image, (b), (c) are Schrödinger transforms. The constant  $at$  is 0.0005, 0.001, respectively.



(a)original image, (b), (c), (d) and (e) are Schrödinger transforms with parameters 0.00001, 0.00005, 0.0005 and 0.001, respectively.

Fig. 4. Schrödinger transforms of a irregular closed image with different parameters  $at$

#### 4. Scale parameter and potential function of Schrödinger transform

It is important to select the parameter  $at$  of Schrödinger transformation while the potential function is selecte. In the section, we'll discuss the effects of scale parameter and potential function for Schrödinger Transform of image.

##### 4.1 Schrödinger transform of rectangle image

Fig. 2 shows the parameter  $at$  is scale parameter of Schrödinger transformation. To make clear the relation between the parameter  $at$  and Schrödinger transformation, without loss of generality, take Schrödinger transform of rectangle image for example. And, the parameter  $at$  is denoted by the parameter  $a$ .

Let a rectangle image be

$$f(m,n) = \begin{cases} 1, & \text{if } M_1 \leq m \leq M_2, N_1 \leq n \leq N_2 \\ 0, & \text{otherwise} \end{cases} \tag{34}$$

Its Fourier transform is

$$F(u,v) = \sum_{m=M_1}^{M_2} \sum_{n=N_1}^{N_2} \exp(-j2\pi(\frac{mu}{M} + \frac{nv}{N}))$$

$$= \frac{(e^{-j2\pi u M_1/M} - e^{-j2\pi u(M_2+1)/M})(e^{-j2\pi v N_1/N} - e^{-j2\pi v(N_2+1)/N})}{(1 - e^{-j2\pi u/M})(1 - e^{-j2\pi v/N})} \tag{35}$$

The transfer function of two-dimensional discrete Schrödinger transform is

$$H(u, v) = e^{-ja((u-M/2)^2 + (v-N/2)^2)} \tag{36}$$

$$G(u, v) = F(u, v)H(u, v) = \frac{\left( e^{-j2\pi u M_1/M} - e^{-j2\pi u (M_2+1)/M} \right) e^{-ja(u-M/2)^2}}{1 - e^{-j2\pi u/M}} \times \frac{\left( e^{-j2\pi v N_1/N} - e^{-j2\pi v (N_2+1)/N} \right) e^{-ja(v-N/2)^2}}{1 - e^{-j2\pi v/N}}$$

Then, Schrödinger transform of rectangle image  $f(m, n)$  is

$$g(m, n) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} G(u, v) \exp(j2\pi(\frac{mu}{M} + \frac{nv}{N}))$$

$$= \frac{1}{M} \sum_{u=0}^{M-1} \frac{\left( e^{-j2\pi u M_1/M} - e^{-j2\pi u (M_2+1)/M} \right) e^{-ja(u-M/2)^2} e^{j2\pi mu/M}}{1 - e^{-j2\pi u/M}}$$

$$\times \frac{1}{N} \sum_{v=0}^{N-1} \frac{\left( e^{-j2\pi v N_1/N} - e^{-j2\pi v (N_2+1)/N} \right) e^{-ja(v-N/2)^2} e^{j2\pi nv/N}}{1 - e^{-j2\pi v/N}}$$

$$= g_{M, M_1, M_2}(m) g_{N, N_1, N_2}(n)$$

Here  $g_{M, M_1, M_2}(m) = \frac{1}{M} \sum_{k=M_1}^{M_2} \sum_{u=1}^M e^{-j2\pi u(k-m)/M} e^{-ja(u-M/2)^2}$ .

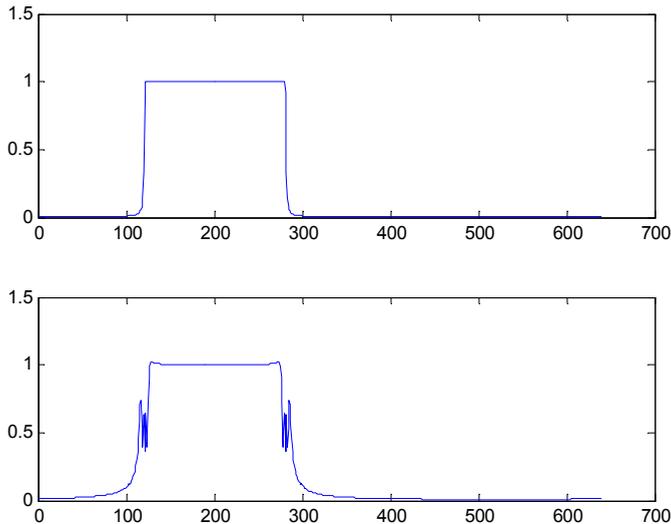


Fig. 5. The function  $|g_{640,120,280}(m)|$ .  $a = 0.00001$  and  $0.0001$ , respectively

**4.2 Scale parameter of Schrödinger transform**

For given threshold  $T(=0.5)$ , denote

$$d = \frac{1}{2} \left( M_1 + 1 - \min \{ m : g_{M, M_1, M_2}(m) > T \} + \max \{ m : g_{M, M_1, M_2}(m) > T \} - M_2 \right) \quad (37)$$

Then, the variable  $d$ , which depend on the constant  $a$ , represents the propagation distance of I-type Schrödinger transform(see Fig. 6). The following experiment shows that the bigger

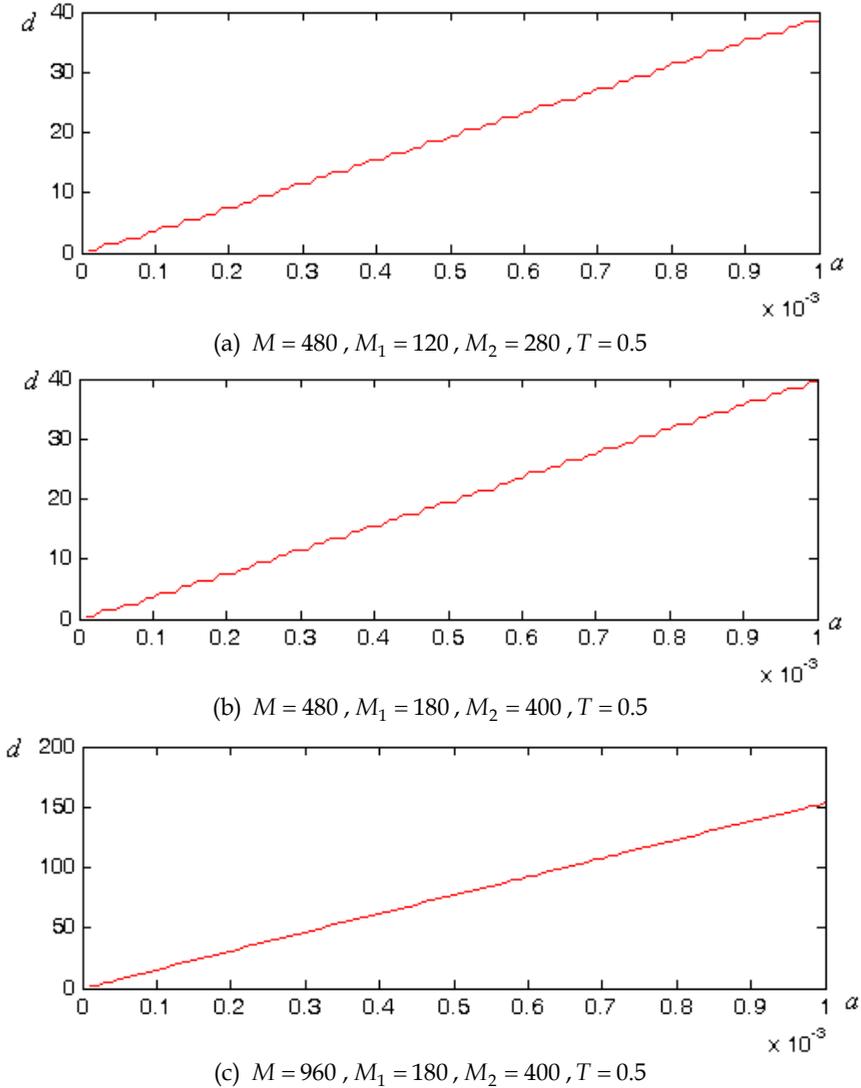


Fig. 6. The relationship between propagation distance  $d$  and scale parameter  $a$ .

$a$  is, statistically speaking, the bigger  $d$  is. So, the constant  $a$  (or  $at$ ) is the Scale Parameter of Schrödinger Transform. In the experiment, we investigate the relation between propagation distance and the constant  $a$ , size of image  $M$ , size of object  $M_2 - M_1$ . The variable  $M$  varies from 480 to 900 with the increment 10,  $M_2$  from  $M_1 + 10$  to  $M - 100$  with the increment 2, and  $a$  from 0.00001 to 0.0001 with the increment 0.00001 while  $M_1 = 170$  and the distance threshold  $T = 0.5$ . We can obtain 88580 samples using Equation (37). The Table 1 lists some descriptive statistics of sample datas. Pearson correlation coefficients between propagation distance  $d$  and  $a$ ,  $M$ ,  $M_2 - M_1$  is 0.823, 0.494 and 0.224, respectively. The multiple regression procedures of SPSS estimate a linear equation of the form:

$$d = -9.625 + 0.13M + 0.01(M_2 - M_1) + 90249.047a \quad (38)$$

And standard error of the estimate, which is less than one pixel, is 0.88556. From Fig. 6 and Equation (38), we know the biggest affect on propagation distance  $d$  is the scale parameter  $a$ , size of image  $M$  takes second place while the propagation distance  $d$  and scale parameter  $a$  have a strong positive linear correlation, the propagation distance  $d$  and size of image  $M$  have a weak positive linear correlation,.

variable	Mean	Std. Deviation
$d$	4.8186	3.15129
$M$	727.38	118.334
$M_2 - M_1$	233.69	146.615
$a$	.00005500	.000028723

Table 1. Descriptive Statistics

**Remark:** Schrödinger Transform of image can be directly computed for a little number  $a$  by using Equation (14) while it needs to use Equation (14) repeatedly for a big unumber  $a$ , and use a little scale parameter every times, so that interference effect of Schrödinger Transform can be avoided for using a big scale parameter  $a$ .

### 4.3 Potential function of Schrödinger transform

It doesn't need to use potential function for I-type Schrödinger Transform of image, which is isotropic. II-type Schrödinger Transform is anisotropic since nonzero potential function is applied. It is necessary that a right potential function is chosed so that we can obtain a perfect deformation processing according to anisotropic property of II-type Schrödinger Transform. However, if the potential function  $D_V$  is real, from Equation (29), we know

$$|\mathbb{S}_V(\varphi)(\mathbf{x})| \approx |\varphi_N(\mathbf{x})| = \left| e^{-itD_V/N} \mathbb{S}(\varphi_{N-1})(\mathbf{x}) \right| = |\mathbb{S}(\varphi_{N-1})(\mathbf{x})| \quad (39)$$

That is, II-type Schrödinger Transform of image  $\varphi(\mathbf{x})$  based on  $v(\mathbf{x})$ ,  $\mathbb{S}_v(\varphi)$ , is only equivalent to implement I-type Schrödinger Transform  $N$  times for image  $\varphi$ . Obviously, that is surely not the result we want. So, the potential function  $D_V$  of II-type Schrödinger Transform must be imaginary so that  $\left| e^{-itD_V/N} \right| \neq 1$ . Meanwhile, if we hope that deformation

processing of II-type Schrödinger Transform stop in some place,  $|e^{-itD_V/N}|$  will be small in the place. Hence, if we hope that deformation processing of II-type Schrödinger Transform will be carried through in homogeneous region of image while it will stop in edge region of image, we can select the potential function of I-type Schrödinger Transform as the following formula:

$$D_V = -ih(G(v(\mathbf{x}))) \quad (40)$$

Where  $G(\cdot)$  is arbitrary gradient or edge-detection operator,  $h(\cdot)$  is a monotone increasing function.

## 5. Application of Schrödinger transform

As a new tool for image analysis, Schrödinger Transform of image can be applied to image smoothing, image enhancement, contour extraction, image inpainting and so on. And Schrödinger Transform of image can also be applied to image segmentation for 3D image. In the section, we'll give some examples.

### 5.1 Constructing high and low pass filter using I-Type Schrödinger transform of image

By the definition and formula of Schrödinger transform of image, transformed image has the same size as the original image. In transformed image, every pixel is related to that of all pixels of original image. So Schrödinger transform is global. High and low pass filter of image, which can detect image edge, enhance and smooth image, can be constructed according to Property 2 and 3 of Schrödinger transform. The filter consider both global and local feature of image while the traditional edge detection operators use only local template.

Suppose Schrödinger transform of  $m \times m$  image  $\phi(\mathbf{x})$  be  $u(\mathbf{x}, t)$ . Then,  $\frac{1}{2}|\phi(\mathbf{x}) + u(\mathbf{x}, t)|$  is the

smoothed image of image  $\phi(\mathbf{x})$  while  $|\phi(\mathbf{x}) - u(\mathbf{x}, t)|$  is the edge image of  $\phi(\mathbf{x})$  if  $at \leq \frac{\pi}{2m^2}$ .

However,  $u(\mathbf{x}, t)$  is a complex-valued image, the result is not good using the above high and low pass filtering directly. Replacing  $|\phi(\mathbf{x}) - u(\mathbf{x}, t)|$  with  $|\phi(\mathbf{x}) - |u(\mathbf{x}, t)||$  has better high pass

filtering effect. Similarly, the smoothing effect of  $\frac{1}{2}|\phi(\mathbf{x}) + |u(\mathbf{x}, t)||$  is better than

$\frac{1}{2}|\phi(\mathbf{x}) + u(\mathbf{x}, t)|$ . The steps of high pass filtering of image by Schrödinger transform are as

follows(the steps of low-pass filtering is similar):

1. Compute Fourier transform  $\phi(\mathbf{y})$  of image  $\phi(\mathbf{x})$ ;
2. Compute Fourier transform  $u(\mathbf{y}, t)$  of  $u(\mathbf{x}, t)$ , according to  $u(\mathbf{y}, t) = \phi(\mathbf{y})e^{-ait|\mathbf{y}|^2}$ ;
3. Compute inverse Fourier transform  $u(\mathbf{y}, t)$  to obtain Schrödinger transform  $u(\mathbf{x}, t)$  of image  $\phi(\mathbf{x})$ ;
4. Compute  $|\phi(\mathbf{x}) - |u(\mathbf{x}, t)||$  to obtain high pass filtered image;
5. Enhancing edge of image  $\phi(\mathbf{x})$  by  $\phi(\mathbf{x}) + |\phi(\mathbf{x}) - |u(\mathbf{x}, t)||$ .

We should pay attention to two issues using the above method to detect image edge, enhance image and smooth image:

1. The parameter  $at$  should be appropriate. If the parameter is too small, the filtering effect would not be obvious. Contrariwise, Schrödinger transform will cause interference which effects the filtering. The parameter  $at$  should not exceed  $\frac{\pi}{2M^2}$ . In the following experiments, Schrödinger transform of image is completed by using Schrödinger transform with smaller  $at$  repeatedly so that the interference can be avoided.
2. The origin of coordinates of frequency domain is the center rather than the top left corner of the image when the formal  $u(\mathbf{y}, t) = \phi(\mathbf{y})e^{-ait|\mathbf{y}|^2}$  is used to compute Fourier transform  $u(\mathbf{y}, t)$  of  $u(\mathbf{x}, t)$ .

We give the results of image edge detection by using Schrödinger Transform in Fig. 7. The experiments show the comparison results of a fan image by several edge detection operators in Fig. 8. According to the comparison results of traditional edge detection operators, the high-pass filter designed by Schrödinger transform can better detect image edge and it would not increase noise simultaneously. In fact, filter designed by Schrödinger transform consider both the local and global feature of image, so filtering effect is better.

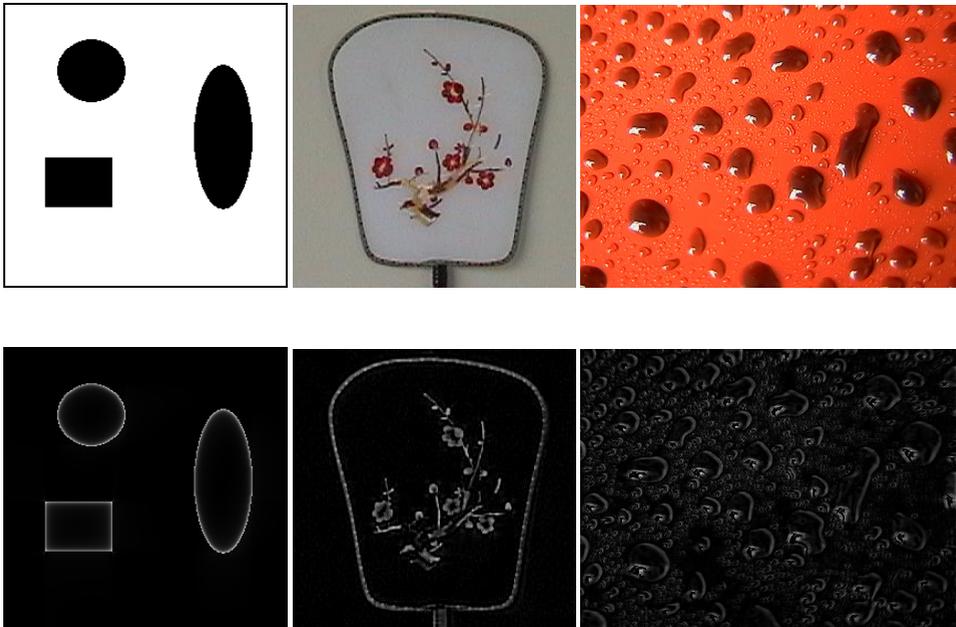
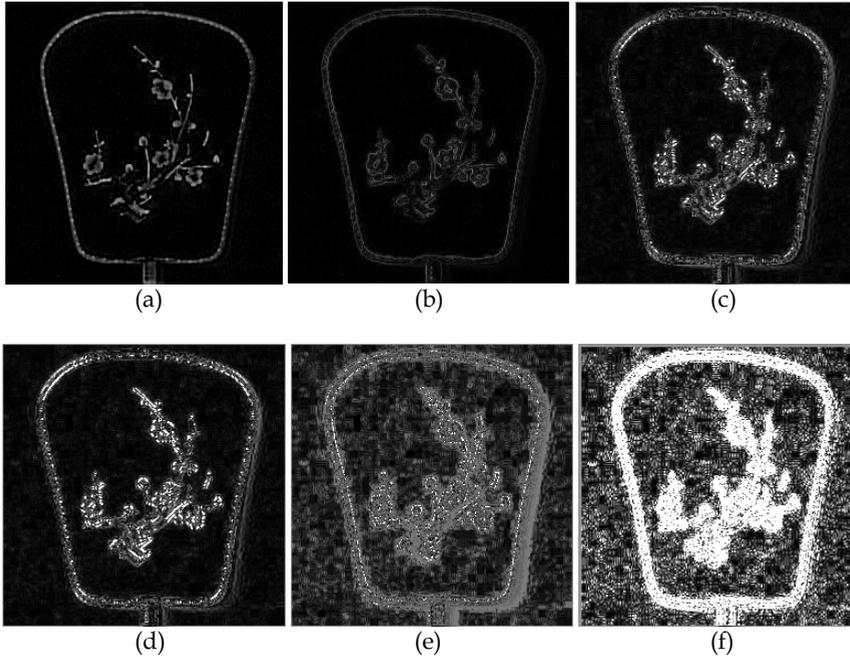


Fig. 7. Edges detected using Schrödinger Transforms of image



(a) High pass filter based on Schrödinger Transform of Image, (b) Robert operator, (c) Prewitt operator, (d) Sobel operator, (e) Canny operator, (f) Laplace operator

Fig. 8. Comparison of several edge detection methods

## 5.2 Automatic contour extraction for multiple objects using I-Type Schrödinger transform of image

### 5.2.1 Exterior and interior points of multiple objects

Assume  $I(x, y)$  be an original image that contains multiple dark objects. We use Schrödinger Transforms of original image  $I(x, y)$  and its inverse image  $255 - I(x, y)$  to obtain exterior and interior of multiple objects. So we can separate dark objects from bright background. Let  $u_1(\mathbf{x}, t)$ ,  $u_2(\mathbf{x}, t)$  be Schrödinger transforms of original image and its inverse image, respectively. To obtain enough exterior and interior points exactly, in the following experiments, we implement Schrödinger transforms of image many times using small parameter  $at$ . Denote

$$I_1(x, y) = |I(x, y) - u_1(\mathbf{x}, t)|$$

$$I_2(x, y) = |255 - I(x, y) - u_2(\mathbf{x}, t)|$$

Then, according to the Proposition 1, these pixels with high gray values in the image  $I_1(x, y)$  are interior points of multiple objects while these pixels with high gray values in the image  $I_2(x, y)$  are exterior points of multiple objects. To obtain interior and exterior points of multiple objects, we convert an intensity image to a binary image using Otsu's threshold method.

### 5.2.2 Contour extraction of multiple objects

After obtaining exterior and interior points of multiple objects, we take the following steps to extract contours of multiple objects:

1. Thinning of sets of exterior and interior points.  
Sets of exterior and interior points need to be thinned because there are too many points in the two sets. Generally speaking, the exterior points surround the interior points for the same object. For each point  $A$  in the set of exterior points(or interior points), we could find the point  $B$  which is the point with the shortest distance between the point  $A$  and the set of interior points(or exterior points). The point  $B$  shall be added into the thinned set of interior points(or exterior points).
2. Find pairs of exterior and interior points with the smallest distance between them.  
The closest pair of points problem is a problem of computational geometry. Using the brute-force algorithm, the closest pair of points can easily be computed in  $O(n^2)$  time. To do that, one could compute the distances between all pairs of points, then pick the pair with the smallest distance. The problem can also be solved in  $O(n \log n)$  time using the recursive divide and conquer approach (Cormen et al., 2001).

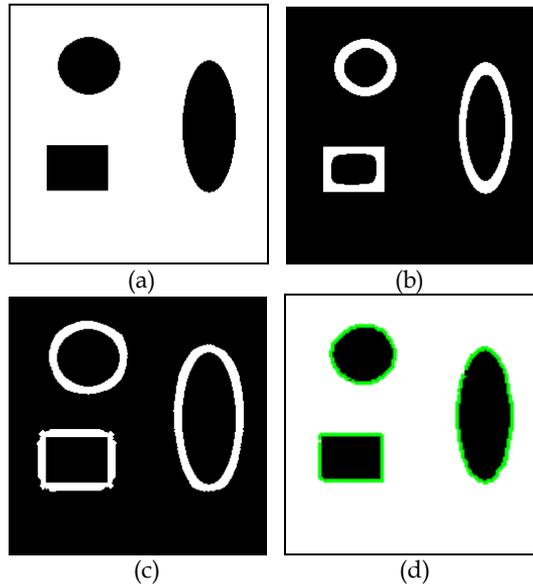
### 5.2.3 Algorithm

The actual operation of the algorithm is as follow:

1. If necessary, convert an original gray image to a binary image using Otsu's threshold method.
2. Compute Schrödinger transforms  $u_1(\mathbf{x}, t)$ ,  $u_2(\mathbf{x}, t)$  of original image and its inverse image.
3. Compute  $I_1(x, y) = |I(x, y) - u_1(\mathbf{x}, t)|$ ,  $I_2(x, y) = |255 - I(x, y) - u_2(\mathbf{x}, t)|$ .
4. Find interior and exterior points of multiple objects by converting images  $I_1(x, y)$  and  $I_2(x, y)$  to binary images using Otsu's threshold method.
5. Thin sets of exterior and interior points.
6. Find pairs of exterior and interior points by using the brute-force algorithm. If the distance between the pair points is larger than some constant, a new object will appear.
7. Obtain contours of multiple objects by connecting all pairs of points belong to the same object in turn.

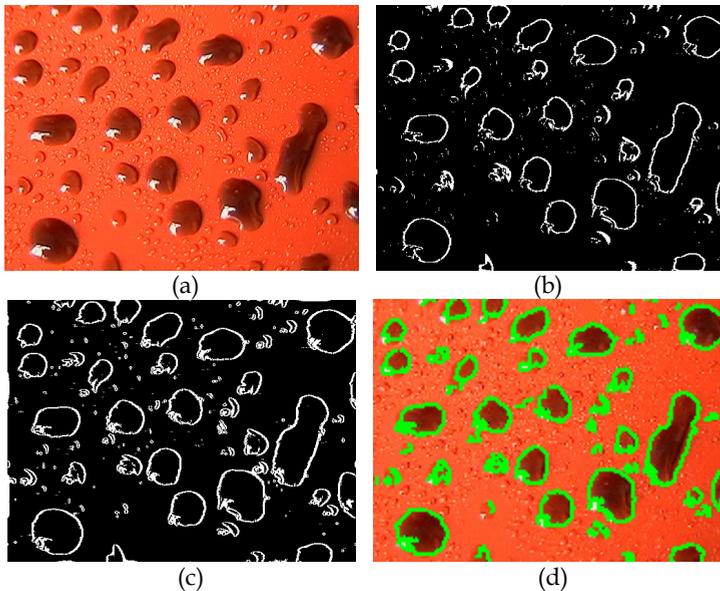
### 5.2.4 Experiments

To evaluate the performance of proposed scheme by experiments, we do Schrödinger transform and contour extraction experiments using the simulated and real images. Schrödinger transform of image can separate boundaries of object from background. Contour extraction experiments for the simulated images consisting of three objects are shown in Fig. 9. With growing pollution levels and an ever growing coastal population, hydrophobicity (water repellence) of insulators has become an important consideration for equipment specifiers. An insulator with hydrophobic properties causes water to bead on its surface and roll off assisting pollution to freely wash away rather than forming a continuous wet sheet or zone and combining with pollutants to form a conductive film. The hydrophobic effect helps to reduce discharge activity and maintenance requirements as well as decreasing the probability of flashover. Hydrophobicity demonstrated by beading on the surface of a new polymer insulator(see Fig. 10(a)). The extracted contours of beads using Schrödinger transform of image are shown in Fig. 10(d).



(a) The original simulated image, (b) The interior points of objects (white pixels), (c) The exterior points of objects (white pixels), (d) the extracted contours.

Fig. 9. The extracted contours of multiple objects using Schrödinger transform of image.

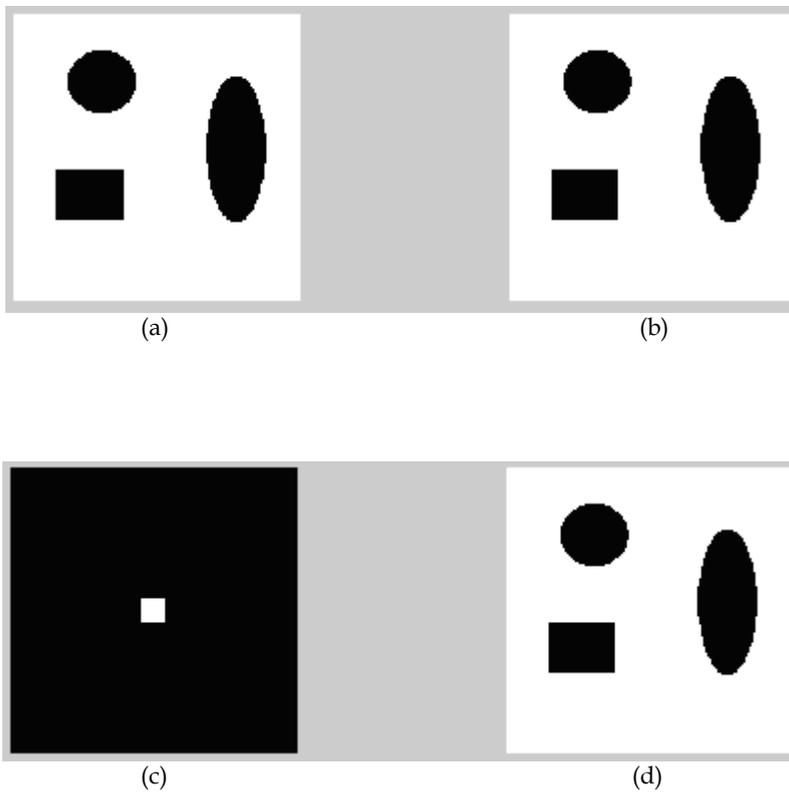


(a) beads on the surface of a polymer insulator, (b) The interior points of objects (white pixels), (c) The exterior points of objects (white pixels), (d) The extracted contours of beads.

Fig. 10. The extracted contours of beads using Schrödinger transform of image.

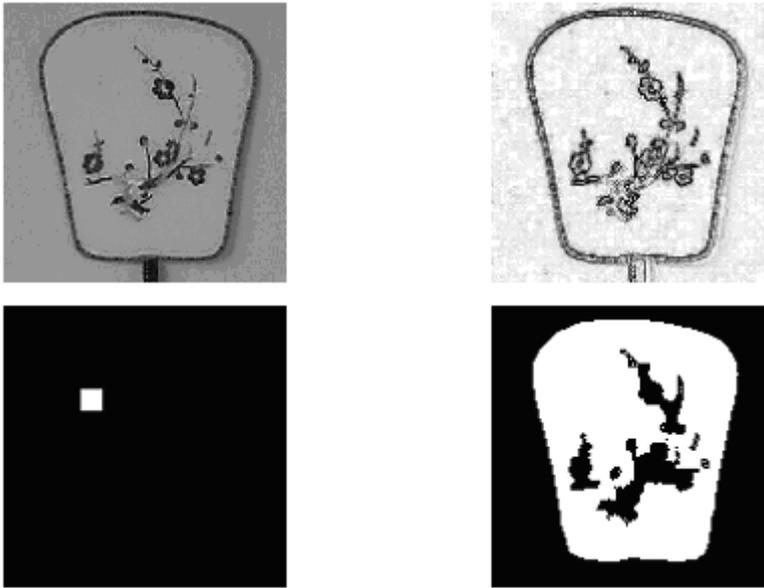
### 5.3 Image segmentation using II-Type Schrödinger transform of image

The following experiments show how to segment the target area by using II-Type Schrödinger Transform. Fig. 11 shows segmentation results of an image with three goals, Fig. 12 shows object segmentation results of a fan image. The two results show that we can extract the contour of object using evolution of II-Type Schrödinger Transform of image. The segmentation results are less depending on the initial contour, as long as be the inside or outside the target area.



(a) original image, (b) image of potential function,  
(c) initial region of object, (d) segmented region of object

Fig. 11. segmented result images of three objects by using II-type Schrödinger transform of image



(a) original image, (b) image of potential function,  
(c) initial region of object, (d) segmented region of object

Fig. 12. segmented result of fan image by using II-type Schrödinger transform of image

## 6. Conclusion

Image segmentation has been, and currently still is, a relevant research area in Computer Vision, and hundreds of segmentation algorithms have been proposed in the last few decades. However, it is well known that elemental segmentation techniques based on edge or region information often fail to produce accurate segmentation results. The chapter attempts to provide a brief introduction of Schrödinger transform of image for the novel segmentation techniques and discussed its potential applications in image analysis and understanding. Schrödinger transform of image is extended from I-type to II-type. Schrödinger Transform of image is applied to image smoothing, image enhancement, contour extraction, and so on. The demonstrate experiment results show the robust of the algorithm.

## 7. Acknowledgment

This work was supported by the National Natural Science Foundation of China under Grant No. 60975011. The research work was also partially supported by the Basic and Frontier Technology Research Programs of the Department of Science & Technology of Henan Province (No. 092300410043), and Supported by Foundation for University Youths Key Teacher by the Education Department of Henan Province; and a grant from the Nature Science Basic Research Programs of the Education Department of Henan Province (No. 2007510023), and Supported by Foundation for key project by the Education Department of Hubei Province(D20101502), China.

## 8. References

- Caselles, V.; Catta, F.; Coll, T. & Dibos, F. (1993). A Geometric model for active contours. *Numer. Math.*, Vol.66, (1993), pp. 1-31, ISSN 0945-3245
- Caselles, V.; Kimmel, R. & Sapiro, G. (1995). Geodesic active contours. *Proc. 5th Intl. Conf. Computer Vision*, pp.694-699, ISBN 0-8186-7042-8, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, June 20-23, 1995
- Chan, T & Vese, LA. (1999). An active contour model without edges, *Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision*, pp. 141-151, ISBN 354066498X, Corfu, Greece, September 26-27, 1999
- Chan, T & Vese, LA. (2002). Active contour and segmentation models using geometric PDE's for medical imaging, In: *Geometric methods in bio-medical image processing*, Ravikanth Malladi (Editor), pp. 63-75, Springer, ISBN 9783540432166. Berlin, German
- Cohen, L. D. & Cohen, I. (1993). Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.15, No.11, (1993), pp. 1131-1147, ISSN 0162-8828
- Cohen, L. D. & Kimmel, R. (1997). Global minimum for active contour models: a minimal path approach. *Int. J. Comput. Vis.*, Vol. 24, No.1, (1997), pp. 57-78, ISSN 0920-5691
- Cohen, L. D. (1991). On active contour models and balloons. *CVGIP: Image Understand.*, Vol.53, No.2, (1991), pp. 211-218, ISSN 1049-9660
- Cormen, T. H.; Leiserson, C. E.; Rivest, R. L. & Stein, C. (2001). *Introduction to Algorithms* (Second Edition), MIT Press and McGraw-Hill, ISBN 0-262-03293-7., Cambridge, England
- Evans, L. C. (1998). *Partial Differential Equations*, American Mathematical Society, ISBN 0821807722, Providence, USA
- Feynman, R. P. & Hibbs, A. R. (1965). *Quantum Mechanics and Path Integrals*, McGraw-Hill Inc., ISBN 0070206503, New York, USA
- Kass, M.; Witkin, A. & Terzopoulos, D. (1987). Snakes: active contour models. *Int. J. Comput. Vis.*, Vol.1, No.4, (1987), pp. 321-331, ISSN 0920-5691
- Lobergt, S. & Viergever, M. A. (1995). A discrete dynamic contour model. *IEEE Trans. Med. Imaging*, Vol.14, No.1, (1995), pp. 12-24, ISSN 0278-0062
- Lou, Liantang & Ding, Mingyue (2007a). Principle and approach of boundary extraction based on particle motion in quantum mechanics. *Optical Engineering*, Vol.46, No.2, (2007), pp. 027005-1- 027005-16, ISSN 0091-3286
- Lou, Liantang; Fu, Zhongliang & Jiang, Si (2007b). Maximal probability method of boundary extraction based on particle motion. *Proc. SPIE, MIPPR 2007*, pp. 67861F.1-67861F.5, ISBN 9780819469502, Wuhan, Hubei, China, November 15-17, 2007
- Lou, Liantang; Zhan, Xin; Fu, Zhongliang & Ding, Mingyue (2008). Method of Boundary Extraction Based on Schrödinger Equation, *Proceedings of the 21th Congress of the International Society for Photogrammetry and Remote Sensing - ISPRS 2008*, B5-2, pp. 813-816, ISSN 1682-1750, Beijing, China, July 3-11, 2008
- Malladi, R. & Sethian, J.A. (1996). Level set and fast marching methods in image processing and computer vision, *Proceedings of the IEEE International Conference on Image Processing*, Vol.1, pp. 489 - 492, ISBN 0-7803-3259-8, Lausanne, Switzerland, Sep., 16-19, 1996

- McInerney, T. & Terzopoulos, D. (1999). Topology adaptive deformable surfaces for medical image volume segmentation. *IEEE Trans. Med. Imaging*, Vol.18, No.10, pp. 840-850, (1999), ISSN 0278-0062
- Osher, S. & Sethian, J. (1988). Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput Phys.*, Vol.79, No.1, pp. 12-49, (1988), ISSN
- Terzopoulos, D.; Witkin, A. & Kass, M. (1987). Constraints on deformable models: recovering 3D shape and nonrigid motion. *Artif. Intell.*, Vol.36, No.1, (1987), pp. 91-123, ISSN 0004-3702
- Williams, L.R. & Jacobs, D.W. (1995). Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Saliency, *Proc. 5th Intl. Conf. Computer Vision*, pp.694-699, ISBN 0-8186-7042-8, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, June 20-23, 1995
- Xu, CY & Prince, JL. (1998). Snakes, shapes, and gradient vector flow. *IEEE Trans Image Process*, Vol.7, No.3, pp. 359-369, (1998), ISSN

# Entanglement in Two and Three Quantum Mechanical Systems

L. M. Arévalo Aguilar, M. M. Méndez Otero and P. Mastranzo Ortega  
*Universidad Autónoma de Puebla*  
*México*

## 1. Introduction

Entanglement is one of most peculiar characteristic of Quantum Mechanics (probably the other most mentioned is the particle tunneling through a barrier) because it entails correlations that supersedes classical correlations. It is conceived as the genuine quantum characteristic by many researchers and quite recently it has emerged as a physical resource to produce non-classical task, like quantum teleportation, quantum cryptography and quantum information. Entanglement could be distributed (efficient quantum communication is equivalent to efficient entanglement distribution (Plenio & Virmani, 2007)), concentrated (given an amount of entanglement, equal to the full content of entanglement in  $n$  pairs of identical entangled pure state systems, it is possible, using local operations on each system, to concentrate the total amount of entanglement into a smaller number  $m$ ,  $m < n$ , of maximally pure entangled state system (Bennett, et al.; 1996)) and used to perform many quantum information tasks useful to overcome technical restrictions present on classical communication.

In this chapter, we review the entanglement of quantum mechanical systems. First, in Section 2, we give a characterization of entanglement in terms of its special features as resource and its mathematical structure. Then, in section 3, we review some of the Bell inequalities. In Sections 4 and 5, we review some of the most important entanglement measures published for two and three entangled systems. After that, in Section 6, we give a characterization of quantum gates and operators by its entanglement power, i. e. the amount of entanglement that they can produce when acting on an arbitrary state. Finally, in Section 7, we briefly review the experimental detection of entanglement.

## 2. What is entanglement?

In this Section we define entanglement by its mathematical structure and talk about the operational entanglement definition. We show some important issues of the mathematical description of entanglement, in terms of density operator, that serves as a base to the development of the following sections of the chapter. The question: What is entanglement? is one of the most quite difficult and subtle one that we can ask, then it is a naive thought to think that it is an easy task to answer it. Then, what we propose as an answer to that question is just a characterization of entanglement by both its mathematical structure and its operational features in terms of nonlocal properties.

Entanglement could be mathematically or operationally defined. Operational definition of entanglement uses the concept of *Local Operations and Classical Communication (LOCC)*

(Bennett, et al.; 1996). LOCC is a restriction to use only local operation and classical communications; that is to say, it is not allowed to exchange any quantum system nor to perform any nonlocal operation (Bennett, et al.; 1996). In this case, entanglement is conceived as the resource that allows to overcome the restriction of LOCC. Keeping this line of thought, classical correlations are those generated by LOCC operations and the quantum ones are those that cannot be simulated classically, see (Plenio & Virmani, 2007) page 3. The operational definition of entanglement is quite useful because it allows both to understand entanglement as a physical resource and gives the background to define some measurements of entanglement, like the entropy of entanglement.

Formally, mathematically, but roughly speaking, we can define entanglement as quantum states that can not be writing as the product of two wave functions, that is, for two two-state systems we have, for example:

$$|\psi\rangle_{12} = |\phi_0\rangle_1 (c_1|\xi_0\rangle_2 + c_2|\xi_1\rangle_2) \rightarrow \text{un-entangled}, \quad (1)$$

$$|\varphi\rangle_{12} = c (|\phi_0\rangle_1|\xi_0\rangle_2 + |\phi_1\rangle_1|\xi_1\rangle_2) \rightarrow \text{entangled}. \quad (2)$$

Where  $|\phi_0\rangle$ ,  $|\xi_0\rangle$ ,  $|\phi_1\rangle$  and  $|\xi_1\rangle$  are basis states, and  $c_1$ ,  $c_2$  and  $c$  are normalization constant. The subindices refer to system 1 and system 2. **We can consider, also, that system 1 and system 2 are situated far away in two different places.** Usually, in quantum information theory, this is stated as: System 1 belongs to Alice and system 2 belongs to Bob. Physically we can say regarding Eq. (1) that each quantum system have its own state, whereas in Eq. (2) each systems does not have its own state, but share (although they are far away) a global state. However, it is possible to assign a density matrix with each system, this could be seen more clearly if we look at the density operator of the state given by Eq. (2):

$$\hat{\rho}_{12} = |\varphi\rangle_{12} \langle\varphi| = |c|^2 \left\{ |\phi_0\rangle_{11}\langle\phi_0| \quad |\xi_0\rangle_{22}\langle\xi_0| + |\phi_0\rangle_{11}\langle\phi_1| \quad |\xi_0\rangle_{22}\langle\xi_1| + \right. \\ \left. |\phi_1\rangle_{11}\langle\phi_0| \quad |\xi_1\rangle_{22}\langle\xi_0| + |\phi_1\rangle_{11}\langle\phi_1| \quad |\xi_1\rangle_{22}\langle\xi_1| \right\}. \quad (3)$$

Now, in order to obtain the density operator for one system, we can take the partial trace over one system of the density operator given in Eq. (3), for instance taking the trace over system 2 we obtain the density operator for system one, as follow:

$$\hat{\rho}_1 = \text{Tr}_2\{\hat{\rho}_{12}\} = |c|^2 \{ |\phi_0\rangle_{11}\langle\phi_0| + |\phi_1\rangle_{11}\langle\phi_1| \}, \quad (4)$$

Eq. (4) is a mixed state. This mean that although systems 1 and 2 does not have its own wave function, as it is stated in Eq. (2), yet it is possible to assign a density operator with each system. In other words, and contrary to Eq. (1), an entangled state seems to each party to be a mixed state Bennett (et al.; 1996).

In fact, any entangled two two-state systems could be writing in terms of its Schmidt form:

$$|\varphi\rangle = \sum_{i=1} \sqrt{\lambda_i} |e_i\rangle \otimes |h_i\rangle \quad (5)$$

where,  $|e_i\rangle$  and  $|h_i\rangle$  are ortogonal vectors in the Hilbert spaces of each system, with dimensions  $m$  and  $n$ , respectively. Hence, the density matrix for each system could be written as, for a proof see (Hughston, et al.; 1993):

$$\hat{\rho}_1 = \sum_{i=1}^m \lambda_i |e_i\rangle \langle e_i|, \quad \hat{\rho}_2 = \sum_{i=1}^n \lambda_i |h_i\rangle \langle h_i|. \quad (6)$$

On the other hand, if  $\hat{\rho}_x$  correspond to a product state, i.e. non-entangled, then

$$\hat{\rho}_x^2 = \hat{\rho}_x, \quad (7)$$

where  $x = 1, 2$ . That is to say, for two systems an entangled state does not fulfill Eq. (7). You can check this using Eqs. (1) and (3), for a general proof see (Popescu & Rohrlich, 1998).

Many intriguing phenomena come from the peculiarities of Eq. (2). For example, in classical mechanics it is not possible to influence the result of a measurement in one system by acting in another system that no longer interacts with the former system, however entangled states make this possible for quantum physical systems. This is the *non-locality* feature of quantum mechanical systems, although entanglement is not the same that non-locality Méthot & Scarani (2007).

### 3. J. S. Bell insight on the physical properties of entangled systems

The physical properties of entangled states represents a high departure from the properties of classical systems. Besides of being the building blocks to construct many non-classical task, entanglement serves as theoretical tool to understand and prove many genuine quantum properties. Then, in this section, we review some of the first noticed physical peculiarities of entanglement based in the insight of J. S. Bell, who first notices the way to test classical correlations against quantum correlations.

In the year 1935 Einstein, Podolsky and Rosen (Einstein, et al.; 1935), introduced an argument that, if is taken seriously, questioned the quantum mechanical description of physical phenomena. This argument is widely know as the EPR paradox. For almost thirty years the EPR argument remains as a philosophical issue, until J. S. Bell deduced some predictions that could be experimentally tested. These predictions were based on the great insight that Bell developed about the quantum correlations that entangled states have.

Briefly, the EPR argument was based on two quite reasonable assumptions (Einstein, et al.; 1935):

1. *If, without in any way disturbing a system, we can predict with certainty (i. e., with probability equal to unity) the value of a physical quantity, then there exist an element of physical reality corresponding to this physical quantity.*
2. *... the absence of an interaction between the two systems.*

The first assumption was aimed at establishing that a theory should be complete; of course it had have strong influence on the conceptions of the physical reality; that is to say, that physical variables have predetermined values independent of measuring them or not. The second, intended to consider as valid only physical process in which the measurement in one quantum system does not affect the result of measurement done on another quantum system when they do not interact and are far away from each other. Based on these assumptions, an in the analysis of measurement's results made on entangled states correlated in position and momentum, Einstein et. al. conclude (Einstein, et al.; 1935): "*While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such description exist. We believe, however, that such a theory is possible.*" This conclusion has been interpreted by physicists, if it is true, as implying that quantum mechanics is not a complete theory and that it must be supplemented with some hidden variables. In short, what Einstein, et al., were looking for was to determine whether or not the measurement result, on two correlated systems, of one observable in system B could be valid predicted using the measurement result of one observable on system A, when the

observables, i. e. the operators representing them, does not commute and when there is not any interaction between the systems A and B.

In his work, Bell first show that previously published arguments concerning the impossibility of hidden variables theories were not well supported (Bell, 1966). Then, he assume some desirable requirements, as the condition of locality, that such theory should have (Bell, 1964) -although he show that such theory should be non-local to reproduce the quantum mechanical statistic results-. These requirements produces some predictions, regarding the statistical correlations of observables, that could be experimentally tested (Bell, 1964). These predictions are established as inequalities equations. Below we write down the Clauser, et al. inequality because it has been one of the most influencing inequalities (Clauser, et al.; 1969),

$$|P(a, b) - P(a, c)| + P(b', b) + P(b', c) \leq 2, \quad (8)$$

in Eq. (8), the correlation  $P(i, j) \equiv \int A(i, \lambda)B(j, \lambda)\rho(\lambda)d\lambda$ ;  $A(i, \lambda)$  and  $B(j, \lambda)$  are the result of the measurement (with values  $\pm 1$ ). The average is taken over many experiment measurements, i. e. inside the total  $\lambda$  space. Locality requires that  $A(i, \lambda)$  must be independent of  $j$  and  $B(j, \lambda)$  independent of  $i$ . Inequality given by Eq. (8) could be deduced by appealing to local hidden variables (Clauser, et al.; 1969) or, as was done by Eberhard (Eberhard, 1977), using just local causes. What Bell shown was that quantum mechanics statistical correlations violate an inequality equation similar to that given by Eq. (8).

As it was stated above, Eq. (8) has been deduced under the assumption of local hidden variable theories. A recent study (Gröblacher, et al.; 2007) obtains a similar inequality for a class of non-local hidden variable theories (which is a generalization of one proposed by Leggett (Leggett, 2003)). This group experimentally test the violation of this inequalities, therefore they shown that a wide varieties of non-local hidden variables theories are in contradiction with quantum mechanics predictions.

#### 4. Physical properties and entanglement measures of two quantum mechanical system

In the current years, there is an agreement between many researcher about the distinguishing features of the physical properties and entanglement measures for just two quantum mechanical system. We can say that this agreement is a universal one. Then, in this Section we review some of the most important entanglement measures for two states systems.

After passing the years of philosophical debates, the most studied entangled state correspond to two-entangled systems. This entangled states were experimentally produced and used to test Bell's inequalities. Some quantum states were named after Bell, nowadays widely know as Bell states:

$$\begin{aligned} |\Psi^+\rangle &= \frac{1}{2} (|0\rangle_1|1\rangle_2 + |1\rangle_1|0\rangle_2), \\ |\Psi^-\rangle &= \frac{1}{2} (|0\rangle_1|1\rangle_2 - |1\rangle_1|0\rangle_2), \\ |\Phi^+\rangle &= \frac{1}{2} (|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2), \\ |\Phi^-\rangle &= \frac{1}{2} (|0\rangle_1|0\rangle_2 - |1\rangle_1|1\rangle_2). \end{aligned} \quad (9)$$

We can say that this are probably the most studied and understood states between the entangled states. They are though as the maximally two-entangled states and they could be used as a basis.

The physical properties of this states are remarkable, it allows teleportation of quantum states between separate locations (Bennett, et al.; 1993), were the background to test Bells inequalities (Popescu & Rohrlich, 1998) and they serve as a reference tool to understand many facets of quantum information theory.

#### 4.1 Two state entanglement measures

Nowadays, there is a general consensus about Entanglement measures for two systems, specially for two two-states systems. In fact, there are some proposed desired requirements for these measurements (Horodecki, et al.; 2000; Vedral, et al.; 1997; Vedral & Plenio, 1998). In this subsection we only refer to three entanglement measures: the entropy of entanglement, the concurrence and the geometric measurement of entanglement. For more complete review, see (Horodecki, et al.; 2009; Plenio & Virmani, 2007).

##### 4.1.1 Entropy of entanglement

One of the first proposed entanglement measures was the entropy of entanglement, which for two entangled quantum mechanical system is defined as the von Neumann entropy of either system, i. e.  $\hat{\rho}_1$  or  $\hat{\rho}_2$ , (Bennett, et al.; 1996):

$$E = -\text{Tr}\hat{\rho}_1 \log_2 \hat{\rho}_1. \quad (10)$$

$E$  gives the amount of entanglement for a given pure quantum entangled state, it is one for maximum entangled states and zero for separable states. For a mixed state, it is possible to use  $E$  as a departure to measure its amount of entanglement. In this case the measure of entanglement is called *entanglement of formation*,  $E_f(\hat{\rho})$ , and it is given by (Bennett, et al.; 1996b):

$$E_f(\hat{\rho}) = \min \sum_j p_j E(\Phi_j), \quad (11)$$

where the minimum must be taken over all pure state decomposition of  $\hat{\rho}$ . The operational meaning of the entanglement of formation is that the  $E_f(\hat{\rho})$  given by Eq. (11) gives the number of singlet states required to create  $\hat{\rho}$  (Plenio & Virmani, 2007; Wootters, 2001). Also, it is worth to mention that neither  $E$  given by Eq. (10) nor  $E_f(\hat{\rho})$  given by Eq. (11) do not increase under local operations and classical communications (Bennett, et al.; 1996b).

##### 4.1.2 The concurrence

Perhaps, one of the most widely recognized entanglement measure is the concurrence (Wootters, 2001) and (Wootters, 1998). The concurrence  $C(\Phi)$  is defined as (Wootters, 2001):

$$C(\Phi) = |\langle \Phi | \tilde{\Phi} \rangle|, \quad (12)$$

where  $|\tilde{\Phi}\rangle$  is the spin-flip operation  $|\tilde{\Phi}\rangle = (\sigma_y \otimes \sigma_y) |\Phi^*\rangle$ , and  $|\Phi^*\rangle$  is the complex conjugate of  $|\Phi\rangle$ . An easy way to calculate the concurrence is by using the magic states, which are defined in terms of the bell states given in Eq. (9) (Hill, 1997):

$$\begin{aligned} |e_3\rangle &= i|\Psi^+\rangle = \frac{i}{2} (|0\rangle_1|1\rangle_2 + |1\rangle_1|0\rangle_2), \\ |e_4\rangle &= -|\Psi^-\rangle = -\frac{1}{2} (|0\rangle_1|1\rangle_2 - |1\rangle_1|0\rangle_2), \\ |e_1\rangle &= |\Phi^+\rangle = \frac{1}{2} (|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2), \\ |e_2\rangle &= -i|\Phi^-\rangle = \frac{i}{2} (-|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2). \end{aligned} \quad (13)$$

in terms of the states defined in Eq. (13), the concurrence is defined as (Hill, 1997):

$$C(\Phi) = \left| \sum_i \alpha_i^2 \right| \quad (14)$$

where  $\alpha_i$  are the expansion coefficients of  $|\Psi\rangle = \sum_i \alpha_i |e_i\rangle$ . Notice that what is squared is the  $\alpha_i$  itself, not their absolute value.

#### 4.1.3 The Geometric Measurement of Entanglement (GME)

The Geometric Measurement of Entanglement (GME) was defined by Shimony (Shimony, 1993). The GME uses ideas of Hilbert space geometry (Shimony, 1993; Wei & Goldbart, 2003) and it is related with the relative entropy of entanglement (Wei, et al; 2004). Given an arbitrary bipartite state  $|\psi\rangle$ , the GME is defined as (Shimony, 1993):

$$E(\psi) = \frac{1}{2} \min \| |\psi\rangle - |\phi\rangle \|^2, \quad (15)$$

where  $|\phi\rangle$  is a normalized product state and the minimum is taken over the set of normalized product state (Shimony, 1993). The Eq. (15) determines the entanglement content of  $|\psi\rangle$  by calculating the minimum distance between  $|\psi\rangle$  and the nearest separable state  $|\phi\rangle$ . Eq. (15) could be interpreted as: the more amount of entanglement in a given state, then further away it will be from its nearest un-entangled approximant (Wei & Goldbart, 2003). The GME reaches its greatest value  $\frac{1}{2}$  for a maximal entangled state and zero for un-entangled states.

## 5. Physical properties and entanglement measures of three quantum mechanical systems

In this Section we review some of the important features of three quantum mechanical entangled systems. First we talk about the Bell's theorem for the three party case, then we review the generalizations of Gisin's theorems for three entangled states. After that, we briefly review some of the most important entanglement measures for three quantum mechanical systems. It is not possible to give a complete review of all the entanglement measures, so we left out some others important entanglement measures like the negativity.

The multi-particle entanglement is more complex than two party entanglement, even for the case of three party quantum mechanical systems. For example, just for the three party entangled state case, there are quantum correlations that could not be explained even in the case of perfect correlations (Greenberger, et al.; 1990); in other words, the three entangled states present features not usually encounter in two entangled states which make them the simplest nontrivial entangled states showing fundamental and strong differences compared to that of two entangled systems. Additionally, whereas for two entangled states it is possible to define entanglement measures by mean of a single number<sup>1</sup> giving the amount of entanglement present in the state, such a single number does not exist for three entanglement measures. In fact, to quantify the amount of entanglement present in a given three quantum mechanical system is yet an open problem.

<sup>1</sup> When the state of two quantum mechanical systems is expressed using the Schmidt decomposition, if it is entangled, then, it contains only two different Schmidt number, i. e.  $\sqrt{\lambda_1}$  and  $\sqrt{\lambda_2}$  that are related by  $\lambda_2 = 1 - \lambda_1$ .

### 5.1 Bell's theorems for three quantum mechanical systems

It was shown by Greenberger, et. al. (Greenberger, et al.; 1990) that three quantum entangled mechanical systems reveal inconsistencies between the EPR premises. In particular the expectation value  $E(\phi_1, \phi_2, \phi_3)$  of the outcomes of a measurement of three observables made in the following state:

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|a\rangle|b\rangle|c\rangle + |a^*\rangle|b^*\rangle|c^*\rangle), \quad (16)$$

is given by (Greenberger, et al.; 1990):

$$E(\phi_1, \phi_2, \phi_3) = \sin(\phi_1 + \phi_2 + \phi_3), \quad (17)$$

this expectation values predicts the following perfect correlations:

$$\begin{aligned} E(\phi_1, \phi_2, \phi_3) &= 1 \quad \text{for} \quad \phi_1 + \phi_2 + \phi_3 = \frac{\pi}{2}, \\ E(\phi_1, \phi_2, \phi_3) &= -1 \quad \text{for} \quad \phi_1 + \phi_2 + \phi_3 = \frac{3\pi}{2}, \end{aligned} \quad (18)$$

these correlations are inconsistent with the premises of EPR (Greenberger, et al.; 1990). To show this, set  $\phi_1 = \frac{\pi}{2}$  and  $\phi_2 = \phi_3 = 0$ ; then there could exist variables  $A(\lambda), B(\lambda), C(\lambda)$  in a classical theory such that for a suitable election of apparatus (Żukowski, et al.; 1998):

$$\begin{aligned} A(\pi/2)B(0)C(0) &= 1, \\ A(0)B(0)C(\pi/2) &= 1, \\ A(0)B(\pi/2)C(0) &= 1, \\ A(\pi/2)B(\pi/2)C(\pi/2) &= -1. \end{aligned} \quad (19)$$

the allowed values of the observables  $A, B, C$  are  $\pm 1$ , therefore multiplying the left hand side of Eqs. (19) we obtain 1, however the right hand side multiplication gives  $-1$  (Żukowski, et al.; 1998). Therefore (Greenberger, et al.; 1990; Żukowski, et al.; 1998), the EPR assumption are inconsistent with quantum mechanics even for the special case of perfect correlations given in Eq. (18).

#### 5.1.1 Gisin's theorem for three qubits

In 1991, N. Gisin (Gisin, 1991) proves that any non-product state, i.e. any entangled state, of two particle systems violate Bell's inequality. Essentially what he proves was that the Clauser, et al. (Clauser, et al.; 1969) inequality is violated for any entangled two-state system. That is to say, using the Schmidt decomposition given in Eq. (5) and the fact that for an entangled state there is at least two different  $\sqrt{\lambda_i}$ , i. e.  $\sqrt{\lambda_1}$  and  $\sqrt{\lambda_2}$ , then:

$$|P(a, b) - P(a, b^*)| + P(a^*, b) + P(a^*, b^*) = 2 \left(1 + 4|\sqrt{\lambda_1}\sqrt{\lambda_2}|\right)^{-1/2}, \quad (20)$$

Eq. (20) is strictly greater than 2 for entangled states. This is Gisin's theorem.

The next question is whether a similar proof can be given for more than two entangled state systems. Gisin himself argue that such a result does not holds. However, as far as we know, there are at least two restrict generalizations of Gisin's theorem for three quantum entangled systems. The first one was given by Chen, et al. (Chen, et al.; 2004); first they study cases for the generalized GHZ states, i.e.  $|\psi\rangle_{GHZ} = \cos(\xi)|000\rangle + \sin(\xi)|111\rangle$  and work out an

inequality of the following type:

$$B(\text{Bell}) \geq \frac{1}{2} + \frac{1}{3} \sqrt{1 + \sin^2(2\xi)}. \quad (21)$$

There is an essential difference between Eq. (20) and Eq. (21), it is that in the case of Gisin's theorem  $P(a, b)$  refers to expectation value whereas in the case of Eq. (21)  $B(\text{Bell})$  is a probability. Hence, what Chen, et. al. prove is the theorem: *All generalized GHZ states of three-qubit system violates a probability Bell's inequality.*

There are other proof of Bell's inequalities for three entanglement quantum mechanical systems. In fact, the search for Bell's type inequalities is an active research field. For example, Wu et. al. (Wu, et al.; 2008) presented a new Bell's inequality in terms of correlation functions and according to the authors this inequality is violated by any pure entangled state of three qubits. Also, Li and Fei (Li & Fei, 2010) present Bell type inequalities for multipartite states of arbitrary dimensions. Additionally, see the paper by Żukowski, et al. (Żukowski, et al.; 2002).

Finally, Ghose, et. al., (Ghose, et. al.; 2009) have shown that the generalized GHZ state does not violate the Svetlichnys inequality for  $\tau \leq 1/2$ , where  $\tau$  is the 3-tangle defined below. Then, showing a way to test for genuine tripartite nonlocal correlations.

## 5.2 Three entanglement measures

Three entangled quantum mechanical systems differ from two entangled systems in many features. One of them is that whereas for two entangled systems it is possible to find local operations (LOCC) that interconverts one entangled state to another such process is not possible for three entangled systems. When it is possible to obtain an entangled state  $|\psi\rangle_{12}$  from other entangled state  $|\phi\rangle_{12}$  using only LOCC, then it is agree that these two entangled states are equivalent because it is possible to carry out the same task with them. In fact, for two two-state entangled quantum mechanical systems it is possible to find such interconversion. However, this is not true for three entangled quantum mechanical systems. In this case it is well know that there are at least two inequivalent kinds of tripartite entanglement (Dür, et al.; 2000). The first one correspond to the GHZ state:

$$|GHZ\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle), \quad (22)$$

the second one correspond to the W state:

$$|W\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle). \quad (23)$$

There is no way to convert  $|GHZ\rangle$  in  $|W\rangle$  using LOCC and viceversa (Dür, et al.; 2000). This fact is related to the problem of the generalization from two quantum states measurement of entanglement to three entanglement measurement. What it is found in the literature is that there is not a single universal measurement of entanglement for the case of three entangled quantum mechanical systems.

### 5.2.1 Residual entanglement

When generalizing the concurrence  $C(\Phi)$  to three quantum mechanical systems Coffman et. al. (Coffman, et al.; 2000) define the *residual entanglement*, called *the 3-tangle*, as:

$$\tau_{ABC} = C_{B(AC)}^2 - C_{BC}^2 - C_{BA}^2, \quad (24)$$

in Eq. (24),  $C_{BC}^2$  is the concurrence between the systems  $B$  and  $C$ ,  $C_{BA}^2$  is the concurrence between the systems  $B$  and  $A$  and  $C_{B(AC)}^2$  is the concurrence between the system  $B$  and the system formed by systems  $A$  and  $C$ , the last taken as if were a single system  $AC$ . Far from be an universal measurement of entanglement, the residual entanglement gives 1 for the  $GHZ$  states and 0 for the  $W$  state (Coffman, et al.; 2000).

### 5.2.2 Geometric Measurement of Entanglement (GME)

The Geometric Measurement of Entanglement (GME) can be extended to define an entanglement measure for three quantum mechanical systems (Wei & Goldbart, 2003). For example, consider a three pure entangled state given by :

$$|\psi\rangle_{123} = \sum_{j_1, j_2, j_3} |e_{j_1}^1 e_{j_2}^2 e_{j_3}^3\rangle. \quad (25)$$

Then, the GME is obtained by the following procedure (Wei & Goldbart, 2003), first minimize  $E(\psi)$

$$E(\psi) = \min | |\psi\rangle_{123} - |\phi\rangle |^2, \quad (26)$$

where  $|\phi\rangle = \otimes_{i=1}^3 |\phi^i\rangle$  and  $|\phi^i\rangle = \sum_{j_i} c_{j_i}^i |e_{j_i}^i\rangle$  (Wei & Goldbart, 2003). Then, find the *entanglement eigenvalue*  $\Lambda_{max}$ . Finally, the GME for three entangled system is given by:

$$E_{\sin^2} = 1 - \Lambda_{max}^2. \quad (27)$$

In some cases it is possible to use the symmetry of the entangled state to alleviate the difficulty of calculations Wei & Goldbart (2003). For the  $GHZ$  state the  $E_{\sin^2} = 1/2$  and for the  $W$  state it is  $E_{\sin^2} = 5/9$  Wei & Goldbart (2003).

### 5.2.3 Schmidt measurement

The Schmidt measurement of entanglement was defined by Eisert and Briegel (Eisert & Briegel, 2001), it is based in the Schmidt representation of quantum states. That is, consider a three state quantum system with parties  $A_1$ ,  $A_2$  and  $A_3$ , then its state is given by:

$$|\psi\rangle = \sum_i^R \alpha_i |\psi_{A_1}^{(i)}\rangle \otimes |\psi_{A_2}^{(i)}\rangle \otimes |\psi_{A_3}^{(i)}\rangle, \quad (28)$$

where  $|\psi_{A_j}^{(i)}\rangle \in \mathbb{C}^{d_j}$ ,  $j = 1, 2, 3$ ;  $\alpha_i \in \mathbb{C}$ ;  $i = 1, \dots, R$ . Let  $r$  be the minimal number of product terms  $R$  in the decomposition of  $|\psi\rangle$ , then the Schmidt measure is defined as (Eisert & Briegel, 2001):

$$P(|\psi\rangle) = \log_2 r. \quad (29)$$

This entanglement measure is a generalization of the concept of Schmidt rank of density matrix for more than two quantum mechanical systems (Sanpera, et al.; 2001). It satisfy almost all the desired properties of a entanglement measure (Eisert & Briegel, 2001). Inclusive, contrary to the 3-tangle, it discriminates between the  $GHZ$  and  $W$  states, for the former  $P(|\psi\rangle) = 1$  and for the latter  $P(|\psi\rangle) = \log_2 3$  (Eisert & Briegel, 2001). See also, (Sperling & Vogel, 2011).

## 6. Entanglement power of operators and quantum gates

In this section we show that an operator or quantum gate could be characterized by calculating the entanglement produced by the operator or gate when acting in a product state. In Kraus and Cirac (Kraus& Cirac, 2001), it was stated that it is possible to give a two-qubit gate characterization in terms of the entanglement that it can produce when acting on a separable state. The Kraus & Cirac goal were the determination of the best separable input state that gives as much entanglement as possible when acting on them by a given operator represented by an unitary operator. One of the operator that they study was  $U_d = e^{-i\alpha S_x} = \cos(\alpha) - i \sin(\alpha)\sigma_x \otimes \sigma_x$ . they found that the best input state to produce the maximal entangled state, as measured by the concurrence, is:

$$\begin{aligned} |\phi\rangle_{AA^*} &= c_a|0\rangle_A|0\rangle_{A^*} + s_a|1\rangle_A|1\rangle_{A^*}, \\ |\psi\rangle_{BB^*} &= s_b|0\rangle_B|0\rangle_{B^*} + c_b|1\rangle_B|1\rangle_{B^*}. \end{aligned} \quad (30)$$

Another possibility to characterize a quantum gate is by calculating the entanglement produced by that gate on an arbitrary separable state. This was done by García Quijas, et. al. (García, et al.; 2011). Additionally, García Quijas and Arévalo Aguilar (García, et al.; 2010) characterize, both the control Z phase gate and the *relative phase gate* for two qubits. Here is shown the amount of entanglement produced by the relative phase gate for three qubits, in order to characterize this gate. Suppose that the relative phase gate acts on an arbitrary three qubit given by  $(\alpha|0\rangle_1 + \beta|1\rangle_1) \times (\delta|0\rangle_2 + \gamma|1\rangle_2) \times (\mu|0\rangle_2 + \nu|1\rangle_2)$ . The relative phase gate for three qubits produces the following change:

$$\begin{aligned} U_r^\theta |0\rangle_1|0\rangle_2|0\rangle_3 &\rightarrow |0\rangle_1|0\rangle_2|0\rangle_3, \\ U_r^\theta |0\rangle_1|0\rangle_2|1\rangle_3 &\rightarrow e^{i\theta}|0\rangle_1|0\rangle_2|1\rangle_3, \\ U_r^\theta |0\rangle_1|1\rangle_2|0\rangle_3 &\rightarrow e^{i\theta}|0\rangle_1|1\rangle_2|0\rangle_3, \\ U_r^\theta |0\rangle_1|1\rangle_2|1\rangle_3 &\rightarrow |0\rangle_1|1\rangle_2|1\rangle_3, \\ U_r^\theta |1\rangle_1|0\rangle_2|0\rangle_3 &\rightarrow e^{i\theta}|1\rangle_1|0\rangle_2|0\rangle_3, \\ U_r^\theta |1\rangle_1|0\rangle_2|1\rangle_3 &\rightarrow |1\rangle_1|0\rangle_2|1\rangle_3, \\ U_r^\theta |1\rangle_1|1\rangle_2|0\rangle_3 &\rightarrow |1\rangle_1|1\rangle_2|0\rangle_3, \\ U_r^\theta |1\rangle_1|1\rangle_2|1\rangle_3 &\rightarrow e^{i\theta}|1\rangle_1|1\rangle_2|1\rangle_3. \end{aligned} \quad (31)$$

Then, when acting in the arbitrary three qubit state, it gives the following entangled state:

$$\begin{aligned} U_r^\theta (\alpha|0\rangle_1 + \beta|1\rangle_1) \times (\delta|0\rangle_2 + \gamma|1\rangle_2) \times (\mu|0\rangle_2 + \nu|1\rangle_2) &\rightarrow \\ \{ \alpha\delta\mu|0\rangle_1|0\rangle_2|0\rangle_3 + \alpha\delta\nu e^{i\theta}|0\rangle_1|0\rangle_2|1\rangle_3 + \alpha\gamma\mu e^{i\theta}|0\rangle_1|1\rangle_2|0\rangle_3 & \\ \alpha\gamma\nu|0\rangle_1|1\rangle_2|1\rangle_3 + \beta\delta\mu e^{i\theta}|1\rangle_1|0\rangle_2|0\rangle_3 + \beta\delta\nu e^{i\theta}|1\rangle_1|0\rangle_2|1\rangle_3 & \\ + \beta\gamma\mu|1\rangle_1|1\rangle_2|0\rangle_3 + \beta\gamma\nu e^{i\theta}|1\rangle_1|1\rangle_2|1\rangle_3 \} & \end{aligned} \quad (32)$$

for  $\theta = \pi$ , Eq. (32) produces the following relative phase state:

$$\begin{aligned} U_r^\pi (\alpha|0\rangle_1 + \beta|1\rangle_1) \times (\delta|0\rangle_2 + \gamma|1\rangle_2) \times (\mu|0\rangle_2 + \nu|1\rangle_2) &\rightarrow \\ (\alpha|0\rangle_1 - \beta|1\rangle_1) \times (\delta|0\rangle_2 - \gamma|1\rangle_2) \times (\mu|0\rangle_2 - \nu|1\rangle_2). & \end{aligned} \quad (33)$$

To calculate the amount of entanglement in the state given by Eq. (32), we calculate the 3-tangle defined above. This measure is given by:

$$\tau_{123} = 4 | d_1 - 2d_2 + 4d_3 |, \tag{34}$$

where:

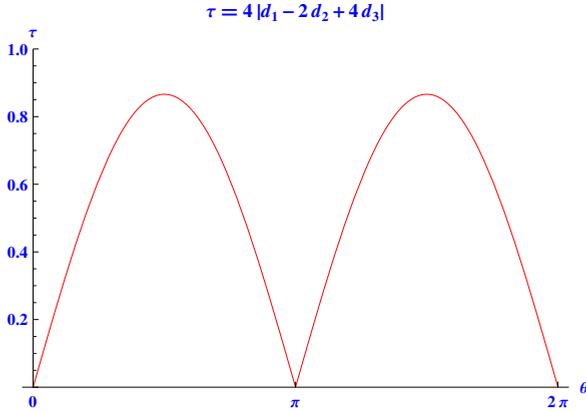


Fig. 1. Plot of formula (36), which gives the 3-tangle measure of entanglement.

$$\begin{aligned} d_1 &= a_{000}^2 a_{111}^2 + a_{001}^2 a_{110}^2 + a_{010}^2 a_{101}^2 + a_{100}^2 a_{011}^2; \\ d_2 &= a_{000} a_{111} a_{011} a_{100} + a_{000} a_{111} a_{101} a_{010} + a_{000} a_{111} a_{110} a_{001} \\ &\quad + a_{011} a_{100} a_{101} a_{010} + a_{011} a_{100} a_{110} a_{001} + a_{101} a_{010} a_{110} a_{001}; \\ d_3 &= a_{000} a_{110} a_{101} a_{011} + a_{111} a_{001} a_{010} a_{100}, \end{aligned} \tag{35}$$

and the coefficients  $a_{ijk}$  are the expansion of the three party entangled states in the standard basis. Comparing with Eq. (32) the coefficients are given by  $a_{000} = \alpha\delta\mu$ ,  $a_{111} = \beta\gamma\nu e^{i\theta}$  and so on. If we take  $\alpha = \beta = \delta = \gamma = \mu = \nu = 1/\sqrt{2}$ , then, using Eq. 32, Eq. 34 and Eq. 35, the residual entanglement for this case is:

$$\tau_{123} = \frac{1}{16} \sqrt{96 - 128 \cos 2\theta + 32 \cos 4\theta} \tag{36}$$

A plot of Eq. (36) is given in Fig. 1. We see that the residual entanglement of the relative entangled state for three qubits, given by Eq. (36), reaches its maximum value twice in the interval  $(0, 2\pi)$ . This corresponds to  $\theta = \pi/2$  and  $\theta = 3\pi/2$ , this agrees with the maximum value for the relative entangled state for two qubits (?).

Therefore, the characterization of a quantum gate (or an arbitrary operator) could be done by calculating the amount of entanglement produced by it when acting on an arbitrary state.

### 7. Experimental determination of entanglement

The experimental detection of entanglement requires the measurement of many physical properties, making this process quite difficult. Then, it is necessary to reduce the quantity of measurements to increase the efficiency of such process. In this section we briefly review

the effort made for many research groups to reduce the quantity of measurements to detect entanglement, both by deducing new definitions for some entanglement measures given in previous section (like concurrence) and by the new approach of proposing entanglement measures in terms of uncertainties of operators.

Nowadays, entanglement is considered a physical resource. Therefore, to use this resource in its potential applicabilities in quantum information theory, it is necessary to realize the following three tasks: i) to produce it, ii) to manipulate it, and iii) to detect it. Then, it is necessary to develop efficient methods to carry these tasks. Recently, there has been a huge effort in the task of detecting entanglement (Audenaert & Plenio, 2006; Gühne, et al.; 2009).

Usually, the principal procedures to experimentally detect entanglement is through Bell's inequalities. Therefore, to experimentally determine entanglement, it is necessary to make measurement in many variables, especially when the amount of entanglement is measured using one of the entanglement measures stated in section 5. An additional problem is the fact that some entanglement measures involves non-physical process, such is the case of the concurrence  $C|\Phi\rangle$  defined by (Hill, 1997) where it is defined using the complex conjugate  $\langle\Psi^*|$  which is an unphysical operation according with Walborn, et. al. (Walborn, et al.; 2006). For this reason, it was necessary to deduce new definitions of entanglement measures or to generalize the previous one taking into account the experimental settings.

To experimentally measure entanglement Walborn, et. al., uses an alternative definition of concurrence given by Mintert, et. al., (Mintert, et. al.; 2005), they use two copies of the qubit and found that the expectation value of an operator  $\hat{A}$ , with respect to these two copies, can capture entanglement properties. Then, the concurrence is defined as (Mintert, et. al.; 2005):

$$c(\Psi) = \sqrt{\langle\Psi| \otimes \langle\Psi|\hat{A}|\Psi\rangle \otimes |\Psi\rangle}. \quad (37)$$

In their setting, Walborn, et. al., store a copy of the state  $|\Psi\rangle$  in the polarization and momentum degrees of freedom of a single photon. Entangled polarization states were created by pumping two perpendicular nonlinear crystals. The concurrence of  $|\Psi\rangle$  is determined by the probability of observing the first photon in the state:

$$|\psi^-\rangle = \frac{1}{2} (|H\rangle|b\rangle - |V\rangle|a\rangle). \quad (38)$$

The probability  $P_A$  is then given by the count rate for the observation of  $|\psi^-\rangle$  normalized by the sum of the count rates of the others Bell states.

Another approach to the experimental quantification of entanglement is based on uncertainty relations, this approach is motivated by the goal of detecting entanglement with just few measurements (Hofmann & Takeuchi, 2003). Then, making the detection process more efficient. This approach was highly influenced by the work of Hofmann and Takeuchi (Hofmann & Takeuchi, 2003), who showed that there are uncertainty limits that are violated by entangled states. In particular, the uncertainty of an observable  $\hat{A}_i$  is given by the statistical variance as follow:

$$(\delta\hat{A}_i)^2 = \langle\hat{A}_i^2\rangle - \langle\hat{A}_i\rangle^2. \quad (39)$$

The uncertainty given by Eq. (39) is zero if the wave function is an eigenfunction of  $\hat{A}_i$ . Then, for a set of commuting observables  $\{\hat{A}_i\}$  there are common eigenfunctions such that  $\delta\hat{A}_i = 0$ . This follows when the expectation values are taken with respect to this common eigenfunction, therefore  $\sum_i (\delta\hat{A}_i)^2 = 0$ . On the other hand, if the set  $\{\hat{A}_i\}$  is a set of non commuting observables, then there must be a lower limit  $U > 0$  such that (Hofmann &

Takeuchi, 2003)  $\sum_i (\delta \hat{A}_i)^2 \geq U$ . The generalization of this equation for two systems is given by:

$$\sum_i (\delta(\hat{A}_i + \hat{B}_i))^2 \geq U_A + U_B. \quad (40)$$

A quantitative amount of entanglement can be defined as (Hofmann & Takeuchi, 2003):

$$C_{LUR} = 1 - \frac{\sum_i (\delta(\hat{A}_i + \hat{B}_i))^2}{2(U_A + U_B)}. \quad (41)$$

$C_{LUR}$  measures the amount of entanglement verified by the violation of local uncertainty (Hofmann & Takeuchi, 2003).

Some criteria must be satisfied by any entanglement measure (Horodecki, et al.; 2000), in particular it must be invariant under local unitary transformation. To apply this criteria to uncertainty relations Samuelsson and Björk (Samuelsson & Björk, 2006) and Kothe and Björk (Kothe & Björk, 2006) defined a new measure of entanglement based on the covariance, i. e.  $C(\hat{A}_i, \hat{B}_i) = \langle \hat{A}_i \hat{B}_i \rangle - \langle \hat{A}_i \rangle \langle \hat{B}_i \rangle$ . This entanglement measure is given by (Kothe & Björk, 2006):

$$G = \sum_{i,j=1}^3 C^2(\hat{\sigma}_i^A, \hat{\sigma}_j^B), \quad (42)$$

where  $\hat{\sigma}_i^A$  is the  $i$ -th Pauli operator for system  $A$  and  $\hat{\sigma}_j^B$  for system  $B$ . The covariance given by Eq. (42) was used by Wang, et. al. Wang (et al.; 2007), to experimentally quantify the entanglement of photons produced by BBO crystals.

A quite related approach was followed by Klyachko 2006, et. al., (Klyachko, et al.; 2006) and Klyachko, et. al., (Klyachko, et al.; 2007) to propose an experimental entanglement measure. In particular they shown that the concurrence is related to the variance by the following relation (Klyachko, et al.; 2007) and (Klyachko, et al.; 2006):

$$C^2(\Psi) = \sqrt{\frac{V(\Psi) - V_{min}}{V_{max} - V_{min}}}, \quad (43)$$

where,  $V(\Psi) = C_H - \langle \psi | X_\psi | \psi \rangle$ ,  $X_\psi = \sum_\alpha \langle \psi | X_\alpha | \psi \rangle X_\alpha$  (Klyachko, et al.; 2007) and  $C_H = \sum_\alpha X_\alpha^2$ .

On the other hand, this approach has been extended to cover the entropic uncertainty relations, see Li, et. al. (Li, et al.; 2010), and Ghune and Lewenstein (Gühne, et al.; 2004). Finally, Fei, et. al., (Fei, et. al.; 2009) has proposed that the experimental measurement of concurrence can be done by a single measurement of the expectation value of tensor products of local observables. In particular, they have shown that the square of the concurrence for two qubit states could be measured just by measuring certain expectation value of the Pauli matrixes.

## 8. Conclusion

In this chapter we have reviewed the concept of entanglement and some issues regarding it. Specially, we have focused in some of the more common entanglement's measures, like the concurrence. Also, we reviewed some interesting issues of the Bell inequalities. Additionally, we showed a way to characterize quantum gates and operators. Finally, we briefly reviewed the experimental procedure to measure the entanglement content of an entangled state.

## 9. Acknowledges

We thanks Consejo Nacional de Ciencia y Tecnologia (CONACYT). L. M. Arévalo Aguilar is grateful to the Programa de Mejoramiento del Profesorado (PROMEP) under grant PROMEP /103.5/10/4959. Also, we acknowledge the support from Vicerrectoria de Investigación y Posgrado VIEP-BUAP under grand ARAL-EXC11-I.

## 10. References

- Audenaert, K. M. R.; Plenio, M. When are correlations quantum?-verification and quantification of entanglement by simple measurements. *New Jour. Phys.*, Vol. 8, No. 11, (November 2006) page numbers (266).
- Bell, J. S. (1964). On the Einstein-Podolsky-Rosen paradox. *Physics*, Vol. 1, No., (month and year of the edition) page numbers (195-200).
- Bell, J. S. (1966). On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.*, Vol. 38, No. 3, (July - September 1966) page numbers (447-452).
- Bennett, C. H.; Brassard, G.; Crépeau, C.; Jozsa, R.; Peres, A.; Wootters, W. K. (1993). Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, Vol. 70, No. 4, (29 March 1993) page numbers (1895-1899).
- Bennett, C. H.; Bernstein, H. J.; Popescu, S.; Schumacher, B. (1996). Concentrating partial entanglement by local operations. *Phys. Rev. A*, Vol. 53, No. 4, (April 1996) page numbers (2046-2052).
- Bennett, C. H.; DiVincenzo, D. P.; Smolin, J.; Wootters, W. K. (1996b). Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, Vol. 54, No. 4, (November 1996) page numbers (3824-3851).
- Chen, J.-L.; Wu, C.; Kwek, L. C.; Oh, C. H. (2004). Gisin's theorem for three qubits. *Phys. Rev. Lett.*, Vol. 93, No. 14, (1 October 2004) page numbers (140407).
- Choudhary, S. K.; Ghosh, S.; Kar, G.; Rahaman, R (2010). Analytical proof of Gisin's theorem for three qubits. *Phys. Rev. A*, Vol. 81, No. 4, (14 April 2010) page numbers (042107).
- Clauser, J. F.; Horne, M. A.; Shimony, A.; Holt, R. A. (1969). Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, Vol. 23, No. 15, (month and year of the edition) page numbers (880-884).
- Coffman, V.; Kundu, J.; Wootters, W. K. (2000). Distributed entanglement. *Phys. Rev. A*, Vol. 61, No. 5, (10 April 2000) page numbers (052306).
- Dür, W.; Vidal, G.; Cirac, J. I. (2000). Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, Vol. 62, No. 6, (14 November 2000) page numbers (062314).
- Eberhard, P. H. (1977). Bell's theorem without hidden variables. *IL Nuovo Cimento*, Vol. 38, No. 1, (month and year of the edition) page numbers (75-79).
- Einstein, A.; Podolsky, B.; Rosen, N. (1935). Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?. *Phys. Rev.*, Vol. 47, No., (month and year of the edition) page numbers (777-780).
- Eisert, J.; Briegel, H. J. (2001). Schmidt measure as a tool for quantifying multiparticle entanglement. *Phys. Rev. A*, Vol. 64, No. 2, (9 July 2001) page numbers (022306).
- Fei, S.-M.; Zhao, M.-J.; Chen, K.; Wang, Z.-X. (2009). Experimental determination of entanglement for arbitrary pure states. *Phys. Rev. A*, Vol. 80, No. 3, (21 September 2009) page numbers (032320).

- García Quijas, P. C.; Arévalo Aguilar, L. M. (2010). A quantum coupler and the harmonic oscillator interacting with a reservoir: Defining the relative phase gate. *Quant. Inf. & Comp.*, Vol. 10, No. 3&4, (March 2010) page numbers (189-199).
- García Quijas, P. C.; Arévalo Aguilar, L. M.; Arroyo-Carrasco, M. L. (2011). Characterizing the relative phase gate. *Rev. Mex. Fis. S*, Vol. 53, No. 3, (July 2011) page numbers (36-41).
- Ghose, S.; Sinclair, N.; Debnath, S; Rungta, P; Stock, R. (2009). Tripartite Entanglement versus Tripartite Nonlocality in Three-Qubit Greenberger-Horne-Zeilinger-Class States. *Phys. Rev. Lett.*, Vol. 104, No. 25, (25 June 2009) page numbers (250404).
- Gisin, N. (1991). Bell's inequality holds for all non-product states. *Phys. Lett. A*, Vol. 154, No. 5,6, (8 April 1991) page numbers (201-202). Elsevier Science Publishers B. V.
- Greenberger, D. M.; Horne, M. A.; Shimony, A; Zeilinger, A. (1990). Bell's theorem without inequalities. *Am. J. Phys.*, Vol. 58, No. 12, (December 1990) page numbers (1131-1143).
- Gröblacher, S.; Paterek, T.; Kaltenbaek, R.; Brukner, Č.; Żukowski, M.; Aspelmeyer, M.; Zeilinger, A. (2007). An experimental test of non-local realism. *Nature*, Vol. 446, No. 05677, (19 April 2007) page numbers (871-875).
- Gühne, O.; Lewenstein, M. (2004). Entropic uncertainty relations and entanglement. *Phys. Rev. A*, Vol. 70, No. 2, (24 August 2004) page numbers (022316).
- Gühne, O.; Tóth, G. (2009). Entanglement detection. *Phys. Rep.*, Vol. 474, No. , (2009) page numbers (1-75). Elsevier.
- Hill, S.; Wootters, W. K. (1997). Entanglement of a Pair of Quantum Bits. *Phys. Rev. Lett.*, Vol. 78, No. 26, (30 June 1997) page numbers (5022-5025).
- Hofmann, H. F.; Takeuchi, S. (2003). Violation of local uncertainty relations as a signature of entanglement. *Phys. Rev. A*, Vol. 68, No. , (2003) page numbers (032103).
- Horodecki, M.; Horodecki, P; Horodecki, R. (2000). Limits for Entanglement measures. *Phys. Rev. Lett.*, Vol. 84, No. 9, (28 February 2000) page numbers (2014-2017).
- Horodecki, R.; Horodecki, P; Horodecki, M.; Horodecki, K.; (2009). Quantum entanglement. *Rev. Mod. Phys.*, Vol. 81, No. 2, (April-June 2009) page numbers (865-942).
- Hughston, L. P.; Josza, R.; Wootters, W. K. (1993). A complete classification of quantum ensembles having a given density matrix. *Phys. Lett. A*, Vol. 183, No. 1, (29 November 1993) page numbers (14-18).
- Kothe, C.; Björk, G. (2006). Entanglement quantification through local observable correlations. *Phys. Rev. A*, Vol. 73, No. 1, (30 January 2007) page numbers (012336).
- Klyachko, A, Öztop, B.; Shumovsky, A. S. (2006). Measurable entanglement. *App. Phys. Lett.*, Vol. 88, No. 12, ( 21 March 2006) page numbers (124102).
- Klyachko, A, Öztop, B.; Shumovsky, A. S. (2007). Quantification of entanglement via uncertainty. *Phys. Rev. A*, Vol. 75, No. 3, (12 March 2007) page numbers (032315).
- Kraus, B.; Cirac, J.I. (2001). Optimal creation of entanglement using a two-qubit gate. *Phys. Rev. A*, Vol. 63, No. 6, (17 May 2001) page numbers (062309).
- Li, M.; Fei, S.-M (2010). Gisin's theorem for arbitrary dimensional multipartite states. *Phys. Rev. Lett.*, Vol. 104, No. 24, (18 June 2010) page numbers (240502).
- Li, C. -F.; Xu, J. -S.; Xu, X. -Y; Guo, G. -C. (2010). Experimental investigation of the entanglement-assisted entropic uncertainty principle. *quant-ph*, Vol. , No. , () page numbers (arXiv:1012.0361).
- Legget, A. J.; . Nonlocal hidden-variable theories and quantum mechanics: An incompatibility theorem (2003). *Found. Phys.*, Vol. 33, No. 10, (October 2003) page numbers (1469-1493).
- Méthot, A. A.; Scarani, V. (2007) An anomaly of non-locality. *Quant. Inf. & Comp.* Vol. 7, No. 1, (2007) page numbers (157-170).

- Mintert, F.; Kuś, M.; Buchleitner, A. (2005). Concurrence of Mixed Multipartite Quantum States. *Phys. Rev. Lett.*, Vol. 95, No. 26, (20 December 2005) page numbers (260502).
- Plenio, M. B.; Virmani, S. (2007). An introduction to entanglement Measures. *Quant. Inf. & Comp.*, Vol. 7, No. 1, (month and year of the edition) page numbers (1-51). Rinton Press.
- Popescu, S. & Rohrlich, D. (1998). The joy of entanglement, In: *Introduction to quantum computation and information*, Lo, H-K.; Popescu, S.; Spiller, T., (Ed. World Scientific), page numbers (29-48), Publisher, ISBN 981023399X, Singapore.
- Samuelson, S.; Björk, G. (2006). Local uncertainty relations serving as measures of entanglement in bipartite two-level system. *Phys. Rev. A*, Vol. 73, No. 5, (12 January 2006) page numbers (012319).
- Sanpera, A; Bruß, D.; Lewenstein, M. (2001). Schmidt-number witnesses and bound entanglement. *Phys. Rev. A*, Vol. 63, No. 5, (9 April 2001) page numbers (050301).
- Shimony, A. (1993). Degree of entanglement. *Ann. New York Acad. Sci.*, Vol. 755, No. 1, (29 November 1995) page numbers (675-679).
- Sperling, J.; Vogel, W. (2011). Determination of the Schmidt number. *Phys. Rev. A*, Vol. 83, No. 4, (12 April 2011) page numbers (042315).
- Vedral, V.; Plenio, M. B.; Rippin, M. A.; Knight, P. L. (1997). Quantifying Entanglement. *Phys. Rev. Lett.*, Vol. 78, No. 12, (24 March 1997) page numbers (2275-2279).
- Vedral, V.; Plenio, M. B. (1998). Entanglement measures and purification procedures. *Phys. Rev. A*, Vol. 57, No. 3, (March 1998) page numbers (1619-1633).
- Walborn, S.P.; Souto Ribeiro, P. H.; Davidovich, L.; Mintert, F.; Buchleitner, A. (2006). Experimental determination of entanglement with a single measurement. *Nature*, Vol. 440, No. 7087, (20 April 2006) page numbers (1022-1024).
- Wang, Z. -W.; Huang, Y. -F.; Ren, X. -F.; Zhang, Y. -S.; Guo, G. -C. (2007). Experimental entanglement quantification and verification via uncertainty relations. *EPL*, Vol. 78, No. 4, (May 2007) page numbers (40002).
- Wei, T.; Goldbart, P. M. (2003). Geometric measure of entanglement and applications to bipartite and multipartite quantum states. *Phys. Rev. A*, Vol. 68, No. 3, (9 October 2003) page numbers (042307).
- Wei, T.-C.; Ericsson, M; Goldbart, P. M.; Munro M. J.T. (2004). Connections between relative entropy of entanglement and geometric measure of entanglement . *Quant. Inf. & Comp.*, Vol. 4, No. 4, (July 5, 2004) page numbers (252-272). Rinton Press
- Wootters, W. K. (1998). Entanglement of Formation of an Arbitrary State of Two Qubits. *Phys. Rev. Lett.*, Vol. 80, No. 10, (9 March 1998) page numbers (2245-2248).
- Wootters, W. K. (2001). Entanglement of Formation and Concurrence. *Quant. Inf. & Comp.*, Vol. 1, No. 1, (June 5, 2001) page numbers (27-44). Rinton Press.
- Wu, C.; Chen, J.-L.; Kweek, L. C.; Oh, C. H. (2008). Correlation-function Bell inequality with improved visibility for three qubits. *Phys. Rev. A*, Vol. 77, No. 6, (5 June 2008) page numbers (062309).
- Żukowski, M.; Zeilinger, A.; Horne, M. A.; Weinfurter, H.; (1998). Quest for GHZ states. *Acta Physica Polonica A*, Vol. 93, No. 1, page numbers (187-195).
- Żukowski, M.; Brukner, Č.; Laskowski, W.; Wieśniak, M. (2002). Do All Pure Entangled States Violate Bell's Inequalities for Correlation Functions?. *Phys. Rev. Lett.*, Vol. 88, No. 21, (27 May 2002) page numbers (210402).

# Quantum Spin Coherent Electron Correlation

Peter G. Burton  
*University of Wollongong*  
*Australia*

## 1. Introduction

A data-driven quantum chemical re-investigation and characterisation of the nature of the singlet-coupled e-pair interaction in weak interactions in physics and chemistry is discussed, in order to bring out its surprisingly broad implications for our understanding of some of the most fundamental aspects of nature.

### 1.1 Outline

The central conceptual difficulty in standard non-relativistic quantum chemistry, as for relativistic quantum mechanics, remains that in the absence of a satisfactory, physically intelligible representation of the phenomenon of spin, spin is represented in our models as a topological constraint on many-particle interactions. All of chemistry is dominated by the phenomenon of spin-pairing, derived from the Aufbau principle of filling electron shells in atoms. However determinantal expression of antisymmetry requirements effectively maintains separation of same spin electrons (exchange effects) while degrading the details of coupling between opposing spin electrons to electrostatics. This chapter investigates quantum mechanically the implications of a semiclassical model of spin proposed and formally defended some time ago (Burton 1988; including a formal modification of coupled pair theory), extending my previous published conjectures (Burton 2009). Its simple new idea touches a vast area of contemporary physics and chemistry.

### 1.2 Basic motivation

The stimulus for this research was a program focused on weak molecular interactions involving small neutral molecules whose interactions are well known to be very sensitive to electron correlation effects (Burton 1977, 1979, 1980; Burton, Gray & Senff 1982; Senff & Burton 1985, 1986, 1989). The particular focus of our investigations was interactions involving He and H<sub>2</sub> molecules involving well-localised e-pairs. Aside from their fundamental relevance to our prehistory in cool interstellar, molecular clouds, from a technical viewpoint these molecules were ideal to minimise the impact of artefacts arising from the various approximations involved in implementing the conventional correlation algorithms we employed at the time.

Interactions involving such well-separated e-pairs should represent ideal candidates for investigation with pair correlation algorithms (Sinanoglu 1961, Hurley 1976) designed to isolate intra-pair and interpair contributions to electron correlation corrections to many-electron wavefunctions, and their simple electronic structures were amenable to relatively

complete Hartree-Fock Self-Consistent-Field modelling. Being confident of control over artefacts from these sources, we carefully investigated basis set completion effects in the pair-correlation ansatz. The weak van der Waals interaction potentials in He...He, He...H<sub>2</sub> and H<sub>2</sub>... H<sub>2</sub> were available experimentally by several routes, including direct scattering in crossed molecular beam experiments, and with these as reference we embarked on a program to test regions of the more polarisable inter-atomic interaction zones for basis set incompleteness.

What we found told us that even very large conventional, nuclear-centred, gaussian basis sets were subject to improved electron correlation energies when augmented by additional s,p probe sets aligned between nuclear centres. Our approach to their disposition was to locate these basis set additions off nuclear centres along the line of interaction, but to keep these located a fixed distance from the nearby atomic centre while intermolecular distance coordinates were varied to sample van der Waals interaction, to minimise basis set extension error (Boys & Bernardi, 1970), an artefact well-known in HF-SCF calculations but possible also in the computation of pair-correlation energies.

The results, admittedly involving small effects, but systematically accumulated over more than a decade, eventually led me to question why the low density region at the polarisable edge of interacting atoms & molecules should be so sensitive to details of basis set disposition. Surely, I had thought, electron correlation is dominated by the dynamics of inter-electron interactions in regions of higher rather than lower electron density<sup>1</sup>. The details of our calculations however were asking me to revisit this assumption: maybe electron correlation is better understood as an *edge effect* of e-pairs?

### 1.3 Reviewing the basic conventions of quantum chemistry

Multi-electron wavefunctions in quantum chemistry are conventionally characterised by aggregating individual electron spin-orbitals into an antisymmetrised, determinantal notation devised in 1929 by J.C. Slater. The spatial and spin parts of multi-electronic wavefunctions fail to factorise except in the special case of a single e-pair, so the burden of the determinantal expansion of electronic spin-orbitals having  $m_s = +\frac{1}{2}$  and  $-\frac{1}{2}$  spins is accepted to ensure that the exclusion principle devised by Wolfgang Pauli in 1925 (Pauli 1945) is satisfied for the electronic structure, even though most molecular electronic structures in chemistry are strongly paired. At least the antisymmetry required of sets of fermions, the wavefunction changing sign with respect to pairwise particle interchange, was guaranteed by the determinantal formalism.

Much of the chemistry of molecules formed of light atoms reflects their electronic structure being comprised of relatively localised e-pairs with close-coupled singlet pairing of spins, reminiscent of the simplified (and pre-quantum mechanical) models of locally-paired electron dot diagrams introduced by Gilbert N Lewis (1916) to explain the covalent bond. This has led to extensive theoretical investigation of pair correlation methods both formally (Sinanoglu 1961, Hurley 1976) and subsequently by a wide range of computational methods in quantum chemistry, including those used in our program of investigating weak van der Waals interactions.

The determinantal notation, which succeeds by construction to keep electrons of the same spin widely separated in the quantum chemistry models, ensures conformance to

---

<sup>1</sup> The weak assumption embedded in density functional approaches to correlation effects, rather than correlation being localised in *polarisable* regions .

antisymmetry and the exclusion principle but its cost is that it de-emphasises the natural tendency to e-pair formation in chemistry. If the close coupling between electrons of opposite spin is more relevant to the physics of the system under investigation, as theorists Sinanoglu and Hurley have chosen to emphasise by treating these structures as clusters of e-pairs, the pair-correlation formalism (rather than the more conventional configuration interaction or CI approach) may be advantageous. By emphasising correlations between e-pairs of opposing spin, and of pair-pair interactions between them, such an approach seemed ideal for investigating weak intermolecular forces, as was the case in our interactions involving He and H<sub>2</sub>.

When one focuses upon weakly interacting systems with one or two closely localised electron pairs (e-pairs), and these are light neutral molecules such as He or H<sub>2</sub>, the interaction energy between the localised pairs is *purely repulsive* when modelled at the uncorrelated level of the Hartree-Fock-Self-Consistent-Field model. HF-SCF approximates each electron (spin-)orbital self-consistently in the time-averaged, 'electrostatic' field of the other electrons. Correcting the approximate HF-SCF wavefunction requires adding electron correlation corrections to the zero-order 'independent electron' wavefunction, by introducing either explicit inter-electron distance parameters (e.g. Hylleraas 1930; Boys and Handy 1969, Gill et al. 2006) or by constructing configuration interaction (e.g. Shavitt 1977, 1981) expansions to implicitly (statistically) model more flexibly the details of electron-pair interaction. Either strategy, at the considerable expense of very extended and diffuse representation of the resulting correlation, and being hard to interpret, can realise excellent agreement between computed and empirical total electronic energy for molecules in their ground states. In particular for our neutral, light molecule systems, total energies can be expected which quite closely mimic the rather weak van der Waals *attraction* between these close-shell molecules as they approach one another. This is a critical achievement, required because otherwise it would be impossible to *liquefy* He(g) or H<sub>2</sub>(g) albeit at temperatures close to zero Kelvin: He boiling point is -268.6 °C (4.55 K); H<sub>2</sub> boiling point (1.013 bar) is -252.8 °C (20.8K). Statistical mechanical models of these phase changes, and molecular beam scattering experiments, are completely consistent with *attractive* van der Waals interaction potentials for the respective molecular interactions. So one can conclude that it is impossible to model the (g) to (l) phase change in these liquids without taking electron correlation effects into account.

#### 1.4 Challenging the conventional approach to representing spin

Our van der Waals interaction calculations, using well-matched electron correlation algorithms and more and more complete basis sets to avoid obvious basis set superposition artefacts, showed some small but unexpected effects which led us to reconsider the conventional wisdom of electron correlation theory. The topological separation of same-spin electrons enforced by the determinantal wavefunction formulation leads to very long multi-configuration expansions for the correlated wavefunction expressions which approach exact ground-state energies of these molecules<sup>2</sup>. The physical interpretation of these correlation effects are best shown from e-pair cluster expansions, which show the correlation corrections are dominated by the principal correlating natural orbitals<sup>3</sup> for each localised e-pair, which

<sup>2</sup> This is already true for *single* two-electron ground-state pairs, such as He (Hylleraas 1930; Pekeris 1958) or H<sub>2</sub> (Kolos et al. 1986), when e-pair correlation is explicitly modelled by  $r_{12}$  expansion.

<sup>3</sup> The significance of natural orbitals lies in the fact that CI expansions based on these orbitals have generally the fastest convergence.

introduce nodal surfaces to intersect the occupied orbitals of the e-pair either angularly or radially. This tends to create an impression that electron correlation effects are derived from regions of higher rather than lower electron density within the e-pair. This long held assumption was now the focus of my concern, because, as with direct conventional Configuration Interaction (CI) expansions, even the PNO-CI expansions of cluster pair theory tend to aggregate many small contributions whose physical nature is difficult to discern.

What our strange off-nuclear-centred basis set augmentations were telling us (in our investigation of the interaction of light non-polar molecules) was that significant regionally specific correlation contributions were arising from low-density, highly polarisable, outer reaches of atoms, well away from the respective nuclear centres which tightly constrain electrons by nuclear attraction.

I began in 1988 to investigate a semi-classical model of correlation in a two-electron, axially symmetric bound-state (Burton 1988), starting from first principles and the helicity equations, and to review Hurley's formal specification of e-pair correlation. In determinantal formulations customary in quantum chemistry, the spin wavefunctions have a purely topological significance consistent with the exclusion principle but what I was coming to realise instead was that spin should be explicitly and physically instantiated *directly* as peripheral wavelet disturbances in the superficial parts of electron orbitals themselves.

The idea of physically embodied but peripheral wave-matter spin, considered semiclassically as a *fractional* electron disturbance, could be a significant effect but remain rather well hidden by the general acceptance of spin being modelled as a determinantal topology. It had just the features to physically explain the Stern-Gerlach 1925 effect in a superior way to that proposed by Uhlenbeck & Goudsmit in 1925 & 1926, which supposed the ultimately untenable notion that the *whole* electron was subject to spin. In semiclassical terms, replacing the electronic charge to mass ratio  $e/m$  in expressions for the electron magnetic moment due to spin by the expression  $\delta e/\delta m$  proposed in Burton 1988, a ratio of identical magnitude for identical fractions of mass and charge, was entirely consistent with fundamental treatments of spin effects (Merzbacher 1970; Landau & Lifshitz 1977). This heretical idea, one certainly not excluded by the history of the concept (Tomonaga 1974/1997), warranted further investigation.

Much later in 2003-5 (Burton 2009), I realised that the dominant or principal correlating orbital for He could in fact be *identified with* the antisymmetrised singlet-coupled spin wavefunction of the spin-orbital coupling itself. Furthermore, in this special case, a single variational parameter corresponding to its radial mixing with the HF-SCF reference (introduced as a conjecture in Burton 1988) was all that remained to exactly determine the He (1s)<sup>2</sup> ground state singlet energy, once the angular interaction effects were *completely specified* by the symmetry of the antisymmetrised singlet coupling formula (Section 2.7; see Figure 1).

## 2. Electron correlation problem in quantum chemistry

During the decade around my hosting an *International Workshop in Atomic and Molecular Physics and Quantum Chemistry* at the University of Wollongong in June 1980 and whose proceedings I edited (Burton 1980), my small research group's activities centred on two ways of testing the validity of then current models of electron correlation in quantum chemistry. We were focused upon investigating models of the electron correlation energy contribution itself at various degrees of approximation for weak van der Waals molecular interactions (Senff & Burton 1985, 1986, 1989), in particular by appraising the impact of

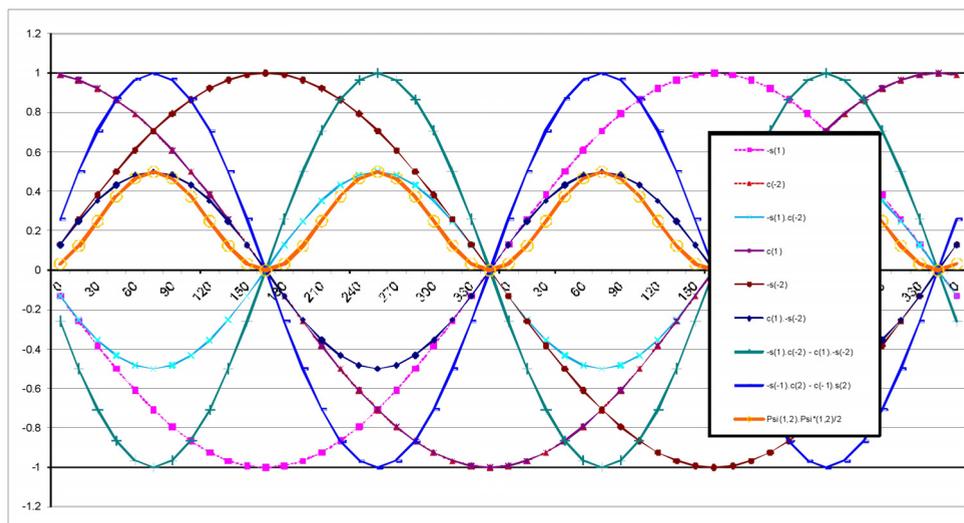


Fig. 1. Angular representation of  $m = \pm 1/2$  spin wavefunctions for He  $(1s)^2$  configuration, in antisymmetrised singlet form, with relative phase optimized to illustrate axially circumferential quadrupolar beats in the periphery of the pair function. The circumferential angular form for a given distance  $\rho$  off-axis is illustrated beginning with explicit sinusoidal (s) or cosine (c) forms for  $m_s = \pm 1/2$  as shown in the legend.

innovative ways of assessing the basis set limitations implicit in those models by the choice of strategy to build more completeness (accessing yet more subtle correlation effects) into those, still approximate, models of the correlation contribution (Burton 1977, 1979, 1980; Burton, Gray & Senff 1982).

## 2.1 Issues in conventional computational approaches

Before Isaiah Shavitt's (1981) '*Graphical Unitary Group Approach*' led to making Full CI calculations practical to exhaust the correlation energy contributions of a given molecular basis set, methods tracing their heritage back to Oktay Sinanoglu (1961) and focused on pair correlation energies were formally analysed for completeness by Andrew Hurley (1976, *Electron Correlation in Small Molecules*) and made practical in for computational quantum chemistry in basis set representation by several groups led by Werner Kutzelnigg, Wilfried Meyer and Reinhart Ahlrichs *et al.* (1975a, 1975b) in Germany.

At that time, through focusing upon relatively localised intra-pair-correlation energies which were known to dominate the correlation energy contributions (corrections) to much simpler molecular Hartree-Fock, Self-Consistent Field estimates of molecular electronic energies, it was left to the relatively weak coupling between the correlation energy effects of *different* pairs to represent the uncertainty of various "coupled pair" methods of assessing total correlation energies within the constraints of a given basis set (Ahlrichs *et al.* 1975A, 1975b).

## 2.2 Interactions involving localised electron pairs

While the transformation to a localised formulation of occupied molecular orbitals in, say, the methane molecule,  $\text{CH}_4$ , generates four equivalent bonding orbitals, the canonical

orbitals in the SCF-MO representation begin symmetry adapted. We preferred to work in a localised orbital representation to assess pair correlation energies to concentrate the nominal impact of intra-pair correlation in a less diffuse, localised orbital representation, consistent with the supposed strength of the formal methods for treating electron correlation we chose to pursue.

In time we narrowed our focus to the simplest of molecules: He, Li<sup>+</sup>, H<sub>2</sub>, H<sub>3</sub><sup>+</sup> having a single (localised) electron pair, and concentrated on assessing the nature of the weak inter-pair correlation energy effects corresponding to the van der Waals interaction between these species, including weakest of interactions of neutral molecules He...He, He...H<sub>2</sub> in various angular alignments and H<sub>2</sub>...H<sub>2</sub> in its various representative configurations.

Molecular beam scattering analyses of these neutral molecule interactions provided the external reference for our internal analysis of systematic errors from the approximations we were forced to make in our calculations because of our access to relatively modest computational resources.

The basis set superposition error (BSSE; Boys & Bernardi 1970) was of particular significance to our approach, since poor representation by way of local basis set limitation led to the correlation ansatz "borrowing" flexibility from nearby centres as they were brought nearer, generating artificially higher (less limited) correlation contributions for more compact molecular forms (close collision) than those more distantly interacting (van der Waals asymptote) molecular forms: this artefact acted apparently to deepen the attractive van der Waals well.

In order to independently appraise the quality of the representations of each counterpart molecule in these molecular interaction calculations, we adopted the pragmatic device, in addition to using large, well-tempered s,p,d,f basis sets on each nuclear centre, of scanning the sensitivity of calculated correlation energies with a free-floating s,p set to probe where in the molecule additional basis set flexibility might still be required.

We reported this approach at the AMPQC workshop in Burton, Gray & Senff (1980), and two of us (myself and Ulrich Senff) subsequently found in a careful series of calculations that in the close interface between weakly interacting neutral systems (in the region where the SCF-MO energies are in fact repulsive, and only correlation effects generate an attractive intermolecular potential energy to qualitatively match experiment) involving He...He, He...H<sub>2</sub> and H<sub>2</sub>...H<sub>2</sub> that even quite large nuclear-centred basis sets appeared systematically inadequate to fully represent the intermolecular correlation energy of these weakly interacting systems.

Our focus on weak interactions between light, non-polar molecules such as He or H<sub>2</sub> allowed us to effectively eliminate the BSSE correction at the HF-SCF level, so the impact of free-floating s,p function specifically in the correlation energy correction could be assessed. What we found consistently (despite correcting for the BSSE itself) was for each atom centre, additional correlation energy arose with s,p sets being located at the 'edge' (say fixed at 50-100pm off-nucleus) of each atom, on the line of centres of their interaction. This remained true as we increased the well-tempered s,p,d,f bases on each centre.

This was a small but persistent effect, and it raised an important question in my mind: Why was the atomic or molecular *edge* of these interacting systems, each just a localised correlated pair of electrons, seemingly so vulnerable to additional degrees of local basis set freedom?

### 2.3 Back to first principles of spin and correlation

By 1988, my reflecting on this question turned active again, but from an entirely different angle. I went back to first principles in the structure of the helicity equations in Landau & Lifshitz (1977) for a free electron and contemplated what they would reduce to for the case of an electron constrained axially by being bound on the attractive nucleus of an atom.

I discovered (Burton 1988) that a semiclassical, fractional electron effect – like ion cyclotron resonance but considered as a fractional electron wavelet circulating circumferentially around a Schrödinger-like delocalised electronic matterwave – could support all the formal equations representing electron spin effects without change if instead of the electronic charge  $e$  and the electronic mass  $m$  occurring everywhere as the electronic charge to mass ratio ( $e/m$ ) I notionally attributed spin effects in these equations to there being a localised disturbance involving a *small fraction* of the electronic matter wave ( $\delta e/\delta m$ ) localised in the axially distant and relatively polarisable periphery (well away and self-shielded from the attraction of the nuclear centre) of each electron distribution.

I also discovered that Andrew Hurley's (1976) careful formal analysis of electron pair coupling had a single minor defect: he had assumed that (single-valued)  $L^2$  orbital representations of basis sets for correlation calculations in molecules could in principle exhaust electron correlation effects in a full-CI formulation. This led him to wrongly conclude that conventional correlation corrections in quantum chemistry should be attributed solely to reduced electron repulsion. As I confirmed directly with my earlier mentor, this conflicted with the need in correcting the energy of an HF-SCF model of an e-pair (e.g. He, or H<sub>2</sub>) to satisfy the Virial Theorem *both* at the level of the HF-SCF model representation (already Virial Theorem compliant at limit basis set), *and* at the level of the additionally correlated representation of the electronic ground-state of that same molecule. If the HF-SCF representation of He satisfies the Virial Theorem, and the true (correlated) ground state of He does also, then changes in *both* the one-electron terms  $\langle T \rangle$  (to elevate one electron kinetic energy terms) and the two-electron terms  $\langle V \rangle$  (to diminish electron interaction energies) of the correlation-corrected total energy expressions relative to the HF-SCF model must be required from the correlation corrections. Assigning the correlation correction solely to a reduction of electron repulsion without a corresponding change to increase one-electron kinetic energy (increased wave-function curvature), as Hurley had concluded, must be wrong.

The implication of this was that no single-centred  $L^2$  basis of reasonable size could expect to rapidly converge correlation energies, even in Full-CI models of that correlation, and that the weakly interacting systems we had been studying, embellished with off-nucleus 'probe' s, p sets, were reflecting *additional physical requirements* for basis set freedom in each atom's periphery.

### 2.4 Identifying the singlet-coupled spin wavefunction with the principal pair correlating natural orbital of each localised e-pair

There is a special case, that of singlet-coupled  $m_s = +1/2$  and  $-1/2$  spins in a two-electron atomic wavefunction, where one can explicitly write the angular behaviour of the spin coupling. Taking the specific case of ground state singlet of atomic He, the e-pair wavefunction can be written:

$$\text{He } (1s)^2 = 1s(1).1s(2) \{a(1).b(2)-b(1).a(2)\}/\text{Sqrt}2 \quad (1)$$

where  $a(1)$  and  $b(1)$  represent spin wavefunctions, each a two-valued angular function periodic in  $4\pi$  corresponding to  $m_s = +1/2$  and  $-1/2$  spins, respectively.

Burton (2009) shows a diagram of each angular contribution of the singlet-coupled two electron wavefunction, and interprets these contributions and their resulting interaction as peripheral surface disturbances oriented about the  $Z=0$  axis of the helicity equation, through the He nuclear position, in this case (Figure 1). Counter-propagating fractional electronic surface disturbances corresponding to  $m_s = +1/2$  and  $-1/2$  spin periodicity (over  $4\pi$ ) were shown there to interact in the physical space domain of their *relative angle* (over  $2\pi$ , i.e. single-valued in physical space) to yield a resultant quadrupolar form of oscillating angular coupling, transverse to the helicity axis, centred on  $Z=0$ .

Conventionally, the separation of spatial and spin variables in forming Slater determinantal representations of anti-symmetrised many-electron wave-functions leads spin to be represented and considered as a topological variable not directly interactive with the underlying spatial orbital of each spin-orbital, exactly as is being portrayed in Eq. 1. What Burton (1988, 2009) discovered instead is that a stable (circumferential) quadrupolar mismatch between the two otherwise spherical orbital forms for each electron arises in singlet coupling: the singlet coupling leads to a *physical resonance* in the counter-propagating angular forms of the two independent electrons in close-matched singlet coupling, where the respective interacting disturbances are built-in to the spin-orbital representation in a multiplicative manner. This resonance corresponds *simultaneously* to an increase in curvature of the underlying one-electron wavefunctions due to the circumferential disturbance, and a systematic circumferential mismatch between the two independent electrons. The former has the effect of increasing  $\langle T \rangle$  for the two electron wave-function (relative to HF-SCF with no spin representation) while the latter has the effect of decreasing the electron interaction energy  $\langle V \rangle$  due to the *systematic persistence* of the resonantly stable circumferential mismatch in the angular behaviour of the two electrons. Hurley's analysis determined the form of the correlation change in its impact on (changes in)  $\langle V \rangle$  but missed the necessary complement to it on (changes in)  $\langle T \rangle$  because he restricted his consideration (in complete consistency with universal practice at the time) to  $L^2$  orbital excitations as the correlating orbitals.

Burton's semiclassical interpretation of electron correlation suggests that one should deny any formal physical separability of spin effects from the underlying spatial wavefunction of each spin-orbital, rather to include them directly as surface modifications to those spatial forms.

## 2.5 Quantum spin coherence using an integrated spin-orbital representation for e-pair

In the models of physical interaction involving two well-separated electron pairs that Ulrich Senff and I considered, very large  $L^2$  orbital basis sets (including off-centred functions) were required to begin to converge the correlation energies of systems even as simple as He...H<sub>2</sub> or H<sub>2</sub>...H<sub>2</sub>, despite the fact that the HF-SCF wave-functions could be well-represented by modest DZ+P nuclear centred basis sets. The startling conclusion we are forced to reach, against eighty years of history in LCAO-MO models of quantum chemistry, is that it is a mistake to topologically separate spin from its physical impact on underlying spin-orbital representations of interacting electrons when these are closely singlet coupled. The implication of this in singlet coupling is that the multiplicative spin-coupled wavefunction form of each pair is *precisely* the form of the principal correlating natural orbital of the HF-SCF version of that pair.

Going back to the helicity equations for this model of He, we find that just a single ( $\rho/a_H$ ) parameter (see below) scales the off-axis form of the multiplicative disturbance due to spin

(Burton 1988, pp. 12-14), and as this varies, the relative *angular* curvature of each independent electron's wavefunction varies but so also does the phasic amplitude mismatch with the co-bound second electron. The optimum value of this parameter serves to restore the required Virial Theorem balance between  $\langle V \rangle_{\text{Exact}}$  and  $\langle T \rangle_{\text{Exact}}$  for the correlated ground state as correlation increases to alter these, in a balanced way but in the opposite sense, from reference HF-SCF values of  $\langle V \rangle_{\text{SCF}}$  and  $\langle T \rangle_{\text{SCF}}$ .

Of course, such a simplified one-parameter optimisation per pair of electron correlation in molecules comprised of relatively localised singlet coupled e-pairs (most of chemistry in fact, and virtually all of protein-dominated biology) would lend itself to very efficiently computed semi-empirical models of explicitly correlated many-electron wave-functions, with the major benefit of needing only HF-SCF level basis set representations, and all other virtual excitations in the CI being dispensed with in favour of a single correlating natural orbital representation for each localised orbital pair.

## 2.6 Coherent e-pair correlation function

Burton (1988) recounts the outcome of the Bernardi & Boys (1973) analysis of the pair correlation function employed in Boys & Handy (1969a, 1969b) exemplars of accurate direct correlation in Ne and LiH using their trans-correlated wavefunction formalism. Roby (1971, 1972) had already shown that the transcorrelated wavefunction and the cluster expansion wavefunctions for the Independent Pair Approximation (IPA) and the Coupled Pair Approximation (CPA) are related for spatially well separated pairs<sup>4</sup> by equivalence between the strong orthogonality constraint of pair theories and the canonical contraction conditions of the trans-correlated approach to the correlated wavefunction, when (see Note):

$$\mu_{ij}(1,2) = [f(1,2)-1] |\Psi_i(1).\Psi_j(2)| \quad (2)$$

Hurley commented on this supposed universal form to pair correlation  $f(1,2)$ :

'One would expect that an accurate description would require quite unrelated functions  $f(1,2)$  for the various pairs and that any restriction of the pair functions would seriously affect the correlation energy' (Hurley 1976; p.254)<sup>5</sup>.

It is precisely this surprising result from the detailed analysis of an effective direct correlated method, coupled with its generalisation from Burton (1988; see Equations 12 through to 23) to include non-single-valued Pair Natural Orbitals for the principal correlation effect circumferential to the axial that we conjoin to propose modelling e-pair correlation *universally*, at least for the case of axially symmetric pairs in  $|L,S\rangle$  representation. The general expression (Burton 1988, Eq.31) for the correlation function is given by Eq. 3:

$$f(1,2) = R_f(\rho_1,\varphi_1,\tau_1) R_f(\rho_2,\varphi_2,\tau_2) \mathbf{S}(\rho_1,\varphi_1,\tau_1; \rho_2,\varphi_2,\tau_2) \quad (3)$$

Burton (2009) advanced this semiclassical analysis suggested by Burton (1988) to *identify* the angular analysis required of the optimised semiclassical correlation interaction  $\mathbf{S}$  in [3] with

<sup>4</sup> Consistent with the van der Waals interactions involving He and H<sub>2</sub> that we were pursuing.

<sup>5</sup> This proposal has seemingly not been directly pursued, although numerical approaches along this line have been cast in terms of an 'inter-particle' distance operator (Boys & Handy 1969; Bernardi & Boys 1973), or seek to model particle position *and* momentum in joint distributions (see Gill, Crittenden et al. 2006).

the fully quantum mechanical, antisymmetrised singlet-coupling of the two-electron singlet spin function from Eq.1.

The crucial point of the argument at this stage is that the function  $\mathbf{S}$  in Eq.3, which is designed to establish and maintain coherence in the opposing sense of Ripple propagation for each electron of the pair, can, at least for stationary state wavefunctions, be **presumed** to have achieved this goal (Burton 1988, p.22). In our chosen context of the He (1s)<sup>2</sup> ground state, to respect the known topology of spin in antisymmetrised e-pairs, means that the angular degrees of freedom for the  $m_s = +/- \frac{1}{2}$  spin components in the e-pair interaction must be constrained. The angular form of the peripheral correlation interaction of the pair is fully determined by the *assumption* of singlet coupling. The resulting angular behaviour around the orienting axis of the 1s(1).1s(2) orbital product for ground state He is given in Figure 1.

### 2.7 Angular characterisation of a singlet-coupled antisymmetrised e-pair

Burton (2009) devised an angular characterisation of the  $m_s = +/- \frac{1}{2}$  spin e-pair interaction. Figure 1 shows the build-up of the resulting angular probability distribution for the circumferential singlet-coupled spin wavefunction  $\Phi(1,2)$  and the corresponding angular density function  $\Phi(1,2).\Phi^*(1,2)$  (shown in orange, showing peripheral quadrupolar form over  $2\pi$ ), where

$$\Phi(1,2) = \{ \sigma(1; \rho, \phi, a'_H) \sigma(2; \rho, -\phi, a'_H) - \sigma(2; \rho, \phi, a'_H) \sigma(1; \rho, -\phi, a'_H) \} / \text{sqrt } 2 \quad (4)$$

and

$$\Phi^*(1,2) = \{ \sigma(1; \rho, -\phi, a'_H) \sigma(2; \rho, \phi, a'_H) - \sigma(2; \rho, -\phi, a'_H) \sigma(1; \rho, \phi, a'_H) \} / \text{sqrt } 2 \quad (5)$$

The incorporation of the singlet-coupled spin wavefunction's angular dependence from Fig.1 leaves (for the ground-state e-pairs of He or H<sub>2</sub> we are considering) just a single axial function  $R_r$  dependent upon  $\rho_1=\rho_2$  being left to optimise. The radial mixing parameter determines the effective fraction of the e-pair in the periphery being subject to the physically (electromagnetically) coupled disturbance between the fractional matter-wave counter-propagations. Therefore, to variationally maximise the pair correlation energy for the pair corresponds to a variational optimisation of the mixing parameter depending upon  $\rho$ . As this function increases off-axis starting from  $\rho=0$  the radial behaviour of  $R_r$  for each electron begins with unit value (Burton 1988, see Equations 8, 9 and 10), and follows a scale parameter  $\rho^2/a_H^2$  which increasingly intersects the polarisable periphery of the e-pair off-axis as the mutually induced *and self-consistent* effective magnetic field  $H$  linking the circulation of the counter-propagating spins increases. Burton 1988 notes that the  $a_H$  parameter has the dimensions of distance and for a laboratory scale field of  $\mathbf{H} = 1$  Tesla has the value of 36.2617 nm. The variation of this single parameter to reach the exact correlation of the pair is therefore a simple variational optimisation.

Three decades after Hurley's mooted universal form of pair-correlation function (Eq. 2), in our realization of a direct matter-wave representation of the pair function (called the Quantum Spin Coherence model), we have an interesting new context in which to re-examine the systematics of electron correlation, including the amazing consistency of (1s)<sup>2</sup> pair correlation energies as diverse in size as in H<sup>-</sup> and U<sup>90+</sup> (around 1.1eV<sup>-1</sup>: see Gill, Crittenden et al. 2006), a long recognised fact of quantum chemistry, that on its own is suggestive of more regularity in the dynamics of electron pair correlation than had yet been appreciated.

$$\tilde{\varepsilon}_{ij} = \left(1 + \langle \hat{\mu}_{ij} | \hat{\mu}_{ij} \rangle\right)^{-1} \left\{ \left\langle \left| \Psi_i \Psi_j \right| \frac{1}{r_{12}} \left| \hat{\mu}_{ij} \right\rangle + \left\langle \hat{\mu}_{ij} \left| \frac{1}{r_{12}} \left| \Psi_i \Psi_j \right\rangle + \left\langle \hat{\mu}_{ij} \left| H_{ij} + E_{ij}^0 - E_0 \right| \hat{\mu}_{ij} \right\rangle \right\} \quad (12)$$

where  $E_{ij}^0$  is the energy of the core of the other (N-2) electrons,

and  $H_{ij}(1,2) = F_{ij}(1) + F_{ij}(2) + \frac{1}{r_{12}}$ , where  $F_{ij} = F - R_i - R_j$  and  $F$  is the Fock operator.

Noting that the form of the correlated wavefunction used in the Appendix is unnormalized and  $\langle \Psi | \Psi \rangle = 1 + \langle X | X \rangle$ . Assuming a normalized correlation function,  $\tilde{\mu}_{ij}(1,2)$ , such that

$$\hat{\mu}_{ij}(1,2) = \lambda_{ij} \tilde{\mu}_{ij}(1,2) \quad (13)$$

we have

$$\tilde{\varepsilon}_{ij} = \left(1 + \lambda_{ij}^2\right)^{-1} \left\{ \lambda_{ij} \left\langle \left| \Psi_i \Psi_j \right| \frac{1}{r_{12}} \left| \tilde{\mu}_{ij} \right\rangle + \lambda_{ij} \left\langle \tilde{\mu}_{ij} \left| \frac{1}{r_{12}} \left| \Psi_i \Psi_j \right\rangle + \lambda_{ij}^2 \left\langle \tilde{\mu}_{ij} \left| H_{ij} + E_{ij}^0 - E_0 \right| \tilde{\mu}_{ij} \right\rangle \right\} \quad (14)$$

Note from Burton 1988

## 2.8 Peripheral field dynamics of surface beats arising from coherent e-pair correlation: Quantum spin coherence

The multiplicative semi-classical pair-correlation function  $f(1,2)$  from Eq. 3 is to be identified in its angular behaviour  $\mathbf{S}$  about an axis defining of spin orientation with the multiplicative fully quantum mechanical, constrained antisymmetric topology of singlet coupling given by  $\Phi(1,2)$  from Eq. 4. Optimisation of the multiplicative interaction of the radial profile of pair correlation function's angular behaviour with the underlying orbital radial profiles is then a single parameter optimisation. This parameter is derived by Burton (1988; Equations 8-10) as  $\rho^2/a_H^2$  where the off-axis distance scale parameter  $a_H$  reflects the both effective semiclassical magnetic moment of the partial electron disturbances linked in coherent counter-propagation in the singlet coupling and the counter-propagation frequencies of that coupling. Since the coupling is magnetic in character, higher effective self-consistent fields in the coupling draw the effective radius of the circumferential counter-propagation inwards to intersect a larger fraction of the electrons of the pair being subject to correlation mismatch.

This entirely explicit and variationally-determined form of peripheral correlation interaction within an e-pair is what we denote Quantum Spin Coherence.

Its primary signature in the introduction of circumferential beats transverse to the spin axis, in the form of a dynamic quadrupolar interaction arising from the spatio-temporal mismatch between the coherently correlated electrons of the pair. This beat phenomenon corresponds in turn to a conservative standing wave whose impact is to effect an oscillating, transverse, quadrupolar magnetic field oriented along the spin axis of the pair (see Figure 1). Arising from singlet coupling for which no magnetic field residue would be expected in first order, this higher order side effect of pair correlation is an unprecedented form of diamagnetic field. Derived from very small fractions of electronic charge subject to interactive coupling in the correlation, the corresponding diamagnetic field is expected to represent one of the weaker fields of atomic physics.

The QSC model has one more surprise: the *ground state* of the pair correlation is what is being determined by the above identification with the conventional singlet coupling topology. Higher harmonics of this correlation stabilisation of the e-pair are possible.

### 2.9 The possibility of a spectral manifold of higher order harmonics of the singlet-coupled spin wavefunction for each e-pair

Burton (2009) saw several implications of the basic QSC model. In particular, as a physically large bound electron pair, He looks capable of subtending an extensive spectrum of singlet-coupled higher spin harmonics, for which evidence may in due course be found.

In the helicity equation model, a uniform magnetic field is assumed to apply to the electronic wavefunction. We can however imagine the two electron version of He represented in this manner beginning with zero spin coupling, in identical spherical orbital forms for each electron  $1s(1)$  and  $1s(2)$ . A minor fluctuation on the surface of one electronic wave-function will have an immediate polarising impact on the independent second electron, to generate a counteracting fluctuation (Maxwell's equations) in the space-time form of the second electron's wave-function. Two counter-propagating disturbances will resolve into a set of interaction harmonics in the circumferentially bounded system, of which the fundamental has the form already given by Eq. 1<sup>6</sup>.

The *higher harmonics* of this ground state singlet spin interaction have to my knowledge never before been contemplated. These will correspond semiclassically to smaller fractional-electron and more angularly intersected two-electron mismatch resonances, converging to the HF-SCF (zero spin) limit, as the effective fractional electron subject to mismatch diminishes asymptotically through the higher manifold states of this harmonic spectrum.

The spectral manifold of standing wave or harmonic representations of pair correlation, is bounded by the exact ground state (conventional or fundamental mode of singlet coupling) of the pair and the completely uncorrelated (HF-SCF) representation, where the effective correlation stabilisation from minuscule coherent correlation mismatch in the higher harmonics converges to zero. These figures are illustrated in Table 1A. where HF-SCF and exact ground state estimates for He  $(1s)^2$  are shown.

The He ground state correlation energy of 1.1439eV or 9227cm<sup>-1</sup> is located as the fundamental spin-coupling mode of the QSC stabilisation w.r.t. the uncorrelated HF-SCF representation of the electron pair (the latter being a purely electrostatic model of the e-pair). On the assumption that higher harmonics intersects less of the bulk of the e-pair distribution, and so represent both engagement of smaller fractional electronic charge and greater radial distance off-axis of the effective correlation, Table 1B. indicates via a semiclassical, first-order series how one might anticipate the spectrum of this manifold.

### 3. History of spin & spinors

Schrödinger (1931) associated the 'Zitterbewegung' (or *jitter*, an intricate internal structure of the electron) of the Dirac formulation with spin. According to Barut & Zanghi (1984) more than fifty years later: 'historically spin was introduced as a classically indescribable two valuedness,' but 'spin is the angular momentum of the Zitterbewegung.' For the

<sup>6</sup> The orientation of the circumferential counter-propagation and the resulting non-spherical quadrupolar beats (Fig. 1) about the axis of an appropriate atomic  $|L, S\rangle$  coupling scheme, implies the need for a compensating sphericalisation of the two-electron wavefunction by a  $J = 0$  rotational wavefunction.

**A. Helium e-pair energies for  $(1s)^2 / mE_h$**

vdW well		-0.063
$\epsilon_C$	-42.04	0.150%
$E_{\text{exact}}$	-2903.74	1.448%
$E_{\text{HF-SCF}}$	-2861.70	0.022%

$\epsilon_C$	42.04 mE <sub>h</sub>
	1.1439 eV
	9227 cm <sup>-1</sup>

$e_{\text{vdW}}$	0.06 mE <sub>h</sub>
	0.0017 eV
	14 cm <sup>-1</sup>

**B. Helium e-pair spin-harmonics for  $(1s)^2 / \text{cm}^{-1}$**

n	Nominal	$\Delta E$ (1h $\nu$ ) excitations	Level spacing	QSC harmonics	e-pair correlation	$\Delta E$ (2h $\nu$ ) excitations
1	-1.000	-		N	S=0 resonances	
2	-0.250	<b>-9226.7</b>		2	-0.250	<b>-9226.7</b> <b>-9.227</b>
3	-0.111	-4100.7	5.126			4100.7
4	-0.063	<b>-2306.7</b>	<b>6.920</b>	4	-0.063	<b>2306.7</b> <b>-2.307</b> 1476.3
5	-0.040	-1476.3	7.750			1476.3
6	-0.028	<b>-1025.2</b>	<b>8.201</b>	6	-0.028	<b>1025.2</b> <b>-1.025</b> 8.201
7	-0.020	-753.2	8.473			753.2
8	-0.016	<b>-576.7</b>	<b>8.650</b>	8	-0.016	<b>576.7</b> <b>-577</b> 8.650
9	-0.012	-455.6	8.771			455.6
10	-0.010	<b>-369.1</b>	<b>8.858</b>	10	-0.010	<b>369.1</b> <b>-369</b> 8.858
11	-0.008	-305.0	8.922			305.0
	{...}			...	...	...
	{SCF}	0.0	9227	{SCF}	0.0	{limit} <b>{9227}</b>

Table 1. **A.** Electron correlation energy of He  $(1s)^2$  electronic singlet configuration. **B.** QSC model of spin-coupled harmonics projected to lie between the uncorrelated SCF limit and the fully quantum-spin-coherent correlation of He  $(1s)^2$ . The lowest bound state of the spin-harmonic manifold (see Fig. 1) can be identified with the optimally correlated pair ( $N=2$ ), providing a pair-correlation ‘stabilisation energy’  $\epsilon_c$  of the constructive interference between ground state singlet coupled spin functions. The distance in energy below the uncorrelated HF-SCF orbital energy of other successively less correlated spin-harmonic states lie in a manifold which converges to the asymptotic state that lacks symmetry breaking i.e. the HF-SCF orbital pair. In the case of circular boundary conditions, the manifold of excited states is determined – to first order – as  $E_N = -\epsilon_c / N^2$ ;  $N$  even.

zitterbewegung itself one has, following Huang (1952): ‘the expectation value of the position (of a localized electron, includes) a time dependent part involving interference terms ...between positive and negative energy states of the electron and represents the Zitterbewegung.’ See also Barut & Bracken (1981) and Barut (1988).

### 3.1 Ongoing confusion about spin & spinors

Going back to the foundations of relativistic quantum mechanics (RQM), Dirac’s (1928a, 1928b) original operator formulation, following Pauli’s matrix algebra lead, produced *ab initio* both antimatter and a technical representation of spin, later named by Ehrenfest, as a *spinor*. In Tomonaga’s (1974/1977) account, p.128-130, the two component wave functions so produced (a *large* part corresponding to the hydrogen-like orbitals, and a *small* part, rather ill-defined, representing the spin) correspond to spinors whose physical interpretation still remains unclear even today: spinors form a “mysterious tribe.” According to M. Atiyah in Goddard (1998) celebrating Dirac: “unlike differential forms, which are related to areas and volumes, spinors have no such simple interpretation. They appear out of some slick algebra, but the geometrical meaning is obscure” (p. 113-4 in Goddard 1998).

Kutzelnigg (1984) attempted to apply orbital basis-set-expansion methods to Dirac’s equations of motion, only to produce non-physical, non-bounded variational states, together

with states low in the continuum<sup>7</sup> which appear equally unphysical, in part due to the matrix representation of the Dirac equation in a finite basis not having the correct non-relativistic limit (Schrödinger limit) for  $c \rightarrow \infty$ .

Mathematically, spinors correspond to *double-valued representations* of the rotation group representative of particles of half-integral spin (i.e. fermions), and are neither vectors nor tensors. This lack of clarity on the physical or geometrical structure of - not just electrons but all fermions - in respect of spin (see again Atiyah in Goddard (1998) and Tomonaga (1974/1977)) follows Schrödinger (1927) himself, reflecting upon what is “missing” in his original hydrogen atom model:

“in the language of the theory of electron orbits, the *angular momentum* of the electron round its axis, which gives it a *magnetic moment*. ...[there is] no doubt that, by the paradoxical yet happy conception of the spinning electron, the orbital theory will be able to master disquieting difficulties which latterly have begun to accumulate [in modeling atomic spectra, including the Paschen-Back effect of line splittings in strong magnetic fields]... We shall be obliged to attempt to take over the idea of Uhlenbeck and Goudsmit (1926) into wave mechanics. I believe that the latter is very fertile soil for this idea, since in it the electron is not considered as a point charge, but is continuously flowing through space, and so the unpleasing conception of a “rotating point charge” is avoided. In the present paper, however, *the taking over of this idea is not yet attempted.*”  
[emphasis at end added; Schrödinger 1927, p.64.]

The continuing lack of a clear physical- or geometrical-model of spin<sup>8</sup> has remained like a cloud over particle physics, and has allowed (R)QM to suffer interminable stress as to its interpretation, no less that that enforced by the ‘Copenhagen’ interpretation being extended from single particle wavefunctions to those for multiple particles. Spin is surely modern physics’ biggest single conceptual barrier: we know how it works; we just do not yet fully appreciate its physics.

### 3.2 Realisation of hidden variable theory with spin as a partial matter-wave

In the spirit of de Broglie & Einstein, Schrödinger (1926), Bohm & Bub (1966) and Bell (1987), our partial matter-wave formulation of spin leads directly to a one-parameter variational model of coherent correlation within each close-coupled e-pair. This model applies the mathematics of a freely propagating electron in a magnetic field to a quantum model of electron pairs. The QM helicity equations given by Landau & Lifschitz (1977) for a free electron in a homogenous field are taken as the template for the *radial* behaviour of the free electron.<sup>9</sup> For the *angular* behaviour, when applied to a non-propagating electron bound to a fixed nucleus, the model is developed with the aid of two assumptions.

The *first*, at last implementing Schrodinger’s suggestion of modeling spin via the electron matter-wave itself, contemplates some to-be-determined ‘volume’ *fraction* of the electron matter-wave being subject to coherent circumferential disturbance. This fraction  $\delta$  should correspond semiclassically to comparable fractions  $\delta e/e$  of the electronic charge and

<sup>7</sup> Where the ‘uncoupled’ or non-coherent spin representations of our Quantum Spin Coherence model would reside.

<sup>8</sup> As opposed to familiar mathematical-, topological-, Pauli matrix-, symmetry representation- or empirical angular momentum vector-representations of spin.

<sup>9</sup> This would generate an effective semi-classical orbiting frequency of 88.04GHz for a laboratory scale field of 1 Tesla, corresponding to a cyclotronic orbital radius of 36.26 nm.

fractions  $\delta\mathbf{m}/\mathbf{m}$  of the electronic mass, so that the *ratio* between these,  $\delta\mathbf{e}/\delta\mathbf{m}$  maintains the same *whole* charge to *whole* mass ratio  $\mathbf{e}/\mathbf{m}$  which (following the Uhlenbeck & Goudsmit hypothesis) enters the helicity equations on the presumption that the *whole* electron is spinning. This fraction  $\delta\mathbf{e}/\mathbf{e} = \delta\mathbf{m}/\mathbf{m}$  is determined variationally by the shape of the radial function off-axis, which is determined to deviate from the multiplicative value of unity, only in the peripheral and polarisable outermost regions of the nuclear-centred electron.

The *second*, based upon two counter-propagating electrons, involves determining the *relative phase* of the lowest stationary state coupling between a  $m_s = +\frac{1}{2}$  and a  $m_s = -\frac{1}{2}$  oriented spin circulations, whose amplitudes are periodic over  $4\pi$ , but whose relative angle values circulate for the pair in the real physical range of  $2\pi$ .

The formal equation for the intra-pair correlation interaction is provided for the QSC model by Eq. 31, p.22 of Burton (1988), where the dynamic correlation is taken (following Hurley's prompting) to conform to a universal form, given here as Eq. 3 for ease of reference.

Consideration of the continuity boundary conditions over  $2\pi$  around the periphery, and variation of the relative phase of the coupling, leads by way of *constructive interference* between the spins to an  $l=2$ , quadrupolar beat phenomenon in the circumferential periphery of the electron pair, orthogonal to the spin axis. The result of the *angular form* of the pair correlation formula for the least energy coupling in axially symmetric situations is essentially available by inspection, as the Figure 1 reveals.

In the absence of the QSC insight, a rather more elaborate computational strategy, via the appropriate RQM equations provided in their quaternion formulation of 'scale relativity' by Celeriere & Nottale (2003, 2005)<sup>10</sup>, might be extended to fully accommodate the spin-spin interaction of QSC so long deleted (in advance) as lacking physical "significance" in Dirac's equation for the spin-spin-interaction of two electrons (Breit, 1932):

"The modification required consists in the removal of terms in  $e^4$  in the result of eliminating the "small" components of the wavefunction. These terms do not involve  $\hbar$  so that their physical significance is doubtful even for this reason alone."(Breit 1932, p.617) ... "The coupling of motions of particles due to their interactions with transverse waves is small compared to that due to electrostatic forces."(p.621)

It remains true that in QSC these effects are small compared to coulomb interactions, but hardly insignificant: thus RQM calculations adopting the Breit Hamiltonian also ignore correlation effects See Whittaker (2007).

#### 4. Interaction fields between QSC-correlated He atoms: Gravity?

Spectral theory suggests that the He spin-manifold is capable of propagating its field of influence relatively far. Burton (2009) suggested that cool astrophysically distributed He gas can mediate a cooperatively aligned physical interaction through the interstellar "aether" of He which orients the quadrupolar resonant form of individual singlet coupled He atoms, so forming a background electromagnetic field in which all 'matter' is bathed. Gravity thus could

<sup>10</sup> While the Celerier & Nottale's 'scale relativity' quaternion framework is not set up in a manner in which the current spinor model has obvious application, or vice versa, the framework admits of *deriving* the Schrödinger and Dirac equations as special cases, and also admits *non-differentiable* functions (such as our two-valued spin functions). I anticipate that the QSC model will represent a highly pre-diagonalised form of basis representation of equations accessible via Celerier & Nottale's equations of motion.

be a side-effect of basic chemistry in cool distributed gas clouds, and so be fundamentally quantum mechanical in origin: if so, it is also *emergent* in the history of the Universe.

#### 4.1 The strange weakness of the gravitational background

It remains of outstanding current interest, many years after Einstein's (1916) formulation of the space-time topology of General Relativity, how gravity can be integrated with quantum mechanics (QM, now eighty five years old: Landau & Lifshitz 1977; Merzbacher 1970, 1992) in general and with the standard model at the *atomic* (QED) and *sub-nuclear* (QCD) levels of particle physics (Fritzsche 1992), also both decades old. The nature of the gravitational field, of relative scale forty orders of magnitude weaker than the coulombic force and - apparently - universally *attractive* between macroscopic bodies, continues to challenge particle theorists, whose 'strings' and 'branes' with 10-dimensional support seek to capture its elusive essence, so far without success. The apparently universal background force field of gravity, empirically identified by Newton more than three centuries ago and re-described as a topology in coupled space-time by Einstein almost a century ago, has remained *unexplained*<sup>11</sup> beyond attributing the action-at-a-distance force of unknown fundamental origin to the interaction of classical 'masses'. What is actually *causing* Einstein's curvature of space-time?

The modern focus of investigations on gravity itself include: (i) galactic modeling, which raises a question of deviation from perfect inverse-square behavior of the gravitational force field, or else the existence of some form of so-far unobserved (or identified) 'dark matter' throughout the universe<sup>12</sup>; (ii) the search for, observation of, and characterization of properties of massive interstellar gravitational 'black holes' which assume that gravity does not change in form at small distances; and (iii) the search for direct observational evidence of 'gravitational waves'.

So far, gravitational waves (GW) have eluded direct detection, although their existence has been inferred from measurement of the orbital decay of the binary neutron stars PSR 1913 + 16 (Taylor & Weinberg 1989). Two huge experimental consortia aimed at directly observing astronomical gravitational waves are currently underway, both focused on observing such *dynamic perturbations* to the gravitational background: LIGO<sup>13</sup> is just attaining its design sensitivity and discrimination to detect gravitational pulsation from binary pulsars such as PSR 1913 +16; and LISA<sup>14</sup>, which spectrally complements LIGO, being readied for 2013 flight and aimed at detecting gravitational waves from galactic binaries and super-massive black hole binaries.

---

<sup>11</sup> As David Hume in 1739 already recognized in *A Treatise of Human Nature*, of Newton's *Principia Mathematica*, that he advanced a description rather than an explanation of gravity. Einstein's topological reformulation of gravity into an observer independent tensor formulation of curved space-time does not escape the same charge.

<sup>12</sup> See Da Rocha & Nottale (2003) and Nottale (2004), who instead propose another explanation.

<sup>13</sup> Laser Interferometer Gravitational Observatory (Gustafson, Shoemaker et al.1999; Allen, Anderson et al. 2001); multiple earthbound laser interferometer pairs of up to 4 kilometre pathlength, positioned thousands of kilometers apart, with high signal-to-noise efficiency in the 10Hz-10kHz region. See Laser Interferometer Gravitational Wave Observatory, LIGO-T0100017-00-Z, [www.ligo.caltech.edu](http://www.ligo.caltech.edu), (23 Feb 2001); see [http://www.ligo-wa.caltech.edu/ligo\\_overview/ligo\\_overview.html](http://www.ligo-wa.caltech.edu/ligo_overview/ligo_overview.html) and <http://www.ligo.caltech.edu/advLIGO/>.

<sup>14</sup> Laser Interferometer Space Antenna (Tinto & Dhurandhar 2005); a giant triangulation of three identical spacecraft  $5 \times 10^6$  km apart, to observe and detect cosmic low frequency GWs not visible from earth, in the spectral region  $10^{-4}$ Hz to 1Hz.

LIGO (version II) and the spaceborne LISA is designed capable of resolving changes of 1 part in  $10^8$  of an interference fringe relative to feedback-controlled dark-fringe resonance in the carrier 180W Nd:YAG 1064nm laser beams, so is capable of resolving differential GW impact on (30kg sapphire) test-masses situated remotely in each arm of the interferometer down to  $10^{-19}$  m, i.e. a rather small part of a single atom's  $10^{-15}$  m nuclear radius. They supposedly may also detect a 'continuous wave' background signal that comprised the Newtonian Background 'gravity gradient' of gravity. This could appear as a 'stochastic' source, a seemingly random waveform generated by a large number of weak, independent sources of gravitational radiation. One possible source of such a signal could be emissions from the early universe – the unthermalised gravitational equivalent of the cosmic microwave background radiation (CMBR). Searches in stochastic backgrounds are conducted by cross-correlating the outputs of two or more interferometers.

The *particle physics* complement of gravity, Higgs bosons exchanging information between masses, remain as yet unidentified<sup>15</sup>, so we lack insight into the frequency domain of the Higgs field<sup>16</sup>, although the *symmetry* of the interaction, by way of an especially intriguing (to this work) 'transverse quadrupolar' field, has been determined. No physical candidate for such dynamics has yet emerged from string theory and other models directed at incorporation of gravity into a unified four-fundamental-force 'theory-of-everything'.

One major issue is the relative strength of gravity in relation to the electromagnetic, weak-nuclear and strong-nuclear forces, given by the dimensionless scaling constants:  $a_g \sim 3 \times 10^{-41}$ ;  $a_e \sim 1/137$ ;  $a_w \sim 1/30$ ,  $a_s \sim 1/3$ . So, while macroscopic inverse-square/curved space-time understanding of gravity is clear enough from its success in planetary and galactic modeling, in the context of fundamental QM and particle physics models, its attractive nature yet its extraordinary weakness remains deeply mysterious to science: one could say, as mysterious as consciousness itself<sup>17</sup>.

#### 4.2 From explicit spinor correlation in atomic He to Meso-Scale interstellar gravitation?

*Rewriting Basic Physics of Spin*: A geometrically specific spinor model in physical, matter-wave terms is devised. A *coherent analytical* model of electron-pair correlation is introduced, invoking the phenomenon of Quantum Spin Coherence (QSC) as a Virial-theorem-compliant – and computationally scalable -- correlation strategy in quantum chemistry. Superficial e-pair spin-coupling *beats* are identified as a *constructive interference* in coherent counterposed spin coupling as a consequence of the model. A spin-harmonic manifold of possible spin-coupled stationary states for (axially) symmetric e-pair is predicted. Novel electromagnetic force-fields, including a 'transverse, quadrupolar' pulsating magnetic field

<sup>15</sup> CERN's \$2.3B Large Hadron Collider (LHC) has two 'general purpose' detectors, *Atlas* and the *Compact Muon Solenoid*, both of whose teams are aiming to identify the elusive Higgs boson. The Higgs boson, 'known as the "God particle" because of its importance to the Standard Model' [BBC commentary], provides an account in particle physics of why all other particles have mass.

<sup>16</sup> The Higgs mechanism is considered to fill all of space with a relatively uniform field, whose energy is regarded as the best candidate for the 'dark energy' of the Universe. 'Dark energy' is thought to make up the greater part (~70%) of the Universe, and be the 'cause' of the accelerating expansion of the Universe, following findings in 1998 by two teams studying supernovae.

<sup>17</sup> Burton, P.G. (2011). *Human Cognition: higher brain function and the science of human consciousness*. ISBN 9781456307400. See [www.createspace.com/3494821](http://www.createspace.com/3494821)

attending coherent e-pair correlation, are adduced, the latter being a form of diamagnetism, *dynamic diamagnetism*, which is a direct implication of the QSC model. This is a novel, attractive and resonant, form of diamagnetism limited to conspecific peer e-pair interactions.

*The Special Role of He:* Atomic helium is a major element in interstellar regions, and present in cool molecular clouds which are astrophysically significant in the formation of new stars. The QSC phenomenon of coherent correlation is particularly significant in the isolated e-pair case of ground state He (1s)<sup>2</sup> because, even though the correlation in the K-shell of all atoms has a similar magnitude (Section 2.8), the small positive charge on the He nucleus means that the electron distribution of this e-pair is especially large, and therefore able to subtend a more extended spectral manifold of harmonics of the fundamental singlet-coupled ground state configuration than from any other atom. This large size and high peripheral polarisability in the He case in turn means that the penetration distance of field interaction involving resonant information exchange between distant conspecific peer atoms is also unusually great.

*He...He interaction:* Isolated He atoms are completely isotropic (Footnote 6). However, once He atoms begin to weakly interact, their natural internal correlated structure, which in QSC is axially oriented, becomes 'exposed' by the interaction. The propagation of the correlation field of each e-pair acts to co-orient and so slightly stabilise the relative interaction. Specifically, the anisotropic interaction of the transverse quadrupolar beat phenomenon tends to co-orient interacting He atoms. Nearby He...He interaction is capable of collisional stabilisation to create a very weakly bound state of He<sub>2</sub> (Luo, McBane et al. 1993) which may amplify orientational alignment in the general He...He interaction field, but the model being proposed does not rely upon this potential amplification of the basic phenomenon. Quantum-spin-coherently sourced, dynamic diamagnetic interactions between e-pair peers are then active even in extremely dilute interstellar He at the direct & local (micro-), proximally co-operative (meso-) and regional source-aggregated (macro-) level of consolidation to form a cooperative and coherently implicative force field.

*What is Mass?--from Spin Coherent Cooperativity to Gravity:* Macro-level dynamic diamagnetic cooperativity sourced *chemically*, between e-pairs in He atoms, acts as a symmetry-compliant source--in fact the *fundamental explanatory source*--of the dynamics of the primary gravitational interaction. Implications for current searches for gravitational waves are considerable and implications of the quantum-spin-coherence model in particle physics, including the removal of 'mass' as an *intrinsic* property of matter, are to be noted.

## 5. Implications of the QSC model of spin & spin-pairing

The adoption and use of coherent partial matterwave electron representation of the spin component of spin-orbitals in chemistry would have many implications. It provides a new conceptual basis to:

1. Satisfy every equation in Landau & Lifschitz or Merzbacher texts on Quantum Mechanics on spin, by substituting fractional  $\delta e/\delta m$  wherever  $e/m$  appears in the formula for the electron-magnetic moment.
2. Variationally optimise electron pair correlation as a coherent matter-wave phenomenon between opposing spins paired in each orbital (i.e. as a purely local, single parameter optimization), with results which, in contrast to most correlation algorithms, are

consistent with the requirements of the Virial theorem, yet scale reasonably with molecular size because of the intrinsically local nature of their impact.

3. Understand the electron magnetic moment at the first semiclassical level as a fractional electron, peripheral circulation effect  $\delta e/\delta m = e/m$ , and at a deeper quantum level, interpret the idea of any fundamental particle having as an empirical attribute something called mass being derived from the strength (low mass) or weakness (high mass) of that particle's interaction from the universal background field mediated astrophysically by distributed yet implicate He atoms.
4. Understand gravity (described but not explained by Einstein, 1916) as an intrinsically coherent, quantum-mechanical indirect interaction between apparent 'masses', whose effect is transmitted onto the local particle from implicate aggregation of field effects due entirely to (intrinsically stable, coherent correlations within) the electronic part of the wave function of remote He-bound spin-pair couplets. In this formulation, mass becomes a secondary and empirical result of engagement of any (electromagnetic) particle with a primary *dynamic diamagnetic* magnetic field interaction between distant but implicate electron peer pairs. The empirical appearance of mass is then understood as the result of implicate cooperation in a background micro-field alignment at the meso-scale, where each micro-field is essentially derived from QSC coupling in correlated e-pairs, and where at large or asymptotic distances the pulsating magnetic field lines are essentially collinear. In this formulation, there is little/no role/contribution of the (electronically shielded) *nuclear* 'hadrons' in aggregating such inertial drag on a particle, at least from nuclear-shielded neutral atoms: the effect is essentially *electronic* in its nature and being derived from atom-based interactions, *chemical* in origin.
5. Tidy up the unification of the four forces of physics, to bring gravity into a formulation of R(QM). This is goal Einstein couldn't see a way to reach, and which is not in principle possible to achieve, without his first having taken QM seriously enough in this context (see also item 12.) to understand that gravity could be *caused* by close-coupling in pair interactions. In this way, gravity is positioned as an *outcome* of quantum entanglement<sup>18</sup>.
6. Convert to heresy the 'Copenhagen' dogma on non-deterministic QM interpretation, substituting a realized form of matter-wave spin following the heritage of Einstein, de Broglie, Schrödinger, Bohm & Bell, and in line with the present conjecture, describe exactly & precisely those few previously 'hidden' spin 'variables' that have been overlooked as a result of the repressively 'dogmatic' Copenhagen view (and along with it the persistent semiclassical, non-determined, determinantal interpretation) of QM.
7. Advance an account of the physical origin of the three colour forces in QCD (Fritzsch, 1992) as deriving from three complementary spatial orientations of spin interactions.
8. Develop new models the nature of the meson in the strong force, the neutrino in the weak force, and understand that the proton must not be immune from decay.
9. Predict the rest mass of the electron-antineutrino (0.575ev), as being displaced or decoupled from its coherent origin as a spin-only, low mass and low charge (virtual) partner in a coherently coupled fermionic correlation of an e-pair.
10. Understand that matter and antimatter is exactly in balance in our universe - that hadrons account for all the original positronic antimatter becoming captured in metastable QSC variant spin couplings with electron counterparts.

---

<sup>18</sup> In QSC, entanglement as *topology* becomes matter-wave *interaction*.

11. Recognise *dynamic* diamagnetism as a new fundamental force of nature, to account for the origin of an apparent, empirically attractive gravitational force, by being responsible ultimately for both internal 'mass' of any particle/ body/'matter'/planet/star/galaxy (i.e. drag w.r.t. the implicate background field) and, by directed difference, apparent gravitational forces between those bodies.
12. Quantum spin-coherence emerges, via the representational power of the 'ripple' model of spin, as a natural, principal and *determinable* solution of the 'small' parts of the quaternionic form of the relativistic quantum equations of motion, as provided by Celerier & Nottale (2003, 2005).

Predictions are also made concerning (a) the existence of, and (b) unique observational fingerprint of, hitherto unsuspected excited harmonics spin-states of electronic singlet coupling. These QSC spin harmonic states, occurring as previously unrecognized structure within the ground state electronic wave-function of an electron pair, are not yet observed. The special case of He (1s)<sup>2</sup>, because of its size subtending an exceptionally large manifold of such harmonics, as illustrated in Table 1, and because of its ubiquity, becomes of outstanding astrophysical interest.

## 6. Directions for future research

The Quantum Spin Coherence model of the phenomenon of electron correlation entails a novel conception of realised electron spin and a new physical and geometric interpretation of the spinors of relativistic quantum mechanics. A simple and direct physical interpretation of the most fundamental Stern-Gerlach experiment arises, originated by the chaotic Zitterbewegung of an originally non-magnetised electron interacting with the inhomogeneous magnetic field to generate its magnetic moment. The fundamental philosophy of modern quantum theory is revised by this model, which defends a realization of the 'hidden-variables' matter-wave interpretation of quantum mechanics. The model is also defensible by consistency with two modern *ab initio* formulations of the applicable quaternion equations of relativistic quantum mechanics.

Three independently disprovable aspects of this reconception of spin and correlation are immediately apparent. Firstly, the ability to model individual filled-shell electron correlation by an essentially exact analytic model with a single variational dimension is available, whereby the *universality* of the pair correlation function  $f(1,2)$  suggested here can be assessed. The simplest test of the model itself is whether its prediction of the achievement of the exact energy of He (1s)<sup>2</sup> by this strategy is met. Secondly, a previously unrecognized quantisation and energy-state manifold in ground-state e-pair coherent spin coupling is predicated, including a manifold of states as I have predicted semiclassically for He. Direct spectroscopic transitions will fail to reveal this previously unsuspected intra-atomic manifold because of forbidden singlet-singlet selection rules, however more condensed states including liquid He (and especially He<sub>2</sub>) may reveal previously unexpected atomic manifold structure in phonon bands in response to shock. Thirdly, the model predicts previously unsuspected coherent magnetic complement of a newly coherent QSC formulation of the weak dispersion interaction between e-pairs, despite their spin-singlet coupling which should (conventionally) close out the possibility of any residual magnetic interaction. Previously, dispersion interaction had been posed as being a rather chaotic interaction between instantaneous dipolar field arising in each structure. Along with a coherent, line-of-sight dipolar *electrical* interaction between remote e-pairs (corresponding to a revised understanding of the van der Waals interaction), a new kind of dynamic magnetic field provides a secondary, transverse, quadrupolar form of e-pair

interaction. This purely quantum phenomenon is called dynamic diamagnetism, and it is fundamentally *attractive* in nature, especially between otherwise identical coupled pairs, in contrast to all previously known forms of diamagnetism which give rise to repulsive forces of interaction, including the Meissner effect in superconductivity. Finally, the model predicts that both complementary fields of any e-pair interaction are fundamentally orientationally anisotropic. [The single exception lies in a hypothetical situation of an e-pair in perfect isolation, where the overriding  $J=0$  rotational homogenization applies.]

In the QSC framework, gravity, apparently the simplest and certainly the first recognized of the four modern forces of nature, is derived as a force of purely quantum-mechanical origin. This result provides the first account of integration of all four forces in nature being quantum mechanical in origin. Einstein's approach (Davies 1995; Feynman 1998; White & Gribbin 2005) to developing a rapprochement between gravity and quantum theory, as though these were fundamentally independent of one another, thus could not have succeeded. Although the incorporation of relativistic concepts into quantum mechanics is *necessary* to explain gravity (through the phenomenon of electron spin and the implications of its singlet coupling) it has, until now, not been *sufficient* to explain gravity. It remains to the future therefore to consider appropriate formulation of the 'equations of matter-wave motion' in which the QSC model provides its insights without requiring the assumption of the empirical phenomenon of particle mass. However, it is extremely unlikely that the LHC design will ever see the ultra-weak signature of bound e-pair interactions.

In the QSC formulation of gravity, gravitational forces cannot be present at the origination of the Universe. Gravitational waves, as a composition of a stabilising gravitational background field and a multiplicity of dynamic perturbations from non-equilibrium meso-scale gravitational activity, are the current target of gravitational wave observatories, based upon laser interferometry. For the LIGO experiment, observations may well detect meso-scale perturbations relative to the background, but not the background itself, there being no gravitational complement of the CMBR in the QSC framework. The additional degrees of freedom in the forthcoming LISA experiment should be capable, subject to certain conditions, of observing the proposed intrinsic anisotropy of QSC gravitational forces at meso-scale.

## 7. References

- Ahlich, R., Lischka, H., Staemmler, V., & Kutzelnigg, W. (1975a). PNO-CI and PNO-CEPA studies of electron correlation effects. *J. Chemical Physics* 62, pp. 1225-1234
- Ahlich, R., Driessler, E., Lischka, H., & Staemmler, V. (1975b). *J. Chemical Physics* 62, 1235.
- Allen, B., Anderson, W., Bose, S., Christensen, N., Daw, E., et al. (2001). Laser Interferometer Gravitational Wave Observatory, LIGO-T0100017-00-Z, 23 Feb 2001. see [www.ligo.caltech.edu](http://www.ligo.caltech.edu).
- Barut, A.O. & Bracken, A.J. (1981). Zitterbewegung and the internal geometry of the electron. *Physical Review D* 23, pp. 2454-2463.
- Barut, A.O. & Zanghi, N. (1984). Classical model of Dirac electron. *Physical Review Letters* 52, 2009.
- Barut, A.O. (1988). Combining relativity and quantum mechanics: Schrodinger's interpretation of  $\Psi$ . *Foundations of Physics* 18, pp. 95-105.
- Bell, J. S. (1987). *Speakable and unspeakable in Quantum Mechanics*, Cambridge, Cambridge University Press (see especially papers 1-5, 7, 8 & 17-19).

- Bernardi, F. and Boys, S.F. Explicit Formula solutions of the contraction conditions for transcorrelated wavefunctions. *Molecular Physics* (1973) 25, 35-44
- Bohm, D. & J. Bub. (1966). A proposed Solution of the measurement Problem in Quantum Mechanics by a Hidden Variable Theory. *Reviews of Modern Physics*, 38, pp. 453-469. (NB: Sections 5, 7 & 8).
- Bohm, A. (1978). The rigged Hilbert space and quantum mechanics. (*Springer*) *Lecture Notes in Physics* 78.
- Boys, S.F. & Handy, N.C. *Proceedings of the Royal Society* (1969) A309, 209; A310, 43
- Boys, S. E, & Bernardi, E. (1970). The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics* 19, 553
- Breit, G. (1932). Dirac's equation and the spin-spin interaction of two electrons. *Physical Review* 39, pp. 616-624.
- Burton P. G. (1977). The computation of intermolecular forces with Gaussian basis functions. Illustration with He<sub>2</sub>. *Journal of Chemical Physics* 67, pp. 4696-4700
- Burton, P. G. (1979). Design of basis sets for precise intermolecular force calculation. Investigation of the He<sub>2</sub> potential curve using CEPA-PNO correlated wavefunctions. *Journal of Chemical Physics* 70, pp. 3112-3118
- Burton, P. G. (1980). The basis set completion problem in potential energy surface calculation. In P.G. Burton (Ed.) *Molecular Physics and Quantum Chemistry: into the 80's. Proceedings of the International Workshop on Atomic & Molecular Physics & Quantum Chemistry*. University of Wollongong pp. 1-79
- Burton, P. G.; Gray, P. D. & Senff, U. E. (1982). Basis set refinement in theoretical modelling of molecular electronic structures. *Molecular Physics* 47, pp. 785-809
- Burton, P. G. (1988). Explicitly Bi-Temporal Electrons. (unpublished manuscript: original<sup>19</sup> pp. 39 manuscript available as a PDF in unmodified form downloadable from archive <http://public.me.com/blinkcentral> in Quantum Spin Coherence (QSC) folder)
- Burton, P. G. (2009). Gravity with Spin: Quantum Spin Coherence as an *Ab Initio* Generator of Gravitational Attraction. *Proceedings of Space, Propulsion & Energy Sciences International Forum: SPESIF*, AIP Conf. Proc. 1103, pp. 267-275. doi:10.1063/1.3115505, see <http://link.aip.org/link/?APCPCS/1103/267/1>
- C el erier M.N. & Nottale L. (2003). A scale-relativistic derivation of the Dirac Equation. *Electromagnetic Phenomena* 9, 83. Special issue dedicated to the 75th anniversary of the discovery of the Dirac equation (arXiv:hep-th/0210027).
- Cel erier, M.-N. and Nottale. L., (2005). Dirac Equation and Scale Relativity. arXiv:hep-th/0112213 (version 14 Nov 2005 pp33)
- Costa, L.F. & Herdeiro, C.A.R. (2008). A gravito-electromagnetic analogy based on tidal tensors. GRG18 presentation arxiv:gr-qc/0612140 v2 21May2007
- Da Rocha, D. & Nottale, L. (2003). Gravitational Structure Formation in Scale relativity. In *Chaos, Solitons and Fractals on "New Cosmology"* 16, pp. 565-595, see arXiv:astro-ph/0310036 (v2 3 Nov 2003)
- Davies, P. (1995). *About Time: Einstein's Unfinished Revolution*. London: Penguin Books.

---

<sup>19</sup> The original Ripple or semiclassical model of electron correlation was presented in Chemistry Department seminars at the University of Wollongong and Monash University late in 1988, and at University College, London in 1989. In Feb 2005, I arrived at predicting the energy of excited harmonics of spin-coupled pairs (i.e. as illustrated in Table 1), so making a first experimental prediction from the new theory at AIP 17 2006 & GRG18 2008 conference presentations.

- Dirac, P.A.M. (1928). The Quantum Theory of the Electron. I. *Proceedings R. Soc. London* A117, pp. 610-624.
- Dirac, P.A.M. The Quantum Theory of the Electron. II. *Proceedings R. Soc. London* (1928) A118, p. 351.
- Einstein, A. (1916). *Annalen der Physik* 49, 769. Translated in *The Principle of Relativity*, Dover (1923, 1952)
- Feynman, R.P., (1948). Space-Time Approach to Non-Relativistic Quantum Mechanics. *Reviews of Modern Physics* 20, pp. 367-387
- Feynman, R. (1985). *QED: the strange theory of light and matter*. Princeton: Princeton University Press
- Feynman, R. P. (1998). *Six Easy Pieces: The Fundamentals of Physics Explained*. London: Penguin Books
- Fritzsch, H. (1992). *Quarks: the stuff of matter*. London: Penguin Books
- Gill, P.M.W.; Crittenden, D.L.; O'Neill, D.P. & Besley, N.A. (2006). A family of intracules, a conjecture and the electron correlation problem. *Physical Chemistry Chemical Physics* 8, pp. 15-25
- Goddard, P. [ed.] (1998). *Paul Dirac: the man and his work*, Cambridge, Cambridge University Press.
- Grant, I.P. (1983). Self-consistency and numerical problems. In G. Malli (Ed) *Relativistic Effects in Atoms Molecules and Solids*. (NATO ASI B87; Plenum) pp.101-113.
- Gsponer, A. & Hurni, J.-P. (2005). Lanczos-Einstein-Petiau: from Dirac's Equation to nonlinear wave mechanics. In W.R Davis, et al., *Cornelius Lanczos Collected Published papers with Commentaries* (Nth Carolina State Uni., Raleigh, 1998) Vol III, pp. 2-1248 to 2-1277. (see arXiv:physics/0508036 v2, orig.)
- Gustafson, E., Shoemaker, D., Strain, K. Weiss, R. LSC white paper on detector research and development, LIGO-T990080-00-D (11 Sep 1999)
- Higgs, P.W., (1964). Broken symmetries and the masses of gauge bosons, *Physics Review Letters* 13, pp. 508-9.
- Huang, K. (1952). On the Zitterbewegung of the Dirac electron. *American Journal of Physics* 20, 479.
- Hurley, A.C. (1976). *Electron correlation in small molecules*. NY: Academic Press.
- Hylleraas, E.A. (1930). Über den Grundterm der Zweielektronenprobleme von H-, He, Li<sup>+</sup>, Be<sup>++</sup> usw. *Zeitschrift Physik* 65, pp. 209-225.
- Hylleraas, E.A. and Midtdal, J. (1956). *Physical Review* 103, 829-830 and Linderberg, J. (1961). *Physical Review* 121, 816-819.
- Kolos, W.; Szalewicz, K. & Monkhorst, H.J. (1986). New Born-Oppenheimer potential energy curve and vibrational energies for the electronic ground state of the hydrogen molecule. *J. Chemical Physics* 84, pp. 3278-3283
- Kutzelnigg, W. (1984). Basis set expansion of the Dirac operator without variational collapse. *Int. J. Quantum Chemistry* 24, pp. 107-129
- Lanczos, C. (1929). *Zeitschrift Physik* 57, 447
- Landau, D.M. & Lifschitz, E. M. (1977). *Quantum Mechanics 3<sup>rd</sup> Ed*. Oxford, Pergamon
- Lewis, G. N. (1916). The Atom and the Molecule. *J. Am. Chem. Soc.* 38, 762-785. doi:10.1021/ja02261a002
- Luo, F., McBane, G. C., Kim, G., Giese, C. F.; Gentry, W. R. (1993). The weakest bond: Experimental observation of helium dimer. *Journal of Chemical Physics* 98, 3564-3567

- Nottale, L. (2004). The Theory of Scale Relativity: Non-Differentiable Geometry and Fractal Space-Time. In CASYS'03 – Sixth International Conference, Liege, Belgium Aug 2003. *AIP Conference Proceedings* (2004) 718, 68-95.
- Roby K. R. (1971). On the theory of electron correlation in atoms and molecules: Relation between cluster expansion theory and the correlated wave functions method. *International Journal of Quantum Chemistry* 5, pp. 119-130
- Roby K. R. (1972). On the theory of electron correlation in atoms and molecules. II. General cluster expansion theory and the general correlated wave functions method. *International Journal of Quantum Chemistry* 6, pp. 101-123
- Schrödinger, E. (1927/1978). *Collected papers on wave mechanics*. NY: Chelsea Publishing. (Translated from second German edition, 2<sup>nd</sup> Ed.: see first paragraph, p. 64. )
- Schrödinger, E. (1931). Zitterbewegung. *Sitzungber. Preuss. Akad. Wiss., Phys.-Math. Kl.* 24, 418; 3, 1.
- Senff, U. E. & Burton, P. G. (1985). An ab initio study of the isotropic and anisotropic potential surfaces of the He-H<sub>2</sub> interaction. *Journal of Physical Chemistry* 89, pp. 797-806
- Senff, U. E. & Burton, P. G. (1986). A CEPA2 investigation of the He-He and He-Li<sup>+</sup> potential functions. *Molecular Physics* 58, pp. 637-645
- Senff, U. E. & Burton, P. G. (1989). A CEPA2 study of the H<sub>2</sub>-H<sub>2</sub> isotropic potential function. *Australian Journal of Physics* 42, pp. 47-58. See [http://www.publish.csiro.au/?act=view\\_file&file\\_id=PH890047.pdf](http://www.publish.csiro.au/?act=view_file&file_id=PH890047.pdf)
- Shavitt, I. (1977). Graph theoretical concepts for the unitary group approach to the many electron correlation problem. *International Journal of Quantum Chemistry* S11, 131-148.
- Shavitt, I. (1981). The graphical unitary group approach and its application to direct configuration interaction calculations. In: Hinze J, ed. *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements (Lecture Notes in Chemistry, No 22)*. Berlin: Springer; pp. 51-99.
- Sinanoglu, O. (1961). Many-Electron Theory of Atoms and Molecules. *Proc. Natl. Acad. Sci. U S A.* 47(8), pp. 1217-1226
- Gerlach, W.; Stern, O. (1922). Das magnetische Moment des Silberatoms. *Zeitschrift für Physik* 9: 353-355. Bibcode 1922ZPhy...9..353G. doi:10.1007/BF01326984
- Taylor, J.H. and Weisberg, J.M. (1989). Tests of relativistic gravity using the binary Pulsar PSR 1913+16. *Astrophys. J.* 345, pp. 434-450
- Tinto, M. & Dhurandhar, S.V. (2005). Time-delay interferometry. *Living Rev. Relativity*, 8.
- Tomonaga, S. I. (1974/1997). *The Story of Spin* (trans. T. Oka). Chicago: Chicago University Press
- Uhlenbeck, G.E. & Goudsmit, S. A. (1925). Erzeugung der Hypothese vom unmechanischen Zwang durch eine Forderung bezüglich des inneren Verhaltens jedes einzelnen Elektrons. *Naturwissenschaften* 13, 953-4
- Uhlenbeck, G.E. and Goudsmit, S. A. (1926). Spinning Electrons and the Structure of Spectra. *Nature* 117 264-5
- Whitaker, M.A.B. (2007). Solomon's Argument on Hidden Variables in Quantum Theory, *Foundations of Physics* 37, 989-997
- White, M. & Gribbin, J. (2005). *Einstein: A Life in Science*. London: The Free Press.

# Toward Automatic Regularization for Feynman Loop Integrals in Perturbative Quantum Field Theory

E. de Doncker<sup>1</sup> and F. Yuasa<sup>2</sup>

<sup>1</sup>*Western Michigan University, Kalamazoo, MI 49008-5466*

<sup>2</sup>*High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801*

<sup>1</sup>*U.S.A.*

<sup>2</sup>*Japan*

## 1. Introduction

The motivations of high energy physics collider experiments include the precise measurement of parameters in the standard model (Passarino & Veltman, 1979; van Oldenborgh & Vermaseren, 1990) and beyond. In recent years, the precision of high energy experiments at colliders has increased significantly as a result of developments in electronics. The detection of any deviations of the experimental data from the theoretical predictions may lead to the study of new phenomena. In modern physics, everything is composed of elementary particles, and there are four basic interactions acting on particles: gravitation, and the weak, electromagnetic and strong interactions. When we consider a scattering process of elementary particles, the cross section reflects the dynamics which govern the motion of the particles, caused by the interaction.

All information pertaining to a particle interaction is contained in the amplitude according to the (Feynman) rules of Quantum Field Theory. Generally, with a given particle interaction, a large number of configurations (represented by Feynman diagrams) are associated. Each diagram represents one of the possible configurations of the virtual processes, and it describes a part of the total amplitude. The square sum of the amplitudes delivers the probability or cross section of the process (by integration). Based on the Feynman rules, it is the goal to obtain the amplitude using the steps listed in Figure 1.

Feynman diagrams are constructed in such a way that the initial state particles are connected to the final state particles by propagators and vertices. Particles meet at vertices according to a coupling constant  $g$  which indicates the strength of the interaction. The amplitude is expanded as a perturbation series in  $g$ , where the leading (lowest) order of approximation corresponds to the tree level of the Feynman diagrams. The evaluation of tree diagrams is well known and analytical formulations exist, which have been developed into automatizations of Figure (1) and are heavily used in high energy physics. For the tree level, packages such as GRACE, COMPHEP, CALCHEP, FEYNARTS/FEYNCALC/FORMCALC, MADGRAPH, FDC, and so on, are available.

- (i) Specify the physics process (external momenta and order of perturbation);
- (ii) draw all Feynman diagrams relevant to the process;
- (iii) determine the contributions to the amplitude.

Fig. 1. Scheme for computation of the cross section amplitude

Without the higher perturbative orders, which require the evaluation of loop diagrams, the theory is inept at modeling the experimental observations precisely. The computation of loop integrals allows the inclusion of higher order terms for perturbation calculations of the amplitude in quantum field theory. For the one-loop order an analytic treatment is established and has been implemented in several automatic computation systems which are in various stages of testing. Currently available packages include FEYNARTS/FEYNCALC/LOOPTOOLS (van Oldenborgh & Vermaseren, 2000), GRACE-1LOOP (Fujimoto et al., 2006; Yasui et al., 2007), XLOOPS-GINAC (Bauer, 2002), GOLEM/SAMURAI (Binoth et al., 2009; Heinrich et al., 2010), HELAC-NLO (van Hameren et al., 2009), and so on. However there is no analytical method for calculating higher order corrections in a systematic way, especially for higher than two-loop electroweak corrections including three or four legs (three-point or four-point) and general mass configurations. Therefore semi-numerical or fully numerical approaches have been an important topic of study for many research groups.

We have developed a novel, fully numerical method, DCM (*Direct Computation Method*), to evaluate loop integrals based on a combination of numerical integration and numerical extrapolation techniques. Developed originally for cases without infrared (IR) or ultraviolet (UV) divergences, DCM uses nonlinear extrapolation to regulate singularities caused by vanishing integrand denominators in the interior of the integration domain. Furthermore, a regularization of IR divergence caused through boundary singularities has been enabled by an extension of DCM.

In this paper we describe the DCM regularization techniques. Subsequently, Section 2 reviews formalism and notations pertaining to Feynman diagrams and loop integrals. After introducing the basic DCM method in Section 3, we discuss IR divergence (and the DCM regularization to handle it) in Section 4, while giving examples and numerical results. Section 5 presents concluding remarks and avenues for future work.

## 2. Feynman diagrams and loop integrals

The general form of an  $L$ -loop integral with  $N$  propagators is given by

$$\mathcal{I} = \int \prod_{j=1}^L \frac{d^D l_j}{i\pi^{D/2}} \prod_{r=1}^N \frac{1}{\mathcal{D}_r} \quad (1)$$

where  $D$  is the space-time dimension, and the integration is over the loop momenta  $l_j$ ,  $j = 1, \dots, L$ . Here  $\mathcal{D}_r$  is the denominator of the  $r$ -th propagator,

$$\mathcal{D}_r = q_r^2 - m_r^2 + i\delta, \quad (2)$$

where  $q_r$  and  $m_r$ ,  $r = 1, 2, \dots, N$  are the momentum of the propagator and the mass of the  $r$ -th internal particle, respectively, and  $i\delta$  is an infinitesimal term (with positive  $\delta$ ), which prevents the integral from diverging if  $\mathcal{D}$  vanishes inside the integration domain. The latter happens in the *physical region*, where the total squared energy  $s$  of the colliding particle system exceeds a threshold so that the reaction can actually take place in the real world. Below this threshold, in the *unphysical region*, the integral satisfies the representation (1) with  $\delta = 0$ . By the generalized Feynman identity (Nakanishi, 1971),

$$\frac{1}{\mathcal{D}_1 \mathcal{D}_2 \cdots \mathcal{D}_N} = (N-1)! \prod_{j=1}^N \left( \int_0^\infty dx_j \right) \frac{\delta(1 - \sum_{r=1}^N x_r)}{(\sum_{r=1}^N x_r \mathcal{D}_r)^N}, \tag{3}$$

where the  $\{x_r\}$  are Feynman parameters. Note that, in view of the  $\delta$ -function and  $x_r \geq 0$  for each  $r = 1, \dots, N$ , we have  $0 \leq x_N = 1 - x_1 - \dots - x_{N-1}$ , so that  $0 \leq \sum_{r=1}^{N-1} x_r \leq 1$ . Thus, if  $x_N$  is eliminated in terms of the other coordinates, the integration region reduces to the  $N - 1$ -dimensional unit simplex,  $\{\mathbf{x} \mid \sum_{j=1}^{N-1} x_j \leq 1, x_r \geq 0 \text{ for } 0 \leq r < N\}$ . However, in the following we will leave the  $\delta$ -function in the integrand, and set the upper bounds of the integration to 1.

After the loop momentum integrations of Eq. (1), the scalar loop integral  $\mathcal{I}$  is given in (Binoth & Heinrich, 2004) as

$$\mathcal{I} = \Gamma(N - DL/2) (-1)^N \int_0^1 \prod_{r=1}^N dx_r \delta(1 - \sum_{r=1}^N x_r) \frac{\mathcal{U}^{N-(L+1)D/2}}{\mathcal{F}^{N-LD/2}}. \tag{4}$$

Eq. (4) is called a Feynman parametric integral. The expressions of  $\mathcal{U}$  and  $\mathcal{F}$  can be constructed according to the topology of the corresponding Feynman diagram. The function

$$\mathcal{U}(\mathbf{x}) = \sum_{T \in \mathcal{T}_1} \prod_{j \in \mathcal{C}(T)} x_j$$

is a sum of monomials, each a product of variables corresponding to the loop lines (edges) which are cut in the graph to form the tree  $T$ . The tree is called a *1-tree*,  $T \in \mathcal{T}_1$  (= the set of all such 1-trees). The set of lines that was cut constitutes a *chord*  $\mathcal{C}(T)$ , which corresponds to a monomial of degree  $L$  in the Feynman parameters. For example, each 1-tree of a diagram consisting of one loop is formed by cutting one line of the loop, thus  $\mathcal{U}(\mathbf{x}) = \sum_{j=1}^N x_j = 1$  for this diagram.

A *2-tree*  $\hat{T}$  consists of two disconnected trees which are obtained by cutting an edge of a 1-tree,  $\hat{T} \in \mathcal{T}_2$  (= the set of all such 2-trees), and the corresponding chords determine monomials of degree  $L + 1$  in the Feynman parameters. Then, the function  $\mathcal{F}_0$  is defined as

$$\mathcal{F}_0 = \sum_{\hat{T} \in \mathcal{T}_2} \left( \prod_{j \in \mathcal{C}(\hat{T})} x_j \right) (-s(\hat{T})),$$

where  $s(\hat{T}) = (\sum_{j \in \mathcal{C}(\hat{T})} p_j)^2$  is a Lorentz invariant. The denominator function  $\mathcal{F}(\mathbf{x})$  is given by

$$\mathcal{F}(\mathbf{x}) = \mathcal{F}_0(\mathbf{x}) + \mathcal{U}(\mathbf{x}) \sum_{j=1}^N x_j m_j^2.$$

In this paper we consider three types of divergences which may occur in the evaluation of loop integrals. The first is a divergence which occurs when the denominator vanishes in the integration region due to the specific configuration of the masses and external momenta. The other two are the *infrared* (IR) divergence, which appears when some of the internal particles are massless, and *ultraviolet* (UV) divergence, which is due to the contribution incurred from very large loop momenta.

The functions  $\mathcal{U}$  and  $\mathcal{F}$  in (4) are positive semi-definite functions of the Feynman parameters. The vanishing of  $\mathcal{U}$  is determined by the topology and is related to the UV subdivergences of the graph. On the other hand, the vanishing of  $\mathcal{F}$  depends not only on the topology but also on the kinematical parameters, and may (or may not) lead to an IR divergence.

### 3. Direct Computation Method

The imaginary part  $i\delta$  in the denominator of the propagator, Eq. (2), is introduced in the formalism as an infinitesimal quantity, to prevent the integral from diverging. In that sense, it plays a role as a regulator. For a numerical computation we consider  $\mathcal{I}$  in (4) as a function of  $\delta$ , and obtain the integral in the limit as  $\delta$  tends to zero.

For this limiting process, we construct a sequence of approximations,  $I(\delta_l)$ ,  $l = 0, 1, \dots$ , for a decreasing sequence of  $\delta_l$  such as  $\delta_l = \delta_0 / A_c^l$ ,  $A_c > 1$ , and extrapolate to the limit. For the extrapolation or convergence acceleration of a sequence  $S(\delta)$  to the limit  $S$  as  $\delta \rightarrow 0$ , we rely on the existence of an asymptotic expansion

$$S(\delta) \sim S + a_1\varphi_1(\delta) + a_2\varphi_2(\delta) + \dots = S + \sum_{j \geq 1} \varphi_j(\delta), \quad (5)$$

where the  $\varphi_j$  functions are arranged in decreasing order of  $\delta$ , so that  $\lim_{\delta \rightarrow 0} \frac{\varphi_{j+1}(\delta)}{\varphi_j(\delta)} = 0$ . A linear extrapolation method can be used if the  $\varphi_j(\delta)$  are known functions of  $\delta$ . It yields solutions to linear systems of the form

$$S(\delta_l) = a_0 + a_1\varphi_1(\delta_l) + \dots + a_\nu\varphi_\nu(\delta_l), \quad l = 0, \dots, \nu, \quad (6)$$

of order  $(\nu + 1) \times (\nu + 1)$  (in the unknowns  $a_0, \dots, a_\nu$ ) for increasing values of  $\nu$  (Brezinski, 1980; Lyness, 1976). For example, the  $\varphi_k(\delta)$  functions may constitute a sequence of integer powers of  $\delta$ , which play the role of the coefficients in the linear system (6). The goal is to combine values of  $S(\delta)$  for decreasing values of  $\delta$ , in order to eliminate terms from the error expansion of  $S(\delta_l) - a_0$ .

Apart from geometric and harmonic sequences of  $\delta$ , we have explored the *Bulirsch sequence* (Bulirsch, 1964), of the form  $\delta = 1/b$  with  $b = 2, 3, 4, 6, 8, 12, 16, \dots$  (consisting of powers of 2, alternating with  $1.5 \times$  the preceding power of 2). The type of sequence selected influences the stability of the process, as for Romberg type integration, which was found to be more stable with the geometric sequence than with the harmonic sequence (the Bulirsch sequence ranging in between) (Lyness, 1976). On the other hand there is a trade-off with the computational expense of  $S(\delta)$ , which may become prohibitive for fast decreasing  $\delta$ , especially in multivariate applications. For the computations we can furthermore use versions of  $b_l$  starting at, e.g.,  $b_0 = 1, 3$  or  $1/8$ .

If the functions of  $\delta$  in the asymptotic expansion (5) are not known, a nonlinear extrapolation may be suitable (Ford & Sidi, 1987; Levin & Sidi, 1981; Sidi, 1979; Wynn, 1956). In that case

we use the  $\epsilon$ -algorithm (Shanks, 1955; Wynn, 1956), which can be applied under more general conditions than those allowing linear extrapolation, if a geometric sequence is used for  $\delta$ , for example with  $\varphi_j$  of the form  $\varphi_j(\delta) = \delta^{\alpha_j} \log^{k_j}(\delta)$ , where  $\alpha_j > 0$  and integer  $k_j \geq 0$ . The actual form of the underlying  $\delta$ -dependency does not need to be specified for the  $\epsilon$ -algorithm. As a disadvantage, the latter does come with additional cost (compared to linear extrapolation) in terms of the number of elements  $S(\delta_l)$  needed for the elimination of parts of the error expansion.

This prescription is implemented as DCM and a schematic view of the program flow of DCM is shown in Fig. 2. For unphysical kinematics, it is sufficient to only carry out the

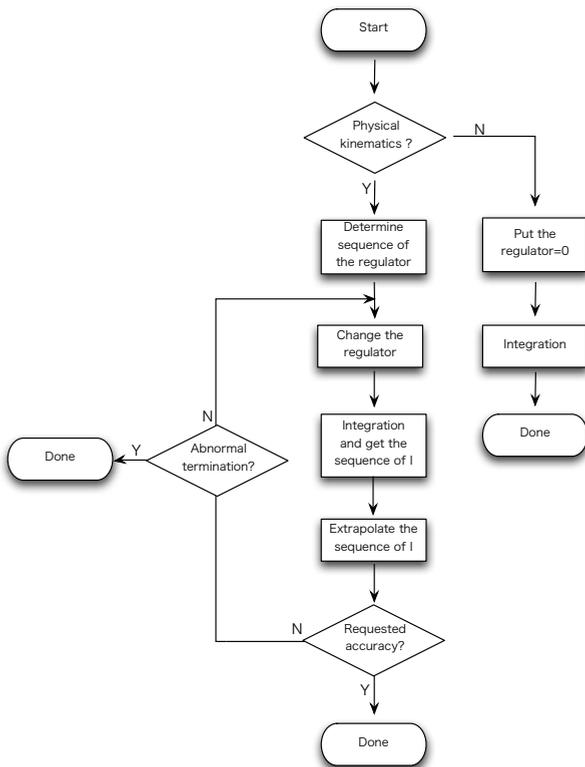


Fig. 2. Program flow of DCM.

multi-dimensional integration with  $\delta = 0$ . While no extrapolation is required in this case, the procedure will nevertheless produce a result, although it is less efficient. If an extrapolation is performed under valid conditions, it is the goal of the implementation in Fig. 2 to produce sequences which converge faster than the original, until convergence is detected within the requested accuracy and maximum number of steps. Otherwise an abnormal termination is flagged.

For the numerical integration, we use the QUADPACK (Piessens et al., 1983) routines DQAGE or DQAGSE. These are 1D automatic, adaptive integration algorithms, which apply

Gauss-Kronrod quadrature rules on the subintervals resulting from the adaptive partitioning. The one-dimensional quadrature algorithms are applied in an iterated scheme for the multi-dimensional integration (de Doncker & Kaugars, 2010; Li et al., 2004).

It is interesting to note that the numerical integration sequence can often start with a fairly large  $\delta$  value, such as  $\delta_0 = 1.2^{40}$  (for a geometric sequence with base 1/1.2), which allows dealing with a less singular behavior of the integrand, and can greatly simplify the integrations through the sequence. We are examining how to set the initial value of  $\delta$  heuristically, depending on the mass values occurring in the integrand denominator.

So far we have applied DCM effectively for one-loop three-, four-, five- and six-point loop integrals with masses (de Doncker, 2003; de Doncker et al., 2010; 2011; de Doncker, Shimizu, Fujimoto & Yuasa, 2004; de Doncker et al., 2006; de Doncker, Shimizu, Fujimoto, Yuasa, Cucos & Van Voorst, 2004; Yuasa et al., 2007; 2008). We have also applied it successfully to two-loop two-, three- and four-point loop integrals with masses (see, e.g., de Doncker et al. 2011; 2006; Yuasa et al. 2011; 2008; 2010). Thus even though the integrands in these cases suffer from detrimental singular behavior due to the configuration determined by the actual kinematical parameters, DCM was clearly able to regulate this type of divergence by *finite*  $\delta$ .

#### 4. IR divergence and regularization

IR divergence corresponds to the presence of massless internal particles like photons or gluons. We will consider two prescriptions to regularize IR divergence. The first is the more natural procedure, classified in a broad sense as *mass regularization* in (Muta, 2010), by introducing a small fictitious mass  $\lambda$  for the massless internal particles; and the other is the *dimensional regularization* technique which was originally proposed to control UV divergence. With respect to the mass regularization prescription, our procedure for the integration of the IR diagram is as given by Fig. 2, applied to configurations where some masses are very small. Since  $\lambda$  in the denominator is fixed, no extension of DCM is needed and the procedure is straightforward. In dimensional regularization, the space-time dimension  $D$  in Eq.(4) is replaced by  $4 + 2\epsilon$  and it is assumed that  $\epsilon \rightarrow 0$ . This results in a Laurent expansion of the integral as a function of  $\epsilon$ . It should be noted that two regulators appear, ( $\delta$ ,  $\lambda$ ) for mass regularization, and ( $\delta$ ,  $\epsilon$ ) for dimensional regularization, in order to handle integrals with divergence due to physical kinematics as well as IR divergence.

It is well-known that IR divergence cancels out in well-defined physical quantities (e.g., (Kinoshita, 1962; Muta, 2010; Nakanishi, 1971)), but the regularization is required to replace each IR divergent integral in the formalization of the physical quantity by a mathematically well-defined quantity. *The regularization is removed after achieving the cancellation of the IR divergence in the physical quantity* (Muta, 2010).

Mass regularization plays a role in verifying calculations of the total cross section. As an example, consider the  $e^+ e^- \rightarrow e^+ e^- \gamma$  process (for only the QED interaction), resulting in eight tree diagrams and 100 one-loop diagrams, of which 80 diagrams are IR divergent. By the cut-off  $k_c$  let us denote an energy threshold of the emitted photon as follows. If the energy of the photon exceeds  $k_c$ , it is a *hard photon*; with energy below  $k_c$ , it is a *soft photon*. The total cross section  $\sigma(\lambda, k_c)$  is a sum of: (1) loop diagram, (2) soft-photon diagram, and (3) hard-photon diagram contributions. Using (1) and (2), the independence of  $\lambda$  can be checked; and independence of  $k_c$  can be verified with (2) and (3).

### 4.1 Regularization by fictitious mass

In the mass regularization based on the procedure of Fig. 2, the dependence on  $\delta$  may dominate the behavior of the denominator in Eq. (4) when  $\lambda$  is very small. In (de Doncker et al., 2005; Yuasa et al., 2007), we investigated the behavior of the results and found that the DCM method breaks down for extremely small values of  $\lambda$  (less than  $10^{-15}$  GeV in double precision arithmetic). The point of deterioration can be moved, to some extent, by performing the calculations in quadruple precision. For example, results of up to 8-figure accuracy are obtained using quadruple precision for the IR divergent four-point diagram ( $\gamma\gamma$  box diagram), with  $m = 0.510^{-3}$  GeV and  $\lambda = 10^{-15}$  GeV and with the center of mass energy  $s = (p_1 + p_2)^2 = 500^2$  GeV<sup>2</sup> (de Doncker et al., 2005), whereas the corresponding double precision results are only accurate to about two figures.

Furthermore using the extended precision library HMLIB (Fujimoto et al., 2006), which is based on the IEEE 754-1985 floating point standard, results are provided in (Yasui et al., 2007; Yuasa et al., 2007) for the one-loop three-point diagram and  $\lambda$  as small as  $\lambda = 10^{-160}$  GeV, and for the one-loop four-point diagram with  $\lambda \geq 10^{-30}$  GeV. The validity range of DCM thus supports the conventional IR regularization technique using extended precision arithmetic for this process.

### 4.2 Dimensional regularization

In the scheme of dimensional regularization, a new idea is required since the second regulator  $\epsilon$  appears in the exponent of the integrand denominator. Let us denote the value of the integral for a fixed  $\delta_l$  and  $\epsilon_k$  by  $I(\delta_l, \epsilon_k)$ . Then we apply DCM of Fig. 2 to calculate  $\lim_{\delta \rightarrow 0^+} I(\delta, \epsilon_k)$ , which we will denote by  $I(0, \epsilon_k)$ . It is our goal to use a sequence of  $I(0, \epsilon_k)$ ,  $k = 0, 1, \dots$  to approximate the coefficients of leading order terms in the Laurent expansion of  $I(0, \epsilon)$  with respect to  $\epsilon$ .

This use of DCM comprises a *double extrapolation* technique where both  $\delta \rightarrow 0$  and  $\epsilon \rightarrow 0$ . A high-level description is outlined in the algorithm of Fig. 3. Its application is demonstrated in subsequent examples. The while loop condition in the algorithm enforces a test on the maximum number of iterations and may implement a check on convergence of the  $\epsilon$ -extrapolation.

```

    Double_extrapolation_algorithm {
    Set initial  $\epsilon$ 
    //  $\epsilon$  extrapolation loop
    while ( $\epsilon$  extrapolation conditions) {
         $I(0, \epsilon) \leftarrow$  DCM( $\epsilon$ )
        Obtain leading term coefficients
        Set next  $\epsilon$ 
    }
    }
    
```

Fig. 3. Algorithm for IR divergent loop integral

### 4.2.1 Example 1

First we present the scalar massless one-shell one-loop four-point diagram ( $L = 1, N = 4$ ) as a simple example. The integral is given by

$$I_4 = \int_0^1 d^4x \frac{\delta(1 - \sum_{j=1}^N x_j)}{(-sx_1x_3 - tx_2x_4 - i\delta)^{2-\epsilon}},$$

with  $s = (p_1 + p_2)^2$  and  $t = (p_1 + p_3)^2$ . Using  $\sum_{j=1}^N x_j = 1$  and  $x_4 \geq 0$ , the integral becomes

$$I_4 = \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \int_0^{1-x_1-x_2} dx_3 \frac{1}{(-sx_1x_3 - tx_2x_4 - i\delta)^{2-\epsilon}}. \quad (7)$$

With the *sector decomposition* technique (Binoth & Heinrich, 2004; Fujimoto & Ueda, 2008), Eq. (7) yields

$$I_4 = 2 \int_0^1 d^3x \frac{x_1^{-1+\epsilon} x_2^{-1+\epsilon} (1 + x_1 + x_2 + x_1x_2x_3)^{-2\epsilon} + (1 + x_1 + x_2(x_1 + x_3))^{-2\epsilon}}{(-s - tx_3 - i\delta)^{2-\epsilon}} \quad (8)$$

$$+ x_1^{-1+\epsilon} \frac{(1 + x_1 + x_2 + x_1x_3)^{-2\epsilon}}{(-s - tx_2x_3 - i\delta)^{2-\epsilon}} + (s \leftrightarrow t),$$

via successive sector decompositions, and elimination of the  $\delta$ -function (symmetrically with respect to the coordinate variables). The ( $s \leftrightarrow t$ ) part consists of the terms listed in the integrand with  $s$  replaced by  $t$  and vice-versa. By expanding the non-singular numerators around  $x_1$  and/or  $x_2 = 0$ , the IR singularity then manifests itself through a pole at  $\epsilon = 0$ . The coefficients of the Laurent expansion

$$I_4 = \frac{C_{-2}}{\epsilon^2} + \frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon) \quad (9)$$

can be collected as sums of multi-dimensional integrals. However, the asymptotic behavior of Eq. (9) as a function of  $\epsilon$  gives a basis for an approximation of the coefficients  $C_0, C_{-1}, C_{-2}$  by extrapolation. For an approximation of  $C_{-2}$  we can use

$$\tilde{C}_{-2} = I(\epsilon) \epsilon^2 \sim C_{-2} + C_{-1}\epsilon + C_0\epsilon^2 + \dots \quad (10)$$

as  $\epsilon \rightarrow 0$ . Thus we proceed by constructing a sequence of  $\tilde{C}_{-2}(\epsilon_l) = I(\epsilon_l) \epsilon_l^2$ , and use the  $\epsilon$ -algorithm (Shanks, 1955; Wynn, 1956) for numerical extrapolation. For  $s, t < 0$  the denominators in Eq. (8) do not vanish and the integral exists for  $\epsilon > 0$ . Results of the computations according to Eq. (10) for  $\tilde{C}_{-2}$  with  $s = t = -1$ , for  $\epsilon_k = 2^{-k}, k \geq 0$ , together with the extrapolated values, are shown in Table 1.

For the evaluation of  $C_{-1}$  we do not require another sequence of integrals to be computed since we can use the  $\tilde{C}_{-2}$  sequence divided by the constant  $\epsilon$ ,

$$\tilde{C}_{-1} = I(\epsilon) \epsilon = \frac{\tilde{C}_{-2}}{\epsilon} \sim \frac{C_{-2}}{\epsilon} + C_{-1} + C_0\epsilon + \dots$$

$k$	$I(\epsilon_k)\epsilon_k^2 = \tilde{C}_{-2}$	Extrapolated ( $C_{-2}$ )
0	0.2467401100272340E+01	
1	0.3915839671912849E+01	
2	0.4356784142166274E+01	0.454976991318E+01
3	0.4322263111559589E+01	0.432476950067E+01
4	0.4203435233964723E+01	0.437091943007E+01
5	0.4113097380184850E+01	0.382664876650E+01
6	0.4059492230026028E+01	0.399594399188E+01
7	0.4030494087959442E+01	0.399987052327E+01
8	0.4015435528831175E+01	0.399999594538E+01
9	0.4007765070784543E+01	0.40000011407E+01
10	0.4003894385147564E+01	0.40000000650E+01
11	0.4001950158104918E+01	0.40000000197E+01
		Exact: 4.0

Table 1. Extrapolation results,  $C_{-2}$  and  $\tilde{C}_{-2}$  of Eq. (10)

Whilst  $\{\tilde{C}_{-1}(\epsilon_k)\}$  is a divergent sequence, the  $\epsilon$ -algorithm is able to extrapolate to the value of  $C_{-1}$  (barring roundoff). Furthermore, we have

$$\tilde{C}_0 = I(\epsilon) = \frac{\tilde{C}_{-2}}{\epsilon^2} \sim \frac{C_{-2}}{\epsilon^2} + \frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon).$$

For this simple example, the analytic formulae are known for  $C_{-2}$ ,  $C_{-1}$  and  $C_0$  (Fujimoto & Ueda, 2008), e.g.,

$$C_{-2} = 4 \int_0^1 dx \frac{1}{(-s - tx - i\delta)^2} + (s \leftrightarrow t).$$

As shown in Table 1, the result by DCM with  $\epsilon$ -extrapolation agrees with the analytic result for  $s = t = -1$ .

### 4.2.2 Example 2

Here we consider the tensor integral of the massless one-loop three-point integral ( $L = 1$  and  $N = 3$ ) of rank  $M \leq 3$  (Kurihara, 2005),

$$T_{\mu \dots \nu}^{(3)} = \sum_i C_{\mu \dots \nu}^i J_3^i(p_1^2, p_2^2, p_3^2; n_x^{(i)}, n_y^{(i)}),$$

where

$$J_3^i(p_1^2, p_2^2, p_3^2; n_x^{(i)}, n_y^{(i)}) = \frac{1}{(4\pi)^2} \frac{\epsilon \Gamma(-\epsilon)}{(4\pi\mu_R^2)^\epsilon} \int_0^1 dx \int_0^{1-x} dy \frac{x^{n_x^{(i)}} y^{n_y^{(i)}}}{\mathcal{D}^{1-\epsilon}}, \tag{11}$$

$$\mathcal{D} = (p_1x - p_2y)^2 - \rho xy - p_1^2x - p_2^2y - i0,$$

$$\rho = p_3^2 - (p_1 + p_2)^2.$$

In the case  $p_1^2 = p_2^2 = 0$  and  $p_3^2 \neq 0$  with  $n_x = n_y = 0$ , this is

$$\begin{aligned} J_3(0, 0, p_3^2; 0, 0; \epsilon) &= \frac{1}{(4\pi)^2} \frac{\epsilon \Gamma(-\epsilon)}{(4\pi\mu_R^2)^\epsilon} \int_0^1 dx \int_0^{1-x} dy \frac{1}{(-p_3^2 xy - i0)^{1-\epsilon}} \\ &= \frac{\epsilon \Gamma(-\epsilon)}{(4\pi)^2} \left( \frac{-\tilde{p}_3^2}{(4\pi\mu_R^2)} \right)^2 \frac{1}{-p_3^2} \frac{B(\epsilon, \epsilon)}{2\epsilon} \end{aligned} \quad (12)$$

where  $\tilde{p}_3^2 = p_3^2 + i\delta$  and  $\mu_R$  is the renormalization energy scale. As  $\epsilon \rightarrow 0$ ,  $J_3$  of Eq. (12) satisfies the expansion

$$J_3(0, 0, p_3^2; 0, 0; \epsilon) \sim \frac{C_{-2}}{\epsilon^2} + \frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon), \quad (13)$$

where  $C_{-2}$ ,  $C_{-1}$  and  $C_0$  are given by

$$\begin{aligned} C_{-2} &= \frac{1}{(4\pi)^2 p_3^2}, \\ C_{-1} &= \frac{1}{(4\pi)^2 p_3^2} \ln(-p_3^2), \\ C_0 &= \frac{1}{(4\pi)^2 p_3^2} \left( -\frac{\pi^2}{12} + \frac{1}{2} \ln^2(-p_3^2) \right). \end{aligned}$$

This result is used subsequently, for the case  $p_1^2 = 0$ ,  $p_2^2 \neq 0$  and  $p_3^2 \neq 0$  with  $n_x = n_y = 0$ . The IR divergent integral

$$\begin{aligned} J_3(0, p_2^2, p_3^2; 0, 0; \epsilon) &= \frac{1}{(4\pi)^2} \times I_3, \\ I_3 &= \frac{\epsilon \Gamma(-\epsilon)}{(4\pi\mu_R^2)^\epsilon} \int_0^1 dx \int_0^{1-x} dy \frac{1}{(-(p_3^2 - p_2^2)xy - p_2^2 y(1-y) - i0)^{1-\epsilon}} \end{aligned} \quad (14)$$

is expressed as

$$J_3(0, p_2^2, p_3^2; 0, 0; \epsilon) = J_3(0, 0, p_3^2; n_x n_y; \epsilon) \times \mathcal{G}_{n_x=0}(z) \quad (15)$$

where  $z = \frac{p_3^2 - p_2^2}{p_3^2}$  and the function  $\mathcal{G}_{n_x=0}(z)$  is defined as

$$\mathcal{G}_{n_x=0}(z) = {}_2F_1(1, 1 - \epsilon; 2; z) \times \epsilon. \quad (16)$$

Here  ${}_2F_1$  is the Gauss hypergeometric function, which has the Euler integral representation,

$$F(\alpha, \beta; \gamma; z) = {}_2F_1(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \int_0^1 \frac{t^{\beta-1}(1-t)^{\gamma-\beta-1}}{(1-tz)^\alpha} dt,$$

where  $\Re\gamma > 0$ ,  $\Re\beta > 0$ .

For  $n_x = n_y = 0$ , Eq. (13) and Eq. (16) are used to expand Eq. (15) as

$$J_3(0, p_2^2, p_3^2; 0, 0; \epsilon) \sim \frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon) \tag{17}$$

with

$$C_{-1} = -\frac{1}{(4\pi)^2 p_3^2} \frac{\ln(1-z)}{z},$$

$$C_0 = -\frac{1}{(4\pi)^2 p_3^2} \left( \ln(-p_3^2) \frac{\ln(1-z)}{z} + \frac{\ln^2(1-z)}{2z} \right).$$

Next we demonstrate the two approaches outlined above for the numerical evaluation of  $J_3(0, p_2^2, p_3^2; 0, 0; \epsilon)$  by a double extrapolation. The first is a direct computation based on Eq. (14), and the second uses Eq. (15) with a hypergeometric function computation as given by Eq. (16).

In the former approach we replace  $i0$  by  $i\delta$  in the denominator of the integrand in Eq. (14), for an application of the DCM scheme given in Fig. 2. We perform the  $\delta$ -extrapolation by the  $\epsilon$ -algorithm (Shanks, 1955; Wynn, 1956), with the geometric sequence of  $\delta = \delta_l = 2^{-8-l}$ ,  $l = 0, 1, \dots$ , which yields a value for the integral  $I(0, \epsilon_k)$  for fixed  $\epsilon_k$ . This delivers a sequence of  $I(0, \epsilon_k)$ ,  $k = 0, 1, \dots$ .

$\nu$	$C_{-1}$	$C_0$
4	-4.2192812666986e-05	-3.712127933292e-04
5	-4.1395889404493e-05	-3.975112609915e-04
6	-4.1448485840595e-05	-3.949340356225e-04
7	-4.1446205836943e-05	-3.951004758891e-04
8	-4.1446278988175e-05	-3.950927950097e-04
9	-4.1446277437321e-05	-3.950930322905e-04
10	-4.1446277461827e-05	-3.950930269725e-04
11	-4.1446277461600e-05	-3.950930270436e-04
Exact	-4.1446277461604e-05	-3.950930270463e-04

Table 2. Extrapolation Eq. (14),  $C_{-1}$  and  $C_0$  of Eq. (17),  $p_2^2 = 40 \text{ GeV}^2$  and  $p_3^2 = -100 \text{ GeV}^2$

In order to handle the IR divergence, a linear extrapolation is applied to the  $I(0, \epsilon_k)$ ,  $k = 0, 1, \dots$ , using the sequence  $\epsilon_k = \frac{1}{b_k}$ , where  $\{b_k\}$  is the Bulirsch type sequence  $\{3, 4, 6, 8, 12, 16, \dots\}$ . Numerical results of the double extrapolation for the real parts of  $C_{-1}$  and  $C_0$  are shown in Table 2 for  $p_2^2 = 40 \text{ GeV}^2$  and  $p_3^2 = -100 \text{ GeV}^2$ . For this computation, the integrations were performed to a requested relative accuracy of  $10^{-15}$ . As the compiler, the Intel Fortran XE Composer was used with flag "-r16" for quadruple precision. The results on row  $\nu$  in Table 2 correspond to solutions of the  $\nu$ -th linear system, of order  $(\nu + 1) \times (\nu + 1)$ , which can be written in the form of Eq. (6) with

$$S(\epsilon_k) = \epsilon_k J_3(0, p_2^2, p_3^2; n_x, n_y; \epsilon_k), \quad \varphi_j(\epsilon_k) = \epsilon_k^j, \text{ for } k = 0, \dots, \nu, \quad j \geq 1.$$

The results are listed from  $\nu = 4$  on. For example,  $\nu = 4$  in the first row gives solutions obtained with a  $5 \times 5$  linear system, using the values of  $b_k = 3, 4, 6, 8, 12$  for  $k = 0, 1, 2, 3, 4$ . In the second approach we replace  $z$  by  $z + i\delta_l$  in the hypergeometric function argument  $z$  of Eq. (16),

$${}_2F_1(1, 1 - \epsilon_k, 2 + n_x; z + i\delta_l), \quad k, l = 1, 2, \dots$$

Here  $\delta$  regulates the non-integrable singularity on the real axis when it occurs inside the integration region, and  $\epsilon$  regulates the IR divergence.

We let  $\delta_l = \frac{1}{b_l}$ , given by the Bulirsch type sequence  $b_l = 0.5, 0.75, 1.5, 2, 3, 4, \dots$ . Numerical results by the second approach are shown in Table 3. The integrals were computed in quadruple precision, to a relative error tolerance of  $10^{-26}$ . In both approaches, we used DQAGSE for an iterated multivariate integration.

$\nu$	$C_{-1}$	$C_0$
4	-4.2192812666950192419664334e-05	-3.7121279333017303070237e-04
5	-4.1395889404540049511801909e-05	-3.9751126098970774666184e-04
6	-4.1448485840565542759037853e-05	-3.9493403562445857754728e-04
7	-4.1446205836959577424975383e-05	-3.9510047588769404693383e-04
8	-4.1446278988182334267440660e-05	-3.9509279500930457847498e-04
9	-4.1446277437326639980970560e-05	-3.9509303229022580430491e-04
10	-4.1446277461859729880198253e-05	-3.9509302696654529617250e-04
11	-4.1446277461602157717453434e-05	-3.9509302704716538311163e-04
12	-4.1446277461604182432081146e-05	-3.950930270462724839680e-04
13	-4.1446277461604171849957952e-05	-3.9509302704627918244479e-04
14	-4.1446277461604171891430927e-05	-3.9509302704627914557531e-04
15	-4.1446277461604171891322514e-05	-3.9509302704627914571332e-04
16	-4.1446277461604171891322881e-05	-3.9509302704627914571267e-04
17	-4.1446277461604171891322692e-05	-3.9509302704627914571315e-04
18	-4.1446277461604171891323096e-05	-3.9509302704627914571171e-04
19	-4.1446277461604171891322859e-05	-3.9509302704627914571292e-04
Exact	-4.1446277461604171891322832e-05	-3.9509302704627914571278e-04

Table 3. Extrapolation Eq. (15),  $C_{-1}$  and  $C_0$  of Eq. (17),  $p_2^2 = 40 \text{ GeV}^2$  and  $p_3^2 = -100 \text{ GeV}^2$

### 4.2.3 Example 3

Analogous with the vertex case, the IR divergent integral  $J_4(p_1^2, p_2^2, p_3^2, p_4^2; n_x, n_y, n_z)$  occurs in the tensor integral of a massless one-loop box of rank  $M \leq 4$ ,

$$T_{\mu\dots\nu} = \sum_k C_{\mu\dots\nu}^k J_4(p_1^2, p_2^2, p_3^2, p_4^2; n_x^{(k)}, n_y^{(k)}, n_z^{(k)})$$

$\nu$	$b_\nu$	TIME (s)	$C_{-2}$	$C_{-1}$	$C_0$
4	6	4.9	4.00152393095	3.993440295	-12.622213
5	8	6.3	3.99820547574	4.046535578	-12.937467
6	12	8.8	3.99974091849	4.009684952	-12.595063
7	16	7.9	3.99997923446	4.001105577	-12.473283
8	24	9.5	3.99999897678	4.000078977	-12.451823
9	32	9.7	3.99999996350	4.000003986	-12.449519
10	48	15.9	3.99999999912	4.000000139	-12.449350
11	64	22.6	3.99999999980	4.000000032	-12.449343
12	96	22.1	4.00000000043	3.999999893	-12.449330
<i>Exact:</i>			4.0	4.0	-12.449341

Table 4. Linear extrapolation for  $I_4(-1, -1, \epsilon)$  of Eq. (19), coefficients of Eq. (20)

and is expressed using dimensional regularization as

$$J_4^k(s, t, p_1^2, p_2^2, p_3^2, p_4^2; n_x^{(k)}, n_y^{(k)}, n_z^{(k)}) = \frac{\Gamma(2 - \epsilon)}{(4\pi)^2 (4\pi\mu_R^2)^\epsilon} \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \frac{x^{n_x^{(k)}} y^{n_y^{(k)}} z^{n_z^{(k)}}}{\mathcal{D}^{2-\epsilon}} \tag{18}$$

where

$$\mathcal{D} = -s x z - t y (1 - x - y - z) - p_1^2 x y - p_2^2 y z - p_3^2 z (1 - x - y - z) - p_4^2 x (1 - y - z) - i0.$$

With all external particles on-shell ( $p_1^2 = p_2^2 = p_3^2 = p_4^2 = 0$ ) and  $n_x = n_y = n_z = 0$  we address the expansion of

$$I_4(s, t; \epsilon) = \frac{(4\pi)^2 (4\pi\mu_R^2)^\epsilon}{\Gamma(2 - \epsilon)} J_4(s, t; 0, 0, 0, 0; 0, 0, 0; \epsilon) \\ = \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \frac{1}{(-s x z - t y (1 - x - y - z))^{2-\epsilon}}, \tag{19}$$

$$\sim \frac{C_{-2}}{\epsilon^2} + \frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon) \tag{20}$$

Apart from a scaling factor, this simplified version of Eq. (18) corresponds to the integral of Example 1, where we used an extrapolation with the  $\epsilon$ -algorithm. Table 4 lists the results of a linear extrapolation, using an iterated (triple) integration with DQAGSE to the target relative accuracy of  $10^{-12}$  (for the outer integration) in double precision. Approximate timings are listed for the computation corresponding to the value of  $\nu$  in the first column, and are spent mainly in the integration of  $I_4$ , as the time for the extrapolation is negligible. The integrand is collected after the sector decomposition of Eq. (7) and integrated over the 3D unit cube.

The sequence  $\epsilon_k = \frac{1}{b_k}$  corresponds to the Bulirsch type sequence 1, 2, 3, 4, 6, 8, 12, ..., and the steps are shown starting at  $\nu = 4$ . The convergence of the coefficients stagnates at  $b = 96$ , i.e.,  $\epsilon = \frac{1}{96}$ , where the obtained accuracy is about 11 digits. This is as expected in view of the accuracy requirement of about 12 digits on the entry sequence for the extrapolation. The results in Table 4 agree with those of (Fujimoto & Ueda, 2008). We obtain the coefficient  $C_{-1}$

of  $\frac{1}{\epsilon}$  in Eq. (20) with the opposite sign since  $\epsilon < 0$  in (Fujimoto & Ueda, 2008), whereas we have  $\epsilon > 0$ .

## 5. Concluding remarks

A large part of this paper is devoted to a fully numerical approach for regulating Feynman loop integrals, which appear in higher order corrections in perturbative quantum field theory. The method is based on combinations of multi-dimensional integration and extrapolation techniques. It is applicable without change to various loop integrals with masses, since DCM is fully numerical. After the initialization of the sequences, the method can be seen as a black-box, which does not rely on a specification of the location of the singularity by the user. At an earlier stage, DCM handled only the divergence appearing in the integration due to the kinematical parameters. In recent work we extended the procedure to address IR divergence as well as the singular behavior of the integrand inside the integration region, by a novel double extrapolation technique. This enables DCM to regularize both divergences, as shown by numerical results for several examples. More work is needed for improvements and further analysis of these strategies. This computation is a step toward a more automatic numerical handling of various types of loop integrals, thereby circumventing the need for a precise knowledge of the structure and the location of the singularities. Since the scheme of the dimensional regularization for the UV divergent diagram is similar to that for the IR case, we are further studying the same techniques for application to the UV cases.

## 6. Acknowledgements

We wish to thank Dr. Y. Kurihara, Dr. T. Ishikawa and Prof. T. Kaneko for discussions and for their valuable comments. We also thank Prof. K. Kato for his interest in DCM and his helpful suggestions. We are gratefully indebted to Prof. Y. Shimizu for his continuous encouragement. This work was supported in part by the CPIS program of Sokendai.

## 7. References

- Bauer, C. (2002). *KEK Proceedings 2002-11*, Vol. [MZ-TH/02-04], pp. 179–185. Do Hoang Son, Ph.D. thesis at the Physics Department, Johannes Gutenberg Univ., 2003.
- Binoth, T., Guillet, J.-P., Heinrich, G., Pilon, E. & Reiter, T. (2009). Golem95: a numerical program to calculate one-loop tensor integrals with up to six external legs, *Comput. Phys. Commun.* 180: 2317.
- Binoth, T. & Heinrich, G. (2004). Numerical evaluation of multi-loop integrals by sector decomposition, *Nucl. Phys. B* 680: 375. hep-ph/0305234v1.
- Brezinski, C. (1980). A general extrapolation algorithm, *Numerische Mathematik* 35: 175–187.
- Bulirsch, R. (1964). Bemerkungen zur Romberg-Integration, *Numerische Mathematik* 6: 6–16.
- de Doncker, E. (2003). On a numerical evaluation of loop integrals, *LoopFest II, Brookhaven National Laboratory*. <http://quark.phy.bnl.gov/loopfest2/doncker.pdf>.
- de Doncker, E., Fujimoto, J., Hamaguchi, N., Ishikawa, T., Kurihara, Y., Shimizu, Y. & Yuasa, F. (2010). Transformation, reduction and extrapolation techniques for Feynman loop integrals, *Springer Lecture Notes in Computer Science (LNCS)* 6017: 139–154.

- de Doncker, E., Fujimoto, J., Hamaguchi, N., Ishikawa, T., Kurihara, Y., Shimizu, Y. & Yuasa, F. (2011). Quadpack computation of Feynman loop integrals, *Journal of Computational Science (JoCS)* p. doi:10.1016/j.jocs.2011.06.003.
- de Doncker, E. & Kaugars, K. (2010). Dimensional recursion for multivariate adaptive integration, *Procedia Computer Science* 1.
- de Doncker, E., Li, S., Fujimoto, J., Shimizu, Y. & Yuasa, F. (2005). Regularization and extrapolation methods for infrared divergent loop integrals, *Springer Lecture Notes in Computer Science (LNCS)* 3514: 165–171.
- de Doncker, E., Shimizu, Y., Fujimoto, J. & Yuasa, F. (2004). Computation of loop integrals using extrapolation, *Computer Physics Communications* 159: 145–156.
- de Doncker, E., Shimizu, Y., Fujimoto, J. & Yuasa, F. (2006). Numerical computation of a non-planar two-loop vertex diagram, *LoopFest V, Stanford Linear Accelerator Center*. <http://www-conf.slac.stanford.edu/loopfestv/proc/present/DEDONCKER.pdf>.
- de Doncker, E., Shimizu, Y., Fujimoto, J., Yuasa, F., Cucos, L. & Van Voorst, J. (2004). Loop integration results using numerical extrapolation for a non-scalar integral, *Nuclear Instruments and Methods in Physics Research A* 539: 269–273. hep-ph/0405098.
- Ford, W. & Sidi, A. (1987). An algorithm for the generalization of the Richardson extrapolation process, *SIAM J. Numer. Anal.* 24: 1212–1232.
- Fujimoto, J., Hamaguchi, N., Ishikawa, T., Kaneko, T., Morita, H., Perret-Gallix, D., Tokura, A. & Shimizu, Y. (2006). Numerical precision control and GRACE, *Nucl. Instr. and Meth. A* 559: 269.
- Fujimoto, J. & Ueda, T. (2008). New implementation of the sector decomposition on FORM, *XII Adv. Comp. and Anal. Tech. in Phys. Res.* PoS (ACAT08) 120; ArXiv:0902.2656v1 [hep-ph].
- Heinrich, G., Ossola, G., Reiter, T. & Tramontano, F. (2010). Tensorial reconstruction at the integrand level, *JHEP* 1010: 105. arXiv:1008.2441 [hep-ph].
- Kinoshita, T. (1962). Mass singularities of Feynman amplitudes, *J. Math Phys.* 3: 650.
- Kurihara, Y. (2005). Dimensionally regularized one-loop tensor integrals with massless internal particles, *Eur. Phys. J. C* 45: 427. hep-ph/0504251 v3.
- Levin, D. & Sidi, A. (1981). Two classes of non-linear transformations for accelerating the convergence of infinite integrals and series, *Appl. Math. Comp.* 9: 175–215.
- Li, S., de Doncker, E. & Kaugars, K. (2004). On the iterated numerical integration method, *Springer Lecture Notes in Computer Science (LNCS)* 3514: 123–130.
- Lyness, J. N. (1976). Applications of extrapolation techniques to multidimensional quadrature of some integrand functions with a singularity, *Journal of Computational Physics* 20: 346–364.
- Muta, T. (2010). *Foundations of Quantum Chromodynamics: An Introduction to Perturbative Methods in Gauge Theories*, World Scientific. Third edition, Lecture Notes in Physics, Vol. 78.
- Nakanishi, N. (1971). *Graph Theory and Feynman Integrals*, Gordon and Breach, New York.
- Passarino, G. & Veltman, M. (1979). *Nucl. Phys. B* 539: 151.
- Piessens, R., de Doncker, E., Überhuber, C. W. & Kahaner, D. K. (1983). *QUADPACK, A Subroutine Package for Automatic Integration*, Springer Series in Comp. Math., Springer-Verlag.
- Shanks, D. (1955). Non-linear transformations of divergent and slowly convergent sequences, *J. Math. and Phys.* 34: 1–42.

- Sidi, A. (1979). Convergence properties of some nonlinear sequence transformations, *Math. Comp.* 33: 315–326.
- van Hameren, A., Papadopoulos, C. G. & Pittau, R. (2009). Automated one-loop calculations: a proof of concept, *JHEP* 0909: 106.
- van Oldenborgh, G. J. & Vermaseren, J. A. M. (1990). New algorithms for one-loop integrals, *Z. Phys. C* C46: 425.
- van Oldenborgh, G. J. & Vermaseren, J. A. M. (2000). Automatic loop calculations with FeynArts, FormCalc and LoopTools, *Nucl. Phys. Proc. Suppl.* 89: 231.
- Wynn, P. (1956). On a device for computing the  $e_m(s_n)$  transformation, *Mathematical Tables and Aids to Computing* 10: 91–96.
- Yasui, Y., Ueda, T., de Doncker, E., Fujimoto, J., Hamaguchi, N., Ishikawa, T., Shimizu, Y. & Yuasa, F. (2007). Status reports from the GRACE group, *International Colliders Workshop LCWS/ILC*. arXiv:0710.2957v1 [hep-ph].
- Yuasa, F., de Doncker, E., Fujimoto, J., Hamaguchi, N., Ishikawa, T. & Shimizu, Y. (2007). Precise numerical results of IR-vertex and box integration with extrapolation, *XI Adv. Comp. and Anal. Tech. in Phys. Res.* PoS (ACAT07) 087, arXiv:0709.0777v2 [hep-ph].
- Yuasa, F., de Doncker, E., Hamaguchi, N., Ishikawa, T., Kato, K., Kurihara, Y. & Shimizu, Y. (2011). Numerical computation of two-loop box diagrams with masses. Submitted.
- Yuasa, F., Ishikawa, T., Fujimoto, J., Hamaguchi, N., de Doncker, E. & Shimizu, Y. (2008). Numerical evaluation of Feynman integrals by a direct computation method, *XII Adv. Comp. and Anal. Tech. in Phys. Res.* PoS (ACAT08) 122; arXiv:0904.2823.
- Yuasa, F., Ishikawa, T., Kurihara, Y., Fujimoto, J., Shimizu, Y., Hamaguchi, N., de Doncker, E. & Kato, K. (2010). Numerical approach to calculation of Feynman loop integrals. arXiv:1109.4213v1 [hep-ph]; PoS (CPP2010), Submitted.