ATLANTIS PRESS

Atlantis Studies in Mathematical Physics: Theory and Applications
Series Editor: Norbert Euler

Gianfausto Dell'Antonio

Lectures on the Mathematics of Quantum Mechanics I

Atlantis Studies in Mathematical Physics: Theory and Applications

Volume 1

Series editor

Norbert Euler, Lulea, Sweden

More information about this series at http://www.springer.com/series/13612

Gianfausto Dell'Antonio

Lectures on the Mathematics of Quantum Mechanics I

Gianfausto Dell'Antonio
Department of Mathematics
Università Sapienza
Rome
Italy

and

Mathematics Area
SISSA
Trieste
Italy

ISSN 2211-8055 ISSN 2211-8063 (electronic)
Atlantis Studies in Mathematical Physics: Theory and Applications
ISBN 978-94-6239-117-8 ISBN 978-94-6239-118-5 (eBook)
DOI 10.2991/978-94-6239-118-5

Library of Congress Control Number: 2015937751

Whether our attempt stands the test can only be shown by quantitative calculations of simple systems

Max Born, On Quantum Mechanics
Z. fur Physik 26, 379–395 (1924)

A Caterina, Fiammetta, Simonetta

Contents

Presentation

These are notes of lectures that I have given through many years at the Department of Mathematics of the University of Rome, La Sapienza, and at the Mathematical Physics Sector of the SISSA in Trieste.

The presentation is whenever possible typical of lectures: introduction of the subject, analysis of the structure through simple examples, precise results in the form of Theorems. I have tried to give a presentation which, while preserving mathematical rigor, insists on the conceptual aspects and on the unity of Quantum Mechanics.

The theory which is presented is Quantum Mechanics as formulated in its essential parts by de Broglie and Schrödinger and by Born, Heisenberg and Jordan with important contributions by Dirac and Pauli.

For editorial reason the volume of Lecture notes is divided into two parts.

The first part, Lecture 1: Elements of the History of Quantum Mechanics I–Lecture 20: Self-adjoint Extensions. Relation with Quadratic Forms. Laplacian on Metric Graphs. Boundary Triples. Point Interaction, contains the essential part of the conceptual and mathematical foundations of the theory and an outline of some of the mathematical techniques that are most useful in the applications. Some parts of these lectures are about topics that are at present subject of active research.

The second volume consists of Lectures 1–17. The Lectures in this second part are devoted to specific topics, often still a subject of advanced research. They are chosen among the ones that I regard as most interesting. Since "interesting" is largely a matter of personal taste other topics may be considered as more significant or more relevant.

At the end of the introduction of both volumes there is a list of books that may be of help for further studies. At the end of each Lecture references are given for self-study.

A remark on the lengths of each Lecture: by and large, each of them is gauged on a two-hour presentation, but the time may vary in relation to the level of preparation of the students. They can also use in self-study, and in this case the amount of time devoted to each Lecture may be vastly different.

I want to express here my thanks to the students who took my courses and to numerous colleagues with whom I have discussed sections of this book for comments, suggestions, and constructive criticism that have much improved the presentation.

In particular, I want to thank Giuseppe Gaeta and Domenico Monaco for the very precious help in editing and for useful comments and Sergio Albeverio, Alessandro Michelangeli, and Andrea Posilicano for suggestions.

Volume I—Basic Elements

Some details of the contents of the Lectures in Volume I:

- *Lectures 1 and 2*: These lectures provide a historical perspective to the beginning of Quantum Mechanics, to its early developments, and to the shaping of present-day formalism.
- *Lecture 3*: An analysis of the mathematical formulation of Quantum Mechanics and of the difficulties one encounters in relating this formalism to the empirical word, mainly for what concerns the theory of measurement.
- *Lecture 4*: Entanglement and the attempts to describe the mathematics of decoherence. An analysis of Bell's inequalities and brief outline of a formalism, originated by de Broglie, in which material points are guided by a velocity field defined by the solution of Schrödinger's equation.
- *Lecture 5*: Groups of transformations of the fundamental quantities in Quantum Mechanics: states and observables. Theorems of Wigner, Kadison, and Segal on implementability with unitary or anti-unitary maps. Continuity of the maps and the basis of Quantum Dynamics.
- *Lecture 6*: Basic facts from the theory of operators in a Hilbert space. Since in Quantum Mechanics these operators represent observables a good control of this formalism is mandatory.
- *Lecture 7*: Elements of the theory of quadratic forms. Quadratic forms are an important tool in the theory of operators on Hilbert spaces and they play a major role in the theory of extensions. Friedrich's extension of a semi-bounded symmetric operator.
- *Lecture 8*: Analytic study of the solutions of the Schödinger equation, beginning with the simple but instructive case of free motion. Propagation inequalities and their relation to the description of the asymptotic properties of a quantum mechanical system. The problem of anholonomy and the geometric phase.
- *Lecture 9*: Elements of the theory of C^*-algebras and von Neumann algebras. This lecture provides some elements of the theory of automorphisms of C^*-algebras and the description of the dynamics of quantum systems.
- *Lecture 10*: Generators, derivations, and in particular the KMS condition for a group of automorphims of a C^*-algebra. Implementation of a group of automorphism by a group of unitary operators. Modular structure of a representation and standard form of a von Neumann algebra.

- *Lecture 11*: Basic elements of the theory of semigroups in Banach spaces and of the theory of dissipations. Markov approximation and conditions for its validity. Elements of a converse problem, the *dilation* of a Markov semigroup.
- *Lecture 12*: Role of positivity and complete positivity in the theory of contraction semigroups on C^*-algebras. Elements of the theory of dissipations and basic facts in the theory of Quantum Dynamical semigroups.
- *Lecture 13*: The problem of *quantization*. Weyl system and Weyl algebra, uniqueness theorem of von Neumall and Weyl. Formalism of *second quantization*. Magnetic Weyl algebra.
- *Lecture 14*: Various representations of the Weyl algebra (real and complex representations of Bargmann and Segal, representations of Fock and Berezin). The case of an infinite number of degrees of freedoms. The real representation and the quantization of the free relativistic field (Segal). van Hove's theorem. Brief outline of deformation quantization and of geometric quantization.
- *Lecture 15*: Formally, for systems for which Plank's constant can be considered very small, and for suitable initial data, the dynamics of a system in Quantum Mechanics can be approximated by the dynamics of the corresponding classical system. One refers to this fact by saying that the system can be described in the semiclassical limit. Basic facts regarding this limit. Role of coherent states and metaplectic group.
- *Lecture 16*: Initial data strongly oscillating in configuration space, stationary and non-stationary phase techniques (WKB method). Role of the Maslov index and the origin of the semi-classical quantization rules.
- *Lecture 17*: Deeper analysis of the theory of operators on a Hilbert space, in particular the self-adjoint ones, that describe time evolution. Comparison theorems, in particular the one of Kato-Rellich. Special classes of potentials: the Rollnik and Stummel classes.
- *Lecture 18*: Weyl's criterium for a self-adjoint operator. Detailed study of the Hydrogen and Helium-like atoms, including properties of the spectrum and the presence of embedded eigenvalues.
- *Lecture 19*: An analysis of techniques to estimate the number of bound states for the Schrödinger operator. The Feshbach method is also presented.
- *Lecture 20*: This more specialized lecture develops the theory of self-adjoint extensions of a symmetric operator and its relation to the theory of quadratic forms. Elements of the method of boundary triples, exemplified with the theory of point interaction. Laplacian on a metric graph.

Volume II—Selected topics

Some *tentative* details of the contents of Lectures 1–17

- *Lectures 1 and 2*: Wigner's functions, Pseudo-differential operators. Other quantization procedures (Berezin-Wick, Kohn-Nirenberg, Shubin)
- *Lecture 3*: Shatten class operators, an anthology of inequalities that are commonly used.
- *Lectures 4 and 5*: Mathematics of periodic structures, in particular crystals. Formalism of Bloch-Floquet, Bloch functions, localized Wannier functions. Topological problems connected with the lattice structure.
- *Lectures 6 and 7*: Feynmann-Kac formula. Relation to the heat semigroup, to Wiener's process, and its extension to the Orstein-Uhlenbeck process.
- *Lecture 8*: A brief and elementary treatment of Brownian motion and of Markov processes in general.
- *Lecture 9*: Analysis of the Friedrichs extension of a closed quadratic form. Connection of Dirichlet forms with the theory of self-adjoint operators.
- *Lectures 10 and 11*: Brief outline of the Tomita-Takesaki theory of the Modular Operator in von Neumann algebras and its relation to Friedrichs's extension. Non-commutative integration and non-commutative extension of the equivalence of measures (existence of a Radon-Nykodyn derivative).
- *Lectures 12 and 13*: Elements of scattering theory in Quantum Mechanics. Time-dependent and time-independent formulations. Outline of a method due to V. Enss, which provides a detailed analysis of the evolution in time of the wave function.
- *Lecture 14*: Propagation estimates and Kato's smoothness theory and application to scattering theory. Outlines of Mourre's theory. Further generalizations (e.g., the method of conjugated operators).
- *Lecture 15*: An outline of the theory of a quantum N-body system, both for its spectral properties and the properties of N-body scattering.
- *Lectures 16*: Completely positive maps (open systems). Contractive Dirichlet forms. Markovian and hyper-contracting semigroups, existence, and uniqueness of a ground state. Connection between Markov processes and strictly contractive Dirichlet forms.
- *Lecture 17*: Canonical anticommutation relations. Pauli equation and Dirac operator.

Bibliography for Volumes I and II

1. Amrein, V., Jauch, J., & Sinha, K. (1977). *Scattering theory in Quantum Mechanics*. Reading, Massachusetts: Benjamin.
2. Baez, J., Segal, I.E., & Zhou, Z. (1992). *Introduction to algebraic and constructive field theory*. Princeton University Press.

3. Bratteli, O. (1986). *Derivations, dissipations and group actions on C*-algebras*. 1229 Lecture Notes in Mathematics, Springer.
4. Bratteli, O., & Robinson, D.W. (1979/87). *Operator algebras and quantum statistical mechanics I, II*. New York: Springer.
5. Brezis, J. (2010). *Functional analysis, sobolev spaces and applications*. Springer Universitext.
6. Cycon, H.L., Froese R.H., Kirsch, W., & Simon B. (1987). *Schroedinger operators with application to quantum mechanics and global geometry*. Text and Monographs in Physics, Berlin: Springer.
7. Davies, E.B. (1976). *Quantum theory of open systems*. Academic Press.
8. Diximier, J. (1969). *Les algebres d'operateurs dan l'espace hilbertien*. Paris: Gauthier-Villars.
9. Doob, J. (1953). *Stochastic processes*. New York: Wiley.
10. Folland, G.B. (1989). *Harmonic analysis in Phase Space*. Princeton, New Jersey: Princeton University Press.
11. Gustafson, S., & Segal, I.M. (2006). *Mathematical Concepts of Quantum Mechanics*. Springer.
12. Hille, E., & Phillips, R.S. (1957). *Functional Analysis and Semigroups*. American Mathematical Society.
13. Hislop, P.D., Sigal, I.M. (1996). *Introduction to Spectral Theory, with application to Schroedinger Operators*. New York: Springer.
14. Kadison, R.V., Ringrose, J.R. (1983/86). *Fundamentals of the theory of operator algebras vol. I–IV*. Academic Press.
15. Kato, T. (1980). *Perturbation theory for linear operators*. Berlin: Springer.
16. Kuchment, P. (1993). *Floquet theory for partial differential equations*. Basel: Birkhauser.
17. Jammer, M. (1966). *The conceptual development of quantum mechanics*. New York: Mc Graw-Hill.
18. Mackey, G.W. (1963). *Mathematical foundations of quantum mechanics*. New York: Benjamin.
19. Maslov, V.P. (1994). *The complex W.K.B. method for non-linear equations I- Linear Theory*. Progress in Physics vol. 16, Birkhauser Verlag.
20. Nelson, E. (1969). *Topics in Dynamics, I: Flows*. Princeton University Press.
21. Pedersen, G.K. (1979). *C*-algebras and their Automorphism Groups*. London: Academic Press.
22. Reed, M., & Simon, B. (1977/78). *Methods of Modern Mathematical Physics vol. I–IV*. New York and London: Academic Press.
23. Sakai, S. (1971). *C*-algebras and W*-algebras*. Berlin: Springer.
24. Shmüdgen, K. (2012). *Unbounded self-adjoint operators on Hilbert space*. Graduate Texts in Mathematics vol. 265. Dordrecht: Springer.
25. Segal, I.E. (1963). *Mathematical Problems of Relativistic Physics*. Providence, Rhode Island: American Mathematical Society.
26. Simon, B. (1971). *Quantum Mechanics for Hamiltonians defined as Quadratic Forms*. Series in Physics, Princeton University press.
27. Simon, B. (1979). *Functional Integration and Quantum Physics*. Academic Press.
28. Sinai, Y.G. (1992). *Probability Theory*. Springer Textbooks, Berlin: Springer.
29. Takesaki, M. (1970). *Tomita's Theory of Modular Hilbert Algebras*. Lecture Notes in Mathematics vol. 128. Heidelberg: Springer.
30. Takesaki, M. (1979). *Theory of Operator Algebras*. Vol 1, 2, New York: Springer.
31. Teufel, S. (2003). *Adiabatic Perturbation Theory in Quantum Dynamics*. Lecture Notes in Mathematics vol. 1821. Springer.
32. von Neumann, J. (1932). *Mathematische Grundlage der Quantenmechanik (Mathematical Foundation of Quantum Mechanics)*. Berlin: Springer. (Princeton University Press 1935)
33. van der Werden, B.L. (Ed.) (1968). *Sources of Quantum Mechanics*. New York: Dover.
34. Yoshida, K. (1971). *Functional Analysis*. Berlin: Springer.

Lecture 1: Elements of the History of Quantum Mechanics I

In this Lecture we review some of the experiments and theoretical ideas which led to establish Quantum Mechanics in its present form.

1 Introduction

At the end of the 19th century Classical Mechanics (Hamiltonian and Lagrangian) and Electromagnetism had reached a high degree of formalization and a high standard of mathematical refinement. It seemed that this theory could account for all phenomena that are related to the motion of bodies and to the interaction of matter with the electromagnetic field.

But new experimental tools available at the atomic level led to a large number of experimental results which don't fit well within this *classical* theory.

• In the years 1888–1909 Rutherford proved that the newly discovered α radiation was composed of doubly ionized helium atoms. He was also able to determine the numerical value (in suitable units) of the ratio between the charge and the mass of the electron. Later experiments (1909) performed by Millikan led to the recognition of the existence of an *elementary quantum of charge*.

A detailed analysis of the structure of atoms was performed by scattering of α particles. Marsden and Geiger recorded a large number of events in which the scattering angle was of the order of magnitude of a radiant. The thickness of sample of material was of order of 6×10^{-5} cm, small enough to expect that multiple scattering be negligible. Rutherford in 1909 suggested that these "large" deviations were due to *a single scattering event*. If this is the case, classical scattering theory gives an order of magnitude of 4×10^{-12} cm for the radius of the scatterer.

On the other hand, the radius of the atom was estimated to be of order of 10^{-8} cm. This estimate was obtained using Avogadro's number under the assumption that the atomic radius is comparable to the inter-atomic distance (as suggested by other experiments). One was therefore led to conclude that the atom was composed of a

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_1

very small *heavy* positively charged nucleus and a number of particles (the *electrons*) which have unit negative charge and a mass much smaller than that of the nucleus.

Rutherford proposed a model in which the electrons "move" within a distance of order of 10^{-8} cm from the nucleus.

• Experiments carried out mainly by Geiger and Marsden confirmed the validity of the model proposed by Rutherford and suggested that the charge of the nucleus is roughly half of its atomic number.

Before the proposal of Rutherford the most accepted model of an atom was that of Thomson, according to which the nucleus is a positive charge density that extends uniformly in a ball of radius $\simeq 10^{-8}$ cm and in which small negative charges (the electrons) move.

On the other hand, assuming Rutherford's model, classical electromagnetism implies that the atom is very unstable: the electrons are attracted by the nucleus and are accelerated, emitting radiation and losing energy. They eventually fall to the nucleus. On the contrary experimental evidence points to the stability of the atom; the mechanism providing stability was unknown.

It is appropriate to point out that the previous argument is weakened by the fact that there was no attempt to solve the equations for the evolution of atom coupled to the electromagnetic field.

• Other experiments revealed aspects of the atomic world that were difficult to reconcile with the classical laws. One of these is the *photoelectric effect*.

When light is flashed on a metallic plate the electrons are emitted "locally" and the current (number of electrons emitted) is proportional to the intensity of the light if its frequency exceeds a threshold ν_0 but is zero if $\nu < \nu_0$ *independently of the intensity*. The constant ν_0 depends on the atom in the metallic plate. The interpretation of this phenomenon proposed by A. Einstein is that in this interaction light *behaves as if* it were composed of a large collection of particles, the *photons* (the name was suggested by Lewis [12]). Light of frequency ν has photons of energy $h\nu$, where h is a universal constant (Planck's constant). According to this interpretation the photoelectric effect is the following. The photon hits the atom which is in a bound state. If E_0 is the *binding energy*, a photon of energy less than E_0 cannot ionize because of energy conservation.

• An assumption of the quantum nature of the energy exchanged in the interaction between matter and the electromagnetic field had already been made by Planck to justify the empirical formula he found for the frequency distribution of the black-body radiation.

A black body can be represented as a cavity with perfectly reflecting walls. Through a very small hole in the walls electromagnetic radiation is inserted. The hole is then closed, and one waits long enough for the radiation to be in equilibrium with the walls of the cavity. The radiation which is emitted when the small hole is opened again is called *black body radiation*.

By general thermodynamic arguments the spectrum $\rho(\nu)$ of the black-body radiation should have the form (Wien's law)

$$\rho(\nu) = \nu^3 \Phi\left(\frac{\nu}{T}\right) \tag{1}$$

where Φ can be a very general function and $\rho(\nu)$ is the probability density that the radiation frequency be ν. Wien's law (1) is derived by studying the variation of the state of the electromagnetic radiation in the cavity when subjected to a Carnot cycle.

If one applies to the photons of the electromagnetic field the law of energy equi-partition familiar for particles from Classical Statistical Mechanics one arrives to the Rayleigh-Wien law $\rho(\nu) = \frac{8\pi}{c^3}\nu^2 KT$ where T is the equilibrium temperature and K a universal constant. Remark that the Rayleigh-Wien law satisfies (1) with $\Phi(z) = \frac{8\pi}{c^3}\frac{1}{z}$.

The law found empirically by Planck is instead

$$\rho(\nu) = \frac{8\pi h}{c^3}\frac{\nu^3}{e^{\frac{h\nu}{KT}} - 1} \tag{2}$$

which satisfies (1) with $\Phi(z) = \frac{8\pi}{c^3}(e^{-\frac{h}{K}z} - 1)^{-1}$ where h is a universal constant and T the equilibrium temperature of the black-box. When ν is small the two distribution laws tend to coincide but they differ much when ν is large.

Let us remark that Wien's law gives equi-partition of energy among the frequency modes of the electromagnetic field in the cavity; it leads to a infinite total energy since the modes are infinite in number.

It should be pointed out that later studies of a system of harmonic oscillators with non-linear coupling showed that the equi-partition law needs not be satisfied in general and even in the cases in which the law is satisfied, if the number of oscillators is very large the time required to reach an approximate equilibrium can be very long (several thousand years in realistic cases according to an estimate by Jeans (1903)).

• If the electromagnetic radiation is made of photons and for the purpose of energy distribution the photons are treated as particles, Planck's law can be derived with a statistical analysis similar to the one made by Gibbs in his formulation of classical statistical mechanics. It gives the number of photons with energy $E_\nu = h\nu$ (and therefore of frequency ν) that are present in the radiation when it is in thermal equilibrium at temperature T. Indeed assuming the relation $E_\nu = h\nu$, Planck's law can be derived from Gibbs' law and the universal constant K is identified with Boltzmann constant. Notice that in a canonical state of equilibrium with energy E there can be only few high-energy photons.

We give here the derivation of Planck's law given by A. Einstein in 1905 under the assumption that photons behave as particles in their interaction with atoms.

The states of a atom are classified by integer numbers and have energies E_m, $m = 1, 2 \ldots$. Consider an atom which is in equilibrium with radiation at temperature T. Let $\Delta w = A_m^n \Delta t$ be the probability of *spontaneous transition* in a time Δt from a state with energy E_m to a state with energy E_n with emission of a photon.

Let $\frac{dw}{dt} = B_n^m \rho$ the probability density for the transition from a state of energy E_m to a state of energy E_n when the atom is exposed to a radiation with density $\rho(\nu)$.

Let $\frac{dw}{dt} = B_m^n \rho$ be the probability density of transition from the state E_n to the state E_m with absorption of radiation with density $\rho(\nu)$.

According to Gibb's laws of statistical mechanics the system atom + radiation will be in equilibrium if *for every frequency* the probability of emission and of absorption are equal. Therefore, denoting by p_n the probability that the atom be in state n one must have

$$p_n e^{-\frac{E_n}{KT}} B_n^m \rho = p_m e^{-\frac{E_m}{KT}} (B_m^n \rho + A_m^n) \tag{3}$$

(the factors $\frac{E_n}{KT}$ are derived form Gibbs' law). This relation must hold at every temperature and every density.

Taking first the limit $\rho \to \infty$ and then the limit $T \to \infty$ one derives

$$B_n^m p_n = B_m^n p_m \tag{4}$$

Substituting in (3) one obtains

$$\rho = \frac{A_m^n}{B_m^n} \frac{1}{e^{\frac{E_n - E_m}{KT}} - 1} \tag{5}$$

Comparing with Planck's law one derives

$$E_m - E_n = h\nu \tag{6}$$

We conclude that if photons follows the laws of classical statistical mechanics, their energy must be given by $E_\nu = h\nu$.

Later, analyzing the thermodynamic relations, including the pressure of the electromagnetic field in a cavity, A. Einstein gave evidence of the fact that also a *momentum* p can be ascribed to a photon and that the relation between energy and momentum is $E^2 = c^2|p|^2$ where c is the speed of light in vacuum. This is the relation which holds in a relativistic theory for a zero-mass particle.

It is interesting to notice that the name *quantum mechanics* does not originate from this property of light to have *quanta of energy*. The name originates from the property of the atomic levels to have a *quantized* structure. The property of light to be composed of photons is at the origin of the *second quantization* for the theory developed independently by W. Heisenberg and W. Pauli and by P.A. Dirac (after a first attempt by P. Jordan) to include the electromagnetic field in Quantum Mechanics. This theory was later generalized to treat *matter fields* and is at the basis of the relativistic theory of quantized fields.

• Further evidence of the difficulties of the classical theory of interaction of matter with the electromagnetic field came from the Compton effect and the emission and absorption of light. The scattering (deflection) of an electron by the electromagnetic radiation was *compatible* with the conservation of energy and momentum only if one assumed that light were composed of *photons* of mass zero. On the other hand one did not find any way to account for this effect assuming the classical laws.

• A large collection of experimental data was available on the frequencies of radiation emitted or absorbed by atoms. For the hydrogen atom Balmer gave the empirical formula

$$k_m = 2\pi R\left(\frac{1}{2^2} - \frac{1}{m^2}\right), \quad m = 1, 2, \ldots \tag{7}$$

where k_m is the wave number ($\equiv \frac{2\pi}{\lambda}$) and R is a *universal constant* (Rydberg constant).

For more complex atoms the empirical formula

$$k_m = K^0 - \frac{2\pi R}{(m+p)^2}, \quad p = 1, 2, \ldots \tag{8}$$

was given, where K^0 is a constant.

Under the photonic hypothesis of the nature of light, (8) and (7) give a rule for the determination of the difference of the energies of the atomic states before and after the emission of a photon of frequency ν.

This experimental evidence, as well as the result of other experiments on the frequency of light absorbed or emitted by atoms in presence of electromagnetic fields, were known since 1905–1909 but had not found a satisfactory explanation.

An attempt had been made by Haas (1910), within Thomson's model, by identifying the potential energy of an electron with its rotation frequency, i.e. making the assumption that also for the electrons within an atom the relation $E = h\nu$ holds. One finds in this way a value of h which is not very different form that obtained from the black-body radiation or the photoelectric effect.

A further step to find a relation between energy of the electron inside the atom and frequency of the emitted or absorbed radiation was taken by Nicholson in 1911 within a *band model* (in which the electrons form bands which rotate around the nucleus).

2 Birth of Quantum Mechanics. The Early Years

The first *organic collection of rules* to determine the energy of atomic states was formulated by N. Bohr in 1913 [2].

Consider first the hydrogen atom. In Bohr's description the stationary states (equilibrium states) are described by periodic orbits of the electrons considered as Newtonian system. If the atom reaches the equilibrium state by emitting homogeneous and mono-chromatic radiation of frequency ν the classical virial theorem (the mean value along the orbit of kinetic and potential energy coincide) gives for the potential energy $W$of the electron

$$W = N\frac{h\nu}{2}, \quad N \in Z \tag{9}$$

From newtonian mechanics one has the relation

$$\nu = \frac{\sqrt{2}}{\pi} \frac{a^{3/2}}{eQ\sqrt{m}} \tag{10}$$

where ν is the frequency of the circular motion, a is the radius of the orbit and Q is the charge of the nucleus.

Comparing (9) and (10) with experimental data provides the possible radius of the orbits, which are *quantized*. The classical Action takes on these orbits values which *are multiples of* $\hbar$.

It is difficult to generalize this analysis for heavier atoms. Even for Helium atoms (three body-problem) one is not able to find all periodic motions and the corresponding value of the action. Therefore this simple rule of quantization must be abandoned; in "Lecture 15: Semiclassical Limit; Coherent States; Metaplectic Group" we will see how these rules are recovered (approximately) in Quantum Mechanics.

The three principles put forward by N. Bohr were

1. *Correspondence to classical states*. The equilibrium states of an atom can be described within classical mechanics. At least for the hydrogen atom they consist of circular orbits of the electron around the nucleus and the permitted radiuses satisfy (10).
2. *Transition between states*. The transition between states *cannot be described by classical mechanics*. The transition is accompanied by emission or absorption of mono-chromatic radiation (photons); the frequency of the radiation is such that energy is conserved. Therefore one has

$$E_{\tau_2} - E_{\tau_1} = h\nu_{\tau_1 \to \tau_2} \qquad \Rightarrow \qquad \nu = \frac{2\pi^2 m e^4}{h^3} \frac{1}{\tau_2^2 - \tau_1^2}$$

   where τ_k are the quantum numbers that characterize each atom.
   For circular orbits, this rule implies that the equilibrium states of an atom are those in which the orbit of the electrons have an action which is multiple of $\hbar \equiv \frac{h}{2\pi}$ (Bohr-Sommerfeld quantization conditions). For other atoms the rule that associates classical periodic motion to atomic states is more complex.
3. *Correspondence principle*. When the quantum number N of a state is very large, the frequency of the radiation which accompanies the transition from a state τ_N to a state τ_{N-1} is (approximately) equal to the orbital frequency of motion of the electron derived by classical electromagnetism. ♣

Remark that of these principles only the third one (*correspondence principle*) is kept in the present day presentation of Quantum Mechanics. Its role is to determine the value of the parameters which enter the theory.

It is important to notice that Bohr's principles assumes the validity of classical kinematics and part of the dynamics, but substitutes another part of the dynamics (absorption and emission of electromagnetic waves) with an ad hoc mechanism that

gives the transition between states. On the contrary, as we shall see, the present day formulation Quantum Mechanics of Born, Heisenberg and Schrödinger introduces a *different kinematics* while dynamics has the same mathematical structure as hamiltonian mechanics.

Further support for Rutherford's model and the principles of Bohr came from the determination of the number of electrons in an atom and of the charge of the nucleus. Franck and Hertz performed ionization measurements on a gas of atoms and their results were in moderately good agreement with the model of Rutherford and Bohr's principles; the agreement was less satisfactory for large atomic numbers.

But also the theoretical predictions were not so accurate for large atomic numbers because they were based on the assumption that only those periodic orbits are to be considered for which the action has values which are multiples of $\hbar$. In the case of large atomic number it was difficult to classify all periodic orbits.

Another set of data had a very important role in the formulation of the theory. Experiments suggested a relation between the energy of atomic states and the frequency and intensity of the radiation emitted and absorbed (such relations are called *dispersion relations*).

This empirical relations can be expressed in terms of matrices (because they refer to pairs of atomic states) and can be compared with similar relations given for a Rutherford atom by classical electromagnetism and classical dynamics. To understand better how this comparison can be made consider the four main assumption that were made:

Assumption A *the adiabatic hypothesis* [9]. Under adiabatic reversible processes permitted motions go over to permitted motions. ◇

This assumption underlines the importance of *adiabatic invariants*. For example, for a periodic motion is adiabatically invariant the ratio $\frac{2\bar{T}}{\nu}$ where ν is the frequency of motion and $\bar{T}$ is the mean value of the energy over a period. In the case of the harmonic motion in one dimension one has $\frac{2\bar{T}}{\nu} = E$.

Ehrenfest compared the adiabatic hypothesis with the hypothesis introduced by Planck, Bohr and Sommerfeld that a harmonic oscillator (that was considered intermediary for the interaction of matter with the electromagnetic field) can only have an energy which is multiple of its frequency.

It follows from the adiabatic hypothesis that the following relation should hold true

$$\frac{2\bar{T}}{\nu} = \iint dq \wedge dp = nh, \quad n \in Z^+ \tag{11}$$

Analogous formulae hold for a rigid rotator and for a magnetic dipole. In general the ratio energy/frequency must be an adiabatic invariant. Therefore the possible periodic motions in an atom must satisfy

$$E = \nu \iint \sum_k dq_k \wedge dp_k \tag{12}$$

and the coefficient ν must take integer values (because this holds for harmonic oscillators).

It was therefore considered adiabatic the (symplectic) transformation from position and momentum variable to action-angle variables that allows to describe the hydrogen atom as a collection of harmonic oscillators.

Assumption B *Bohr's postulates* [2]. There are denumerably many stationary states in an atom, each with (internal) energy E_n, $n \in N$. The frequency of the radiation emitted or absorbed by an atom when passing from a stationary state to the other satisfies $\nu = \frac{|E_n - E_m|}{\hbar}$. ◇

For each atom A there must be a positive function f_τ^A such that

$$\nu_{n,m} = |f_\tau^A(n) - f_\tau^A(m)|, \qquad E_n^A = -f_\tau^A(n) \tag{13}$$

where τ varies over a denumerable set. Therefore the state of an atom *can be described by two parameters*; one (n) is integer-valued and the other (τ) can take values in a denumerable set.

For the hydrogen atom the energy is function only of n. This agrees with hamiltonian mechanics since for this system the energy is a function of a single action variable. The presence of the parameter τ reflects the complexity of the other atomic systems and the complexity of the quantization rules.

Assumption C *The correspondence principle*. The new mechanics should give results which are (almost) the same as those given by classical mechanics when the number $\frac{n-m}{n}$ is small (in particular when n is very large and $n - m$ is small). ◇

Assumption D *Principle of mechanical transformability*. The laws of classical mechanics hold for observables associated to adiabatic invariants. ◇

Notice that Assumption (D) coincides essentially with Ehrenfest's adiabatic hypothesis and that Assumption (C) is needed to find the function f_τ.

In the years 1917–1923 an attempt to construct a new theory [3] was made assuming that atoms, *in their interaction with the electromagnetic field*, behave as if they were *virtual* harmonic oscillators with frequencies in one-to-one correspondence with the absorption and emission frequencies [8, 11]. Slater [19] speaks of *virtual field of radiation* and considers the interaction of the atom with the electromagnetic field as *mediated by virtual harmonic oscillators*. This line of research was concentrated on finding more refined empirical formulae in order to find the function f_τ^A. Another line of research was the study of the *dispersion relations*, i.e. the relation between the frequency of the radiation which was absorbed (or emitted) and its intensity. This line of research proved to be very important for the future development of the theory, since in this way one finds relations which had a counterpart in the

classical theory of the interaction matter-electromagnetic field, and led to discover important analogies and differences.

The simplest *dispersion relation* in the classical theory is the following. Consider a charged harmonic oscillator which interacts with a time-dependent mono-chromatic electromagnetic field. We approximate the interaction with a constant conservative force proportional to the intensity of the field and a dissipative term, linear in the velocity of the oscillator, which describes the loss of energy due to emission of radiation.

The classical equation of motion for the damped oscillator (that should be compared with the dipole moment of the atom) is

$$\frac{d^2P}{dt^2} + \gamma\frac{dP}{dt} + (2\pi\nu_0)^2 P = \frac{e^2}{m}E(t), \qquad E(t) = E_0 e^{2\pi i\nu t} \tag{14}$$

where ν_0 is the frequency of the oscillator, ν is the frequency of the electromagnetic field and γ is the dissipation constant.

A solution of (14) is

$$P(t) = \alpha E(t), \quad \alpha = \frac{e^2}{m}\left[4\pi^2\nu_0^2 - 4\pi^2\nu^2 + 2i\pi\nu\gamma\right]^{-1} \tag{15}$$

under the assumption $\nu \neq \nu_0$ (non resonance condition).

Recall now that the atom in its interaction with the electromagnetic field was considered *equivalent* to a collection of harmonic oscillators; therefore (15) should be compared with the law that describes the variation in time of the dipole momentum of an atom in an external time-dependent homogeneous mono-chromatic electromagnetic field.

Experiments conducted to determine this variation in time led to the following *empirical formula* [10]

$$P = E\frac{e^2}{4\pi^2 m}\sum_k f_k\left[\nu^2 - \nu_k^2\right]^{-1} - E\frac{e^2}{4\pi^2 m}\sum_j g_j\left[\nu^2 - \nu_j^2\right]^{-1} \tag{16}$$

where ν_k are the absorption frequencies, ν_j the frequency of emission (induced or spontaneous) and the coefficients f_k and g_j are the probabilities of absorption and emission.

Comparison of (16) with (15) suggests that the dissipation coefficient γ be negligible: its presence is an indication that the proper frequencies of the system can be only approximately defined (the atomic levels correspond only approximately to constant energy). Equation (16) must be supplemented by the suggestion by Einstein that the frequencies are given as

$$\nu_{k,n} = \frac{1}{h}[E(k) - E(n)] \tag{17}$$

where $E(k)$ is the energy of the k^{th} atomic state.

One can therefore imagine to describe the atoms as *a virtual collection* of harmonic oscillators and the transition between energy levels as due to the *interaction between the virtual oscillators and the electromagnetic field*.

3 Birth of Quantum Mechanics 1. The Work of de Broglie

The work of de Broglie, started roughly in 1923, is an attempt to reconcile the two "quantum" phenomena that were known at that time, namely the apparently *dual nature of radiation* (corpuscolar and wave-like) and the *(Bohr-Sommerfeld) quantization* of atomic levels. The theory is elaborated in several papers [5–7]. It was later taken up by E. Schrödinger (and modified in an essential part).

We have seen that photons describe light, and therefore they have the structure of waves, but at the same time behave as particles in the photoelectric effect. de Broglie introduced the hypothesis that *particles (e.g. electrons) must have a similar dual nature*. The diffraction of X-rays had been observed by von Laue in 1914 [20] and de Broglie *predicted* that a similar phenomenon occurs for electrons.

If, in analogy with the photon, an electron has also a wave-like structure and if, in analogy with the photon, the relation between its momentum and the frequency of the wave is $|p| = h\nu$ (in units in which the velocity of light is set to 1), *diffraction of electrons* should be seen if a beam of electrons goes past a crystal of appropriate lattice spacing.

In 1928, experiments carried out by Thompson and by Davisson and Germer confirmed the intuition of de Broglie. Only much later (in the 40s) it was possible to measure neutron diffraction and the relation was again confirmed. The difficulty in making diffraction experiments with neutrons is that, due to their large mass, the neutrons must move at very low speed (thermal neutrons) in order that the wave associated to them have a wave length comparable to the lattice spacing of the diffracting material. The development of the technology of thermal neutrons due to E. Fermi in the late thirties allowed to measure the diffraction of neutrons and to confirm also for neutrons de Broglie's hypothesis.

It is remarkable that de Broglie actually *predicted* the phenomenon of interference patterns in the diffusion of electrons on a crystal.

To give strength to his hypothesis de Broglie noted that the determination of stable orbits the electrons led to integer numbers, as is the case for the eigen-vibration of a string. He then proposed that to a particle (electron) be associated *also a wave* and that the quantization conditions for the bound states be regarded as condition for the associated wave to remain in phase with the position of the electron in its motion within the atom.

For an electron not bound inside an atom de Broglie proposed instead that its path follows the normal to the surface of equal phase of the associated ray, as is the case in the description of motion in Classical Mechanics as flow of lagrangian manifolds.

This is a radical change of view on the dynamics, abandoning the traditional (newtonian) view of the motion as due to a force, and substituting it with a dynamics

in which the motion of the particle is determined by a vector field which depends on the wave associated to the particle. We will refer to the de Broglie scheme as *pilot wave theory*.

As de Broglie pointed out, the new dynamics stands to classical dynamics in the same relation as wave optics to geometrical optics. In the high-frequency limit the predictions of the two theories tend to coincide. In the new theory *force is a derived entity*, identified (modulo a multiplicative constant) with the time derivative of the velocity (for sufficiently regular motions).

The interference fringes in the scattering of electrons on a crystal are seen as a consequence of the interaction of the atoms of the crystal *with the pilot wave of the electrons*. This interaction is responsible for the fact that trajectories of the electrons *wiggle* and arrive at the screen in such a way to *reproduce* the interference patterns (two trajectories never intersect so that the equations of motion are well posed).

For de Broglie the *fundamental idea of quantum theory is the impossibility of considering an isolated quantity of energy E (de Broglie speaks rarely of "particles")* without associating to it a wave. He stated: I was led to assume that for a given total energy of the moving body (and therefore for a given frequency of the associated wave) the possible trajectories of the body coincide with the possible rays of the associated wave (Chap. 2 of de Broglie thesis).

In order to associate a wave of frequency ν_0 to an *energy fragment* of proper mass m_0 (in its reference frame) de Broglie generalized the quantum relation $h\nu = E$ (which is supposed to hold for photons) to obtain $h\nu_0 = m_0c^2$.

A guiding line in de Broglie's research is an attempt to an unification of classical mechanics and the new theory of quanta through the fundamental variational principles: Maupertuis Principle in Mechanics and Fermat Principle in optics. It is in fact easy to see that for fixed energy Hamilton Principle reduces to Maupertuis'. And if one considers a sinusoidal wave Fermat Principle can be written as $\delta \int_P^Q d\phi = 0$ where $\phi(t, x)$ is the phase of the wave. de Broglie proposed that the wave associated to a particle satisfies $d\phi = 2\pi v_\mu dx^\mu$ (for a particle of energy E and momentum $p \in R^3$) one has $p_i = hv^i$. This leads to a complete analogy between the two variational principles.

It can be seen that the speed of the crest of the (monochromatic) wave is greater than the speed of light. The same is true in Hamilton's formulation of mechanics as flow of lagrangian manifolds in configuration space. The velocity of the representative point is perpendicular to the manifold of constant value for the lagrangian action, inversely proportional to the local velocity of the lagrangian manifold and equal to the *group velocity* properly defined. In the same way de Broglie compared the velocity of the particle with the *group velocity of the wave*.

De Broglie remarked also that it is possible to *artificially re-introduce* in the theory the newtonian point of view of forces by computing the acceleration of the "fragment of energy" and multiplying by its mass. This lead to the introduction of an artificial *Quantum Potential*; this has led later to view Quantum Mechanics as Classical Mechanics with the addition of the Quantum Potential.

In de Broglie's interpretation the wiggling of trajectories in a two-slit experiment must be ascribed to the interaction of the guiding wave with the atoms of the lattice (or with the borders of the slits).

Later de Broglie extended these ideas to system composed of N particles described by points in configuration space R^{3N} which was associated to a pilot wave ψ with phase $S(x_1, \dots, x_n; t)$. The motion of the points is determined by the equation

$$m_k \frac{dx_k}{dt} = \nabla_k S \qquad S = \hbar\, Im(\log|\psi|^2) \tag{18}$$

To complete the laws of dynamics one must consider also the case in which the electron is bound to a nucleus and prove that in this the stationary states of the electron have an energy that can assume only a discrete set of values. This can be done if one has an equation for the wave ψ and one sets suitable boundary conditions (as is the case for a string with extremal points fixed).

Hints to the form of this equation were given by de Broglie who had shown that in presence of an electrostatic potential ϕ the phase of an electron of charge e and velocity v would have frequency $\nu = \frac{mc^2+V}{h}$ and phase velocity $v_{phase} = \frac{mc^2+V}{mv}$ where $V = e\phi$ and $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$. It was Schrödinger who attempted to combine the formulation of de Broglie with the Klein-Gordon equation to describe the dynamics and the structure of the energy levels in the hydrogen atom [13, 15]. For further comments and analysis of de Broglie's work one can consult [1].

4 Birth of Quantum Mechanics 2. Schrödinger's Formalism

The idea that electrons are in some way tied to a wave, together with the relation between energy and the frequency of the wave, led to recognize that the energy of the electrons inside a nucleus must be quantized. Indeed their stationary wave must satisfy boundary conditions and this sets limitations on its wave-length.

This idea was developed by E. Schrödinger using at first the stationary form of the Klein-Gordon equation modified by the electrostatic potential as suggested by de Broglie. For the electron in the hydrogen atom the equation would be

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \Delta\psi - e(x)\psi, \quad \Delta \equiv \sum_k \frac{\partial^2}{\partial x_k^2} \tag{19}$$

where $\psi(x)$, $x \in R^3$ is twice differentiable and satisfies $\|\psi\|^2 \equiv \int_{R^3} |\psi(x)|^2 dx = 1$. This normalization results from the fact that, according to Born's postulate, the probability of finding the particle somewhere in the entire space must be one.

The results did not fit experimental data. Schrödinger modified then the equation taking (formally) a non relativistic limit. Considering the nucleus fixed and denoting

by x the coordinate of the electron the wave function of electron for the stationary problem at energy E satisfies the equation (e is the charge of the electron)

$$-\frac{\hbar^2}{2m}\Delta u(x) - \frac{e^2}{|x|}u(x) = Eu \qquad \psi_E(x;t) = e^{-iEt}u(x) \tag{20}$$

Schrödinger *solved exactly* this equation. The condition that the solution be normalized ($\|\psi\| = 1$) forces the parameter E to take only a *discrete* set of values. These *energy levels* (bound states) for the Hydrogen atom are in complete agreement with the experimental data.

Equation (20) can be easily generalized for the case of an atom with more that one electron, including also the interaction among the electrons. It is now much more difficult the find an explicit solutions. The problem is simper if the interaction among the electrons is neglected; as a consequence the agreement with empirical results becomes less satisfactory as the atomic number increases.

Equation (20) can be solved exactly also in the case $E > 0$. In this case the solutions behave asymptotically in time as distorted plane waves (distorted because the Coulomb potential decreases slowly in space). This allows to describe scattering of one electron by the nucleus; the result is in complete agreement with the formula of Rutherford scattering and is confirmed by experiments.

These findings on the hydrogen atom gave further stimulus to the idea that the electrons were somehow connected to wave. Schrödinger stressed the need of having *a wave mechanics* for material bodies but put aside the idea of de Broglie that the electrons were *at the same time waves and particles* and considered only the wave aspects.

We shall see in "Lecture 3: Axioms, States, Observables, Measurement, Difficulties" that the ideas of de Broglie have been revived by D. Bohm; later by elaborating on them, D. Dürr, S. Goldstein and collaborators have constructed a rigorous mathematical theory that gives the same results as Schrödinger's Quantum Mechanics for measurements of functions of position.

It must be emphasized that in Schrödinger's theory the waves are *complex-valued functions* and describe *probability amplitudes*. Still, according to the theory, they *can interact*. Indeed we shall see later that according to the probabilistic interpretation given by Born [4], measurements of physical quantities are described by sesquilinear functions of the wave.

Notice that this happens also in the theory of electromagnetism: observable quantities (energy, polarization, Poynting vector) are sesquilinear functions of the electromagnetic field; but in classical electromagnetism, contrary to what happens in Schrödinger's Quantum Mechanics, the equations can be written as real valued equations.

While de Broglie and later Schrödinger stressed the analogy with the variational principles of mechanics and optics, it is worth remarking the basic equation (20) can be obtained by substituting in the classical equation for the energy levels

$$E = \frac{1}{2m}p^2 - \frac{e^2}{|q|} \tag{21}$$

the representation given by de Broglie of the momentum in terms of vibration of waves, i.e. $p = \frac{\hbar}{i}\nabla$. This *derivation* is stressed in most textbooks in quantum mechanics (following P.A.M. Dirac) while the connection with the variational principles is hardly mentioned.

The first results of Schrödinger were on stationary states [14]. Later Schrödinger considered the equation that describes the motion of a particle. Again by taking *formally* the non-relativistic limit he arrived at the equation [18]

$$i\hbar\frac{\partial\phi(t\ \cdot\ x)}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi(t,x) + V(t,x)\psi \tag{22}$$

where $x \in R^3$ and $V(t,x)$ is a generic potential.

The resulting equation (now generally known a Schrödinger equation) can be easily generalized to a system of an arbitrary number of particles. To include the presence of an electromagnetic field Schrödinger generalized (22) to

$$i\hbar\frac{\partial}{\partial t}\psi(t,x) = -\frac{\hbar^2}{2m}(-i\nabla + eA)^2\psi(t,x) + V\psi(t,x) \tag{23}$$

where V is a scalar potential and A is a vector potential associated to the magnetic field [16, 17].

Equations (22) and (23) of Schrödinger's *wave mechanics* and its natural extension to the case of a many-body system is to the present days at the basis of Quantum Mechanics and of its applications. It is *a deterministic equation* and allows to predict completely the wave function at time $t \neq 0$ of an isolated system of N particles when the wave function is known at time $t = 0$, provided the interaction potential satisfies suitable regularity assumption to guarantee existence and uniqueness of the solution to the P.D.E. (22) and (23).

Notice that (22) has complex valued coefficients and its solutions are *complex valued functions*. Therefore the solution $\psi(t,x)$ cannot have a direct physical interpretation. We shall discuss in Chap. 2 the relation between the wave function and measurement.

The possibility to make use of methods of the theory of functions, in which much progress had been made in those years, makes Schrödinger's approach very efficient in the solution of concrete problems.

Schrödinger was aware of the peculiarities of his *wave theory without point particles* and in particular of the *entanglement* which is one of its characteristic feature: if a wave function which describes a composite system is not a product of wave functions of the subsystems, complete knowledge of the state of system as a whole *does not imply even approximate knowledge of the state of its parts*. For this reason Schrödinger was unsatisfied of the conceptual aspects of his own theory [18] and hoped that it would become part of a conceptually more refined theory.

References

1. Baccigaluppi, G., & Valentini, A. (2009). *Quantum theory at the crossroad*. Cambrigde: Cambrigde University Press.
2. Bohr, N. (1917). *Memoirs Academy Roy Dan*.
3. Bohr, N., Kramers, H., & Slater, J. (1924). *Philosophical Magazine*, *47*, 785–802.
4. Born, M. (1924). *Zeitschrift fur Physik*, *26*, 379–395.
5. de Broglie, L. (1927). La mechanique ondulatoire e la stucture atomique de la matiere. *Le Jornal de physique*, *8*, 225–243.
6. de Broglie, L. (1923). Ondes et quanta. *Comptes rendus de l'Academie des Sciences*, *177*, 548–550.
7. de Broglie, L. (1925). Recherches sur la theorie des quanta. *Annales de Physique*, *3*, 22–128.
8. Eckart, P. (1926). *Physical Review*, *28*, 711–726.
9. Ehrenfest, P. (1917). *Philosophical Magazine*, *33*, 500–524.
10. Kramers, H. A. (1924). *Nature*, *113*, 673–676.
11. Lademburg, R. (1921). *Zeitschrift fur Physik*, *4*, 451–468.
12. Lewis, G. (1926). *Nature*, *118*, 874–875.
13. Schrödinger, E. (1926). *Annalen der Physik*, *79*, 734–756.
14. Schrödinger, E. (1926). *Annalen der Physik*, *79*, 361–376.
15. Schrödinger, E. (1926). *Annalen der Physik*, *79*, 489–527.
16. Schrödinger, E. (1926). *Annalen der Physik*, *80*, 437–490.
17. Schrödinger, E. (1926). *Annalen der Physik*, *81*, 109–139.
18. Schrödinger, E. (1926). *Physical Review*, *28*, 1049–1070.
19. Slater, J. (1924). *Nature*, *113*, 307–320.
20. von Laue. (1914). *Annalen der Physik*, *44*, 1197–1212.

Lecture 2: Elements of the History of Quantum Mechanics II

1 Birth of Quantum Mechanics 3: Born, Heisenberg, Jordan

We mentioned in chapter "Lecture 1: Elements of the History of Quantum Mechanics I" that a line of research was centered on finding relations between the structure of the energy levels and the frequency of the emitted radiation and on describing the scattering from a dipole. A preliminary theoretical analysis of the polarization problem was done by van Velt [3, 23].

Later Born [2] made a more detailed study with the purpose of *presenting the results in a form that may give suggestions on how to make the transition from Classical Mechanics to a New Mechanics*. In this paper Born notices that a completely integrable system (such as the Coulomb system) when presented in action-angle variables *appears as a system of harmonic oscillators*. By means of perturbation theory (developed by Hamilton) this description can be carried over to the same system in interaction with the electromagnetic field. If one describes these *virtual oscillators* according to Bohr's rules, their frequencies must satisfy

$$\nu(n, n') = \frac{1}{h}[E(n) - E(n')] \tag{1}$$

where h is Planck's constant and $E(n)$ is the energy of the nth state.

At the conclusions of this fundamental paper Born outlines the guiding lines of the search for a new mechanics and predicts that in this new mechanics finite difference equations will substitute the ordinary differential equations of the old mechanics. Strictly speaking Born's suggestion proved to be incorrect, but a trace of *finite difference scheme* can be seen in the formulation of the theory in terms of Hilbert space operators (infinite matrices).

Bohr's correspondence principle states that if k and n are very large and $\frac{|k-n|}{k}$ is very small, formula (Eq. (13) in chapter "Lecture 1: Elements of the History of Quantum Mechanics I") for the frequency of the radiation emitted or absorbed in the transition between atomic levels must have a classical analogue [1, 3]. According to

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_2

Bohr's correspondence principle the parameters should be determined by comparison with the classical theory.

Each stationary state σ_n of the atom is regarded by Born as *equivalent* to a collection of harmonic oscillators (virtual resonators) with frequency $\nu_{n;m},\ m \in \mathcal{N}$. One has $\nu_{n;m} = K|\nu_n - \nu_m|$ with a universal constant K. Remark that harmonics do not appear in general because $\frac{E_n - E_m}{E_n - E_h},\ k \neq m$ are not in general rational numbers. This avoids in perturbation theory the problem of the *small denominators* that makes perturbation analysis difficult in Classical Hamiltonian Mechanics.

The interaction with an electromagnetic field results in a change in frequencies in the oscillators.

It is convenient at this point to recall some elements of the description in classical mechanics of the interaction between an electric dipole and the electromagnetic field, of which Eq. (14) is an approximation in the case of damped harmonic oscillator. The problem can be stated as follows. A completely integrable system with N degrees of freedom is described in action-angle variables by a hamiltonian $H_0 = F(J)$ where $J \equiv \{J_1, \dots J_N)\}$ are the action variables. We denote by $\theta_k,\ k = 1, \dots N$ the corresponding angular variables.

We set

$$\omega_k = \frac{1}{2\pi}|\nu_k|, \qquad \nu_k = \frac{\partial H_0}{\partial J_k} \tag{2}$$

In hamiltonian mechanics the interaction of an electric dipole with the electric field is described by an interaction hamiltonian $H^{int} \equiv P \cdot E(t),\ P \in R^3$, where P is the total momentum of the system and $E(t)$ is the electric field. We denote by

$$P_0 = \sum_{k_i \in \mathcal{N},\ i=1,\dots N} A_{k_1,..k_n}(J) e^{2i\pi(k_1\omega_1 + \cdots k_N\omega_N)} \tag{3}$$

the electric dipole of the unperturbed system.

We assume that the field is weak (in order to apply perturbation theory) and express this by multiplying the interaction term by a small factor ϵ to keep track of the order of approximation. The interaction hamiltonian is then

$$H^{int} = \epsilon P_0 \cdot E(t) = \epsilon \sum_{k_i} E(t) \cdot A_{k_1,..k_N}(J) e^{\{2i\pi(\omega_1 k_1 + \cdots + \omega_N k_N)\}} \tag{4}$$

To determine (to first order in perturbation theory) the dipole momentum of the system one performs a canonical transformation $J_k, \theta_k \to J_k^\epsilon, \theta_k^\epsilon$ chosen so that the total hamiltonian $H^\epsilon \equiv H_0 + \epsilon H^{int}$ is written in the new canonical variables as

$$H_0(J) + \epsilon H^{int}(J, \theta) = H_\epsilon(J^\epsilon) + \epsilon^2 K(J^\epsilon, \theta^\epsilon, \epsilon) \tag{5}$$

where K is a suitable function, which is assumed to be regular.

By Hamilton's perturbation theory the function H^ϵ in the new canonical variables takes the form

$$H^\epsilon(J^\epsilon \cdot \theta^\epsilon) = H_0(J^\epsilon) + E \cdot p_1(J^\epsilon) + E \cdot p_0(J^\epsilon) + O(\epsilon^2) \tag{6}$$

where for each (small) value of ϵ, $J^\epsilon \equiv \{J_1^\epsilon, \dots J_N^\epsilon\}$ are a new system of variables in involution and $p_0(J^\epsilon)$ is the average of p_0 over the angles θ_k. Notice that H^ϵ in (6) depends on θ^ϵ at order ϵ^2 only.

The term $E \cdot p_1(J^\epsilon) + E \cdot p_0(J^\epsilon)$ represent the effective energy of the system (to first order in ϵ) and therefore

$$p_1(J^\epsilon) + p_0(J^\epsilon) \tag{7}$$

is the *effective electric dipole* of the system (to first order in ϵ). Notice that (6) describes the system in term of *adiabatic invariants* (the action variables).

If $E(t)$ is mono-cromatic with frequency ν_0 the average of the function $H(J, \theta)$ over the torus is zero. The canonical transformation which produces (6) is obtained through a generating function $S(\theta, J^\epsilon, \epsilon)$ by setting

$$\theta_k^\epsilon \equiv \frac{\partial S}{\partial J_k}, \qquad J_k^\epsilon \equiv \frac{\partial S}{\partial \theta^k} \tag{8}$$

The function S is a solution of the Hamilton-Jacobi equation to first order in ϵ. The new momentum is to first order in ϵ

$$p_0 + \epsilon p_1; \quad p_1 = \sum_k \left[\frac{\partial p_0}{\partial J_k} \frac{\partial S}{\partial \omega_k} - \frac{\partial p_0}{\partial \omega_k} \frac{\partial S}{\partial J_k} \right] \tag{9}$$

Using (8), (9)

$$p_1 = -E \cos 2\pi\nu_0 t \sum_k \sum_{\nu\cdot\tau>0} \tau_k \frac{\partial}{\partial J_k} \left(\frac{2|A_\tau|^2 \, \nu \cdot \tau}{(\nu \cdot \tau)^2 - \nu_0^2} \right) \tag{10}$$

where $\nu^k = \frac{\partial H_0}{\partial J^k}$ are the frequencies of the oscillators (for each value of k the index τ_k runs over the integers.

A hint on the structure of the *new mechanics* should be obtained comparing (10) with the empirical formula obtained by Kramers.

By construction the sequence $\{E_n\}$ is increasing and has a limit which we conventionally take to be zero (it is the ionization threshold). For n large and $n-m$ of order one $E_m - E_n$ is *infinitesimal* with respect to E_n. The correspondence principle states that near the ionization threshold the quantum laws should be comparable to the classical laws. Therefore the parameters should be chosen through a comparison of (10) with Kramer's empirical formula [17].

The correspondence rule that emerges by the analysis by Born is

$$\nu \cdot \tau \to \nu_{n,m}, \qquad J \to nh \tag{11}$$

The second arrow in (11) is Bohr's correspondence principle.

To understand better the role of the first arrow notice that one can envisage an artificial adiabatic process under which the system goes from the state n to the state m through (fictitious) very small intermediate steps. One has then

$$\nu \cdot \tau = \sum_k \frac{\partial H_0}{\partial J_k} \tau_k \sim = \frac{1}{h} \sum_k \frac{\partial H_0}{\partial J_k} \frac{\partial J_k}{\partial \mu} = \frac{1}{h} \frac{\partial H_0}{\partial \mu} \tag{12}$$

This approximation is better justified if μ is small and therefore if $n - m$ is of order of magnitude αh where α is very small.

On the other hand, by Einstein's rule, if $m = n + \tau$ one has

$$\nu_{m,n} = \frac{1}{h} |E(n + \tau) - E(n)| \tag{13}$$

Comparing (12) with (13) one sees that the operation performed in passing from the classical rules to the quantum ones consists in *substituting differentials with finite difference quotients* (the finite differences being of order h). This allows the use of the correspondence principle to obtain a relation between the classical and quantum coefficients.

Consider the case $\frac{\tau}{n} << 1$. Then

$$\sum_k \tau_k \frac{\partial \Phi}{\partial J_k} \quad \to \quad \int_0^1 d\mu \sum_k \tau_k \frac{\partial \Phi}{\partial J_k} \tag{14}$$

From Bohr's quantization rule, if $\frac{\tau}{n}$ is sufficiently small (and therefore $\frac{\tau h}{J}$ sufficiently small) one can consider $\tau_k d\mu \simeq d J_k$ and therefore

$$\int_0^1 \sum_k \frac{\partial \Phi}{\partial J_k} \tau_k d\mu \simeq \int_0^1 \sum_k \frac{\partial \Phi}{\partial J_k} d J_k = \frac{1}{h} (\Phi(n + \tau) - \Phi(n)) \tag{15}$$

This identification must hold for any quantum observable in the limit $\frac{\tau}{n} \to 0$. Moreover, since $|A_\tau(J)|^2 = A_\tau(J) A_{-\tau}(J)$ and $A_{-\tau}(J) = A^*_\tau(J)$ one must have

$$|A_\tau(J)|^2 \equiv \Gamma(n, m) = \Gamma(m, n), \qquad m = n + \tau \tag{16}$$

Performing in (10) the substitutions indicated in (13)–(15) and recalling the definition of Γ in (16) one obtains

$$p_1 = E\cos(2\pi\nu_0 t)\frac{1}{h}\sum_{\tau_k>0}[\frac{2\Gamma(n+\tau,n)\nu_{n+\tau,n}}{\nu^2_{n+\tau,n}-\nu^2_0} - \frac{2\Gamma_{n,n-\tau}\nu_{n,n-\tau}}{\nu^2_{n,n-\tau}-\nu^2_0}] \qquad (17)$$

This expression must be compared with Kramer's empirical rule (Eq. (16) in Lecture 1). There is a fair agreement if one chooses

$$-\frac{e^2}{4\pi^2 m}f_{n,m} = \frac{1}{h}2\Gamma(n,m)\nu_{n,m}, \qquad \Gamma(n,m) \equiv |A_{m-n}|^2 \qquad (18)$$

It follows from (18) that the knowledge of the emission and absorption frequencies *is not sufficient* to determine the matrix elements $A_{n,m}$. Only their absolute values are determined; to determine the phases it is necessary to go further in the order of approximation or to study a problem in which the coupling to the electric field has a different expression (e.g. Heisenberg approximated the atom by an anharmonic oscillator).

Recall that $A_\tau(J)$ are the elements of the series expansion of the momenta p_k as functions of the angles θ_k. Therefore in accordance with Kramers's formula in quantum mechanics the momentum is represented by a quantity which *depends on two indices* m and n, namely *by a matrix*. When $m_k = n_k + h\tau_k$ and τ_k is small with respect to n_k the quantity $A_{n+h\tau_k,n}$ plays the role of $A_\tau(J)$ in the classical case, where J is the set of action variables associated to the state n.

2 Birth of Quantum Mechanics 4. Heisenberg and the Algebra of Matrices

The next step in the construction of the new mechanics was taken by W. Heisenberg. He analyzed the connections among the quantities of type $A_{n,m}$ associated to classical quantities other that momentum. These could be obtained studying other types of interactions with the electromagnetic field e.g. making use of interactions in which the classical form of the interaction hamiltonian is of the type $x \cdot E$ (e.g. the formulas for the polarization in a slowly varying electric field).

In this way one can determine the matrices which the new mechanics associates to the observable position and more generally to observables that in hamiltonian mechanics are described by polynomials in the canonical variables position and momentum.

Heisenberg [10] did a detailed analysis of the corresponding formulas and in particular of those that refer to the anomalous Zeeman effect, which is described in classical theory by the equation $\ddot{x} = -\omega^2 x - \epsilon\, x^4$. This led to establish the following correspondence (to the left the classical case, to the right the quantum case)

$$\begin{aligned} \nu(n\tau) \equiv \tau\nu(n) \equiv \tau\frac{1}{h}\frac{\partial E}{\partial n} \;&\rightarrow\; \nu_{n,n-\tau} = \frac{1}{h}(E(n) - E(n-\tau)) \\ \nu_{n,\tau_1} + \nu_{n,\tau_2} = \nu_{n,\tau_1+\tau_2} \;&\rightarrow\; \nu_{n,n-\tau} + \nu_{n-\tau,n-\tau-\tau_1} = \nu_{n,n-\tau-\tau_1} \end{aligned} \qquad (19)$$

On this basis Heisenberg stated that the correspondence

$$a_\tau(n)\, e^{it(\omega(n)\cdot\tau)} \to A_{n,n-\tau}\, e^{it\omega(n,n-\tau)} \tag{20}$$

holds for any classical observable a which can be expressed in the form $a = \sum_\tau a_\tau(n) e^{it(\omega(n)\cdot\tau)}$.

Heisenberg found the following relation between the quantum representatives $\hat{a}, \hat{b}$ of the classical observables $a,\ b$

$$\hat{b}_{n,n-\tau} = \sum_\alpha \hat{a}_{n,n-\alpha}\hat{a}_{n-\alpha,n-\tau} \tag{21}$$

This is to be compared to the classical case

$$(a^2)(n,t) = \sum_\alpha a^2_\beta(n) e^{it(\omega(n)\cdot\beta)} = \sum_{\alpha,\beta} a_\alpha(n) a_{\beta-\alpha}(n) e^{it[(\omega(n)\cdot\alpha)+(\omega(n)\cdot(\beta-\alpha)]} \tag{22}$$

Remark that (21) is the product rule for matrices, extended to the case of matrices of infinite rank.

This analysis was summarized in two papers by Born and Jordan [4] and by Born et al. [5]. The latter paper is nicknamed *the three men's work*. In this paper the Authors state that, rather then adapting the formalism of Classical Mechanics in an artificial way, *an organic description of a new theory is established* and a mathematically coherent theory is presented which *describes the properties which is characteristic of quantum phenomena and at the same time shows a remarkable analogy with Classical Mechanics*.

In this [5] the Authors speak for the first time explicitly of a *symbolic quantum geometry* which tends for small values of h to the *visualizable geometry of Classical Mechanics*. The Authors also speak of *relations among observables* and state that any observable *can be represented by an infinite matrix* (i.e. a linear operator in an infinite dimensional Hilbert space). The Authors stress that these matrices are not of the same type as those which Hilbert was studying in the same years (at that time both the Authors and D. Hilbert were working in Göttingen). The operators that Hilbert was studying are now called Hilbert-Schmidt operators; we will verify that the operators which represent position and momentum *cannot be of Hilbert-Schmidt type*.

In [4, 5] the Authors develop the quantum matrix calculus, establish perturbation theory in Quantum Mechanics (in strict analogy with hamiltonian perturbation theory) and develop in detail the formalism up to second order. The example treated in more detail is that of the anharmonic oscillator with a fourth order anharmonic term. The results were in good agreement with experimental data. It is curious to observe that in this case the perturbation series does not converge. The series is however asymptotic, therefore for small perturbations the analysis to second order gives a satisfactory answer.

The contents of these important papers establish *the essential part* of the present day algebraic-axiomatic formulation of Quantum Mechanics. One can evidence

- the use of methods of *simbolic differentiation* that (in a present days language) substitutes the vector field of Classical Mechanics with the algebraic expression *commutator of two matrices*. In this respect a special role is taken by the matrices $\hat{q}_k$, $\hat{p}_k$ which are associated to the coordinates in phase space.
- the writing of the equations in the form

$$\frac{d\hat{q}_k}{dt} = i[\hat{H}, \hat{q}_k] \qquad \frac{d\hat{p}_k}{dt} = i[\hat{H}, \hat{p}_k] \quad k = 1, \dots N \tag{23}$$

  where the matrix $\hat{H}$ describes the interactions present in the system and N is the number of degrees of freedom of the classical system.
- the proof that the Eq. (23) are (formally) variational equation for the functional

$$\int [(\hat{p}, \hat{q}) - \hat{H}(\hat{q}, \hat{p})] dq dp \qquad q = \{q_1, \dots q_N\} \quad p = \{p_i, \dots p_n\}$$

  (if the definition of integral is properly interpreted).
- the *demonstration* that the quantum counterpart of

$$1 = 2\pi \sum_{\tau} \{q_\tau^h, p_\tau^h\}, \quad p_\tau^h \equiv \nabla_\tau(q_\tau^k) \tag{24}$$

  (where $\{q, p\}$ denotes Poisson brackets) is the identity

$$[\hat{p}_k, \hat{q}_h] = \frac{h}{2i\pi} \delta_{k,h} \, I \tag{25}$$

  where I is the (infinite dimensional) identity matrix.

Still it should be remarked that the other identities

$$[\hat{p}_k, \hat{p}_h] = [\hat{q}_k, \hat{q}_h] = 0, \qquad k \neq h \tag{26}$$

is postulated by Born, Heisenberg and Jordan without strictly convincing arguments.

From this brief analysis it is clear that in the new mechanics a special role is reserved to infinite matrices $\hat{q}_k$, $\hat{p}_h$ which satisfy, at least formally, the commutation relations

$$[\hat{q}_k, \hat{p}_h] = i\hbar \delta_{h,k} I, \quad h, k = 1 \dots N \quad \hbar = \frac{h}{2\pi} \tag{27}$$

all other commutators being set equal to zero.

Notice that in view of (27) the matrices that appear in quantum mechanics are complex valued. We shall see that the natural space for their action is the linear space of sequences of complex numbers $\phi \equiv \{c_{n,m}, \quad m \in Z\}$ such that $\sum_{m=1}^{\infty} |c_{n,m}|^2 = 1$.

From (2) follow Heisenberg's *uncertainty relations* i.e. that there is no basis in which the matrices $\hat{q}_k$ and $\hat{p}_k$ can be simultaneously diagonalized (we shall see later a more precise statement).

Later we shall see that the context of the theory led to interpret $|c_k|^2$ as the probability that system be in the kth (atomic) state.

It follows that $\sum |c_k|^2 = 1$. If one endows the linear space with the scalar product

$$(\phi^1, \phi^2) \equiv \sum_k (c_k^1)^* \, c_k^2 \tag{28}$$

the space becomes the Hilbert space $l^2(N)$.

We shall see that in order to represent observable quantities the matrices must have real eigenvalues. This forces these matrices to be hermitian. In fact a stronger requirement must be satisfied, i.e. acting on $l^2(N)$ they must be self-adjoint (in the course of these lecture we will explain the difference).

A relevant contribution to the success of the new Quantum Mechanics was the analysis made by Pauli [20] of the spectrum of the Hydrogen atom using only the algebraic rules of matrix mechanics, i.e. only the commutation relations between the operators (matrices) obtained by using the algebraic rules for the generators of the rotation group and the Runge-Lenz vector (which are constants of motion) for fixed values of the hamiltonian. We shall sketch the analysis of Pauli at the end of this lecture and give more details in chapter "Lecture 18: Weyl's Criterium, Hydrogen and Helium Atoms".

The analogy of the new formalism with hamiltonian Dynamics permits also a description in the new theory of the interactions of particles in space, in particular scattering. In the same way as Pauli did for the hydrogen atom one can use the algebraic rules of matrix mechanics to describe Rutherford's scattering by an atomic nucleus.

3 Birth of Quantum Mechanics 5. Born's Postulate

A very important step in the formulation of Quantum Mechanics was taken by M. Born. He noticed that all waves corresponding to atomic states were square integrable while in classical mechanics one has integrability of the (real valued) distribution of charge, masses etc.

This suggests that $|\phi(x)|^2$ have a role similar to density. Since the particle is not a fluid Max Born assumed that the real positive function $|\phi(x)|^2$ represents a density of probability i.e. a *probability density*.

In particular, *if one performs a measurement of position*, $\int_\Omega |\phi(x)|^2 dx$ is the probability that the particle be found in Ω. It follows that

$$\int_{R^3} F(x)|\phi(x)|^2 dx \tag{29}$$

gives the average of the results that are obtained measuring the observable $F(x)$ in a state described by ϕ.

In the same way, according to Born, if $\hat{\phi}(p)$ is the Fourier transform of ϕ, the quantity $|\hat{\phi}(p)|^2$ represents the probability density that, *if one measures the momentum of the particle*, the integral

$$\int_{R^3} G(p)|\hat{\phi}(p)|^2 dp \tag{30}$$

gives the average result that one obtains making a measurement of the observable $G(p)$ in a state described by ϕ.

In the Hilbert space terminology, Born's postulate takes the form

$$\bar{F}_\phi = (\phi, F(x)\phi), \qquad \bar{G}_\phi = (\hat{\phi}, G(\hat{p})\hat{\phi}) \tag{31}$$

and can be extended to any other quantum observable if one were able to associate to any observable a an operator A. The mean value of a in the state ϕ will then be $\bar{a}_\phi = (\phi, A\phi)$. In order to obtain real numbers the operator A must symmetric.

4 Birth of Quantum Mechanics 6. Pauli; Spin, Statistics

Returning now to Schrödinger's formulation of Quantum Mechanics we remark the solution to evolution equation

$$i\hbar \frac{\partial \phi}{\partial t} = H\phi \tag{32}$$

where ϕ belongs to a complex Hilbert space $\mathcal{H}$ and H is a suitable operator is (at least formally)

$$\phi(t) = e^{-i\frac{t}{\hbar}H}\phi(0) \tag{33}$$

By Born's postulate the map $\phi(0) \to \phi(t)$ must be unitary and this implies that the operator H is self-adjoint (we will explain later the difference between "self-adjoint" and "closed and symmetric"; in the finite-dimensional case there is no difference).

Moreover Born's postulate implies the wave functions that differ by a constant phase *represent the same state*. This property implies that symmetries of the system

under a group $\mathcal{G}$ of transformations are described by *projective representations* of $\mathcal{G}$ in the Hilbert space. An important example is symmetry under rotations and the introduction of the *spin*.

To account for the hyperfine structure of the spectrum of the atoms, in particular of the helium atom, Pauli [15, 16] *postulated* the existence of particles *of spin* $\frac{1}{2}$, introducing thereby *the spin*, a quantity extraneous to Classical Mechanics.

The wave function of a particles of *spin* $\frac{1}{2}$ transforms under rotations according to a faithful representation of the SU_2 group (which is a double covering of the rotation group) acting in a two-dimensional complex Hilbert space. This is a projective representation of the rotation group. It is by definition *a spinor*. The name *spin* is somewhat connected with spinning, i.e. set something in rotation. The notation $\frac{1}{2}$ has its origin in the fact that the product of two two-dimensional representations of SU_2 contains the (real) vector representation of the rotation group which corresponds to angular momentum one. Therefore in some sense these particles have half unit of angular momentum.

The presence of spin doubles the number of atomic levels, and the coupling of spin degree of freedom with the magnetic field accounts for a small difference in energy between the two levels in a pair, i.e. for the *hyperfine structure* of the emission lines.

The doubling of the number of degrees of freedom has an another important consequence.

Consider

$$\mathcal{H} \equiv L^2(R^3) \times C^2, \quad \Phi \in \mathcal{H} \equiv \{\phi_1, \phi_2\} \quad \phi_i \in L^2(R^3) \tag{34}$$

as a Hilbert space with scalar product

$$(\phi_k, \phi_h) = \int \bar{\phi}_k(x) \cdot \phi_h(x) dx \qquad k, h = 1, 2 \tag{35}$$

In this space which is isomorphic to $L^2(R) \otimes C^2$ one can define, following Pauli, a first order differential operator $\tilde{\nabla}$ whose "square" is minus the Laplacian times the identity. It is defined by the matrix-value differential operator

$$\tilde{\nabla}\Phi = \sum_{h=1}^{3} \sigma_h \nabla \Phi \tag{36}$$

where the 2×2 hermitian matrices σ_i (Pauli matrices) are such that

$$[\sigma_k \sigma_h] = i\epsilon_{k,h,j}\sigma_j, \qquad k, h, j = 1, 2, 3 \tag{37}$$

Here ϵ is the Ricci symbol, taking value zero if two of the indices are equal, plus one if the permutation is even and minus one if the permutation is odd. One has $Tr(\sigma_k) = 0 \; \forall k$ (the matrices σ_k have trace zero).

It is also easy to verify that

$$(\tilde{\nabla})^* \tilde{\nabla} = -\Delta \times I \tag{38}$$

Therefore the operator $\tilde{\nabla}$ deserves the name of "square root" of minus the Laplacian (notice that the Laplacian is a negative operator as one sees taking Fourier transforms).

Notice now that from (38) it follows that the solution of the equation

$$i \frac{\partial \Phi(t,x)}{\partial t} = (-\Delta \times I)\Phi(t,x) \qquad \phi \in L^2(R^3) \otimes C^2 \tag{39}$$

is a spinor with components *that satisfy the free Schrödinger equation*.

In the interacting case, one distinguishes between interaction with the electromagnetic field and other type of interactions. As in the classical case, the interaction with the electromagnetic field is described by adding a vector potential A to the momentum and a scalar term (electrostatic potential) to the potential.

In the Schrödinger equation this results in substituting $i\tilde{\nabla}$ with $i\tilde{\nabla} - A(t,x) \times I$ (recall that $i\nabla$ is a symmetric operator) and adding a term $U(t,x)$ to the interaction potential.

Therefore Eq. (39) becomes

$$i \frac{\partial \Phi}{\partial t} = (i\tilde{\nabla} + A \times I)^2 \Phi(t,x) + V(x)\Phi(t,x) \tag{40}$$

for some potential V (that includes the electrostatic potential).

Another very important consequence of the fact that wave function which differ only by a constant phase represent the same state is the possibility to introduce the *statistics of identical particles*. The (elementary) particles are subdivided in two distinct categories: that of *bosons* for which the permutation of the indices of identical particles does not alter the wave function (Bose-Einstein statistics) and that of *fermions* in which this operation multiplies the wave function by a factor -1 (Fermi-Dirac statistics).

It follows that the wave function of a state which describes two identical fermions cannot be the product of the corresponding wave function: two fermions cannot be in the same state (Pauli *exclusion principle*) [16]. This is at the basis of the properties of the spectra of atoms and molecules and is also responsible for the *stability of matter*.

To the contrary the wave function $\psi(x_1, \ldots x_N)$ of any number of bosons may be the product of the same wave function, i.e.

$$\psi(x_1, \ldots x_N) = \phi(x_1) \cdots \phi(x_N) \tag{41}$$

i.e. two identical bosons are allowed to occupy the same state.

This is at the root of the properties of the black-body radiations if one assumes that the photons satisfy the Bose-Einstein statistics; it also at the root of the *Bose-Einstein condensation*, a phenomenon predicted by Bose and Einstein and observed experimentally only in recent years.

The ability to account in a simple way for phenomena which don't have a classical counterpart has contributed to the success of Quantum Mechanics.

Notice that *it is a matter of fact* that in nature the wave function of a bosons transform under rotation according to a representation of the rotation group while the wave function of a fermion transforms according to a representation of $SU(2)$.

Notice that there is a strict correspondence between the symmetry properties under exchange of indices and the behavior under the rotation group. This fact has *no explanation* within non-relativistic Quantum Mechanics although is true that only in the case of particles with spin $\frac{1}{2}$ one is naturally led to use a Hilbert space that admits a non-trivial representation of the permutation group. In the relativistic Quantum Field Theory the connection between spin and statistics is a consequence of locality and positivity of energy.

5 Further Developments: Dirac, Heisenberg, Pauli, Jordan, von Neumann

Soon after the proposals of Wave Mechanics by Schrödinger and of matrix mechanics by Born, Heisenberg and Jordan, the *equivalence*, at least at a formal level, of the two formulations was noticed by Schrödinger [21], Eckart [9], Jordan [13], Lademburg [18], Pauli [19], Dirac [6].

It is indeed easy to verify that the canonical commutation relations (27) are satisfied (at least formally) by operators that act on spaces of function on $L^2(R^n)$ as follows

$$\hat{q}_k\psi(x) \equiv x_k\psi(x), \qquad \hat{p}_k \equiv -i\hbar\frac{\partial\psi(x)}{\partial x_k} \quad k = 1\ldots n \tag{42}$$

Notice that the operator $-i\hbar\frac{\partial\psi(x)}{\partial x_k}$ has as (generalized) eigenvectors the de Broglie's states of definite values of momentum. From this point of view the indeterminacy relations between position and momentum, justified by Heisenberg considering *virtual experiments*, is a consequence of the properties of the Fourier transform.

The similarity between the formulation of dynamics in Hamiltonian Mechanics and in Quantum Mechanics has its roots in the fact that the basic elements of the two formalisms have *the same algebraic structure*. If the hamiltonian in the classical case is taken to be $H_{class}(q, p)$ the evolution of the observables is given by

$$\frac{dA_{class}}{dt} = \{A_{class}, H_{class}\} \tag{43}$$

where $\{f, g\}$ is Poisson bracket.

Formally the solution of the Schrödinger equation $i\hbar\frac{d\phi}{dt} = H\phi$ is $\phi(t) = e^{-itH}\psi$. By duality it follows that the evolution of the observables $\hat{A}$ (operators in the Hilbert space in which the system is described) is described by the equation

$$\frac{d\hat{A}}{dt} = i[\hat{A}, H] \tag{44}$$

where $[A_1, A_2] \equiv A_1 A_2 - A_2 A_1$ is the commutator of the matrices A_1, A_2.

The operation on the right hand side in (43) (which acts on functions on phase space) and that in (44) (which acts on bounded operators in a complex Hilbert space) have the same algebraic structure: they are *-derivations, i.e. the commute with taking adjoints, satisfy Leibnitz's rule and Jacobi's identity. Notice that in Quantum Mechanics the imaginary unit i takes the role of the symplectic structure J (recall that J is the imaginary unit in the presentation of the symplectic structure as complex structure). This algebraic homeomorphism was particularly emphasized by P. Dirac.

Soon after the formulation of Matrix Mechanics P. Dirac, who had become aware of those results without probably knowing many details, develops the *Quantum Algebra* [7]. In this very important paper Dirac introduces explicitly the terms *Quantum Algebra*, *Quantum Differentiation*, *Quantum Poisson Brackets* and remarks that the relation between Hamiltonian Mechanics and Quantum Mechanics lies in the isomorphism of the underlying algebraic structures.

Still, same care has to be taken in exploiting this relation because the algebra of functions on phase space is well defined while the algebra of unbounded operators in a Hilbert space must be treated with care. This leads to some difficulties if one attempts to formulate precisely the correspondence between the two theories.

6 Abstract Formulation

The formalism of Quantum Mechanics was later described in more mathematical terms by von Neumann [22], without restriction to the quantization of the canonical variables and connecting the formalism to the theory of algebras of operators in a (separable) Hilbert space. Very important contributions in this direction were made by A. Weyl; these Authors proved that, under suitable assumption, all representations of of the canonical commutation relations are unitary equivalent. Later in these Lectures we will come back to this point and make more precise the statement of unitary equivalence.

Research on these more abstract aspects of Quantum Mechanics have led to major developments in the theory of C^*-algebras and of partial differential equations. It is safe to say that the main progress in these fields came through deepening the answers to questions that arise in Quantum Mechanics.

The exploitation of the structure of Quantum Mechanics and the need to refine the mathematical instruments led quickly to a distinction between researches on Quantum Mechanics: those with primary interest in the mathematical aspects and those with main interest in the properties of specific systems: atoms, molecules, crystals, semiconductors. The analysis of these systems have an independent mathematical interest and have greatly contributed to the field of Quantum Chemistry and Solid State Theory.

The first line of research has led to an axiomatic formulation of the theory and has laid the bases for the mathematical treatment of systems with infinitely many degrees of freedom (Quantum Field Theory, Algebraic Quantum Theory and Quantum Statistical Mechanics). This line of research favors in general algebraic structures.

The second line of research uses mostly Schrödinger's representation, and therefore its mathematics is mostly in the field of functional analysis and partial differential equations. This line of research benefits greatly from the *visualization* associated to the use of configuration space. The extraordinary success of Quantum Mechanics in the field of advanced technology comes from this line of research.

Both lines of research have put little emphasis on conceptual problems, e.g. on the *theory of measurement*, which is considered trivially solved in Classical Mechanics and is up to now far from being solved in Quantum Mechanics. We shall come back to this point in chapter "Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories".

7 Quantum Field Theory

Soon after the writing of Quantum Mechanics in quantum canonical variables, the structure was extended, at least formally, to system with an infinite number of degrees of freedom by Heisenberg and Pauli [11], Jordan and Pauli [15], and by Dirac [7].

The extension was naturally accomplished by choosing a basis of functions in the Hilbert space $L^2(R^3)$ and promoting the functions in the basis chosen to be "quantum coordinates" satisfying Heisenberg commutation relations. A natural field of application is given by the Maxwell equations (quantum electrodynamics) and the Klein-Gordon equation (particle physics). Classically these systems are described by P.D.E. (partial differential equations) with a natural symplectic structure. The choice of a basis of functions $f_k(t,x)$ turns this formalism into a system of differential equation in infinite dimension. These equations can be written, introducing a (formal) symplectic form, as Hamilton equations for an infinite set of harmonic oscillators.

One can *quantize* the system, at least formally, selecting a basis in $L^2(R^3)$ and introducing quantum coordinates $\hat{q}_k$, $\hat{p}_k$ associated to the elements of the basis chosen. One can then define *quantum fields* $\phi(t,x)$, $\pi(t,x)$ by

$$\hat{q}_k \equiv \int \phi(t,x) f_k(x) d^3x, \quad \hat{p}_k \equiv \int \pi(t,x) f_k(x) d^3x \tag{45}$$

The basis can be chosen in such a way that the resulting commutation relations for the fields be (at least formally)

$$[\phi(t,x), \pi(t,y)] = i\delta(y-x) \quad [\phi(t,x), \phi(t,y)] = [\pi(t,x), \pi(t,y)] = 0 \quad x, y \in R^3 \tag{46}$$

where if $z \in R^3$ one defines $\delta(z) \equiv \delta(z_1)\delta(z_2)\delta(z_3)$. The symbol $\delta(w)$, $w \in R^1$ (invented by Dirac) [7] is defined by

$$\int \delta(w - w')f(w')dw' = f(w) \tag{47}$$

for any continuous function f. In treating electrodynamics one has to pay special attention to gauge invariance. Notice that through (45)–(47) we have introduced *quantum fields* through which we describe the quantum mechanical version of the classical fields.

The formulation (46) through the use of Dirac distributions evidences *local properties* of the fields. It has set the basis for the treatment of the Quantized Electromagnetic Field and the development of Quantum Electrodynamics. An equivalent formulation has been given by Heisenberg and Pauli and later by Pauli and Wigner using proper bases in the Hilbert space. Although Jordan and Klein [12, 14] proved the equivalence of the two types of field quantization, Dirac's approach is remarkable for simplicity and clarity of exposition and has remained a milestone in Quantum Mechanics. It is reported in almost all textbooks on Quantum Field Theory and in almost all research papers. Only occasionally one finds reference to the more mathematically correct quantization which uses an orthonormal basis in the Hilbert space.

The fact that the system has now an infinite number of degrees of freedom gives rise to formal difficulties. These difficulties can be overcome in the case of a free field theory but if one introduces a *relativistic local* interaction between the fields one rums into very serious difficulties, mainly due to the distributional properties of the fields and to the need to control convergence of the formal series. The former can be attacked with appropriate redefinitions of products of distributions; convergence of the series is more difficult; in favorable cases one prove Borel summability.

Very soon Fock, and later Dirac, noticed that the *quantization of the fields* could be given a different form. Take for simplicity the case of the wave equation (massless Klein-Gordon equation) in one dimension

$$\frac{\partial^2}{\partial t}\phi(t, x) = \frac{\partial^2}{\partial x}\phi(t, x) \tag{48}$$

which in Fourier basis can be written

$$\frac{\partial^2}{\partial t}\hat{\phi}(t, k) = k^2\hat{\phi}(t.k) \tag{49}$$

i.e. as an infinite (non-denumerable) collection of harmonic oscillators.

On the other hand, setting

$$a(k) = \frac{1}{\sqrt{2}}(\hat{p}(k) + i\hat{q}(k)), \qquad a^* = \frac{1}{2}(\hat{p}(k) - i\hat{q}(k)) \tag{50}$$

it is easy to verify for each k that the operator $a^*(k)a(k)$ is the hamiltonian of a quantum harmonic oscillator in $L^2(R)$ with elastic constant $|k|$. This operator is selfadjoint and has (simple) eigenvalues $(n+\frac{1}{2})k^2,\ n = 0, 1, \ldots$; the corresponding normalized eigenfunctions $\psi_n(k)$ are the Hermite polynomials.

We have therefore obtained two representations of the scalar field of mass zero: one through the operators $\hat{Q}(k),\ \hat{P}(k)$ and one through the operators $a(k), a^*(k)$.

The first is a quantization of the solutions of the wave equation. To each configuration of the classical field corresponds in this quantization a quantum state. This state can be written as superposition (with prescribed phases) of elements which an increasing number of particles (coherent states). This representation of the (quantized) free electromagnetic field which is commonly used in Quantum Optics.

The representation through the operators $a_k,\ a_k^*$ is particularly convenient when one introduces an interaction that does not preserve the number of particles. The quantization usually employed to describe the interaction of particles with the electromagnetic field (modulo difficulties connected with gauge invariance).

8 Anticommutation Relations

The Dirac-Fock representation is also interesting because it suggests how to provide in a simple way a Fock space for fermions through the introduction of anticommutation relations. Indeed if f, g are test functions and $a(f) \equiv \int a(x)f(x)dx$ we require that the following relations (anti-commutation relations) be satisfied

$$\{a(f), a^*(f)\} = |f|^2, \quad \{a(f), a(f)\} = a^2(f) = 0 \quad \{a^*(f), a^*(f)\} = (a^*(f)) = 0 \tag{51}$$

where $\{b, c\} \equiv bc + cb$. It follows that two fermions cannot be in the same state and $n_f \equiv a^*(f)a(f)$ can be either zero or $|f|^2.I$.

This field quantization through *anticommutation relations* has the advantage of incorporation "Pauli statistics" (in an unpublished manuscript Jordan described what is now called Fermi-Dirac statistic and attributed it to Pauli) [1, 12, 13]. We shall discuss the anti-commutation relations in the second part of these Lecture Notes.

Jordan returned later to this field quantization in a paper together with Wigner [16] and gave an alternative derivation of the anti-commutation relations. Jordan arrived at the description of anti-commuting fields through an attempt to overcome the ambiguity in the order of operators in Quantum Mechanics. He defined a new product for operators that is symmetric the *Jordan product*. For two operators A, B it is defined as

$$A \cdot B \equiv \frac{1}{2}(AB + BA) \equiv \frac{1}{2}\{A, B\} \tag{52}$$

This product is commutative but not associative. It satisfies however a weak form of associativity

$$(A^2 \cdot B) \cdot A = A^2 \cdot (B \cdot A), \qquad A^2 \equiv A \cdot A \tag{53}$$

Consider the bilinear operation on hermitian operators

$$[A, B] = \frac{i}{2}(AB - BA) \tag{54}$$

(the imaginary unit is there to assure hermiticity). This antisymmetric bilinear operation satisfies the Jacobi identity

$$[[A, B]C] + [[B, C]A] + [[C, A], B] = 0 \tag{55}$$

and therefore defines a Lie algebra structure. The lack of commutativity of the Jordan product is linked to this Lie structure by

$$(A \cdot B) \cdot C - A \cdot (B \cdot C) = [B, [A, C]] \tag{56}$$

and is is easy to see that bilinear forms in the operators $a(f), a^*(f)$ satisfy (54)–(56) *if and only if* the anticommutation relations (51) are satisfied. Recalling the in classical electromagnetism the observable (currents, Pointing vector) are quadratic in the field this led naturally Jordan and Wigner to introduce fields that obey Fermi-Dirac statistics.

In this way Jordan in an unpublished manuscript arrived through his *Jordan algebra* to a formulation of what is now called *Fermi-Dirac statistics* for particles which satisfy Pauli's exclusion principle.

It should also be remarked that Jordan [13, 14] was the first to generalize the new Quantum Mechanics to a system with infinitely many degrees of freedom by *quantizing the wave function* and introducing a (formal) functional calculus for a collection of functions on $\oplus_{n\in N} L^2(R^d)$. This is a quantization adapted to the classical Lagrangian field formalism. One may say that the purpose of Jordan was to introduce as *configuration space* a space in which the point are substituted with suitable (generalized) functions. The quantized fields would appear as (non-commuting) functionals on this space. It was therefore more akin to the quantization of the solutions of the wave equation introduced later by I. Segal.

9 Algebraic Structures of Hamiltonian and Quantum Mechanics. Pauli's Analysis of the Spectrum of the Hydrogen Atom

We noticed that the basic structures with which time evolution is generated *in both theories by derivations* i.e. operations which are linear and satisfy Leibnitz's rule. In Hamiltonian Dynamics they are given by Poisson brackets on the algebra of

functions on phase space. In Quantum Dynamics they are given by fs on the algebra of observables. These structures satisfy Leibnitz' rule and Jacobi's identity and are algebraically isomorphic. This algebraic isomorphism allows to set up easily a perturbation theory in analogy with hamiltonian perturbation.

It was soon realized that in quantum mechanics perturbation theory does not have the difficulties which plagued the classical case and which were emphasized by H. Poincaré, namely the *small denominators problem*. This is due to the fact that in general no harmonics appear in the frequencies of an atom. Perturbation theory as developed by Born and Heisenberg in analogy with Hamiltonian dynamics is still at present at the basis of dynamics in quantum mechanics.

Pauli made use of this algebraic isomorphism to determine the energy spectrum of the hydrogen atom. As already mentioned this had a relevant role in the acceptance of Quantum Mechanics by the community of researchers in Physics.

The classical equation for the energy a particle with mass m attracted by a particle of mass M by the coulomb force is hamiltonian. The hamiltonian is in cartesian coordinates

$$H = \frac{p^2}{2\mu} - \frac{e}{r}, \quad r = |x| \tag{57}$$

where μ is the reduced mass and e the charge.

The *stationary* Schrödinger equation for the Hydrogen atom is

$$\hat{H}\phi(x) = E\phi(x) \quad \hat{H} \equiv -\frac{\hbar^2 \Delta}{2\mu} - \frac{e}{|x|} \tag{58}$$

The quantum hamiltonian is derived from the classical one using the quantization we have described. In the analysis that follows we don't pay attention to domain problems (most of the operators considered are unbounded); a more refined analysis will be given in chapter "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method".

Equation (58) had been solved by Schrödinger [21] providing the spectrum of the Hydrogen atom and the eigenfunctions. Following Pauli we use the correspondence between Poisson brackets and commutators to determine *the discrete part* of the spectrum of the operator (58).

In classical mechanics it is known the invariance of the Hamiltonian under spacial rotation implies that the angular momentum $L \equiv q \wedge p$ is conserved. Correspondingly in quantum mechanics the momentum $\hat{J} = \hat{x} \wedge \hat{p}$ is conserved (commutes with the Hamiltonian).

It is known in Hamiltonian mechanics that in the case of Coulomb (=Kepler) system there is a further vector which is constant of motion, the Runge-Lenz vector

$$R = \frac{p}{\mu} \wedge L - e^2 \frac{x}{|x|} \tag{59}$$

In hamiltonian mechanics the presence of this further integral of motion implies hat the energy can be made to depend on only one action variable; the value of the two other constants of motion are determined by the parameters of the theory. This leads to a particularly simple solution of the equation of motion.

Pauli investigated the commutation relations of the hermitian part of the operator which is obtained form R applying the rules of the quantum correspondence. This hermitian part is

$$\hat{M} = (2\mu)^{-1}\hat{p} \wedge \hat{J} - (2\mu)^{-1} J \wedge \hat{p} - e^2 \frac{x}{|x|} \tag{60}$$

By using formally Heisenberg's commutation relations one can see that $H,\ M_k,\ J_k$ satisfy (formally) the following commutation relations

$$\begin{aligned} &[H, \hat{J}_k] = 0, \qquad [H, \hat{M}_k] = 0 \qquad [\hat{J}_i, \hat{J}_k] = i\hbar\epsilon_{i,k,l}\hat{J}_l \\ &[\hat{J}_m, \hat{M}_k] = i\hbar\epsilon_{i,k,l}\hat{M}_l \qquad [\hat{M}_j, \hat{M}_k] = \frac{2\hbar}{i\mu}\epsilon_{j,k,l}\hat{J}_l H \end{aligned} \tag{61}$$

$$\hat{J}.\hat{M} = \hat{M}.\hat{J} = 0, \qquad (\hat{M})^2 - e^2 = \frac{2}{\mu}\hat{H}(\hat{J})^2 + \hbar^2 \tag{62}$$

Since $\hat{J}$ and $\hat{M}$ commute with the hamiltonian it is possible to restrict oneself to a representation in which the hamiltonian is diagonal and negative (bound states).

Consider a subspace in which the Hamiltonian has value $E < 0$. Introducing the operator $\tilde{M}_i = \frac{\mu}{2|E|}^{\frac{1}{2}}\hat{M}_i$ one verifies that the vector valued operators $\frac{\hat{J}\pm\tilde{M}}{2}$ commute and the components of each pair satisfy the commutation relations of angular momentum. Hence

$$(\hat{J} + \tilde{M})^2 = 4\hbar^2 j_1(j_1 + 1) \qquad (\hat{J} - \tilde{M})^2 = 4\hbar^2 j_2(j_2 + 1) \tag{63}$$

where j_1, j_2 may take integer or half-integer values. But $\hat{J}.\tilde{M} = \tilde{M}.\hat{J} = 0$ and therefore

$$(\hat{J} + \tilde{M})^2 = (\hat{J} - \tilde{M})^2 \tag{64}$$

and $j_1 = j_2 \equiv j$.

Assume for the moment that $2j$ can take any integer value. We derive from (62)

$$\frac{2e}{\mu}(\tilde{M}^2 + \hat{J})^2 + m\hbar^2 = m - e^2 \tag{65}$$

and then

$$(\tilde{M}^2 + \hat{J}^2 n + \hbar^2)4[(\tilde{M} + \hat{J})^2 + (\tilde{M} - \hat{J})^2] + \hbar^2 = \hbar^2(2j + 1)^2$$

It follows

$$E = -\frac{\mu e^2}{2\hbar^2}\frac{1}{(2j+1)^2} \tag{66}$$

The procedure we have followed shows also that the degeneracy of the levels is $(2j+1)^2$.

The commutation relations (61) and (62) are those of the generators (properly normalized) of the Lie algebra $sO(4)$. Since $SO(4)$ is semisimple and the representation of $sO(4)$ is irreducible the representation lifts to an irreducible representation of $SO(4)$.

This analysis can be done also in the Schrödinger representation but is more cumbersome (the operator $\tilde{M}$ is in the Schrödinger representation a second-order differential operator).

The work of Pauli, based on the newly established homeomorphism between commutators and Poisson brackets had a relevant role in the acceptance of the new Mechanics by the community of theoretical physicists.

10 Dirac's Theorem

For the description of dynamics the formulation given by Schrödinger and by Heisenberg depend heavily on analogies with hamiltonian dynamics. The systematic use made by von Neumann of a Hilbert space structure and of the structure of the algebra of operators acting on this space makes no longer necessary for the description of dynamics to choose in advance a representation (a special presentation of the Hilbert space). We shall see in chapter "Lecture 5: Automorphisms; Quantum Dynamics; Theorems of Wigner, Kadison, Segal; Continuity and Generators" that the structure of Quantum Mechanics as described by its axioms has in itself a natural definition of evolution (of dynamics), *independent of classical analogies*. Of course specific problems, that have a classical analogue, benefit form analogies with the hamiltonian formalism of classical dynamics. And then the best choice of representation is determined by the specific problem at hand.

While quantum dynamics can be obtained without reference to the isomorphism between commutators and Poisson brackets, the structural rigidity of the connection of dynamics in Quantum Mechanics with the hamiltonian structure Classical Dynamics is well described by the following theorem due to Dirac.

Recall that in Hamitonian Dynamics the vector field on the space of differentiable function on phase space that describes the evolution under the action of a hamiltonian H is a *derivation* (satisfies Leibnitz' rule and the Jacobi indentity).

An analogous structure exists for matrices. Let M_n be the algebra of rank n matrices; a linear operator δ on M_n is called *a derivation* if it satisfies Leibniz's rule for the product. It is a star-derivation if $\delta(a^*) = (\delta(a))^*$ for $a \in M_n$, where a^* is the Hermitian conjugate of a. We shall say that the derivation δ is an *inner derivation* if there exists $h \in M_n$ such that $\delta(a) = i[h, a]$.

Theorem (Dirac) *Every *-derivation of M_n is inner (and the matrix h can be chosen to be hermitian).*

We shall give in the course of these Lectures a proof of Dirac's theorem placing it in the more general setting of operator theory on infinite dimensional Hilbert spaces and C^*-algebras (Bratteli). We remark here that Dirac's theorem implies that every linear dynamics on M_n has the form

$$i\frac{da}{dt} = ha - ah \tag{67}$$

where the matrix h is hermitian and unique modulo addition of a term proportional to the identity $I \in M_n$. The solution of (67) is $a(t) = e^{-ith}a\, e^{ith}$. One can interpret Dirac's theorem by saying that (in finite dimensions) every quantum vector field is hamiltonian and every evolution is unitary.

References

1. Baccigaluppi, G., & Valentini, V. (2009). *Quantum theory at the crossroad*. Cambrigde: Cambrigde University Press.
2. Born, M. (1924). *Zeitshrift fur Physik*, *26*, 379–395.
3. Bohr, N., Kramers, H., & Slater, J. (1924). *Philosophical Magazine*, *47*, 785–802.
4. Born, M., & Jordan, P. (1952). *Zeitshrift fur Physik*, *34*, 858–888.
5. Born, M., Jordan, P., & Heisenberg, W. (1926). *Zeitshrift fur Physik*, *35*, 557–581.
6. Dirac, P. A. M. (1926). *Proceedings of the Royal Society A*, *112*, 661–676.
7. Dirac, P. A. M. (1927). *Proceedings of the Royal Society A*, *114*, 243–265.
8. The Principles of Quantum Mechanics Oxford University Press 198.
9. Eckart, P. (1926). *Physics Review*, *28*, 711–726.
10. Heisenberg, W. (1925). *Zeitshrift fur Physik*, *33*, 879–893.
11. Heisenberg, W., & Pauli, W. (1927). *Annalen der Physik*, *81*, 109–139.
12. Jordan, P. (1925). *Zeitshrift fur Physik*, *30*, 297–319.
13. Jordan, P. (1927). *Zeitshrift fur Physik*, *45*, 765–783.
14. Jordan, P. & Klein, O. (1927). *Zeitshrift fur Physik*, 75–94.
15. Jordan, P., & Pauli, W. (1928). *Zur Quantenelekrodynamic ladunsfreie Felder Zeits. fur Physik*, *47*, 151–173.
16. Jordan, P., & Wigner, P. (1928). Uber das Paulische Aequivalenzeverbol. *Zeitshrift fur Physik*, *47*, 631–655.
17. Kramers, (1924). *Nature*. *113*, 673–676.
18. Lademburg, R. (1921). *Zeitshrift fur Physik*, *4*, 451–468.
19. Pauli, W. (1926). *Letter to Jordan*.
20. Pauli, W. (1926). *Zeitshrift fur Physik*, *36*, 336–363.
21. Schrödinger (1926). *Annalen der Physik*, *81*, 109–139.
22. von Neumann, J. (1927). *Nachrichten von der Gesellschaft der Wissenschaften zu Gottingen*, 1–57.
23. van Velt. (1924). *Physics Review*, *24*, 330–365.

Lecture 3: Axioms, States, Observables, Measurement, Difficulties

1 Introduction

In "Lecture 1: Elements of the History of Quantum Mechanics I" and "Lecture 2: Elements of the History of Quantum Mechanics II" we have presented experimental results, theoretical analysis, models, analogies and insights that were accumulated for the description of physics at the scale of atoms. We described also the theoretical formulations that were developed in very few years and are the basis of a new theory for the dynamics of atomic systems: Quantum Mechanics.

Let us recall that *a mathematical model*, according to J. von Neumann, is a *mathematical construction* that, supplemented by a *verbal language of correspondence*, provides a coherent basis for the description of a large class of physical phenomena. A model is based on a combination of experimental evidence, theoretical intuitions and mathematical analysis. When the class of phenomena which are organized in a unified way by a model covers an entire field of physics one speaks of *a theory*.

A theory in general provides a different perception of what is relevant for the comprehension of an entire class of experimental results, indicates further question that should be meaningfully asked and experiments that are worth performing to deepen the understanding of nature (one refers often to this as a *paradigm*).

The passage from a *model* to a *theory* is also shaped by cultural background, by the versatility of the model to adapt to applications and also by pre-judgements. When a theory reaches acceptance by a majority of people working in the field any attempt to construct an alternative model is dismissed as "irrelevant".

Form this point of view Quantum Mechanics deserves the name of theory. It has changed our approach towards the physics at microscopic scale and the perception of the world at atomic scale and provides a unified physical and mathematical picture at this scale. It is at the basis of the new technology and it has stimulated the development of a relevant part of modern mathematics.

Quantum mechanics has become so familiar to people working in the physics at atomic scale that they tend give reality to all symbols of the theory, in spite of

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_3

the fact the theory itself does indicate that this is true. Quantum Mechanics is a theory that distinguishes clearly between the symbols that enter the description of the mathematical structure and the quantities that can be measured in an experiment. Still it is natural that people working in Quantum Mechanics (and teaching it) tend to speak e.g. of wave function as if it were a real object as real as planets in astronomy and waves in fluid dynamics and not a *complex probability amplitude* as suggested by Born.

As emphasized by N. Bohr this is due, at least in part, to the fact that the language with which we describe quantum mechanics, theory and experiments, is the same language we use to describe macroscopic objects. We have no direct experience of the word at atomic scale (N. Bohr went to the extreme of stating that classical physics is *necessary* to describe Quantum Mechanics).

This leads inevitably to ambiguities and misunderstandings. For example, in the mathematical formulation of solid state physics, one studies the *topological properties* of the wave function. From the point of view of the mathematics of Quantum Mechanics this is fully legitimate, but one should keep in mind that the wave function is not an observable.

The mathematical appeal and the success of quantum mechanics both in its theoretical aspects and in the applications and in technology has lead researchers in Quantum Mechanics to idealize the theory and to think that it is a perfect theory free of difficulties. This leads to regard as irrelevant any research on the foundations of the theory and to dismiss alternative theories as "mental constructions".

Still there are conceptual problems in Quantum Mechanics, due in part to the fact that language which is used is borrowed from classical physics and in part from the difficulty one has to reconcile the *intrinsic probabilistic character* of Quantum Mechanics with the deterministic features which we associate to macroscopic physical phenomena. Quantum Mechanics is a theory which is robust and self-consistent from the mathematical view-point, very effective in its applications but *not conceptually complete*.

2 The Axioms of Quantum Mechanics

We have stated that the mathematical construction of a model (and of a theory) requires

- the statement of axioms (or postulates), derived in general from phenomenology, from partially successful previous theories and from historical and cultural background.
- the deduction from the axioms of some non-trivial mathematical constructions, typically under the form of theorems and equations.
- the choice of a *connecting language* which associates measurable quantities to the mathematical structures. This description is made in every-day language (in the language of Classical Physics) and links the theory to experimental data.

For example, in Classical Physics the mathematical constructions consists of the variational principles, the equations of hamiltonian or lagrangian dynamics for material points, the equations of the dynamics of the continuum and the equations of Electromagnetism as formulated by Maxwell.

The (scientific) verbal language, i.e. the correspondence between mathematical entities and quantities that can be measured, *is given for granted* in Classical Physics. This common agreement is a result a common cultural background and of centuries of "experience"; this is possible because we have *a daily experience of the classical world*. No one doubts of the objective meaning of terms such as *measure of a velocity* or *measure of a magnetic field* and we regard the result of a measurement as independent in principle from the apparatus used and from the observer.

For the phenomena at atomic scale as described by Quantum Mechanics this "objectivity" fails and, as we will see, *the very concept of* "*measurement*" *can become problematic*. One can try to overcome this problem by stating that macroscopic objects, *such as a measurement apparatus*, must be regarded as classical objects, obeying the laws of classical physics. But this would imply a division of the world of experiments in two separate *incompatible* parts, and it would be difficult to say where the cut should be placed. It would in particular be difficult to understand the extraordinary success of Quantum Mechanics in describing *macroscopic* phenomena (e.g.superconductivity) that have no classical description.

Several reasons have been proposed why in (most) macroscopic bodies *one does not perceive* the typical structures of Quantum Mechanics. In particular, why it is difficult to perceive (outside specialized laboratories) the *superposition principle* (the sum of two wave functions does not correspond to the sum of probabilities) and the *entanglement* (the wave function of a two-component system is not in general the product of wave functions of the components and this prevents extracting information about the components).

Many efforts have been made to solve the *measurement problem* i.e. the detailed description in Quantum Mechanics of the process of measurement (which is a link between the microscopic world to the macroscopic one). Some of these attempts have led to a better understanding of the conceptual structure of the formalism and of its interpretation, but a satisfactory answer is still to be found.

From the empirical point of view Quantum Mechanics has had outstanding success in organizing, describing and also in some cases *predicting* results of experiments in its range of validity, namely the (non-relativistic) physics of atoms and molecules and their aggregates. The formalism of Quantum Mechanics has also contributed greatly to the development of modern mathematics. Still it must be remembered that the theory has its own *range of validity*, in particular that it is non-relativistic and it is not applicable to phenomena which occur at very high energies.

In this Lecture and in the following ones we shall describe the mathematical structure of Quantum Mechanics and its verbal language. It is convenient to begin by defining state, observable, evolution and describing the dynamics. In this context we shall state existence and uniqueness theorems. The formalism we present in these Lectures is at the root of the description of such diverse processes as the diffraction

of electrons by a crystal, the discrete structure of emission spectra of an atom, the diamagnetic properties of a crystal and the scattering of two particles.

We begin our analysis with a mathematical description of states and observables.

3 States and Observables

In general terms, a state of a system is the result of a preparation procedure. In order to construct a model, one must think of an idealized procedure that results in a well defined state.

One may want in any case that the definition of a state be such that when a system is in state σ any experiment aimed at find out whether a property **a** is satisfied gives the same result if performed by different observers. One may want also that the measurement can be done, at least in principle, without altering the state of the system. Moreover one may want that the definition of a state be such that one can describe the state of a composite system by describing separately the state of its parts.

This is true in Classical Physics. To give an example, the "state of the solar system" can be described by giving the position and motion of each the planets. And the determination of this state does not depend on the observer. An idealization of this setting has led in Hamiltonan Mechanics to consider as elementary objects (pure states) the points in *phase space* $\mathcal{M}$ of a dynamical system. An observable is then *by definition* characterized *by its value on each state*: observables are *represented* by (real) continuous functions on phase space.

In Hamiltonian Mechanics one convenes that if $m \in \mathcal{M}$ and if f is continuous the number $f(m)$ represents the result of the measurement of the observable described by f when performed on a system in the state described by m and that this result is independent of the observer. It is given for granted that such measurement can always be performed and that it gives the same result independently of the instrument used.

From a mathematical point of view therefore the pure states in Hamiltonian Mechanics are elements of the dual of the space of continuous functions (with its natural topology). The duality is given by $\{m, f\} \to f(m)$. The *duality* in Hamiltonian dynamics between states and observables permits to have a dual version of the laws of dynamics, as Newtonian laws and as Koopman formalism.

In Classical Statistical Mechanics one introduces also states represented by *densities* or more precisely by positive measures μ that are absolutely continuous with respect to Lebesgue measure with density (Radon-Nikodym derivative) ρ (i.e. $\mu(dm) = \rho(m)dm$). In this representation pure states are represented by Dirac measures.

These states are linear continuous positive functionals on essentially bounded functions. As a consequence one can include in the theory a larger class of observables i.e. function in $L^1(\mathcal{M})$. The correspondence is now given by $\{\rho, f\} \to \int f(m)\rho(m)dm \equiv (f, \rho)$. The states that we have added are not *pure (indecomposable) states* since any positive function ρ in L^1 can be written as $\rho = \rho_1 + \rho_2$ with $\rho_k \in L^1,\ k = 1, 2$. Notice that the correspondence

$$\rho,\ f \to (f,\rho) = \int f(m)\rho(m)dm$$

is continuous in f for the L^∞ topology and continuous in ρ for the L^1 topology.

As can be seen from this brief reminder, the definition of *pure state* in Classical Mechanics is linked to the possibility of considering continuous functions as *observables*. Moreover dynamics in Hamiltonian Mechanics is described by means of differential equations for functions in phase space; this formulation of dynamics is therefore adapted to a subclass of observables, those which are *differentiable* (in order to define Poisson Brackets).

When one tries to develop Quantum Mechanics and its dynamics keeping some analogy with Hamiltonian Mechanics the first problem one faces is that in Quantum Mechanics an equivalent of phase space *does not exist* and therefore it is difficult to decide a-priori how to describe a pure state and characterize an observable.

The two basic formulations of Quantum Mechanics (we shall call them *Schrödinger Quantum Mechanics* and *Heisenberg Quantum Mechanics*) differ by the choice of which objects are more fundamental: states or observables. This corresponds in Hamiltonian Dynamics to give more emphasis to phase space or to function on phase space.

4 Schrödinger's Quantum Mechanics

In the formulation of Quantum Mechanics due to Schrödinger the primitive elements are the (pure) states which are represented by (normalized) vectors in the separable complex Hilbert space $\mathcal{H} \equiv L^2(R^d)$ where d is the number of degrees of freedom of the corresponding classical system. In what follows all vectors are understood as having Hilbert norm one.

This interpretation makes explicit use of the analogy between $|\phi(x)|^2,\ x \in R^d$ and the classical distribution $\rho(x)$ which is suggested by Born's law. In particular both are positive and belong to $L^1(R^d)$.

In Schrödinger's formulation the observables are *a derived (dual) structure*; they are *represented* by operators on $\mathcal{H}$. If $\mathcal{H}$ is concretely represented as $L^2(R^d)$ the observables are represented by operations on functions, typically by multiplication by another function and by differential operators. In view of the analysis made by de Broglie the operator $-i\frac{\partial}{\partial x_m}$ should be identified with the momentum p_m. Since Fourier transform provides an isometry between L^2 spaces, also functions on momentum space represent observables.

One can expect to be able to give a concrete experimental apparatus and a measurement procedure for the measurement of *each of these observables*.

In the Schrödinger representation Dynamics is introduced through the Schrödinger equation; it is described by one parameter group of unitary operators $U(t)$ defined by $\psi(t) = U(t)\psi(0)$ if $\psi(t)$ is the solution at time t of the Schrödinger equation with initial datum $\psi(0)$. The dynamics on the observables is defined by duality

$$\psi(t) = U(t)\psi(0) \Rightarrow A(t) = U^*(t)A(0)U(t) \tag{1}$$

so that $(\psi(t), A(0)\psi(t)) = (\psi(0), A(t)\psi(0))$.

Notice that the set of operators which are sum of a function of position and of a function of momentum *is not invariant under this dual flow* (unless it is free flow). In order to be able to describe the dynamics *one is forced to increase the number of observables* to include a set which is invariant under the dual flow for interacting systems.

According to Born's rule the average of the result of the measurement of the observable **a** represented by the operator A in the state represented by $\psi \in \mathcal{H}$ is $(\psi, A\psi)$. If we require that the average value of any observable in any state be a real number, we should restrict ourselves to hermitian operators. A more detailed analysis leads to consider only self-adjoint operators in order to have a functional calculus; in "Lecture 5: Automorphisms; Quantum Dynamics; Theorems of Wigner, Kadison, Segal; Continuity and Generators" we shall briefly recall some elements on the theory of linear operators on a separable Hilbert space.

This class of operators is invariant under the dual flow. One is therefore led to assume that the observables are in *one-to-one correspondence* with self-adjoint operators, in spite of the fact that for a generic self-adjoint operator A one *may be not able* to exibit the experimental apparatus and the experimental procedure which may be used to *measure* the observable represented by A.

It should be noted that there may be *superselection rules* which forbid transitions between suitable classes of states. A typical superselection rule is given by charge conservation. Consequently not all self-adjoint operators are acceptable as being associated to observables but only the ones that are compatible with the superselection rules (e.g. they commute with the operator which represents the charge).

5 The Quantization Problem

We will see that for self-adjoint operators one can define a *functional calculus* and therefore to state the axioms it is possible to restrict the analysis to *bounded self-adjoint operators*. These operators generate an algebra over the complex field; if no super-selection rule is present, the resulting algebra is $\mathcal{B}(\mathcal{H})$, the algebra of all bounded operators on the Hilbert space $\mathcal{H}$.

Contrary to the case of classical mechanics, $\mathcal{B}(\mathcal{H})$ *is not a commutative algebra*. This leads to a difficulty in associating an operator to classical observable (a function in phase space). As a simple example, consider a prescription to associate an operator to the function qp. If, according to the de Broglie (and Schrödinger) prescription, we associate p with $\hat{p} \equiv -i\frac{d}{dx}$ and q with $\hat{x}$ (multiplication by x), one has an ambiguity, since whereas $pq = qp$ one has $\hat{p}\hat{x} \neq \hat{x}\hat{p}$.

A natural choice would the symmetric product (Jordan product) $\frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})$. But for higher order polynomials the choice is not easy since the Jordan product is

not associative. For functions which are not analytic the choice is problematic. This leads to the *quantization problem*.

Several different prescriptions have been given. For the functions $G(q)p^M$ M. Born and P. Jordan favored the mean over all permutations i.e

$$\frac{1}{M+1}\sum_{k=0}^{M}\hat{p}^k G(\hat{q})\hat{p}^{M-k} \tag{2}$$

Notice that $G(\hat{q})$ is well defined because in Schrödinger's representation it is the operator of multiplication by the function $G(x)$.

Another quantization that we shall discuss, *Weyl quantization*, introduces a twisted (symplectic) Fourier transform using the unitary operators $e^{ia\hat{q}}$ and $e^{ib\hat{p}}$. Weyl's quantization is particularly useful when studying operators which are projections on a subspace of the Hilbert space, e.g. projections on a part of the spectrum of a Hamiltonian (this is common when one studies space-periodic systems). These operators *cannot be expressed* in terms of polynomials in the elementary observables $\hat{q}_k\hat{p}_h$.

Other quantizations will be described in the following Lectures together with a brief analysis of the quantization procedure. For the moment we shall neglect this problem, and *assume* that every observable which has finite expectation value in every state is represented by a bounded self-adjoint operator.

We have remarked that according to Born's rule $\int_{\Omega\subset R^3}|\phi(x)|^2dx$ represents the probability that, performing a position measurement of a particle in the state described by $\phi(x)$, the outcome be that the particle in localized in the region Ω. This implies that if the observable **a** is represented by the function $A(x)$ in configuration space then

$$(\phi, A\phi) \equiv \int |\phi(x)|^2 A(x)dx$$

is the average of the results one obtains if one measures the outcomes of a measurement the observable **a**.

In the same way, if the observable **b** is represented by the function $B(p)$ in momentum space, then

$$(\hat{\phi}, B\hat{\phi}) \equiv \int |\hat{\phi}|^2(p)|B(p)dp$$

(where $\hat{\phi}(p)$ is the Fourier transform of $\phi(x)$) is the average of the results of the measurements of the observable **b**. For other observables the explicit form of the expectation value depends on the quantization chosen.

By polarization one obtains the value of $(\phi, A\psi)$ for any pair $\phi,\ \psi \in \mathcal{H}$.

6 Heisenberg's Quantum Mechanics

In the formulation given by Born, Jordan and Heisenberg (matrix mechanics) the primitive elements are the observables, represented by infinite hermitian matrices, i.e. linear symmetric operators on a separable Hilbert space. For example in the Heisenberg representation of atomic systems the off-diagonal elements of some of these matrices determine the probability of transition from an atomic state u_n to another state u_m under the influence of a external field or under *spontaneous decay*. The full matrix (i.e. also the diagonal elements) can be recovered by using different external fields.

We shall soon make more precise these statements (in particular that the operators must be self-adjoint).

The structure of the states plays a lesser role in this formulation of Quantum Mechanics. Mathematically the states belong to the dual of the algebra $\mathcal{B}(\mathcal{H})$ which represents the observables and are therefore characterized by the value they give for the expected value of the observables. If it is required that the duality between the observables and the states be continuous in the topology of $B(\mathcal{H})$ the states can be represented by *density matrices* i.e. positive operators of trace class with trace equal to one. Denoting by σ a generic element of this space and by A a generic element of $\mathcal{B}(\mathcal{H})$ the duality is expressed by $\sigma,\ A \to Tr(\sigma\ A)$ (the symbol Tr indicates the trace). In this correspondence pure states are in one to one correspondence with projection operators on one-dimensional subspaces.

Interference effects, one of the distinctive features of Quantum Mechanics, are not easy to describe in the Heisenberg formalism and a "concrete" analysis of entanglement and interference is difficult without reference to the Schrödinger representation.

The formalism of Heisenberg can be extended to consider as observables elements of more general structures (e.g. of C^*-algebra) and correspondingly as states the elements in the dual. We shall discuss in these Lectures some of these generalizations. Of course one can rely on the fact (we shall discuss it in the next Lectures) that the observables, when considered as a normed algebra, admit a faithful representation as bounded self-adjoint operators on a Hilbert space $\mathcal{H}$.

7 On the Equivalence

We have already seen that the representations of Schrödinger and Heisenberg are *equivalent in the mathematical sense* and they correspond *mathematically* to dual structures. The mathematical instruments used are different, mostly algebraic in Heisenberg's presentation, mostly function-theoretical in Schrödinger's. The bridge between the two formulations of is given by *Born's rule*. In this rule the states and the observables play an equally important and symmetric role.

Schrödinger's formulation has special properties that come from the fact that representating of the Hilbert space as space functions permits the use of the very powerful structure of Functional Analysis. Of course *all separable Hilbert spaces*

are isomorphic (in their structure as Hilbert space) but this special presentation tends to introduce elements which at the same time make the Schrödinger representation *more visualizable in space* (one can draw on the blackboard the "shape of the wave function") and more exposed to misunderstanding since the complex-valued wave function are *representatives* of abstract objects (probability *amplitudes*). The only "physical" meaning that can be attributed to them is in connection with Born's rule, i.e. as instruments to determine probabilities of real events.This misunderstanding is at the root of the conceptual problems in Quantum Mechanics.

The possibility of this *fake visualization* is also at the origin of the preference that is given in most textbooks (and research papers on Non Relativistic Quantum Mechanics) to the Schrödinger representation. The possibility to *visualize* is of great help for the imagination of a *macroscopic* researcher e.g. in the description of the microscopic structure of a crystal or in the search of new aspects of the many-body problem.

This visualization is also at the root of the use of geometric and topological methods in Quantum Mechanics, considering the wave function as a *geometrical object* and not as a *probability distribution*. And in this visualization one speaks of *topological properties* of the wave function and associates them with them measurable physical effects (e.g. whether a material is a good conductor of electricity).

We shall see instances of this in the following Lectures when we shall discuss the *Berry phase* and *the topological properties* of the wave functions for systems subjected to a periodic potential e.g. a crystal (and in general the role of topology in Solid State Physics). In fact we have already seen an example of this visualization in the use of the word *spin* in some way connected with rotations (but experimentally seen through the hyperfine splitting of the atomic spectra).

8 The Axioms

After these preliminaries we can state the axioms of Quantum Mechanics. In choosing the order of the axioms, we shall follow the point of view of Schrödinger. The first two axioms refer to the mathematical structure, the last three refer to the connection with measurable quantities.

Axiom 1 Pure states are represented by unit vectors in a separable Hilbert space $\mathcal{H}$. Vectors that differ by a phase represent the same pure state. We assume here that all Hilbert space vectors represent states, i.e. there are no "superselection rules". The collection of all vectors of norm one that differ only by a phase is called *ray*; therefore a pure state is represented by a ray.

An equivalent statement is: pure states are represented by one dimensional projection operators P_ϕ. ♣

A Hilbert space is a vector space and the *superposition principle* holds: if $\phi,\ \psi \in \mathcal{H}$ also $a\phi + b\psi \in \mathcal{H}$ where a, b are complex numbers. If ϕ and ψ do not belong to the same ray the state $a\phi + b\psi$ (properly normalized) represents *a pure state* that is *not a mixture* of the pure states represented by ϕ and ψ.

Also in Quantum mechanics one can introduce non-pure states, called statistical mixtures. They are represented by sums of projection operators

$$\sigma = \sum_n c_n P_{\phi_n} \quad c_n > 0, \qquad \sum_n c_n = 1 \tag{3}$$

The positive trace class operator with trace 1 defined by (3) are called *density matrices*. Their relation with the pure states *is the same as in classical mechanics*.

Contrary to what happens in Classical Mechanics *no pure state is dispersion free for all operators*; this is due to the fact that their dual, the algebra $B(\mathcal{H})$, is not commutative.

Recall that the dispersion of a state σ relative to a (symmetric) operator A is defined as

$$\delta_\sigma(A) \equiv \sigma(A^2) - (\sigma(A))^2 \tag{4}$$

A state σ is *dispersion free relative to* A iff $\delta_\sigma(A) = 0$. For comparison notice that in Classical Mechanics, where the role of $\mathcal{B}(\mathcal{H})$ is taken by continuous functions, *all pure states (Dirac measures) are dispersion free with respect to every observable*.

The dual of the pure states, under the duality given by P_ϕ, $A \to Tr P_\phi A$, is $\mathcal{B}(\mathcal{H})$, the set of bounded closed operators $\mathcal{H}$.

Axiom 2 The observables in Quantum Mechanics are represented by the *self-adjoint operators* on a separable complex Hilbert space $\mathcal{H}$.

One can restrict attention to the algebra $\mathcal{B}(\mathcal{H})$, the algebra of bounded operators on the complex Hilbert space $\mathcal{H}$. We will come back in this Lectures to the properties of this algebra. ♣

Notice that if $A \in \mathcal{B}(\mathcal{H})$ the correspondence

$$A \to Tr(\sigma\, A) \tag{5}$$

defines a linear continuous functional on $\mathcal{B}(\mathcal{H})$. Therefore we could start, as in the Heisenberg point of view, with the definition of the observables as the *real part* of the algebra of all bounded operators on a Hilbert space $\mathcal{H}$ and consider the states as derived quantities by the duality (5).

We shall call *concrete* C^*-algebra a uniformly closed algebra of bounded self-adjoint operators in the Hilbert space $\mathcal{H}$. We shall call *Jordan algebra* an algebra of linear operators one a Banach space with the product rule $A \circ B = \frac{1}{2}(AB + BA)$ and we shall call *Jordan* C^**-algebra* (denoted by JC^*) a uniformly closed subalgebra of a C^*algebra endowed with a Jordan product. The observables are therefore represented by a realization (representation) of the self-adjoint elements of a Jordan algebra JC^*. It is interesting to notice that a JC^*-algebra is *characterized* by being the self-adjoint part of a C^*-algebra [4]. There is therefore no loss of generality in regarding the observables as (bounded) self-adjoint operators in a Hilbert space. The extensions to unbounded operators is more delicate; we postpone it to the Lecture in which we define self-adjoint operators.

In this interpretation one can prove that the states we have introduced are *normal* i.e. if an increasing filter A_α in $\mathcal{B}(\mathcal{H})$ has B as extremum, the filter $Tr(\sigma\ A_\alpha)$ has $Tr(\sigma\ B)$ has extremum.

Notice in general

$$\frac{(\phi, A\phi)}{|\phi|^2} + \frac{(\psi, A\psi)}{|\psi|^2} \neq \frac{(\phi+\psi), A(\phi+\psi)}{|\phi+\psi|}$$

this elementary fact goes under the name *superposition principle*; it is a purely quantum property. It follows that if we interpret $\phi(x)$ and $\psi(x)$ as *probability waves*, these waves show *interference*. Interference is well known e.g. in the theory of electromagnetic waves but in Quantum Mechanics it seem to go beyond common sense since they refer to probabilities amplitudes and not waves that carry energy-momentum. It points to the fact that the probability waves should not be confused with classical (Kolmogorov) probabilities.

The duality states-observables can be extended to more general frameworks, e.g. to the case in which the observables are represented by a the real elements of a Banach algebra $\mathcal{B}$ and the states are represented by its dual $\mathcal{B}^*$. We shall see an instance of this the course of these Lectures; in this case it is natural to define first the representatives of the observables, i.e. to take the Heisenberg point of view.

Axioms 1 and 2 describe the mathematical content of Quantum Mechanics. They were put in precise mathematical form by Jordan et al. [2].

We introduce now three axioms that represent the *verbal* part of the model, i.e. the rules to associate measurable quantities to the mathematical entities introduced in Axioms 1 and 2.

Axioms 3, 4 and 5 connect the mathematical formalism to the outcome of laboratory experiments.

Axiom 3 The *mean value* of the measurements of the observable a represented by the bounded symmetric operator A in the state represented by P_ϕ (the projection operator on the one dimensional subspace spanned by the vector ϕ) is given by

$$\langle A\rangle_\phi \equiv (\phi, A\phi) \equiv Tr(AP_\phi) \quad |\phi| = 1 \tag{6}$$

where the symbol Tr stands for trace, a function defined as usual for finite rank matrices and extended by sum convergence to infinite matrices.

In the case of a statistical mixture σ prescription (6) is replaced by

$$\langle A\rangle_\sigma \equiv Tr(\sigma\ A) \equiv \sum_n c_n Tr(P_{\phi_n} A) \tag{7}$$

◇

In Axiom 3, a generalization of Born's rule, nothing is said about the outcome of a *single measurement*. And we did not speak of the effect that has, on a state described by σ, the measurement of an observable a described by an operator A.

Strictly speaking, since Axioms 3, 4 and 5 are meant to connect the mathematical formalism with the results of experiments, one should add an analysis of the *preparation of states* in Quantum Mechanics. This is a foundational problem as is the analysis of the actual process of measurement. Very little theoretical and mathematical work has been devoted to this subject, although on the experimental side very accurate empirical procedures have been developed involving precise control of macroscopic apparatuses. This is a field in which more research would be needed.

Axiom 4 Let the operator A describe the observable a and assume that A is self-adjoint and has purely discrete simple spectrum: the eigenvalues are different from each other and the eigenfunctions ψ_i^A form a complete orthonormal basis. If one performs a measurement of the observable a in a state represented by a vector $\phi \in \mathcal{H}$, $|\phi|_2 = 1$ the *only outcome* can be one of the eigenvalues a_k of A. The *probability* of the outcome a_k is $p_k^A = |(\psi_k^A \cdot \phi)|^2$, where ψ_k^A is the eigenfunction of the operator A associated to the k^{th} eigenvalue. ◇

Notice that this statement is *compatible with Born's rule*. Indeed from Axiom 3 it follows that the average of the results of the measurements of a when the state is described by the vector ϕ is $(\phi, A\phi) \equiv Tr(P_\phi A)$. If the system is in the state described by the density matrix σ the average result will be $Tr(\sigma A)$.

For observables which are represented by operators with partly continuous spectrum the formulation of Axiom 4 is slightly more complicated; we don't detail here the obvious modifications.

Remark that Axiom 4 is probably too ambitious as formulated. Given a generic self-adjoint bounded operator A it is difficult *even in principle* to give a prescription for the construction of a measuring instrument which measures the observable associated to A. For example it is difficult to indicate the instrument that measures the observable associated to $\xi(\Omega)\hat{\xi}(\Sigma)\xi(\Omega)$ where $\xi(\Omega)$ is the operator of multiplication by the indicator function of the domain Ω in configuration space and $\hat{\xi}(\Sigma)$ is the multiplication by the indicator function of the domain Σ in Fourier space.

Notice that generally speaking in Quantum Mechanics the evolution of functions that depend at time t_1 only on positions or only on momenta does not give at time $t_2 > t_1$ functions that depend only on positions or on momenta. For *some systems* it can be proved that the algebra of functions each of which *at some time* depends only on position is dense in $\mathcal{B}(\mathcal{H})$ but this is not a general property.

Axiom 4 refers to the *possible results* of a measurement and to the *probability* with which they are obtained. There is no indication about the state of the system *after measurement*.

In Classical Mechanics it is *assumed* that, at least in principle, it is possible to perform measurements on a system *without altering its state*. In other words, the measurement gives information about the state in which the system was *before the measurement*. In Quantum Mechanics *this is not possible*. The interaction with the measuring apparatus *alters the state of the system in a way that cannot in general be predicted*.

Axiom 5 Let a_i be a non degenerate eigenvalue, with eigenfunction ψ_i^A, of the operator A associated to the observable a. If the measurement of a has given a_i as result, *immediately after the measurement* the state of the system is described by the vector ψ_i^A. ◇

The condition *immediately after*, although imprecise, takes into account the fact that the operator A may not commute with the hamiltonian that represents time evolution and therefore the eigenstates of A may not be invariant in time. Since the evolution under the Schödinger equation is a continuous process in time, this effect is negligible if the time elapsed between two measurements is negligible.

We notice that Axiom 5 is needed to give objective meaning to the measurement process. To see this, consider measurement made at times $t_0,\ t_0+\epsilon,\ t_0+2\epsilon$; the parameter ϵ is chosen so small that we can neglect the dynamics in the interval $[t, t+2\epsilon]$. At times t_0 and $t_0+2\epsilon$ we measure A, at time $t_0+\epsilon$ we measure B, with $AB - BA \neq 0$ (and both have discrete simple spectrum).

Let $\{a_n\}, \{\phi_n\}$ be the eigenvalues and eigenvectors of A, $\{b_n\}, \{\psi_n\}$ those of B. Suppose that at time t_0 we measure the observable described by A. Let the result be a_1.

According to postulate (4) the state of the system *immediately after the first measurement* is represented by ϕ_1. If it were not so, a measurement of A after a further negligible amount of time would not give the result a_1 with probability one, in spite of the fact that the state has not (essentially) changed. This would imply that measurements does *not codify an objective property* of the system *immediately after measurement*. We assume therefore the validity of Axiom 5.

But Axiom 5 has far-reaching consequences since the algebra of the observables is non abelian in Quantum Mechanics.

At time $t_1 = t_0 + \epsilon$ with ϵ "arbitrary small" we perform a second measurement now of the observable b associated to the operator B which does not commute with A. According to the rules of Quantum Mechanics we obtain the result b_k with probability $|(\phi_1, \psi_k)|^2$; if we obtain b_k as a result we conclude, by Axiom 5, that the system immediately after the new measurement is in state ψ_k.

Now we perform again a measurement of a at time $t_2 = t_1 + \epsilon$; the result will be $a_h,\ h \neq 1$ with probability $|(\psi_k, \phi_h)|^2$.

Consider all possible outcomes that we had at time t_1 (e.g. if we do not read the results, or even we don't know that a measurement of b has been performed). It is easy to see that if B and A do not commute, in general $\sum_i |(\psi_k, \phi_1)|^2 < 1$. Therefore now the system has a finite probability to be in a state *different from* ϕ_1.

This implies that it is impossible to determine (even approximately) the state of the system *if one does not have complete control of the environment*. Indeed we reached the following conclusion: the measurement of the observable a performed at time $t_0+2\epsilon$ gives a result which is different from the result obtained at time t_0 if in the meanwhile a measurement of b has taken place, *even if we are unaware of this measurement*. If the commutator $[A, B]$ is relatively large, the second measurement of a may give a result very different from the first.

Axiom 5 gives objective meaning to a result of an experiment but it does not guarantee that at later times this information is kept. Moreover it states that it may be impossible to construct a state that gives definite results at time t for the observables a and b if the corresponding operators do not commute. One should keep in mind that the operator x and $U_t(x)$ (where U_t stands for time evolution for a time $t > 0$) *do not commute* even for free motion.

For comparison notice that in Classical Mechanics all observables (i.e. all functions on phase space) take *at any given time* a definite value on all pure states. Moreover one can conceive measurements on a pure state which do not alter the state (non-demolishing measurements).

In Quantum Mechanics some systems admit the existence of observables whose representative (operators) *commute with the representative of all other observables*. One speaks then of *superselection rules*. A typical observable that gives rise to a superselection rule is the charge. We do not discuss further this topic, but we remark that the algebra of operators that give superselection rules can be itself non-commutative.

9 Conceptual Problems

We comment briefly on Axioms 3, 4 and 5 to see the conceptual problems that arise. *We shall discuss the Schrödinger version of Quantum Mechanics*.

Axiom 5 is the *reduction postulate*. It implies that, whatever is the state of the system before the measurement, *as a the result of the measurement* of an observable a represented by an operator A with non degenerate point spectrum the system will be *projected* in one of the eigenstates ϕ_k of A.

We leave to the reader the modification to this statement that must be made if there is some degeneracy or if the spectrum of A is partly continuous.

The *reduction postulate* is at the root of the conceptual difficulties that follow from Axioms 3, 4 and 5 in Schrödinger's formulation. These difficulties disappear if one accepts that *Quantum Mechanics is an intrinsically probabilistic theory*.

We shall see in "Lecture 5: Automorphisms; Quantum Dynamics; Theorems of Wigner, Kadison, Segal; Continuity and Generators" that Quantum Dynamics gives rise to continuous and *deterministic* variation of the states, as in classical Mechanics. On the other hand the reduction of the wave function as presented here is *instantaneous* and essentially a random process.

To reconcile the two statements one may try to regard the reduction of the wave function as an *effective phenomenon*, governed by the deterministic laws of Quantum Mechanics and which takes place in a fraction of time *very small* compared to the variation time of the macroscopic physical system that enter the measurement process. In this way the reduction would be seen as *instantaneous for all practical purposes* and the probabilistic rule would be attributed to the complexity of the interaction between the observed object (microscopic) and the measurement apparatus (macroscopic and very unstable).

In "Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories" we will discuss the role of decoherence and show there that even this *mild interpretation* is not tenable. Quantum Mechanics *is an intrinsically probabilistic theory*. What could be achieved by resorting to decoherence, under favorable circumstances and within suitable approximations, is to reduce *the intrinsic probabilistic aspects of quantum mechanics* to classical probability if one considers *a macroscopic observer*.

The role of the reduction principle is clearly spelled out in the analysis of the emission of an α particle by a radioactive source as observed in a Wilson cloud chamber. This is an apparatus filled by a supersaturated vapor (usually etilic alchol) that may undergo a phase transition (producing droplets) under a very small perturbation. What one sees is ionization *tracks* that are rectilinear or, in presence of a magnetic field, curved as one expects of a charged particle in the magnetic field. Only one track is seen in every decay event; the direction of the track varies randomly with a spherically invariant distribution.

Mott [3] observed that it is difficult to understand how a spherical wave should produce rectilinear tracks; one rather expects that ionization be produced at random. We will see that the presence of tracks can be explained with the axioms of Quantum Mechanics *including the projection postulate*.

It is sufficient to use the Schrödinger equation and pay attention to the *the initial conditions* and to the following facts: the interaction is of very short range, the time of interaction is very small and the frequencies of the α wave are very large compared to the binding energy of the atoms to be ionized.

It is instructive to have a quantum-mechanical description of the process, in order to see the role of the reduction postulate. In Quantum Mechanics the initial state resulting from the radioactive emission of an α-wave, is described as a spherical wave concentrated in a narrow spherical shell with center O (the emission point) with high radial momentum directed opposite to O. It is convenient to present it in the form

$$\phi_0(|x|) = F_0^{\hbar}(|x|) \int_{S^2} d\omega e^{i \frac{M v_0(\zeta(\omega),x)}{\hbar}} \tag{8}$$

where ω is the angular variable on the unit sphere and $\zeta(\omega)$ is the unit normal to the sphere at the point ω pointing outwards. The function $F_0^{\hbar}(\rho)$ is supported in a neighborhood of the origin of linear size $\sqrt{\hbar}$ (in natural units, so that $\sqrt{\hbar}$ is very small).

Under free propagation (i.e. using the free Shrödinger equation) at a later time $T > 0$ the wave function can be seen, using stationary and non stationary phase theorems and under the assumption that $M v_0$ is very large, to have the form

$$\phi_T(|x|) = F_T^{\hbar}(|x|) \int_{S^2} d\omega e^{i \frac{M v_0(\zeta(\omega),x)}{\hbar}} + R_T(x) \qquad x \in R^3 \tag{9}$$

where the function $F_T^{\hbar}(\rho)$ is supported in a neighborhood of $v_0 T$ of linear size $\sqrt{\hbar}$. The residual term $R_T(x)$ is of order $\sqrt{\hbar}$ and decreases fast in time (it is due to the

dispersive properties of the Schrödinger equation); we will neglect this term in the following analysis.

The α wave (9) interacts with the atoms in the cloud chamber producing ionization; the gas is supersaturated and this acts as enhancing mechanism to produce a *macroscopic* result. Hence the gas acts as *macroscopic apparatus*.

The initial state of the system is represented by a function in R^{3N+1} where N is the number of atoms that can be ionized. The initial wave can be seen as a *coherent superposition* of small fragments of the spherical corona (that we call wavelets) each localized in a region of linear size comparable to that of each atom. This description has important similarities with the description in geometrical optics of a high frequency spherical wave as a collection of light rays. There are of course differences, the α wave is a probability wave, the light wave is an energy-momentum wave. The equations are also different, although in the high momentum regime there is less difference.

As in geometric optics, *due to the properties of the initial wave function* the wave *before the interaction with the atoms* can be *mathematically* decomposed in small wavelets each localized in configuration space to order $\sqrt{\hbar}$ along a radial direction $\frac{x_0}{|x_0|}$, $x_0 \in R^3 - 0$ and localized in momentum space to order $\sqrt{\hbar}$ along the same direction and with very high average momentum. This is a *virtual* decomposition similar to the one makes when one speaks of light rays in wave optics; it is virtual since it is made now for *probability waves*.

In the presentation of the α-wave given in (9) the fragments we consider are obtained by multiplying the integrand by elements of a partition of the unit sphere (smoothed characteristic functions). Each element of the partition is a function with support of linear size $O(\sqrt{\hbar})$. Each fragment has initial data at time T

$$\phi_\xi(x, T) = \Phi_\xi(x, T) + R_T$$

$$\Phi_\xi(x, T) = \int_{S^2} \xi(\omega) e^{i \frac{M v_0(\zeta(\omega), x)}{\hbar}} d\omega \qquad x \in R^3 \quad \omega \in S^2 \tag{10}$$

where ξ is an element of the partition. R_T is a correction term with $|R_T|_{L^2} = O(\hbar)$ and $F_T(|x|)$ has support in a neighborhood of $v_0 T$ of linear dimension $O(\sqrt{\hbar})$.

It is not difficult to see that the solution with initial datum $\Phi_\xi(x, T)$ is localized together with its (quantum) Fourier transform in a domain of linear size $\hbar$. In the Lecture devoted to semiclassical analysis we shall see that under the evolution described by the free Schrödinger equation waves with such initial data remain sharply localized both in configuration space and in momentum space (the dispersion remains of order $\sqrt{\hbar}$ for a long time) and their barycenter follows a classical path. We shall call them *semiclassical wavelets*.

We stress that this decomposition before the interaction with the atoms *is a mathematical exercise*. Before the interaction the wavelets move coherently (the decomposition is artificial).

Since the interaction is local, under the assumption that the atoms are sufficiently separated from each other *only one of the semiclassical wavelets* interacts with each

atom. When the wavelet interacts with an atom (in its ground state) leading to ionization the final state of the wavelet gets intertwined (entangled) with the final state of the atom. Since this is orthogonal to the ground state the combined wave function of the wavelet and of the atom is orthogonal to the initial state. An operator which depends only on the degrees of freedom of the wavelet or of the atom does not have matrix elements between the initial and the final state. In this sense each outgoing wavelet is incoherent with respect to the remaining wavelets.

It is difficult to describe in full detail the interaction of a wavelet with an atom. This would require taking into account the structure of the atomic wave function. We choose to schematize this process of interaction as *inelastic* scattering of the wavelet against an atom with a loss of energy (the ionization energy); since the momentum of the incoming wavelet is very much larger that the momentum exchanged with the atom, the resulting final state is made of a free electron, an ionized atom and a wavelet with large momentum approximatively equal to the momentum of the incoming wavelet (by momentum conservation: the atom can be considered very massive and therefore at rest before and after the interaction).

Recall now that the initial wave as well as the wavelet in which it has been decomposed are *probability waves*. The result is a *virtual* final state after the interactions *but before the measurement* in which each probability wavelet is *entangled* with the probability wave function of an ion. We call this a *virtual pair*. By virtual we mean that, according to Quantum Mechanics, it represents the probability that the corresponding ion be produced *if the measurement is performed*.

According to quantum mechanics *when the experiment (of the position of the first droplet) is performed* only one of the ions is selected to interact with the nearby atoms and to trigger an amplification mechanism that leads to the production of a droplet (recall that the vapor is supersaturated). Notice the difference between *measurement* and *interaction*. It is only the measurement that *selects* one atom to be the origin of amplification process. The measurement process *is not described* by the Schrödinger equation. Understanding it is *the Measurement Problem* which has had so far no solution in spite of its relevance.

One may suppose that the corresponding wavelet remains in the description of the physical system and that its probability (L^2 norm) is now one. From the point of view of the wave the process of measurement is *non local* and unitary (probability preserving). The other probability wavelets *are no longer part of the description of the physical system*.

This shows explicitly the difference between *interaction* and *measurement*.

The selected wavelet (now it may be called α-ray or α-particle) continues on its (classical) path and may ionize further atoms that it encounters.This gives rise to a sequence of droplets (the track).

This analysis of Mott's problem should clarify also the role of the *entanglement* and of the *reduction postulate*. The role of entanglement is that to transfer to the α-wavelet the effect of the reduction postulate which in principle applies only to the wave function of the ion, since it is the position of the ion that "is measured" by the macroscopic devise.

10 Information-Theoretical Analysis of Born's Rule

We conclude this Lecture with a comment on postulate 3 (Born's rule). Combined with the projection postulate Born's rule says that when one knows that a measurement corresponding to an observable a associated to a symmetric operator A with discrete spectrum has taken place without knowing the result, the following rule applies. If the initial state is described by density matrix ρ_{in}, then the density matrix ρ_{fin} of the final state is given by

$$\rho_{fin} = \sum_i Tr(P_i \rho_{in} P_i) \tag{11}$$

where $A = \sum_i \lambda_i P_i$ is the spectral decomposition of the operator A. One has by definition $\sum_i P_i = I, \quad A = \sum_i \lambda_i P_i$.

This formulation does no longer require that initial state of the system be pure.

It is interesting to notice [1] that (11) *can be interpreted in information theory* as saying that ρ_{fin} represents *the most probable state* that one may have after a measurement of the observable a in the state described by ρ_{in}.

We clarify what this statement means.

According to von Neumann information is measured by relative entropy, and the most probable state is the state which has maximal entropy relative to the initial state. In fact, following Wiener, we consider the amount of information to be the negative of the quantity defined as entropy. We take therefore as measure of the relative information content of the quantum state σ from the quantum state ρ the negative of the relative entropy function [5]

$$D(\rho, \sigma) \equiv tr(\rho ln \rho - \rho ln \sigma) \tag{12}$$

The function D is non negative and can be considered a measure of the error if given state σ we believe it to be ρ. It is a *non symmetric* distance. The most probable state is by definition the state which minimizes the distance $D(\rho, \sigma)$ for all states σ which can be obtained in a measurement of an observable **a** (reference states).

If the observable **a** is represented by the operator $A = \sum \lambda_i P_i$ the reference states are the density matrices which belong to the set

$$\Sigma_A \equiv \{\sigma : [P_i, \sigma] = 0\} \quad \forall P_i \tag{13}$$

Equation (13) is equivalent to $[A, \sigma] = 0$.

We must minimize $D(\rho, \sigma)$ over Σ_A. This amounts to selecting the quantum state that is least distinguishable from the original state among all the states that satisfy the constraint of being produced by the measurement of **a**.

We consider only the case in which the Hilbert space is finite-dimensional. The same results are obtained if A is compact.

The set Σ_A is defined by a linear relation so it is a simplex. $D(\cdot,\cdot)$ is jointly convex in both arguments so that $(\cdot,\sigma)$ is convex for all σ. Due to the fact that the problem is finite-dimensional the following holds: if the function f is differentiable and strictly convex on a simplex, and the directional derivatives at a point b are all zero, then b is the *global* minimum of f.

We can parametrize Σ_A noticing that every element is of the form

$$\sigma = U\Lambda U^* \tag{14}$$

where Λ is a trace one matrix with positive entries and U is a unitary operator $U = \pi_i U_i$ where U_i is the identity on the range of $(I - P_i)$. This means that $[U, P_i] = 0 = [\Lambda, P_i]$. Therefore writing σ_i for σ restricted to the range of P_i we have for every function f on Σ_A

$$f(\sigma) = \oplus_i f(\sigma_i) \tag{15}$$

(i.e. functions act blockwise on Σ_A).

Consider first the variation along the directions parametrized by U. We look for the variation in the direction parametrized by one-parameter subgroups. Call L the generator. We compute then

$$\frac{d}{dt}\phi_t^* tr AU_{t=0} = \frac{d}{dt} tr(Ae^{tL})U_{t=0} = \sum_{i,j} A_{i,j} L U_{i,j} = tr(ALU) \tag{16}$$

where we have denoted by ϕ_t^* the adjoint action. In the same way one computes

$$\frac{d}{dt}\phi^* Tr(AUBU^*)_{t=0} = Tr(ALUBU^*) - tr(AUBU^*L) \tag{17}$$

It is easy to verify that

$$[L_i, P_j] = 0 \quad L_i P_j = \delta_{i,j} L_i \tag{18}$$

The derivatives take the form

$$\partial_{L_i} \sum_j tr(P_j \rho P_j U_j \; ln\lambda_j \; U_j ln\lambda_j \; U_j^*) = tr L_i [ln\sigma_i, \rho_i] \tag{19}$$

If σ and $\sum P_i \rho P_i$ can be diagonalized simultaneously the derivative vanish. This is also a necessary condition since the commutator $[ln\sigma_i, \rho_i]$ is traceless and L_k, iL_k span the all traceless matrices in the in the i^{th} block.

Consider next the variation with resect to Λ restricting to the case when σ and $\Sigma_i P_i \rho P_i$ can be simultaneously diagonalized. Let μ_k^σ and μ_k^σ be the eigenvalues of σ and of $\sum_i P_i \sigma P_i$.

If $\mu_i^\rho \neq 0$ and $\mu_i^\sigma = 0$ one has $D(\rho, \sigma) = \infty$ so that this cannot be a minimum. One has

$$\partial_{\lambda_k^\sigma} - \partial_{\lambda_k^\mu} \sum_m \lambda_m^\rho \ln \lambda_m^\sigma = 0 \tag{20}$$

which implies that

$$\frac{\lambda_k^\rho}{\lambda_k^\mu} \tag{21}$$

is independent of k. So the ratio of the eigenvalues of σ and of $\sum_i P\sigma_i P$ is fixed. Since they are both of trace one, they coincide.

It follows that the state $\sigma = \sum P_i \rho P_i$ is the unique minimum of the relative entropy i.e. it is the *unique* state that is least distinguishable form the original state among all state which are compatible with the observation of the observable a.

References

1. Hellmann, F., et al. arXiv:1407.7766 [quant-ph].
2. Jordan, P., von Neumann, J., & Wigner, E. (1934). *Annals of Mathematics*, *35*, 29–45.
3. Mott, M. (1929). *Proceedings of the Royal Society*, *124*, 375–380.
4. Stormer, E. (1965). *Transactions of the AMS*, *120*, 438–447.
5. Vedral, V. (2002). *Reviews of Modern Physics*, *74*, 197–212.

Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories

Since a pure state is represented by a vector in a Hilbert space (or rater by a ray of vectors) entanglement with respect to a division of a system in two separate parts is a natural property of Quantum Mechanics. Consider a Hilbert space $\mathcal{H}$. Assume that there has a "natural" description of $\mathcal{H}$ as $\mathcal{H} \equiv \mathcal{H}_1 \times \mathcal{H}_2$ e.g. there are two particles in the system and $\mathcal{H}_k,\ k = 1, 2$ is the space used to describe particle $k,\ k = 1, 2$.

If the initial system is prepared in a state represented by a vector $\Phi \equiv \phi_1 \times \phi_2,\ \phi_k \in \mathcal{H}_k$ and if the hamiltonian H cannot be written as $H = H_1 \times I + I \times H_2$ (i.e. if there is a coupling between the two systems) the system after time t cannot in general be described by a product vector but rather is described by a sum

$$\Phi(t) = \sum_i c_i \phi_{1,i}(t)\phi_{2,i}(t),\ \phi_{k.i} \in \mathcal{H}_k \tag{1}$$

Notice that in (1) the $\{\phi_{n,i}\},\ n = 1, 2$ are not necessarily an orthonormal set.

We express this situation by saying that the final state is *entangled* (with respect to the decomposition chosen). This property, that we may also call *coherence*, is an advantage e.g. in the theory of quantum computation but is unwelcome when we want to describe the motion of a particle α interacting with an environment made of very may very (light) particles of type β. But *we expect* that when the number N of particles of type β is very large it is no longer possible to perceive the entanglement with the wave function of particle α with *every single particle*.

1 Decoherence. I

We rather expect to be able to describe the motion of the particle α, at least approximately, with an *effective equation* that describes the result of the collective action of the particles β. We expect this to be true in particular if the interaction is

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_4

very weak (so that only the collective action can be perceived) and if particles β have a mass m_β much smaller then the mass m_α of particle α; in this case the momentum transfer is very small in each collision. We also expect that this approximation is valid if the collision time is very short (compared to the time in we monitor the state of particle α) and therefore the collisions can be regarded as independent. We neglect also re-collisions and collisions among the β particles. Notice that these assumptions are used to prove the validity of the Boltzmann equation in the description the interaction of a classical particle with a thermal bath of lighter particles.

This is the setting of the theory of *decoherence*. One expects that this theory has relevance in the theory of measurement performed using *macroscopic apparatuses;* we may expect to find approximate *effective* equations which can be used to describe the evolution of measurable quantities.

Model of this type have been proposed, with different degree of success. Significant contribution have been made by Blanchard and Olkiewicz [3], Sewell [19]. Hepp [13], van Kampen [18]. The first two Authors use operator-algebras techniques, the last two use the Schrödinger formalism.

To a similar circle of ideas belong all the theories that attribute to decoherence to the impossibility of observing the superposition principle when one considers macroscopic systems (e.g. Schrödinger's cat). We shall discuss briefly in this lecture a laboratory experiment done in Paris by the group of Haroche. This experiment indicates how decoherence may be measured and indicates also that, even in a controlled environment of a laboratory, coherence is not seen after a very short time.

It must be noticed again that the wave functions are defined in the configuration space of the total system, including the environment and the *superposition principle holds in this space*.

Coherence should be regarded as a consequence of a continuous process of correlations between two or more quantum systems. Realistic quantum systems are never fully isolated. When a quantum systems which is described, if isolated, in a Hilbert space of dimension n interacts with another quantum system (environment or measuring apparatus) which is described, if isolated, in a Hilbert space of dimension N with $N \gg n$ the waves functions of the two system get *entangled*. This implies that we can no longer describe independently the status of the two system.

The theory is described in a Hilbert space of very large dimensions $N+n$ and it is unrealistic to expect to be able to keep track of the dynamics of the small system in such a huge setting. It is hopeless to try to give the exact solution of the Schrödinger equation when $N+n$ is very large. Decoherence is an attempt to describe the evolution in a space of a limited number of dimensions (hopefully n). Decoherence must be regarded not as a theory but rather *as a set of prescription with a given range of validity*.

When one measures a system S with n degrees of freedom (where n is very small) with a macroscopic apparatus which has N degrees of freedom (N very large) on can only hope, *and seldom prove*, that the Hilbert space of the entire system (measuring apparatus + small system) con be written with sufficient accuracy as

tensor product of two spaces, of dimensions n and N and that the smaller space carries all the information on the measurement. One expects that the space of dimension n be spanned by the eigenstate of the operator associated to the observable that is measured.

All the degrees of freedom of the measuring apparatus enter a complete description of the process; we can select n of them to play an effective role. One can *effectively* describe the process of measurement with an operator that has small matrix elements between the n-dimensional space and the one of dimension N. The *effective* operators of the small space are called *stable* because they are not much perturbed by the real process that takes place in the huge Hilbert space. It is expected (but seldom proved) that the matrix elements of the evolution operator that corresponds to the process of measurement between the *small space* and the *big space* are *averaged out* by a sort of quantum analog of the classical mean value theorem.

Notice that decoherence *does not imply the projection postulate*. It only implies that measurement is described by a map between n dimensional Hilbert spaces. Needless to say, the proof that this spitting can be done in realistic system is far from complete.

Another example of a system where decoherence is expected to be of use is the interaction of a heavy particle α (of mass M) with a huge (N) number of particles β of mass $m \ll M$. The origin of the decoherence in this case can be seen as follows. The initial state of the particle is represented by a wave function ϕ_{in}. In this case, under the assumption that the interaction is very weak and that $\delta \equiv \frac{m}{M}$ is very small one can expect to substitute, with a negligible error, the evolution with its asymptotics (the scattering matrix) and take the product of the scattering matrices, each of which differs little from the identity. One has to take into account that the description of scattering is usually done in the reference system of the barycenter and this differs (although little) form the barycenter of the wave describing the particle.

To first order in ϵ and δ the modification of the wave of the heavy particle due to the a single scattering is a phase shift and therefore to first order its state is not changed. This is also true under the product of the scattering matrices. To second order there is small spreading of the wave function of the heavy particle due the mismatch mentioned above between the two barycenters and a small modification of the shape of the wave function of the heavy particle due the properties of th scattering matrix. If $\epsilon \simeq \delta \simeq \frac{1}{N}$ is very small, *and if the scattering evens are independent* one *expects* that the final state of the heavy particle can be described by a density matrix σ with density $\rho(x)$ in the Schrödinger representation where the support of ρ is larger that of ϕ_{in}. Notice that the final state is represented by a density matrix and not by a vector because the phases introduced by the scattering matrices in the second order term are essentially random and the best approximation to this term is given by a density matrix.

One should also prove convergence of the perturbation series. This process is called *dynamical decoherence* or *environment induced decoherence*.

2 Decoherence. II

From a mathematical point of view decoherence is linked to *partial trace* or *conditioning*.

Conditioning in Quantum Mechanics has properties similar to those of the operation with the same name in Classical Probability Theory. But one should notice that, contrary to what happens in classical probability theory, in Quantum Mechanics *a complete information about the system does not imply knowledge of each of its components*.

Suppression of information relative to the environment should lead, according to the point of view of decoherence theory, to write *effective equations* for *a subset* of the measurable quantities of the subsystem that we consider, associated to commuting operators. The choice of the subset depends on the structure of the interaction. The effective dynamics that one obtains should describe the evolution of these observables independently of the evolution of the environment *for almost all its configurations* and for sufficient long time.

As a warning, it should be noticed that in the published paper on the theory of decoherence many results are accepted on qualitative or even intuitive ground and very few of them are mathematically proved. Further analysis, mainly in the direction of proving rigorously some of the key results, is required in this important subject.

While the role of coherence in quantum mechanics (and the associated difficulties) was evident to Schrödinger, the role of decoherence was appreciated only relatively late. The ubiquitous presence of environment entanglement in quantum mechanics was emphasized in the 70th by Zeh [21] and the formalism of decoherence was developed ten years later by Zurek [22] and was further analyzed in [15]. Decoherence by scattering of a large number of light particle is described in [20] by Tegmark. On the mathematical side results are very scarce and refer to very simplified systems. Among the latter we can quote [5, 8].

There are at present three books on decoherence [3, 10].

Roughly speaking the mechanism of decoherence is as follows. The state of the total system *observed object plus environment* is

$$\Psi \in \mathcal{H}_{tot} = \mathcal{H}_{obs} \otimes \mathcal{H}_{env} \quad \Psi = \psi \otimes \phi, \quad \psi \in \mathcal{H}_{obs} \quad \phi \in \mathcal{H}_{env} \tag{2}$$

If one measures an observable $A \in \mathcal{B}(\mathcal{H}_{obs})$ the mean of the values that is obtained is $(\Phi, (A \otimes I)\Phi)$ (we have used the natural immersion of $\mathcal{B}(\mathcal{H}_{obs})$ in $\mathcal{B}(\mathcal{H})$).

If the hamiltonian of the total system is H, a time $t > 0$ the measurement of A in the state Φ will give

$$(\Phi, e^{itH}(A \otimes I)e^{-itH}\Phi) \tag{3}$$

Due to the interaction between the two systems, *there does not exist* in general an operator $K \in \mathcal{B}(\mathcal{H}_{obs})$ such that for all $A \in \mathcal{B}(\mathcal{H}_{obs})$ and for a generic state Φ one has

$$(\Phi, e^{itH}(A \otimes I)e^{-itH}\Phi) = (\phi, e^{itK}Ae^{-itK}\phi) \tag{4}$$

One can only expect that if the environment has a large number N of degrees of freedom and the interaction hamiltonian has suitable properties a mean value (ergodic) theorem holds, at least approximately.

One can therefore at most expect that the following be approximately true: for a suitable class $\mathcal{A}$ of observables in $\mathcal{H}_{obs}$ (this class depends on the properties of the interaction of the system with the environment) extended to $\mathcal{H}_{tot}$ by the operator $A \times I$ and for a suitable class of initial states $\Psi \in \mathcal{H}_{tot}$ one can find an operator K on $\mathcal{H}_{obs}$, a state σ_0 and a time t_0 such that for $t \geq t_0 > 0$

$$|(\Phi, e^{itH}(A \otimes I)e^{-itH}\Phi) - Tr\sigma_0(e^{itK}Ae^{-itK})| \leq \epsilon_N, \qquad lim_{N\to\infty}\epsilon_N = 0 \qquad (5)$$

This first step has much in common with the Born-Oppenheimer approximation in molecular Physics, originating form the huge difference in the time scale between the motion of the electrons and the much heavier nuclei. In a first approximation one may assume that the electrons move in the force field of the nuclei considered as fixed points, and then considers the motion of the nuclei due to the interaction with the electrons. We shall come back in the course of these Lectures to the Born-Oppenheimer approximation.

The second step is to prove that the long-time effect of evolution is to *cancel the effects of decoherence* for some quantum mechanical states, so that *after some time* these states can be represented by classical Liouville distributions relative to some *preferred observable* (in general, the position). Notice that in general the preferred observable, if it exists, is in general uniquely determined by the interaction.

Attempts in this direction have been made in particular by Joos-Zeh [15] and by Tegmark [20]. The latter Author has given strong qualitative arguments, using quantum mechanical scattering theory. that decoherence can be produced by the interaction of a system with a large number of light particles. In this case position variables emerge as "pointer basis" (variables which have a classical behavior). We remark that, in the one-dimensional case, for a specific form of the interaction and for the scattering of only light particles the analysis in [15, 20] has been made rigorous by Dürr et al. [5]. One finds a panoramic view on these considerations e.g. in [10] and [14].

Arguments have been given also to describe the decoherence for a quantum system *in a thermal bath*, i.e. interacting with a large number of particles in thermodynamic equilibrium at fixed temperature. In this approach one makes use of the approximation in which the dynamics of the system under observation is described by a Markov semigroup. An approach to the mathematical description of decoherence has been pursued in [3] in the framework of the algebraic formulation of Quantum Mechanics (which we will describe briefly the course of these Lectures).

We have described what one *expects* but we have not been precise about the meaning of the adjective *suitable* neither for how long one has to wait. In fact, from the point of view of mathematics decoherence theory is still in its infancy and *much work is required in this important subject*.

Other attempts to define or find *decoherence free subspaces* or *pointer subspaces* (usually decoherence is used to describe the outcome of measurements in which the measuring apparatus has a pointer) may be found in the references to this Lecture.

From this point of view decoherence can be viewed as a filter on the space of quantum states of a given system that *for a given Quantum Dynamics* singles out those states that can be stably prepared and maintained. This filter effectively excludes *non-classical states* e.g. those popularized as superposition of dead and alive cat states. In this sense decoherence is at the heart of the problem of the relation between the quantum world and the classical one. In spite of these developments the mathematical theory of decoherence is still in its infancy. Only special cases have been treated rigorously and only *strong qualitative arguments* have been given in sufficiently general cases.

Recall that the Schrödinger equation is a *deterministic equation for probability waves*. It is deterministic in the sense that it is well posed, i.e. given the initial datum the solution exists and is unique; in particular the solution is identically equal to zero if the initial datum has this property. The measurement *changes the initial datum* (since the initial datum is a probability amplitude it can be changed by acquiring more information). Decoherence aims at substituting the probability wave with a classical probability distribution and substituting (approximatively) the Schrödinger equation with a well posed *effective equation*.

In the case of interaction with a measuring apparatus decoherence tries to describe how, under suitable assumptions *and for a selected number of observables that depend on the performed measurement*, it is possible to interpret the outcome of an experiment according to the rules of *classical* probability theory in spite of the fact that the operators associated to these observables do not commute with the hamiltonian. In this case one expects that Classical Mechanics describes the evolution of the *selected observables*; the only trace of Quantum Mechanics is that a pure state of the system is now described now by a Liouville distribution obtained by Born's rules. One expects that this be possible because due to the interaction after a very short time one can neglect the correlations between spectral projection on disjoint intervals of the spectra of the operator corresponding to these selected observables.

It is useful to stress again that decoherence *does not solve the measurement problem* in Quantum Mechanics i.e. the selection of a pure final state. It may provide a mechanism through which the final state can be described by a statistical mixture but it *does not* provide the mechanism through which a specific result is obtained.

3 Experiments

From the point of view of experiments relevant progress has been made in the measurement of coherence and of *its disappearance due to interaction with the environment*. Very refined experiments have been performed in particular by the group of S. Haroche at E.N.S.

A typical experiment is described in [12, 16, 17]. Reference [16] gives a particularly readable account of the experiment.

The experiment is aimed at probing the coherence and decoherence of a *photonic cat*. The photonic cat is composed of photons. It is prepared in a *classical (coherent)* state, coherent superposition of states with different number of photons. This state is called *classical* because it is an eigenstate of the quantized electromagnetic field and its wave function satisfies the equations of classical electromagnetism. An example of non-classical state of the photonic cat is a state with a fixed number of photons.

The photonic cat is prepared in a classical state in a cavity (photon box) with perfectly reflecting walls. The lifetime of photonic cat in this cavity is about 130 ms; in this time the photons travel in a folded trajectory for about 40,000 kms and survive long enough for a thousand of atoms to cross the cavity one-by-one and interact with the photons.

In the experiment in order to probe the state of the photons field (the photonic cat) rubidium atoms (atomic mice) are injected in the photon box; the atoms are in a Rydberg state of quantum number 50 (called f) or 51 (called g).

These states have natural life-time of the order of 30 ms, which is roughly the same as the photon lifetime in the cavity. This justifies neglecting the atomic decay process during the interaction with the cavity field. The (circular) Rydberg states have another important property: these states have very small electric dipole and therefore interact weakly with the photons.

In order to create an electric dipole (that disturbs the photonic cat) a pulse of resonant microwave is applied to the atomic mouse. This brings it (the rubidium atom) to be in a coherent superposition of the 50 and 51 states. The photonic cat is prepared in a classical quasi-monochromatic state with frequency resonating with that of the transition between the 50 and 51 states of the rubidium mouse. Since the de Broglie wave of the states of the atomic mouse have wave numbers that differ by one unit, the superposition results in an electric dipole rotating in the orbital plane at 51 GHz. This dipole field acts on the photonic cat and changes its structure, in particular its average number of photons (the cat becomes aware of the presence of the atom-mouse).

The photon cat in now in a state that is a superposition of coherent states. The wording reminds us of the Schrödinger cat suspended between the state of being dead and that of being alive; its fate is controlled by a single atomic mouse. One cannot any longer speak of a well defined state of the photonic cat (neither of the atomic mouse).

The environment is in this case represented by the walls of the cavity. The rubidium 51 atoms are injected one at a time in the open cavity with reflecting walls (photon box). The cavity can keep for 1 ms a specified number of photons of wave length 6 mm. The electromagnetic field in the cavity is prepared in a state of 8 up to 10 photons; this state approximates reasonably *a coherent state*. Through a very sophisticated apparatus one can measure the number and polarization of the photons in the cavity at any time.

The decoherence time δ_T is the time it takes to the quantum cat to relax to a coherent (classical) state with a relative error of 1×10^{-3}. The experiment measures

the time of decoherence for the photonic cat by changing the delay with which the second atom is injected into the cavity. The results of the experiment which we have roughly described above indicate that the decoherence time in the case of the photonic box is of the order of 0.1 ms.

In this case the coherence of the photonic cat persists for a fraction of a second *in presence of a very limited number of photons* and in a very strictly controlled environment. One can expect that in the case of the *real* Schrödinger cat the survival time of coherence be several orders of magnitude smaller and therefore *coherence cannot be seen under normal every-day's life conditions* (i.e. for true cats).

This and related experiment show that a result that seems to be counter-intuitive when one has little control of the environment can be observed in a laboratory *in which maximum control is possible*.

4 Bell's Inequalites

We have seen that one faces serious difficulties in the interpretation of Quantum Mechanics as a causal theory in view of the probabilistic nature of Born's postulate. And with the probabilistic interpretation one has difficulties in making precise the instant in which the deterministic dynamics described by the Schrödinger equation should be substituted by the reduction postulate and Born's rule. The reference to a *macroscopic observers* lacks precision because there is no clear description of what *macroscopic* means. Therefore Quantum Mechanics in spite of its extraordinary success is conceptually incomplete.

One may be led to attribute this to the incompleteness of the description of a state of a system in Quantum Mechanics; if this is the case, the addition of new variables (that are *hidden* in the formulation of Quantum Mechanics) may cure the conceptual problems and lead to a complete and consistent theory. That this is not the solution is proved by means of inequalities described by Bell and therefore called *Bell inequalities*. These are inequalities for the results of measurements performed on a suitable system by two separate observers traditionally called *Alice* and *Bob*.

These inequalities are satisfied in Quantum Mechanics but *cannot be satisfied* if one assumes that one has only incomplete knowledge of the system because some degrees of freedom are hidden, and complete knowledge also of these *the hidden variables* would lead to a description according to Classical Theory.

For simplicity consider a quantum system described in a complex four-dimensional Hilbert space $\mathcal{H}$ e.g. a system of two identical particles of spin $\frac{1}{2}$ neglecting their configuration in space.

We write $\mathcal{H}$ as

$$\mathcal{H} \equiv \mathcal{H}_1 \otimes \mathcal{H}_2 \tag{6}$$

where $\mathcal{H}_i$, $= 1, 2$, are the Hilbert space that describe the state of each particle.

Denote by $\sigma_\eta \;\; \eta \in S^3$ the operator (2×2 matrix) that each observer associate to the spin. Choose an orthonormal frame and write S_i the generator of rotations in the ith direction. One has the following commutation relations

$$[S_i, S_k] = \sum_l \epsilon_{i.k,h} S_h \tag{7}$$

where $\epsilon_{i,k,h}$ is the completely antisymmetric Ricci symbol.

S_k is also the generator of rotations around the kth axis in the projective representation of the rotation group induced in $\mathcal{H}$ by the unitary representation of $SU(2)$.

Denote by $|+>, |->$ the basis defined by the eigenstates of S_3. Suppose that Alice and Bob both perform the measurement when the system is in the state described by the vector

$$\psi \equiv \frac{1}{\sqrt{2}}[(|+> \;\; |->) - (|-> \;\; |+>)] \tag{8}$$

Consider an experiment in which Alice measures the spin component along $\hat{x} \in R^3$ and Bob measures the one in the direction $\hat{y} \in R^3$. The corresponding operator is $S_x \otimes S_y$. Its expectation value in ψ is $(\psi, S_x \otimes S_y \; \psi) = -(\hat{x}, \hat{y})$.

If Alice chooses to make two measurements of the spin in the directions x_1 and x_2 and Bob chooses to make two measurements in the directions y_1 and y_2 the four possible results correspond to the 4×4 matrix

$$(\psi, S_{x_i} \otimes S_{y_i} \psi) = -(\hat{x}_i, \hat{y}_j) \qquad i, j = 1, 2 \tag{9}$$

If one chooses the unit vectors

$$x_1 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), \quad x_2 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0) \quad y_1 = (0, -1, 0), \quad y_2 = (1, 0, 0) \tag{10}$$

the corresponding quantum expectation satisfy

$$-\frac{1}{2} x_1 \,.\, y_1 - \frac{1}{2} x_1 \,.\, y_2 - \frac{1}{2} x_2 \,.\, y_1 + \frac{1}{2} x_2 \,.\, y_2 = \sqrt{2} \tag{11}$$

Consider now the outcome of the same experiment in a hidden variables theory that is *local* i.e. in which the direction chosen by Alice can be regarded as *independent* from that chosen by Bob. The theory assigns to the physical system under measurement a (hidden) parameter λ whose value determines the result of the experiment. Let $X_i(\lambda)$ the result obtained by Alice when she measures in the direction x and the parameter has value λ. By definition one must have $X_i(\lambda) = \pm 1$. Let $Y_j(\lambda)$ the result obtained by Bob.

If the hidden variable theory should reproduce the quantum correlations between the various result of the measurements one must have

$$E(X_i\ Y_j) = \int X_i(\lambda)\ Y_j(\lambda) d\mu(\lambda) = (\psi, S_{x_i} \otimes S_{y_j}\psi) = -(x_i, y_j) \tag{12}$$

where we denoted by E the expectation with respect to a probability measure μ in the hidden-variable space Λ.

We remark now that, since the functions $X_i(\lambda)$ and $Y_j(\lambda)$ take only the values ± 1, for any choice of μ the following inequality holds

$$\frac{1}{2}X_1Y_1 + \frac{1}{2}X_1Y_2 + \frac{1}{2}X_2Y_1 - \frac{1}{2}X_2Y_2 \leq 1 \tag{13}$$

and therefore, independently from the choice of μ

$$\frac{1}{2}E(X_1Y_1) + \frac{1}{2}E(X_1Y_2) + \frac{1}{2}E(X_2Y_1) - \frac{1}{2}E(X_2Y_2) \leq 1 \tag{14}$$

A simple comparison of (12) with (14) shows that the quantum result *cannot be obtained* within a (naive) hidden variables theory.

Notice that the choices of directions in (14) *maximize* the algebraic expression

$$g(\hat{a}, \hat{b}, \hat{c}, \hat{d}) \equiv (\hat{a}, \hat{b}) + (\hat{a}, \hat{c}) + (\hat{d}, \hat{a}) - (\hat{d}, \hat{c}) \tag{15}$$

where $\hat{a}, \hat{b}, \hat{c}, \hat{d}$ are unit vectors in R^3 and $(\hat{a}, \hat{b})$ is the scalar product of $\hat{a}$ with $\hat{b}$.

Experiments performed by Aspect et al. [11] have shown that in the conditions the result for $g(a, b, c, d)$ is a number strictly greater than one, and recent results of experiments performed with more refined techniques have given a result close to $\sqrt{2}$ in accordance the prediction of Quantum Mechanics.

It is interesting to compare Quantum Mechanics and Hidden Variables Theory when one considers only the case in which Alice and Bob perform measurements in any direction but always the same for both.

In Quantum Mechanics one will always have $(\psi, S_i \otimes S_i\psi) = -1$. In the hidden variables theory this implies $X_i(\lambda) = -Y_i(\lambda)$ for all values of λ. Therefore

$$E(X_iX_j) = -E(X_iY_j) = (\hat{x}_i, \hat{x}_j) \tag{16}$$

On the other hand for variables can only take value ± 1 the following relation holds

$$-X_1\ X_2 - X_1\ X_3 - X_2\ X_3 \leq 1 \tag{17}$$

and therefore independently of the measure μ

$$-E(X_1\ X_2) - E(X_1\ X_3) - E(X_2\ X_3) \leq 1 \tag{18}$$

Taking $\hat{x}_1,\ \hat{x}_2, \hat{x}_3$ coplanar and forming and angle of $\frac{2\pi}{3}$ one obtains

$$-(\hat{x}_1, \hat{x}_2) - (\hat{x}_2, \hat{x}_3) - (\hat{x}_3, \hat{x}_1) = \frac{3}{2} \tag{19}$$

A comparison of (19) with (18) leads again to the conclusion that Quantum Mechanics is incompatible with a (naive) hidden variables theory.

Notice that in the state ψ is antisymmetric for exchange of Alice and Bob. This state can also be written

$$\psi \equiv \frac{1}{\sqrt{2}}[(|+>_\beta\ \ |->_\beta) - (|->_\beta\ \ |+>_\beta) \tag{20}$$

where β is *any* unit vector in R^3 and $|\pm>_\beta$ are the eigenvectors corresponding to the value ± 1 of σ_β.

From this it is easy to deduce that if Alice and Bob perform a measurement of the spin *in the same direction*, the measurements give *always opposite values*, as expected since the total spin of the state is zero! One is led to the conclusion that the property to have the spin directed in any chosen direction ξ *is not a property of the state*. This correlation is due to the fact that the state ϕ is *entangled* i.e. it cannot be written as $\psi \otimes \phi$ with $\psi,\ \phi \in \mathcal{K}$.

5 Alternative Theories

The conceptual difficulties of Quantum Mechanics have led to a search for possible alternatives.

It is natural to think that the statistical nature of the theory is due to incomplete knowledge of the state of the system. The system has further degrees of freedom that are *hidden*; a complete determination of their values would allow to make sure predictions for any sequence of measurements. We saw that this *naive* hidden variables hypothesis is ruled out by *Bell's inequalites* which are experimentally verified. Bell's inequalities follow from the formalism of Quantum Mechanics but are violated by naive hidden variables theories.

A further deep analysis done by Bell [1] shows that experimental evidence is incompatible with a theory which is *both local and causal*. To preserve locality one may assume [9] that Schrödinger's equation does not provide a complete description of quantum dynamics, but must be supplemented by a random process in configuration space which provides the collapse of the wave function ψ to x with a probability density given by $|\psi(x)|^2$.

No experimental evidence for the need of this modification of the Schrödinger equation has been found so far but the experiments that may give evidence of these effects are very difficult.

A theory alternative to Schrödinger's Quantum Mechanics is the *Pilot Wave Theory*, which we describe now. It is a deterministic theory but it is not based on the assumption of *local* hidden variables. In some way a *hidden variable* exists but it is a field (the pilot wave) which can be felt only through its action on the particles. As we remarked in "Lecture 1: Elements of the History of Quantum Mechanics I", the photoelectric effect (and the theory of photonic nature of light) suggested that the phenomena at atomic scale could be formulated as a theory of material points. Pilot Wave Theory takes up this view.

Pilot Wave Theory was formulated by de Broglie in the years 1923–1925 and presented at the Solvay conference in 1927. It was praised by Einstein and Lorenz (Einstein himself had earlier tried to see the electromagnetic field as *guiding field* for the photons). At that time some mathematical problems remained connected with the singularity of the field. The theory was shadowed by the successes and consensus that meanwhile was having the new Mechanics of Schrödinger and Heisenberg.

To the dismissal of Pilot Wave Theory contributed also the fact that it is not a Newtonian theory. It is a theory described by equation of first order in time in the variables position and therefore has no evident connection to Lagrangian and Hamiltonian dynamics. Moreover it is more cumbersome: one has to solve Schrödinger 's equation for a system of N particles in order to determine the structure of the vector field and once this problem is solved one has still to solve a singular non-linear differential equation in $3N$ variables.

The theory was taken up again by Bohm in 1952, [4] who emphasized a remark by de Broglie that the equation given by the vector field can be interpreted as Newton's equation if one added a potential force. Bohm gave the name *Quantum Potential* to this potential that has to be added *to take into account the quantum properties of the system*. One can introduce a Quantum Potential (making de Broglie's theory more similar to Hamiltonian Dynamics) by writing the solution $\phi(t, X)$ of the Schrödinger equation

$$i\frac{\partial\phi(t,X)}{\partial t} = -\sum_{k=1,\dots N}\frac{m_k}{2}\Delta_k\phi(t,X) + \sum_{k,j}V^N_{k,h=1}(x_k,x_j)\phi(t,X) \qquad (21)$$

for a N-particle system in polar form

$$\phi(t,X) \equiv R(t,X)e^{\frac{iS(t,X)}{\hbar}} \quad \rho(t,X) \geq 0 \qquad X \in R^{3N} \qquad (22)$$

where S and R are real valued function over R^{3N}.

The Schrödinger equation becomes then a system of two partial differential equations: the *continuity equation* for the density $\rho \equiv |R(t,X)|^2$ and an equation of Hamilton-Jacobi type that differs form the usual one encountered in Hamiltonian Mechanics for the addition of a potential-like term (*Quantum Potential*)

$$U(x) = -\sum_{n=1}^{N}\frac{\hbar^2}{2m_n}\frac{1}{R(x)}\Delta_n R \qquad (23)$$

We shall come back to these equations in "Lecture 16: Semiclassical Approximation for Fast Oscillating Phases. Stationary Phase. W.K.B. Method. Semiclassical Quantization Rules" devoted to the semiclassical approximation of Quantum Mechanics and there we shall discuss their mathematical structure. Notice that the phase is ill defined when $\rho(t, X)$ is zero and that the phase in (22) has very strong oscillations since the parameter $\hbar$ is very small.

In this way one recovers formally a theory in which force plays a role; de Broglie was aware of this but considered it too artificial.

D. Bohm instead emphasized this aspect of the theory, and was led to regard the theory as a theory of a *quantum fluid*. As a consequence in most textbooks in Physics in which Bohm's theory is mentioned and in most research papers in Solid State Physics the theory has became known as *Bohmian Mechanics* and is presented either as a modification of Hamiltonian Mechanics through the introduction of a Quantum Potential or as a theory that describes a quantum fluid. Rarely the original formulation by de Broglie is mentioned.

The equations for the quantum fluid as formulated by Bohm can be solved form an Eulerian point of view. One solves the Hamilton-Jacobi equation with the addition of the Quantum Potential by the introduction of trajectories in configuration space which are solutions of the (classical) equation of motion

$$\frac{dQ_{n,k}}{dt} \equiv \frac{1}{m_n}\frac{\partial S}{\partial x_{n,k}} \quad N = 1, ..N, \quad k = 1, 2, 3 \tag{24}$$

where m_n is the mass of particle m.

One may regard this formulation of the equations of Quantum Mechanics only as a formal analogy used mostly to exploit the amount of results available on the solutions of the Hamilton-Jacobi equation. In doing so one encounters the following difficulty.

From the continuity equation

$$\frac{\partial \rho}{\partial t} + div\, j = 0 \qquad j = Re\frac{\partial \Phi(t, X)}{\partial t}\Phi(t, X) \tag{25}$$

one derives that if $\Phi(t, X)$ satisfies Schrödinger's equation the vector field in (24) is given by

$$\frac{dQ_n}{dt} = \frac{\hbar}{m_n} Im\left(\frac{\bar{\Phi}\nabla_n \Phi}{|\Phi|^2}\right)(Q_1, .., Q_N) \tag{26}$$

where $Q_n \equiv \{q_{n,k}\ k = 1 \ldots 3\}$.

If magnetic field is present the gradient in (26) must be replaced by the magnetic gradient.

It follows from (26) that the vector field is singular at the points in which $\Phi(t \cdot X)$ is zero. In Bohm's formulation of Quantum Mechanics as Quantum fluid this problem is often overlooked. The formal resolution of this difficulty is attributed to the fact that by definition the particle has zero probability of being there.

The modern formulation of de Broglie's pilot wave theory has been developed by S. Goldstein, D. Dürr and collaborators in the early nineties. Through their work the theory of de Broglie was revived and brought to be a fully mathematical theory. Good reference are [6, 7].

In the formulation of Goldstein and Dürr the *primitive observables* are the coordinates of N material points

$$q_1(t), ..q_N(t), \quad q_n \in R^3, \quad n = 1, ..., N \tag{27}$$

The system is composed of these point and in addition of a complex-valued *pilot wave*

$$\Phi(t, X) \qquad X = \{x_1, ..x_N\} \quad x_n \in R^3 \tag{28}$$

defined in the configuration space of the particles.

In this formulation Pilot Wave Theory places itself half-way between Quantum Mechanics and the naive hidden variable theories. There is no hidden degree of freedom but the wave is *hidden*: it can only be perceived through its action on the particles. There have been attempts to consider the Pilot Wave as an *effective* result of the interaction between the N particles in (27) and every other particle in the universe but this point has not been clarified yet.

The dynamics of the pilot wave is given by

$$i\hbar \frac{\partial \Phi(X)}{\partial t} = (H\Phi)(X) \tag{29}$$

where H is the (non-relativistic) Schrödinger operator. This dynamics is *independent* of the position and velocity of the points.

Equations (28) and (29) are invariant under time inversion and Galilei transformations.

This is the simplest system of equations that guarantees the equivalence with Quantum Mechanics for position measurements. This equivalence is due to the fact that the vector field in (26) is proportional to $\frac{J(t,X)}{\rho(t \cdot X)}$, the ratio between current density and probability density in Quantum Mechanics. Recall that the theory is not newtonian and by definition the vector field is proportional to the current.

By construction for measurement of the position of the particles at a given time the pilot wave theory gives *the same results* as Quantum Mechanics. Indeed if at time t_0 the configuration of the system is a random variable with density $|\Phi(t_0, X)|^2$, according to Pilot Wave Theory at a later time t the density will be $|\Phi(t, X)|^2$. Momentum is a derived observable (mass time the velocity of the point). For trajectories that are not differentiable in time momentum is not defined. In Pilot Wave Theory the observables are functions of the coordinates; one can measure different coordinates at different times and one can consider as observables the exit times of the points from a given domain Ω. This quantities *do not correspond to observables in Quantum Mechanics*.

As we have pointed out, Eq. (25) is not free of difficulties, the principal being that the vector field is divergent at the points in which $\Phi(t, X) = 0$; in a neighborhood of these point the equation must be handled with care. On the other hand this configuration of material points occurs with zero probability. A careful analysis shows that the dynamics is well posed (uniqueness and continuity of the solutions), in spite of the singularity. This is one of the major achievements of the modern version of the Pilot Wave Theory which has acquired a status of mathematically complete theory.

Due to this mathematical difficulties Bohm's theory had little success among researchers more mathematically inclined and was mostly used to find semiclassical solutions (the added potential term U is of second order in $\hbar$ and this suggests the use of an iteration scheme). In Pilot Wave Theory this difficulties are overcome by a detailed mathematical analysis of the solutions of (25).

The conceptual difference between the two theories is that Quantum Mechanics states that *the entire knowledge about the system* at a given time is given by the wave function Ψ and $|\Psi(t, X)|^2$ is the *probability density* of the configuration X *when one performs at time t a position measurement*. In Pilot Wave Theory the entire knowledge of the system at time t is provided by *the position* X of N particles ($X \in R^{3N}$) *and the pilot wave* which gives its dynamics. The momenta of the particles are not primitive quantities in this theory, which is described *by first order differential equations*. The velocity of a particle is given by component of the vector field relative to the degrees of freedom of the particle evaluated at the position of the particle (recall that the vector field has a many components as as the number of degrees of freedom of the system). The momentum of a particle at a given point is by definition the product of this vector field and of the mass of the particle and is a derived quantity.

At any time the evolution of the pilot wave is given as a function of the positions of the particles by the N-body Schrödinger equation. The particles at time t have a physical reality *independently from the presence of an observer who measures their positions*.

From (25) one derives that the pilot field does not change if the wave function is multiplied by a constat phase (it depends only on the ray) and is zero if the wave function can be taken real. In particular it is constant in time for eigenvalues of Schrödinger's Hamiltonian and it is zero if the eigenvectors are simple: to these eigenvectors correspond in Pilot wave theory *equilibrium distributions*. In particular this applies to the ground state. If the number of particles is very large one may hope to regard *the system as a thermodynamic N-body system described by the laws of thermodynamics*. The distribution of the particles (statistical ensemble) is given by $|\Psi(t, X)|^2$ and this distribution remains constant if the system is in equilibrium.

The epistemological interest of Pilot wave theory lies in the fact that the position of the points have a well defined status and the result of their measurement does not depend on the observer and on the previous measurements. Note that the observer itself is a collection of points. The non-local character of Quantum Mechanics is reflected in Pilot Wave Theory in the fact the dynamics of the points, depending as it does on the solution of the Schrödinger equation, cannot be written solely in terms of their positions at least in a local way.

In Pilot wave theory the interference patterns in two slit experiment are recovered by carrying out an accurate analysis of the trajectories of each material point: the solutions of the system (27) satisfy the uniqueness property and therefore two trajectories cannot touch each other. One proves that starting with initial data distributed uniformly in an initial plane one finds in this way the maxima and minima that Quantum Mechanics attributes to interference phenomena.

By closing one of the two slits in a two-slit experiment one modifies the boundary conditions for the Schrödinger equation, therefore one modifies the solution and consequently one modifies the vector field. The probability distribution of point of arrival on the screen is radically changed and all "interference" effects disappears.

One of the difficulties is that if the pilot wave is real, the vector field is zero; this occurs e.g. for the ground state of the hydrogen atom. In this state the points do not move. On the other hand, it must be noted that the wave function is real when one considers the hydrogen atom isolated and in absence of the electromagnetic field. In presence of the electromagnetic field and of other atoms the wave function is no longer real and the vector field is no longer zero (the particles move). In Quantum Mechanics this interaction gives the transition from one atomic state to the other, in Pilot Wave Theory it gives to the material point the possibility to change position. Ionization corresponds to the fact that one or more points can be displaced and move freely.

In Quantum Mechanics if one consider a system of N points composed of two subsystems composed of N_1 and N_2 if the interactions are of sufficiently short range and the configurations of the sub-system are separated enough the motion of the two subsets of points are essentially independent. The same occurs in Pilot WaveTheory as long as the Pilot wave factorizes.This is not the case if the wave functions are entangled; in this case *it is as if* there were an action at a distance between the two sets of points.

In the formulation of the Pilot Wave Theory we have described the wave is a structure superimposed to the system of particles to provide their dynamics. There are attempts to regard the wave in some sense as the result of global structure of the system composed of all the particles in the Universe and in this sense only a bookkeeping devise. The wave would be *of statistical origin* and it would reflect the properties of a *typical configuration* of the entire system.

A difficult point in Pilot Wave Theory is to justify the assumption that the probability density that the points in an atomic state be in position $X(t_0) \equiv \{x_k(t_0)\}$ at time t_0. There have been attempts to relate this problem to the initial conditions and to an equilibrium average and *typicality* of the configurations but so far a decisive argument is still missing. In this context it is still problematic the justification of Born's rule as initial distribution of all the material points. This is the obscure point in Pilot Wave Theory.

We shall not discuss any further the pilot wave theory and we refer to a review paper of Berndl et al. [2] and the book [6].

From a conceptual point of view the Pilot Wave Theory has, compared with Quantum Mechanics, the advantage of introducing *as observables* only the positions of material points, and avoids the problems connected with measurement (also the

measuring apparatus is composed of material points). The theory is fully deterministic: given at time 0 the shape of the pilot wave and the position of the points, the future configuration of the pilot field and the position of the points are uniquely determined.

References

1. Bell, J. (1987). *Speakable and unspeakable in quantum mechanics*. Cambridge: Cambridge University Press.
2. Berndl, V., et al. (1995). *Il Nuovo Cimento 110 B*, pp. 737–750.
3. Blanchard, P., & Olkiewicz, R. (2003). *Reviews in Mathematical Physics, 15*, 217–243.
4. Bohm, D. (1952). *Physical Review, 85*, 166–193.
5. Dürr, D., Figari, R., & Teta, A. *Journal of Nonlinear Mathematical Physics*.
6. Dürr, D., & Teufel, R. S. (2009). *Bohmian Mechanics*. New York: Springer.
7. Dürr, D., Goldstein, S., & Zanghi, N. (1992). *Journal of Statistical Physics*, *67*, 843–907.
8. Figari, R., et al. (2004). *International Journal of Modern Physics B,* 667–674.
9. Ghirardi, G. C., Pearle, P., & Rimini, A. (1990). *Physical Review A*, *42*, 78–89.
10. Giulini, D., et al. (1996). *Decoherence and the appearance of a glassical word in quantum theory*. New York: Springer.
11. Grangier e, P., & Roger, G. (1982). *Physical Review Letters, 49*, 91.
12. Haroche, S., et al. (2007). *Nature*, *446*, 297–300.
13. Hepp, K. (1972). *Helvetica Physica Acta, 45*, 237.
14. Hornberger, K. *Introduction to decoherence theory*, arXiv:quantum-ph 0612118 [quant-ph].
15. Joos, E., & Zeh, H. (1985). *Zeitschrift fur Physik B*, *59*, 223–243.
16. Raimond, J. M. (2011). *Exploring the quantum world Les Houches*, Session XVVI.
17. Raimond, J. M., Brume, M., & Haroche, S. (2001). *Reviews of Modern Physics*, *73*, 565–592.
18. Sclosshauer, M. (2007). *Decoherence and the quantum-to-classical transiton*. New York: Springer.
19. Sewell, G. (2005). *Reports on Mathematical Physics*, *56*, 271.
20. Tegmark, M. (1993). *Foundations of Physics Letters*, *6*, 571–590.
21. Zeh, H. D. (1970). *Foundation of Physics, 1*, 69.
22. Zurek, W. H. (1981–1982). *Physical Review D, 24*, 1516–1534, 26 1862–1885.

Lecture 5: Automorphisms; Quantum Dynamics; Theorems of Wigner, Kadison, Segal; Continuity and Generators

Recall that as compared to Hamiltonian Mechanics in Quantum Mechanics the role of real bounded functions is taken by the symmetric elements of the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators A on a Hilbert space $\mathcal{H}$ (observables). The role of Lagrange densities ρ (positive functions of class L^1 on phase space with integral equal to one) is taken by positive trace-class operators σ with trace 1 (density matrices).

Recall also Born's rule: if A is the operator that describes the observable a and σ is the density matrix that represents the state S, then $Tr(\sigma A)$ is *the mean value* of the results which one obtains measuring the observable a when the system is prepared in the state S. Both in Classical Mechanics and in Quantum Mechanics the states so defined are *normal* (the supremum of the values assigned to an increasing filter of observables coincides with the value assigned to the supremum of the filter of the corresponding operators).

We shall use this correspondence between states and observables to define the evolution in Quantum Mechanics.

1 Short Summary of Hamiltonian Mechanics

In Classical Mechanics the evolution of a point *in phase space* with coordinates $z \in R^{2n}$ is given by a vector field $f(z)$ through

$$\dot{z} = f(z) \quad z(0) = z_0 \quad z = \{z_i\} \quad i = 1 \dots 2n \quad z_i = q_i, \quad z_{i+n} = p_i, \tag{1}$$

The evolution of a continuous function ϕ is given by duality by

$$G_t(z_0) = G(\phi_{-t}(z(t, z_0))). \tag{2}$$

Denote by $\phi_t(z)$ the trajectory (the solution of (1)); we assume that it is defined for all initial data z_0 and for all times t. If G is differentiable (2) is equivalent to

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_5

$$\frac{dG}{dt} = f \cdot \nabla G. \tag{3}$$

The definition (2) can be extended by duality to measures

$$t \to \phi_t^*(\mu), \qquad \phi_t^*(\mu) \equiv \mu(G_t) \tag{4}$$

The (weak) differential version of (4) is

$$\frac{d\mu_t(G)}{dt} = \frac{d}{dt}\mu(G_t) \tag{5}$$

If the measure is absolutely continuous with respect to Lebesgue measure, with density ρ, (5) takes the form

$$\frac{d\rho_t}{dt} = -div(f\ \rho_t) \tag{6}$$

In the case of a Hamiltonian dynamics with Hamiltonian H one has

$$\dot{z} = JdH \qquad z \in R^{2n} \qquad \frac{dG}{dt} = f \cdot \nabla G \equiv \{H, G\} \tag{7}$$

where J is the standard symplectic matrix and $\{.,.\}$ is the associated Poisson bracket. The map

$$G \to \partial_H(G) \equiv \{H, G\}$$

for H real defines a *-derivation since it is linear, preserves complex conjugation and satisfies Leibnitz's rule for products. Recall that a derivation can be seen also as a *vector field* and therefore can be the generator of the flow $\phi(t) :\ G \mapsto G(t)$ where $G(t)$ is the solution of

$$\frac{dG}{dt} = \{H,\ G\} \tag{8}$$

Moreover the *-derivation defined by H satisfies the (Jacobi) identity

$$\partial_H(FG) = \{H, FG\} = \{H, F\}G + F\{H, G\} = \partial_H F.G + F\partial_H G \tag{9}$$

This identity can be interpreted as the statement that the one-form dual to the vector field is exact.

The domain of this derivation is the collection of differentiable functions.

The evolution described by (8) defines a *one-parameter group of automorphisms* of the abelian algebra C^0 of continuous functions. By duality it defines also a one-parameter group of automorphism of the Banach space L^∞.

2 Quantum Dynamics

As we have seen in "Lecture 2: Elements of the History of Quantum Mechanics II", since the beginning of Quantum Mechanics structural analogies with Hamiltonian Mechanics were evident; in fact these analogies played a relevant role and are often stressed.

We want now to see if there are deeper reasons for this analogies.

In this context it is important the following theorem, proved by Dirac for matrices; as we shall see the proof extends to the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a separable Hilbert space.

Theorem (Dirac) *Let $\mathcal{H}$ be a separable Hilbert space and let δ be a norm-continuous *-derivation with domain $\mathcal{B}(\mathcal{H})$.*

There exist (unique modulo an additive constant) a bounded symmetric operator A such that

$$\delta(B) = i[A, B] \qquad \forall B \in \mathcal{B}(\mathcal{H}). \tag{10}$$

We shall give in "Lecture 8: Properties of Free Motion, Anholonomy, Geometric Phase", under some further hypotheses, a proof of this important theorem in the case when $\mathcal{H}$ is infinite-dimensional and the *-derivation is defined on a dense subset of $\mathcal{B}(\mathcal{H})$ and is continuous in the weak topology.

The operator A will be in general unbounded and therefore defined only on a subset of $\mathcal{H}$. This result lifts, under suitable assumptions, to the algebra of operators in a separable Hilbert space. This gives a natural structure to the dynamics, *independently from analogies* with Hamiltonian Dynamics.

We shall now prove that the formulation of the dynamics in Quantum Mechanics is a consequence of the Axioms 1 and 2 and some natural requirements on automorphisms of states and observables.

Given the axioms (and therefore accepting that in Quantum Mechanics observables are described by self-adjoint operators in a separable Hilbert space $\mathcal{H}$ and states by density matrices), in order to develop Quantum Dynamics we must

(1) study the automorphisms of the algebra $\mathcal{B}(\mathcal{H})$;
(2) study the automorphisms of the trace-class operators.

Since there is a duality between states and observables, there is a duality between (1) and (2).

It is remarkable that this leads to generators of dynamics which have the same structure as derivations in Hamiltonian Dynamics.

Definition 1 (**-derivations*) In analogy with the classical case we shall call *derivation* on $\mathcal{B}(\mathcal{H})$ a linear map δ that is linear and satisfies Leibnitz's rule

$$\delta(A\ B) = \delta(A)\ B + A\ \delta(B), \qquad A,\ B \in \mathcal{B}(\mathcal{H}) \tag{11}$$

We shall say that it is a *-derivation if it commutes with hermitian conjugation.

In Hamiltonian Mechanics to the Hamilton is associated a vector field on phase space. We may ask therefore that to every element $A \in \mathcal{B}(\mathcal{H})$ be associated a derivation δ_A. One has therefore for any triple A, B, C in $\mathcal{B}(\mathcal{H})$

$$\delta_A(BC) = \delta_A(B)\ C + B\ \delta_A(C) \tag{12}$$

This relation is certainly satisfied if $\delta_A B = [A, B]$ because in this case (11) is the Jacobi identity. Equation (11) takes then the form of a cocycle condition; the corresponding statement in Hamiltonian Mechanics is that the dual form is exact.

We shall see that this implies also that $(\delta_A)^2$ is dissipative i.e. it generates a semigroup of contractions in the Hilbert space (compare with the (improper) case of $A \equiv i\frac{d}{dx}$ with $(i\frac{d}{dx})^2 = -\Delta$).

3 Automorphisms of States and Observables

We start describing dynamics in Quantum Mechanics by considering first the discrete case, i.e. the automorphisms of the states and observables. Pure states are represented by *rays* i.e. equivalence classes of normalized $\mathcal{H}$ which differ by a phase. We shall indicate with $[\phi]$ the ray that contains the element ϕ.

Pure states can be also represented by one-dimensional orthogonal projection P_ϕ on the ray $[\phi]$. Statistical mixtures are represented by density matrices (positive trace-class operators with trace 1). We shall use the notation $\mathcal{I}_{1,+}$ to denote density matrices. We assume always that $\mathcal{H}$ is separable.

Notice that this representation of the states generates some ambiguity since in this way states are represented by operators (and therefore can be associated to an observable).

One should however remark that the topology on the states as density matrices is the trace topology (continuity of the map $\sigma \mapsto Tr(P_\phi \sigma)$ for $\sigma \in \mathcal{I}_{1,+}$) and the states in this topology form a separable Hilbert space with scalar product $(\sigma_1, \sigma_2) = Tr(\sigma_1 \sigma_2)$. This topology is *stronger* than the topology induced on $\mathcal{I}_{1,+}$ by the topology of $\mathcal{B}(\mathcal{H})$. As a consequence the *abstract dual* of $\mathcal{B}(\mathcal{H})$ is larger then $\mathcal{I}$ (its positive elements *may contain* states which are not normal). This leads to the Singer-Kadison ambiguity in the definitions of extension of pure states. This is relevant if one considers norm-closed subalgebras of $\mathcal{B}(\mathcal{H})$ e.g. (C^*-algebras) and their duals.

Definition 2 (*Wigner automorphisms*) Let $\mathcal{P}$ be the collection of one-dimensional orthogonal projections in the Hilbert space $\mathcal{H}$. A *Wigner automorphism* is a linear map α, *continuous in the trace topology*, such that, for each pair $P_1,\ P_2 \in \mathcal{P}$ one has

$$Tr(P_1 P_2) = Tr(\alpha(P_1)\alpha(P_2)). \tag{13}$$

Notice that, since $Tr(P_\phi P_\psi) = |(\phi, \psi)|^2$, condition (13) is the requirement that *transition probabilities be invariant under the map* α.

For Wigner automorphisms the following important theorem holds [13].

Theorem 1 [13] *If the Hilbert space has dimension at least two and if α is a Wigner automorphism, there exists (unique modulo a phase factor) an operator U_α, either unitary or antiunitary, such that for all projection operators one has*

$$\alpha(P) = U_\alpha^* \; P \; U_\alpha \tag{14}$$

We shall give an elementary proof of this theorem, following V. Bargmann. Further details can be found in [4].

A slightly different definition of state automorphism has been considered by R. Kadison. Instead of invariance of the transition probabilities it is now required that the map α can be extended to *an affine map* of $\mathcal{I}_{1,+}$, the positive trace class operators. This condition refers therefore only to properties of the states. Notice that the elements of $\mathcal{I}_{1,+}$ are the normal states of the system as defined in Axiom 1.

Definition 3 (*Kadison automorphisms*) The map $\sigma \mapsto \beta(\sigma)$ of $\mathcal{I}_{1,+}$ is a *Kadison automorphism* if, for every $0 \leq s \leq 1$ and $\sigma_1, \; \sigma_2 \in I_{1,+}$ one has

$$\beta(s\sigma_1 + (1-s)\sigma_2) = s\beta(\sigma_1) + (1-s)\beta(\sigma_2) \tag{15}$$

(i.e. β preserves the *affine structure* of the normal states).

One has for Kadison automorphisms.

Theorem 2 [6] *If β is a Kadison automorphism there exists a unitary or anti-unitary operator U_β such that for all $\sigma \in \mathcal{I}_{1,+}$*

$$\beta(\sigma) = U_\beta \; \sigma U_\beta^* \tag{16}$$

We will give a proof of Kadison's theorem based on Wigner's theorem. A different proof is in [6].

Consider now *the automorphisms on the observables*; in analogy with the classical case, we require that automorphisms preserve the algebraic structure associated to symmetric operators. Notice that the product of two symmetric operators is non-symmetric in general. A product that preserves symmetry is the *Jordan product*

$$A \circ B \equiv \frac{1}{2}(A \; B + B \; A). \tag{17}$$

This product does not satisfy associativity but it commutes with taking the adjoint and therefore it is a *natural product structure for self-adjoint operators*. Moreover it reduces to the ordinary product for commuting operators and preserves the order structure for projection operators.

The Jordan product has other interesting properties. In particular one can verify the following identities:

$$(A^2 \circ B) \circ A = A^2 \circ (B \circ A), \quad A^2 = A \circ A, \quad (A \circ B) \circ C - A \circ (B \circ C) = [B, [A, C]].$$

Moreover the Jordan product is a natural structure for the product of bilinear forms of anti-commuting operators (this is the way Jordan was led to consider anti-commuting variables).

Definition 4 (*Segal automorphisms*) The bijection $\gamma : \mathcal{B}(\mathcal{H}) \leftrightarrow \mathcal{B}(\mathcal{H})$ is called *Segal automorphism* if

$$\gamma(A \circ B) = \gamma(A) \circ \gamma(B) \qquad \forall A, B \in \mathcal{B}(\mathcal{H})_{sym} \tag{18}$$

where $\mathcal{B}(\mathcal{H})_{sym}$ is the collection of bounded symmetric operators on $\mathcal{H}$.

The following theorem of Segal [7] holds. We shall give a proof based on Wigner's theorem. The proof given in [7] is different.

Theorem 3 [7] *If γ is a Segal automorphism there exists an orthogonal projection $P \in \mathcal{B}(\mathcal{H})$, a unitary operator U on $P\ \mathcal{H}$ and an anti-unitary operator V on $(I - P)\ \mathcal{H}$ such that*

$$\gamma(A) = W\ A\ W^*, \qquad \forall A \in \mathcal{B}(\mathcal{H})_{sym} \tag{19}$$

with $W = U \oplus V$.

4 Proof of Wigner's Theorem

We give now a proof of Wigner's theorem on the implementability of Wigner automorphisms by unitary or antiunitary maps when the Hilbert space has (complex) dimension greater or equal to two. We shall follow closely the presentation given by Bargmann [2].

Denote by $[\phi]$ the ray associated to the unitary vector ϕ (the collection of vectors in $\mathcal{H}$ which differ from ϕ by a phase vector). Define $\langle [\phi], [\psi] \rangle \equiv |(\phi, \psi)|$; this is conventionally called *transition probability* from the pure state $[\psi]$ to the pure state $[\phi]$. Notice that a unitary or anti-unitary map preserves $\langle ., . \rangle$. Wigner's theorem states that *the converse is also true*.

A *symmetry* of the system is by definiton a map $T : [\phi] \mapsto [\psi]$ with the following properties:

(a) T is defined for every ray;
(b) $\langle T[\phi], [\psi] \rangle = \langle [\phi], T[\psi] \rangle \quad \forall \phi, \psi \in \mathcal{H}$;
(c) T is one-to-one;
(d) T is surjective.

With these notation Wigner's theorem may stated as follows.

Theorem (Wigner) *If the map T satisfies (a), ..., (d) there exists a map $U : \mathcal{H} \to \mathcal{H}$ such that*

$$U\phi \in T[\phi] \tag{20}$$

with the properties

(1) $U(\phi + \psi) = U\phi + U\psi$;
(2) $U(\lambda\phi) = \xi(\lambda)\phi$;
(3) $(U\phi, U\psi) = \xi((\phi, \psi))$

where for all λ either $\xi(\lambda) = \lambda$ or $\xi(\lambda) = \bar{\lambda}$. Therefore U is either linear or antilinear. A map U which satisfies (20) is said to be compatible with T.

Proof The statement is trivial when the complex dimension of $\mathcal{H}$ is one.

Let the dimension be ≥ 2. Consider three rays $[\phi_1]$, $[\phi_2]$, $[\phi_3]$ that are *not mutually orthogonal*. The number

$$\Delta([\phi_1], [\phi_2], [\phi_3]) \equiv (\phi_1, \phi_2)(\phi_2, \phi_3)(\phi_3, \phi_1) \tag{21}$$

does not depend on the choice of the phases of the three vectors and is therefore a function of the rays. Denote by γ_T the function defined on C by

$$\Delta(T[\phi_1], T[\phi_2], T[\phi_3]) = \gamma_T(\Delta([\phi_1], [\phi_2], [\phi_3])). \tag{22}$$

It is convenient to extend T from the unital rays to all rays defining

$$T(c[\phi]) \equiv cT([\phi]), \qquad c > 0. \tag{23}$$

The map so extended (still denoted by T) has the following properties

$$\langle T[\phi_1], T[\phi_2]\rangle = \langle [\phi_1], [\phi_2]\rangle, \quad |T[\phi]| = |[\phi]|. \tag{24}$$

Notice now that if $\{[\phi_1], \ldots, [\phi_N]\}$ is a collection of mutually orthogonal rays ($\langle [\phi_i], [\phi_j]\rangle = \delta_{i,j}$) then also the transformed rays are ortogonal. And if

$$[\phi] = \sum_{n=1}^{N} c_n[\phi_n], \quad c_n > 0 \quad \forall\, 1 \leq n \leq N$$

then

$$T[\phi] = \sum_{n=1}^{N} c_n' T[\phi_n], \quad |c_n'| = |c_n| > 0 \tag{25}$$

Consider now a unital ray $[\phi]$. Let $[\phi]' \equiv T[\phi]$ and choose representatives ϕ in $[\phi]$ and ϕ' in $[\phi]'$.

The unitary group acts transitively on the Hilbert space and therefore there exists a unitary map U that maps ϕ' to ϕ. We can therefore assume $\phi' = \phi$.

Every vector $u \in \mathcal{H}$ has a unique decomposition

$$u = d\phi + \zeta, \qquad (\zeta, \phi) = 0. \tag{26}$$

The strategy of the proof is to begin by constructing the map U for a generic pair ϕ, ϕ' in the case $d = 2$ making use of the properties of the action of T on the rays associated to the one-dimensional space O_ϕ of the vectors orthogonal to ϕ. We shall prove that this produces an *isometric map* V from O_ϕ to $O_{\phi'}$ which is either linear or antilinear depending on the transformation properties of Δ under T.

The transformation induced on the entire space (i.e. for an arbitrary choice of the dimension d) is then obtained by linear extension or anti-linear extension. It is easy to verify, making use of the triangular inequalities, that if $u = \phi + \zeta,\ \zeta \in O_\phi,\ |\zeta| \neq 0$, there exists *a unique* ζ' such that $\phi + \zeta' \in T[u]$. This uniqueness property *is an essential part* of the construction given by Wigner.

Denote by V the map defined by $\zeta' \equiv V\zeta$ and define

$$U(\phi + \zeta) = \phi + V\zeta \quad \zeta \in O_\phi \tag{27}$$

If $\zeta = 0$ we set $V\zeta = 0$. Notice that

$$|(\phi + Vw, \phi + Vx)|^2 = |(\phi + w, \phi + x)|^2 \quad \forall x;\ w \in O_\phi$$

implies

$$Re(Vw, Vx) = Re(w, x) \tag{28}$$

and that if (w, x) is real then $(Vw, Vx) = (w, x)$. This proves that V acts as an isometry on O_ϕ.

We must prove that the map V is linear in O_ϕ and that $V(\alpha\, x) = \xi(\alpha)V(x)$ where $|\xi(\alpha)| = |\alpha|$.

If the Hilbert space $\mathcal{H}$ has complex dimension 2, then O_ϕ has complex dimension one. In this case $V(cw) = \xi(c)w,\ |c| = 1$. Consider in O the vectors w and iw. It follows that $Re(Viw, Vw) = 0$ and therefore $\xi(i) = \pm i$ (the sign depends on the map V and therefore on T). From this the following alternative follows

$$\xi(i) = i \quad \rightarrow \quad V(cw) = cV(w), \qquad \xi(i) = -i \quad \rightarrow \quad V(cw) = \bar{c}V(w) \tag{29}$$

(notice that $V(bw) = bV(w)$ if b is real). This proves the theorem if $d = 2$.

If the Hilbert space has complex dimension greater than two, and therefore O has complex dimension greater than one, let f_1 and f_2 be orthogonal, and consider the vector $w = f_1 + f_2$. By orthogonality $Vw = f_1' + f_2'$ and $V(\alpha w) = \xi_1(\alpha)f_1' + \xi_2(\alpha)f_2'$; on the other hand $V(\alpha w) = \xi_\omega(\alpha)\ w$. It follows $\xi_1(\alpha) = \xi_2(\alpha)(= \xi_\omega(\alpha))$.

Let us remark that

$$\xi(\alpha+\beta)=\xi(\alpha)+\xi(\beta),\quad \xi(\alpha\,\beta)=\xi(\alpha)\xi(\beta),\quad \xi(\bar{\alpha})=\bar{\xi}(\alpha) \tag{30}$$

We have therefore proved that the map V has the following properties:

(a) $V(x+w)=V(x)+V(w)$;
(b) $V(\alpha x)=\xi(\alpha)\,x$;
(c) $(Vx, Vw)=\xi((x,w))\,(x,w)$

where ξ has the properties described in (30).

We can now extend V to a map U defined on $\mathcal{H}$ and satisfying properties (a), (b), (c). Define, for any complex α

$$U(\alpha\,\phi+\zeta)=\xi(\alpha)\,\phi+V\zeta,\quad \zeta\in O \tag{31}$$

If $\alpha=1$ we obtain the old definition and it easy to see that conditions (a), (b), (c) are satisfied. This concludes the proof of Wigner's theorem. ♡

It is easy to verify that the map we have constructed is the unique map with the desired properties in the following sense: in a Hilbert space of complex dimension greater than one if two additive maps U_1 and U_2 ($U(x+y)=U(x)+U(y)\ \ \forall x,\ y\in\mathcal{H}$) are compatible with a given map T between rays, then there exists a complex number $\omega,\ |\omega|=1$ such that $U_2=\omega U_1$.

5 Proof of Kadison's and Segal's Theorems

We now reduce the proofs of the theorems on Kadison and Segal automorphisms to the case of Wigner automorphism [10]. Denote as before by $\mathcal{I}_{1,+}$ the space of density matrices in $B(\mathcal{K})$

Lemma 4 *Let β be a Kadison automorphism. Then for any two-dimensional subspace $\mathcal{K}\subset\mathcal{H}$ there is a two dimensional space $\beta(\mathcal{K})$ such that $\beta(\mathcal{I}_{1,+})=\mathcal{I}_{1,+}(\beta(\mathcal{K}))$.*

Proof Recall that a convex subset F of $\mathcal{I}_{1,+}$ is a *face* if $\rho\in F,\ \omega\in\mathcal{I}_{1,+}(\mathcal{K}),\ \omega\leq k\rho,\ k>0$ implies $\omega\in F$.

$\mathcal{I}_{1,+}(\mathcal{K})$ is a face of the convex set $\Sigma(\mathcal{H})$ with the property to be the smallest face in $\mathcal{I}_{1,+}(\mathcal{H})$ that contains two given extreme points $u,\ v$. Since β is a convex automorphism it preserves the structure of faces. ♡

This Lemma extends with the same proof to all subspaces of $\mathcal{H}$.

Lemma 5 *Given a Kadison automorphism β there is a Wigner automorphism α with $\beta(P_\phi) = P_{\alpha(\phi)}$ for every one-dimensional projection P.*

Proof Since P_ϕ are extreme point in $\mathcal{I}_{1,+}(\mathcal{H})$ one has $\beta P_\phi = P_{\alpha(\phi)}$ for any one-dimensional projection and any map α on the rays. We must prove that α preserves inner products in $P\ \mathcal{H}$.

Given ϕ and ψ let $\mathcal{K}$ be the linear span of ϕ and ψ. Let β_U the map induced by the unitary operator that maps $\beta(\mathcal{K})$ into $\mathcal{K}$. Such unitary operator exists by Wigner's theorem. By composing β with the Kadison automorphism β_U we obtain that $\tilde{\beta} = \beta_U \circ \beta$ leaves $\mathcal{K}$ invariant.

By the result for the two-dimensional case, $\tilde{\beta}$ is induced by a unitary operator so there exist a unitary or anti-unitary operator $V : \mathcal{K} \to \beta(\mathcal{K})$ such that $\beta(\rho) = V\rho V^*$ for $\rho \in \sigma(\mathcal{K})$. This implies that β preserves $Tr(P_\phi P_\psi)$ i.e. $\langle \alpha[\phi], \alpha[\psi] \rangle = \langle [\phi], [\psi] \rangle$. ♡

Proof of Kadison's theorem By Wigner's theorem and Lemma 18 we can find U, unitary or anti-unitary, so that $\beta_U \cdot \beta$ is the identity on all P_ϕ, $\phi \in \mathcal{H}$. But a Kadison automorphism that leaves each extreme point fixed is the identity. ♡

We prove now the theorem about Segal automorphisms by reducing it the Wigner's theorem. We again achieve the proof trough a sequence of Lemmas.

Lemma 6 *Let γ be a Segal automorphism. Then γ is order preserving, norm preserving and takes projections into projections.*

Proof To prove that it is order preserving notice that for B positive one has $\gamma B = \gamma(\sqrt{B})\gamma(\sqrt{B})$. The map γ takes projections into projections because $\gamma(P) = \gamma(P \cdot P) = (\gamma(P))^2$.

Finally $\|A\| \pm A \geq 0$ implies $\|\gamma(A)\| \leq \|A\|$. Since γ is invertible and γ^{-1} is a Segal automorphism, $\|A\| \leq \|\gamma(A)\|$. ♡

One dimensional projection are *minimal* projection so that the order preserving map γ must take them one into another. Therefore a Segal automorphism takes one-dimensional projection into one-dimensional projections and therefore induces on the projections on one-dimensional subspaces of $\mathcal{H}$ a map α_γ such that $\gamma(P_\phi) = P_{\alpha_\gamma(\phi)}$ is a Wigner automorphism.

Lemma 7 *If γ is a Segal automorphism and $\gamma(P_\phi) = P_\phi$ holds for all vectors ϕ then γ is the identity.*

Proof Let P be any projection and denote by $R(P)$ its range. Since $\phi \in R(P)$ if and only if $P_\phi \leq P$ and since γ is order preserving, one has $R(\gamma(P)) = R(P)$ i.e. γ leaves all finite dimensional projections invariant. Since γ is continuous in norm and each element in $\mathcal{B}(\mathcal{H})$ is norm limit of finite linear combinations of finite dimensional projections we conclude that γ is the identity.

This concludes the proof of Theorem 3 about Segal automorphisms. ♡

6 Time Evolution, Continuity, Unitary Evolution

Consider now the evolution in time. We consider time-dependent Wigner or Kadison automorphisms if we refer to the evolution of the states, and consider time-dependent Segal automorphisms if we consider the evolution of the observables. We must specify in which topology we require continuity. One has the strongest result for Wigners's automorphisms.

Theorem 8 (Wigner 2) [7] *If a family of Wigner automorphisms* $t \to \alpha_t$ *is* measurable *as a function of* t *for the weak topology (i.e. for every* $\phi \in \mathcal{H}$ *the function* $(\phi, \alpha_t \phi)$ *is measurable) and if the Hilbert space* $\mathcal{H}$ *is separable it is possible to choose the collection of unitary operators* $U(t)$ *in such a way that they become a one-parameter group of unitary operators continuous in* t in the strong operator topology *such that for every projection operator* P

$$\alpha_t(P) = U(t) P U^*(t) \tag{32}$$

Analogous results hold for the automorphisms of Kadison and Segal. However in both cases one must require continuity in t *in a stronger topology* [5, 12].

Recall that the weak topology for trace-class operators is defined by the requirement that for every $A \in \mathcal{B}(\mathcal{H})$ the function $\sigma \to Tr(\sigma A)$ be continuous.

On $\mathcal{B}(\mathcal{H})$ the weak topology is defined by the requirement that the function $A \to (\psi, A\psi)$ be continuous for every $\psi \in \mathcal{H}$, i.e. it is the topology induced by the dual space.

A *weaker* topology is the weak-* topology (also called *ultraweak* or *vague*) induced by *the predual* $\mathcal{I}$. Notice that $\mathcal{B}(\mathcal{H})$ is the dual of $\mathcal{I}$ through the map $\sigma \mapsto Tr(\sigma A)$.

Every $\sigma \in I_{1,+}$ can be written as

$$\sigma = \sum_k \lambda_k P_k, \qquad \lambda_k \geq 0, \quad \sum_k \lambda_k = 1$$

where P_k are the orthogonal projections on one-dimensional subspaces.

Therefore *on bounded subsets* of $\mathcal{B}(\mathcal{H})$ the ultraweak topology is characterized by the continuity of rank-one projections. For one-dimensional projections P_ϕ one has $Tr(P_\phi A) = (\phi, A\phi)$; therefore on bounded subsets of $\mathcal{B}(\mathcal{H})$ the ultraweak topology coincides with the weak operator topology.

Notice that in $\mathcal{B}(\mathcal{H})$ one can introduce another topology, called *strong topology*, intermediate between the uniform and the weak ones. This topology is defined by the requirement that for every $\phi \in \mathcal{H}$ be continuous the map $A \to |A\phi|$. Also this topology is equivalent to the weak topology *on bounded subsets of* $\mathcal{B}(\mathcal{H})$.

For one-parameter groups of automorphisms of Kadison and Segal one has

Theorem 9 (Wigner, Kadison) *If* $t \mapsto \alpha_t$ *is a one-parameter group of Kadison automorphisms continuous in the norm topology of* $\mathcal{I}$*, there exists a one-parameter*

group of unitary operators $U(t)$*, continuous* in the strong operator topology *such that*

$$\alpha_t(\sigma) = U(t)\sigma U^*(t) \qquad \forall \sigma \in \mathcal{I}_{1,+} \tag{33}$$

An analogous result holds for Segal automorphisms.

Theorem 10 (Kadison, Segal) *If* $t \mapsto \gamma_t$ *is a one-parameter group of Segal automorphisms continuous in the strong topology of* $\mathcal{B}(\mathcal{H})$ *there exists a one-parameter group of unitary operators* $U(t)$ continuous in the strong operator topology *such that*

$$\gamma_t(A) = U(t)AU^*(t) \qquad \forall A \in \mathcal{B}(\mathcal{H}) \tag{34}$$

Finally, if one assumes that each automorphism preserves the algebraic stucture of $\mathcal{B}(\mathcal{H})$ (*and not only its Jordan product structure*) one obtains the same result under *weaker continuity properties*. This is due to the fact that one can use the properties of the algebra $\mathcal{B}(\mathcal{H})$ which we will describe in "Lecture 8: Properties of Free Motion, Anholonomy, Geometric Phase".

Theorem 11 (Kadison, Sakai [7]) *If* $t \mapsto \gamma_t$ *is a one-parameter group of automorphisms of* $\mathcal{B}(\mathcal{H})$ *continuous in* the weak operator topology*, there exists a one-parameter group of unitary operators* $U(t)$ continuous in the weak topology *such that*

$$\gamma_t(A) = U(t)AU^*(t) \qquad \forall A \in \mathcal{B}(\mathcal{H}) \tag{35}$$

We shall give a proof of Theorem 8. We don't give here a proof of Theorems [9]. One can reduce their proof to the proof of Theorem 8 as we have done in the case of maps. A different proof that relies more on the structure of $\mathcal{B}(\mathcal{H})$ as algebra of operators can be found in [7].

It is relevant to notice that in Theorem 8 only *weak measurability* is assumed, and apparently different assumptions about continuity are made in Theorems 9, 10 and 11.

It is easy to see that the assumptions of Theorems 9, 10 and 11 are comparable. Indeed the norm continuity in Theorem 11 refers to maps of $I_{1,+}$ and one can use the duality between $\mathcal{B}(\mathcal{H})$ and $I_{1,+}$ and the fact that all maps are between bounded sets. Notice also that for unitary operators the weak and strong operator topology coincide.

Indeed denoting by $\mathcal{U}$ the collection of unitaries in $\mathcal{H}$, if $\mathcal{N}_V$ is a neighborhood of $V \in \mathcal{U}$ then $\mathcal{N}'_I \equiv V^{-1}\mathcal{N}_V$ is a neighborhood of the identity. It is sufficient therefore to consider continuty in a neighborhood of the identity I. If U is unitary

$$|(U - I)\phi|^2 = 2|\phi|^2 - 2Re(\phi, U\phi) = 2Re(\phi, (I - U)\phi)$$

and convergence of the left hand side implies convergence of the right hand side.

The assumption of Wigner Theorem are at first sight much weaker, as only weak measurability of the family of maps is required. The apparently weaker assumption is sufficient because the structure of the Hilbert space is exploited in von

Neumann's Theorem (Theorem 20) which enters in the proof of Theorem 8 (and has an independent interest). This theorem makes use of the properties of unitary operators in a separable Hilbert space.

We shall now give a proof of Wigner's theorem 8. The result was later generalized by Bargmann [1] and Diximier [5] to semisimple groups.

The proofs of Theorems 9 (Kadison, Wigner) and 10 (Kadison, Segal) follow by reducing them to Theorem 8 as we have done for the case of maps. One must assume strong continuity because in the proof of thesis Theorems one does not fully exploit the Hilbert space structure. A different proof is given in [7].

The proof of Theorem 11, which assumes only *weak continuity* can be reduced to the proof of Wigner's theorem but requires results from the theory of the algebra $\mathcal{B}(\mathcal{H})$.

Definition 5 (*weak measurabilty*) A map $t \mapsto U(t)$ from the reals to the unitary operators in a separable Hilbert space $\mathcal{H}$ is called *weakly measurable* if $t \mapsto (\psi, U(t)\phi)$ is measurable for all $\phi, \psi \in \mathcal{H}$. A map $t \mapsto \phi(t)$ from the reals to $\mathcal{H}$ is called *weakly measurable* if the scalar product $(\psi, \phi(t))$ is measurable for each $\psi \in \mathcal{H}$.

Proof of Wigner's theorem 8 We shall give the proof through a sequence of lemmas.

Lemma 12

(1) Let $U_1(t)$, $U_2(t)$ be weakly measurable, then so is the product $U_1(t)U_2(t)$.
(2) If $\phi(t)$ is a weakly measurable vector-valued function and $U(t)$ is weakly measurable also $U(t)\phi(t)$ is weakly measurable.
(3) If $\phi(t)$ and $\psi(t)$ are weakly measurable, their scalar product $(\phi(t), \psi(t))$ is measurable.

Proof For point (1), let $\{\phi_n\}, n = 1, \ldots, \infty$ be an orthonormal basis. The function

$$(\psi, U_1(t)U_2(t)\psi) \equiv \sum_{n=1}^{\infty} (\psi, U_1(t)\phi_n)(\phi_n, U_2(t)\psi_n) \tag{36}$$

is measurable as limit of measurable functions. Statements (2) and (3) are proved similarly. ♡

Lemma 13 *There exists a unitary operator valued function $U_{\phi,\eta}$ defined on pairs of unit vectors such that:*

(1) if ψ is orthogonal both to ϕ and to η then $U_{\phi,\eta}\psi = \psi$;
(2) $U_{\phi,\eta}\eta = \phi$;
(3) if $\phi(t)$ and $\eta(t)$ are weakly measurable, then $U_{\phi(t),\eta(t)}$ is weakly measurable.

Proof If $\phi = a\eta$ for some $a \in C$ define $U_{\phi,\eta} = 1 + (a-1)P_\eta$. If ϕ is orthogonal to η define $U_{\phi,\eta}$ by

$$U_{\phi,\eta}\psi = \psi + (\phi, \psi)(\eta - \phi) + (\eta, \psi)(\phi - \eta). \tag{37}$$

If $(\phi, \eta) \neq 0$ write $(\phi, \eta) = |(\phi, \eta)|e^{i\theta}$. One verifies that with these definitions (1) and (2) hold, one has $U_{\phi,\eta}\phi = e^{-2i\theta}\eta$ and measurability holds. ♡

Denote as before by $[\phi]$ the ray associated to the vector ϕ.

Lemma 14 *If $t \mapsto \alpha_t$ is a measurable family of Wigner automorphisms with $\alpha_t\alpha_s = \alpha_{t+s}$ then each α_t is induced by a unitary operator. Moreover for each ϕ one can choose $\eta(t)$ weakly measurable such that $\alpha_t([\phi]) = [\eta(t)]$.*

Proof Since $\alpha_t = (\alpha_{\frac{t}{2}})^2$ it follows from Wigner's theorem that α_t is induced by a unitary operator. Let $\{\psi_k\}$ be an orthonormal basis. Let Ξ_k be the interval

$$([\psi_k], \alpha_t[\phi]) \neq 0 \quad \forall t \in [0, \tau] \tag{38}$$

Each Ξ_k is measurable, and we can choose $\eta(t)$ measurable in each Ξ_k. Choose $\eta(t)$ on Ξ_k such that $(\psi_k, \eta_t) > 0$ and $\alpha_t([\phi]) = [\eta(t)]$.

Let $f_j(t) = (\psi_j, \eta(t))$, j/k. We must show that each $f_j(t)$ is measurable. This follows because, as in the proof of Wigner's theorem, f_j is determined by $|(\psi_j, \eta(t))|$, $|(\psi_j + \psi_k, \eta(t))|$, $|(\psi_k + i\psi_k, \eta(t))|$. ♡

Lemma 15 *Let the Hilbert space $\mathcal{H}$ have dimension two and let $\phi \in \mathcal{H}$ be fixed. Let $t \mapsto \alpha_t$ be measurable and induced by a unitary operator with $\alpha_t([\phi]) = [\phi]$ for any t. Choose $U(t)$ inducing α_t so that $U(t)\phi = \phi$. Then $U(t)$ is measurable in t.*

Proof $\mathcal{H}$ is isomorphic to C^2. Let $[\phi]$ correspond to $(0, 1)$. α_t corresponds then to a rotation by an angle $\theta(t)$ in the plane. A unitary rotation is measurable. ♡

Proof of Theorem 8 By Lemma 14 one can find α_t measurable so that $[\eta(t)] = \alpha_t[\phi]$.

Let

$$\hat{\alpha}_t = \alpha_t U_{\phi,\eta(t)}\alpha_t. \tag{39}$$

Notice that $U_{\phi,\eta(t)}$ is a unitary map that acts on ϕ as α^{-1}. Then $\hat{\alpha}_t$ is measurable and $\hat{\alpha}_t[\phi] = [\phi]$. Choose $\hat{U}(t)$ inducing $\hat{\alpha}_t$ and such that $\hat{U}(t)\phi = \phi$. We need to prove that $\hat{U}(t)$ is measurable. It is sufficient to prove that $\hat{U}(t)\psi$ is measurable if $\psi \perp \phi$.

Choose $\xi(t)$ measurable so that $\hat{\alpha}_t([\psi] = [\xi(t)]$ and let $\beta_t = \alpha_t U_{\psi,\xi}\hat{\alpha}_t$. By construction β_t is a measurable family which leaves invariant the set $\{[\psi], [\phi]\}$. Therefore there is a measurable $V(t)$ on $\{\psi, \phi\}$ such that $V(t)\phi = \phi$. Since $U_{\psi,\xi}(t))\hat{U}(t)\phi = \phi$ we conclude that $\hat{U}\psi = U^{-1}_{\phi,\xi(t)}V(t)\phi$ so $\hat{U}(t)\phi$ is measurable. ♡

Consider now a one parameter group α_t, $t \in R$ of Wigner automorphisms, continuous in the strong topology. We recall that an operator is unitary resp. antiunitary if it satisfies $U^*U = UU^* = I$ and is linear resp. antilinear. In particular if a strongly continuous group is made of operators each of which is either unitary or antiunitary, it follows that it is a group of unitary operators.

By construction the operators U_{t+s} and $U_t \cdot U_s$ induce the same automorphism. From Wigner's theorem we conclude that there exists a continuous phase function $\omega(t,s) = e^{i\xi(t,s)}$ such that

$$U_{t+s} = \omega(t,s) U_t U_s \tag{40}$$

If all operators are unitary it follows from the associativity property that the function ω must satisfy the *cocycle condition*

$$\omega(t,s)\omega(t+s,w) = \omega(t,s+t)\omega(s,w) \tag{41}$$

It is therefore by definition a *Borel multiplier* on R.

Definition 6 (*multiplier*) A multiplier ω is a function from $R \otimes R$ to an algebra $\mathcal{A}$ that satisfies

$$\omega(a,b)\omega(a+b,c) = \omega(a+b,c)\omega(b,c), \qquad a,b,c \in \mathcal{A}$$

Given a measurable invertible function $\lambda : R \to \{\alpha,\ |\alpha| = 1\}$ we define a multiplier $\partial\lambda$ by

$$\partial\lambda(t,s) = \lambda(t+s)\lambda(t)^{-1}\lambda(s)^{-1} \tag{42}$$

Theorem 16 [12] Every Borel multiplier on the algebra R is of the form $\omega = \partial\lambda$ for some function λ. Therefore a function $\xi(t,s)$ with the property (41) always exists.

As a consequence of this theorem, since all the unitary operators U_t are definded modulo a phase factor, it is always possible to modify the choice of ξ as follows

$$\xi(t,s) \to \xi(t,s) - \beta(t+s) + \beta(t) + \beta(s) \tag{43}$$

where $\beta(t)$ is an arbitray continuous function. Therefore if there exists a function $\gamma(t)$ for which

$$\xi(t,s) = \gamma(t+s) - \gamma(t) - \gamma(s) \tag{44}$$

then one can choose a representative U_t in such a way that

$$U_{t+s} = U_t \cdot U_s \tag{45}$$

The symbol $\partial\lambda$ is motivated by the connection with the cohomology groups. Theorem 16 is the statement that the first cohomology group on the real line with coefficients in the circle group is trivial. Remark that while R has no trivial multiplier, $R^2 \equiv C^1$ does have non trivial multipliers. For example

$$\omega(a+ib, c+id) = e^{\frac{i}{2}(ad-bc)} \tag{46}$$

is the multiplier associated to the Heisenberg group.

Proof of Theorem 16 We shall give the proof of Theorem 16 through a sequence of Lemmas. It is easy to see that there is no loss of generality in requiring $\omega(0, s) = \omega(s, 0) = 1$ for all s.

Lemma 17 *Without loss of generality we may suppose* $\omega(t, -t) = 1 \ \ \forall t$.

Proof By the definition it follows

$$\omega(t, -t)\omega(0, t) = \omega(t \cdot 0)\omega(-t, t) \tag{47}$$

so that $\omega(t, -t) = \omega(-t, t)$. Define $\lambda(t) = \sqrt{\omega(t, -t)}$ where we choose the argument in $[0, \pi)$. Then

$$\partial\lambda(t, -t) = \lambda(0)\lambda(t)^{-1}\lambda(-t)^{-1} = \omega(t, -t)^{-1}$$

so that defining $\tilde{\omega} = \omega \circ \partial\lambda$ one has $\tilde{\omega}(t, -t) = 1$. ♡

Notice now that for any multiplier $\omega \in R$ if $\mathcal{H} = L^2(R, dm)$ (dm a Lebesgue measure) defining $(U(t)f)(s) = \omega(t, s)f(t + s)$ the family $U(t)$ satisfies $U(t)U(s) = \omega(t, s)U(t + s)$

Lemma 18 *If* ω *is a multiplier in R, then* $\omega(t, s) = \omega(s, t) \ \ \forall t, s \in R$.

Proof Define $q(t, s) = \frac{\omega(t,s)}{\omega(s,t)}$ and let U some representation with multiplier ω.

Then $U(t)U(s)U(t)^{-1} = q(t, s)U(s)$. From the properties of $\omega(t, s)$ one derives $(\omega(t, s)U(t + s))^{-1} = \omega(t, s)^{-1}U(-t - s)$ and hence $q(t + w, s) = q(t, s)q(w, s)$.

Measurability of $q(t, s)$ implies then $q(t, s) = e^{2i\pi t}f(s)$ for some measurable function f. Since $q(t, t) = 1$ it follows $f(t) = t^{-1}n(t)$ where $n(t)$ is an integer. Continuity allows to deduce $n(t) = 0$ i.e. $q(t, s) = 1 \ \forall t, s$. ♡

Lemma 19 *For any multiplier* $\omega \in R$ *there is an irreducible family of unitary operators in some Hilbert space* $\mathcal{H}$ *such that* $U(t + s) = \omega(t, s)U(t)U(s)$.

Proof If we knew that $U(t)$ is continuos, to prove the lemma we could use the theory of locally compact groups. Since we did not assume continuity but rather we want to prove it, we use an argument which is used in the theory of compact groups.

We say that a function F on R is ω-positive definite if $F \in L^\infty(R)$ and for all $f \in L^1(R)$

$$\int \bar{f}(t)f(s)F(t - s)\omega(-t, s)dtds \geq 0 \tag{48}$$

On these functions consider the inner product

$$(f, g) = \int \bar{f}(t)g(s)F(t - s)\omega(-t, s)dtds$$

It is easy to see that the map $(U(t)f)(s) = \omega(t, s-t)f(s-t)$ satisfies

$$U(t)U(s)f = \omega(t,s)U(t+s)f \quad U(0) = I, \quad (f, U(t)g) = (U(-t)f, g) \tag{49}$$

so they define a ω-representation.

The set of ω-positive functions is a compact convex set of L^∞ in the weak-* topology of L^1 so there are extremal points. Such extremal points lead to irreducible representations. In particular ω can be chosen to be a function with values in S^1. ♡

We can now conclude the proof of Theorem 16.

Proof of Theorem 16 By Lemma 18 $U(t)$ is a commuting family so the irreducible representations given by Lemma 15 is one-dimensional (Schur's Lemma). Thus the representation is multiplication by λ and $\lambda(t)\lambda(s) = \omega(t,s)\lambda(t+s)$. ♡

We have proved that the map $t \to U(t)$ is weakly measurable and can be made to satisfy $U(t)U(s) = U(t+s)$. The next step is to prove that it is strongly continuous.

This is the content of von Neumann's theorem.

Theorem 20 [8, 12] *Let $t \mapsto U(t)$ be a* weakly measurable *map from the real axis to the unitary operators on a separabe Hilbert space $\mathcal{H}$. Suppose $U(t)U(s) = U(t+s)$. Then $t \to U(t)$ is* strongly continuous.

Proof Since the $U(t)$ are bounded it suffices to find a total subset (subset whose linear combinations are dense) of ϕ's such that $t \mapsto U(t)\phi$ is continuous. Choose an orthonormal subset $\{\phi_n\} \subset \mathcal{H}$ and define a conjugate linear functional of norm $\leq a$ by

$$\eta \mapsto \int_0^a (\eta, U(t-n)\phi_n)dt \equiv f_{n,a}(\eta), \qquad \eta \in \mathcal{H} \tag{50}$$

By Riesz theorem there is a vector $\phi_n(a)$ such that $(\eta, \phi_n(a)) = f_{n,a}$. By construction the vector $\phi_n(a)$ can be written

$$\phi_n(a) = \int_0^a U(t)\phi_n dt$$

A simple argument proves that

$$U(s)\phi_n(a) = \int_s^{a+s} U(t)\phi_n dt$$

and therefore

$$\|(U(s) - U(w))\phi_n(a)\| \leq \|U(t)\int_s^w U(\tau)\phi_n d\tau\| + \|\int_{a+s}^{a+w} U(t)\phi_n dt\| \leq 2|s-w| \tag{51}$$

It follows that $t \mapsto U(t)\phi_n(a)$ is continuous for all n, a. The collection $\phi_n(a)$ for all n and a is a total set. To prove that this set is total we show that if ψ is orthogonal

to all elements of this set, then $\psi = 0$. In fact for each n one has $(\psi, U(t)\phi_n) = 0$ a.e. and therefore there must exist t_0 such that $(\psi, U(t_0)\phi_n) = 0$ for each value of n. It follows that $U(t_0)^{-1}\psi = 0$ and therefore $\psi = 0$. ♡

7 Time Evolution: Structural Analogies with Classical Mechanics

The analysis presented so far shows that, under mild conditions and reasonable assumptions, time evolution in Quantum Mechanics is represented by a one-parameter group of unitary operators which generates a one-parameter group of automorphisms of the observables (Heisenberg representation) or of the states (Schrödinger representation).

In Classical Mechanics the dynamics is given by a vector field which in Hamiltonian Mechanics is derived from a real function on phase space (the Hamiltonian) which is *the infinitesimal generator*. The following theorem indicates that there is a corresponding structure in Quantum Mechanics.

Theorem (Stone) [11] *The map $t \mapsto U(t)$ provides a weakly continuous representation of the additive group R^+ through unitary operators on a Hilbert space $\mathcal{H}$ if and only if there exists a self-adjoint operator H such that if $\phi \in D(H)$ one has*

$$i\frac{dU(t)}{dt}\phi = HU(t)\phi \tag{52}$$

By analogy with the notation used in the classical case the operator H is called *hamiltonian* of the systems. The reciprocal is also true: if $U(t)$ is a one-parameter group of unitary operators which satisfies (52) then (14), (16), (19) are the one-parameter group of automorphisms that describe the dynamics generated by the *Quantum* hamiltonian H.

Of course the axioms cannot indicate which is the operator that has to be chosen to describe the dynamics of a specific system.

Definition 7 (*adjoint*) Let A be a closed symmetric operator on a Hilbert space $\mathcal{H}$; the adjoint of A is the operator A^* whose domain of definition $D(A^*)$ is the collection of elements $\phi \in \mathcal{H}$ such that $(\phi, A\psi)$ is continuous in ψ in the topology of $\mathcal{H}$. Riesz's representation theorem implies that there is a vector ξ in $\mathcal{H}$ such that $(\phi, A\psi) = (\xi, \psi)$ for all $\psi \in D(A)$. On its domain A^* is defined by $A^*\phi = \xi$.

Definition 8 (*self-adjoint*) An operator is called *self-adjoint* if it coincides with its adjoint (*same domain* and same action).

In the next Lecture we shall give some basic definitions and theorems related to the theory of Hilbert space operators. We remark here that if the operator A is densely defined and bounded the domain of A^* is the entire space $\mathcal{H}$. If in addition A

is symmetric then it is closable and its closure $\bar{A}$ exists and is a selfadjoint operator. In this case the operator is called *essentially self-adjoint*.

The following Theorem, which will be proved in the next Lecture, characterize the self-adjoint operators

Theorem (spectral theorem) *The operator A on the Hilbert space*$\mathcal{H}$ *is self-adjoint if and only if there exists a measure space* (X, μ), *a measurable function* f_A on X *and an invertible isometry* U *between* $\mathcal{H}$ *and* $L^2(X, \mu)$ *such that*

$$AU = Uf_A.$$

where f_A. *stands for the operator of multiplication by* f_A.

If the operator H exists, one writes

$$U(t) = e^{-iHt}$$

where the right hand side is defined through functional calculus. Notice that in the representation in which the Hilbert space is realized as $L^2(M, d\mu)$, where μ is a measure on the (locally compact) space M, and the operator H is represented by multiplication by the function $f_H(m)$ one has

$$(U(t)\phi)(m) = e^{-itf_H(m)}\phi(m), \qquad m \in M, \quad \phi \in L^2(M, d\mu).$$

Denoting by Π_ϕ the orthogonal projection on $\phi \in \mathcal{H}$, (52) takes the form (on a suitable domain)

$$i\frac{d}{dt}\Pi_{\phi_t} = [H, \Pi_{\phi_t}] \tag{53}$$

In general, if the operator B leaves the domain of H invariant, (53) implies

$$i\frac{d}{dt}\gamma_t(B) = [H, \gamma_t(B)] \tag{54}$$

where $B \to \gamma_t(B)$ is the map dual to $\phi \to U(t)\phi$. Formally

$$\gamma_t(B) = U(t)BU(t)^*$$

The identity $[H, AB]\phi = [H, A]B\phi + A[H, B]\phi$ is valid if all terms are well defined. The map

$$B \to [H, B] \tag{55}$$

is linear and satisfies all requirements to be *derivation*; it is therefore *structurally homeomorphic* to a vector field.

Dirac was among the first to stress the homeomorphism, as algebraic structures, of Poisson Brackets in Classical Mechanics with the commutator on $\mathcal{B}(\mathcal{H})$ and to

suggest a natural way to write the generator of the dynamics in Quantum Mechanics for systems which have a classical Hamiltonian analogue

$$H_{cl}(q,\, p) = \frac{1}{2}p^2 + V(q), \qquad q,\ p\ \in R^n \tag{56}$$

where $V(q)$ is a sufficiently regular function.

Dirac, Schrödinger and Pauli stressed the importance of finding self-adjoint operators $\hat{q}_k,\ \hat{p}_i$ which satisfy, on a suitable domain,

$$\{\hat{q}_i\,, \hat{q}_k\} = \{\hat{p}_i\,, \hat{p}_k\} = 0, \qquad \{\hat{q}_i\,, \hat{p}_k\} = \delta_{i,k} I \tag{57}$$

There is a "natural" representation of (57) in $L^2(R^n)$ in which $\hat{q}_k$ is multiplication by the coordinate x_k and $\hat{p}_k$ is the differential operator $i\frac{\partial}{\partial x_k}$ (we have here used units in which $\hbar = 1$).

If the Hamiltonian separates as in (56) Heisenberg's equations take then the form, in the representation in which $\hat{p}_k \equiv i\frac{\partial}{\partial q_k}$

$$\frac{d\,\hat{p}_k}{dt} = \frac{\partial V(q)}{\partial q_k}\Big|_{q_k \to \hat{q}_k}, \qquad \frac{d\hat{q}_k}{dt} = \hat{p}_k \tag{58}$$

The corresponding dual equation is Schrödinger equation

$$i\frac{\partial}{\partial t}\phi(x,t) = -\frac{1}{2}\Delta\phi(x,t) + V(x)\phi(x,t) \tag{59}$$

Remark that Schrödinger arrived at (59) following the ideas of De Broglie and having in mind the formulation of Classical Mechanics through the Hamilton-Jacobi equation.

One can then define a generator (in the sense of Stone's Theorem)

$$H = \frac{1}{2}\hat{p}^2 + V(\hat{q}) \tag{60}$$

The operator $V(\hat{q})$ is defined through the Spectral Theorem and functional calculus.

If the classical Hamiltonian does not have the form (56), and in general *it is not* the sum of a function that depends only on one half of each pair of canonical variables and a function that depends only on the complementary set of conjugate variables, the prescription for constructing a quantum generator is not straightforward nor unique, due to the non-commutativity of the operators $\hat{q}_k$ and $\hat{p}_k$.

In these Lectures we shall discuss several prescription for *quantization*. Of course one must verify also that the operator H be self-adjoint; notice that the sum of two self-adjoint unbounded operators A and B is a priori defined only on $D(A) \cap D(B)$ and this domain may be too small to lead to the proof that the sum is a self-adjoint operator.

For general functions a and b

$$[a(\hat{p}), b(\hat{q})] \neq \sum_{i,k} [\hat{q}_k, \hat{p}_i] \left(\frac{\partial a}{\partial q_k} \frac{\partial b}{\partial p_i} \right)_{q_k \to \hat{q}_k \; p_i \to \hat{p}_i} \tag{61}$$

If the operator H is unbounded, as is the case in (57), the analysis we have given requires further justification. In particular we should prove that equation (60) has a unique solution which is global in time for every initial datum and that the evolution is unitary. In Lecture 12 we shall give conditions on V that provide global existence, uniqueness and unitarity.

In the following Lectures we shall also discuss the case of a group of automorhisms $\alpha_g, \; g \in G$ of the pure states, where G is a Lie group. In particular we will state the conditions under which there exists a continuous unitary representation U_g of G for which $\alpha_g(P) = U_g \; P \; U_g^*$. We shall see that this is always possible [5] if the group is semisimple (it contains no invariant abelian subgroup). A typical example is the group of rigid motions in R^d or the Lorentz group. The Galilei group is not semisimple since it has time translation as abelian invariant subgroup.

8 Evolution in Quantum Mechanics and Symplectic Transformations

The structural isomorphism between the commutator of operators on a complex Hilbert space $\mathcal{H}$ and Poisson brackets of functions on classical phase space leads to *a formal analogy* between evolution equations in Quantum Mechanics and one-parameter groups of symplectic transformations.

According to this analogy the evolution in Quantum Mechanics is represented by unitary transformations which leave invariant the antisymmetric form $\langle \phi, \psi \rangle \equiv Im(\phi, \psi)$.

If the complex space $\mathcal{H}$ is regarded as direct sum of two real Hilbert space

$$\mathcal{H}^0 = \mathcal{H}_r \oplus \mathcal{H}_i \tag{62}$$

with scalar product

$$(\phi, \psi) = \langle \phi_r, \psi_r \rangle \; + \; i \langle i \phi_r, \psi_i \rangle, \quad \phi = \phi_r + i \phi_i \tag{63}$$

and the complex structure of H is regarded as symplectic structure of $\mathcal{H}^0$, then quantum evolution corresponds *formally* to a one-parameter group of symplectic transformations *in an infinite-dimensional space*.

The *formal* generator is associated with the quadratic form

$$K(\psi) = \frac{1}{2}(\psi, -\Delta\psi) + (\psi, V\psi), \qquad \psi = \psi_r + i\psi_i \tag{64}$$

Indeed, *formally*,

$$i\frac{d\psi}{dt} = J\frac{\partial K}{\partial \bar{\psi}}, \qquad \psi = \{\psi_r + i\psi_i\}, \quad \bar{\psi} = \{\psi_r - i\psi_i\} \tag{65}$$

This analogy may be misleading. One must notice that K is *only formally differentiable* (for the Frechet structure obtained regarding $\mathcal{H}$ as a Banach space).

On the Hilbert space $\mathcal{H}$ one can find *other symplectic structures* which can be associated to a quantum Hamiltonian H providing equations which correspond to classical field theories (often called *hydrodynamic* since they are similar to field equations in hydrodynamics).

For example in $L^2(R^1)$ one can use as simplectic form the antisymmetric operator $J \equiv \frac{\partial}{\partial x}$. This operator is closable but unbounded; its spectrum is continuous and covers the entire imaginary axis and therefore it does not admit a bounded inverse. The symplectic form defined by J is singular, and therefore it must be handled with some care.

It can be verified that with this symplectic form the KdV (Kortweg—de Vries) equation

$$\frac{\partial u}{\partial t} + \frac{\partial^3 u}{\partial x^3} + u\,\frac{\partial u}{\partial x} = 0 \tag{66}$$

is Hamiltonian and

$$H(u) = \int \left[\frac{1}{2}\frac{\partial u}{\partial x}^2 + \frac{1}{6}u^4(x)\right] dx \tag{67}$$

In this example the definition of Poisson brackets require functional derivatives and must be treated with care.

The expression becomes somehow less formal when the same system is restricted to $[0, 2\pi]$ with periodic boundary conditions. By discrete Fourier transform the system can be realized in $l^2(ZZ)$ and the symplectic form has a one-dimensional null space (but remains unbounded). Writing $u(x) = \sum_{-\infty}^{\infty} u_n e^{inx}$ one can choose as canonical variables $q_n \equiv \frac{u_n}{n}$, $p_n \equiv \frac{u_{-n}}{n}$, $n > 0$. u_0 is a constant of motion and can be chosen to be zero. With this choice the Poisson bracket is

$$\{F, G\} = \frac{i}{2\pi}\sum_{n=1}^{\infty} \frac{\partial F}{\partial q_n}\frac{\partial G}{\partial p_n} - \frac{\partial F}{\partial p_n}\frac{\partial G}{\partial q_n} \tag{68}$$

The classical Hamiltonian system associated to the KdV equation is *completely integrable*: it admits a complete system of integrals of motion I_k, $k \in Z$ in involution (since there are infinitely many degrees of freedom, the term *complete* must be

taken with care). For a dense set of initial data $\phi_0(x) \in L^2(0, 2\pi)$ the solution $\phi(t, x) \in L^2(S^1)$ can be written as

$$\phi(t, x) = F_{\phi_0}(\{I_k\}, \{\theta_k(t)\}) \qquad F_{\phi_0}(\{I_k\}, \{\theta_k(0)\}) = \phi(t, 0) \tag{69}$$

where F_{ϕ_0} is a function of the action variables I_k and of the canonically associated angles, and the latter satisfy linear differential equations with constant coefficients.

These initial data are dense in $L^2(0, 2\pi)$ but they have in general special regularity properties as functions of x; therefore the solution of (69) as completely integrable system with an infinite number of degrees of freedom cannot in general be obtained by standard methods of functional analysis. The quantum analog of a system which satisfies the KdV equation has not been treated so far.

9 Relative Merits of Heisenberg and Schrödinger Representations

We have seen in "Lecture 2: Elements of the history of Quantum Mechanics II" that an important point in the *mathematical structure* of Quantum Mechanics is the *mathematical* equivalence between the *Schrödinger representation* and the *Heisenberg representation*. In the latter the choice of some concrete Hilbert space *is not needed* and the observables are represented by *operators*. In the former the Hilbert space $\mathcal{H}$ is taken to be $L^2(R^N, dx)$ where N is the number of degrees of freedom (defined to be the number of degrees of freedom of the corresponding classical system).

Although *all separable Hilbert spaces are equivalent* each representation has additional structures: the category of differentiable functions can be defined in the representation of the Hilbert $\mathcal{H}$ as square integrable functions on a manifold but it is difficult to define it in the realization of $\mathcal{H}$ as $l^2(Z)$.

In this sense the specific choice that leads to Schrödinger's Quantum Mechanics is *an essential part* of this theory, that has no counterpart in Heisenberg's formulation. The possibility to make use of the theory of partial differential equation and in general of the theory of functions in the Euclidian space (or on manifolds) makes the Schrödinger representation best suited for the analysis of the evolution of the states (represented by functions) and explains why in non relativistic Quantum Mechanics one mostly uses this representation.

In a representation which uses the specific realization of the Hilbert space as space of functions on the configuration space of a classical system one can meaningfully ask questions related to the topology of the configuration space.

Moreover one can ask questions about the *relative phases* of two functions that represent a state (each state is defined by a ray); we shall see an instance of this in the definition of *Berry phase* and in the *adiabatic approach*. Schrödinger's representation also leads to *entanglement and superposition*, notions that are more difficult to introduce in the Heisenberg representation.

Schrödinger's and Heisenberg's representations share the property to admit the existence of a collection of operators $\hat{q}_k,\ \hat{p}_h,\ k, h = 1, \ldots, N$ satisfying (at least on a suitable domain) the *canonical commutation relations* (*c.c.r.*)

$$[\hat{q}_h, \hat{p}_k] = i\delta_{h,k}, \qquad [\hat{q}_k, \hat{q}_h] = 0 = [\hat{p}_h, \hat{p}_k] \qquad h, k = 1, \ldots, N \tag{70}$$

It is natural therefore to inquire whether all the representations of the *canonical relations* (70) are equivalent i.e. whether if $\pi(\hat{q}_h, \hat{p}_k)$ and $\pi'(\hat{q}_h, \hat{p}_k)$ are two irreducible representations on Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, there exists an invertible isometry $V : \mathcal{H} \to \mathcal{H}'$ such that

$$\pi'(\hat{q}_h, \hat{p}_k)V = V\pi(\hat{q}_h, \hat{p}_k) \tag{71}$$

Under some regularity assumptions we shall prove that for finite N given two (irreducible) representations of the *c.c.r.* (70) (more precisely an integrated version of them) on two separable Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ there is an isometric invertible map from $\mathcal{H}_1$ to $\mathcal{H}_2$ intertwining the two representations.

This equivalence of the representations *no longer holds* for systems with an infinite number of degrees of freedom ($N = \infty$). Therefore in the analysis of systems with an infinite number of degrees of freedom (Quantum Statistical Mechanics and Relativistic Quantum Field Theory) *the choice of a representation is mandatory*. This is a major difficulty because the interactions that one considers *formally* cannot in general be described *mathematically* within any one of the representations. Neglecting this fact leads to the problems that plague the formal application of perturbation theory to these systems.

If $N = \infty$ it is (formally) easier to work with Heisenber's algebraic structure; it is more difficult to use Schrödinger's formalism because an analogue of Lebesgue's measure does not exist in R^∞ and the theory of partial differential equations in Banach spaces (i.e. for functions of infinitely many variables) is still in its infancy. For this reason when treating the quantization of systems with an infinite number of degrees of freedom Heisenberg's *algebraic formalism* is mostly used (Dirac quantization procedure, Fock space analysis) [3].

For systems with infinite number of degrees of freedom the lack of function-analytic instruments in the Heisenberg representation leads in general to use perturbation theory. The results are then expressed as series in a small parameter with little control on its convergence.

On the other hand one can replace Lebesgue measure by Gauss measure, which is a *probability measure* and can be defined on Banach spaces of functions of infinitely many variables. One can make use of methods of functional analysis and probabilty theory and set up a generalization of Schrödinger representation to achieve a quantization of systems with an infinite number of degrees of freedom and prove, in special cases, the existence of a Quantum Dynamics in infinite-dimensional function spaces.

Notice in infinite dimensions Gaussian measures with different covariances *are not equivalent* (as probability measures).

Apparently Jordan was the first to conceive this strategy, but at that time the theory of integration on function space was not developed enough.

References

1. Bargmann, V. (1954). *Annals of Mathematics*, *59*, 1–46.
2. Bargmann, V. (1964). *Journal of Mathematical Physics, 5*, 862.
3. Born, M., & Fock, V. (1928). *Zeitschrift fur angewandte Mathematik und Physik*, *51*, 165–169.
4. De Vito, G., Kahti, E., & Levrero, P. J. (1998). *The theory of symmetric actions in quantum mechanics*. London: Springer.
5. Diximier, J. (1969). Les Algebres de von Neumann Gauthier-Villars.
6. Kadison, R. (1965). *Topology*, *3*(Suppl. 2), 178–198.
7. Kadison, R., & Ringrose, M. (1967). *Communications in Mathematical Physics*, *1*, 32–63.
8. Mackey, G. (1963). *The mathematical foudations of quantum mechanics Benjamin*.
9. Segal, I. (1954). *Annals of Mathematical, 59*, 1–46.
10. Simon, B. *Bargmann Festschrift* Princeon 76.
11. Stone, M. (1929). *Proceedings of the National Academy of Sciences, 15*, 198–200.
12. von Neumann, J. (1932). *Mathematische Grundlagen der Quantum Mechanik*. New York: Springer.
13. Wigner, E. (1939). *Annals of Mathematics*, *40*, 149–204.

Lecture 6: Operators on Hilbert Spaces I; Basic Elements

In this Lecture we give a brief review of basic elements and basic terminology of the theory of linear operators on a Hilbert space $\mathcal{H}$.

Definition 1 (*graph*) Let A be an operator on a Hilbert space $\mathcal{H}$ with domain $D(A) \subset \mathcal{H}$. The *graph* of A (denoted by $\Gamma(A)$) is the ordered pair $\{\phi,\ A\phi\}$.

The operator A is *closed* if $\Gamma(A)$ is a closed subset of $D(A) \oplus \mathcal{H}$ in the graph topology given by $|\psi| + |A\psi|$ Here $|.|$ denotes the Hilbert space norm. The operator A is *closable* if it admits a closed extension.

We give an example of *non-closable operator*.

Let $\{\phi_n\}$ an orthonormal complete basis in $D(A)$, and let $f \in \mathcal{H}$, $f \neq 0$. Assume that f is not a finite linear combination of the ϕ_k. Let D be the set composed of f and of the finite linear combinations of the ϕ_k and set $D(A) = D$ with the following action

$$A\left(af + \sum_{k=1}^{N} d_k \phi_k\right) = af \tag{1}$$

By construction $\{f,\ f\} \in \Gamma(A)$ and $A(\sum_1^N d_k\phi_k) = 0, \quad \forall\{d_k\}, \quad \forall N$.

Due to the completeness of the basis $\{\phi_k\}$ there exist $\{\delta_k\}$ with $\sum_k |\delta_k|^2 < \infty$ such that $f = lim_{N\to\infty}\delta_k\phi_k$. Therefore $\{f, 0\} \in \Gamma(A)$. But then $\Gamma(A)$ is not a graph since it associates to an element of $\mathcal{H}$ two elements of $\mathcal{H}$ namely 0 and f. ♡

Definition 2 (*symmetric operator*) The operator A is *symmetric* if $(\psi, A\phi) = (A\psi, \phi)$ for all $\phi, \psi \in D(A)$.

Theorem 1 (closed graph theorem) *An operator A defined on the entire Hilbert space $\mathcal{H}$ is closed iff it is bounded.*

We do not give the proof of this important (and non trivial) theorem. It can be found in any good book on Functional analysis, e.g. [2, 6]. We give instead the proof of the

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_6

Corollary *Let A be symmetric and let $D(A) = \mathcal{H}$. Then A is bounded*

Proof We prove that if A is symmetric and defined on all $\mathcal{H}$, then it is closed. Let the sequence $\{x_N,\ Ax_n\}$ converge to $\{x,\ y\}$. We must prove that $x \in D(A)$ and that $Ax = y$. The first statement is trivial, as $D(A) = \mathcal{H}$. Let $z \in \mathcal{H}$. To prove the second statement remark that, using the symmetry , one has

$$(z, y) = \lim_{n\to\infty} (z, Ax_n) = \lim_{n\to\infty} (Az, x_n) = (Az, x) = (z, Ax)$$

♡

Definition 3 (*adjoint operator*) The *domain* of the *adjoint* A^* of the operator A are the elements $\phi \in \mathcal{H}$ such that $(\phi, A\psi)$ is continuous in ψ in the topology of $\mathcal{H}$.

By the Riesz representation theorem it follows that if $\phi \in D(A)$ there exist ξ_ϕ such that $(\psi, A\phi) = (\xi_\phi, \phi)$. The element ξ_ϕ is uniquely defined if $D(A)$ is dense in $\mathcal{H}$. Otherwise it is unique as element of the subspace $\bar{D}(A)$. On this domain the action of A^* is defined by

$$(A^*\psi, \phi) = (\psi, A\phi) \qquad \forall \phi \in D(A) \tag{2}$$

In general there is no relation between $D(A)$ and $D(A^*)$. If A is symmetric and closed, $D(A) \subset D(A^*)$. One can also verify in this case that $(A^*)^*$ is the closure of A in the graph norm.

We leave as an exercise the proof of the following Lemma.

Lemma 2

(i) If A is densely defined the operator A^* is closed.
(ii) The operator A is closable if $D(A^*)$ is dense in $\mathcal{H}$. In this case $\bar{A} = (A^*)^* \equiv A^{**}$.
(iii) If A is closable, then $(\bar{A})^* = A^*$.

Corollary 3 *It follows from (ii) that if A is symmetric then it is closable.*

Proof If A is symmetric one has $D(A) \subset D(A^*)$. Therefore if $D(A)$ is dense also $D(A^*)$ is dense. ♡

Definition 4 (*resolvent set*) The *resolvent set* $\rho(A)$ is by definition the set of complex numbers λ for which $(\lambda I - A)$ has a bounded inverse.

The resolvent set is open. In fact if $\lambda_0 \in \rho(A)$ and if ϵ is sufficiently small one has

$$((\lambda + \epsilon)I - A)^{-1} = \sum_k (\lambda_0 I - A)^{-1} (\epsilon(\lambda_0 I - A)^{-1})^k \tag{3}$$

For ϵ sufficiently small $\|\epsilon(\lambda_0 I - A)^{-1}\| < 1$, the series is absolutely convergent and defines a bounded operator.

Definition 5 (*spectrum*) The *spectrum* $\sigma(A)$ of the operator A is the complement in C of the resolvent set.

From the definition it follows that for every $K > 0$ the closed set

$$\sigma(A) \cap \{\lambda : \ |\lambda| \leq K\} \tag{4}$$

is compact.

There need not be any relation between the spectrum of A and the spectrum of A^*.

Definition 6 (*spectral radius*) The *spectral radius* $r(A)$ of the operator A is by definition

$$r(A) \equiv \sup_{\lambda \in \sigma(A)} |\lambda| \tag{5}$$

Lemma 4 *One has*

$$r(A) = \|A\| \tag{6}$$

Proof In general for operators on Banach space $r(A) = \liminf_{n\to\infty}(|A^n|)^{\frac{1}{n}}$. For operators on a Hilbert space the following inequality holds: $\|A^{n+m}\| \leq \|A\|^m \|A\|^n$. From this inequality Lemma 4 follows. ♡

The definitions of resolvent, spectrum, spectral radius extend easily to operators on Banach spaces. Let X be a Banach space A a closed operator on X. Let $\phi \in X^*$. Then A^* defined by $A^*\phi(x) \equiv \phi(Ax)$ is a linear operator on X^*.

More generally let X and Y be two Banach spaces and A be a map $X \to Y$. The map dual to A is defined by

$$A^* : \ Y^* \to X^*, \qquad (A^*\phi)(x) = \phi(Ax) \qquad \phi \in Y^* \tag{7}$$

Define $\|A\|_{L(X,Y)} = \sup_{x\in X, \|x\|=1} \|Ax\|_Y$. Then

Lemma 5

$$\|A\|_{L(X,Y)} = \|A^*\|_{L(Y^*,X^*)} \tag{8}$$

Proof

$$\begin{aligned} \|A\|_{L(X,Y)} &= \sup_{x\in X, \|x\|=1} \|Ax\|_Y = \sup_{\|x\|\leq 1} \sup_{\|\phi\|\leq 1} |\phi(Ax)| \\ &= \sup_{\|x\|\leq 1} \sup_{\|\phi\|\leq 1} |A^*\phi(x)| = \sup_{\|\phi\|\leq 1} \|A^*\phi\| = \|A\|_{L(Y^*,X^*)} \end{aligned} \tag{9}$$

♡

The spectrum of an operator A on a Hilbert space $\mathcal{H}$ may be an arbitrary subset of the complex plane. We give an example of an operator which has as spectrum the entire complex plane and of one which has empty spectrum.

Example 1 Consider the operator $i\frac{d}{dx}$ defined on absolutely continuous functions in $[0, 1]$. The Hilbert space is $\mathcal{H} = L^2([0, 1])$. One has $\rho(A) = \{0\}$. Indeed for every $\lambda \in C$ one has

$$(A - \lambda I)e^{-i\lambda x} = 0 \tag{10}$$

and the function $e^{-i\lambda x}$ is absolutely continuous. It follows that $\sigma(A) = C$.

Example 2 We consider the same differential operator but now restricted to absolutely continuous functions on $[0, 1]$ which vanish at the origin. The Hilbert space is still $\mathcal{H} = L^2([0, 1])$.

In this case $\rho(A) = C$. To see this for each $\lambda \in C$ and $g \in \mathcal{H}$ define

$$(S_\lambda g)(x) = i \int_0^x e^{-i\lambda(x-y)} g(y) dy \tag{11}$$

The operator S_λ is bounded (it is an integral operator with kernel bounded in L^1) and one verifies that S_λ leaves invariant the set of absolutely continuous functions in $[0, 1]$. Moreover it is a right and left inverse of $A - \lambda I$:

$$(A - \lambda I)S_\lambda = S_\lambda(A - \lambda I) = I \tag{12}$$

Therefore $A - \lambda I$ is invertible for every $\lambda \in C$ and $\sigma(A) = \emptyset$.

Definition 7 (*point spectrum*) The *point spectrum* of A is the collection of $\lambda \in C$ for which $A\phi = \lambda\ \phi$ has at most finitely many solutions $\phi^k \in \mathcal{H}$.

Definition 8 (*essential spectrum*) The *essential spectrum* of the operator A is the complement in $\sigma(A)$ of the point spectrum.

Notice that with this definition if the Hilbert space is infinite-dimensional the operator I (identity) has 1 as essential spectrum and its point spectrum is empty.

Definition 9 (*Compact operator*) An operator K is *compact* if it maps to unit ball of $\mathcal{B}(\mathcal{H})$ into a compact subset of $\mathcal{B}(\mathcal{H})$. Equivalently it is compact is for every weakly convergent bounded sequence $\{\phi_k,\ k = 1, \ldots\}$, $\phi_k \in \mathcal{H}$ the sequence $\{K\phi_k\}$ is *strongly convergent*.

Weyl's Lemma The essential spectrum of an operator A does not change if one adds a bounded operator B. In particular if the Hilbert space is infinite dimensional 1 is in the essential spectrum of the operator $I + B$ if the operator B is bounded

We analyze further the spectrum of a special class of operators, the self-adjoint ones, that we introduce now.

Definition 10 (*self-adjoint operator*) An operator A is self-adjoint if

$$D(A) = D(A^*) \qquad \text{and} \qquad A^*\phi = A\phi \quad \forall \phi \in D(A) \tag{13}$$

Definition 11 (*essentially self-adjoint operator*) A densely defined symmetric operator A is *essentially self-adjoint* if its closure in the graph norm of $\bar{A}$ is a self-adjoint operator.

The property of essential self-adjointness is important in practice: the self-adjoint operators of interest in Physics are usually defined initially on *natural domains* which are in general *smaller* than their full operator domains.

Consider for example the operator $A \equiv \frac{d^2}{dx^2} \equiv \Delta$ on $L^2(R)$ defined initially on functions that belong to $C_0^\infty(R)$ (infinitely many times differentiable and with compact support). The domain of its closure in the graph norm is the collection of functions that are square integrable together with their second derivative. On this domain one has $\bar{A}\phi(x) = \frac{d^2}{dx^2}\phi(x)$. Integrating by parts twice one verifies that $\bar{A}$ is a self-adjoint operator and that it is uniquely determined by its restriction to $C_0^\infty(R)$.

Consider now the operators Δ and $\frac{1}{|x|^\alpha}$. They are both self-adjoint. We have already discussed the domain of Δ. The domain $D(\frac{1}{|x|^\alpha})$ is the collection of functions f which are square integrable and such that $\int \frac{1}{|x^{2\alpha}|}|f(x)|^2 < \infty$. Both domains are dense.

The sum $H_{\alpha,C} = \Delta + C\frac{1}{|x|^\alpha}$ $C \in R$ is defined *in the intersection of the domains*, which is a domain on which both are *essentially* self-adjoint. We shall prove that if $\alpha < \frac{3}{2}$ *their sum* is essentially self-adjoint in this domain and that *the closure* of their sum is a self-adjoint operator. But for $\alpha > 2$ or $\alpha = 2$ and C large enough the operator $H_{\alpha,C}$ on its natural domain of definition is symmetric *but not self-adjoint*.

Every self-adjoint operator is symmetric. We shall analyze in chapter "Lecture 17: Kato-Rellich Comparison Theorem. Rollnik and Stummel Classes. Essential Spectrum" the obstructions to the converse statement.

If A is symmetric and not self-adjoint, A^* *extends* A (in the sense that the domain is not smaller and on the common domain the two operators coincide). In general A^* *is not a symmetric operator*; we shall see that the self-adjoint extensions of A are restrictions of A^* to a suitable sub-domain.

1 Characterization of the Self-adjoint Operators

We shall now characterize the self-adjoint operators.

Theorem 6 *Let A be closed densely defined and symmetric. The following conditions are equivalent:*

(i) *A is self-adjoint;*
(ii) *$Ker(A^* \pm iI) = \{0\}$;*
(iii) *$Range(A \pm iI) = \mathcal{H}$.*

We have defined

$$Ker\ A \equiv \{\phi \in D(A),\quad A\phi = 0\} \qquad Range\ A \equiv \{A\phi,\quad \phi \in D(A)\} \tag{14}$$

Proof $(i) \Rightarrow (ii)$ Let $\phi \in Ker\ (A^* + i)$ i.e. $(A^* + i)\phi = 0$. Since A is self-adjoint also $A\phi = -i\phi$ holds. Therefore

$$i\|\phi\|^2 = i(\phi, \phi) = (A^*\phi, \phi) = (\phi, A\phi) = -i(\phi, \phi) = -i\|\phi\|^2 \tag{15}$$

and $\phi = 0$.

$(ii) \Rightarrow (iii)$ We have assumed that A is densely defined therefore A^* is closed and $Range\ (A \mp iI)^\perp = Ker\ (A^* \pm iI)$. We must therefore prove that $Range(A \pm iI)$ is closed (being closed and dense it coincides then with $\mathcal{H}$).

We give the proof for $(A - iI)$. Let $\phi \in D(A)$. Then

$$\|(A - iI)\phi\|^2 = \|A\phi\|^2 + \|\phi\|^2 + i(\phi, A\phi) - i(\phi, A\phi) = \|A\phi\|^2 + \|\phi\|^2 \tag{16}$$

We have used $D(A) \subset D(A^*)$ and the fact that A^* extends A because A is symmetric.

Consider a sequence $\phi_n \in D(A)$, $(A - iI)\phi_n \to \psi_0$. Suppose that $\{\psi_n\}$ converges to $\psi \in \mathcal{H}$ and that also the sequence $A\phi_n$ converges. Since A is closed, there exists $\xi \in \mathcal{H}$ such that $A\phi_n \to \xi$. For every n one has $(A - iI)\phi_n - A\phi_n + i\phi_n = 0$ therefore $\psi = A\xi - i\xi$, $\xi \in D(A)$. This proves $\psi \in Range(A - iI)$.

$(iii) \Rightarrow (i)$ Let $\phi \in D(A^*)$. By assumption $Range(A - iI) = \mathcal{H}$, therefore

$$\exists \eta \in D(A), \qquad (A - iI)\eta = (A^* - iI)\phi \tag{17}$$

Since $D(A) \subset D(A^*)$ it follows $(A^* - iI)(\phi - \eta) = 0$. But $Ker(A^* - iI) = Range(A + iI)^\perp = \{0\}$. Therefore $\phi = \eta$ and $\phi \in D(A)$ and therefore $A^* \subset A$. Since $A \subset A^*$ (A is symmetric and closed) $A = A^*$ follows. ♡

We remark that there is nothing sacred in the choice of the complex number i in Theorem 6. One could as well have formulated points (ii) and (iii) in Theorem 6 in the following way: for any complex number z with $\mathcal{I}mz \neq 0$ one has

$$Ker(A^* - zI) = 0, \qquad Range(A - zI) = \mathcal{H} \tag{18}$$

Indeed we shall prove that if A is symmetric the dimension of $Ker(A - zI)$ is constant both in the upper and in the lower half planes (but it may not be the same in different half planes).

From Theorem 1 one concludes that if A is essentially self-adjoint the operators $(\bar{A} - iI)^{-1}$ and $(\bar{A} + iI)^{-1}$ are bounded, commute and

$$(\bar{A} \pm iI)^{-1} = [(\bar{A} \mp iI)^{-1}]^* \tag{19}$$

Therefore $(\bar{A} + iI)^{-1}$ and $(\bar{A} - iI)^{-1}$ are *normal* (they are bounded and commute with the adjoint). It follows that if A is essentially self-adjoint the operator

$$U_{\bar{A}} \equiv (\bar{A} - iI)(\bar{A} + iI)^{-1} \tag{20}$$

is unitary and one has

$$\bar{A} = i\frac{U_{\bar{A}} - I}{U_{\bar{A}} + I} \tag{21}$$

Conversely if $U_{\bar{A}}$ is unitary, from (21) one derives that $\bar{A}$ is selfadjoint.

Definition 12 (*Cayley operator*) The unitary operator U_A is the *Cayley unitary operator* (or *Cayley tranform*) of the self-adjoint operator A.

Note that even if $\bar{A}$ is not self-adjoint, its Cayley transform can be defined, but in this case it is not a unitary operator but only *a partial isometry*.

The definition of Cayley tranforms is *not unique*. For any complex number z with $Imz \neq 0$ one could have chosen as unitary operator

$$U_z = \frac{\bar{A} - zI}{\bar{A} - \bar{z}I} \tag{22}$$

and then $\bar{A} = i\frac{U_z - (Rez)I}{U_z + (Rez)I}$. Independently on the definition chosen, the operator A is bounded iff the spectrum of its Cayley operator does not contain -1.

Definition 13 (*essential spectrum*) The real number λ belongs to the *essential spectrum* $\sigma_{ess}(A)$ of the self-adjoint operator A if for every $\epsilon > 0$ the spectral projection $E_{[\lambda-\epsilon,\lambda+\epsilon]}$ has infinite-dimensional range.

Definition 14 (*discrete spectrum*) The *discrete spectrum* $\sigma_{disc}(A)$ of the self-adjoint operator A is the complement of $\sigma_{ess}(A)$ in $\sigma(A)$.

Notice that an eigenvalue λ of the operator A belongs to the *discrete spectrum* iff it has finite multiplicity. For example the operator $A \equiv -\frac{d^2}{dx^2}$ defined on functions that belong to $L^2(R)$ together with their second (distributional) derivative is a self-adjoint operator with absolutely continuous spectrum that has multiplicity two and coincides with R^+. The unitary operator U which *diagonalizes* A can be chosen to be in this case the Fourier transform operator and in the new representation the operator becomes multiplication by k^2. The spectral multiplicity is two because k^2 corresponds to the two values $+k$ and $-k$ of the coordinate.

The operator I (identity) has purely point spectrum consisting in the single value 1; if $\mathcal{H}$ is infinite-dimensional the corresponding subspace is infinite-dimensional. Therefore in this case the discrete spectrum of I *is empty* and the number 1 belongs to the essential spectrum of I. The spectral measure of I is the infinite direct sum of replicas of the Dirac measure concentrated at 1.

2 Defect Spaces

Theorem 7 *Let the operator A be closed, symmetric and densely defined. Then the dimension of* $Ker(A^* + \lambda I)$ *is constant (as a function of* λ*) both in the upper and in the lower complex plane (the dimension may be different in the two half-planes).*

Outline of the proof The main part of the work is the proof that if $|Im\ \eta| < \gamma|\ Im\ \lambda|$ with γ sufficiently small, then

$$Ker(\lambda + \eta - A^*) \cap Ker(\lambda - A^*)^{\perp} = \{0\} \tag{23}$$

It follows from (23) that the two spaces have the same dimension (possibly infinite). Given (23) the proof of the theorem follows by iteration, covering each of the two half planes with open sets that do not intersect the real axis. To prove (23) we remark that, under the hypotheses made on λ and γ, this equation is equivalent to

$$(Range(\lambda - A))^{-} \cap (Range(\lambda + \eta - A))^{\perp} = \{0\} \tag{24}$$

Let $\phi \in Range\ (\lambda - A)^{-}$. Then there exist a sequence u_n such that

$$\lim_{n\to\infty} u_n = \phi, \qquad u_n = (\lambda - A)\xi_n \qquad \|\xi_n\| > c\|\phi\| \tag{25}$$

and it suffices to prove $Range(\lambda - A) \cap (Range(\lambda + \eta - A))^{\perp} = \{0\}$.

If $u = (\lambda - A)\zeta$ one must have

$$((\lambda - A)\zeta, (\lambda + \eta - A)\psi) = 0 \qquad \forall \psi \in D(A) \tag{26}$$

Choosing $\psi = \zeta$ this gives

$$\|(\lambda - A)\zeta\|^2 + \eta((\lambda - A)\zeta, \zeta) = 0. \tag{27}$$

But this is impossible for $|\eta| = \gamma|Im\ \lambda|$ with γ sufficiently small since $\|(\lambda - A)\zeta\| > c\|\zeta\|$. ♡

We remark that A is bounded below (i.e. if for every $\phi \in D(A)$ one has $(A\phi, \phi) \geq M\|\phi\|^2$) one can make use of

$$\|(A - \lambda)\phi\| \geq (M - \lambda)\|\phi\| \tag{28}$$

and repeat the previous analysis in a complex neighborhood of $\{\lambda : Im\ \lambda = 0, \quad Re\ \lambda < M\}$.

Corollary I *Let A be closed, symmetric bounded below. Then* $dim\ Ker\ (\lambda I - A^*)$ *is constant as a function of* λ *in the complement of the set* $Im\ \lambda = 0, \quad Re\ \lambda \geq M$ *where* $M \equiv \inf_{\|\phi\|=1}(\phi, A\phi)$.

Corollary II *If A is closed, symmetric and densely defined and Range $(\lambda - A) = \mathcal{H}$ for at least one real value of λ, then A is self-adjoint.*

Definition 15 (*Defect space*) The space $Ker\ (\lambda - A^*)$ for $\lambda \notin \sigma(H)$ is the defect space for the operator A^*. In chapter "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method" we shall see the role of the defect space in the determination and classification of the possible self-adjoint extensions of the operator A.

For a closed operator A we have defined the *resolvent set* $\rho(A)$ as the open set of those $\lambda \in C$ for which $A - \lambda$ is boundedly invertible. If A is self-adjoint $\rho(A)$ is contained in the real line and $\rho(A)$ is an open set. By definition the *spectrum* of A is the complement of $\rho(A)$ in C.

We remark that the following *resolvent identity* holds. If A is self-adjoint and $\lambda_1,\ \lambda_2$ are in the resolvent set of A the following identity holds

$$\frac{1}{A - \lambda_1\, I} = \frac{1}{A - \lambda_2\, I} + \frac{1}{A - \lambda_1\, I}(\lambda_2 - \lambda_1)\frac{1}{A - \lambda_2\, I} \tag{29}$$

One can verify this identity noting that it is a relation between closed bounded operators and is verified when both sides act on the dense domain spanned by vectors of the form $(A - \lambda_1 I)(A - \lambda_2\, I)\psi$.

Two characterizations of a self-adjoint operator are important and will be used in these Lectures. One is a *spectral decomposition property*; the other is the characterization as *generator of a strongly continuous one-parameter group of unitary operators* (Stone's theorem). They were briefly mentioned in chapter "Lecture 5: Automorphisms; Quantum Dynamics; Theorems of Wigner, Kadison, Segal; Continuity and Generators".

Theorem (spectral theorem) *The operator A in the Hilbert space $\mathcal{H}$ is self-adjoint iff there exists a measure space $\{X,\ \mu\}$, a real measurable function $a(m)$ on X and a unitary operator (isometric and with isometric inverse) U from $L^2(X, \mu)$ to $\mathcal{H}$ such that $AU = Ua$ (we have denoted by a the multiplication operator by the function $a(m)$). If A is bounded one has $\|A\| = \|a\|_\infty$.*

Theorem (Stone's theorem) *The operator A is self-adjoint iff it is the generator of a one-parameter group of unitary operators $U(t)$ strongly continuous in t. We will write $U(t) = e^{itA}$.*

One sees from Stone's theorem that in Quantum Mechanics self-adjointness plays the same role as completeness of a vector field in Classical Mechanics. Together with the theorems of Wigner and Wigner-von Neumann, Stone's theorem associates a dynamics to maps that preserve transition probabilities and are measurable as a function of time.

We begin with the proof of the spectral theorem. For the proof we shall make use of the following theorems; for their proof we refer to any textbook in Functional Analysis.

Theorem (Riesz-Fischer theorem) *$L^2(M, \mu)$ is a (complete) Hilbert space.*

Theorem (Riesz theorem) *Let I be an interval of the real axis and let $C(I)$ be the class of continuous real valued functions on I. If $\tilde{\mu}$ is a linear continuous functional on $C(I)$ then there exists a Borel measure μ on I such that*

$$\tilde{\mu}(f) = \int_I f(m) d\mu(m), \qquad f \in C(I) \tag{30}$$

(n.b.: every continuous function is measurable for any Borel measure).

3 Spectral Theorem, Bounded Case

Using the theorems of Riesz and Riesz-Fisher we now prove the spectral theorem. We begin with the case A is closed, bounded and symmetric.

Proof of the spectral theorem for A closed, bounded, symmetric The *if* part of the theorem follows since the operator of multiplication by a real bounded function $a(x)$ is a self-adjoint operator on $L^2(X, \mu)$ for any Borel measure and a unitary map preserves the property of being self-adjoint.

We prove now the *only if* part. By assumption A acts on a separable Hilbert space, is bounded, closed and symmetric, therefore self-adjoint. Denote its norm by $\|A\|$ and consider the interval $I \equiv [-\|A\|, \|A\|]$. If P is a polynomial of degree N

$$\|P(A)\| \leq \sup_{\lambda \in I} |P(\lambda)| \tag{31}$$

Indeed if $P(x) = \sum_{n=1}^{N} c_n x^n$ one has $P(A) = \sum_{n=1}^{N} c_n A^n$. To prove (31) choose any $\phi \in \mathcal{H}$ and notice that if we denote by E the orthogonal projection to the subspace (at most $(N+1)$-dimensional) spanned by $\{\phi,\ A\phi, \ldots, A^N\phi\}$, $\phi \in \mathcal{H}$, one has, by construction

$$P(EAE)\phi = P(A)\phi \tag{32}$$

This holds for any $\phi \in \mathcal{H}$ and $\mathcal{H}$ has a denumerable basis. But $EAE\phi$ is a finite rank matrix and therefore (31) follows from

$$\|P(EAE)\| \leq \sup_{\lambda \in I} |P(\lambda)| \tag{33}$$

It follows from (31) that $f(\|A\|) = \|P(A)\|$ is a continuous function on on the interval I. By a lemma of Weierstrass every continuous function on I can be approximated arbitrary well by polynomials. Denote by $P_n(x)$ a sequence of polynomials that approximate uniformly the function $f(x)$ in I. According to (33) the sequence $P_n(A)$ converges in norm; therefore there exists an operator $f(A) \in \mathcal{B}(\mathcal{H})$ such that

$$\lim_{n\to\infty} P_n(A) = f(A) \tag{34}$$

From uniform convergence one has that

$$f(x) = \lim_{n\to\infty} P_n(x), \qquad g(x) = \lim_{n\to\infty} Q_n(x), \tag{35}$$

implies

$$f(A)g(A) = (fg)(A) = (gf)(A) \tag{36}$$

It follows that the continuous functions from I to $\mathcal{B}(\mathcal{H})$ defined by (37) form a commutative normed algebra. Therefore the map

$$C(I) \ni f \mapsto f(A) \in \mathcal{B}(\mathcal{H}) \tag{37}$$

is a *-homeomorphism of commutative normed algebras (it preserves conjugation $f(A)^* = \bar{f}(A)$ because this relation is satisfied by polynomials).

We want to show that this homeomorphism is *realized by a unitary map* between $\mathcal{H}$ and a space $L^2(I, \mu)$ for a suitable choice of the measure μ. Let $\phi \in \mathcal{H}$, $\|\phi\| = 1$. For every continuous positive function f on I define

$$\mu_\phi^A(f) \equiv (\phi, f(A)\phi) \tag{38}$$

Since f is positive there exists a continuous real valued function g such that $f(x) = g^2(x)$. Therefore

$$(\phi, f(A)\phi) = \|g(A)\phi\|^2 \geq 0 \tag{39}$$

By Riesz's theorem there exists a Borel measure μ_ϕ on I such that

$$(\phi, f(A)\phi) = \int_I f(\lambda)\, d\mu_\phi \tag{40}$$

Denote by $\mathcal{H}_\phi$ the subspace spanned by $\{\phi,\ A\phi, A^2\phi, \ldots\}$. By construction $\mathcal{H}_\phi$ is left invariant by the action of A and of any of its powers. Since A is self-adjoint one has

$$(A\psi, \zeta) = 0 \qquad \psi \in \mathcal{H}_\phi, \qquad \zeta \in \mathcal{H}_\phi^\perp \tag{41}$$

so that the operator is reduced by $\mathcal{H}_\phi$. We must prove that there exists an invertible isometry V_ϕ such that

$$\mathcal{H}_\phi \xrightarrow{V_\phi} L^2(I, \mu_\phi) \tag{42}$$

$$V_\phi \phi = \iota, \qquad V_\phi A_R V_\phi^{-1} = \lambda\, I \tag{43}$$

We have denoted by A_R the restriction of A to the invariant subspace $\mathcal{H}_\phi$ and with ι the function which is constant on I and such that $\mu_\phi(\iota) = 1$. Remark that if $f(A)\phi = 0$ then

$$0 = \|f(A)\phi^2\| = \int_I |f(\lambda)|^2 d\mu_\phi(\lambda) \tag{44}$$

and therefore the function $f(A)\phi$ is the zero element of $\mathcal{H}_\phi$. Therefore V_ϕ is defined as a map from a dense set of $\mathcal{H}_\phi$ to a dense set in $L^2(I, \mu_\phi)$. Moreover one has for any two continuous functions f and g

$$(g(A)\phi, f(A)\phi) = \int_I \bar{g}(\lambda) f(\lambda) d\mu_\phi \tag{45}$$

It follows that V_ϕ is an isometry and extends to a unitary operator from $\mathcal{H}_\phi$ to $L^2(I, \mu_\phi)$.

If $\mathcal{H}_\phi = \mathcal{H}$ (i.e. if ϕ is cyclic under the repeated action of A) we have proved the thesis of the theorem and the function $a(m)$ is the coordinate in I. If $\mathcal{H}_\phi$ is a proper subspace of $\mathcal{H}$ choose a second element $\phi_2 \in \mathcal{H}_\phi^\perp$ and repeat the procedure considering now the restriction of A to the subspace generated by the action on ϕ_2 of polynomials in A. Notice that by construction $\mathcal{H}_{\phi_2}$ is perpendicular to $\mathcal{H}_\phi$.

We obtain thus an invertible isometry between $\mathcal{H}_{\phi_2}$ and $L^2(I, \mu_{\phi_2})$ where μ_{ϕ_2} is

$$\mu_{\phi_2}(f) = (\psi, f(A)\phi_2), \qquad f \in C(I) \tag{46}$$

If $\mathcal{H}_\phi \oplus \mathcal{H}_{\phi_2} = \mathcal{H}$ we have found an invertible isometry V between $\mathcal{H}$ and $L^2(I, \mu_\phi) \oplus L^2(I, \mu_{\phi_2}) = L^2(I, \mu_\phi + \mu_{\phi_2})$ such that

$$VAV^{-1} = \lambda\cdot \tag{47}$$

where $\lambda\cdot$ denotes multiplication by the coordinate λ.

Introducing on I a Borel measure μ such that $\int_B d\mu = \int_B d\mu_\phi + \int_B d\mu_\psi$ for every Borel set B we can write the isometry as isometry between $\mathcal{H}$ and $L^2(I, \mu)$.

If $\mathcal{H}_\phi \oplus \mathcal{H}_{\phi_2} \neq \mathcal{H}$ we can repeat the procedure with a third element $\phi_3 \in \mathcal{H}$ that is perpendicular to $\mathcal{H}_\phi \oplus \mathcal{H}_{\phi_2}$. By an induction procedure ($\mathcal{H}$ is separable) we obtain an invertible isometry U between $\mathcal{H}$ and $L^2(I, \mu)$ for a Borel measure $\mu(\lambda)$ which is the sum of the Borel measures we have obtained associated to the sequence. The measure μ does not depend on the sequence of vector $\phi, \phi_2, \phi_3 \ldots$ that were used in its construction. One has an invertible isometry operator U from $\mathcal{H}$ to $L^2(I, d\mu)$ such that

$$UAU^{-1} = \lambda\cdot \tag{48}$$

The measure $d\mu$ associated to the self-adjoint operator A is called the spectral measure of A. This completes the proof of the spectral theorem for bounded self-adjoint operators. ♡

Notice that the measure which is constructed in this way may be degenerate. For example the measure corresponding to the identity correspond to infinitely many copies of the Dirac measure on the point on R of coordinate 1.

It is convenient to introduce the *spectral multiplicity*. One can show that one can decompose uniquely the spectrum in the denumerable sum of Borel subsets σ_i in such a way that in each of them the spectral multiplicity is an integer $N_i, 0 \leq N_i \leq \infty$. As an example the Laplacian on R has as spectrum the negative real line with multiplicity 2 (complex conjugation intertwines the two components).

One can say equivalently that the spectral theorem gives for any self-adjoint operator A an isometric invertible map between $\mathcal{H}$ and $L^2(\sigma(A), K)$ where K a infinite dimensional space such that the operator A is mapped into the operator $x P_x$ where $x \in \sigma(A)$ and P_x is a projection in $\mathcal{B}(\mathcal{K})$.

The spectral theorem allows us to extend the functional calculus to bounded closed operators. Indeed every such operator can be written as sum over the complex field of self-adjoint bounded operators.

It is also possible to show that the L^∞ topology introduced on functions on I corresponds to the weak topology of bounded operators on $\mathcal{H}$. In particular projection operators correspond to indicator functions of Borel sets and belong to the weak closure of $C(A)$. The extension to L^∞ can be obtained as follows. Let $\{\phi_\alpha, \alpha = 1, \ldots\}$ be a collection of orthogonal vectors such that

$$\bigoplus_\alpha \mathcal{H}_\alpha = \mathcal{H}, \qquad \mathcal{H}_\alpha \equiv \mathcal{H}_{\phi_\alpha} \tag{49}$$

and suppose that the operator A is reduced by all $\mathcal{H}_\alpha$.

Let f be a bounded function measurable for μ_α and define on $\mathcal{H}_\alpha$ the operator

$$\hat{f}_\alpha \equiv V_\alpha^{-1} f V_\alpha \tag{50}$$

Define

$$f(A) = \bigoplus_\alpha \hat{f}_\alpha \tag{51}$$

This definition is well posed and induces a correspondence between essentially bounded functions on R and bounded self-adjoint operators on $\mathcal{H}$.

Since $L^\infty(R)$ is in the dual of Borel measures the correspondence between L^∞ and $\mathcal{B}(\mathcal{H})$ does not depend on the choice of the functions f_α. If ξ_B is the indicator function of the Borel set $B \subset M$ the operator

$$E_B \equiv V \xi_B V_{-1} \tag{52}$$

is an orthogonal projection.

One verifies that if $B_1 \subset B_2$ then $E_1 \leq E_2$ and that $E(I) = I$. One has thereby found a *projection-valued* Borel measure associated to the self-adjoint operator A.

Definition 16 (*spectral family*) The family of projection operators in $\mathcal{B}(\mathcal{H})$ associated by the operator A to the Borel sets in $[-\|A\|, +\|A\|]$ is called *spectral family* of A.

Often one makes use of the notation $E_\lambda \equiv E(-\infty, \lambda]$ and one writes $f(A) = \int f(\lambda) d\mu(\lambda)$; one uses also the notation

$$(\phi, f(A)\phi) = \int f(\lambda) d\mu_\phi(\lambda) \tag{53}$$

for $\phi \in D(f(A))$.

The construction of a spectral measure can be extended to the case of N self-adjoint operators $A_1, A_2, \ldots, A_N$ which are *mutually commuting* in the sense that bounded Borel measurable function $f(A_k)$ and $g(A_h)$ pairwise commute. The spectral measure is now a projection-valued Borel measure on R^N. The functional calculus for their continuous functions is defined as in the case of a single operator. We make use of Weierstrass's lemma to approximate by polynomials the set of continuous compactly supported functions in N variables.

In this case the compact set is

$$[-\|A_1\|, \|A_1\|] \times [-\|A_2\|, \|A_2\|] \times \cdots \times [-\|A_N\|, \|A_N\|] \tag{54}$$

The functional calculus is established also in this case by considering a generic element $\phi \in \mathcal{H}$ and setting $\mathcal{H}_\phi = \bigcup_P P(A_1, \ldots, A_n)\phi$. Since the operators A_k commute $\mathcal{H}_\phi$ reduces all A_k and proceeding by induction one constructs a measure μ on R^N such that for every continuous function f one has

$$\int f(m_1, \ldots, m_N) d\mu_\phi = (\phi, f(A_1, \ldots, A_N)\phi) \tag{55}$$

By continuity one defines $g(A_1, \ldots, A_N)$ for every L^∞ function on R^N. Passing to indicator functions of Borel sets in R^N one can define spectral projections and spectral families. Using the notation

$$E_{\lambda_1, \ldots, \lambda_N} \equiv E((-\infty, \lambda_1] \times \cdots (-\infty, \lambda_N]) \tag{56}$$

one can write the spectral decomposition

$$f(A_1, \ldots, A_N) = \int f(\lambda_1, \ldots, \lambda_N) dE_\lambda \qquad \lambda \equiv \{\lambda_1, \ldots, \lambda_N\} \tag{57}$$

i.e. for every $\psi \in \mathcal{H}$ one has $(\psi, f(A_1, \ldots, A_N)\psi) = \int f(\lambda_1, \ldots, \lambda_N) d\mu_\psi$.

One can prove also here that the spectral family does not depend on the elements of $\mathcal{H}$ used in the construction. Notice that the formulas above remain valid also in case the operators A_k are functionally dependent. For example if $N = 2$ and $A_2 = g(A_1)$ one has

$$f(A_1, A_2) = \int f(\lambda_1, \lambda_2) dE_\lambda \tag{58}$$

but for for every $\psi \in \mathcal{H}$ the measure $(\psi, E_\lambda \psi)$ is supported by the domain in R^2 defined by $\lambda_2 = g(\lambda_1)$.

To analyze further the spectral structure of the self-adjoint operator A on the Hilbert space $\mathcal{H}$, recall that according to the spectral theorem there exists a smallest integer $M \in (N \cup \infty)$, M measures μ_i, $i = 1, \ldots, M$ on the real line and a unitary operator

$$U : \mathcal{H} \to \oplus_i L^2(R, d\mu_i) \tag{59}$$

such that $UAU^* = \lambda I$ where λ is the coordinate used on R and I is the identity in R^N. The operator A has *spectral multiplicity m* in the interval I_0 if in I_0 only m of the measures have weight different from zero.

A regular Borel measure on the real line decomposes uniquely in three disjoint measures: a purely point part $\mu_{p.p.}$, a part $\mu_{a.c.}$ which is absolutely continuous with respect to the Lebesgue measure and a continuous part μ_{sing} that gives zero measure to any denumerable collection of point and is singular with respect to the Lebesgue measure. Correspondingly one has an orthogonal decomposition

$$L^2(R, d\mu) = L^2(R, d\mu_{p.p.}) \oplus L^2(R, d\mu_{a.c.}) \oplus L^2(R, d\mu_{s.c.}) \tag{60}$$

and for any self-adjoint operator A on $\mathcal{H}$ one has the orthogonal decomposition

$$\mathcal{H} = \mathcal{H}_{p.p.,A} \oplus \mathcal{H}_{a.c.,A} \oplus \mathcal{H}_{s.c.,A} \tag{61}$$

4 Extension to Normal and Unbounded Self-adjoint Operators

Definition 17 (*normal operator*) A closed operator A is normal if it satisfies

$$A\, A^* = A^*\, A \tag{62}$$

(in particular every unitary operator is normal).

One can extend the definition of spectral family to the case of closed bounded *normal operators.* Defining

$$C \equiv \frac{1}{2}(A + A^*), \qquad B \equiv \frac{i}{2}(A - A^*) \tag{63}$$

one verifies that operators C and B are self-adjoint and commute; therefore a spectral family can be defined for them. The assertion follows since $A = C + B$, $A^* = C - B$.

If $\phi \in D(A)$ one has

$$((iI - A)\phi, (iI - A)\phi) \geq \|\phi\|^2 \tag{64}$$

and therefore $(iI - A)$ is injective. Since by assumption A is closed, also $Range(iI - A)$ is closed. Moreover $Range(iI - A)$ is dense in $\mathcal{H}$ since A is self-adjoint. Therefore

$$Range(iI - A) = \mathcal{H} \quad \Rightarrow \quad (iI - A)^{-1} \in \mathcal{B}(\mathcal{H})$$

In the same way one proves

$$(iI - A^*)^{-1} \in \mathcal{B}(\mathcal{H}) \tag{65}$$

and

$$(iI - A)^{-1}((iI - A^*))^{-1} = (iI - A^*)^{-1}(iI - A)^{-1}. \tag{66}$$

Therefore $(iI - A)^{-1}$ is bounded and normal and there exists an invertible isometry V from $\mathcal{H}$ to $L^2(M, \mu)$ such that

$$V^{-1}(iI - A)^{-1}V = c(m)\cdot \tag{67}$$

where $c(m)$ is bounded and μ-measurable. In the same way one proves $V^{-1}(iI + A)^{-1}V = -\bar{c}(m)$.

Recalling the definition of Cayley operator it follows that if we define $u(m) \equiv -\frac{\bar{c}(m)}{c(m)}$ and $a(m) \equiv i\frac{u(m)-1}{u(m)+1}$ one has

$$A = Va(m)V^{-1}, \quad D(A) = \left\{ \int |f(m)a(m)|^2 d\mu(m) < \infty, \quad V\phi = f(m) \right\} \tag{68}$$

In the separable case the measure μ can be regarded as a measure on R^K with $K \leq +\infty$. The least value of K in the interval I is called *spectral multiplicity* of the operator in the interval I. If one can choose $K = 1$ the operator has *simple spectrum*.

Notice that the Cayley operator U_A can be defined also in the case when A is closed but not self-adjoint. In that case it is a partial isometry from $Ker(A^* + iI)$ to $Ker(A^* - iI)^{\perp}$. We shall see in chapter "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method" the role of the Cayley operator for the construction of possible self-adjoint extensions of A.

5 Stone's Theorem

We now prove Stone's theorem. Recall that a operator A with dense domain $D(A)$ is called *generator* of a one-parameter family $V(t)$ of operators if $V(t)$ solves $i\frac{d}{dt}V(t) = AV(t)$ on suitable dense domains.

Theorem 8 (Stone) *A one parameter group of unitary operators $U(t)$ in $\mathcal{H}$ is strongly continuous iff its generator A is a self-adjoint operator.*

One should notice that for groups of unitary operators weak and strong continuity coincide. Indeed

$$\|(U(t) - U(s))\phi\|^2 = 2 - Re(\phi, U(t-s)\phi) \tag{69}$$

Proof of Stone's theorem The condition is necessary. Indeed from the spectral theorem one has

$$V e^{itH} V^{-1} = e^{itm} \qquad \text{on } L^2(R, d\mu) \tag{70}$$

But then

$$(\phi, [e^{itH} - I]\phi) = \int (e^{itf(m)} - 1) d\mu_\phi \qquad \mu_\phi(R) = 1$$

and due to the dominated convergence theorem $\lim_{t\to 0} \int (e^{itf(m)} - 1) d\mu_\phi = 0$.

We prove now the the condition is sufficient. Define

$$B\phi \equiv \lim_{t\to\infty} i \frac{U(t) - I}{t} \phi \tag{71}$$

for every ϕ for which the limit exist and denote by $D(B)$ the collection of such vectors. We prove first that this set in dense. Let $f \in C_0^1$ and define

$$\Phi_f = \int_{-\infty}^{\infty} f(t) U(t) \phi dt \qquad \phi \in \mathcal{H}. \tag{72}$$

This set of vectors is dense in $\mathcal{H}$. For them we can compute the action of B:

$$B\phi_f == \lim_{s\to 0} \frac{U(s) - I}{s} \phi_f = i\phi_{-f'} \tag{73}$$

On these vectors the operator B is symmetric: it is easy to prove that

$$(B\phi_f, \phi_g) = (\phi_f, B\phi_g)$$

We must prove that B is essentially self-adjoint.

Suppose to the contrary that there exist $\psi \in D(B^*)$ such that $B^*\psi = i\psi$. Then for every $\phi \in D(B)$ one has

$$\frac{d}{dt}(U(t)\phi, \psi) = (iBU(t)\phi, \psi) = -i(U(t)\phi, B^*\psi) = (U(t)\phi, \psi) \tag{74}$$

But this would imply $(U(t)\phi, \psi) = e^t(\phi, \psi)$ and since it must remain bounded we derive $(\phi, \psi) = 0$. Since the domain is dense, one has $\psi = 0$. A similar analysis

shows that $B^*\psi = +i\psi$ has no solution. Therefore the defect spaces are empty and the operator B is essentially self-adjoint. ♡

In view of the importance of self-adjointness it will be convenient to give criteria that imply this property for a Schrödinger operator. In chapter "Lecture 17: Kato-Rellich Comparison Theorem. Rollnik and Stummel Classes. Essential Spectrum" and "Lecture 18: Weyl's Criterium, Hydrogen and Helium Atoms" we shall discuss some of them, related to perturbations of the Laplacian.

6 Convergence of a Sequence of Operators

We now discuss briefly the convergence of a sequence of operators.

Definition 18 (*resolvent convergence*) A sequence of operators A_n on a Hilbert space $\mathcal{H}$ is *strongly resolvent convergent* to A ($A_n \to_{s.r.} A$) if for $z \notin (\bigcup_n \sigma(A_n)) \cup \sigma(A)$ one has

$$(A_n - z)^{-1} \to (A - z)^{-1} \tag{75}$$

where convergence is in the strong operator topology.

Strong resolvent convergence is a property that is very useful in the study of the spectrum of operators which are given as limits of sequences of operators with known spectra. Indeed, if strong resolvent convergence holds, the spectrum of the limit operator is not bigger that the limit of the spectra (it may be smaller!). The following proposition gives a criterion for strong resolvent convergence.

Proposition 9 *If there exists a subset $E \subset D(A_n)$ dense in $D(A)$ such that for every $f \in E$ the sequence $A_n f$ converges uniformly to Af, then the sequence converges strongly in resolvent.*

Proof If $f \in E$ set $h \equiv (A - z) f$. Then

$$(A - z)^{-1}h - (A_n - z)^{-1}h = (A_n - z)^{-1}(A_n - A) f \tag{76}$$

and for a dense set $[(A_n - z)^{-1} - (A - z)^{-1}]h \to 0$ since $(A_n - A) f \to 0$ and $(A_n - z)^{-1}$ is bounded. The operators $(A - z)^{-1}$ and $(A_n - z)^{-1}$ are equi-bounded. Therefore

$$\|[(A_n - z)^{-1} - (A - z)^{-1}]\phi\| \to 0 \quad \forall \phi \in \mathcal{H} \tag{77}$$

♡

This result holds under weaker assumptions if A_n and A are bounded below. Without loss of generality one may assume that they are bounded below by the identity I. The subset E may now be not a subset of $D(A)$ but must be dense in the domain of the quadratic form $(\phi, A\phi)$ associated to A. This is the content of Theorem 11 that has an independent interest. This theorem is particularly important

in case the bounded operators are resolvents of self-adjoint operators. Before stating this theorem we shall state and prove a preliminary Proposition.

Proposition 10 *Let the sequence of bounded and closed operators* $\{A_n\}$ *satisfy* $A_n \leq A$ *where* A *is closed and bounded below. Then*

$$(A_n - A) \to_w 0 \quad \leftrightarrow (A_n - A) \to_s 0 \tag{78}$$

Proof For bounded and closed operators the functional calculus holds. Therefore the square-root of a bounded positive operator is well defined. One has

$$|(A_n - A)\phi| \leq \|(A - A_n)\||(A - A_n)^{1/2}\phi| \forall \phi \in \mathcal{H} \tag{79}$$

But $|(A - A_n)^{1/2}\phi|^2 = (\phi, (A - A_n)\phi)$. Therefore

$$\|(A_n - A)\phi\|^2 \leq \|(A - A_n)\|(\phi, (A - A_n)^{1/2}\phi)\| \qquad \forall \phi \in \mathcal{H} \tag{80}$$

The assertion is proved. ♡

Proposition 10 can be extended to unbounded A_n on a common dense domain.

Theorem 11 *Let the operators* $\{A_n\}$, A *be closed, with* $A_n \geq A \forall n$ *and* A *bounded below. Denote by* $Q(A)$ *the set of vectors* $\phi \in \mathcal{H}$ *such that* $(\phi, |A|\phi) < \infty$.

If there exists a set E *dense in* $\mathcal{H}$ *and such that* $E \subset Q(A_n)$ $\forall n$ *then* $(f, A_n f) \to (f, Af)$ $\forall f \in E$ *implies* $A_n \to_{s.r.} A$.

Proof There is no loss of generality in assuming $A \geq I$ and thereofore also $A_N \geq I$. By assumption $A_n^{-1} f$ is bounded in $Q(A)$.

Therefore A_n^{-1} converges weakly to A^{-1} and from Theorem 5.8 A_n^{-1} strongly converges to A^{-1}. ♡

An interesting consequence of Theorem 11 is the following:

Theorem 12 *Define*

$$\mathcal{H} = L^2(R^N), \quad H_0 = -\Delta, \quad U(x) \geq 0, \quad W(x) < 0 \tag{81}$$

Let $\mathcal{K} \equiv Q(H_0) \cap Q(U)$ *be dense in* $\mathcal{H}$*) and for* $\phi \in \mathcal{K}$ *let*

$$-((\phi, W\phi) \leq a(\phi, (H_0 + U)\phi) + b(\phi, \phi) \qquad a < 1 \tag{82}$$

Denote by U_N *the function* U *and by* W_R *the function* W *truncated at height* N *and* $-R$ *respectively. Then*

(1) $H = H_0 + U + W$ *is essentially self adjoint, bounded from below and for all vectors* ϕ *such it is finite*

$$(\phi, H\phi) = (\phi, H_0\phi) + (\phi, U\phi) + (\phi, W\phi) \tag{83}$$

(2)

$$\overset{s.r.}{\lim_{N\to\infty}} \overset{s.r.}{\lim_{R\to\infty}} H_{N,R} = H, \qquad H_{N,R} = H_0 + U_N + W_R$$

Proof Let $H_N = H_0 + U_N + W$. Notice thate $H_{N,R} < H_N \quad \forall R$ and that $H > H_N \quad \forall N$. Furthermore

$$\lim_{N\to\infty} (\phi, H_N\phi) = (\phi, H\phi), \quad \phi \in Q(H_0) \cap Q(U),$$

$$\lim_{R\to\infty} (\psi, (H_N + W_R)\psi) = (\psi, H_N\psi) \quad \psi \in Q(H_0) \tag{84}$$

Then we can apply Theorem 11 first to the sequence $H_0 + U_N + W_R$, $R \to \infty$ and then to the sequence H_N, $N \to \infty$. ♡

7 Ruelle's Theorem

We prove now an important theorem which is relevant for the study of the spectral properties of a self-adjoint operator A and will play a role in the theory of scattering in Quantum Mechanics.

Before discussing it, we recall the following identity

$$\lim_{\epsilon\to 0} \frac{1}{2i\pi} \int_a^b \left(\frac{1}{A - \lambda - i\epsilon} - \frac{1}{A - \lambda + i\epsilon} \right) d\lambda = \frac{1}{2}(E_{[a,b]} - E_{(a,b)}) \tag{85}$$

which follows immediately from the isomorphism that the spectral representation induces between bounded function of A and $L^\infty(R)$ and from the fact that, setting

$$f_\epsilon(x) \equiv \frac{1}{2i\pi} \int_a^b \left(\frac{1}{x - \lambda - i\epsilon} - \frac{1}{x - \lambda + i\epsilon} \right) d\lambda \tag{86}$$

one has that $\lim_{\epsilon\to 0} f_\epsilon$ is 0 if x does not belong to $[a, b]$, is equal to $\frac{1}{2}$ if $x = a$ or $x = b$ and is equal to one if $x \in (a, b)$.

Theorem 13 (Ruelle) *Let A be self-adjoint, $\phi \in \mathcal{H}$. If there exists $p > 1$ such that*

$$\lim_{\epsilon\to 0} \int_a^b |Im(\phi, R(\lambda + i\epsilon)|^p d\lambda < +\infty, \qquad R(z) = (z - A)^{-1} \tag{87}$$

then $E_{(a,b)}\phi \in \mathcal{H}_{a.c.,A}$.

Proof Let S be an open set, and let

$$\bigcup_{i=1}^{N}(a_k, b_k) \subset S, \qquad b_k < a_{k+1}, \qquad N < +\infty$$

From Hölder inequality for every pair of positive numbers $p,\ q$ with $\frac{1}{q} + \frac{1}{p} = 1$ one has

$$(\phi, E_S\phi) \leq \lim_{\epsilon\to 0}\left[\int_a^b |Im(\phi, R(\lambda + i\epsilon)|^p d\lambda\right]^{\frac{1}{p}} \nu(S)^{\frac{1}{q}} \tag{88}$$

where $\nu(S)$ is the Lebesgue measure of S. We must prove that if $\nu(I) = 0$ then $(\phi, E_I\phi) = 0$.

Since Lebesgue measure is regular there exists a sequence S_i^N, $i = 1, \ldots, N$, of disjoint intervals such that $I \subset \bigcup_k S_k^N$. From $\nu(I) = 0$ follows

$$\inf_n \sum_k \nu(S_k^N)^{\frac{1}{q}} = 0, \qquad q < 1 \tag{89}$$

and from this one derives $(\phi, E_S\phi) = 0$. ♡

We close this Lecture with the definition of *spectral density*, much used in Solid State Physics. Like any density it is a measure of the *relative weight* of different interval of the spectrum. In the case of Schrödinger operators an intrinsic definition of *spectral density* is given as follows.

Consider the Schrödinger operator $H = -\Delta + V(x)$ in R^d and denote by K_L the hypercube of side $2L$ centered in the origin. Assume that $V(x)$ is sufficiently regular. Let Δ_L be the Laplacian with Dirichlet boundary condition at ∂K_L. This operator has compact resolvent and pure point spectrum. If the potential V is sufficiently regular the same is true for $H_L \equiv \Delta_L + V$ defined in K_L with the same boundary conditions.

Denote by $N_I(H_L)$ the number of eigenvalues of H_L in the interval $I \subset R$ counting multiplicity.

Definition 19 (*spectral density*) Define the *weight* $p_H(I)$ of the interval I with respect to H as

$$p_H(I) = \lim_{L\to\infty} \frac{N_I(H_L)}{(2L)^d} \tag{90}$$

if the limit exists.

If $p_H(I)$ extends to a measure $d\nu(\lambda)$ on the Borel sets of R and this measure is absolutely continuous with respect to the Lebesgue measure on R we define *spectral density* the positive measurable function ρ defined by $d\nu = \rho(\lambda)d\lambda$.

References

1. Dunford, N., & Schwartz, J. T. (1963). *Linear operators II*. New York: Interscience Publishers.
2. Kato, T. (1966). *Perturbation theory for linear operators*. Berlin: Springer.
3. Lieb, E., & Loss, M. (1997). Analysis American Mathematical Society Providence.
4. Reed, M., & Simon, B. (1975). *Methods of modern mathematical physics II*. New York: Academic Press.
5. Riesz, F., & Nagy, B. Lecons d'Analyse Fonctionelle Akademiai Kiado Budapest 72.
6. Yoshida, K. (1965). *Functional Analysis*. Berlin: Springer.
7. Weidemann, J. (1980). *Linear operators in Hilbert Spaces*. New York: Springer.

Lecture 7: Quadratic Forms

A quadratic form $q(\phi)$ is a map form a complex Hilbert space $\mathcal{H}$ to $(-\infty, +\infty]$ which satisfies

(1)
$$q(\phi + \psi) + q(\psi - \phi) = 2q(\phi) + 2q(\psi)$$

(2)
$$q(\lambda\phi) = |\lambda|^2 q(\phi) \qquad \lambda \in C$$

Its *domain* $Q(q)$ (*the form domain* of q) is the collection of ϕ such that $|q(\phi)| < +\infty$.

We will consider only forms that are *bounded below*, i.e. satisfy

(3)
$$q(\phi) \geq a\|\phi\|^2, \qquad a \in R$$

There is a strict relation between forms and sesquilinear maps. A sesquilinear map S is a map $\{\phi, \psi\} \mapsto C$ that is linear in the second argument and antilinear in the first.

A sesquilinear map S defines uniquely a quadratic form by $q_S(\phi) \equiv S(\phi, \phi)$. Conversely property (2) implies by polarization that a quadratic form q uniquely defines a sesquilinear form S_q by $q(\phi) = S_q(\phi, \phi)$. Therefore a semibounded quadratic form defines a pre-Hilbert space product by

$$((\phi, \psi)) = S_q(\phi, \psi) + (a + 1)(\phi, \psi)$$

If the form is closed, this defines a Hilbert space associated to the form.

The largest a for which (1) holds is called *lower bound* of q and is denoted by $\gamma(q)$. We will prove (KLMN theorem) that semibounded closed forms are in one-to-correspondence with semibounded self-adjoint operators. If the symmetric unbounded operator A has many self-adjoint extensions, different extension

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_7

correspond to different quadratic forms. Therefore different extensions can be characterized by their quadratic forms rather that by their action in the defect space. A theorem of Krein, proved in "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method", states that the difference between these quadratic form is a form associated to a self-adjoint operator in the defect space.

For example if B is a second order elliptic operator $\sum_{i,j=1}^{d} \partial_i A_{i,j} \partial_j$, where $A(x)$ is a smooth function on R^d with value in strictly positive matrices, defined in the interior of a closed domain $\Omega \subset R^n$ with smooth boundaries the associated forms are *Dirichlet forms.* The *standard Dirichlet form* can be expressed ad $\int_\Omega (\nabla\phi, A(x)\nabla\phi)dx$ is associated to the self-adjoint operator defined by Dirichlet boundary conditions ad $\partial\Omega$ (i.e. $\phi(x) = 0$ at $\partial\Omega$). All other forms are obtained by adding to the standard one a negative term which is a bilinear form on the boundary data (function and normal derivative). In this case different extensions define different Sobolev spaces.

Definition 1 (*closed*) The form q is *closed* if $Q(q)$ is complete in the norm

$$[q(\phi) + C)\|\phi\|^2]^{\frac{1}{2}} \tag{1}$$

for $C > 0$ sufficiently large,

It can be proved that this condition is equivalent to lower semi-continuity.

1 Relation Between Self-adjoint Operators and Quadratic Forms

There is a natural way to associate quadratic forms to symmetric closed operators that are bounded below. There is no loss of generality in assuming that the operator satisfies $B \geq I$. If this is not the case, consider the operator $B + b$ for a sufficient large positive b.

If B is self-adjoint and strictly positive ($\sigma_B \subset [b, \infty)$, $b > 0$) consider the quadratic form $q_B(\phi) = \infty$ if $\phi \notin D(|B|^{\frac{1}{2}})$ and otherwise

$$q_B(\phi) = (|B|^{\frac{1}{2}}\phi, |B|^{\frac{1}{2}}\phi), \quad \phi \in D(|B|^{\frac{1}{2}}) \tag{2}$$

Notice that if $\phi \in D(B)$ one has $q_B(\phi) = (\phi, B\phi)$.

In general if B is positive and symmetric let q_B be the positive closed form associated to B. Let $\mathcal{M}$ be the domain of q_B. We extend q_B to all of $\mathcal{H}$ by setting $q_B(\psi) = +\infty$ if $\psi \in \mathcal{H} - \mathcal{M}$.

The KLMN theorem states that the converse is also true, i.e. there is a one-to-one correspondence between closed semi-bounded quadratic forms and self-adjoint semi-bounded operators on closed subspaces of $\mathcal{H}$: for any such form q there exists B such that $q = q_B$.

Notice that we could think of associating a quadratic form $\tilde{q}_A$ with domain $Q(A)$ (which is smaller in general than $D(\sqrt{A})$) to a symmetric positive operator A by

$$q_A(\psi) = (\psi, A\psi), \qquad \psi \in D(A) \tag{3}$$

With this definition the form $\tilde{q}_A$ *may not be closed*. But it follows from the KLMN theorem that if A is symmetric and *bounded below*, the form $\tilde{q}_A$ defined on $D(A)$ is closable. We shall discuss further in these Lectures the relation between *closed* quadratic forms and self-adjoint extensions of a symmetric operator.

Quadratic forms have a natural role in Quantum Mechanics [8]. For Schrödinger operators on compact domains in R^d they are useful to distinguish between different boundary conditions. Quadratic forms enter in the proof of self-adjointness of Schrödinger operators with singular potentials, in particular in cases when the potential is not Kato-small with respect to the Laplacian. This is the case, e.g. of forces which have support in a set of Lebesgue measure zero but finite Newtonian capacity.

Moreover Dirichlet forms by their very nature lend themselves to use of mini-max methods (for example in the estimate of the number of bound states) and can be very useful in the study of convergence of a sequence of operators.

We will prove that a positive *closed* quadratic form associated to a symmetric closed operator form defines uniquely a self-adjoint positive extension, the *Friedrichs extension*, which corresponds to the largest form among those which are associated to self-adjoint extensions of the given symmetric operator. The Friedrichs extension is important also for its connection to the Tomita-Takesaki duality theory for von Neumann algebras $\mathcal{M}$ and for the Tomita duality for positive cones. We will give the basic elements of this theory in the second part of these Lectures. The connection occurs through the following property of the Friedrichs extension.

Let A be a positive symmetric closed operator in the Hilbert space $\mathcal{H}$ and let U be a unitary operator on $\mathcal{H}$ which commutes with A in the following sense: U leaves the domain of A invariant and on this domain $UA = AU$. Let A_F be the Friedrichs extension of A. The U commutes with the spectral projections of A_F. Therefore if $\mathcal{M}$ is a weakly closed algebra of operators on $\mathcal{H}$ which is generated by unitary operators which commute with the symmetric operator A (in the sense made precise above) then the spectral projections of A_F are in $\mathcal{M}'$, the commutant of $\mathcal{B}$. This plays a role in the construction of the *modular operator* associated to a cyclic and separating vector of a von Neumann algebra $\mathcal{M}$.

2 Quadratic Forms, Semi-qualitative Considerations

We shall begin the study of quadratic forms, in particular those associated to a Schrödinger operator, with a few semi-qualitative considerations.

We study first the one-dimensional case. Let $\rho(x)$ be a strictly positive function on the real line, Lebesgue integrable with integral one.

Consider in $L^2(R)$ (with respect to Lebesgue measure) the quadratic form defined on C_0^∞ by

$$q_\rho(f,\ g) \equiv \int_{-\infty}^{\infty} \frac{d\,\bar{f}}{d\,x}.\frac{d\,g}{d\,x}\rho(x)\ dx + \int \bar{f}(x)\ g(x)\rho(x)\ dx, \qquad f,\ g \in C_0^\infty(R) \tag{4}$$

This form can be extended to a closed form $\bar{q}_\rho$, with domain

$$D(\bar{q}_\rho) \equiv \{f\ :\ \int |\frac{d\ f}{dx}|^2\rho(x)dx + \int |f(x)|^2\rho(x)\ dx < \infty\}$$

Since $\rho \in C^1$, one can integrate by parts to obtain

$$q_1(f,\ g) = \int \bar{f}[-\frac{d^2g}{dx^2} - \frac{d\ \log\ \rho(x)}{dx}\frac{dg}{dx}]\rho(x)\ dx + \int \bar{f}\ g\rho(x)\ dx \tag{5}$$

Equation (5) defines, at least formally and in weak sense, an operator on L^2_ρ which can be interpreted as

$$-\frac{d^2}{dx^2} - \frac{d\ \log\ \rho}{dx}.\frac{d}{dx} + \rho \tag{6}$$

If $\rho(x)$ is the indicator function of the interval (0, 1), one obtains, at least formally, a symmetric operator which acts as $-\frac{d^2}{dx^2} + 1$ on twice differentiable functions with support in the open interval (0, 1). We will see in "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method" that this operator is symmetric but not self-adjoint has defect indices 2, 2 and therefore admits a four-parameter family of self-adjoint extension.

In this chapter we shall prove that to any *closed* quadratic form bounded below corresponds a *unique* self-adjoint operator, called *Friedrichs extension associated to the given form*.

A natural question to ask is: is there a maximal one among the forms which are associated to a positive operator? And a minimal one? In particular which is the maximal form for the symmetric operator $-\frac{d^2}{dx^2} + 1$ defined on twice differentiable function with support in (0, 1)?

We shall see that the extensions correspond to the choice of different boundary conditions and that the maximal extension is obtained by choosing Dirichlet boundary conditions. The other extensions correspond to forms that are obtained by subtracting positive bilinear forms defined on the boundary.

If ρ is not differentiable the manipulations in (5) are only formal; still the form q is positive and, under favorable circumstances, closable. It defines therefore a self-adjoint operator. In this case, in which sense one has to interpret (6)?

Other interesting problem are the following.

Consider the sequence of normalized isolated ground states of self-adjoint operator A_n. Assume that the sequence converges to a function ρ in some topology. Does the sequence of self-adjoint operators A_n converge *in some other topology* to a self-adjoint operator A which has ρ as isolated ground state?

Consider the operator Δ_{ρ_n} associated to the form $\int |\frac{d\phi}{dx}|^2 + \int \rho_n(x)|\phi|(x)^2$ and assume that the sequence ρ_n converges in L^1 to the characteristic function of the interval $[0, 1]$. It follows that on functions that are zero at the boundary with zero directional derivative the operators converge to the symmetric operator $-\frac{d^2}{dx^2} + 1$. Does the sequence Δ_{ρ_n} converges to one of the self-adjoint extensions?

For the moment we suppose that $\rho \in C^2$, bounded and with support the entire real line. Consider the isometry $L^2(R, \rho(x)\, dx) \xrightarrow{\Phi} L^2(R, dx)$ defined by $\Phi(f) = \rho^{1/2} f$. One has

$$\int |\Phi(f)(x)|^2 dx = \int |f(x)|^2\, \rho(x)\, dx$$

and Φ can be extended to a unitary map between $L^2(R)$ and $L^2_\rho(R)$.

In what follows, unless indicated otherwise, we shall denote by the same symbol quadratic forms which differ by such unitary transformation, and we shall denote by $Q(q)$ their domains. We will consider only closable forms.

Set $F = \Phi(f)$, $f = \Phi^{-1}(F)$; one has

$$q(f, g) \equiv \tilde{q}(F, G) = \int \rho(x) \frac{d}{dx}(\bar{F}(x)\rho^{-\frac{1}{2}}) . \frac{d}{dx} G(x)\rho^{-\frac{1}{2}})\, dx \quad G, F \in L^2(R, dx) \quad (7)$$

Integrating by parts, under the assumption that $\rho \in C^2$ is strictly positive and $F,\ G \in C^2$ are compactly supported,

$$q(F) = \int \bar{F}\ (-\frac{d^2}{dx^2} F + \frac{1}{2}\frac{\rho''}{\rho}\ F - \frac{1}{4}\frac{(\rho')^2}{\rho^2}\ F) dx$$

Defining

$$V(x) \equiv \frac{1}{2}\frac{\rho''}{\rho} - \frac{1}{4}\frac{(\rho')^2}{\rho^2} \quad (8)$$

the quadratic form is given for all $F \in C^2(R) \cap D(q)$ by

$$q(F) = \int \bar{F}\ (-\frac{d^2}{dx^2} + V)\ F(x)\, dx \quad (9)$$

The quadratic form is therefore associated to the operator $-\Delta + V$ in $L^2(R, dx)$ or at least to the restriction of this operator to C_0^2. If the form *defines* a self-adjoint operator, it must be one of the self-adjoint extensions of $-\Delta + V$ defined on C_0^2.

Conversely, one expects that, given $V(x)$, if (8), *regarded as an equation for* $\rho(x)$, has a positive integrable solution, then $H = -\Delta + V$ can be mapped by a unitary transformation into the operator

$$H_\rho = -\frac{d^2}{dx} - (\frac{d}{dx}\ \log\ \rho(x)).\frac{d}{dx} \tag{10}$$

defined on a suitable (dense) subset of L^2_ρ. Notice that the function ι identically equal to one satisfies at least formally

$$H_\rho \iota = 0$$

If the solution ρ of (8) is positive and integrable the function ι belongs to $L^2(\rho(x)dx)$, hence the operator H_ρ in (10) has a zero eigenvalue and the corresponding eigenfunction is the constant function ι. Since the map is unitary, the operator $-\Delta + V$ has a zero eigenvalue with eigenfunction $\rho^{1/2}$. If $\rho(x)$ satisfies further assumptions, in particular to be strictly positive, the zero eigenvalue is not degenerate.

Through the study of the Dirichlet form one can also determine sufficient conditions in order that the spectrum of $-\Delta + V$ be $\sigma \equiv \{0\} \cup [a, \infty)$ with $a > 0$. Notice that if $(-\Delta + V)\phi_0 = 0,\ \phi_0 \geq 0$ one has

$$V \equiv \frac{\Delta\ \phi_0}{\phi_0} \tag{11}$$

so that the potential V is determined by ϕ_0 (see (9)). Setting $\phi_0 \equiv \rho^{1/2}$ one sees that (8) coincides with (11).

Conversely, given a Schrödinger operator

$$H = -\frac{d^2}{dx^2} + a(x)\ \frac{d}{dx}$$

if there exists a positive function $\rho(x) \in L^1$ such that $a(x) = \frac{d\rho(x)}{dx}$, then H is a positive symmetric operator on L^2_ρ and $-\frac{d^2}{dx^2} + V(x)$ (where $V(x)$ is defined in (8)) is a positive symmetric operator on L^2. Moreover H has zero as eigenvalue and $\sqrt{\rho(x)}$ is the corresponding eigenfunction. If $\rho(x) \in C^2$ and strictly positive the operator H is essentially self-adjoint on C_0^∞.

This analysis leads to the following conclusion. Let $-\Delta + V$ be a positive self-adjoint operator and let zero be an eigenvalue. Let ϕ be the corresponding eigenfunction and define $\rho(x) \equiv |\phi(x)|^2$. Then $-\Delta + V$ corresponds to the closed quadratic form

$$q_V(f) = \int |\frac{df(x)}{dx}|^2 \rho(x)\ dx + \int V(x)|f(x)|dx \tag{12}$$

This remarks can be easily extended to more than one dimensions. An operator of the form $H = -\Delta + \mathbf{f}.\nabla$ where $\mathbf{f}(x) = \nabla(\log \rho(x))$, $\rho(x) \geq 0$ and $\rho \in L^1 \cap L^2$, is symmetric in $L^2_\rho(R)$ and unitary equivalent to

$$H = -\Delta + V, \qquad V(x) \equiv \frac{1}{2}\frac{\Delta\rho}{\rho} - \frac{1}{4}\frac{(\nabla\rho)^2}{\rho^2} \tag{13}$$

If ρ is sufficiently regular and strictly positive, then H is essentially self-adjoint on C_0^∞.

If $U(x) = \frac{x^2}{4} - \frac{1}{2}$ one has

$$\phi(x) = (2\sqrt{\pi})^{-\frac{1}{2}} e^{-\frac{x^2}{4}}, \qquad \rho(x) = (2\sqrt{\pi})^{-1}\, e^{-\frac{x^2}{2}}$$

We conclude that the operator $-\frac{d^2}{dx^2} + \frac{x^2}{4} - \frac{1}{2}$ on $L^2(R^n, dx)$ is unitary equivalent to $-\frac{d^2}{dx^2} + \mathbf{x}.\nabla$ on $L^2(R, C_n e^{-\frac{x^2}{2}} d^n x)$ (C_n is a normalization constant).

An interesting problem, within the theory of quadratic forms, is the following: let $\{V_n\}$ be a sequence of potentials with the properties given above, and let ϕ_n be the eigenfunctions of $\Delta + V_n(x)$ corresponding to the lowest eigenvalue. We have seen in (11) that ϕ_n determines the potential $V_n(x)$.

A natural question is: is it possible to study the convergence (in semigroup, in resolvent sense…) of the sequence of operators $-\Delta + V_n$ by studying the convergence in L^1 of the sequence $\rho_n \equiv |\phi_n|^2$?

The analysis of quadratic forms will lead also to a better understanding of the relation between elliptic operators and Markov processes. An important role is played by the *Dirichlet forms* which we will discuss in the course of these Lectures.

3 Further Analysis of Quadratic Forms

After these heuristic motivations, we come back to the theory of quadratic forms. Recall that by polarization a quadratic form q with domain $D(q) \subset \mathcal{H}$ can be seen as a symmetric sesquilinear map $\tilde{q}$ from $D(q) \times D(q)$ to R. One has $q(\phi) = \tilde{q}(\phi, \phi)$. Conversely symmetric sesquilinear forms correspond to quadratic forms. From now on we shall use the same symbol for quadratic and for sesquilinear forms.

Definition 2 (*core*) If $D \subset D(q)$ is dense in $D(q)$ for the Hilbert topology defined by $\tilde{q}$ then we say that D is *a core* for q.

A quadratic form q with domain $D(q)$ is *closable* if it is the restriction of a closed form $\bar{q}$ and $D(q)$ is dense in $D(\bar{q})$ in the topology induced by $\bar{q}$. In this case we call $\bar{q}$ *the closure* of q.

We do not give the easy proof of the following:

Lemma 1 *The quadratic form q defined on $D(q)$ is closed iff the following is true: if*

$$\phi_n \in D(q), \quad \lim_{n\to\infty} |\phi_n - \phi| = 0, \quad \lim_{n\to\infty} q(\phi_n - \phi) = 0 \tag{14}$$

then

$$\phi \in D(q), \quad \lim_{n\to\infty} q(\phi_n - \phi) = 0$$

Example

(a) If $\mathcal{H} \equiv R^n$, $\tilde{q}$ is represented by a matrix Q; if Q is strictly positive, $\tilde{q}$ defines a scalar product.

(b) If $H = L^2(R, dx)$ define $D(q) \equiv \{f \in L^2,\ xf(x) \in L^2\}$ and, for $f,\ g\ \in D(\tilde{q})$

$$\tilde{q}(f, g) \equiv \int \bar{f}(x)\ g(x)x^2 dx$$

Then q is positive and closed and is the quadratic form of the operator that act as multiplication by the function x^2. Its domain is the domain of the operator multiplication by $|x|$.

(c) If $\mathcal{H} \equiv L^2(R)$ and $f \in H^1$ define a positive quadratic form by

$$q_\alpha(f, g) = \int (\nabla f, \nabla g) + (f, g) + \alpha \bar{f}(0)g(0), \quad \alpha \in R \tag{15}$$

Notice that q_α *is closed* and the topology it defines is equivalent to the topology of the Sobolev space H^1. Indeed functions in H^1 are absolutely continuous and therefore their evaluation at a point is a continuous operation.
The form q_α is therefore associated to a self-adjoint operator. It is natural to use for this operator the *formal* notation

$$H = -\frac{d^2}{dx^2} + \alpha\delta(x) + 1 \tag{16}$$

Notice that in dimension $d > 1$ the form $\int (\nabla f, \nabla g) + (f, g) + \alpha \bar{f}(0)g(0)$ *is not closed* and therefore in $d > 1$ *it does not define a self-adjoint operator.*

Not all forms which are bounded below are closable. For example

$$q_0(f) \equiv |f(0)|^2, \qquad D(q) = C_0(R)$$

is not closable in $L^2(R)$.

4 The KLMN Theorem; Friedrichs Extension

Theorem 2 (KLMN theorem) *Let q be a closable quadratic form on $D(q) \subset \mathcal{H}$* bounded from below.

There exists unique *a self-adjoint operator A, bounded below, such that*

$$q(\phi, \psi) = (\phi, A\psi) \qquad \psi \in D(A), \qquad \phi,\ \psi \in D(q) \tag{17}$$

Moreover if $q(\phi) \geq -M$ then A satisfies $(\phi, A\phi) \geq -M$. If q is positive then $D(q) = D(A^{1/2})$. This operator is called Friedrichs extension *of the form q.*

The name of the Theorem is composed of the initials of the names Kato, Lax, Milgram, Neumark.

As we remarked above this extension was described by Friedrichs in the case the standard Dirichlet form, and is now called the Friedrichs extension also in the general case. In the case of a standard Dirichlet form defined on functions in the interior of a compact domain $\Omega \subset R^3$ the Friedrichs extension correspond in Electrostatics to setting the potential to be zero on the boundary. For this case the proofs given by K. Friedrichs precedes the proof of the KMLN theorem; the second proof of the KLMN we give below follows the lines of the original proof of Friedrichs. The result of Friedrichs was later re-obtained by Sobolev and provides the definition of Sobolev space for domains with boundaries.

We can restrict our analysis to strictly positive quadratic forms. Indeed if q is bounded below by $-M$ we can consider the strictly positive quadratic form q' defined by $q'(\phi) = q(\phi) + (M+1)|\phi|^2$.

The theorem gives the existence of an associated self-adjoint operator A and it is easy to see that the operator $A - (M+1)\ I$ is associated to the form q. The restriction to strictly positive quadratic forms is needed because in the course of the proof we want to associate to the form a Hilbert space *contained* (strictly if the form is unbounded) in the Hilbert space $\mathcal{H}$ on which the form is defined.

We will give *two proofs* of Theorem 2. The two proofs are very similar but each has an independent interest and emphasizes different aspects of the Theorem.

First proof Set

$$(\phi, \psi)_1 \equiv q(\phi, \psi) + (\phi, \psi), \qquad \phi,\ \psi \in D(q) \tag{18}$$

By assumption $Q \equiv D(q)$ is a Hilbert space for the scalar product (18). We denote by $|\phi|_1$ the corresponding norm.

The topology of $\mathcal{H}$ is weaker than the topology induced by q. Therefore, by Riesz's theorem, there exists unique an element $\xi \in \mathcal{H}$ such that $(\xi,\ \psi) = (\phi, \psi)_1$.

Denote with B the map $Q \to \mathcal{H}$ defined by

$$B\ \phi = \xi \tag{19}$$

Since the form is symmetric the operator B is a symmetric operator in the Hilbert space $\mathcal{H}$. It is also bounded because the topology induced by the closed positive form q is stronger that the topology of $\mathcal{H}$.

We remark now that the range of B is dense in $\mathcal{H}$. From the theorem of Hellinger and Toeplitz we conclude that B (the closure of B) is self-adjoint in the Hilbert space Q *and bounded*.

Also $B > I$ and therefore B is invertible; notice that B^{-1} maps Q into $\mathcal{H}$. For all vectors $\phi \in D(B^{-1})$ one has

$$((B^{-1} - I)\phi, \psi) = (\phi, \psi)_1 - (\phi, \psi) = \tilde{q}(\phi, \psi)$$

Set $A \equiv B^{-1} - I$ (an operator with self-adjoint closure). One has

$$(A\phi, \psi) = \tilde{q}(\phi, \psi) \qquad \forall \phi \in D(A) \cap Q, \quad \forall \psi \in Q \tag{20}$$

This ends the first proof of Theorem 2. ♡

Second proof This proof is based on the introduction of a scale of Hilbert spaces. In the case of the Dirichlet form these will be Sobolev spaces. The advantage of this proof is that it permits to clarify the structure of the operator $A^*\ A$ if A is *closed but not self-adjoint*.

As in the previous proof, we remark that Q is a complete Hilbert space, that we shall denote with $\mathcal{H}_1$. By construction $\mathcal{H}_1 \subset \mathcal{H}$.

Denote with j the natural inclusion $\mathcal{H}_1 \subset \mathcal{H}$ obtained by identification, and with $\mathcal{H}_{-1}$ the space dual of $\mathcal{H}_1$ with respect to the scalar product in $\mathcal{H}$. This notation originates from the fact that if $\tilde{q}(f, f) = \int \nabla f . \nabla f + |f|_2^2$ the spaces $\mathcal{H}_{-1}$ and $\mathcal{H}_1$ are the Sobolev spaces of order -1 and 1 respectively.

For each $\phi \in \mathcal{H}$ the map $\psi \mapsto (\phi, \psi)$ defines a linear continuous functional and therefore $\mathcal{H} \subset \mathcal{H}_{-1}$. For each $\psi \in \mathcal{H}$ and $\phi \in \mathcal{H}_1$ one has

$$|j(\phi)\ \psi| \leq |\phi|\ |\psi| \leq |\phi|_1 |\psi| \tag{21}$$

and therefore $|j| \leq 1$ ($|j|$ is the norm of j as a map $\mathcal{H} \to \mathcal{H}_{-1}$).

Define now the operator $\hat{A}$ from $\mathcal{H}_1$ to $\mathcal{H}_{-1}$ as follows

$$(\hat{A}\ \xi)(\phi) = q(\phi, \xi) + (\phi, \xi), \qquad \phi, \quad \xi\ \in \mathcal{H}_1 \tag{22}$$

Notice that $\hat{A}$ is an *isometry* between $\mathcal{H}_1$ and $\mathcal{H}_{-1}$. Indeed

$$(\hat{A}\phi)(\phi) = |\phi|_1^2, \quad \sup_{|\phi|_1 = 1} |\hat{A}\xi(\phi)|| = \sup_{|\phi|_1 = 1} (\phi, \xi) = |\xi|_{-1} \tag{23}$$

and therefore the norm of $\hat{A}\ \xi$ as an element of $\mathcal{H}_{-1}$ coincides with the norm of ξ as an element of $\mathcal{H}_1$. Let the operator $\tilde{A}$ be defined by

$$D(\tilde{A}) = \{\phi \in \mathcal{H},\ \hat{A}\ \phi \in Range\ j\} \qquad \tilde{A}\ \phi = j^{-1}.\hat{A} \tag{24}$$

In this way $\tilde{A}$ is a map from the range of $\tilde{j}$ to $\mathcal{H}$.

Remark that the range of $\tilde{j}$ is dense in $\mathcal{H}$. Indeed let us assume that there exists $\xi \in \mathcal{H}_{-1}, \xi \neq 0$ such that $\xi(j(\phi)) = 0\ \forall \phi \in \mathcal{H}$. According to Riesz's theorem there must exist $\hat{\xi} \in \mathcal{H}_1$ such that

$$0 = \xi(j\ \phi) = (\tilde{\xi}, \phi) \quad \forall \phi \in \mathcal{H} \tag{25}$$

But this is impossible since H is dense in $\mathcal{H}_1 = (\mathcal{H}_{-1})^*$. Therefore $\tilde{A}$ is densely defined. It is symmetric since

$$(\phi, \tilde{A}\psi) = \tilde{q}(\phi, \psi) + (\phi, \psi) = (\tilde{A}\,\phi, \psi) \tag{26}$$

Consider now the operator $C \equiv (\tilde{A})^{-1}.\ j\ : \mathcal{H} \to \mathcal{H}_1$. It is symmetric (due to the symmetry of $\tilde{A}$) and everywhere defined in $\mathcal{H}$.

Let $\hat{j}$ be the identification map $\mathcal{H}_1 \to \mathcal{H}$ and consider the operator $\hat{j}C$ on $\mathcal{H}$. It is bounded and everywhere defined. Therefore by the theorem of Hellinger and Toeplitz its closure defines a self-adjoint operator. Due to the spectral theorem also C^{-1} is self-adjoint. One has $C^{-1} \equiv \tilde{A}$. Therefore $\tilde{A}$ is self-adjoint and has domain the range of C.

Set $A = \tilde{A} - I$. Also A is self-adjoint and has the same domain as $\tilde{A}$. If $\phi,\ \psi \in D(A)$ it follows from the definition of $\tilde{A}$ that $(\phi, A\psi) = \tilde{q}(\phi, \psi)$. Moreover $D(A)$ is dense in $\mathcal{H}_1$ and therefore also in $\mathcal{H}$.

This completes the second proof of Theorem 2. ♡

In the finite dimensional case, a strictly positive quadratic form together with the choice of a complete basis defines a positive definite matrix and then a positive hermitian operator. The choice of the complete basis is irrelevant since any two strictly positive matrices can be transformed one into the other by a *linear* transformation. One can consider the Friedrichs extension as an infinite-dimensional analogue of this property. We shall come back to this point in the Lecture in which we shall discuss the modular operator.

Example Let $\tilde{q}(f, g) \equiv \int \frac{df}{dx}.\frac{dg}{dx}\ dx = \frac{1}{2\pi} \int \bar{f}(k)\ g(k)\ k^2 dk$. One has

$$\mathcal{H}_1 \equiv \{f\ :\ \int |f(k)|^2 (k^2+1)^{-1} dk < \infty\},$$
$$\mathcal{H}_{-1} \equiv \{f\ :\ \int |f(k)|^2 (k^2+1) dk < \infty\} \tag{27}$$

$$D(A) \equiv \{f\ :\ \int |f(k)|^2 (k^2)^2 dk < \infty\} \tag{28}$$

and the operator $\hat{A}\ :\ \mathcal{H}_1 \to \mathcal{H}_{-1}$ after Fourier transform becomes multiplication by $(k^2 + 1)$.

The operator $\tilde{A}$ is the restriction of $\hat{A}$ to those elements of $D(\hat{A})$ whose image $\hat{A}$ belongs to $\mathcal{H}_1$. Notice that in this example the spaces $\mathcal{H}_1$ and $\mathcal{H}_{-1}$ coincide with the Sobolev spaces H^1 and H^{-1}.

A simple corollary of Theorem 2 is the following

Proposition 3 *Let A be a closed operator densely defined on a Hilbert space $\mathcal{H}$. Set*

$$D(A^*A) \equiv \{\phi \in D(A) \ : \ A\phi \in D(A^*)\}, \quad (A^*A)\phi = A^*(A\phi), \quad \phi \in D(A^*A) \tag{29}$$

Then A^*A is self-adjoint on $D(A^*A)$.

Proof Consider on $D(A) \times D(A)$ the sesquilinear form

$$q_A(\phi, \psi) \equiv (A\phi, A\psi) \tag{30}$$

One has $q_A \geq 0$ and $q(A)$ is closed since A is closed.

Denote by B the Friedriches extension of q_A. We prove that B coincides with $A^*\ A$ as defined in (29). Let

$$\mathcal{H}_1 \subset \mathcal{H} \subset \mathcal{H}_{-1} \tag{31}$$

the scale of Hilbert spaces associated to the form q_A. Define

$$A' \ : \ \mathcal{H} \to \mathcal{H}_{-1}, \quad (A'\phi)(\psi) = (\phi, A\psi), \quad \psi \in D(A) \tag{32}$$

(recall that $D(A) \subset D(A^*)$).

It is obvious that A' extends A^*; A' is the restriction of A^* to those elements in its domain such that the image under A' belongs to $\mathcal{H}$. One has therefore

$$D(A') \equiv \{\phi \ : \ A'\phi \in \mathcal{H}\} \tag{33}$$

Let $\hat{B}$ be defined as

$$\hat{B} \ : \ \mathcal{H}_1 \to \mathcal{H}_{-1}, \quad (\hat{B}\phi)(\psi) = q_A(\phi, \psi) = (A\phi, A\psi), \quad \phi, \ \psi \in \mathcal{H}_1 \tag{34}$$

Then $D(B) = \{\phi \in \mathcal{H}_1, \quad \hat{B}\phi \in \mathcal{H}\}$ and therefore $\hat{B}$ extends B.

Choose now $\phi, \ \psi \in \mathcal{H}_1$. Then

$$(A'(A\phi), \psi) = (A\phi, A\psi) = (\hat{B}\phi)(\psi) \Rightarrow B' = A'\ A \tag{35}$$

On the other hand

$$D(B) = \{\{\phi \in \mathcal{H}_1, \ : \ A^*(A\phi) \in \mathcal{H}\} = \{\phi \in \mathcal{H}_1, \ : \ A\phi \in D(A^*)\} = D(A^*\ A) \tag{36}$$

and moreover

$$B = \hat{B}\big|_{D(B)} = \hat{B}\big|_{D(A^*\,A)} = A^*\,A\big|_{D(A^*\,A)} = A^*\,A$$

This concludes the proof of Proposition 3. ♡

In the same way one proves

Proposition 4 *If A is closed and symmetric and A^2 is densely defined, then $A^*\,A$ is the Friedrichs extension of the quadratic form q corresponding to the sesquilinear form $\tilde{q}(\phi, \psi) = (\phi, A^2\psi), \quad \psi \in D(A^2)$.*

Example

(I) Let $A = i\frac{d}{dx}, \quad D(A) = \{\phi \,:\, \phi \in AC[0,1] \quad \phi(0) = \phi(1) = 0\}$ and therefore

$$D(A^*) = AC[0,1], \qquad A^* = i\frac{d}{dx}$$

($AC[0,1]$ is the set of absolutely continuous functions on $[0,1]$).
Then

(a) $A^*\,A$ coincides with $-\frac{d^2}{dx^2}$ with Dirichlet boundary conditions.

(b) $A\,A^*$ coincides with $-\frac{d^2}{dx^2}$ with Neumann boundary conditions.

(II) Let $a_k(x) \in L^2(R^d)_{loc}$, $k = 1, \ldots, d$, be real valued functions. Let τ_k be the closure of $(i\nabla + ea(x))_k$ on $C_0^\infty(R^d)$. Then $H \equiv \sum_{k=1}^d \tau_k^*\,\tau_k$ is self-adjoint on $\bigcap_k \{\psi \in D(\tau_k), \tau_k\psi \in D(\tau_k^*)\}$ and moreover the domain of the form q associated to H is $D(q) = \cap_k D(\tau_k)$. Formally one has $H = \sum_k (-i\frac{d}{dx_k} - ea_k(x))^2$.

5 Form Sums of Operators

The relation between closed semi-definite quadratic forms and self-adjoint operators bounded below allows the definition of a sum of two self-adjoint operators A and B by means of the sum of the corresponding quadratic forms (we will speak of *form sum*). In this way the sum can be defined also in cases when $D(A) \cap D(B) = 0$.

It is convenient to introduce a notion of *smallness* similar to the notion introduced by T. Kato for operators (we shall discuss this notion in "Lecture 17: Kato-Rellich Comparison Theorem. Rollnik and Stummel Classes Essential Spectrum").

Definition (*Kato smallness for forms*) The closed quadratic form q_1 is *Kato small* with respect to the closed, positive, symmetric quadratic form q_2 if one can find $0 < a < 1, b > 0$ for which

$$|q_1(\phi,\phi)| \leq a q_2(\phi,\phi) + b|\phi|^2 \qquad \forall \phi \in Q(q_2) \tag{37}$$

In this case we shall use the notation $q_1 \prec q_2$. If one can choose a arbitrary small (provided b is chosen sufficiently large), the form q_1 is *infinitesimal* with respect to q_2. We shall write $q_1 \prec\prec q_2$.

If $q_1 \prec q_2$ *it is not true in general* that the corresponding operators satisfy $A_1 \prec A_2$ in the sense of Kato. The converse is instead true if A_2 is strictly positive.

Theorem 5 [4] *Let A be a strictly positive self-adjoint operator and let q be a quadratic form which is Kato-small with respect to the form associated to A:*

$$|q(\phi,\phi)| \leq a q_A(\phi,\phi) + b|\phi|^2 \quad \phi \in D(A) \quad 0 < a < 1 \quad b > 0 \tag{38}$$

Then the form $q' = q_A + q$ defines uniquely a self-adjoint operator B with $D(B) \subset D(A)$ such that

$$(\phi, B\phi) = (\phi, A\phi) + q(\phi,\phi) \quad \forall \phi \in D(A) \tag{39}$$

The self-adjoint operator B is by definition the form sum *of the self-adjoint operators A and B.*

Moreover every domain on which A is essentially self-adjoint is a core for the quadratic form q_B associated to B (q_B is the closure of its restriction to $D(A)$).

Proof From (38) it follows that the topology defined by q' is weaker than that defined by $q_A(\phi) + c_1|\phi|^2$. Let $q(A)$ be the domain of the quadratic form associated to the positive operator A. By mini-max it follows that there exist λ such that $(\phi, B\phi) + \lambda(\phi,\phi) > 0$. It follows that the quadratic form

$$(\phi, B\phi) = (\phi, A\phi) + q(\phi)$$

defines a self-adjoint operator. If the quadratic form $q(\phi)$ is the quadratic form of an operator V the operator B is *the form sum* $A + V$.

Notice that since the domain of the forms are larger then those of the corresponding operators it may occur that the intersection $D(A) \cap D(V)$ is not dense in $\mathcal{H}$ and may even be the empty set. ♡

6 The Case of Dirichlet Forms

We call Dirichlet form on R^d the form $\int (\nabla\phi(x), A(x)\nabla\phi(x)) d\mu(x)$ on $L^2(R^d, d\mu(x))$ where μ is a measure on R^d and $A(x)$ is a function with values in positive definite matrices. In the second part of these Lectures we treat the general case and prove the connection of Dirichlet form with stochastic processes. Here we will limit ourselves to that case in which μ is the Lebesgue measure and the matrix-valued $A(x)$ is constant.

As an example of the results one can obtain by using Sobolev embeddings we prove that the form q_α on $L^2(R)$ defined by

$$q_\alpha(f) = \alpha |f(0)|^2$$

is Kato-infinitesimal with respect to the form q defined by

$$q(f) = \int \frac{|d\bar{f}}{dx|^2} + \int |f(x)|^2 dx \tag{40}$$

so that the form sum defines a self-adjoint operator. Notice that the form q_α is not even closed as a form in the Hilbert space $L^2(R)$. Indeed one can find converging sequences $\{f_n\}$ for which

$$lim_{n\to\infty} q_\alpha(f_n) \neq q_\alpha(lim_{n\to\infty} f_n)$$

The reason why the sum of the forms is closed is that those sequences do not converge in the topology defined by the sum of the forms.

The proof is simplified by taking the Fourier transform. We must prove that there exist positive constants $a_\lambda,\ b_\lambda$ with a_λ arbitrary small, for which if $\hat{\phi}(k) \in L^1$ and $\int (1+k^2)|\hat{\phi}(k)|^2 dk < \infty$ (so that $\phi \in Q(A)$) one has

$$|\lambda|^2 \int \hat{\phi}\, dk \leq a_\lambda \int |\hat{\phi}(k)|^2 (k^2+1) dk + b_\lambda \int |\phi(k)|^2 dk \tag{41}$$

Divide the integration range in two parts, $|k| \leq k_0$ and $|k| > k_0$. We estimate separately the two integrals, and then choose k_0 conveniently. For the integration on $|k| > k_0$ one has

$$\begin{aligned} \int_{|k|>k_0} \hat{\phi}(k) dk &\leq \int_{|k|>k_0} \frac{1}{\sqrt{1+k^2}} \sqrt{1+k^2}\, \hat{\phi}(k) dk \\ &\leq [\int_{|k|>k_0} \frac{1}{\sqrt{1+k^2}}^{-1}]^{1/2} . [\int_{|k|>k_0} (1+k^2)|\hat{\phi}(k)|^2 dk]^{1/2} \\ &\leq C(k_0) \int (1+k^2)|\hat{\phi}(k)|^2 dk \end{aligned} \tag{42}$$

where $C(k_0) \to 0$ if $k_0 \to \infty$.

For the integration on $|k| \leq k_0$ one has the following estimates

$$\begin{aligned} |\int_{|k|\leq k_0} \hat{\phi}(k) dk| &\leq [\int_{|k|\leq k_0} dk]^{1/2} \int_{|k|\leq k_0} |\hat{\phi}(k)|^2 dk \\ &\leq (2\, k_0)^{1/2} \int_{|k|\leq k_0} |\hat{\psi}(k)|^2 dk]^{1/2} \end{aligned} \tag{43}$$

It follows that, provided one takes k_0 sufficiently large, for any value of the parameter λ one can satisfy (41) with $a(\lambda)$ arbitrary small. This proves that for any real λ the quadratic form $q_A + \lambda\, q$ defines a self-adjoint operator. ♡

Notice that in the previous proof we have used *in an essential way* the fact that *the space is one-dimensional*; this implies $(1+k^2)^{-\frac{1}{2}} \in L^2(R)$. An analogous result *does not hold* in R^d if $d \geq 2$; indeed in R^d, $d \geq 2$, a function with square integrable gradient *needs not be continuous*.

For $d = 2$ and $d = 3$ it is still possible to define a self-adjoint operator which in some sense represents a perturbation of the Laplacian supported on a point but the construction is entirely different. We shall discuss it in "Lecture 20: Self-adjoint Extensions. Relation with Quadratic Forms. Laplacian on Metric Graphs. Boundary Triples. Point Interaction".

Notice also that in the proof we have used Fourier transform as an isometric map. The result *is no longer true in* R^+.

As a further example let $V \geq 0$, $V \in L^2_{loc,\Gamma}(R)$ (i.e. $|V(x)|$ is integrable outside a closed set Γ of Lebesgue measure zero). Let Q_Γ the set of functions in $L^2(R)$ which have support disjoint from Γ and define in Q_Γ the quadratic form

$$q_\Gamma(\phi) = \int |\phi(x)|^2 V(x) dx, \qquad \phi \in Q_\Gamma \tag{44}$$

Let

$$q_0(\phi) = \int (|\nabla\phi|^2 + |\phi|^2) dx \quad Q_0 \equiv \{\phi \, : \phi \in L^2, \quad \int |\nabla\phi|^2 dx < \infty\} \tag{45}$$

Then $Q_0 \cap Q_\Gamma$ is dense and $q \equiv q_0 + q_V$ is a form which can be easily proved to be positive and closed. Therefore q corresponds to a self-adjoint operator H. Formally

$$H = H_0 + V \qquad H_0 = -\frac{d^2}{dx^2} + 1 \tag{46}$$

(this identification is true in a strict sense on $D(H) \cap D(H_0) \cap D(V)$).

7 The Case of $-\Delta + \lambda|x|^{-\alpha}$, $x \in R^3$

We want now to use Theorem 2 to prove

Lemma 6 *If $\alpha < 2$ the expression $-\Delta + \lambda|x|^{-\alpha}$, $x \in R^3$ defines a self-adjoint operator for every real λ.*

In "Lecture 17: Kato-Rellich Comparison Theorem. Rollnik and Stummel Classes. Essential Spectrum" we will discuss perturbation theory for operators and we will see that this result in obtained only for $\alpha < 3/2$. In "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method" we shall prove a criterion, due to Weyl, according to which the preceding statement *is not true* if $\alpha = 2$ and $\lambda < -\frac{3}{2}$.

For the proof of Lemma 6 we recall the following inequality.

Lemma 7 (Hardy's inequality) *For each* $\phi \in L^2(R^3)$ *one has*

$$\int_{R^3} \frac{1}{4|x|^2}|\phi(x)|^2 d^3x \leq \int_{R^3} |\nabla\phi|^2 d^3x \tag{47}$$

or, equivalently,

$$(\phi, |\hat{p}|^2\phi) \geq (\phi, \frac{1}{4|x|^2}\phi) \tag{48}$$

Proof One has $\nabla(|x|^{1/2}\phi) = |x|^{1/2}\nabla\phi + \frac{1}{2}|x|^{-3/2}x\phi$ and therefore pointwise

$$\begin{aligned} |\nabla\phi|^2 &= \tfrac{1}{4}|x|^{-2}|\phi|^2 + \tfrac{1}{|x|^2}\textstyle\sum_k [\frac{\partial}{\partial x_k}(|x|^{1/2}\phi)]^2 - Re\,\frac{1}{|x|^{3/2}}(\bar{\phi}x_k \frac{\partial}{\partial x_k}(|x|^{1/2}\phi) \\ &\geq \tfrac{1}{4}|x|^{-2}|\phi|^2 - Re\,\tfrac{1}{|x|^3/2}(\bar{\phi}x_k \tfrac{\partial}{\partial x_k}(|x|^{1/2}\phi) \end{aligned} \tag{49}$$

We can now integrate over R^3 and obtain

$$\int_{R^3} |\nabla\phi|^2 d^3x \geq \int_{R^3} \frac{1}{4}|x|^{-2}|\phi|^2 d^3x - Re \int_{R^3} \frac{1}{|x|^{3/2}}\bar{\phi}\cdot(x\cdot\nabla)(|x|^{1/2}\phi)d^3x \tag{50}$$

For functions of compact support which are regular at the origin the right-hand side is

$$-\frac{1}{2}\int_0^\infty dr \frac{\partial}{\partial r}\int_{S^2} r|\phi(x)|^2 d\Omega = 0$$

and this proves (47). By continuity the same is true for all functions for which the right-hand side is finite. ◇

We prove now Lemma 6.

Proof of Lemma 6 We prove that in R^3 if $0 < \alpha < 2$ the (multiplication) operator $|x|^{-\alpha}$ is infinitesimal with respect to $-\Delta$ in the sense of quadratic forms.

Let $\phi \in C_0^\infty$ and, given $a > 0$, choose ϵ in such a way that $|x|^{-\alpha} < a\,|x|^{-2}$ for $|x| < \epsilon$.

Divide the integration domain in two parts, $|x| < \epsilon$ and $|x| \geq \epsilon$. By (47)

$$\begin{aligned} \textstyle\int_{R^3} |x|^{-\alpha}|\phi(x)|^2 dx &\leq a \textstyle\int_{|x|<\epsilon} |x|^{-2}|\phi(x)|^2 dx + \int_{|x|\geq\epsilon} |\phi(x)|^2 dx \\ &\leq a \textstyle\int |\nabla\phi|^2 dx + \frac{1}{\epsilon^\alpha}\int |\phi(x)|^2 dx \end{aligned} \tag{51}$$

By considering a sequence of approximating C_0^∞ functions (in the $L^2(R^3)$ topology) the inequality extends to all functions in $L^2(R^3)$. Therefore

$$q_{|x|^{-\alpha}}(\phi) \leq a q_0(\phi) + b|\phi|^2$$

where a can be chosen arbitrary small if one chooses b sufficiently large. It follows that

$$H = -\Delta + \lambda|x|^{-\alpha}, \qquad 0 < \alpha < 2$$

interpreted as sum of quadratic forms defines a self-adjoint operator whose form domain coincides with the form domain of the Laplacian. ♡

In the same way one can treat the following examples.

Example 1 Let $\mathcal{H} = L^2(R, dx)$, $H_0 = -\frac{d^2}{dx^2} + 1$. Let μ be a measure on the real axis whose Fourier transform (in a distributional sense) satisfies $\hat{\mu} \in L^\infty$ (remark that $\tilde{\mu}(k) = 1\ \forall k$ is equivalent to $\mu = \delta(x)$). Define an operator $W : Q_0 \to Q_0$ through

$$q_W(f, g) \equiv (f, W\ g) = \int \hat{f}(k)\ \hat{\mu}(k - p)\ \hat{g}(p)dp \tag{52}$$

If $\hat{\mu} \in L^\infty$ it is easy to prove that $\hat{f} \in L^1 \cap L^2$ implies that for every $\epsilon > 0$ one can find $b(\epsilon) > 0$ for which

$$|\hat{f}|_1 \leq \epsilon|k\ \hat{f}|^2 + b(\epsilon)|\hat{f}|^2$$

Therefore q_W is infinitesimal (as a form) with respect to q_0 and the Friedrichs extension $q_0 + q_W$ defines a self-adjoint operator. If $V \in L^1$ then $\hat{V} \in L^\infty$. But this example leads to more general perturbations of $\frac{d^2}{dx^2}$ since $\{\mu \in L^1\} \subset \{\tilde{\mu} \in L^\infty\}$.

Example 2 Let $H_0 = -\Delta + 1$ on $L^2(R^d, dx)$, $V = U + W$, $U \geq 0$, $W \in L^p + L^\infty$ with $p \geq d/2$ if $d \geq 3$, $p > 1$ if $d = 2$ and $p \geq 1$ if $d = 1$. Then the sum $H = -\Delta + 1 + V$ *understood in the sense of quadratic forms* defines a self-adjoint operator which is bounded below.

Proof We prove that W is infinitesimal in the sense of quadratic forms with respect to $-\Delta$ and therefore also to $-\Delta + V$ since V is positive.

In the decomposition $W = W_p + W_\infty \in L^p + L^\infty$ for each $\epsilon > 0$ one can choose $|W_p|_p < \epsilon$. Indeed every function in L^p is bounded in the complement of a domain of arbitrary small Lebesgue measure. On the other hand, using Hölder inequality

$$(f, Wf) \leq |W|_p|f^2|_q, \qquad \frac{1}{p} + \frac{1}{q} = 1 \tag{53}$$

and under our assumptions $\frac{1}{2q} \geq \frac{1}{2} - \frac{1}{d}$. It follows from Sobolev inequality

$$|f|_{2q}^2 \leq a(f, (-\Delta + b)\ f), \qquad a < \epsilon, \qquad b > b(\epsilon)$$

and hence

$$(f, Wf) \leq a|W|_p(f, (-\Delta f) + ab\ |W|_p|f|_{2q}^2 \tag{54}$$

♡

Example 3 Let $\hat{\mu}(k) = i\pi\ sign(k)$. For $\epsilon > 0$ the Fourier transform of $i\pi\ sign(k)\ e^{-\epsilon|k|}$ is $\frac{x}{x^2+\epsilon^2}$. Define the quadratic form

$$q(f) \equiv lim_{\epsilon\to 0} \int |f(x)|^2 \frac{x}{x^2+\epsilon^2} dx = P_v \int |f(x)|^2 \frac{1}{x} dx \tag{55}$$

(we have denoted by P_v the principal value of the integral). Notice that this quadratic form is not associated to the multiplication operator by $\frac{1}{|x|}$ nor to the multiplication operator by $\frac{1}{x}$. The sequence

$$q_\epsilon(f) \equiv \int |f(x)|^2 \frac{x}{x^2+\epsilon^2} dx \tag{56}$$

is composed of quadratic forms which are infinitesimal uniformly in ϵ with respect to the form associated with $-\frac{d^2}{dx^2} + 1$. Therefore also q is infinitesimal and the self-adjoint operator $-\frac{d^2}{dx^2} + P(\frac{1}{x})$ (P stands for *principal part*) is well defined *as sum of quadratic forms*.

Notice that in Theorem 5 one cannot substitute $a < 1$ with $a = 1$. One can see this through an example. Let

$$H_0 = -\frac{d^2}{dx^2}, \qquad V \equiv \frac{d^2}{dx^2} + \delta(x) \tag{57}$$

If $a < 1$, V is form-small with respect to H_0. (Recall that $\delta(x)$ is form-infinitesimal with respect to $-\Delta$.) Therefore if $a < 1$

$$H \equiv H_0 + V = -(1-a)\frac{d^2}{dx^2} + \delta(x)$$

is essentially self-adjoint on C_0^∞. But for $a = 1$ it is not even defined as closed form. This is to be contrasted with the similar case for self-adjoint operators (Wurst theorem) that we will discuss in "Lecture 14: A Theorem of Segal. Representations of Bargmann, Segal, Fock. Second Quantization. Other Quantizations (Deformation, Geometric)".

8 The Case of a Generic Dimension *d*

One can derive from Sobolev inequalities the conditions under which the expression $-\Delta + V$, understood in the sense of sum of the corresponding quadratic forms, defines a self-adjoint operator. Notice that in the case of interaction among N points in R^3 the Hilbert space is $L^2(R^{3N})$.

By inequalities of Sobolev and Hölder type [8, 9] one proves that in R^d, $d \geq 1$ there exist positive constants $C(p,d)$ for which for any function f

$$c|f|_p + |\nabla f|_p \geq C(p,d)|f|_q, \qquad \frac{1}{q} = \frac{1}{p} + \frac{1}{d} \tag{58}$$

(if the right-hand side is $+\infty$ the inequality is considered to be trivially satisfied). This inequality expresses the fact that if $f \in L^p$ is such that $\nabla f \in L^p$ the L^q norm of f is finite for some $p > q$.

If moreover f has mean zero the following (Poincaré) inequality holds

$$|f|_{\frac{2d}{d-2}} \leq \frac{d-1}{d-2} \prod_{k=1}^{d} |\frac{\partial f}{\partial x_k}|_2^{\frac{1}{d}} \tag{59}$$

The coefficient $C(p,d)$ in (58) diverges when $d \to \infty$. In this limit one has $q \to p$ in (58) and $\frac{2d}{d-2} \to 2$ in (59). Therefore in the limit $d \to \infty$ there is no gain in the L^q norm.

This can be seen as a consequence of the fact that there does not exist on R^∞ a regular measure that is quasi-invariant under translation (i.e. equivalent to its translates). Notice that translations are generated by the gradient.

On the contrary Gauss measures can be defined as probability measures also when $d = \infty$. It is therefore interesting to find inequalities in the case of a Gauss measure which correspond to Sobolev inequalities for the Lebesgue measure [8].

Let $d\nu(x)$ be the Gauss measure in R^d of mean zero and variance one. One must recall that in the case of the Gauss measure the standard Dirichlet form corresponds to the operator $-\Delta - x \cdot \nabla$. Using the explicit form of the density one verifies that *independently of the dimension d* the following inequality holds:

$$\int |\nabla f|^2 d\nu(x) \geq C \int |f(x)| \frac{\log f(x)}{\|f\|_\nu} d\nu(x) \tag{60}$$

This inequality holds also for $d = \infty$ if one uses a proper definition of Gauss measure (weak normal distribution in the terminology of I. Segal). Inequality (60) is the *logarithmic Sobolev inequality*; we shall discuss it further in these Lectures.

The logarithmic Sobolev inequality shows that, *independently from the dimension of the space*, the norm $|\nabla u|_\nu$ controls (with a dimension-independent constant) a norm which is *slightly stronger* than the L^2 norm (what is gained is only a logarithmic factor). The presence of the logarithm in (60) suggests that in estimating the strength of $V(x)$ compared to $-\Delta$, if one looks for a result which is valid also in R^∞, it may be convenient to compare $e^{-V} f$ with $e^{\Delta} f$.

Indeed we shall see that there exists a positive constant C such that

$$e^{-t} \leq \sqrt{\frac{q-1}{p-1}} \Rightarrow \|e^{-tH_0}\|_{q,p} < C \tag{61}$$

where H_0 corresponds to the quadratic form $q_{H_0}(f, f) = \int |\nabla f|^2 d\nu(x)$.

We have denoted by $|A|_{q,p}$ the norm of A as operator from L^q to L^p with $q < p$.

It is instructive to compare (61) valid for any $d \leq \infty$ with the inclusion for $d < \infty$

$$f \in L^2(R^d, dx) \Rightarrow e^{-tH_0} f \in C^\infty \quad \forall t > 0 \tag{62}$$

Therefore for any finite d the operator e^{-tH_0} has a very strong regularization property for $t > 0$ but this regularization is not independent of the dimension d and does not hold for $d = +\infty$.

9 Quadratic Forms and Extensions of Operators

We have seen (KLMN theorem) that a quadratic form q on a Hibert space $\mathcal{H}$ which is densely defined, closed and bounded below defines *uniquely* a self-adjoint operator A_q. In the case of Dirichlet forms this is the Friedrichs operator associated to q [3]. From the proof of Friedrichs theorem it follows that if the operator A_q is associated to the form q, the operator $H + cI$ is associated to the form $q + c|\phi|_2^2$, $c \in R$.

It is therefore sufficient to restrict the analysis to strictly positive quadratic forms and then to strictly positive operators. In this case the domain $Q(q)$ has the structure of a Hilbert space. If A_q is strictly positive, one has $D(A_q) \subset Q(q)$ with strict inclusion if the form is unbounded.

If the operator A is self-adjoint let q_A the corresponding quadratic form let A' be another extension and q' the associated quadratic form. In "Lecture 18: Weyl's Criterium, Hydrogen and Helium Atoms" we will prove that

$$q' = q_A + W \tag{63}$$

where W is a closed quadratic form on the defect space $\mathcal{K}$ considered as a Hilbert space with a norm equivalent to the graph norm of the adjoint A^*. Therefore W is the quadratic form associated to a self-adjoint operator B in the Hilbert space $\mathcal{K}$.

The quadratic forms associated to the extensions of a symmetric operator are therefore partially ordered (recall that the domains of the extensions have no partial ordering).

It can be seen [6] that if q is semibounded there is always a largest extension A_∞ (we will see later the origin of this notation), which turns out to be the Friedrichs extension, and a smallest extension A_0 (sometimes called the *Krein-von Neumann* extension).

In this notation the self-adjoint extensions of A are classified by the operator B and are self-adjoint if and only if B is self-adjoint (this is true also if $\mathcal{K}$ is infinite-dimensional). Call A_B the self-adjoint extension associated to B. One has

$$q(A_B) = q(A) + q_B(\eta) \qquad \eta \in \mathcal{K} \tag{64}$$

where $q(B)$ is a positive quadratic form $q_B(\psi) = (\psi, B\psi), \psi \in \mathcal{K}$. For $\phi \in Q(A_\infty)$, $\eta \in D(B)$ one has

$$((\phi + \eta), A_B(\phi + \eta)) = (\phi, A_\infty \phi) + (\eta, B\eta) \tag{65}$$

The smallest extension, denoted A_0, corresponds to $B = 0$ and the maximal extension A_F is corresponds to $B = \infty$, i.e. $q_B(\eta) = 0$ for all $\eta \in \mathcal{N} \equiv Ker(A^* - iI)$. The notation minimal and maximal refer to the fact that their quadratic form are minimal resp. maximal among the quadratic forms associated to extensions of the given symmetric operator.

One has [1, 2, 6, 7]

$$D(A_0) = \{\phi + A_\infty^{-1}\eta, \quad \phi \in D(B), \quad \eta \in \mathcal{N}\} \equiv D(A) \dot{+} \mathcal{N} \tag{66}$$

and

$$A_B\phi + A_\infty^{-1}(B\xi + \eta) + \xi = A\phi + B\xi + \eta, \quad \phi \in D(A),\ \xi \in D(B), \quad \eta \in \mathcal{N} \cap (D(B))^\perp \tag{67}$$

These results can be obtained by a careful analysis of the construction we have described. We shall come back to this topic in "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method". There we will discuss the theory of extensions of a symmetric operator both from the point of view of von Neumann (isometric maps between the two deficiency spaces) and from the point of view of Milman and Krein (extension of quadratic forms).

10 A Simple Example

A simple example may illustrate the problems treated above. We analyze it both from both from the point of view of operators and from that of quadratic forms [1].

Consider the operator $A = -\frac{d^2}{dx^2} + 1$ defined in R^+ on twice differentiable functions having compact support strictly contained in $(0, \infty)$. This operator is symmetric but not self-adjoint; it has defect spaces of dimension one, and we construct explicitly the corresponding elements.

From the definition of A^* one sees that the Kernel N of is made of the square integrable solutions of

$$-\frac{d^2u}{dx^2} + u = 0 \tag{68}$$

i.e. multiples of e^{-x}.

The condition on the domain D such that restriction of A^* to D be self-adjoint is obtained by integration by parts and the requirement that the contribution from the

boundary terms vanish (this contribution is a bilinear form which is not continuous in the topology of $\mathcal{H}$). This is seen to require

$$\sin\alpha\ \phi(0) + \cos\alpha\ \phi'(0) = 0 \tag{69}$$

for some $0 \leq \alpha \leq \frac{\pi}{2}$.

For $\alpha = 0$ this gives $\phi'(0) = 0$ (Neumann boundary conditions); for $\alpha = \frac{\pi}{2}$ it gives $\phi(0) = 0$ (Dirichlet b.c.). In the other cases one has $\phi'(0) = \tan\alpha\phi(0)$ (Robin b.c.).

Call A_α the corresponding operators. They are self-adjoint with R^+ as absolutely continuous spectrum with multiplicity and are represented by k^2 in the representation of the Hilbert space as

$$\mathcal{H} \ni \phi = \{\psi(k) = \int_R (\cos\alpha kx - \sin\alpha kx)\phi(x)dx,\ \ \phi \in L^2\} \tag{70}$$

We consider now the corresponding quadratic forms. An arbitrary function in $Q(A_\infty) + N$ is decomposed as

$$\phi(x) = \phi(x) - \phi(0)e^{-x} + \phi(0)e^{-x}$$

and one has

$$(\phi, A_B\phi) = \int_0^\infty |\phi(x) - \phi(0)e^{-x}|^2 dx + \int_0^\infty |\phi'(x) + \phi'(0)e^{-x}|^2 dx + B|\phi(0)|^2 \tag{71}$$

where B is a multiplication operator with spectrum in $[2, \infty)$. One has

$$(\phi, A_\alpha\phi) = \int_0^\infty [|\phi'(x)| + |\phi(x)|^2 + (\alpha - 2)|\phi(0)|^2 \tag{72}$$

In this example the upper bound for B is $\alpha = \infty$ (i.e. $\phi(0) = 0$) and corresponds to Dirichlet boundary conditions. The lower bound $\alpha = 2$ and corresponds to Neumann b.c.

We will prove in "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method" that there is a relation between the order structure of quadratic forms and the order structure of the resolvents [4, 8, 9]. For any two strictly positive self-adjoint operators $H_1,\ H_2$ denote by $q_1,\ q_2$ the corresponding quadratic forms. One has, for any $\lambda \in R$, $\lambda < \inf q_1(\phi, \phi)$, $\phi \in Q(q_1)$, $|\phi| = 1$

$$q_1 < q_2 \leftrightarrow (H_2 - \lambda)^{-1} \leq (H_1 - \lambda)^{-1} \tag{73}$$

where $\lambda \in R^-$ and where on the right the $\leq$ sign is understood in the ordinary sense for bounded operators.

Inequality (73) is very useful because it allows to control resolvent convergence with the convergence of quadratic forms, i.e. a liner problem.

The following characterization of the minimal and maximal extensions are often useful.

The Friedrichs extension A_F has the property that the domain of its square root $(A_F)^{\frac{1}{2}}$ consists of all vectors $\phi \in \mathcal{H}$ for which there is a sequence $\phi_n \in D(A)$ such that

$$\lim_{n\to\infty} |\phi_n - \phi| = 0, \qquad \lim_{n,m\to\infty} (A(\phi_n - \phi_m), (\phi_n - \phi_m)) = 0$$

On its domain A_F is the restriction of A^* to $D(A^*) \cap D(A_F)^{\frac{1}{2}}$.

The Krein-von Neumann extension A_0 has the property that its domain are the vectors $\phi \in \mathcal{H}$ such that

$$|A_0\phi|^2 = \sup_{\psi\in D(A)} \frac{|(A\psi, \phi)|^2}{(A\psi, \psi)} < +\infty \tag{74}$$

One can verify that if A is densely defined with closed range $R(A)$, its von Neumann extension A_0 is defined by

$$A_0(\psi - \xi) = A(\psi) \quad \psi \in D(A), \quad \xi \in R(A)^{\perp} \tag{75}$$

The following result is often useful [2, 6, 7, 9].

For $a > 0$ define

$$A_a \equiv (A + a)_0 - aI$$

Then in resolvent sense

$$\lim_{a\to 0} A_a = A_0, \qquad \lim_{a\to\infty} A_a = A_F \tag{76}$$

This is the reason why the Krein extension is called A_0 and often the Friedrichs extension is called A_∞.

References

1. Alonso, A., & Simon, B. (1980). The Birman-Krein-Visik theory of self-adjoint extensions. *Journal Operator Theory*, *4*, 251–270.
2. Ando, T., & Nishio, K. (1970). Positive self-Adjoint extensions of positive symmetric operators Tohoku. *Mathematics Journ*, 22, 65–75.
3. Friedrichs, K. (1934). Spectraltheorie halbeschranker operatoren. *Mathematics Annals*, *109*, 465–487.
4. Kato, T. (1978). Trotter's product formula for contracton semigroups topics in functional analysis. In I. Godberger & M. Kac (Eds.), (pp. 185–195). New York: Academic Press.
5. Kato, T. (1980). *Perturbation theory for linear operators*. Heidelberg: Springer.

6. Krein, M. (1947). The theory of self-adjoint extensions of semibounded operators. *Record Mathematics*, *20*, 431–495.
7. Simon, B. (1971). Quantum Mechanics for Hamiltonians defined as Quadratic Forms Princeton Series in Physics.
8. Reed, M., & Simon, B. (1978). *Methods of mathematical physics* (Vols. 1 & 2). New York: Academic press.
9. Yoshida, K. (1981). *Functional Analysis*. New York: Springer.

Lecture 8: Properties of Free Motion, Anholonomy, Geometric Phase

In this Lecture we start the analysis of the Schrödinger equation and describe the properties of some of its solutions. We begin with the case $V(x) = 0$ (free Schrödinger equation) because in this case the analysis is particularly simple and instructive. The properties of the solutions will be important also for the case $V(x) \neq 0$.

Notice that in the case $V(x) = 0$, $x \in R^d$, existence and continuity in any dimension d are guaranteed for initial data that belong to spaces *which admit Fourier transform as continuous map*. In particular they hold in any dimension for initial data in $\mathcal{S}'$, the space of tempered distributions, dual to the space $\mathcal{S}$ of C^∞ functions in R^d which decrease at infinity faster than any inverse power of $|x|$. But they do not hold for initial data in $\mathcal{D}'$ (space of distributions that are not tempered) dual to the space of C^∞ functions with compact support.

Indeed, taking Fourier transform, and denoting by $\hat{\phi}(k)$ the Fourier transform of $\phi(x)$, when $V = 0$ the Schrödinger equation reads, in units in which $\hbar = 1$

$$i\frac{d\hat{\phi}(k;t)}{dt} = \frac{1}{2}|k|^2\hat{\phi}(k;t) \quad k \in R^d \tag{1}$$

The solution (unique and global in time) is for $\phi \in \mathcal{S}'$ (and therefore $\hat{\phi} \in \mathcal{S}'$)

$$\hat{\phi}(k;t) = e^{-it\frac{|k|^2}{2}}\hat{\phi}(k;0) \tag{2}$$

This proves that for $V = 0$ the solution exists unique for initial data in $\mathcal{S}'$ and that it is continuous in t for the (Frechet) topology of this space. Moreover the L^2 norm is not modified by the multiplication by a phase factor and Fourier transform is an isometry in $L^2(R^d)$. Therefore, if one denotes by ϕ_t the solution of (1) with initial datum $\phi_0 \in L^2(R^d)$, the map $\phi_0 \to \phi_t$ is unitary for every choice of $t \in R$ and continuous in t. The group property is satisfied since for every k

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_8

$$e^{it\frac{|k|^2}{2}}\ e^{is\frac{|k|^2}{2}} = e^{i(t+s)\frac{|k|^2}{2}} \tag{3}$$

We conclude that for $V = 0$ the solution of the Schödinger equation (1) defines a one-parameterone-parameter group of unitary operators on $L^2(R^d)$ continuous in the strong topology. It is not continuous in the norm topology in t because the function $|k|^2$ in the phase factor in (2) is not a bounded function.

Remark that when $V = 0$ we have studied existence and uniqueness of the solution regarding (1) as a differential equation *regardless of the role in Quantum Mechanics*. In particular we have not used the theory of operators on a Hilbert space.

The simple form of Eq. (1) in Fourier space allows for a *detailed analysis* of the properties of its solutions; this will be very useful in the study of the case $V(x) \neq 0$. It is convenient to go beyond the $L^2(R^d)$ setting. For initial datum $\phi_0 \in \mathcal{S}'$ the solution of

$$i\frac{\partial}{\partial t}\phi(x,t) = -\frac{1}{2}\Delta\phi(x,t) \quad \phi(x,0) = \phi_0(x) \quad x \in R^d \tag{4}$$

is uniquely given for each time by

$$\phi(x,t) = (\frac{1}{2\pi})^{\frac{d}{2}} \int e^{ixk - it\frac{k^2}{2}}\hat{\phi}(k,t)dk \tag{5}$$

where we have denoted by $\hat{u}$ the Fourier transform of $u \in \mathcal{S}'$.

Definition 1 (*free propagator*) We call *free propagator* the family $U_0(t),\ t \in R$ of linear operators on $\mathcal{S}'$ defined by

$$(U_0(t)\phi_0)(x) \equiv \phi(x,t) \tag{6}$$

We shall denote by $\mathcal{L}(K)$ the collection of linear operators on the Frechet space K and by $H^\sigma(R^d)$ the Sobolev space of order $\sigma \in R$ (the collection of function $u(x)$, $x \in R^d$, for which $\int |\hat{u}(k)|(1+|k|^2)^{2\sigma}dk < \infty$). In particular $L^2(R^d) \equiv H^0(R^d)$.

Note that $\mathcal{S}'(R^d) = \cup_{\sigma<0}H^\sigma(R^d)$ and $\mathcal{S}(R^d) = \cap_{\sigma>0}H^\sigma(R^d)$. In particular the Sobolev norms are seminorms in the (Frechet) space $\mathcal{S}(R^d)$.

Theorem 1 *The free propagator satisfies*

(1) $U_0(t) \in \mathcal{L}(\mathcal{S}')$ and the map $\phi \to (U_0\phi, \psi)$ is C^∞ for every pair $\phi,\ \psi \in \mathcal{S}'$
(2) $U_0(t)$ satisfies the group property $U_0(t)U_0(s) = U_0(t+s)$ and $U_0(0) = I$.
(3) For every value of σ, $U_0(t)$ is a one-parameter group of unitary maps on $H^\sigma(R^d)$.
(4) $U_0(t)\mathcal{S} = \mathcal{S}$, $U_0(t)\mathcal{S}' = \mathcal{S}'$ and the map $t \to U_0(t)\psi$ is continuous in t.
(5) If $\psi \in \mathcal{S}$ one has

$$(U_0\psi)(x,t) = e^{\mp i\frac{d\pi}{4}}(\frac{1}{2\pi|t|})^{\frac{d}{2}} \int e^{-\frac{i(x-y)^2}{2t}}\psi(y)dy \quad \pm t > 0 \quad x \in R^d \tag{7}$$

Proof Properties 1–4 are an immediate consequence of the properties of Fourier tranform. Property 5 is a consequence of the following easy result: let $Re\ a \geq 0$ and $Ima \neq 0$. If one chooses the branch of $\sqrt{a}$ in such a way that $\sqrt{a} > 0$ if $a > 0$ one has

$$(\mathcal{F}e^{-a\frac{x^2}{2}})(k) = a^{-\frac{d}{2}}e^{-\frac{k^2}{2a}} \qquad \heartsuit$$

The operator $G_0(t)$ defined by

$$(G_0 f)(t) \equiv \int_0^t U_0(t-s)f(s)ds, \qquad t \in I \subset R_+ \tag{8}$$

has the following properties. For every $k \in R$

$$f \in L^1(I, H^k) \Rightarrow G_0 f \in AC(I, H^k), \qquad i\frac{d}{dt}Gf = -\Delta\, Gf \tag{9}$$

(AC is the set of absolutely continuous functions).

This regularity property still holds if the operator $-\Delta$ is substituted with $-\Delta + V(x)$ with a potential V which belongs to a suitable regularity class. This fact will be useful in the study of the Schrödinger equation.

In particular one has the following regularity result.

Lemma *If $f \in L^1(I, H^k)$ and g is a solution of $\frac{dg}{dt} = i\Delta g + f$ with $g_0 \in H^{k+2}$, one has*

$$g \in AC(I, H^{k+2}), \quad g(t) = U_0(t)g_0 + G_0 f. \tag{10}$$

1 Space-Time Inequalities (Strichartz Inequalities)

While $U_0(t)$ is an isometry in L^2 for each fixed value of t, if one considers its dependence on time one derives useful *interpolation properties* between spaces of functions of *space and time*.

These properties are called *dispersive* since they correspond roughly to the fact that in the course of time the solution becomes more regular while its *essential support* grows linearly. Therefore some average of the function over a time interval may be more regular then the function itself. Recall that the essential support of $\psi \in L^2(R^d)$, $|\psi|_2 = 1$ can be defined as the minimal radius R_ϵ of a ball B_R^d for which

$$\int_{R^d - B_R}^d |\psi(x)||^2 dx = \epsilon \tag{11}$$

when ϵ is very small.

Reference to the *essential support* is necessary because the flow described by the Schrödinger equation is *dispersive*: even if the initial datum ψ_0 has compact support, the support of $U_0(t)v_0$ is not contained in any finite ball of R^d for any $t > 0$.

The following important *space-time inequality* (Strichartz inequality) plays an important role in scattering theory. If $p \geq 2$ there is a constant C_p such that

$$\|U_0(t)\psi\|_p \leq C_p|t|^{-d(\frac{1}{2}-\frac{1}{p})}\|\psi\|_q, \qquad \frac{1}{p}+\frac{1}{q}=1 \tag{12}$$

This bound provides a stronger decay in time if p increases. Notice that this requires q to decrease since $\frac{1}{p}+\frac{1}{q}=1$. The bound (12) is obtained by interpolation of (7) in R^{d+1}: the operator $U_0(t)$ is isometric in $L^2(R^d)$ and for $t > 0$ *it is bounded from* L^1 to L^∞ since its kernel has the form $\frac{g_d(t)}{t^{\frac{d}{2}}}$ where $g_d(t)$ is bounded.

We shall prove that similar estimates hold for the propagator $e^{it(H_0+V)}$ under suitable conditions on the potential V, in particular that it decays fast enough at infinity. In this Lectures we will give some elements of scattering theory and we will see the role of Strichartz inequalities.

2 Asymptotic Analysis of the Solution of the Free Schrödinger Equation

From the structure of the free propagator one derives the asymptotic behaviour in time of the solutions of the free Schrödinger equation. We shall see that, on a suitable scale, it differs little from the free classical motion along the directions of momentum given by the Fourier transform of the initial wave function. This property will be very useful in scattering theory.

For $t \neq 0$ define the operators $M(t)$ and $D(t)$ as follows

$$M(t)f(t)(x) = e^{-\frac{ix^2}{2t}}f(x), \qquad D(t)f(x) = |t|^{-\frac{d}{2}}f(\frac{x}{t}) \tag{13}$$

The symbol D has been chosen because it corresponds to time dilation. M represents a phase which is invariant under scaling $t \to \lambda^2 t, x \to \lambda x$.

Lemma 2

(a) For $|t| \neq 0$ both $M(t)$ and $D(t)$ are isomorphisms of $\mathcal{S}'$ and of $\mathcal{S}$ and unitary operators on $L^2(R^d)$.

(b) The propagator $U_0(t)$ is given by

$$U_0(t) = e^{\mp i\frac{d\pi}{4}}M(t)D(t)\mathcal{F}M(t) \tag{14}$$

where we have denoted with $\mathcal{F}$ the Fourier transform.

Proof Point (a) is obvious. For point (b) use the identity

$$e^{i\frac{(x-y)^2}{2t}} = e^{i\frac{x^2}{2t}} e^{-i\frac{xy}{t}} e^{i\frac{y^2}{2t}}$$

and the explicit representation of $U_0\psi$ given above. ♡

Theorem 3 *Define $T(t)$ for $\pm t > 0$ as*

$$(T(t)\phi)(x) = e^{\mp\gamma(d)} \frac{e^{i\frac{x^2}{2t}}}{t^{\frac{d}{2}}} \hat{\phi}(\frac{x}{t}) \qquad \gamma_d > 0. \tag{15}$$

The operators $T(t)$ are unitary on $L^2(R^d)$ and for every $\phi \in L^2(R^d)$ one has

$$\lim_{t\to\infty} \|[(U_0(t) - T(t))]\phi\|_2 = 0 \tag{16}$$

Before giving the proof of Theorem 3 we stress its consequences for the description of dynamics in Quantum Mechanics. According to Born's interpretation $|\phi(x_0)|^2$ and $|\hat{\phi}(k_0)|^2$ are the probability densities to obtain x_0 (respectively k_0) as the result of the measurement of position (resp. momentum) on the system described by the wave function ϕ (normalized to one).

Theorem 3 states that the probability distribution in momentum space and also in configuration space (after a suitable scaling) approaches asymptotically when $t \to \infty$

$$\frac{1}{t^d}|\tilde{\phi}(\frac{x}{t})|^2 dx = |\hat{\phi}(\xi)|^2 d\xi, \qquad \xi = \frac{x}{t} \tag{17}$$

This is the distribution in position at time t of *an ensemble of classical particles* placed at time 0 at the origin with momentum distribution $|\hat{\phi}(\xi)|^2 d\xi$.

Proof of Theorem 3 It is sufficient to give the proof in case $\phi(x) \in \mathcal{S}(R^d)$. Indeed, since $\mathcal{S}(R^d)$ is dense in $L^2(R^d)$ and both $U_0(t)$ and $T(t)$ are unitary operators, (16) holds for all $\phi \in L^2(R^d)$ by the theorem of Banach-Steinhaus.

To prove (16) if $\phi \in \mathcal{S}(R^d)$ we remark that in this case we can expand the exponentials (15) and (16) to any order with a uniform estimate on the remainder. One verifies that the expansion of the difference in powers of t satisfies to any order n the bound

$$\|(U_0(t) - T(t))\phi_n\| \leq \frac{1}{n!}\frac{ix^2}{2t}^{n+1} \int_0^1 \|(1-\theta)^n e^{\frac{ix^2\theta}{2t}} \phi(x)\| d\theta \tag{18}$$

For every $\phi \in \mathcal{S}(R^d)$ the right hand side of (18) converges to zero. Therefore (16) holds in $S(R^d)$ and therefore holds in L^2 since both terms are uniformly bounded. ♡

3 Asymptotic Analysis of the Solution of the Schrödinger Equation with Potential V

One can use the asymptotic behavior of the solution of the free Schrödinger equation to derive results on the asymptotics of the solutions of the equation

$$i\frac{\partial}{\partial t}\phi = H_0\phi + V\phi$$

under the assumption that the potential $V(x)$ decreases fast in space. This is a first step to establish Scattering Theory, a subject that will be treated in detail in the second part of these Lectures.

Here we shall make use of the dispersive properties of the integral kernel of the operator $U_0(t)$ to study the behavior of the solutions of Schrödinger's equation with Hamiltonian $H_V = -\Delta + V(x)$ under the assumption that the potential $V(x)$ is smooth and has compact support. These assumptions on $V(x)$ can be greatly relaxed.

We shall use the notation

$$H_0 = -\frac{1}{2}\Delta, \qquad H_V = H_0 + V$$

We already proved that the operators e^{itH_V} form a strongly continuous group of unitary operators. We want to prove that for each $\psi_0 \in L^2(R^3)$ there exists *a unique* $\phi_\pm \in L^2(R^3)$ such that

$$\lim_{t\to\pm\infty} |e^{-itH_0}\psi_0 - e^{-itH_V}\phi_\pm|_2 = 0 \tag{19}$$

This means that for any initial datum $\psi_0 \in L^2(R^3)$ there exists two *unique* elements $\phi_\pm \in L^2(R^3)$ with the property that if one takes $\phi_\pm$ as initial data for the free motion and ψ_0 as initial datum for the true motion, the difference of the two vectors tends to zero in L^2 for $t \to \pm\infty$.

Since we have seen that $\limsup_{t\to\infty} \int_\Omega |e^{it\Delta}\psi(x)|^2 dx = 0$ for every compact $\Omega \in R^3$ it is difficult to visualize the meaning of (19) since both terms tend to zero locally when $t \to \infty$. It is convenient to make use of the fact that both e^{itH_0} and e^{itH} are unitary operators to study instead of (19) the problem of existence and uniqueness of $\phi_\pm \in L^2(R^3)$ such that

$$s - \lim_{\pm t\to\infty} |\phi_\pm - e^{itH_V}e^{itH_0}\psi|_2 = 0 \tag{20}$$

($s-$ lim stands for strong limit) or the existence and uniqueness of $\psi_\pm$ such that

$$s - \lim_{\pm t\to\infty} |\psi_\pm - e^{itH_0}e^{itH_V}\phi|_2 = 0 \tag{21}$$

In turn these problems can be stated as the existence, on suitable domains, of the operators

$$W_{\pm}(H_V, H_0) = \lim_{t \to \pm\infty} e^{itH_V} e^{itH_0} \tag{22}$$

and

$$W_{\pm}(H_0, H_V) = \lim_{t \to \pm\infty} e^{itH_0} e^{itH_V} \tag{23}$$

These operators, *if they exist*, do not have in general the entire space $L^2(R^3)$ as their domain.

For example, if $\phi_E \in L^2(R^3)$ satisfies $H_V \phi_E = E\phi_E$ (it is a *bound state* of $H_0 + V$) the strong limit in (23) does not exist (the weak limit is the null vector). Indeed, due to the dispersive property of e^{-itH_0}, the norm $e^{-itH_0} e^{itE} \phi_E$ restricted to any compact set tends to zero for $t \to \infty$. It follows that the domain of the operator $W_{\pm}(H_0, H_V)$ cannot contain the bound states of $H + V$.

We shall come back later in more detail to the conditions that the potential V must satisfy to guarantee existence and suitable properties of the operators $W_{\pm}(H_V, H_0)$ and $W_{\pm}(H_0, H_V)$ (called wave operators), and in general to provide a solution to the *Scattering Problem* in Quantum Mechanics. We limit ourselves here to prove the existence of $W_{\pm}(H_V, H_0)$ when V is bounded and has compact support.

We do this by making use of the dispersive properties we have proved for e^{-itH_0}. The proof follows the lines of the proof of a theorem of Cook-Kuroda which holds under more general conditions.

Let us remark that the following conditions are satisfied.

(1) For a dense set $\mathcal{D}$ in $L^2(R^3)$ one has

$$\forall \phi \in \mathcal{D},\ \exists t_0(\phi)\ :\ e^{-itH_V} \phi \in D(H_0) \cap D(H_V), \quad t_0 \le t \le \infty \tag{24}$$

where $D(H)$ is the domain of H.

One can take, e.g. as $\mathcal{D}$ the space of twice differentiable functions with compact support.

(2) If $\phi \in \mathcal{D}$ then $V e^{-itH_0} \phi$ is a continuous function of in t in $t_0 \le t \le \infty$.

(3) If $\phi \in \mathcal{D}$ then

$$\int_{t_0}^{\infty} |V e^{-itH_1} \phi|_2 dt < \infty. \tag{25}$$

To prove (25) remark that from the Strichartz estimates we have provided one derives

$$|(e^{-itH_0}\phi)(x)| \le [\frac{1}{4\pi|t|}]^{\frac{3}{2}} \int |\phi(y)| dy \tag{26}$$

and therefore for all t_0 and for all $\phi \in L^1$

$$\int_{t_0}^{\pm\infty} |V e^{-itH_0} \phi|_2 dt \le |\phi|_1 \int_{t_0}^{\pm\infty} (\frac{1}{4\pi|t|})^{\frac{3}{2}} |V|_2 dt < \infty \tag{27}$$

We have therefore proved (25) if $\phi \in \mathcal{D}$ but the proof extends to L^2 since $\mathcal{D}$ is dense in L^2 and all operators are bounded.

The existence of the limits that define $W_\pm(H_V, H_0)$ follows from (1), (2), (3) and the fundamental theorem of calculus applied to the operator-valued function $e^{-itH_V}e^{-itH_0}$ (by integrating its derivative between t_0 and $\pm\infty$).

4 Duhamel Formula

If V is a bounded operator, $\|V\| < C$, the solution of the Schrödinger equation can be written as a convergent series. It is easy to see that $\phi_t \equiv e^{-it(H_0+V)}\phi_0$ solves

$$\phi_t = e^{-itH_0}\phi_0 + i\int_0^t e^{-i(t-s)H_0}Ve^{-isH}\phi_0 ds \tag{28}$$

(consider the equation for the function $e^{-itH_0}\ e^{it(H_0+\epsilon\ V)}\phi_0$).

Equation (28) can be written in the equivalent form

$$e^{-itH_0}\phi_t \equiv \psi_t \qquad \psi_t = \psi_0 + i\int_0^t e^{isH_0}Ve^{-isH_0}\psi_s ds \tag{29}$$

and provides the *interaction representation* frequently used in Theoretical Physics.

Equation (29) (namely the fundamental equation of Calculus) is known in the mathematical literature as *Duhamel formula*. By iteration one obtains the formal series

$$\phi_t e^{-itH_0}\phi_0 + \sum_{k=1}^{\infty} i^k \int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_k} ds_k e^{-i(t-s_1)H_0}V$$

$$e^{-i(s_1-s_2)H_0}V \cdots e^{-i(s_{k-1}-s_k)H_0}Ve^{-is_kH_0}\phi_0 \tag{30}$$

called *Duhamel's series*.

If V is a bounded operator the series (30) converges in the topology of $\mathcal{H}$. Indeed one has

$$\|e^{-i(t-s_1)H_0}Ve^{-i(s_1-s_2)}V \cdots e^{-i(s_{k-1})-s_k)H_0}e^{-is_kH_0}\phi_0\| \le v^k\|\phi_0\|, \quad v = \|V\|$$

therefore

$$\int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{k-1}} ds_k e^{-i(t-s_1)H_0}Ve^{-i(s_1-s_2)H_0}V \cdots VH_0e^{-is_kH_0}\phi_0$$

$$\le v^k\|\phi_0\| \int_0^t ds_1 \int_0^{s_1} ds_2 .... \int_0^{s_(k-1)} ds_k = \frac{t^k\ v^k}{k!}\|\phi_0 \tag{31}$$

Using the convergent series (31) in (ϕ_t, ϕ_t) one obtains $(\phi_t, \phi_t) = (\phi_0, \phi_0) \quad \forall t$; for each value of t the solution is isometric for all t. In the same way one verifies the group property, and continuity follows from the fact that a uniformly convergent series of continuous functions defines a continuous function.

We have therefore proven that if $V \in \mathcal{B}(\mathcal{H})$ the solution of Schrödinger equation gives a strongly one-parameter group of unitary operators, with generator $H \equiv H_0 + V$. But notice that the existence of the wave operators cannot be proved in this way. In fact one cannot determine through Duhamel's formula the asymptotic behavior of the solution since Duhamel's series is not uniformly convergent in time.

In most problems in Physics the potential V is unbounded and *it is not possible* to follow the approach through Duhamel's formula.

Later in these Lectures we shall discuss estimates that allow to conclude, for a wide class of potentials V, that the Hamiltonian $H = H_0 + V$ is the generator of a one-parameter group of unitary operators and that the solutions of the corresponding Schrödinger equation exist for all times and is unique. We shall also show that for potentials that are sufficiently regular and decrease fast enough at infinity the asymptotic behavior of the solution is similar to the one for $V = 0$.

Remark that an unbounded Hamiltonian induces an unbonded derivation on $\mathcal{B}(\mathcal{H})$; also in Hamiltonian mechanics Hamiltonian vector fields may correspond to unbounded derivations in phase space (Poisson brackets are bounded only for differentiable functions) and the ensuing differential equations do not admit always unique global solutions. In particular if the vector field in not of Lipshitz class, the solution may be not unique.

To compare the conditions on the potential for the existence of solutions, remark that in Quantum Mechanics existence and uniqueness of the solution holds if V is only bounded and measurable and *therefore its gradient is not even defined*. On the other hand, if the potential is smooth *almost everywhere* but has strong local singularities, the family of unitary operators $U(t)$ may not exists, while the classical equation may have global solutions *for almost all initial data*.

5 The Role of the Resolvent

There is a further estimate which can be derived from (21) and proves useful in studying the properties of the solutions of the Schrödinger equation; it is an estimate on the *resolvent*.

Recall that for a selfadjoint operator H on the Hilbert space $\mathcal{H}$ the resolvent is the operator-valued function

$$R(z) = \frac{1}{H - z} \qquad z \notin \sigma(H)$$

where $\sigma(H)$ is the spectrum of H, a subset of the real line.

Taking Laplace transforms it easy to see that if the generator is $-\frac{1}{2}\Delta$ the resolvent defines a function, with values in bounded operators, which is analytic in the complex plane cut along the positive real line.

One can verify that the boundary values on the positive real line, both from above and from below, *are not* continuous operator-valued functions. But from (12) one can prove that the integral

$$\int_0^\infty e^{-it(H_0-z)}dt \qquad z = a + i\epsilon, \quad \epsilon > 0$$

is absolutely convergent when $\epsilon \to 0$ if it is regarded as an operator between the spaces $\mathcal{K}_p$ and $\mathcal{K}_{-p}$ where

$$\mathcal{K}_p = \{\psi \in L^2(R^d), \quad \int_{R^d} (1 + |x|^2)^p |\psi(x)|^2 dx < \infty\}$$

and p is taken sufficiently large (depending on the dimension d).

A similar analysis in the lower half plane allows us to conclude that if p is sufficiently large the following limits exist as bounded operators on $L^2(R^d)$:

$$\lim_{\epsilon \to \pm 0} M_p R_0(x \pm i\epsilon) M_p \qquad R_0 = \frac{1}{\Delta - x - \pm i\epsilon} \tag{32}$$

where M_p is multiplication by $(1 + |x|^2)^{-p}$.

At least formally,

$$\lim_{\epsilon \to 0} \frac{1}{a + i\epsilon} - \frac{1}{a - i\epsilon} = 2\delta(a) \tag{33}$$

Equation (33) gives information on the spectrum of H_0. In the case of H_0 this information is trivial, but this procedure is relevant for similar estimates in the case of $H = H_0 + V$, $V \neq 0$. Indeed when $V \neq 0$ existence and uniqueness cannot be proven so simply, and it is not in general possible to give an explicit form for the solution for large times. Also the spectral properties of H are not easy to obtain.

6 Harmonic Oscillator

An important case in which the solution of the Schrödinger equation can be written in closed form is the harmonic oscillator. Its equation is (in suitable units)

$$i\frac{\partial \phi}{\partial t} = -\frac{1}{2}\Delta\phi + \frac{1}{2}|x|^2\phi, \qquad \Delta = \sum_{i=1}^{N} \frac{\partial^2}{\partial x_i^2} \tag{34}$$

and it corresponds to the operator $H_{osc} = -\frac{1}{2}\Delta + \frac{1}{2}|x|^2$ defined on a dense subset of $L^2(R^N)$. It is easy to verify that the operator H_{osc} is self-adjoint and has discrete spectrum.

The eigenstates of H_{osc} are products of Hermite functions

$$h_{k_1}(x_1)h_{k_2}(x_2)....h_{k_N}(x_N) \quad k_j = 1, \ldots, N$$

and the corresponding eigenvalues are $\sum_{j=1}^{N} \sum_{n_j=1}^{\infty}(n_j + 1)$.

Recall that the Hermite functions $h_n(x), x \in R$, are a complete orthonormal basis of $L^2(R)$ and are given by

$$h_n(x) = C_n(x - \frac{d}{dx})^n e^{-\frac{x^2}{2}}, \quad n = 0, 1, 2, \ldots \tag{35}$$

where C_n is a normalization constant chosen such that $\int h_n(x)^2\, dx = 1$. From this one can derive the explicit solution for any initial datum in $L^2(R^N)$.

For other unbounded potentials the proof of existence of solutions requires estimates derived from the regularization properties of the integral kernel of the operator $(-\Delta + I)^{-1}$ (i.e. the fact that $F_f(x) \equiv \int(-\Delta + I)^{-1}(x, y)f(y)dy$ has stronger regularity properties than $f(x)$). Notice that $\hat{F}_f(p) = (p^2 + 1)^{-1}\hat{f}(p)$.

7 Parallel Transport. Geometric Phase

Anholonomy is a *geometric property*. If a system depends on some external parameters which vary periodically some of the quantities that enter in the description of the system may not return to their initial values after a complete cycle.

This fact is well known in Geometry and in Classical Mechanics. A typical example is the parallel transport of a unit vector for which the base point traces a closed curve on a sphere in R^3.

Let $\{\mathbf{r},\ \mathbf{e}_1,\ \mathbf{e}_2\}$ be a reference frame given by a triple of orthogonal unit vectors in which $\mathbf{r}$ is the unit vector from the center of the sphere to the base point P and $\mathbf{e}_1,\ \mathbf{e}_2$ are tangent to the sphere at P. In the reference frame which turns with constant angular velocity ω directed along $\mathbf{r}$ the law of motion is

$$\dot{\mathbf{e}} = \omega \wedge \mathbf{e}$$

Parallel transport is given by the law $\omega = c\mathbf{r} \wedge \frac{d\mathbf{r}}{dt}$ where c is constant. On a sphere this reads

$$\frac{d\mathbf{e}}{dt} = (c(\mathbf{r}) \wedge \frac{d\mathbf{r}}{dt}) \wedge \mathbf{e} = -c(\mathbf{e}, \frac{d\mathbf{r}}{dt})\mathbf{r} \tag{36}$$

Set $\mathbf{e}(t) = e_1(t)\mathbf{e_1} + e_2(t)\mathbf{e_2}$ and define a complex unit vector ϕ as $\phi = \frac{1}{\sqrt{2}}(e_1 + ie_2)$. One has

$$(\phi, \dot{\phi}) = 0 \tag{37}$$

Choose now on the unit sphere a local basis and assign to each point the coordinates u and v and therefore the complex number of norm one $n = u + iv$. It is easy to see that

$$\phi(t) = e^{-i\theta(t)}\, n(\mathbf{r}(t)) \tag{38}$$

where θ is the angle by which the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ must be rotated to coincide with $\{\mathbf{e}_1(t), \mathbf{e}_2(t)\}$. We therefore describe the motion as a transformation of coordinates in a rotating frame. From (37) and (38) one verifies that $\dot{\phi} = Im\ (\bar{n}, \dot{n})$. Integrating along the closed curve $\mathcal{C}$ traced by the base point on the sphere

$$\theta(\mathcal{C}) = Im \oint_{\mathcal{C}} (n, dn) = -\oint_{\mathcal{C}} (\nu, d\nu) \tag{39}$$

where we have denoted by ν the normal to the curve $\mathcal{C}$ on the sphere oriented opposite to the center. We now make use of Stokes' theorem to write

$$\theta(\mathcal{C}) = \int\int_{\mathcal{S}} Im\ (d\bar{n} \wedge n) ds \tag{40}$$

Here $\wedge$ is the wedge operation on $C \equiv R \oplus R$. We remark that the *holonomy* $\theta(\mathcal{C})$ *does not depend* on the basis chosen and is represented by a two-form in the space of the parameters θ and t. In order to have a non-zero holonomy the space of parameters must be *at least two dimensional*.

We choose a basis of unit vectors

$$u = \frac{\mathbf{r} \wedge \mathbf{e}_z}{|\mathbf{r} \wedge \mathbf{e}_z|}, \quad \mathbf{v} = \frac{\mathbf{r} \wedge \mathbf{u}}{|\mathbf{r} \wedge \mathbf{u}|}$$

where $\mathbf{e}_z$ is a pre-chosen vector. The motion takes place on a spherical surface of radius $|r|$. We have

$$\theta(\mathcal{C}) = \int\int_{S^3} \frac{(\mathbf{r}, dS)}{|\mathbf{r}|^3} \tag{41}$$

From (41) one sees that $\theta(\mathcal{C})$ is the solid angle defined by $\mathcal{C}$ with respect to the origin (it is also the flux across $\mathcal{C}$ of the magnetic field generated by a unit dipole placed at the origin). Formula (40) can be generalized by deformation to the case of a geodesic motion on a compact non contractible surface.

8 Anholonomy and Geometric Phase in Quantum Mechanics

A similar analysis can be done in Quantum Mechanics *if one regards the wave function as a geometrical object and not as a probability distribution*.

The vector field is now generated by a Hamiltonian which depends on parameters $\alpha_1, \ldots, \alpha_M$; the role of the unit vectors in R^2 is taken by unit vectors in a Hilbert space (Bloch sphere). We consider time as one of the parameters and make the assumption that the dependence on time of the other parameters dbeifferentiable and periodic with period τ.

Let us suppose moreover that for each value of $t \in [0, \tau]$ the Hamiltonian $H(t) \equiv H(\alpha_1(t), \ldots, \alpha_M(t))$ has an *isolated simple eigenvalue* $\lambda(t)$ which is continuous in t. This is the case e.g. if the Hamiltonian is obtained by a small periodic perturbation of a time independent Hamiltonian H which has a simple isolated eigenvalue λ with eigenvector ϕ_0.

To fix notation, we consider the case

$$H_\epsilon(t) = -\Delta + V + \epsilon V_1(t) \tag{42}$$

where the potential $V_1(t)$ is bounded and periodic in time with period τ and ϕ_0 is an eigenvector of H_0 relative to the isolated eigenvalue λ. Without loss of generality we choose $V_1(0) = 0$.

Let $\phi_\epsilon(t)$ be the solution of

$$i\frac{d\phi_\epsilon(t)}{dt} = H_\epsilon(t)\phi_\epsilon(t), \qquad \phi_\epsilon(0) = \phi_0 \tag{43}$$

We take ϵ so small that for all $t \in [0, \tau)$ the Hamiltonian $H_\epsilon(t)$ has an eigenvalue $\lambda(t)$ isolated from the rest of the spectrum and such that $\lambda(0) = \lambda$.

Let $\hat{\phi}_\epsilon(t)$ the eigenvector of $H_\epsilon(t)$ corresponding to the eigenvalue $\lambda(t)$. Under these conditions we will prove that, *up to order* ϵ^2, $\hat{\phi}_\epsilon(\tau)$ represents *the same pure state as* ϕ_0. Therefore *to this order in* ϵ these two vectors differ only by a phase $\theta_\epsilon(\tau)$, called the *Berry phase*. It is also called *geometric phase* because it is connected with the geometry of the problem as in Classical Mechanics (the interaction representation plays the role of the rotating frame).

The result *holds only to order* ϵ; the vector $\hat{\phi}(t)$ has a component $O(\epsilon^2)$ in the space orthogonal to $\phi(t)$. This *leaking out* in a space with infinite dimension is a purely *quantum phenomenon due to the dispersive properties of the Schrödinger equation*.

Since the space orthogonal to $\phi(t)$ is infinite-dimensional in general the leaking is complete for $t \to \infty$: if one waits a sufficiently long time the vectors $\phi(t)$ and $\phi(0)$ *tend to become orthogonal*. If the Hilbert space has finite dimension we expect rather to have recurrence properties.

In the case of a Schrödinger equation, if $V(t, x)$ is periodic in time with period T the function $\phi_\epsilon(t) \in L^2$ solution to

$$i\frac{\partial \phi_\epsilon(t)}{\partial t} = H_\epsilon(t)\phi_\epsilon(t) \qquad H_\epsilon(t) = -\Delta + V(x) + \epsilon V(t, x), \qquad \phi_\epsilon(0) = \phi(0) \quad (44)$$

generically has the property that, for any positive R,

$$\lim_{t\to\infty} \int_{|x|<R} |\phi(t, x)|^2 dx = 0 \quad (45)$$

This can be seen e.g. by considering the Floquet Hamiltonian

$$K_T = i\frac{\partial}{\partial t} + H(t) \quad (46)$$

defined (and self-adjoint) on $L^2([0, T], L^2(R^3))$. The point spectrum of this operator is generically empty (see e.g. [3]) and this leads to (45).

One may however expect to be able to prove that if ϵ (the perturbation) is small enough the solution $\phi_\epsilon(t)$ differs from $e^{itH_0}\phi$ only by a phase *for times of order* $\epsilon^{-\alpha}$ for $\alpha < 1$. In other words we expect to be able to prove that in Quantum Mechanics the difference in phase that we describe is an *adiabatic* property. The definition of Berry phase is therefore *strictly connected to an adiabatic limit*, i.e. to a *very slow* variation of the parameters *in a very long, but finite time interval*.

Notice that in the case of classical Hamiltonian systems if the Hamiltonian $H + \epsilon K = H_\epsilon$ admits for each sufficiently small ϵ an isolated periodic solution the determination of the variation of the period $\tau(\epsilon)$ as a function of ϵ is done with the Lyapunov-Schmidt method through the search for a fixed point of the Lagrange action functional. This efficient method is not available in Quantum Mechanics because there is no phase space; we will limit ourselves to a result that holds only adiabatically.

9 A Two-Dimensional Quantum System

Before describing in general the adiabatic limit we shall consider in detail a system with two degrees of freedom. The Hilbert space is C^2 and therefore we expect that it is possible to make use of recurrence and averaging. This system can be used to describe the motion of a spin-$\frac{1}{2}$ particle in a rotating magnetic field. It is simple enough to permit a rather complete analysis and at the same time relevant because it plays a crucial role in Quantum Information Theory (it represents a *q-bit*).

The systems we describe has had an experimental realization by Bitter and Dubbers by means of a neutron beam, initially polarized along the z axis, that follows a rectilinear guide along the z axis. The spatial motion has uniform speed ϵ. In this case the space orthogonal to any given vector is one-dimensional. Since the time

evolutionis given by unitary operators and the "leaking space" is compact we expect convergence in the mean.

In the reference frame in which the neutrons are at rest the magnetic field rotates slowly around an axis placed at an angle θ with respect to the z axis:

$$B(t)_x = B\sin\theta\cos(2\pi\epsilon t),\ \ B(t)_y = B\sin\theta\sin(2\pi\epsilon t)\ \ B(t)_z = B\ \cos\theta \tag{47}$$

where ϵ is a small parameter. We will consider the *adiabatic limit* $\epsilon \to 0$. The vector $\hat{B}$ moves on a cone along a curve $\mathcal{C}$ with solid angle $\Omega(\mathcal{C}) = 2\pi(1-\cos\theta)$.

Let us notice that this quantum system can be described as follows: the vector representing the state moves in $C^2 \simeq R^4$ on a sphere of radius one (Bloch sphere). If the magnetic field were constant the representative point on the sphere would trace a maximal circle $\mathcal{C}_\theta$ and the phase would depend linearly on time.

The quantum Hamiltonian which describes the system for $\epsilon = 0$ is $H_0 = B\cdot\sigma$ where σ is the vector which has as components the Pauli matrices. The Hilbert space is two-dimensional and the hamiltonian has two simple eigenvalues $\pm|B|$.

The motion in presence of a varying magnetic field is rather complicated but since the Bloch sphere is compact there is in general a unique invariant measure and the mean value theorem can be applied.

The Berry phase $\Theta(\beta)$ is the average in time of difference of the phase of the wave function for $\epsilon = 0$ and $\epsilon \neq 0$

$$\Theta(\beta) = \lim_{N\to\infty}\frac{1}{N}\theta_\epsilon(N\tau) - \beta\tau \tag{48}$$

where τ is the period for $\epsilon = 0$.

10 Formal Analysis of the General Case

We shall begin the discussion of more general cases, in which the Hilbert space is not finite-dimensional, with a non-rigorous analysis which assumes the applicability of the averaging method. This *non-rigorous analysis* serves the purpose of evidencing the relevant assumptions and procedures. This is also a suggestion to verify the applicability of the method in the case in which the variation of the magnetic field is random (with a given probability law); the ergodic properties of the random system can provide a substitute for the mean value theorem.

This formal analysis is the starting point of the approach that was used by M. Born and V. Fock [2] in their proof that an adiabatic result holds in case the Hamiltonian has only discrete non degenerate eigenvalues increasingly separated and under a condition of non-resomance among the eigenvalues.

We shall later give a rigorous version.

We subdivide a given interval (of time) $[0, T]$ in equal parts of length $\frac{T}{N} \equiv \delta$. We will take the limit $N \to \infty$ first and then $T \to \infty$. We can think of approximating the motion described by $H(t)$ in $[0, T]$ by a motion generated in the intervals

$$[((n-1)\delta, n\delta), n = 1, \ldots, N, \qquad K\delta = T$$

by the *time independent* Hamiltonians $H_n^\epsilon \equiv H_\epsilon((n-1)\delta)$.

This approximate evolution is then expressed as

$$\psi(T) = e^{-i\delta H_N^\epsilon} I e^{-i\delta H_{N-1}^\epsilon} \ldots e^{-i\delta H_2^\epsilon} I e^{-i\delta H_1^\epsilon} \phi_0 \tag{49}$$

Assume that for every ϵ the Hamiltonians H_k^ϵ have discrete spectrum and simple isolated eigenvalues (as is the case of the simple model described above). For each time $k\delta$ denote by $\xi_{n,1}^\epsilon$ the simple isolated eigenvalue $E_{n,1}^\epsilon$ of the Hamiltonian H_k^ϵ.

Insert in (49) at each time $n\delta$ the unit operator written as the sum of projections the eigenvectors ξ_{n,m_n}^ϵ, $m_n = 1, \ldots$, of H_k^ϵ. The vector $\psi(T)$ in (49) becomes then the product of sums of terms each of which corresponds to a choice of time $n\delta$ and an element of the basis. Each term is the expectation value of $e^{-i\delta H_n^\epsilon}$ between elements of the basis H_n^ϵ and elements of the basis of H_{n+1}^ϵ and has therefore the form

$$e^{-i\delta E_{n,m}^\epsilon} (\xi_{n+1,m}^\epsilon, \xi_{n,m}^\epsilon) \tag{50}$$

In the simple two-dimensional model described above a crucial simplifying feature was that the basis consists of only two elements, and therefore the sum contains only 2^N elements. In this case it is less difficult to make rigorous the formal manipulations we will perform. In the general case the Hilbert space is infinite-dimensional and when N increases (and correspondingly δ decreases) it is more difficult to give an estimate of (49) sufficient to take the limit when $\delta \to 0$.

11 Adiabatic Approximation

If one applies *formally* the averaging method, and if each Hamiltonian H_n has non-degenerate discrete spectrum without accumulation points, one can verify that the terms in the series which correspond for some value of n to elements of the basis different from ξ_{n,m_1}^ϵ are of order δ^2.

Assume that their sum converges to zero when $T \to \infty$ for a convenient choice of $N = N(T)$. This assumption is supported by the use of formal estimates but it is difficult to give a rigorous proof since it is difficult to verify the convergence of the series uniformly in N.

Convergence of partial sums is easier, but in an infinite-dimensional space the convergence of the projection of a sequence of vectors on each element of a complete orthonormal basis *does not imply convergence of the sequence*.

The contribution of the terms in which for each value of n one chooses the element $\xi^\epsilon_{n,1}$ is

$$e^{-i\sum_n \delta E^\epsilon_{n,1}}\,\Xi^\epsilon(T), \qquad \Xi^\epsilon(T) \equiv \prod_{n=1}^{N}(\xi^\epsilon_{n+1,1}, \xi^\epsilon_{n,1}) \tag{51}$$

Assume that initially the system is in the state ξ_1. The *non-diagonal* term provide a reduction of value of the ξ_1 component since the total result is normalized to one (the evolution is unitary). They represent therefore a *dissipation* of this component. The purpose of the adiabatic hypothesis is to cancel the components along $\xi_k, k > 1$, to order ϵ^2. If these terms do not contribute, the final result is a variation of the phase of the component along e_1. The mean difference between this variation and the *dynamical variation* itE_1 is the Berry phase $\delta(t)$.

If the eigenvalues $E^\epsilon_{n,i}$ are simple and there is no accumulation point the contribution of the off-diagonal terms is of order ϵ^2 [2].

To estimate $\Xi^\epsilon(T)$ we use

$$\begin{aligned}\langle \xi^\epsilon_{n+1}, \xi^\epsilon_n\rangle &= \delta\langle \xi^\epsilon_n + \epsilon\dot{\xi}^\epsilon_n, \xi^\epsilon_n\rangle + O(\epsilon^2) = 1 + \epsilon\langle \dot{\xi}^\epsilon_n, \xi^\epsilon_n\rangle + O(\epsilon^2)\\ &= e^{\epsilon\langle \dot{\xi}^\epsilon_n(t), \xi^\epsilon_n(t)\rangle} + O(\epsilon^2)\end{aligned}$$

Since $\xi^\epsilon_n(t)$ has constant length the exponential factor is purely imaginary and therefore

$$\langle \xi^\epsilon_{n+1}, \xi^\epsilon_n\rangle = e^{iRe(\epsilon\langle \dot{\xi}^\epsilon_n(t), \xi^\epsilon(_nt)\rangle)} + O(\epsilon^2)$$

Notice that for $\epsilon \to 0$, $\dot{\xi}^\epsilon$ gives the evolution law for $\xi(t)$ under the time-dependent Hamiltonian $H_\epsilon(t)$. Therefore in the limit $\delta \to 0$ one has

$$\phi^\epsilon(\tau) = W^\epsilon e^{-i\int_0^\tau E^\epsilon_{n_0}(t)dt}, \qquad W^\epsilon = e^{i\int_0^\tau Re(\langle \dot{\xi}^\epsilon(t), \xi^\epsilon(t)\rangle)dt}\phi(0) \tag{52}$$

The geometric phase factor $\Omega^\epsilon \equiv -i\log\ W^\epsilon$ can be written as

$$\Omega^\epsilon = \int_0^\tau Re\langle d\xi^\epsilon(t) \wedge \xi^\epsilon(t)\rangle dt \tag{53}$$

The differential form is defined in a space which has as coordinates time and the parameters in the Hamiltonian $H_\epsilon(t)$. Notice that the assumption of totally discrete spectrum that has no accumulation points rules out to order ϵ the dispersive character of the Schrödinger equation. If the eigenvalues we consider are dense there is no control of the convergence of the sum in (51) and the estimate fails.

The adiabatic method can be viewed as the search of the time-independet dynamics that *gives the best approximation* to the time-dependent perturbed one.

It is not limited to the case in which the subspace of interest is one-dimensional and can be applied also in case of subspaces of dimension ≥ 2 under the condition that the one-dimensional subspaces cross each other a finite number of times ad remain well separated from the remaining part of the spectrum (one refers to this in the literature by saying that one assumes *a gap condition*). Without this condition one obtains much weaker results.

If the gap condition holds, Berry's phase is substituted by a unitary operator U_{Berry} on the subspace considered. One can show that, *to first order in the small parameter*, the projection operator P on the subspace considered is taken to itself after a period under the Heisenberg flow. Therefore there is a unitary operator U in $P\mathcal{H}$ such that $\sigma(T) = U\sigma(0)$ for every density matrix acting on $P\mathcal{H}$. A finite-dimensional unitary operator can be diagonalized; let ψ_k be the eigenvectors of U, where $k = 1, \ldots, d$ and d is the number of eigenvalues in the collection considered.

It follows that there are phase factors $e^{i\theta_k}$ which are Berry phases if one starts with the state described by ψ_k.

From a mathematical point of view the adiabatic evolution is a parallel transport associated to the spectral subspaces of a time-dependent Hamiltonian. It is important that the spectral subspaces be separated from each other for all times and therefore the method applies usually to small perturbations of a time independent Hamiltonian which has an isolated eigenvalue.

12 Rigorous Approach

Rigorous results can be obtained if one substitutes the attempt to have an approximation valid for all times by estimates up to some (possibly large) power of ϵ for a time of order $\epsilon^{-\beta}$, $\beta > 0$. We give now an outline of this rigorous analysis and we shall give references for a more detailed description.

If the time-dependent interaction is switched off after time of order $\epsilon^{-\beta}$, if the eigenvalue considered is isolated and under further assumptions we will see that the error term in the approximation can be made smaller than any power of ϵ. In the case of level crossing case the degree of approximation is worst, typically polynomial in the small parameter ϵ.

We shall treat only the case in which the spectral subspace of interest is one-dimensional. The Hamiltonian has the form

$$H_\epsilon(t) = H_0 + H_1(t) \tag{54}$$

where H varies slowly in time and $H_1(T) = H_1(0)$ (or $H(t)$ is periodic in time with very large period T). Set $\tau = \epsilon T$. Schrödinger equation takes the form

$$i\frac{d}{d\tau}\phi = \epsilon H_0\phi + \epsilon H_1(\tau)\phi \tag{55}$$

Denote by $P_\epsilon(t)$ the orthogonal projection on an eigenstate of $H_\epsilon(t)$ that varies in time but remains at all times isolated with respect to the remaining part of the spectrum. One has $\dot{P}_\epsilon(t) \equiv O(\epsilon)$. We seek a time-independent hamiltonian that gives an evolution, parametrized by τ and such that $\frac{dP}{d\tau}$ is of order ϵ^2 so that the corresponding evolution approximates the real one to order ϵ^2 and therefore gives an approximation to order ϵ for times of order $\frac{1}{\epsilon}$.

If P is a projection operator one has the following identities

$$\frac{1}{2}[[\dot{P}, P]P] = \dot{P} + \dot{P}P - P\dot{P}P, \qquad \dot{P} = \frac{1}{2}(P\dot{P} + \dot{P}P) \tag{56}$$

Notice (this remark is due to T. Kato) that for any function $f(H)$ the evolution of the projection onto the state we consider is given by the hamiltonian $f(H) - i[\dot{P}, P]$, where f is an arbitrary continuous function. Indeed one has

$$i[f(H) - i[\dot{P}, P], P] = [\dot{P}, P], P = \dot{P}P^2 - 2P\dot{P}P + P^2\dot{P} = \dot{P}P + P\dot{P} = \dot{P} \tag{57}$$

It follows that if we choose in particular as new Hamiltonian the adiabatic approximation

$$H^\epsilon_{ad}(t, P(t)) = H_\epsilon(t) - \frac{1}{2}i[\dot{P}_\epsilon(t), P_\epsilon(t)] \tag{58}$$

we obtain the desired error of order ϵ^2 in the vector field. Under sufficiently general conditions the operator $H^\epsilon_{ad} - H_\epsilon(t)$ is bounded.

This form of the *adiabatic theorem* was rigorously proved by M. Born and V. Fock for the case of discrete non-degenerate spectrum [2]. The proof of the generic case was sketched by Kato [5].

Denote by $U^\epsilon_{ad}(t)$ the evolution operator associated to $H^\epsilon_{ad}(t)$. We make the following assumptions:

(i) $H_\epsilon(t)$ is a family if self-adjoint operator on the Hilbert space $\mathcal{H}$ uniformly bounded below and with domain independent of t.
(ii) For each ϵ the family $H_\epsilon(t)$ is k-times differentiable in t in the strong operator topology.
(iii) For each value of t the spectrum of $H_\epsilon(t)$ has gaps and $P_\epsilon(t)$ is the projection on a band of the spectrum which is separated by an amount $\delta(t)$ from the remaining part of the spectrum, with $\inf_t \delta(t) = \delta_0 > 0$

Under these assumptions, for any interval I of the real axis containing the origin and with the notation

$$P_\epsilon(t) = U_\epsilon(t)\ P\ U^*_\epsilon(t), \quad P(t) = U_{ad}(t)\ P\ U^*_{ad}(t) \tag{59}$$

one can prove ([1], see also [4])

Theorem 4 *For ϵ sufficiently small one has*

(a)

$$P_\epsilon(t) - P(t) = \epsilon C_\epsilon(T), \quad t \in I = [0, T] \tag{60}$$

(b) If in the assumption (ii) one can take $k = +\infty$ then

$$P_\epsilon(t) - P(t) = O(\epsilon^\infty), \quad t \in I - Supp(\frac{dP}{dt}) \tag{61}$$

The constant $C_\epsilon(T)$ in (60) is an increasing function of T and in general $\epsilon C_\epsilon(\epsilon^{-1}) = O(1)$.

We shall give an outline of the proof, since it is *a prototype of adiabatic and multiscale theorems*. Recall that we seek estimates which are valid up to times of order ϵ^{-1}; it is therefore convenient to scale time and write the equation in the form

$$i\frac{\partial U_\epsilon(t)}{\partial t}\phi = \epsilon^{-1} H(t)\ U_\epsilon(t) \quad U_\epsilon(0) = I \tag{62}$$

Under our assumptions this equation has a unique global solution which gives a strongly continuous one parameter family of unitary operators $U_\epsilon(s)$ (it is not a group since the Hamiltonian is time-dependent).

Moreover $U_\epsilon(s)\phi$ is strongly differentiable in s if $\phi \in D(H_s)$ (this domain is independent of s by assumption). The same is true for the *adiabatic* equation

$$i\frac{\partial U_{ad}(t)}{\partial t}\phi = \epsilon^{-1} H_{ad}(t, P)(t)\ U_{ad}(t) \quad U_{ad}(0) = I \tag{63}$$

where H_{ad} is defined in (58).

Let us remark that the evolution $U_{ad}(t, P)$ induced by $H^\epsilon_{ad}(t, P(t))$ decouples $P(t)\ \mathcal{H}$. Indeed

$$U_{ad}(t, P)\ P(t) = P(t)\ U_{ad}(t, P) \tag{64}$$

This relation can be proved noticing that it is true at $t = 0$ and that the time derivative of the difference between the two term is zero for each time t.

The next step is to obtain a-priori estimates on

$$\Omega(t) = U^*_{ad}(t, P)\ U_\epsilon(t) \tag{65}$$

Since $U_{ad}(t, P)$ is unitary, an estimate of Ω provides an estimate of the difference between the two dynamics, as in scattering theory. Using the fundamental theorem of calculus and taking into account that $\Omega(0) = I$ one verifies that $\Omega(t)$ satisfies the Volterra-type equation

$$\Omega(t) = I - \int_0^t K_\epsilon(s, P)\Omega(s)ds, \quad K_\epsilon(s, P) = U^*_{ad}(t, P)\ [\dot{P}(s), P(s)]\ U_{ad}(s, P) \tag{66}$$

By iteration one has, for $t \in [0, T_0] \equiv I_0$

$$\Omega(t) - \sum_{n=1}^{N} \Omega_n(t) = O(\epsilon^N), \quad \sup_{t \in I_0} \|\Omega_n(t)\| = O(\epsilon^{n-1}) \tag{67}$$

(recall that on the right-hand side of Eqs. (62) and (63) there is a factor ϵ^{-1}).

The expression (66) looks like a Duhamel series, but it does not provide a power series expansion in ϵ because the operator $K_\epsilon(s, P)$ depends on ϵ. To obtain a power series expansion recall that

$$P(t) = \frac{1}{2i\pi} \int_\gamma R(z, t) dz, \quad R(z, t) = (H(t) + z)^{-1} \tag{68}$$

where γ is a closed path that encircles the part of the spectrum on which $P(t)$ projects for all times $t \in I$. It is also convenient to make use of the identity which follows from (66):

$$Q(t)\ X(t)\ P(t) = -Q(t)\ ([H_{ad}(t), \tilde{X}(t)] + i\epsilon[\dot{P}(t), \tilde{X}(t)])P(t) \tag{69}$$

where $Q(t) = I - P(t)$.

We will introduce the following notation: for $X \in \mathcal{H}$

$$\tilde{X}(t) \equiv \frac{i}{2\pi} \int_\gamma R(z, t)\ X(t)\ R^{-1}(z, t) dz \tag{70}$$

In this way one obtains an expansion of $\Omega(t)$ as power series in ϵ; the first two terms are

$$\begin{aligned} \Omega_1(t) &= -i\epsilon(U_{ad}^*(t)\ \dot{P}(t)\ U_{ad}(t)) \\ \Omega_2(t) &= -i\epsilon[\int_0^t U_{ad}^*(t)\ \dot{P}(t)\ U_{ad}(t) dt]^2 (P - Q) \end{aligned} \tag{71}$$

With these estimates part (a) of Theorem 4 follows since

$$\dot{\Omega}(t) = -K_\epsilon(t, P)\Omega(t), \quad \Omega(0) = I \tag{72}$$

The proof of part (b) requires better estimates which allow the iteration of the Volterra equation, and will not be given here; for a detailed proof see [4].

References

1. Avron, J., Seiler, R., & Yaffe, L. (1987). *Communications in Mathematical Physics*, *110*, 33–49.
2. Born, M., & Fock, V. (1928). *Zeitschrift für Physik*, *51*, 165–169.
3. Enss, V., & Vesel, K. (1983). *Annales de l'Institut Henri Poinocaré*, *39*, 159–181.
4. Klein, M., & Seiler, R. (1990). *Communications in Mathematical Physics*, *128*, 141–160.
5. Kato, T. (1958). *Journal of the Physical Society of Japan*, *5*, 435–439.

Lecture 9: Elements of C^*-algebras, GNS Representation, Automorphisms and Dynamical Systems

Quantum Mechanics associates to observables self-adjoint operators on a Hilbert space, to pure states orthogonal projections on one-dimensional subspaces and describes the evolution as a one-parameter group of unitary operators.

This simple characterization is insufficient to describe systems with infinitely many degrees of freedom as is the case e.g. of Quantum Statistical Mechanics or Quantum Field Theory. It is therefore convenient to extend the formalism to cover also these cases. The resulting structures are C^*-algebras, their states and their automorphisms [2, 11, 18].

1 Elements of the Theory of C^*-algebras

Notice that the bounded operators $\mathcal{B}(\mathcal{H})$ on the Hilbert space $\mathcal{H}$ form an algebra on which are defined an involution $a \to a^*$ and a norm $a \to |a|$ where $|a| = sup_{\phi\in\mathcal{H};\ |\phi|=1}|a\phi|$.

This norm has the property $|a^*a| = |a|^2,\ |a| = |a^*|,\ |ab| \leq |a||b|$. This property will be at the basis of the new structure.

Definition 1 (*Banach Algebras*) A Banach *-algebra is an algebra $\mathcal{B}$ (with elements $b \in \mathcal{B}$) that admits a Banach space structure with norm $|b|$ and an involution $b \to b^*$ such that $(\lambda b)^* = \bar{\lambda} b^*$ and $(b_1 b_2)^* = b_2^* b_1^*$. ◇

We will describe Quantum Systems by means of a special class of Banach algebras characterized by conditions on the norm. These conditions are satisfied by the algebra $\mathcal{B}(\mathcal{H})$ and are so restrictive that any Banach algebra of this class admits a *faithful* representation as a norm closed subalgebra of $\mathcal{B}(\mathcal{K})$ where $\mathcal{K}$ is a suitable Hilbert space (which need not be separable).

We shall often assume that the algebra $\mathcal{B}$ admits an unity. This condition is not essential: it serves the purpose of simplifying statements and proofs. The Banach

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_9

algebras $\mathcal{B}$ we will consider admits *always* an approximate left unit, i.e. a sequence of elements $\{b_n\},\ n \in \mathcal{N}$ and of positive constants $\{c_n\},\ n \in \mathcal{N}$ such that

$$\forall b \in \mathcal{B} \quad sup_{b\in\mathcal{B}}|b_n - b| \leq c_n, \quad lim_{n\to\infty}c_n = 0.$$

Equivalently one can assume the existence of an approximate right unit or an approximate bilateral unit.

Notice that if $\mathcal{B}$ admits a unit e then $\{e, \cdots, e, \cdots\}$ is a bilateral approximate unit. As an example, the abelian B^*-algebra $l^\infty(Z)$ has no unit but the sequence

$$\{b_m^N\} \quad b_m^N = 1, \quad m \leq N, \quad b_m^N = 0 \quad n > N \tag{1}$$

is a bilateral approximate unit.

Definition 2 (*C^*-algebras*) We define C^*-algebra $\mathcal{A}$ as a Banach *-algebra $\mathcal{B}$ with norm that satisfies

(i)
$$|ab| \leq |a||b| \qquad \forall a,\ b \in \mathcal{A}$$

(ii)
$$|a^*a| = |a|^2 \qquad \forall a \in \mathcal{A}$$

(iii)
$$|e| = 1$$

where e is the unit of the algebra.

If $\mathcal{A}$ does not have a unit one requires $lim_{n\to\infty}|e_n| = 1$ for every approximate unit. ◇

From (i), (ii) one derives $|a^*| = |a|$.

Indeed from $|a^2| = |a^*a| \leq |a^*||a|$ follows $|a| \leq |a^*|$ if $|a| \neq 0$. Inverting the role of a and a^* one derives $|a^*| \leq |a|$ for $|a^*| \neq 0$.

On the other hand $|a^*| = 0 \leftrightarrow |a| = 0$.

We remark that Kadison [9] proved that condition (ii) implies (i) and (iii).

We will prove later in this Lecture that a C^*-algebra has a *faithful* representation as operators on a Hilbert space. Since a Hilbert space is a linear space endowed with a scalar product (and therefore a topology) and this scalar product can be used to define a distance (and therefore a metric: we shall later define the *Bures distance* between states) one sees that in some sense the algebraic condition (ii) originates a metric structure on a linear space.

Conversely, the definition of a scalar product (of a Hilbert space structure) on a linear space Σ leads to the definition of self-adjoint operators acting on Σ and therefore to C^*-algebra of operators.

Examples of C^*-algebras are

(1) The algebra of *canonical anticommutation relations* (fermionic algebra)

$$a_i a_k + a_k a_i = 0, \qquad a_i a_k^* + a_k^* a_i = \delta_{i,k} \quad k, i = 1 \dots N \quad N \leq \infty$$

(2) The Kuntz algebra (spin algebra) with N generators satisfying

$$s_i^* s_j = \delta_{i,j} I \qquad \sum_{i=1}^{N} s_i s^* = I$$

(3) The rotation algebra

$$UV = e^{i\theta} VU \qquad \theta \notin Q$$

A concrete C^*-algebra is a realization of a C^*-algebra structure by means of operators. Simple examples of concrete C^*-algebras are

(1) The algebra of complex-valued $N \times N$ matrices, with conjugation $A \to A^*$ (the adjoint of A) and norm $|A| \equiv sup_{u \in C^N, |u|=1}(u, Au)$.
(2) The algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$ with the natural definition of conjugation and norm.
(3) The algebra of complex-valued bounded, continuous functions $C_b(\Omega)$ on an open domain $\Omega \in R^N$ with conjugation given by complex conjugation and norm $|f| \equiv sup_{x \in \Omega} |f(x)|$.
(4) The algebra generated by the identity and by the compact operators in a separable Hilbert space $\mathcal{H}$.
(5) The Weyl algebra

$$e^{ixQ} e^{ipP} = e^{ixp} e^{ipP} e^{ixQ}$$

where Q and P are self-adjoint operators on a Hibert space $\mathcal{H}$.

Notice that the algebra of canonical commutation relations

$$[Q_h, P_k] = i\delta_{h,k}, \quad [Q_h, Q_k] = [P_h, P_k] = 0 \quad k = 1 \dots N \quad N \leq \infty$$

is not a C^*-algebra since it does not have a norm (every realization is through unbounded operators).

Definition 3 Denote by e the unit of $\mathcal{A}$; we shall call it often the *identity* of $\mathcal{A}$. An element $a \in \mathcal{A}$ is said to be *normal* if $a^* a = a a^*$, *hermitian* if $a = a^*$, *unitary* if $a\ a^* = a^*\ a = e$, *positive* if there exists $b \in \mathcal{A}$ such that $a = b^*\ b$, *projector* if $a = a^*$ and $a^2 = a$. The *inverse* of $b \in \mathcal{A}$ is the element of $\mathcal{A}$ (unique if it exists) which satisfies $a\ b = b\ a = e$. ◇

If the C^*-algebra $\mathcal{A}$ does not have a unit the definitions of unitary and inverse are suitably modified using the approximate units.

Definition 4 (*-*automorphisms*) A *-automorphism of $\mathcal{A}$ is an invertible map on $\mathcal{A}$ which preserves the algebraic structure, the identity (or the class of approximate identities) and commutes with the involution. ◇

Definition 5 (*states*) A *state* of a C^*-algebra $\mathcal{A}$ is a continuous linear functional which takes non-negative values on positive elements of $\mathcal{A}$ and value one on the unit (or which tends to one on any approximate unit). It is therefore an element of $\mathcal{A}^*_{+,1}$. In general we will use the symbol ρ to denote a generic state. ◇

By construction the collection of states has an affine structure. If $\mathcal{A} = \mathcal{B}(\mathcal{H})$ Definition 5 coincides with the definition of state that we have given in the Axioms of Quantum Mechanics.

Definition 6 (*pure states*) The *extremal elements* of $\mathcal{A}^*_{+,1}$ are called *pure states*. A theorem of Choquet assures that every state belong to the convex hull of the pure states. ◇

In the case of the (concrete) commutative algebra $C_b(\Omega)$ (continuous bounded functions on a compact space Ω with the topology of the sup-norm) the states are the Radon measures of total mass one. Pure states are Dirac measures.

Every positive function $g(\omega) \in L^1(\Omega, d\mu)$ with $\int_\Omega g(\omega)d\mu = 1$ defines a state of $C_b(\Omega)$ by

$$\rho_g \; : \; f \to \int f(\omega)g(\omega)d\mu, \qquad \forall f \in C_b(\Omega). \tag{2}$$

where μ is a probability measure.

Also in the case of a generic C^*-algebra (2) holds: every state defines uniquely a (probability) measure on the pure state

A first classification of states comes from their behavior under filtration.

Recall that a *filter* $\mathcal{F}$ in a C^*-algebra $\mathcal{A}$ is a partially ordered set of elements, with partial order denoted by $>$, in which for any two elements $a,\ b \in \mathcal{A}$ there is element $c \in \mathcal{A}$ such that $c > a, c > b$. Moreover there exists in a maximal element C in $\mathcal{F}$.

Here we consider the partial order on the real elements in $\mathcal{A}$ defined by $c > a \equiv c - a > 0$.

Definition 7 (*Normal states*) A state ρ is said to be *normal* if

$$sup_{a\in\mathcal{F}}\rho(a) = \rho(C). \tag{3}$$

◇

All states of the algebra of $N \times N$ complex valued matrices are normal.

For the algebra of continuous functions on $[-1, 1]$ the states that correspond to evaluation of the function in one point *are not normal*. Indeed consider the filter of

continuous positive functions $f(x)$ such that $f_\alpha(x) \leq 1$, $f_\alpha(0) = 0$. The maximal element is the function identically equal to one but the evaluation at the origin of this function does not coincide with the limit of the evaluation of the functions in the filter, which is zero. In this example the normal states are described by probability measures which are absolutely continuous with respect to the Lebesgue measure. These states can be extended to normal states on the algebra of essentially bounded functions on $[-1, 1]$, denoted by L^∞.

Normal states on the commutative algebra $C[-1, 1]$ can be extended to normal states on a larger C^*-algebra, namely $L^\infty[-1, 1]$, and they form its *predual* (recall that the *predual* B_* of a Banach space B is a topological space such that B can be identified with the space of continuous linear functionals on B_*).

Concrete C^*-algebras may have special properties in regard to the states. For example all the states of the algebra $\mathcal{B}(\mathcal{H})$ are normal. Other properties in the non-commutative case that lead to the definition of W^*-algebras and von Neumann algebras.

For a C^*-algebra $\mathcal{A}$ we shall use the following notation.

$\mathcal{A}^*_{1,+}$ are the states (the index 1 denotes normalization, + denotes positivity).

If $\mathcal{A}$ admits a predual $\mathcal{A}_*$, then the elements of $\mathcal{A}_{*,\,1,+}$ are normal states.

2 Topologies

One can define on a C^*-algebra $\mathcal{A}$ two natural topologies.

Definition 8

(a) *Uniform Topology*. A basis of neighborhoods of $a \in \mathcal{A}$ is given by $\{b \in \mathcal{A},\ |b - a| < \epsilon\}$ for all $\epsilon > 0$.

(b) *Weak topology*. A basis of neighborhoods of $a \in \mathcal{A}$ is given by $\{b \in \mathcal{A}, |\rho(b) - \rho(a)| < \epsilon\}$ for all $\epsilon > 0$ and $\rho \in \mathcal{A}^*_{1,+}$. Therefore a sequence $\{a_n\} \in \mathcal{A}$ converges weakly if for any $\rho \in \mathcal{A}^*_{1,+}$ one has

$$lim_{n,m\to\infty}|\rho(a_n) - \rho(a_m)| = 0. \tag{4}$$

$\diamond$

Notice that the uniform topology is finer that the weak one (it contains a larger number of neighborhoods). Uniform convergence implies weak convergence. The two topologies are equivalent iff $\mathcal{A}$ is a finite dimensional space.

For a C^*-algebra the following holds

$$|a|^2 \leq sup_{\rho\in\mathcal{A}_{1,+}}[\rho(a^*a) + \rho(a\ a^*)]. \tag{5}$$

Therefore uniform convergence is equivalent to uniformity of weak convergence (hence the name).

If the algebra $\mathcal{A}$ has a predual $\mathcal{A}_*$ one can define another topology, the weak* (or *vague*) topology.

Definition 9 (*Weak*-topology*) The weak*-topology has as basis of neighborhoods of $a \in \mathcal{A}$

$$\{b \in \mathcal{A}, \quad |\rho(a) - \rho(b)| < \epsilon\} \tag{6}$$

for $\epsilon > 0$ and $\rho \in \mathcal{A}_*$. ◇

By construction the weak* topology is weaker than the weak topology (it has fewer neighborhoods). The two topologies coincide only if the algebra satisfies $\mathcal{A} = (\mathcal{A}^*)^*$

The importance of the weak* topology comes from the following theorem

Theorem 1 (Banach-Alaoglu) *Let X be a Banach space and let X^* be its dual. The unit ball in X^* (denoted X_1^*) is compact for the topology induced by X.* ◇

Proof The weak * topology for X^* is by definition the one under which all elements of X are continuous and is therefore the product topology of $B \equiv \Pi_{x \in X} C_x$ where C_x are copies of the complex plane C. Let $\dot{C}$ be the one-point compactification of C. By construction $\dot{C}$ is compact and therefore $\Pi_{x \in X} \dot{C}_x$ is compact in the product topology.

Since B is closed in this topology, B is compact. Also X_1^* is closed in B in the weak* topology. Indeed if $l_n(x) \to l(x)$, $|l_n(x)| \leq |x| \ \forall n$ then $|l(x)| \leq |x|$. By definition $|l| \equiv sup_{x \neq 0} \frac{|l(x)|}{|x|} \leq 1$. Moreover $l_n(e) = 1, \ \ \forall n \Rightarrow l(e) = 1$. Therefore X_1^* is compact and so is $X_{1,+}^*$ because a converging sequence of positive elements converges to a positive element. ♡

Since states are by definition the positive part of the unit ball $\mathcal{A}^*$ we conclude that the set of states is compact for the topology induced by $\mathcal{A}$ i.e. for the topology in which a sequence $\{\rho_n\} \ \ \rho_n \in \mathcal{A}_{1,+}^*$ is convergent iff for every $a \in \mathcal{A}$

$$lim_{m,n \to \infty} |\rho_n(a) - \rho_m(a)| = 0. \tag{7}$$

Therefore from every sequence of normalized states ρ_n which satisfies (7) it is possible to extract a convergent subsequence. The limit point ρ is a normalized state since $\rho(e) = lim_{n \to \infty} \rho_n(e) = 1$ and for any $b \in \mathcal{A}^*$ one has $\rho(b^*b) \geq 0$ since $\forall n \ \ \rho_n(b^*b) \geq 0$.

It is important to notice that if a sequence is all made of normal states, *the limit state may not be normal*. A simple example is given by a sequence of states on $C_0(R)$ defined by positive elements of $L^1(R)$ that converge *in the sense of measures* to a measure concentrated in one point.

In the same way one can construct a sequence of density matrices σ_n on an infinite dimensional Hilbert space $\mathcal{H}$ which converge in the sense that (the symbol Tr denotes the trace)

$$\forall A \in B(\mathcal{H}) \quad lim_{n,m \to \infty} |Tr(\sigma_n A) - Tr(\sigma_m(A))| = 0 \tag{8}$$

but there does not exist a density matrix σ such that for all $A \in \mathcal{A}$ one has $lim_{n\to\infty}\sigma_n(A) = \sigma(A)$.

In what follows we are going to use frequently the following theorem

Theorem 2 (Hahn-Banach) *Let X be a vector space, p a real-valued convex functional on X*

$$p(\alpha x + (1-\alpha)y) \leq \alpha p(x) + (1-\alpha)p(y), \quad 0 < \alpha < 1 \tag{9}$$

Let Y be a subspace of X and let ρ be a positive functional on X with $\rho(x) \leq p(x) \forall x \in X$. Then ρ can be extended to a linear functional $\tilde{\rho}$ on X dominated by p. ◇

Proof Let $z \in X, z \notin Y$; we will extend ρ to the subspace generated by z and Y. The extension to X is obtained by induction.

Notice that by linearity if $y_1, y_2 \in Y$ and $a, \ b > 0$ then

$$\begin{aligned} b\rho(y_1) + a\rho(y_2) = (a+b)\rho(\frac{b}{a+b}y_1 + \frac{a}{a+b}y_2) \leq (a+b)p(\frac{b}{a+b}y_1 \\ + \frac{a}{a+b}y_2) \leq (a+b)p(\frac{b(y_1-az)}{a+b} + \frac{a(y_2+bz)}{a+b}) \\ \leq bp(y_1 - az) + ap(y_2 + bz) \end{aligned}$$

Dividing by ab one has

$$a^{-1}[\rho(y_1) - p(y_1 - az)] \leq b^{-1}[p(y_2 + bz) - \rho(y_2)]$$

We extend now ρ to z by defining $\rho(z) = b$ where b is chosen in such a way that

$$sup_{a>0,\ y\in Y} a^{-1}[\rho(y) - p(y-az)] \leq b \leq inf_{a>0,\ y\in Y} a^{-1}[p(y+az)) - \rho(y)] \tag{10}$$

By construction $\rho(x) \leq p(z)$ and therefore extending ρ to $z \cup Y$ and taking into account that p is convex one has $l(\xi) \leq p(\xi)$ for every ξ in the subspace generated by z and Y. ♡

In applications p is often a norm (e.g. the norm defined on a Banach space X) and Y can be a one-dimensional subspace.

In this case one has the following important corollary

Corollary 1 *Let $\mathcal{A}$ be a C^*-algebra with unit, a positive element of $\mathcal{A}$. There exists a pure state ρ such that $\rho(a) = |a|$. Therefore* pure states separate $\mathcal{A}$. ◇

Proof Let $\mathcal{A}_0$ be the subspace generated by the identity e and by a^*a. Define on $\mathcal{A}_0$ the following linear functional ρ

$$\rho(\alpha e + \beta a^*a) \equiv \alpha + \beta|a|^2 \quad \alpha, \ \beta \geq 0, \quad \alpha + \beta = 1. \tag{11}$$

By construction $\rho \in (\mathcal{A}_0)_{1,+}$. Using the triangular inequality one sees that for every $y \in \mathcal{A}_0$ one has $\rho(y) \leq |y|$.

Choosing $p(a) \equiv |a|$ the Hahn-Banach theorem proves the existence of an element $\tilde{\rho} \in \mathcal{A}_{1,+}$ with $\tilde{\rho}(a) \leq |a|$ and

$$\tilde{\rho}(\alpha e + \beta a^* a) = \alpha + \beta |a|^2. \tag{12}$$

For $\alpha = 0$ one derives

$$\tilde{\rho}(a^* a) = |a|^2. \tag{13}$$

Let us now consider the convex set Z_a of the states that satisfy (13). This set is not empty due to Hahn-Banach theorem. Since Z_a is convex it admits extremal elements due to a theorem of Choquet. Each of these states is pure and satisfies (13). ♡

It is worth remarking that pure *normal* states may not separate the algebra [10].

The following corollaries of the Hahn-Banach theorem are also worth noting.

Corollary 2 *Let $Y \subset X$, $\rho \in Y^*$. There exists $\tilde{\rho} \in X^*$ such that $\tilde{\rho}$ restricted to Y coincides with ρ and moreover $|\tilde{\rho}|_{X^*} = |\rho|_{Y^*}$. The proof makes use of the (convex) function $p(a) \equiv |a|\rho|_{X^*}$.* ◇

Corollary 3 *For any $x \in X$, $x \neq 0$ there exists $\rho \in X^*$ such that $\rho(x) = |x||\rho|_{X^*}$ For the proof use corollary (1) applied to the subspace $\{cx,\ c \in C\}$.* ◇

Corollary 4 *Let Z be a subspace of X and let $y \in X$ be at finite distance from Z. There exists $\rho \in X^*$, $|\rho| = 1$ $\rho(y) \neq 0$ which satisfies $\rho(x) = 0$ $\forall x \in Z$.* ◇

3 Representations

A representation of a C^*-algebra $\mathcal{A}$ as operators in a Hilbert space $\mathcal{H}$ is a homeomorphism ϕ of $\mathcal{A}$ in $\mathcal{B}(\mathcal{H})$. If $\mathcal{A}$ has an identity e one must have $\phi(e) = I$.

Every representation is continuous. Indeed if $\mathcal{A}$ has an identity from $a^* a \leq |a|^2 e$ it follows $\phi(a)^* \phi(a) \leq |\phi(a)|^2 I$. If $\mathcal{A}$ has does not have an identity this estimate is used with approximate identities.

A representation ϕ is *faithful* if it is injective. If it is not injective, its *kernel* $\mathcal{N}^\phi \in \mathcal{A}$ is a bilateral ideal. Recall that the Kernel of ϕ is the subalgebra of $\mathcal{A}$ of all element whose image under ϕ is the zero element of $\mathcal{B}(\mathcal{H})$.

On the quotient $\mathcal{A}/\mathcal{N}^\phi$ one can define a norm

$$|\tilde{a}|_0 \equiv inf_{b \in N} |a + b|$$

(we have denoted by $\tilde{a}$ the equivalence class of a). With this norm $\mathcal{A}/\mathcal{N}_\prec$ is a C^*-algebra.

4 The Gel'fand-Neumark-Segal Construction

We prove now that a *-algebra admits a faithful representation as norm-closed subset of $\mathcal{B}(\mathcal{K})$ where the Hilbert space need not be separable [8, 12, 16].

To each state ρ we associate a representation (the GNS representation). This representation may not be faithful.

Let ρ be a state of $\mathcal{A}$. The *kernel* of $\mathcal{A}$ relative to the state ρ, denoted $\mathcal{N}_\rho$, is the closed subalgebra defined by

$$\mathcal{N}_\rho \equiv \{a \in \mathcal{A} \ : \ \rho(a^*a) = 0\}. \tag{14}$$

It is a bilateral ideal. Indeed if $a \in \mathcal{A}_\rho$

$$\rho((ba)^*ba) = \rho(a^*b^*ba) \equiv \rho_a(b^*b) \leq |b^*b|\rho_a(e) = |b^*b|\rho(a^*a) = 0 \tag{15}$$

where we have used the fact that also ρ_a is a positive bilinear functional.

Consider now the sesquilinear form $\langle b, a\rangle \equiv \rho(b^*a)b, a \in \mathcal{A}$. It is well defined on $\mathcal{A}/\mathcal{N}$ and non-degenerate. Indeed denoting by $\tilde{\rho}$ the state defined by ρ on $\mathcal{A}/\mathcal{N}_\rho$ and by $\tilde{a}$ the equivalence class of a, the identity

$$\tilde{\rho}(\tilde{a}\tilde{b}) = 0 \quad \forall b \in \mathcal{A} \tag{16}$$

implies $\rho(ab) = 0 \ \ \forall b \in \mathcal{A}$ and therefore $a \in \mathcal{N}$ and $\tilde{a} = 0$.

We shall denote by $\mathcal{H}_\rho$ the Hilbert space obtained completing $\mathcal{A}/\mathcal{N}_\rho$ in the topology of the scalar product $\langle b, a\rangle$. By construction $\mathcal{A}/\mathcal{N}_\rho$ is dense in $\mathcal{H}_\rho$. Denote by i_ρ the identification of elements of $\mathcal{A}$ with elements of $\mathcal{H}_\rho$.

Definition 10 (*G.N.S. representation*) The G.N.S. representation of $\mathcal{A}$ associated to the state ρ is the homeomorphism ϕ_ρ defined by

$$\mathcal{A} \ni a \Rightarrow \phi_\rho(a) \in B(\mathcal{H}), \qquad \phi_\rho(a)\phi_\rho(b) = \phi_\rho(ab) \tag{17}$$

$\diamond$

By construction $\phi_\rho(a)| \leq |a|$ and therefore $\phi_\rho(a)$ can be extended to a bounded closed operator on $\mathcal{H}_\rho$. Notice that ϕ_ρ is faithful iff $\mathcal{N}_\rho = \emptyset$.

Theorem 3 (Dixmier) *Every C^*-algebra $\mathcal{A}$ admits a faithful representation as algebra of operators on a Hilbert space $\mathcal{K}$ (which may not be separable).* $\diamond$

Proof The states separate $\mathcal{A}$ (if $\rho(a) = 0 \forall \rho$ then $a = 0$). There will be a collection of states $\{\rho_\alpha\}$ (α may be non denumerable) such that $N \equiv \cap_\alpha N_{\rho_\alpha} = \emptyset$. Since $\phi_\rho(a) = 0$ implies $a \in N_\rho$ the representation $\oplus_\alpha \phi_{\rho_\alpha}$ acting on $\mathcal{K} \equiv \oplus \mathcal{H}_{\rho_\alpha}$ is faithful. $\heartsuit$

If the algebra is separable (it is generated by a countable set of elements) the Hilbert space $\mathcal{K}$ may be chosen separable. Remark that if the state is pure, the corresponding representation is irreducible.

If the state ρ is not cyclic the ideal $\mathcal{N}_\rho$ is not empty and one obtains a representation of the subalgebra $\mathcal{A}/\mathcal{N}$.

As an example, let $\mathcal{A}$ be the algebra of continuous functions on [0, 2] and choose ρ as

$$\rho(f) = \int_0^1 f(x)dx \qquad f \in \mathcal{A}. \tag{18}$$

In this case $\mathcal{N}_\rho$ is the collection of continuous functions vanishing in [0, 1] and the bilinear form is

$$\langle f, g\rangle = \int_0^1 \bar{f}(x)g(x)dx \tag{19}$$

The Hilbert space $\mathcal{H}_\rho$ is therefore the space of equivalence classes of square-integrable functions on [0, 1], one has on this functions $\phi_\rho(f) = f$ and the representation is faithful only on the continuous functions that vanish outside [0, 1].

In this example ρ is a normal state (it maps norm-convergent sequences in convergent sequences of numbers).

Consider instead the non-normal state $\rho_1(f) = f(1)$. The representation space is C considered as a Hilbert space. The representation cannot be extended to the von Neumann algebra of bounded measurable functions on [0, 2]. And the representation $\phi_\rho(a)$ is not faithful on the sub-algebra $\mathcal{A}_1$ of continuous functions which vanish in the point $x = 1$. Therefore if one wants to use the non-normal states $\rho_y(f) = f(y),\ 0 < y < 2$ a faithful representation is obtained in $\oplus_{0<y<1} C_y$, a non separable Hilbert space.

This example can be easily generalized to an arbitrary C^* algebra and points out the difference between the representations associated to normal states and those associated to non-normal states: the former can be continued to representation of the von Neumann algebra $(\pi_\rho(\mathcal{A})'')$.

We have noticed that for any C^*-algebra $\mathcal{A}$ the states are a separating set. *This is not true generically if one considers only normal states* (recall that for a normal state ρ one has $\rho(sup a_\alpha) = sup\rho(a_\alpha)$ for any increasing filter $\{a_\alpha\}$).

Definition 11 (*W^*-algebras*) A W^*-algebra is a C^*-algebra that has the following two (equivalent) properties.

(1) Every bounded increasing filter converges to its supremum (therefore in particular every state is normal).
(2) For every $a \in \mathcal{A}$ there exists a *normal state* ρ such that $\rho(a) \neq 0$ (the algebra is separated by *normal states*). ◇

Referring to the abelian example given above of the C^* algebra of the continuous function on [0, 1] notice that the algebra of L^∞ of the essentially bounded function in [0, 1] is a W^*-algebra. The (normal) states are the L^1 function on [0, 1].

5 von Neumann Algebras

A class of representation of W^*-algebras which is particularly useful in Quantum Mechanics are the *von Neumann algebras* [15]. We need one more definition.

Definition 12 (*commutant*) For a representation π of a C^*-algebra $\mathcal{A}$ in $B(\mathcal{H})$ denote by $\mathcal{A}'_\pi$ the commutant of $\pi(\mathcal{A})$, i.e. the subset of $B(\mathcal{H})$ of all bounded operators which commute with all elements of $\pi(\mathcal{A})$. ◇

It is easy to verify that $\mathcal{A}'_\pi$ is a C^*-algebra closed in the weak operator topology. Indeed if $b_n\ a = a\ b_n$ for every n and b_n converges to b in the weak operator topology one has for every $\phi \in H$

$$(b_n^*\psi, a\psi) = \psi, b_n a\psi = (a^*\psi, b_n\psi) \tag{19}$$

Taking the limit $n \to \infty$ one has $(b^*\psi, a\psi) = (a^*\psi, b\psi)$ and this implies $ab = ba$.

The kernels of the representations of $\mathcal{A}$ admit a partial ordering by inclusion. The same is true for the relative commutants $\mathcal{A}'_\pi$. The *commutant* $\mathcal{A}'$ of $\mathcal{A}$ is the union of the relative commutants over all representations.

Definition 13 (*maximal abelian*) If $\mathcal{A}$ is abelian one has always $\pi(\mathcal{A}) \subset \pi(\mathcal{A})'$; the algebra is called *maximal abelian* if $\pi(\mathcal{A}) = (\pi(\mathcal{A}))'$ holds for every representation π. π.

Definition 14 (*von Neumann algebras*) An algebra of operators on a Hilbert space $\mathcal{H}$ is called (a concrete) *von Neumann algebra* if it is weakly closed, closed under conjugation and contains the identity. Every von Neumann is a C^*-algebra with the natural definition of norm and adjoint. We shall reserve the symbol $\mathcal{M}$ for a von Neumann algebra. ♣

The relation between W^*-algebras (which are defined abstractly) and von Neumann algebras (which are operator algebras) is given by the following theorem.

Theorem 4 (Sakai) [14] *Every representation of a W^*-algebra is a von Neumann algebra. In particular every W^*-algebra has a faithful representation as von Neumann algebra.* ◇

Notice that von Neumann algebras of operators are characterized by having nice regularity properties with respect to filtrations [15].

Theorem 5 (Kadison) *A C^*-algebra contained in $\mathcal{B}(\mathcal{H})$ is a von Neumann algebra iff every increasing filter of* positive elements *has an extremal element.* ◇

For the proof one can see [11].

A typical example of a C^*-algebra of operators which is not a von Neumann algebra is the algebra generated by the identity operator I and the compact operators in a Hilbert space $\mathcal{H}$. It is not a von Neumann algebra because it is closed in norm but not weakly closed. Its weak closure is the von Neumann algebra $\mathcal{B}(\mathcal{H})$.

Definition 15 (*Centre*) The *centre* of a von Neumann algebra $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is the intersection of $\mathcal{M}$ with its commutant $\mathcal{M}'$ (the collection of operators in $\mathcal{B}(\mathcal{H})$ which commute with all elements of $\mathcal{M}$). ♣

6 von Neumann Density and Double Commutant Theorems. Factors, Weights

The following theorems have an important role in the theory of von Neumann algebras.

Theorem 6 (von Neumann density theorem) *If $\mathcal{A}$ is a *-closed subalgebra of $\mathcal{B}(\mathcal{H})$ which contains the unit, then $\mathcal{A}$ is strongly dense in $\mathcal{A}''$.* ◇

Proof We need to show that if $a \in \mathcal{A}$ and Ω is an open neighborhood of a in the strong operator topology, then $\mathcal{A} \cap \Omega \neq \emptyset$.

Since we have assumed that $\mathcal{H}$ is separable, we may assume that

$$\Omega \equiv \{b \in \mathcal{B}(\mathcal{H}),\ |(a-b)\xi_n| \leq \epsilon\} \qquad n = 1 \ldots N \tag{20}$$

for some integer N. We give the proof of the theorem for $N = 1$. The proof of the general case follows by induction.

Define $S \equiv \{a\xi,\ \in \mathcal{A}\}$ and let P be the orthogonal projection onto the subspace $\bar{S}$. Since $\bar{S}$ is invariant under $\mathcal{A}$ it follows that P commutes with $\mathcal{A}$ and therefore leaves $\bar{S}$ invariant. Since $\xi \in \bar{S}$ it follows that $a\xi \in \bar{S}$. In particular there exists $a \in \mathcal{A}$ such that $|a\xi| < \epsilon$ as desired. ♡

From Theorem 6 one derives

Theorem 7 (von Neumann double commutant theorem) *The following conditions on a self-adjoint subalgebra $\mathcal{M}$ of $\mathcal{B}(\mathcal{H})$ are equivalent*

(1) $\mathcal{M} = \mathcal{M}''$
(2) *$\mathcal{M}$ is closed in the strong topology*
(3) *$\mathcal{M}$ is closed in the weak topology.* ◇

Notice that the inclusion $\mathcal{M} \subset \mathcal{M}''$ follows trivially because commuting is a reflexive property. The identity is a consequence of Theorem 6.

One can define an *abstract von Neumann algebra* $\mathcal{M}$ as a C^*-algebra $\mathcal{A}$ which admits a *predual* $\mathcal{A}_*$ in the sense that it isometrically isomorphic to the Banach dual space $(\mathcal{M}_*)^*$.

It is easy to verify that *concrete* von Neumann algebras, i.e. subalgebras of $\mathcal{B}(\mathcal{H})$, are indeed abstract von Neumann algebras.

Definition 16 (*factor*) If the center of a von Neumann algebra is trivial, i.e. it is composed only of multiples of the identity in $\mathcal{B}(\mathcal{H})$, the algebra $\mathcal{M}$ is called *a factor*. We shall indicate factors with the symbol $\mathcal{F}$.

Since a von Neumann algebra is closed in the weak topology it contains projection operators and in fact *it is generated by its projections*.

Definition 17 (*weight*) On the projections of $\mathcal{F}$ one can define a positive *weight function* which is additive on the classes of projections which are orthogonal to each other.

The factor $\mathcal{F}$ is said to be of type:

(1) I_N if the weight takes value in the integers form 1 to $N < \infty$.
(2) I if the weight takes value in the positive integers $N \leq +\infty$.
(3) II_1 if the weight takes value in [0, 1].
(4) II_∞ if the weight takes value R^+.
(5) III if the weight of every projector is $+\infty$.

Definition 18 (*trace*) A weight is called a *trace* (and denoted by the symbol Tr Trace) if it can be extended as an affine map on the positive elements which satisfies $Tr(ab) = Tr(ba), \quad \forall a,\ b\ \in \mathcal{F}$. The map Tr has range in $R^+ \cup +\infty$.

Factors of type $I_N,\ II_1\ II_\infty$ have a *tracial state*.

Type I factors are isomorphic to $B(\mathcal{H})$ where $\mathcal{H}$ is separable; in this case the weight of a projection operator P coincides with the dimension of the space $P\mathcal{H}$ and $Tr(A)$ is the usual trace defined in for operators in a Hilbert space.

For a factor $\mathcal{F}$ which is not of type I any projection $P \in \mathcal{F}$ projects on a space of infinite dimension and the operation Tr is not the Hilbert space trace.

An example of type II_1 factor is the inductive limit of $2^n \times 2^n$ matrices which we will discuss in the following Lecture. In the Nth step of the induction the operator $Tr(A)$ is defined as $\frac{1}{N} TraceA$ so that $Tr(I) = 1 \forall N$. As a results in the limit trP for a projection operator $P \in \mathcal{F}$ can have any value in the interval (0, 1].

One does not encounter factors of type II and III in the usual treatment of non-relativistc Quantum Mechanics for systems with a finite number of degrees of freedom. The have instead an important role in Quantum Statistical Mechanics and in Relativistic Quantum Field Theory [3].

The algebras introduced in Quantum Statistical Mechanics are of type II and are inductive limits of algebras (often finite-dimensional) associated to subsets of an infinite-dimensional lattice with the property that operators associated to disjoint subsets commute.

The algebras introduced in Relativistic Field Theory are *local algebras* associated to finite *open* space-time domains Ω with the property that subalgebras commute if they are associated to open space-time domains that are located at space-like distance from each other. They have the property that if Ω_1 and Ω_2 are open and relatively space-like and their closure are not disjoint then the algebra associated to $\Omega_1 \cup \Omega_2$ is strictly larger than the algebra generated by $\mathcal{A}_{\Omega_1}$ and $\mathcal{A}_{\Omega_2}$. This algebras are of type III.

Factors of type I are generated by their finite dimensional projections. As a consequence their normal states have an important property

Lemma 8 [5] *The limit points in the weak-* topology of a net of normal states of a type I factor are normal states. This property* characterizes *type I factors.* ◇

A result which holds for every von Neumann algebra is the following. We establish first notations.

Definition 19 (*normal maps*) Let $\mathcal{M}$ and $\mathcal{N}$ be von Neumann algebras. A map $\mathcal{M} \to_\beta \mathcal{N}$ which sends $\mathcal{M}^+$ to $\mathcal{N}^+$ is said to be *normal* if it maps a filter $\{x_\alpha\}$ with upper limit x to a filter $\beta(x_\alpha)$ that has $\beta(x)$ as upper limit. ◇

Notice that from the definition we have given of *normal state* it follows that normal states are normal maps $\mathcal{N} \equiv C$.

We state without proof the following theorem; its proof can be found in the references indicated.

Theorem 9 [11]

(1) Every isomorphism of von Neumann algebras is normal.
(2) For every projector $p \in \mathcal{M} \cap \mathcal{M}'$ the von Neumann algebras $\mathcal{M} \in \mathcal{B}(\mathcal{H})$ and $p\mathcal{M}$ are isomorphic. ◇

7 Density Theorems, Spectral Projection, Essential Support

Important density theorems hold for the C^*-algebras; together with Theorem 6 they often allow to determine the regularity of a map by restricting it to a smaller set on which it has a simpler form.

Theorem 10 (Kaplanski) [10] *Let $\mathcal{M}$ be a von Neumann algebra which is the weak closure of the representation π of a C^*-algebra $\mathcal{A}$ in $\mathcal{B}(\mathcal{H})$. Then*

(a) The unit ball of $\pi(\mathcal{A})$ is dense on the unit ball of $\mathcal{M}$.
(b) The self-adjoint part and the positive part of $\pi(\mathcal{A})$ are dense in the corresponding parts of $\mathcal{M}$.
(c) If $\mathcal{A}$ has a unit, then the unitary operators in $\pi(\mathcal{A})$ are dense in the unitary operators of $\mathcal{M}$. ◇

Density theorems for generic C^*-algebras have a correspondence with classical theorems for the commutative case, i.e. for C^*-algebras of functions on locally compact space.

If the space X is compact and has the Hausdorff property the collection of continuous functions is a C^*-algebra which has as weak closure the collection of L^∞ functions. In this context a classic density result is the following.

Theorem 11 (Lusin) *If X is locally compact and Hausdorff, and μ is a Radon measure, for any $f \in L^\infty$ and any $\epsilon > 0$ there exists a Borel set $Y \subset X$, with $\mu(X/Y) < \epsilon$, and a function $g \in C_0(X)$ such that $g \equiv f$ in Y.* ◇

The corresponding result in the non-commutative case is the following [11].

Theorem 12 (Powers) *Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a (concrete) C^*-algebra and set $\mathcal{A}'' \equiv \mathcal{M}$. Then for any $x \in \mathcal{M}$, any projection operator $p_0 \in \mathcal{M}$, any choice of $n \in Z$ and of $\{\xi_1, \ldots, \xi_n\} \in \mathcal{H}$ it is possible to choose a projection operator $p \in \mathcal{M}$ (with $p < p_0$ and $|(p - p_0)\xi_k| < \epsilon, \quad k = 1, \ldots, n$) and an element $y \in \mathcal{A}$ in such a way that*

$$xp = py \qquad |y| \leq |xp| + \epsilon \tag{21}$$

Moreover if $x \in \mathcal{M}_{s.a.}$ one can choose $y \in \mathcal{M}_{s.a.}$ but then and can prove only

$$|y| \leq min\ \{2|x||p_0|, \quad |x|\} + 2\epsilon \tag{22}$$

(*we have denoted by $\mathcal{M}_{s.a.}$ the set of self-adjoint elements of $\mathcal{M}$*). ◇

If the G.N.S state is a standard trace on $\mathcal{B}(\mathcal{H})$ the ideal considered in the G.N.S construction is equal to $\mathcal{S}_1(\mathcal{H})$ i.e. trace class operators on $\mathcal{H}$ with trace one.

In case $\mathcal{M}$ admits a faithful normal semi-definite trace τ a proper generalization of these properties can be obtained by choosing a representation $(\mathcal{H}, \pi)$ which is semi-standard, i.e. there exists a conjugation operator J such that

$$JaJ = a^* \qquad \forall a \in \mathcal{M} \tag{23}$$

The existence of J follows from the fact that the operation * from $\mathcal{M}$ to $\mathcal{M}$ is isometric for the inner product defined by the trace. However if $\mathcal{M}$ does not admit such a trace such construction fails. An operator T intertwining $\mathcal{M}$ with $\mathcal{M}'$ can still be constructed but the construction is more elaborate (Tomita-Takesaki construction) [18]. We shall describe it in the second part of these Lectures.

In the case of finite (type II) C^*-algebras, the G.N.S. representation associated to a state ρ which is a trace has the structure of a Hilbert-Schmidt space $\mathcal{S}_2 \equiv \mathcal{H}_{H.S.}$ (the space of Hilbert-Schmidt operators on a Hilbert space $\mathcal{H}$; this space is called occasionally *superspace*). This space can be considered as a non-commutative analogue of the L^2 space.

In this case every $\omega \in C^*_+$ (positive functional normalized by $\omega(I) = 1$) corresponds *uniquely* to a vector ξ_ω in the cone of positive elements of $\mathcal{S}_2$

$$\xi_\omega = \rho_\omega^{\frac{1}{2}} \in \mathcal{S}_2 \leftrightarrow \rho_\omega \in \mathcal{S}_1 \tag{24}$$

where $\mathcal{S}_1$ is the space of trace-class (nuclear) operators.

In this correspondence one has for $\omega \in A^*$

$$\omega(a) = Tr\rho_\omega\pi_\omega(a) \tag{25}$$

This construction, in which the algebra can be seen dual to a vector space, can be extended to the case of W^* algebras and permits to construct analogues of the Hilbert-Schmidt space for an arbitrary W^* algebra $\mathcal{N}$. We shall use the symbol $\mathcal{N}$ (normal) to denote an arbitrary W^*-algebra; recall that W^* algebras are characterized by being

dual of a Banach space and therefore they satisfy $W = (W^*)^*$. Of course every finite dimensonal C^*-algebra is a von Neumann algebra,

A notion that plays an important role in applications is the *essential support*.

Definition 20 (*spectral projection, essential support*) If $a \in \mathcal{A} \subset \mathcal{B}(\mathcal{H})$ its *spectral projection* denoted with $[a]$ is the projection on the closed subspace generated by $\{a\phi, \quad \phi \in \mathcal{H}\}$. The subspace $[a]\mathcal{H}$ is called the essential support of the element a.

If a is a projection, $a = \pi$, one has $[\pi] = \pi$ and the essential support coincides with the support. ◇

Theorem 13 *Let $\mathcal{A}$ a C^*-algebra contained in $B(\mathcal{H})$. Then*

(1) $[a] \in \mathcal{A}''$
(2) $[a] = [a^*a]$. ◇

Proof Relation (2) follows from the definition. Therefore it is sufficient to consider the case $a > 0$. In this case one can verify that

$$[a] = w - lim_{n\to\infty}[\xi_n] \qquad \xi_n \equiv (\frac{1}{n} + a)^{-1}\,[a] \tag{26}$$

is a projection operator. If ϕ is such that $(\phi, a\psi) = 0, \quad \forall\psi$ one has $(\xi_n\phi, \psi) = 0 \ \ \forall n$ and therefore $P_a\phi = 0$. From this point (1) follows. ♡

Theorem 13 has the following corollary [13].

Theorem 14 (polar decomposition in $\mathcal{M}$) *For every $a \in \mathcal{M}$ there exists unique $U \in \mathcal{M}$ with $U^*\,U = 1, \quad U^*\,U = [a]$ such that $a = U\sqrt{(a^*a)}$.* ◇

Sketch of the Proof For the proof of Theorem 14 one starts with a polar decomposition of a with a partial isometry V such that $a = V\sqrt{(a^*a)}$. One notices then that V is defined modulo the multiplication by a unitary element $W \in \mathcal{M}'$. This permits to show that it is possible to choose W in such a way that $U^*\,U \in \mathcal{M}$ with $U = WV$. ♡

8 Automorphisms of a C^*-algebra. C^*-dynamical Systems

In the formalism of C^*-algebras dynamics is described by a one parameter group of automorphisms. We shall denote by $Aut(\mathcal{A})$ the group of *-automorphisms (i.e. $(\alpha(a))^* = \alpha(a^*)$).

The group $Aut(\mathcal{A})$ is a topological group endowed with different topologies. We shall use two of them.

(a) The *uniform topology* defined by the norm

$$|\alpha| = sup_{a\in\mathcal{A},|a|\leq 1}|\alpha(a)|.$$

(b) The *strong topology* defined by the complete set of neighborhoods

$$N(\alpha_1; a_1, \ldots, a_n, ; \epsilon) \equiv \{\alpha \in Aut(\mathcal{A}), \quad |\alpha(a_j) - \alpha_1(a_j)| < \epsilon, \quad j = 1, \ldots, n\} \tag{27}$$

Let G be a topological group and let $G \ni g \to \alpha(g)$ be a homeomorphism of G in $Aut(\mathcal{A})$.

Definition 21 (C^*- *dynamical system*) [12] The triple $\{\mathcal{A}, \ G, \ \alpha\}$ is called C^*-*dynamical system* if $g \to \alpha(g)$ is continuous as a function of g in the strong topology; it is called uniform dynamical system if $g \to \alpha(g)$ is continuous in the uniform topology. ◇

A relevant case is when the C^*-algebra considered is a W^*-algebra $\mathcal{W}$ or a von Neumann algebra $\mathcal{M}$ and therefore the dual of a Banch space $\mathcal{B} \equiv \mathcal{W}_*$. In this case it is interesting to consider on $\mathcal{W}$ the weak* topology. Setting

$$\alpha^* b(a) = b(\alpha(a)), \quad a \in \mathcal{W} \quad b \in \mathcal{W}^* \tag{28}$$

by duality one defines a group structure on $\mathcal{W}_*$ and a topology on $Aut(\mathcal{W})$ which we shall call weak-* topology.

Every automorphism of $\mathcal{W}$ is continuous in the weak-* topology and it can be proved that $g \to \alpha(g)$ extends to a weak-* continuous map of G in $Aut(\mathcal{W})$. The triple $\{\mathcal{W}, G, \alpha\}$ is called a W^*-*dynamical system*.

We have discussed in "Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories" the theorem of Wigner and Kadison which implies that if $\mathcal{A} = \mathcal{B}(\mathcal{H})$, and the one-parater group of automorphisms G is weakly measurable, then the automorphisms are implemented by of is a one-parater group of unitary operators. This theorem *does not hold* if $\mathcal{A}$ is a proper subalgebra of $\mathcal{B}(\mathcal{H})$.

If the dynamical systems is *norm-continuous* one has a theorem of Dixmier that we quote here without proof.

Theorem 15 (Dixmier) [7] *Let G be a semisimple fnite-dimensional Lie group and let $G \ni g \to \alpha_g$ a norm continuous representation of G in the group of automorphisms of $\mathcal{W}$. Then there exists a norm-continuous representation $g \to U(g)$ in the group of unitary operators in $\mathcal{W}$ such that*

$$\alpha_g(a) = U(g) a U^*(g), \qquad a \in \mathcal{W}, \quad g \in G \tag{29}$$

We shall say that the group $U(g)$ implements *the dynamical system.* ◇

Definition 22 (*inner dynamical systems*) Given a von Neumann algebra $\mathcal{M}$ on the Hilbert space $\mathcal{H}$, and a locally compact Lie group G, a strongly continuous dynamical system $\{\mathcal{M}, \ G, \ \alpha_g\}$ is called *weakly inner* if the automorphisms are implemented by a group of unitary operators. It is called *inner* if one can choose in $\mathcal{M}$ the implementing unitary operators. ◇

In general the group of automorphisms of $\mathcal{W}$ is not spacial i.e. it cannot be induced by a group of unitary operators. In the next lecture we will see that a group of automorphims is spacial if there is in $\mathcal{H}$ a vector which is cyclic both for $\mathcal{M}$ and for its commutant $\mathcal{M}'$ (Tomita-Takesaki theorem) [18].

The condition that the vector be cyclic for the commutant can be substituted with the condition that the vector corresponds to a state invariant under the dual action of the group on the states [6].

Theorem 16 (Diximier) *The representation of the group of automorphisms is spacial if there exists in $\mathcal{H}$ a cyclic vector ρ_0 that corresponds to an invariant state ϕ_0.*

$$\alpha_g^* \rho_0(a) \equiv \rho_0(\alpha_g(a)) = (\alpha_g(a)\phi_0, \phi_0) = (a\phi_0, \phi_0) = \rho_0(a) \quad \forall a \in \mathcal{M}. \tag{30}$$

$\diamond$

Proof Define

$$U(g)a\phi_0 \equiv \alpha_g(a)\phi_0, \qquad g \in G \tag{31}$$

This definition is well posed and $U(g)$ is an isometry on the closure of $\{\mathcal{M}\phi_0\}$. Indeed one has

$$|U_g a\phi_0|^2 = (\alpha_g(a^*)\alpha_g(a)\phi_0, \phi_0) = (a^* a\phi_0, \phi_0). \tag{32}$$

Since ρ_0 is cyclic, this isometry extends to a unitary operator that we denote by $U(g)$. The group property follows from the definition and if $g_n \to e$ (the unit of the group) $\alpha_{g_n}(a)\phi_0$ converges weakly to $a\phi_0$. Therefore $U(g) \to I$ weakly and also strongly since $U(g)$ are unitaries. One has

$$U(g)aU^*(g)b\phi_0 = U(g)a\alpha_g^{-1}(b)\phi_0 = \alpha_g(a)b\phi_0. \tag{33}$$

and, making use of the cyclicity of ϕ_0

$$U(g)aU^*(g) = \alpha_g(a) \qquad a \in \mathcal{M}, \quad g \in G \tag{34}$$

$\heartsuit$

In general, even in case there exists a representation of G in the Hilbert space $\mathcal{H}$ by means of unitary operators $U(g)$, *it may not be true* that one can choose $U(g) \in \mathcal{M}$ (except for the case $\mathcal{M} \equiv \mathcal{B}(\mathcal{H}) \otimes M_N$).

Recall (Wigner theorem) that if $\mathcal{M} = \mathcal{B}(\mathcal{H})$ and $G = R$ such a weakly continuous unitary group always exists. The extension of this result to general von Neumann algebras has been proven by H. Araki.

Theorem 17 [1] *Let $t \to \alpha_t$, $t \in R$ be a weakly continuous group of automorphisms of a von Neumann algebra $\mathcal{M}$ on a Hilbert space $\mathcal{H}$, with cyclic vector ϕ_0.*

If the generator H of the corresponding group of unitary operators is strictly positive *then $U(t) \equiv e^{itH}$ belongs to $\mathcal{M}$.* ◇

Proof From $U_t \mathcal{M} U_t^* = \mathcal{M}$ follows $U_t \mathcal{M}' U_t^* = \mathcal{M}'$.

Since the action of the group $U(t)$ leaves invariant $\mathcal{M}$ it leaves also invariant $\mathcal{M}'$ and therefore for any $b \in \mathcal{M}'$ also $U_t b u_t^* \in \mathcal{M}'$.

Let $x \in \mathcal{M}'$, a, $b \in \mathcal{M}$ and consider the function

$$f(t) \equiv (U_t x U_t^* a\phi_0, b\phi_0) = (ae^{itH} x\phi_0, b\phi_0) \tag{35}$$

(we have taken into account that $U_t^* \phi_0 = \phi_0$). Since by assumption $H \geq 0$ setting

$$f(t + is) = (ae^{(it-s)H} x\phi_0, b\phi_0) \tag{36}$$

the function f in (34) can be extended to a bounded function analytic in the upper half-plane. On the other hand $f(t)$ can also been written

$$f(t) = (a\phi_0, bU_t x^* U_t^* \phi_0) = (a\phi_0, x^* e^{-itH} b\phi_0) \tag{37}$$

and therefore f can also be extended as analytic function to the lower half-plane. We conclude that the extension is a bounded entire function and therefore is a constant.

One has then $(U_t x U_t^* a\phi_0, b\phi_0) = (xa\phi_0, b\phi_0)$ and, since ϕ_0 is cyclic $U_t \in \mathcal{M}'' = \mathcal{M}$. ♡

The proof of Araki's theorem can be extended to the case in which $H \geq 0$ and there does not exists a invariant cyclic vector. The proof becomes substantially more difficult and will not be given here. It can be found e.g. in [6, 17].

Let $\mathcal{A}$ be a separable von Neumann algebra, $\mathcal{G}$ a group of *-automorphisms of $\mathcal{A}$ and denote with $a \to g(a)$. Let $\mathcal{B}$ the fixed point algebra; it is easily seen that it is a von Neumann algebra.

Theorem 18 [12] *Assume that for each a the weak closure of the convex set $conv(\mathcal{G}; \mathcal{A})$(generated by the action of $\mathcal{G}$ on $\mathcal{A}$) has not empty intersection with the von Neumann algebra $\mathcal{B}$. Then the conditional expectation Φ from $\mathcal{A}$ to $\mathcal{B}$ exists. If the intersection consists of one point, then $\Phi(a) = g(a)$, $\forall g \in \mathcal{G}$.* ◇

This theorem has an important corollary.

Corollary *If $\mathcal{C}$ is the center of the von Neumann algebra $\mathcal{M}$ and $\mathcal{G}$ is the group of* inner automorphisms *then the intersection of $\mathcal{C}$ with $conv(\mathcal{G}(\mathcal{M}))$ is not empty. In this case the conditional expectation Φ is called* central trace. *It is a* tracial state *if $\mathcal{M}$ is a factor.* ◇

9 Non-commutative Radon-Nikodim Derivative

We return now to the G.N.S. construction. In the commutative case the algebra $\mathcal{A}$ is an algebra of functions and the states are measures. In that case one has a natural definition of Radon-Nykodym derivative [12]. We want now to see to which extent this notion can be extended to the non-commutative case.

We have seen that the GNS construction provides a bijection between the states of a C^*-algebra $\mathcal{A}$ and its representations. The following theorem of Schur is important. Recall that if $\mathcal{M}$ is a subset of $\mathcal{B}(\mathcal{H})$ its *commutant* $\mathcal{M}'$ consists of the operators $B \in \mathcal{B}(\mathcal{H})$ which commute with all elements in $\mathcal{M}$.

Theorem 19 (Schur) *Let $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be a representation. The following are equivalent*

(1) π is irreducible.
(2) The commutant $\pi(\mathcal{A})'$ is one-dimensional. ◇

Proof If π is not irreducible one has $\pi = \pi_1 \oplus \pi_2$. Then $\pi_k \pi(a) = \pi(a)\pi_k$ for $k = 1, 2$. Hence $\pi(\mathcal{A})'$ has more than one dimension.

Conversely, suppose that $\pi(\mathcal{A})'$ has more that one dimension. Let $X \in \pi(\mathcal{A})'$ and consider $X + X^*$. One has $X + X^* \neq cI$ since by assumption $\pi(\mathcal{A})'$ has more than one dimension. Therefore $X + X^*$ has non trivial spectral projections $P(E)$ and $I - P(E)$. Since both of them commute with $\pi(\mathcal{A})$ it follows that π is not irreducible. ♡

When specialized to the GNS representations associated to given state σ this is equivalent to saying that for all positive functionals the condition $\tau \leq \sigma$ is equivalent to $\tau = \lambda\sigma$ for some $\lambda \geq 0$ [15, 17].

Recall that for the GNS representation constructed from a state σ the state σ itself is represented by a vector Ω_σ in a Hilbert space and therefore it is *normal* for this representation. Given a C^*-algebra the states $S(\mathcal{A})$ form a compact convex subset of the unit ball in the dual $\mathcal{A}^*$.

Let $\mathcal{A}_+$ the set of positive elements in $\mathcal{A}$. Given $\sigma \in S(\mathcal{A})$ let τ be a linear positive functional on $\mathcal{A}$. By $\tau < \sigma$ we mean $\tau(a) < \sigma(a)$ for all $a \in \mathcal{A}$.

We look for a relation between τ and the commutant of $\pi(\mathcal{A})$.

Lemma 20 (Schur-Sakai-Nykodym) *Let τ be a positive functional, and let σ be a state. Let π_σ the G.N.S representation associated with σ. There is a bijection between τ satisfying $0 \leq \tau \leq \sigma$ and a posititve self-adjoint operators B in the commutant of $\pi_\sigma(\mathcal{B})$ with $0 \leq B \leq I$. This relation is given by*

$$\tau(a) = (\Omega, \pi_\sigma(a) B \Omega) \tag{37}$$

where Ω is the vector associated to σ by the G.N.S construction. ◇

We remark that this can be regarded as an extension to the non commutative setting of the classical Radon Nykodim derivative. We may write $A = \frac{d\tau}{d\sigma}$.

Proof

(1) Suppose $B \in (\pi_\sigma(\mathcal{A})'$ and $0 \leq B \leq I$. We need to show that $\tau(a) = (\Omega, \pi(a)B\Omega) \leq \sigma(a)$ for all $a \in \mathcal{A},\ a \geq 0$. Let $a = b^2$. Since $B \in \pi_\sigma(\mathcal{A})'$ also $\sqrt{B}$ belongs to this set. Therefore

$$\begin{aligned}\tau(a) = (\Omega, \pi_\sigma(a)\Omega) = (\Omega, (\pi_\sigma b)^* \pi_\sigma(b) B\Omega) = (\pi_\sigma(b)\Omega, B\pi_\sigma(b)\Omega)\\ \leq (\pi_\sigma(b), \pi_\sigma(b)\Omega) = (\Omega, \pi_\sigma(a)\Omega) < \sigma(a)\end{aligned} \tag{38}$$

(2) Conversely suppose $\tau \leq \sigma$. Then for all $a \geq 0$ one has

$$\tau(a) \leq \sigma(a) = (\Omega, \pi_\sigma(a)\Omega)0 = \|\pi_\sigma b\|^2 \tag{39}$$

By Riesz's theorem there is a unique η such that

$$\tau(a) = (\pi_\sigma(b)\Omega, \eta)$$

Let $a = b^2$. Then

$$\tau(b^2) \leq \sigma = \sigma(b^2) = \|\pi_\sigma(b)\Omega\|^2 \tag{40}$$

i.e. $\pi_\sigma(b)\Omega \to \tau(b^2)$ is a bounded quadratic form. Therefore there exosts a unique $B \geq 0$ such that

$$\tau(b^2) = (\pi_\sigma(b)\Omega, B\pi_\sigma(b)\Omega) \tag{41}$$

It is easy to see that $0 \leq B \leq I$ and that B belongs to the commutnat of π_σ $(\mathcal{A})$. ♡

As a corollary of Lemma 20 we prove now

Lemma 21 *Let σ be a state, and $\{\Omega, \pi_\sigma, \mathcal{H}\}$ the corresponding G.N.S. construction. The following are equivalent*

(1) For all positive linear functional τ with $\tau \leq \sigma$ one has $\tau = \lambda\sigma$ with $0 \leq \lambda \leq 1$.
(2) π_σ is irreducible. ◇

Proof Notice that by Lemma 20 $\tau \leq \sigma$ iff there is a self-adjoint operator $B \in \mathcal{A}'$ so that $\tau(a) = (\Omega, \pi_\sigma(a)\Omega)$ for all $a \in \mathcal{A}$. Therefore $\tau = \lambda\sigma$ iff $A = \lambda I$.

$1 \to 2$. Suppose that $\tau \leq \sigma$ implies $\tau = \lambda\sigma$ for some $\lambda > 0$. Then π_σ must be irreducible since otherwise there exists $B \in \mathcal{A}'$ with $B \neq cI$ and $\tau(a) = (\Omega, \pi_\sigma(a)B\Omega),\ a \in \mathcal{A}$ defines a linear positive functional. One has $\tau \neq \lambda\sigma$ a contradiction

$2 \to 1$. By Schur's theorem $(\pi_\sigma\mathcal{A})'$ is one-dimensional, i.e. for all $B \in \mathcal{A}'$ one has $B = \lambda I$ for some $\lambda \in C$. Therefore if $\tau \leq \sigma$ by Sakai's theorem one ha $\tau(a) = (\Omega, \pi_\sigma B\Omega)\ a \in \mathcal{A}$. It follows $\tau = \lambda\sigma$ for some $\lambda \geq 0$. ♡

Lemma 22 *The following are equivalent:*

(1) $\tau \leq \sigma$ *implies* $\tau = \lambda\sigma$ *for some* $\lambda \geq 0$.
(2) π_σ *is irreducible.*
(3) σ *is a pure state.* ◇

Proof We have already proved (2) → (1). We will show (1) → (3) and (3) → (2).
(1) → (3). Suppose $\tau \leq \sigma$ implies $\tau = \lambda\sigma$ for some $\lambda \geq 0$. If σ is not pure then $\sigma = c\sigma_1 + (1-c)\sigma_2$ for $c \in (0,1)$. By assumption $\sigma_1 \leq \sigma$ implies $\sigma_1 = \lambda\sigma$. It follows $\sigma_1 = \sigma_2 = \sigma$.
(3) → (2). Suppose π is not irreducible. Then there is a non trivial projector P in $\pi(\mathcal{A})'$. Let $\Omega_1 = P\Omega$, $\Omega_2 = (I-P)\Omega$ then

$$\sigma(a) = (\Omega_1 \oplus \Omega_2, \pi(a)(\Omega_1 \oplus \Omega_2) = (\Omega_1, \pi(a)\Omega_1) + (\Omega_2, a\Omega_2)$$

$$= \|\Omega_1\|^2 (\frac{\Omega_1}{\|\Omega_1\|}, \pi(a)\frac{\Omega_1}{\|\Omega_1\|}) + (1-\|\Omega_1\|^2)\frac{\Omega_2}{\|\Omega_2\|}\pi(a)\frac{\Omega_1}{\|\Omega_1\|} = \lambda\sigma_1(a) + (1-\lambda)\sigma_2(a). \tag{42}$$

Hence σ is not a pure state. ♡

References

1. Araki, H. (1964). *Progress of Theoretical Physics*, *32*, 956–965.
2. Arveson, A. (1976). *Invitation to C^*-algebras*. New York: Springer.
3. Bratteli, O. (1986). *Derivations, dissipations and group action on C^*-algebras LNM*. New York: Springer.
4. Christensen, A. (1979). *Journal of the London Mathematical Society*, *20*, 358–376.
5. Dell'Antonio, G. F. (1967). *Communications on Pure and Applied Mathematics*, *20*, 413–429.
6. Diximier, J. (1981). *von Neumann Algebras*. Amsterdam: North-Holland.
7. Dixmier C^*-algebres Gallimard Paris.
8. Gel'fand, I., & Neimark, M. (1943). *Matematicheskii Sbornik*, *12*, 197–217.
9. Kadison, J. R. (1951). *Annals of Mathematics*, *54*, 325–338.
10. Kadison, R. (1963). *Annals of Mathematics*, *83*, 273–298.
11. Kadison, R., & Ringrose, J. *Introduction to C^*-algebras* (vol I, II)
12. Pedersen, G. K. (Ed.). (1979). *C^*-algebras and their automorphism groups*. St Louis: Academic Press.
13. Ringrose, J. (1972). *London Mathematical Society*, *5*, 432–450.
14. Sakai, H. (1960). *Tohoku Mathematical Journal*, *12*, 31–59.
15. Saunders, V. S. (1986). *Invitationto von neumann algebras*. New York: Springer.
16. Segal, I. (1947). *Bulletin of the American Mathematical Society*, *53*, 73–88.
17. Störmer, E. (1965). *Transactions of the AMS*, *120*, 438–447.
18. Takesaki, M. (2002, 2003). *Theory of operator Algebras I, II, III Encyclopaedia of Mathematical Sciences* (Vols. 124, 125, 127). Berlin: Springer.

Lecture 10: Derivations and Generators. K.M.S. Condition. Elements of Modular Structure. Standard Form

1 Derivations

Under suitable assumptions a group of automorphisms of a C^*-algebra has an infinitesimal generator which is a *-*derivation* i.e. a linear operation which commutes with taking the adjoint and satisfies Leibnitz's for the derivation of a product. Indeed one has

$$(\alpha_t(ab) - a\ b) \equiv (\alpha_t(a) - a)\ b + \alpha_t(a)(\alpha_t(b) - b). \tag{1}$$

Dividing by t and defining, whenever this strong limit exists

$$\delta(a) = lim_{t\to 0} t^{-1}(\alpha_t(a) - a) \tag{2}$$

one has

$$(\delta(a))^* = \delta(a^*), \quad \delta(a\ b) = \delta(a)\ b + a\ \delta(b). \tag{3}$$

In case $\mathcal{A} \equiv \mathcal{B}(\mathcal{H})$ if $\alpha_t(a) = U(t)aU^*(t)$

$$\delta(a) = i[h, a], \quad U(t) = e^{iht} \tag{4}$$

with domain $D(\delta)$ the collection of operators $a \in B(\mathcal{H})$ such that $[h, a]$ is densely defined. If h is bounded, $\alpha_t(a)$ is norm continuous in t and $|\delta(a)| \leq 2|h||a|$.

The converse also holds.

Theorem 1 (Kadison, Sakai, Kaplanski) [7, 8] *Let $\mathcal{A}$ be a C^*-algebra and let δ be a derivation which is closed in the norm topology.*

Then in every representation π of $\mathcal{A}$ in the Hilbert space $\mathcal{H}$ the derivation δ extends to the weak closure $\pi(\mathcal{A})^-$ of $\pi(\mathcal{A})$. Therefore there exists an operator h_π in $\pi(\mathcal{A})^-$

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_10

such that $\delta(a) = i[h_\pi, a]$. *The spectrum of the operator* h_π *depends in general on the representation* π.

For each representation π *the operator* h_π *is bounded and self-adjoint and therefore* e^{ith_π} *defines a one parameter norm- continuous group of unitary operators.* ◇

When the derivation δ is continuous *only in the weak topology* in general there is no group of automorphisms associated to it. An intermediate situation is given by derivations that are generators of one-parameter group of automorphisms but this group *is not implementable* by a continuous group $U(t)$ of unitary operators. A further more special case is that in which the generator of the group $U(t)$ is a self-adjoint operator *affiliated* to the algebra $\mathcal{A}$ (this means that its spectral projection are in the weak closure of $\pi(\mathcal{A})$ [13].

Definition 1 (*inner and weakly inner derivations*) The derivation δ is *inner* in the representation π if there exists $b \in \pi(\mathcal{A})$ such that $\delta\pi(a) = [b, \pi a]$. It is *weakly inner* (or *spacial*) in the representation π on the Hilbert space $\mathcal{H}_\pi$ if there exists a linear operator b_π on $\mathcal{H}_\pi$ such that $\pi(\delta(a)) = [b_\pi, \pi(a)] \ \ \forall \pi(a) \in D(\delta)$. We shall call b_π the generator of the group of automorphisms in the representation π. ◇

Notice that a derivation of a C^*-algebra may be inner in one representation and not inner in another representation.

In the remaining part of this Lecture we shall assume that the algebra $\mathcal{A}$ has the identity e and that $e \in D(\delta)$. In this case from Leibnitz' rule on derives $\delta(e) = 0$.

The following theorems are important, although the hypothesis are often too stringent for many physical systems. Recall that if $\mathcal{A}$ is an algebra of functions of a real variable and if f is differentiable then its derivative vanishes in the maximum and minimum points. The following result is the analogous property for C^*-algebras.

Recall that according to the Hahn-Banach theorem, for every $a \in \mathcal{A}$ there exists a state ϕ_a such that $\phi_a(a) = |\phi||a|$.

Lemma 1 *Let* δ *be a derivation of a* C^**-algebra* $\mathcal{A}$. *One has*

$$\phi(a) = |a||\phi|, \quad a \in D(\delta) \Rightarrow \phi(\delta(a) = 0. \tag{5}$$

◇

Proof Without loss of generality assume $|a| = |\phi_a| = 1$. Set $y = (I - a)^{\frac{1}{2}}$. Then

$$|\phi(\delta(a))|^2 = |\phi(\delta(I - y^2))|^2 = |\phi(\delta(y^2))|^2$$
$$= |\phi(y \cdot \delta(y)) + \phi(\delta(y) \cdot y)| \leq 4|\phi(y^2)\frac{1}{2}\phi(\delta(y)^2)|^{\frac{1}{2}} = 0 \tag{6}$$

since $\phi(y^2) = \phi(I - a) = |\phi| - \phi(a) = 0$. ♡

In the notation which we will use in chapter "Lecture 13: Weyl System, Weyl Algebra, Lifting Symplectic Maps. Magnetic Weyl Algebra" for the semigroups

we can express the relation (5) as follows: δ and $-\delta$ are *dissipative* (and δ is it conservative) relatively to the state ϕ.

Theorem 2 (Sakai) [12] *Every derivation δ such that $D(\delta) \equiv \mathcal{A}$ is a bounded derivation, and therefore inner by the Kadison-Sakai theorem.* ◇

For completeness we give here without proof two further important theorems, with the references for their proofs.

Theorem 3 (Ringrose) [11] *If $D(\delta) = \pi(\mathcal{A})$ in each representation π, there exists a derivation $\hat{\delta}_\pi$ of $\pi(\mathcal{A})^-$ which extends δ.* ◇

Theorem 4 (Christensen) [4] *If δ is as in Theorem 3, then it is weakly inner in each representation and it is possible to choose the generator b_π (such that $\delta\pi(a) = [b_\pi, \pi(a)]$) in the convex closure of the elements of the form U, V^* where U, V are unitary operator which belong to $\pi(\mathcal{A}')'$.* ◇

In the application it is important to know the assumptions under which one can apply the rules for derivation of composite functions.

For comparison recall the following lemma in the commutative case (chain rule).

Lemma 5 *Let $\mathcal{A}$ be an abelian C^*-algebra and let δ be a closed derivation.*
If $a \in D(\delta)$ and $g \in C^1(R)$, then $g(a) \in D(\delta)$ and one has

$$\delta(g(a)) = g'(a)\delta(a). \tag{7}$$

◇

Proof Equation (7) holds if g is a polynomial. According to the Stone-Weierstrass theorem if $g \in C^1$ one can find a sequence of polynomials $P_n(t)$ such that

$$sup_{|t|\leq|a|}|g(t) - P_n(t)| \to_{n\to\infty} 0 \quad sup_{|t|\leq|a|}|g'(t) - P_n'(t)| \to_{n\to\infty} 0$$

Then $g(a) = lim_{n\to\infty} P_n(a)$, $\quad lim_{t\to\infty}\delta(P_n(a)) = lim_{n\to\infty} P_n'(a) \cdot \delta(a) = g'(a) \cdot \delta(a)$.

By assumption δ is closed and therefore

$$g(a) \in D(\delta), \quad \delta(g(a)) = g'(a)\delta(a).$$

Therefore for an abelian C^*-algebra the domain of a derivation is closed under C^1 functional calculus. ♡

In the non-commutative case there is only *a weaker result.* In particular it is necessary to require further regularity properties of $g(t)$ and the statements hold only for the real elements of $\mathcal{A}$. One can give examples of function of class C^1 over a non-commutative C^*-algebra for which the C^1 functional calculus does not hold.

The weaker result is the following [2, 3].

Theorem 6 (Powers) *Let δ be a *-derivation of a C^*-algebra $\mathcal{A}$, and let δ be closed in the uniform topology. Let $f(t)$ be a real-valued function with Fourier transform which satisfies*

$$\int |p||\hat{f}(p)|dp < +\infty. \tag{8}$$

Let a be a real element of $\mathcal{A}$. Then $f(a) \in D(\delta)$ and

$$\delta(f(a)) = \frac{i}{2\pi} \int p\hat{f}(p)(\int_0^1 e^{itpa}\delta(a)e^{i(1-t)pa}dt)dp. \tag{9}$$

$\diamondsuit$

Proof We will give the proof in several steps.

1. Let $a = a^* \in D(\delta)$ and λ not in the spectrum of a. Then $a(\lambda I - a)^{-1} \in D(\delta)$. In this case the proof is obtained using the spectral representation of a and the fact the δ is closed.
2. If $e \in \mathcal{A}$ then $\delta(e) = 0$. Indeed if a is positive invertible one has when $\epsilon \to 0$

$$a[\epsilon\, e + a]^{-1} \to e, \qquad \delta(a[\epsilon\, e + a]^{-1}) \equiv -\epsilon[\epsilon\, e + a]^{-1}\delta(a)[\epsilon\, I + a]^{-1} \to 0 \tag{10}$$

   Since δ is closed, $\delta(e) = 0$ follows.
3. If $a = a^* \in D(\delta)$ then $e^{ipa} \in D(\delta)$ and

$$\delta(e^{ipa}) = ip \int_0^1 e^{itpa}\delta(a)e^{i(1-t)pa}dt \tag{11}$$

   Indeed from spectral theory one has

$$e^{ipa} = lim_{n\to\infty}(1 - \frac{ipa}{n})^{-n} \tag{12}$$

   and Leibniz's rule for polynomials

$$\delta((I - \frac{ipa}{n})^{-n}) = ip\sum_n n^{-1}(I - \frac{ipa}{n})^{-n})^{\frac{m+1}{n}}\delta(a)((I - \frac{ipa}{n})^{-n})^{1-\frac{m+1}{n}}. \tag{13}$$

   From this (4.30) follows. $\heartsuit$

2 Derivations and Groups of Automorphisms

In important cases the derivations are generators of one-parameter groups of automorphisms.

Definition 2 (*pre-generator*) The operation δ is said to be *a pre-generator* if there exists a one-parameter group of automorphisms γ_t of $\mathcal{A}$ such that for every $a \in D(\delta)$ one has $\delta(a) = lim_{t\to 0} t^{-1}(\gamma_t(a) - a)$. The closure of $\bar{\delta}$ is the *generator* of the group γ_t. $\diamond$

In many cases of interest in Physics one deals with unbounded derivations which are the (inductive) limit of bounded closed ones, and are therefore "internal" in a generalized sense. In important cases it is possible to prove that they are pre-generators and sometimes even to prove that they are weakly inner.

As an intermediate step we can look for conditions under which a symmetric operator h can be associated to the derivation δ in such a way that $\delta(a) = [h, a]$ holds on $D(\delta)$ without requiring that h be associated to $\mathcal{A}$.

A useful condition that is often verified is given by the following theorem, generalization of Dirac's Theorem that we have mentioned in chapter "Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories" [2, 3].

Theorem 7 (Bratteli) *Let δ be a symmetric derivation of a separable closed subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ ($\mathcal{H}$ separable). Let $\Omega \in \mathcal{H}$ be cyclic for $\mathcal{A}$ and set $\omega(a) = (\Omega, a\Omega)$.*

The following condition are equivalent

$$(i) \quad \exists C > 0 \ : \ |\omega(\delta(a))| \leq C[\omega(a^*a) + \omega(aa^*)]^{\frac{1}{2}} \quad \forall a \in D(\delta) \tag{14}$$

(ii) there exits a closed symmetric operator h with domain $D(h) = \mathcal{A}\Omega$ such that $\forall a \in \mathcal{A}, \ \forall \phi \in D(h)$

$$\delta(a)\phi = i[h, a]\phi \tag{15}$$

Moreover if $I \in \mathcal{A}$ then $2|h\Omega|^2 \leq C^2$. $\diamond$

Proof $(ii) \to i$. If $a \in \mathcal{A}$, then

$$\omega(\delta(a)) = (h\Omega, a\Omega) + (a^*\Omega, h\Omega) \leq |h\Omega|(|a\Omega| + |a^*\Omega|) \tag{16}$$

$(i) \to (ii)$ Set $\tilde{H}$ be the closure of $\{a\Omega, a^*\Omega\} \quad \forall a \in \mathcal{A}$. Define η on $\tilde{H}$ by

$$\eta\{a\Omega, a^*\Omega\} = i\omega(\delta(a)) \tag{17}$$

Notice that the definition (17) is well posed. Indeed if $a\Omega = 0$ one derives form (i) that $\omega(\delta(a)) = 0$. The map η is continuous in the topology of H (since by construction ω is a normal state) and therefore there exist $\{\phi_1, \phi_2\} \in \tilde{H}$ such that

$$i\omega(\delta(a)) = (\phi_1, a\Omega) + (a^*\Omega, \phi_2) \tag{18}$$

By assumption one has $(\delta(a))^* = \delta(a)^*$. Therefore taking difference of adjoints in (18)

$$2i\omega(\delta(a)) = (\phi_1 - \phi_2, a\Omega) - (a^*\Omega, \phi_1 - \phi_2) = 0$$

Define h as follows

$$D(h) \equiv \mathcal{A}\Omega, \qquad ha\Omega = i\delta(a)\Omega + \frac{a(\phi_1 - \phi_2)}{2}. \tag{19}$$

Since $I \in \mathcal{A}$ one has $h\Omega = \frac{\phi_1 - \phi_2}{2} \equiv \Omega_\delta$.

Let us verify that the definition (19) is well posed. If $a\Omega = 0$ from Eq. (31) in chapter "Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories" one derives $2i\omega(\delta(a)) + (a^*\Omega, \phi_1 - \phi_2) = 0$ and therefore $ha\Omega = 0$.

The operator h is densely defined due to the cyclicity of ω. It is symmetric

$$\begin{aligned}(ha\Omega, b\Omega) - (a\Omega, hb\Omega) &= i\omega(\delta(a^*b)) + (\Omega_\delta, a^*b\Omega) - (\Omega_\delta, a^*b\Omega)\delta \\ &= i\omega((\delta(a^*b)) - \frac{i}{2}[\omega((a^*b)) + \omega(\delta(a^*b))] = 0\end{aligned} \tag{20}$$

Moreover

$$\delta(a)b\Omega = [\delta(ab) - \delta(b)]\Omega = ihab\Omega - ab\Omega_\delta - iahb\Omega + ab\Omega_\delta = i[h, a]b\Omega \tag{21}$$

and therefore (20) holds on a dense domain. If $\mathcal{A}$ contains the identity

$$|h\Omega|^2 = |\Omega_\delta|^2 \leq \frac{1}{2}|\eta|^2 \leq \frac{C^2}{2}. \tag{22}$$

♡

One can notice that in the proof of Theorem 7 we have introduced a doubling of the Hilbert space and we have considered the action of a, respectively a^* on the two spaces. This can also be expressed introducing an *anti-involution* J which leaves Ω invariant. Since the two copies of $\mathcal{A}$ act in a direct sum of Hilbert spaces the copy acting on the second space can be regarded to be part of the commutant of the first copy of $\mathcal{A}$. With this interpretation, the vector Ω is cyclic both for $\mathcal{A}$ and for its commutant and Theorem 7 can be seen as a particular case of the Tomita-Takesaki theorem [14] that we will state in this Lecture and discuss in the second part of the Lectures.

Sometimes is may not know a-priori that δ is closed or at least closeable. In this case the following theorem may be useful.

Theorem 8 *Let $\mathcal{A}$ be a C^*-algebra and δ a densely defined * -derivation. Assume that there exists a collection Λ of states such that*

$$|\omega(\delta(a))| < c_\omega |a| \quad \forall \omega \in \Lambda, \qquad c_\omega < \infty \tag{23}$$

and that the representation $\oplus_{\omega\in\Lambda}\pi_\omega$ is faithful, where Π_ω is the G.N.S. representation associated to ω. Then δ is closable. ◇

Proof Set $\omega_{x,z}(a) \equiv \omega(xaz), \quad x,\ z \in D(\delta)$. Then $\omega_{x,z} \in D(\delta^*)$. Indeed Leibniz's rule gives

$$\omega(x\delta(a)z) = \omega(\delta(xaz)) - \omega(\delta(x)az) - \omega(xa\delta(z))$$

By (23) the first term to the right is bounded in norm by $c_\omega|xaz|$ and is therefore continuous in the topology of $\mathcal{A}$. The second and third terms are continuous since $x, z \in D(\delta)$ and ω is a state. By construction

$$\omega_{x,z}(a) = (\pi_\omega(x^*)\Omega_\omega, \pi_\omega(a)\pi_\omega(z)\Omega_\omega).$$

By assumption $\oplus_{\omega\in\Lambda}\pi_\omega$ is faithful and $D(\delta)$ is dense. Therefore the collection of states $\omega_{x,z}$ is separating for $\mathcal{A}$ and their linear hull is dense in the W^*-topology in the space of states of $\mathcal{A}$.

Therefore δ^* is densely defined and δ is closable. ♡

As a particular case, δ is closable if it is densely defined and there exists a faithful state ω such that

$$\omega(\delta(a)) = 0 \quad \forall a \in \mathcal{A} \tag{24}$$

Remark that if δ is a pre-generator the relation (24) implies that ω is invariant under the dual action of the one parameter group of automorphisms associated to δ.

It is also important to find conditions under which δ is a pre-generator. We shall discuss more in detail this problem in this chapter in the framework of semigroup theory.

Here we only mention a result of E. Nelson, frequently used.

3 Analytic Elements

Definition 3 (*analytic elements*) The element $a \in \mathcal{A}$ is analytic for δ if there exists $t > 0$ for which

$$\sum_{n=0}^{\infty} \frac{1}{n!} t^n |\delta^n(a)| < \infty \tag{25}$$

◇

Theorem 9 (Nelson) *If the derivation δ is densely defined and closable, then the following three statements are equivalent.*

(i) the closure of the operator δ is the generator of a norm continuous one-parameter group of automorphisms.
(ii) δ has a dense set $\{\sigma\}$ of analytic vectors and is conservative (i.e. $\sigma(\delta(a)) = 0$). ◇

Proof (i) $\rightarrow$ (ii) The generator of a one-parameter norm continuous group of automorphisms α_t of a C^*-algebra $\mathcal{A}$ is conservative.

Indeed if $\sigma(a) = |\sigma||a|$ one derives

$$\sigma(\delta(a)) = lim_{t\rightarrow\infty}(\sigma(\alpha_t(a)) - \sigma(a)) = lim_{t\rightarrow\infty}(\sigma(\alpha_t(a)) - |\sigma||a|). \quad (26)$$

From $|\alpha_t(a)| = |a|$ one derives $|\sigma(\alpha_t(a))| \leq |\sigma||a|$ and therefore $\sigma(\delta(a)) \leq 0$ whenever a is real. On the other hand also $-\delta$ is a generator and therefore $\sigma(\delta(a)) = 0$ for every real $a \in \mathcal{A}$ and then by linearity on all of $\mathcal{A}$.

We prove now that there exists a dense set of analytic vectors. Let $\alpha_t = e^{t\delta}$ and for all $a \in \mathcal{A}$ set

$$a_n \equiv \sqrt{\frac{n}{\pi}} \int_{-\infty}^{\infty} e^{-t^2} \alpha_t(a) dt. \quad (27)$$

It is easy to verify that a_n is entire analytic for every n. In particular defining $\alpha^f(a) = \int_{-\infty}^{\infty} f(t)\alpha_t(a)dt$ one has $\delta(\alpha^f(a)) = -\alpha^{f'}(a)$. Moreover $norm-lim_{n\rightarrow\infty} a_n = a$. Therefore the collection of the a_n is dense in $\mathcal{A}$.

(ii) $\rightarrow$ (i) This is a consequence of the Hille-Yosida theorem. A direct proof can be obtained noting that a construction similar to (25) provides entire elements starting form analytic ones and the set of elements which are obtained is still dense.

On entire elements the series $\sum_n \frac{1}{n!} t^n \delta^n(a) \equiv a_t$ converges uniformly and defines a one parameter group which is norm-continuous in t for every a. By continuity the map $a \rightarrow T_t(a)$ extends for each value of t to the entire algebra and the extension is still a group. If δ is conservative T_t for each value of t is an automorphism . ♡

4 Two Examples from Quantum Statistical Mechanics and Quantum Field Theory on a Lattice

We make use of Nelson's theorem to discuss the following examples which are of interest in Quantum Statistical Mechanics and in Quantum Field Theory.

4.1 Example 1

Let X be a lattice in R^d, let $\mathcal{F}$ denote the finite subsets of X. Let Λ be an element of $\mathcal{F}$. Let x be a point in the lattice and for each x set $\mathcal{A}_{\{x\}} \equiv M_n$ (the algebra of $n \times n$ matrices). Define

$$\mathcal{A}_\Lambda \equiv \otimes_{x\in\Lambda}\mathcal{A}_x \tag{28}$$

Notice the natural inclusions

$$\Lambda \subset \Sigma \to \mathcal{A}_\Lambda \otimes I_{\Sigma\ominus\Lambda} \subset \mathcal{A}_\Sigma. \tag{29}$$

Call $\mathcal{A}$ the inductive limit (with respect to inclusion) of $\mathcal{A}_\Lambda$, $\Lambda \in \mathcal{F}$.

Definition 5 (*interaction*) Define *interaction* a function Φ from $\mathcal{F}$ to $\mathcal{A}$ with the following properties
(a) $\forall \Lambda \in \mathcal{F}, \quad \Phi(\Lambda) \in \mathcal{A}_\Lambda$
(b) $\Phi^*(\Lambda) = \Phi(\Lambda)$ ◇

For $\Lambda \in \mathcal{F}$ and for all $a \in \cup_{\Lambda\in\mathcal{F}}\mathcal{A}_\Lambda$ define

$$\delta_\Lambda(a) = i[H_\Lambda, a], \qquad H_\Lambda = \sum_{W\in\Lambda} \Phi(W).$$

By construction

$$\Lambda \cap \Sigma = \emptyset \to \mathcal{A}_\Lambda \in \mathcal{A}'_\Sigma.$$

Assume that for every $\lambda > 0$

$$||\Phi||_\lambda \equiv \sum_n e^{\lambda n}[sup_{x\in X} \sum_{X\in\Lambda,\ |\Lambda|=n+1} |\Phi(\Lambda)|] < \infty \tag{30}$$

where $|\Lambda|$ is the number of lattice points in Λ.

Define for $\Lambda \in \mathcal{F}$

$$\delta_\Lambda(a) = i[H_\Lambda, a]. \tag{31}$$

Remark that according to (29) for every $a \in \cup_{\Sigma\in\mathcal{F}}\mathcal{A}_\Sigma$ the operator $[H_\Lambda, a]$ *does not depend on* Λ if Λ is sufficiently large. Therefore (31) defines a derivation δ on $\cup_{\Sigma\in\mathcal{F}}\mathcal{A}_\Sigma$.

The algebra $\mathcal{A}_\Lambda$ is a matrix algebra for every $\Lambda \in \mathcal{F}$ and therefore admits a trace (normalized to one on the identity of the algebra); the trace is a normal state. This construction is compatible with the inclusion and therefore extends to $\mathcal{A}$ and defines a state ω_0 which has the property

$$\omega_0(ab) = \omega_0(ba), \qquad \forall\ a,\ b\ \in \mathcal{A} \tag{32}$$

We extend by continuity ω_0 to a normal state ω on $\mathcal{A}$. On finite matrices A, B one has $Tr[A, B] = 0$ and therefore since the state is normal

$$\omega(\delta(a)) = 0 \qquad \forall a \in D(\delta) \tag{33}$$

Since ω is separating for $\mathcal{A}$ in the representation π_ω associated to ω the derivation δ is closable. By an explicit computation, using (30), it is possible to verify that each element of $\cup_{\Sigma\in\mathcal{F}}\mathcal{A}_\Sigma$ is analytic for δ. Therefore the assumptions of Nelson's theorem are satisfied and δ is a pre-generator.

The one-parameter group of automorphisms of $\pi_\omega(\mathcal{A})$ generated by the derivation δ is defined as (inductive) limit of *local dynamics* generated by the Hamiltonians H_Λ, $\Lambda\in\mathcal{F}$. From (34) one derives that there exists in the weak closure of $\Pi_\omega(\mathcal{A})$ a closed densely defined symmetric operator $\hat{H}$ such that

$$\delta(\pi_\omega(a)) = i[\hat{H}, \pi_\omega(a)]. \tag{34}$$

The domain of definition of $\hat{H}$ contains $\pi_\omega(D(\delta))\Omega$ (where Ω is the vector that the G.N.S. construction associates to ω). Formally

$$T_t = e^{it\hat{H}} \tag{35}$$

so that, *always formally*

$$\alpha_t(\pi_\omega(a)) = \sum_n \frac{i^n t^n}{n!}[\hat{H}, ...[\hat{H}, \pi_\omega(a)]..] = T_t\pi_\omega(a)T_{-t}. \tag{36}$$

By construction $T_t\Omega = \Omega$ and therefore (35) defines T_t as convergent series on the entire elements of $\Pi_\Omega(\mathcal{A})$. Using the fact that δ is conservative one verifies $|T_t\Omega| = |\Omega|$ and that T_t is defined as an invertible isometry on a dense domain in $\mathcal{H}$ (the Hilbert space of the representation π_ω), and can be extended by continuity to a unitary operator.

In the same way one verifies the group property $T_t\cdot T_s = T_{t+s}$. Strong continuity is verified noting that the convergence of the series (36) on the entire elements provides strong continuity in t of $T_t(a)$ when a is an entire element, and by continuity on $\pi_\omega(\mathcal{A})$. Therefore the map $t\to T_t$ is implemented by a strongly continuous one-parameter group of unitary operators on $\mathcal{H}$. By Stone's theorem $\hat{H}$ is self-adjoint.

It is worth noting that $\hat{H}$ is the generator of the limit dynamics defined in the G.N.S. representation associated to the tracial state ω but *it is not the formal limit* of $\pi_\omega(H_\Lambda)$ for $\Lambda\to X$; this limit *does not exists*. However it can be shown that one obtains a limit *compensating* with a family of operators $Z_\Lambda\in\pi'_\omega(\mathcal{A})'$. One can choose Z_Λ in such a way that

$$lim_{\Lambda\to\infty}[\pi_\omega(H_\Lambda), \pi_\Omega(a)] = [\hat{H}, \pi_\omega(a)] \tag{37}$$

◇

4.2 Example 2

Another example of application of Nelson's theorem, of interest in the theory of quantized fields, is the following. The structure is similar to the first example but specific properties are used to construct a representation based on a state *different from the tracial state*.

In the preceding example we choose $n = 2$ i.e. we associate to each point of the lattice the algebra B_2 of 2×2 matrices. Consider the automorphism defined in B_2 by

$$\alpha_t(a) \equiv e^{iNt} a e^{-iNt}, \quad N = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{38}$$

and extended to $\mathcal{A}$ by direct product.

The algebra B_2 is generated by the identity and by $A,\ A^*$ where

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad A^* = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{39}$$

It is easy to verify that $\alpha_t(A) = e^{it} A$, $\alpha_t(A^*) = e^{-it} A^*$. The infinitesimal generator of

$$\alpha_t^{\Lambda} = \otimes \alpha_t^k, \qquad k \in \Lambda \tag{40}$$

is $N_\Lambda \equiv \sum_{k \in \Lambda} N_k$.

It is possible to define the limit automorphism group α_t but it is not norm continuous in t (α_t^{Λ} is continuous but not continuous uniformly in Λ). Still if π_0 is the G.N.S. representation associated to the state $\omega_0 \equiv \otimes \xi_n$ one has $N\xi_n = 0 \ \ \forall n$ and it can be verified that the resolvent of $\pi_0(N_\Lambda)$ converges for $\Lambda \to X$ to the resolvent of a self-adjoint operator $\hat{N}$ with spectrum the integer numbers and

$$lim_{\Lambda \to X} \pi_0(\alpha_t^{\Lambda}) = \pi_0(\alpha_t(a)) = e^{it\hat{N}} \pi_0(a) e^{-it\hat{N}} \tag{41}$$

Of course, since this is a particular case of the previous example, the limit automorphisms group has also a pregenerator in the representation associated to the tracial state. But the formal generator is unbounded below and has a continuous spectrum. ◇

5 K.M.S. Condition

A particularly important relation in the theory of C^*-dynamical systems is the K.M.S. condition (after Kubo, Martin and Schwinger); we now describe it briefly. We will come back to the K.M.S. condition in connection with modular theory.

The K.M.S. condition [3] was introduced in Classical Statistical Mechanics by Dobrushin, Lanford and Ruelle as a boundary condition on an increasing family of sets; it assures uniqueness of the solution of an infinite family of differential equations that describe the thermodynamic equilibrium state of a system of infinitely many particles with a given density. Later Haag, Hugenholtz and Winnik noticed that the K.M.S. condition has a much wider range of applicability and could take the place of Gibbs postulate as a condition for the existence of an equilibrium (lowest energy) state at temperature T for a thermodynamical system in the framework of Quantum Statistical Mechanics.

While the Gibbs postulate refers to a finite system, and can be extended to infinite systems only through a limit procedure, the K.M.S. condition can be formulated directly for systems with infinitely many degrees of freedom. In the context of Quantum Statistical Mechanics the parameter β that enters in the K.M.S. condition has the role of inverse of the temperature.

In view of its abstract nature the K.M.S. condition can be applied to many different situations. It is applied in Quantum Field Theory to prove existence of a ground state and the implementability with unitary operators of the inhomogeneous Lorentz group of automorphisms. In this context the parameter β takes the role of parameter of acceleration. The analysis of the continuous representations of $O(3, 1)$ in Field Theory allow for a continuation to complex values of the parameter t and for imaginary values it is possible to consider the action of the *boosts* (transformation to a frame moving with constant speed).

We shall place later the K.M.S. condition in the context of Modular Theory and we shall establish a non-commutative analog of the Radom-Nykodim derivative for states [9, 10, 14]. Here we use it to describe conditions under which a one-parameter group of automorphisms of a C^*-algebra can be implemented by a group of unitary operators.

To introduce the K.M.S. condition, consider a C^*-dynamical system which we will denote by $\{\mathcal{A},\ \alpha_t\}$. Recall that an element $x \in \mathcal{A}$ is *analytic* for α_t if the map $t \to \alpha_t(x)$ has an extension to an entire analytic map $z \to \alpha_z(x)\ \ z \in \mathcal{C}$.

If $x \in \mathcal{A}$ define

$$x_n \equiv \frac{\sqrt{n}}{\sqrt{\pi}} \int \alpha_t(x) e^{-n^2 t} dt \tag{42}$$

For any integer n the element x_n is an element of $\mathcal{A}$ analytic in the entire plane and $lim_{n\to\infty}|x_n - x| = 0$. Therefore the set $\mathcal{A}^a$ of analytic vectors is a norm dense *-subalgebra of $\mathcal{A}$. The same conclusions is reached if one considers a W^*-dynamical system or a dynamical system with values in a von Neumann algebra.

Definition 6 (*K.M.S. condition*) Given a dynamical system $\{\mathcal{A}, \alpha_t\}$ denote by $\rho(t) \equiv \rho(\alpha_t)$ the action of the group α_t on the state ρ.

We say that the state ρ satisfies the K.M.S condition for the value β of the parameter, $0 < \beta < \infty$ if for all $a \in \mathcal{A}$ the function $\rho_t(a)$ it the boundary value on the real line of a function $\rho_z(a)$ analytic in the strip $0 < Imz < \beta$ and continuous at the boundary. Moreover the following relation between the boundary value at $=$ and at

$Imz = \beta$ must be satisfied

$$\rho_\beta(y\ \alpha_{t+i\beta}(x)) = \rho_\beta(\alpha_t(x)\ y), \qquad \forall x,\ y \in \mathcal{A} \quad t \in R. \tag{43}$$

♣

To see the connection with Quantum Statistical Mechanics consider first a system with a finite number of degrees of freedom described by $\mathcal{A}$, an algebra of $N \times N$ matrices. Let the evolution of the system be described by the Hamiltonian H. Consider the state ρ_β of $\mathcal{A}$ defined by

$$\rho_\beta(a) \equiv \frac{Tre^{-\beta H}a}{Tre^{-\beta H}} \quad a \in \mathcal{A} \tag{44}$$

where Tr denotes the Hilbert space trace and α_t denotes the group of automorphisms with H as generator

$$\alpha_t(a) \equiv e^{itH}ae^{-itH}. \tag{45}$$

It easy to see that ρ_β is invariant under the dual flow, and that it is an equilibrium state in the sense that the correlations

$$\rho_\beta((\alpha_t(a)\alpha_s(b))$$

depends only on $t-s$. On recognizes that it is the Gibbs for the system at temperature $T = \frac{1}{\beta}$.

By direct inspection the state ρ_β satisfies the K.M.S. condition (43).

Notice that if $\beta = 0$ the state is a trace and of $\beta = +\infty\ \ e^{-\beta H}$ is the projection on the state with energy equal to zero (ground state).

Consider now the first example given above, and let $\mathcal{A}_\Sigma$ the algebra associated to a finite region Σ. Consider only the interactions within Σ and summarize the interaction with the system which is outside Σ with boundary condition γ at the boundary $\partial\Sigma$. Let $H_{\Sigma,\gamma}$ the resulting Hamiltonian and $\rho_\beta^{\Sigma,\gamma}$ be its Gibbs state.

The K.M.S. condition is satisfied for any size of the region Σ and for any boundary condition γ. It therefore natural to attribute the K.M.S. condition to the entire system. In this more general context it not a priori clear whether the group of automorphisms has a generator (a hamiltonian) and whether, if this operator exists, it is bounded below and can be considered to be the limit of the generators associated to a collection of increasing regions Ω_n that cover asymptotically the entire space.

The interest of this example lies in the fact that if the algebra $\mathcal{A}$ is the inductive limit of matrix algebras $\mathcal{A}_\Sigma$ for which the K.M.S. holds trivially. Under suitable assumption the condition is satisfied also in the limit $n \to \infty$ and therefore can be used to prove that the limit automorphisms group admits a generator. This generator *is not in general* the limit of the generators for the matrix algebra and it is interesting to study its spectrum. This in some cases leads to the proof of existence and uniqueness of an equilibrium state at temperature $\frac{1}{\beta}$.

We extend the definition of K.M.S. state to cover also the cases $\beta = 0$ and $\beta = \infty$. We will say that ρ_0 satisfies the K.M.S. condition for the group α_t at $\beta = 0$ if

$$\rho_0(y\ \alpha_\zeta(x)) = \rho_0(\alpha_\zeta(x)\ y) \quad \forall x \in \mathcal{A}^a \quad y \in \mathcal{A}. \tag{46}$$

We will say that ρ_∞ satisfies the K.M.S. condition for the group α_t at infinity if for any $x \in \mathcal{A}^a$ and every $y \in \mathcal{A}$ the analytic function $f(\zeta) \equiv \rho_\infty(y\ \alpha_\zeta(x))$ satisfies

$$|f(\zeta)| \leq \|x\|\ \|y\| \quad if \quad Im\zeta \geq 0. \tag{47}$$

In this case the state ρ_∞ is said to be *ground state* relative to the group of automorphisms α_t that represents translation in time. Notice that the case $\beta = \infty$ corresponds in the Statistical Mechanics to zero temperature; this is the origin of the name *ground state* [3].

Proposition 10 *Let $\{\mathcal{A}, \alpha_t\}$ be a C^*-dynamical system and let $0 \leq \beta \leq \infty$. The following conditions on a state ρ are equivalent*
(1) ρ is β-K.M.S. state
(2) ρ satisfies the α_t K.M.S. condition for a dense set of elements $x \in \mathcal{A}^a$.
(3) For any pair $x \in \mathcal{A}^a$, $y \in \mathcal{A}$ there exists a function $f_\rho(\zeta)$ bounded continuous in the strip

$$\Omega_\beta \equiv \{\zeta \in \mathcal{C},\ 0 \leq Im\zeta \leq \beta\} \tag{48}$$

holomorphic in the interior of Ω and satisfying the boundary conditions

$$f_\rho(t) = \rho(y\ \alpha_t(x)), \quad f_\rho(t + i\beta) = \rho(\alpha_t(x)\ y). \tag{49}$$

If $\beta = \infty$ the last condition takes the form

$$f_\rho(t) = \rho(y\ \alpha_t\ (x)) \quad t \in R \quad \|f_\rho\| \leq \|x\|\ \|y\|. \tag{50}$$

◇

Proof The implications (1) → (2) → (3) are evident. To prove (3) → (1) assume that the condition is satisfied for all y and for all x in a dense subset of $\mathcal{A}^a$. In the case $\beta < \infty$ let $\{x_n\} \in \mathcal{A}^a$ be a sequence converging to x. This provides a sequence of analytic functions

$$f_n(\zeta) = \rho(y\ \alpha_\zeta(x_n), \quad f_n(\zeta + i\beta) = \rho(\alpha_\zeta(x_n)\ y) \tag{51}$$

all of which are bounded in Ω_β.

From the Pragmen-Lindelöf theorem

$$\begin{aligned} |f_n(\zeta) - f_m(\zeta)| &\leq \sup_{\zeta \in \partial\Omega_\beta} |f_n(\zeta) - f_m(\zeta)| \\ = \sup\{|\rho(y\ \alpha_t(x_n - x_m)|, |\rho(\alpha_t(x_n - x_m)\ y)|\} &\leq \|y\|\ \|x_n\ - x_m\| \end{aligned} \tag{52}$$

Therefore the sequence $\{f_n\}$ converges uniformly to a function $f \in C^b(\Omega)$ holomorphic in the interior of Ω and which satisfies the desired boundary conditions.

In case $\beta = \infty$ define $f_n(\zeta) \equiv \rho(y)\ \alpha_\zeta(x_n)$ and notice the K.M.S. condition at infinity provides

$$|f_n(\zeta) - f_m(\zeta)| \leq \|x_n - x_m\|\ \|y\| \tag{53}$$

when $\zeta \geq 0$. It follows that $\{f_n\}$ converges to a function $f(\zeta) \in C^b(\Omega)$ which is holomorphic in the interior. Moreover $f(t) = (y\ \alpha_t(x))\quad \forall t$ and $\|f\| \leq \|x\|\ \|y\|$. ♡

An important property of the K.M.S. condition is provided by the following Proposition [5, 13, 14].

Proposition 11 *Let $\{\mathcal{A}, \alpha_t\}$ be a C^*-dynamical system and let ρ_β be a state which satisfies the α_t-K.M.S. condition for a value β of the parameter* $(0 \leq \beta \leq \infty)$. *Then* ρ_β is invariant *under the dual action of the automorphisms group α_t. Moreover it is a equilibrium state, in the sense that for any value of β it has the property under the one-parameter group of automorphisms α_t*

$$\rho_\beta(\alpha_t(a)\ b) = \rho_\beta(a\alpha_{-t}(b)) \tag{54}$$

◇

Proof If $0 < \beta < \infty$ choose in (53) $y = e$ (the identity of the algebra; if the algebra does not have an identity, use a sequence of approximate identities).

From (50) for every $x \in \mathcal{A}^a$ one has $\rho_\beta(\alpha_{\zeta+i\beta}(x)) = \rho_\beta(\alpha_\zeta(x))$. Therefore the analytic function $f(\zeta) \equiv \rho_\beta(\alpha_\zeta(x))$ is bounded in the strip Ω_β and periodic of period β. It follows that it is bounded in $\mathcal{C}$ and therefore a constant according to Liouville's theorem.

Since $\mathcal{A}^a$ is dense in $\mathcal{A}$ it follows that ρ_β is invariant under the action of α_t. The property to be an equilibrium state is verified similarly.

If $\beta = 0$ the state is invariant by definition. If $\beta = \infty$ choosing for y the identity for every $x \in \mathcal{A}^a$ the function $1\ f\ :\ \zeta \to \rho_\infty(\alpha_\zeta(x))$ satisfies $|f(\zeta)| \leq \|x\|$ for $Im\zeta \geq 0$.

Taking into account $\rho_\infty = \rho_\infty^*$ one has $\alpha_\zeta(\bar{x}) = \phi(\alpha_\zeta(x^*)$. It follows that also when $\zeta \leq 0$ one has $|f(\zeta)| \leq \|x\|$ and f is an entire bounded function and therefore a constant. Also in this case ρ_∞ is invariant under α_t. ♡

One requires usually that the state ρ_β be separating ($\rho_\beta(a) > 0$ if $a > 0$) and cyclic. In the GNS representation π_{ρ_β} both the algebra $\mathcal{A}$ and its commutant $\mathcal{A}'$ have a cyclic and separating vector. This gives a natural conjugation between the algebra and its commutant and the K.M.S. condition provides a unique hamiltonian H that is the generator of the group of automorphisms. This hamiltonian has the form $H = H_1 - H_2$; the spectral projections of H_1 belong to $\mathcal{A}$ and those of H_2 belong to $\mathcal{A}'$.

Notice that for the algebra generated by the concrete von Neumann algebra $\mathcal{A}$ and its commutant the spectrum of the Hamiltonian is symmetric with respect to the origin and therefore is unbounded both above and below if the group of automorphisms is only weakly continuous.

Recall that in Theorem 7 we have constructed a one parameter group of inner automorphisms by doubling the Hilbert space providing therefore a natural isomorphic map (a natural conjugation J) between the von Neumann algebra $\mathcal{A}$ and its commutant. The vector Ω defines a state on the algebra generated by $\mathcal{A}$ and its commutant $\mathcal{A}'$ and is invariant under this map. The construction given by Bratteli correspond therefore to the case $H = 0$.

If one applies the Bratteli construction stating with a KMS state but giving an extra "twist" (equivalent to the multiplication by $e^{-\beta H}$) one arrives at a state that is cyclic and separating (as in Bratteli) and satisfies the KMS condition with hamiltonian H and parameter β.

This can be regarded as a particular instance of the theorems of Tomita and Tomita-Takesaki that we will state without proof in what follows.

6 Modular Structure

Recall that for a von Neumann algebra $\mathcal{M}$ on a Hilbert space $\mathcal{H}$ the bicommutant theorem holds $\mathcal{M} = \mathcal{M}''$ where

$$\mathcal{M}' \equiv \{a \in \mathcal{B}(\mathcal{H}), : ab = ba \ \ \forall b \in \mathcal{M}\} \tag{55}$$

Definition 7 (*cyclic; separating*) A vector $\phi \in \mathcal{H}$ is *cyclic* for the algebra $\mathcal{M}$ if the subspace $\{a\phi, \ a \in \mathcal{M}\}$ is dense in $\mathcal{H}$. It is *separating* for $\mathcal{M}$ if $a\phi = 0$ for some $a \in \mathcal{M}$ implies $a = 0$. ◇

Definition 8 (*modular vector*) A vector $\phi \in \mathcal{H}$ is *modular* for $\mathcal{M}$ if it is both cyclic and separating for $\mathcal{M}$. ◇

A vector state ω_ϕ is separating for $\mathcal{M}$ iff the corresponding normal state $\omega_\phi(a) = (\phi, a\phi)$ is faithful. The *support* s_ω of a normal state ω is by definition the smallest orthogonal projection $P \in \mathcal{M}$ such that $\omega(P) = 1$. It follows that $\omega(a^*a) = 0$ if and only if $s_\omega = 0$. The state ω is *faithful* iff $s_\omega = I$.

The support of the vector state ω_ϕ is the orthogonal projection on the closure of $\mathcal{M}'\phi$. Therefore a vector $\phi \in \mathcal{H}$ is separating for $\mathcal{M}$ iff it is cyclic for $\mathcal{M}'$ and that $\phi \in \mathcal{H}$ is separating for $\mathcal{M}$ only if it cyclic for $\mathcal{M}'$.

Let ω be a state on the C^*algebra $\mathcal{A}$. Denote by $(\mathcal{H}, \pi_\omega, \Omega_\omega)$ the G.N.S representation of $\mathcal{A}$ defined by ω.

Definition 9 (*modular state*) The state ω is *modular* if the vector Ω_ω is modular (cyclic and separating) for the von Neumann algebra $\mathcal{A}_\omega \equiv (\pi_\omega(\mathcal{A}))''$.

Note that the state ω is modular if the vector Ω_ω is faithful on $\mathcal{A}_\omega$.

The following result links modular theory with K.M.S states [14]:

Theorem 12 *Let $\mathcal{A}, \tau$ be a C^* dynamical system. Any (τ, β)-KMS state, $\beta \in R$, is modular.* ◇

Let ϕ be a modular vector for $\mathcal{M}$. Since ϕ is separating, the map

$$a\phi \to a^*\phi \tag{56}$$

defines an antilinear involution S_0 of $\mathcal{M}\phi$. Recall that ϕ is cyclic for $\mathcal{M}'$. This implies that S_0 is closable and its closure S is involutive; it has a densely defined adjoint since ϕ is cyclic for $\mathcal{M}$.

Let Δ be the self-adjoint operator S^*S and define J by the polar decomposition $S = J\sqrt{\Delta}$. Since S is injective and has dense range the operator J is anti-unitary. From $I = S^2 = J\sqrt{\Delta}J\sqrt{\Delta}$ it follows

$$J\sqrt{\Delta} = \sqrt{\Delta}^{-1}J^+, \qquad J^2\sqrt{\Delta} = J\Delta^{-\frac{1}{2}} \tag{57}$$

The unicity of the polar decomposition yields $J^2 = I$ and then $J = J^*$.

Definition 10 (*modular operator, modular conjugation*) The positive self-adjoint operator Δ is the *modular operator* and the anti-unitary operator J is the *modular conjugation* of the pair $\mathcal{M}, \phi$. ◇

Theorem 13 (Tomita-Takesaki) [14] *Every normal faithful semi-finite weight ϕ on a W^*-algebra $\mathcal{M}$ determines a one-parameter group of $*$- automorphisms with generator $log\Delta$ and an anti-linear $*$-automorphism with a conjugation operator J_ϕ*

$$\begin{aligned} J_\phi &: \mathcal{M} \to J_\phi \pi_\phi(\mathcal{M}) J_\phi \in \mathcal{M}' \\ J_\phi a J_\phi = a^* \quad \forall a \in \mathcal{M} \cap \mathcal{M}' \quad & Ad(e^{itlog\Delta})a = a \end{aligned} \tag{58}$$

for all a in the quotient space with respect to the kernel of the representation. Notice that $e^{itlog\Delta} = \Delta^{it}$. ◇

We will discuss the Tomita-Takesaki theorem in the second part of these Lectures.

Definition 11 (*modular group*) The group of automorphisms defined by $\sigma_t(a) = \Delta^{it}a\Delta^{-it}$ is the *modular group* of the pair $\mathcal{M}, \phi$. ◇

The modular group can be defined more generally if ω is a faithful normal state of $\mathcal{M}$ and Δ is the modular operator for $\{\pi_\omega(\mathcal{A}, \Omega_\omega\}$ then $\sigma_\omega(t)(a) = \pi_\omega^{-1}(\Delta^{it}\pi_\omega(a)\Delta^{-it})$ is the modular group of ω.

The main property of the modular group is the following result

Theorem 14 (Takesaki) [14] *Let ω be a faithful normal state of the von Neumann algebra $\mathcal{M}$. Then ω is a K.M.S. state for the modular group σ_ω at inverse temperature β^{-1}. Moreover the modular group is the only dynamical group acting on $\mathcal{M}$ for which ω has this property.* ◇

If the normal semi-finite weight is a trace, then Δ_π is the unit operator in the representation π.

This leads to the construction of an analog of the Hilbert-Schmidt representation, called *standard representation* developed for countably additive W^*-algebras by Araki [1] and for arbitrary W^*-algebras by Haagerup [6].

Recalling the definition of K.M.S. condition, we see that as a result of the theory of Tomita and Takesaki every semi-finite weight on a W^* algebra satisfies the K.M.S. condition with respect to its modular operator. We will return to the the Tomita-Takesaki theorem in the second part of these Lectures. Its importance there will be to construct a non-commutative analog of the Laplacian, the Dirichlet forms and the positivity preserving contraction semigroups.

7 Standard Cones

Other central objects in Modular Theory are standard cones. We shall discuss the properties of these cones in the second part of these lectures. Here we point out only the following properties [8, 14]

A subspace $\mathcal{K} \subset \mathcal{H}$ is called a *cone* iff $\lambda\xi \in \mathcal{K}$ for all $\xi \in \mathcal{K}$ and for all $\lambda > 0$.

A cone is called *self-polar* iff

$$\mathcal{K} = \{\zeta \in \mathcal{H} | (\xi, \zeta) \geq 0 \quad \forall \xi \in \mathcal{D}\} \tag{59}$$

Every self-polar cone is *pointed* ($\mathcal{K} \cap -\mathcal{K} = \emptyset$), spans linearly $\mathcal{H}$ and determines uniquely a conjugation J in $\mathcal{H}$ and a unique order on the set $\mathcal{H}^{s.a.} = \{\xi \in \mathcal{H},\ J\xi = \xi$ given by

$$\xi \leq = \zeta \leftrightarrow \xi - \zeta \in \mathcal{K} \quad \forall \xi,\ \zeta \in \mathcal{H}^{s.a.} \tag{60}$$

Definition 19 (*standard cone*) The *standard cone* associated to the pair $\mathcal{M}, \omega$ is the closed subset of $\mathcal{H}$ defined by

$$\mathcal{H}_+ = \{aJaJ\phi,\ a \in \mathcal{M}\}^{cl} \tag{61}$$

$\diamond$

A closed convex self-polar cone in $\mathcal{H}$ is a standard cone if, for a given vector Ω cyclic and separating with respect to a von Neumann algebra $\mathcal{M}$ on $\mathcal{H}$ one has

$$\mathcal{H}^+_\Omega = \Delta^{\frac{1}{4}}_\Omega \mathcal{M}^+ \Omega \tag{62}$$

where Δ_Ω, J_Ω are the modular operator and modular conjugation associated with $\mathcal{M}, \Omega$.

Theorem 15 [14] *The standard cone $\mathcal{H}_+$ is self-dual*

$$\mathcal{H}_+ = (\mathcal{H}_+)^* \equiv \{\Omega \in \mathcal{H} \mid (\phi, \Omega) \geq 0\} \quad \forall \phi \in \mathcal{H}_+ \tag{63}$$

Moreover

$$J\phi = \phi \quad \forall \phi \in \mathcal{H}_+ \qquad aJaJ \in \mathcal{H}_+ \quad \forall a \in \mathcal{M} \tag{64}$$

◇

8 Standard Representation (Standard Form)

The theory of Tomita-Takesaki provides a refinement of the G.N.S. representation called *standard representation* [13, 14]. It is particularly adapted to the setting of W^*-algebras.

If $\mathcal{N}$ is W^* algebra, $\mathcal{H}$ is a Hilbert space, $\mathcal{H}^+ \subset \mathcal{H}$ is a self-polar cone, π a non-degenerate faithful normal representation of $\mathcal{N}$ on $\mathcal{H}$ and J is a conjugation on $\mathcal{H}$ then the quadruple $\mathcal{H}, \pi(\mathcal{N})J, \mathcal{H}^+$ is called *standard form* of $\mathcal{N}$ iff
(1)

$$J\pi(\mathcal{N})J = \pi(\mathcal{N})' \tag{65}$$

(2)

$$\xi \in \mathcal{H}^+ \to J\xi = \xi \tag{66}$$

(3)

$$\pi(a)J\pi(a)J\mathcal{H}^+ \subseteq \mathcal{H}^+ \tag{67}$$

(4)

$$a \in C_{\mathcal{N}} \to JaJ = a^* \tag{68}$$

In (4) we have denoted by $C_{\mathcal{N}}$ the center of $\mathcal{N}$.

In such a case J is called *standard conjugation* while $\mathcal{H}^+$ is a standard cone. If the elements of $\mathcal{N}$ are identified with the elements of $\pi(\mathcal{N})$ acting on $\mathcal{H}$ then one says that $\mathcal{N}$ is *in standard form*. The standard form representation satisfies the following properties [13]
(1)

$$\forall \phi \in \mathcal{N}_*^+ \quad \exists! \ \xi_\pi(\phi) \in \mathcal{H}^+ \quad \phi(a) = (\xi_\pi(\phi), \pi(a)\xi_\pi(\phi) \tag{69}$$

The vector $\xi_\pi(\phi)$ is called *standard vector representative* of ϕ.

(2) the map $\xi_+^\pi : \mathcal{H}^+ \ni \xi \to \phi_\xi \in \mathcal{N}_*^+$ is a bijective norm continuous homeomorphism. The map $\phi \to \xi_\pi(\phi)$ preserves order and

$$\|\xi - \zeta\|^2 \leq \| \phi_\xi - \phi_\zeta \| \leq 2\|\xi - \zeta\|^2 \tag{70}$$

(3)

$$\xi \in \mathcal{H}^+ \to J\xi \in \mathcal{H}^* \tag{71}$$

(4) For all $\zeta \in \mathcal{H}$ there exist *a unique* $\xi \in \mathcal{H}^+$ and a unique partial isometry $v \in \pi(\mathcal{N})$ such that $\zeta = v\xi$ and $v^*v = P(\xi)$ where $P(\xi)$ denotes the projection onto the closure of $\mathcal{N}'\xi$.
(5) $\mathcal{H}^+$ is closed and convex
(6) $\cup_{a\in\pi(\mathcal{N})}(aj(a)\xi_\pi(\phi)$ is dense in $\mathcal{H}^+$, $\xi_\pi(\phi)$ is cyclic and separating for $\pi(\mathcal{N})$ and $\mathcal{H}^+$ is a natural cone which satisfies

$$\Delta_{\xi_\pi}^{it}(\phi)\mathcal{H}^+_{\xi_\pi(\phi)} = \mathcal{H}^+_{\xi_\pi(\phi)} \tag{72}$$

Every W^*-algebra $\mathcal{N}$ has a faithful representation π such that $\pi(\mathcal{N})$ is in standard form and this representation is unique up to unitary equivalence.

The following statements are equivalent
(1) ω is faithful
(2) ϕ_ω is separating for $\pi(\mathcal{N})$
(3) ϕ_ω is cyclic for $\pi(\mathcal{N})$.

Definition 12 (*polar decomposition*) The relation $\zeta = v\xi$ is called *polar decomposition* of ζ. ♣

9 Standard Liouvillian

Given a W^* dynamical system $\mathcal{N}$, R, $\alpha(t)$ and a standard representation $\mathcal{H}$, π, J, $\mathcal{H}^+$ there is a *unique* representation of the group of automorphisms $\alpha(t)$ by a group of unitary operators and a unique self-adjoint generator. This generator is called *standard Liouvillian* [13]. It is not called *Hamiltonian* because in general it is not bounded either above or below.

For a W^*-algebra this property of representability derives from the uniqueness property of representation of the positive elements of $\mathcal{N}_*$ in terms of a standard cone of a standard representation of $\mathcal{N}$. Any $\alpha \in Aut(\mathcal{N})$ defines a unique map $u : \mathcal{H}^+ \to \mathcal{H}^+$ by

$$u\xi_\pi(\phi) = \xi_\pi(\alpha_*(\phi)) \quad \forall \phi \in \mathcal{N}_*^+ \tag{73}$$

This map is linear, can be extended to a unitary operator on $\mathcal{H}$ and satisfies

$$u\pi(a)u^* = \pi(\alpha(a)\forall a \in) \tag{74}$$

These unitary operators can be chosen to form a strongly continuous one-parameter group. Thus if $\mathcal{N}$, R, α is a W^* -dynamical system with $\mathcal{N}$ in standard form, there exists a strongly continuous group of unitary operators $V_\alpha(t),\ t \in R$ and a unique self-adjoint operator L_α on $\mathcal{H}$, called *standard Liouvillian* such that
(i)

$$V_\alpha(t) = e^{itL_\alpha}$$

(ii)

$$e^{itL_\alpha}\mathcal{H}^+ = \mathcal{H}^+$$

(iii)

$$JL_\alpha + L_\alpha J = 0$$

Moreover $V_\alpha(t)$ satisfies

$$e^{itL_\alpha}\mathcal{N}'e^{-itL_\alpha} = \mathcal{N}'$$

The definition of the standard Liouvillian L_α *depends only* on the W^*-dynamical system and a standard representation α of the W^*-algebra.

As an example [3] consider a W^*-algebra $\mathcal{N}$ and its standard representation $\mathcal{H}, \pi, J, \mathcal{H}^+$. The algebra $\mathcal{N}$ is equipped with group of unitary operators $U(t) = e^{itH} \in \pi(\mathcal{N})$ with $t \in R$ and $H \in \pi(\mathcal{N})^{s.a.}$ which is a unitary implementation of $R \ni t \to \alpha(t) \in Aut(\mathcal{N})$

$$\pi(\alpha_t(a)) = e^{itH}\pi(a)e^{-itH} \quad \forall t \in R \quad \forall a \in \mathcal{N} \tag{75}$$

If $\omega \in \mathcal{M}_+^*$ let $\xi_\pi(\omega)$ be its standard vector representative. Then by definition

$$\omega(\alpha_t(a)) = e^{-itH}\xi_\pi(\omega), \pi(a)e^{-itH}\xi_\pi(\omega) \quad \forall t \in R, \forall a \in \mathcal{N} \tag{76}$$

However generically $\xi_\pi(\omega)(t) \notin \mathcal{H}^+$.

On the other hand, the group $V(t)$ of unitary operators, *uniquely determined* by the condition $V(t)\mathcal{H}^+ = \mathcal{H}^+$ also implements α_t in the representation π:

$$V_\alpha(t)\pi(a)V_\alpha(t)^* = \pi(\alpha_t(a)) = U(t)aU(t)^* \quad \forall a \in \mathcal{M}, \forall t \in R \tag{77}$$

But $V_\alpha(t) \neq U_\alpha(t)$; indeed

$$V_\alpha(t) = U(t)JU(t)J \tag{78}$$

If the concrete representation of $\mathcal{N}$ is a semifinite von Neumann algebra $\mathcal{M}$, equation (78) implies that the standard Liouvillian L^{α} generator of $V_{\alpha}(t)$ is related to the generator H of $U(t)$ by

$$L_{\alpha} = H - JHJ \tag{79}$$

Hence $L_{\alpha} \notin \pi(\mathcal{M}^{s.a})$. The spectrum of L_{α} is given by

$$sp(L_{\alpha}) = \{\lambda_1 - \lambda_2 \quad \lambda_1, \lambda_2 \in spH\} \tag{80}$$

The standard vector representative of the evolved state $\omega(t)$ is given by

$$\xi_{\omega(t)} = (\rho_{\omega}(t))^{\frac{1}{2}} e^{-itH} \rho_{\omega}^{\frac{1}{2}} e^{itH} \neq \xi_{\omega}(t) \tag{81}$$

If $\mathcal{M}$ is a type I algebra, the standard Liouvillian takes the form

$$L_{\alpha} = H \otimes I - I \otimes H \tag{82}$$

and acts on $L_2(\mathcal{B}(\mathcal{H}), tr)$.

References

1. Araki, H. (1974). *Pacific Journal of Mathematics*, *50*, 309–354.
2. Bratteli, O. (1986). *Derivations, dissipations and group action on C*-algebras* (Vol. 1229). Lecture notes in mathematics. Berlin: Springer.
3. Bratteli, O., & Robinson, D. W. (1986). *Operator algebras quantum mechanics 1, 2* Springer, New York.
4. Christensen (1979). *Journal of the London Mathematical Society*, *20*, 358.
5. Dixmier, J. *C*-algebres*. Paris: Gallimard.
6. Haagerup, U. (1975). *Mathematical Scandinava*, *37*, 271–283.
7. Kadison, R. (1966). *Annals of Mathematics*, *83*, 273–296 and 83 (1966), 280–299.
8. Kadison, R., & Ringrose, J. *Introduction to C*-algebras* (Vols. I & II).
9. Pedersen B. E., & Takesaki, M. The Radon-Nykodym derivative for von Neumann algebras.
10. Pedersen, G. K. (1979). *C*-algebras and their automorphism groups*. St Louis: Academic Press.
11. Ringrose, J. R. (1972). *Journal of the London Mathematical Society*, *5*, 432.
12. Sakai, H. (1960). *Tohoku Mathematical Journal*, *12*, 31.
13. Stratila, S., & Szido, L. (1975). *Lectures on von Neumann algebras*, Bucharest: Editura Academiei.
14. Takesaki, M. (1979, 2002). *Theory of operator algebras* (Vols. I & II) Berlin: Springer.

Lecture 11: Semigroups and Dissipations. Markov Approximation. Quantum Dynamical Semigroups I

In "Lecture 10: Derivations and Generators. K.M.S. Condition. Elements of Modular Structure. Standard Form" we have discussed some aspects of quantum dynamics when it is represented by a one-parameter group of automorphisms of a C^*-algebra. This dynamics is *reversible* and it can describe isolated systems, in analogy with Hamilton's dynamics in Classical Mechanics.

In Classical Mechanics one can describe also *dissipative* systems, i.e. systems that typically evolve towards an equilibrium position. These systems *dissipate* energy through interaction with the environment; it is assumed in general that the environment in not affected by the interaction (a typical example is a thermostat) and that the evolution is described by a *semigroup*, e.g. the heat semigroup that has the Laplacian as generator.

In Classical Mechanics, under appropriate assumptions, one proves that by *averaging* over the degrees of freedom of the environment one can describe the system by a semigroup with a precise estimate of the error made. One obtains in this way an equation for Liouville distributions (in configuration space). A typical example is the derivation of the Boltzmann equation for a classical gas. The assumption in this case is the *stossenzahlansatz* i.e. the assumptions that the interactions may be considered statistically independent.

In this Lecture we shall show that also in Quantum Mechanics the description of the evolution of an open system as a semigroup requires an approximation (*Markov approximation*). This places strong restrictions on the interaction with the environment, which are not always satisfied for physical systems.

In Classical Mechanics it is possible (and sometimes more convenient) to describe the evolution of a dissipative system by giving a measure on the possible trajectories that the system can follow. In this approach the motion of the dissipative system is regarded to be so random that it is not possible to define the velocity of a single point (in configuration or in phase space).

The role of the ordinary differential equations is now taken by *stochastic differential equations* and correspondingly the evolution is a *stochastic process*. Typical

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_11

examples are Brownian motion and the Ornstein-Uhlembeck process. We will discuss them briefly in the second part of these Lectures.

For these systems the trajectories are *continuous but not differentiable*; the probability that a given trajectory be followed is given by a probability distributions that by duality satisfies a parabolic equation.

In Quantum Mechanics a similar approach has been proposed (Nelson Stochastic Mechanics) but has not been completed so far. For an indication of the route followed and for several interesting results one can consult e.g. [7, 8]. We shall not discuss further this interesting approach.

In this Lecture we give some elements of the theory of semi-groups on Banach spaces. One can find an extensive analysis on semigroups on function spaces e.g. in [1, 5, 9, 12]. In the second part of the Lecture we shall begin the analysis of the quantum case.

In the theory of Classical Dynamical Systems the Banach space considered is the space $C(X)$ of continuous function on configuration space. In Quantum Mechanics its role will be taken by the Banach space $\mathcal{B}(\mathcal{H})$ of bounded operator on a (separable) Hilbert space $\mathcal{H}$. Of course both $C(X)$ and $\mathcal{B}(\mathcal{H})$ have further structures as C^*-algebras. In the second part of this Lecture we shall specialize to semigroups on C^* algebras i.e. the case of interest in Quantum Mechanics.

1 Semigroups on Banach Spaces: Generalities

We recall elementary notions about semigroups and about their infinitesimal generators, i.e. *dissipations*. They have a role analogous to that of derivations in the conservative case.

Definition 1 (*semigroup on a Banach space*) A *semigroup* on a Banach space Banach X is a homeomorphism $\{T_t\}$ of the semigroup R^+ in the bounded operators on X. ◇

We will consider continuity in the three following topologies; notice that for semigroups of bounded operators it is enough to assume continuity at $t = 0$. We denote by X^* the topological dual of X (linear functions on X continuous in the norm topology)

(1) *Uniform continuity*: $\lim_{t\to 0} \sup_{x\in X,|x|\leq 1} \|T_t x - x\| = 0$;
(2) *Strong continuity*: $\lim_{t\to 0} \|T_t x - x\| = 0, \quad \forall x \in X$;
(3) *Weak continuity*: $\lim_{t\to 0} \phi(T_t x - x) = 0, \quad \forall x \in X, \quad \phi \in X^*$.

Uniform continuity is almost never verified for system of interest for Physics. This follows form the following theorem.

Theorem 1 *Denote by $\mathcal{B}(X)$ the collection of all bounded linear closed operators on the Banach space X and with $\|A\|$ the norm of $A \in \mathcal{B}(X)$.*

For a semigroup on X the following conditions are equivalent:

(i) $T_t = \sum_n \frac{t^n}{n!} A^n$ *is norm convergent;*
(ii) T_t *is uniformly continuous in* t*;*
(iii) $x \mapsto T_t x$ *is uniformly continuous at the origin in* $R^+ \times X$*;*
(iv) There exists $A \in \mathcal{B}(X)$ *such that in norm* $\lim_{t\to 0} t^{-1}(T_t - I) = A$. ◇

Proof The implications (i) → (ii), (iii), (iv); (ii) → (iii), (iv) → (ii); (iii)→ (ii) are obvious (the last one due to the semigroup property).

We prove (ii) → (i). By the assumption of uniform continuity

$$\tau^{-1} \int_0^\tau T_t dt \equiv I + \tau^{-1} \int_0^\tau (T_t - I) dt \tag{1}$$

defines, for τ sufficiently small, a linear operator on X which differs little form the identity and is therefore invertible. Indeed one has

$$\left\| \tau^{-1} \int_0^\tau (T_t - I) dt \right\| \leq \sup_{0\leq t\leq \tau} \|T_t - I\|, \qquad \lim \sup_{t\to 0} \|T_t - I\| = 0$$

Define

$$A_\tau \equiv (T_\tau - I) \left(\int_0^\tau T_t dt \right)^{-1}. \tag{2}$$

Using the semigroup property one verifies $A_{2\tau} = A_\tau$ and by iteration $A_{N\tau} = A_\tau$. By continuity A_τ is independent from τ. Setting $A \equiv A_\tau \ \forall \tau$ one has

$$T_\tau = I + A \int_0^\tau T_t dt \tag{3}$$

and from this one proves (i) by iteration (the series is absolutely convergent since A is bounded). ♡

The operator A associated to a norm-continuous semigroup T_t satisfies

$$\frac{d}{dt} T_t(x) = A T_t(x), \qquad \forall x \in X. \tag{4}$$

Consider now the case in which the semigroup is continuous only in the strong topology. There still exist an operator A which satisfies (4) on its (dense) domain but this operator *is in general unbounded*.

Let us define also in this case

$$A_\tau x = x - \tau^{-1} \int_0^\tau T_t x dt \tag{5}$$

This operator is still defined for all $x \in X$ but may be non-invertible. For each $x \in X$ there exits τ_x such that $I - \tau_x^{-1} \int_0^{\tau_x} T_t \, dt$ is invertible, but it may happen that $\inf_x \tau_x = 0$ (notice that X is not compact in general).

It is still true however that there exists a dense set Y in X on which the operator (5) is invertible. For each $x \in X$ define

$$y_{\tau,x} \equiv \tau^{-1} \int_0^{\tau} T_t x \, dt, \qquad \tau > 0, \quad x \in Y. \tag{6}$$

Because of continuity with respect to t the set of all $y_{\tau,x}$ is dense in X. Moreover

$$\left(\tau^{-1} \int_0^{\tau} T_t dt\right)^{-1} y_{\tau,x} = x,$$

therefore the map $x \to y_{x,\tau}$ is invertible, and it is easy to verify that

$$\lim_{t \to 0} h^{-1}(T_h - I) y_{x,\tau} = A \, y_{x,\tau} \tag{7}$$

Hence the operator A is densely defined and

$$Ay = (T_\tau - I)\left(\int_0^{\tau} dt T_t\right)^{-1} y \qquad y \in D(A).$$

Therefore

$$T_\tau y = y + A \int_0^{\tau} T_t y \, dt. \tag{8}$$

♡

Definition 2 (*generator*) We call the operator A *generator* of the semigroup T_t.

◇

Remark that, if the semigroup is not norm-continuous, in general (8) *can be iterated indefinitely only in a subset of* Ω. This poses a problem analogous to that encountered in the definition of a generator for reversible systems: given an operator A on the Banach space X with domain $D(A)$, find the conditions on A under which (6) has a unique solution. This gives the condition for existence of a semigroup $x \mapsto T_t x$, continuous in t for each $x \in X$ which satisfies (6) in a dense domain (and therefore has A as generator).

2 Contraction Semigroups

Definition 3 (*contractions*) A semigroup $t \mapsto T_t, t \geq 0$, on a Banach space X is a *contraction semigroup* if for every $t \in [0, \infty)$ the map T_t is a contraction $(\forall x \in X, \quad \|T_t x\| \leq \|x\|)$. ◇

We remark that all automorphism groups are contraction semigroups. Moreover if $\|T_t x\| \leq \|e^{\alpha t} x\|$ for some $\alpha > 0$ then $T'_t \equiv e^{-\alpha t} T_t$ defines a contraction semigroup. The same holds if $\|T_t x\| \leq c\|x\|, c > 1$.

An important element in the study of contraction semigroups is the *resolvent operator*.

Definition 4 (*resolvent*) Let T_t be a contraction semigroup on the Banach space X. The collection of operators R_z with parameter z defined by

$$R_z x \equiv -\int_0^\infty e^{-tz} T_t x \, dt, \qquad \mathrm{Re}(z) > 0 \quad x \in X \tag{9}$$

is called *resolvent* of the semigroup. ◇

It is easy to verify that

$$\|R_z x\| \leq \int_0^\infty e^{-t\mathrm{Re}(z)} \|T_t x\| \, dt \leq \frac{1}{\mathrm{Re}(z)} \|x\| \tag{10}$$

and therefore

$$\|R_z\| \leq \frac{1}{|\mathrm{Re}(z)|}. \tag{11}$$

It follows that R_z, $\mathrm{Re}(z) > 0$, is a bounded operator with dense range. Moreover

$$(A - zI)R_z = I, \qquad R_z(A - zI)x = x, \quad x \in D(A). \tag{12}$$

Indeed

$$AR_z x = \lim_{h \to 0} h^{-1}(T_h - I)R_z x =$$

$$\lim_{h \to 0} h^{-1}(1 - e^{-hz}) \int_0^\infty dt e^{-zt} T_t x + h^{-1} e^{zh} \int_0^\infty dt e^{-zt} T_t x = zR_z x + x \tag{13}$$

which proves the first part of (12). The second is proved similarly by choosing $x \in D(A)$.

We have therefore proved that (11) and (12) are *necessary conditions* for the existence of a contraction semigroup with A as generator. We shall prove now that these conditions are also *sufficient*.

Theorem 2 (Hille–Yosida) [1, 12] *If an operator A on a Banach space X is such that $(A - \lambda I)^{-1}$, $\lambda > 0$, defines a map from $X \to D(A)$ and the following inequality holds*

$$|A - \lambda I|^{-1} \leq \lambda^{-1}, \qquad \lambda > 0, \tag{14}$$

then A is the generator of a unique contraction semigroup that will be denoted by $t \to \exp tA$. ♢

Proof Set $A_\lambda \equiv -\lambda I - \lambda^2(A - \lambda I)^{-1}$, $\lambda \in R^+$. This collection operators is called *Yoshida approximant* of A. One verifies that, if $\phi \in D(A)$, then $\lim_{\lambda\to\infty} A_\lambda\phi = A\phi$.

The operators A_λ are bounded, commute among themselves and each of them generates a norm-continuous contraction semigroup. Indeed

$$\left\| e^{tA_\lambda}\phi \right\| = e^{-t\lambda} \left\| e^{t\lambda^2(A-\lambda I)^{-1}}\phi \right\| \leq e^{-t\lambda}e^{t\lambda}\|\phi\| = \|\phi\|$$

where we have made use of $\lambda\|(A - \lambda I)^{-1}\| \leq 1$.

Moreover for each $\phi \in D(A)$ one has $\|e^{tA_\lambda} - e^{tA_\mu}\| \leq t\|A_\lambda - A_\mu\|$ and then

$$\frac{d}{ds} e^{tsA_\lambda + (1-s)tA_\mu} = t(A_\lambda - A_\mu)e^{tsA_\lambda + (1-s)tA_\mu}. \tag{15}$$

It follows that if $\phi \in D(A)$ then $\lim_{\lambda\to\infty} e^{tA_\lambda}\phi = S_t\phi$ where S_t is a contraction semigroup. Since $D(A)$ is dense in X the family of operators S_t extends to X and is continuous in t for every $x \in X$ (since it the continuous limit of continuous functions). Moreover if $\phi \in D(A)$ then

$$S_t\phi - \phi = \lim_{\lambda\to\infty} (e^{t\,A_\lambda}\phi - \phi) = \lim_{\lambda\to\infty} \int_0^t e^{sA_\lambda} A_\lambda\phi \, ds = \int_0^t S_s A\phi \, ds \tag{16}$$

Let B be the generator of S_t. Dividing (16) by t and taking the limit $t \to 0$ one has that $D(B) \supseteq D(A)$ and B restricted to $D(A)$ coincides with A. On the other hand $D(B) = (I - B)^{-1}X$ and therefore $D(B) \subseteq D(A)$. Therefore $D(A) = D(B)$ and the operators A and B coincide. ♡

We point out the following simple application of the Hille-Yosida theorem. Denote with $BU(0, \infty)$ the family of bounded functions which are uniformly continuous on $(0, \infty)$ with respect to the sup-norm. Set $(T_t f)(s) \equiv f(t + s)$, $t \geq 0$.

One verifies that T_t is a contraction semigroup and that its generator A exists and satisfies $Af = f'$. For every value of $\lambda \in C$ the equation $(\lambda I - A)\phi = 0$ has as unique solution $\phi_\lambda(t) = e^{\lambda t}$. Therefore if Re $\lambda > 0$ the operator $(A - \lambda I)$, $\lambda > 0$ is well defined from X to $D(A)$.

On the other hand, for any differentiable function f in $BU(0, \infty)$ the Laplace transform gives $\lambda\|f\| \leq \|(\frac{d}{dt} - \lambda)f\|$ and therefore if $\phi = (\frac{d}{dt} - \lambda)f$ then $\|(a - \lambda I)^{-1}\phi\| \leq \lambda^{-1}\|\phi\|$. It follows from the Hille-Yosida theorem that the semigroup has generator A.

We now give a *useful characterization* of the generators of contraction semigroups. This characterization is given by the Lumer-Philips theorem; the property described corresponds to the maximum principle in the case of elliptic operators.

We assume still that the range of $(A - \lambda I), \lambda > 0$ is the entire space X but now substitute the condition $|A - \lambda I|^{-1} \leq \lambda^{-1}$ with a linear property that is easier to verify. We shall use the following corollary of the Hahn-Banach theorem (see "Lecture 8: Properties of Free Motion, Anholonomy, Geometric Phase").

Corollary of the Hahn-Banach Theorem *Let X be a Banach space and $x \in X$. There exists $l_x \in X^*$ such that*

$$\|l_x\| = \|x\|, \quad l_x(x) = \|x\|^2. \tag{17}$$

◇

Notice that the element l_x with this property is in general not unique. We shall call *face associated to x* (denoted it by F_x) the collection of those elements l_x of X^* which satisfy (17). As an example let X be the continuous functions on a compact $\mathcal{K}$. In this case the face F_f is the collection of Borel measures concentrated on the maxima of f.

Definition 5 (*dissipation*) An operator Δ on a Banach space X is called *dissipative* (or equivalently *accretive* or *monotone*) if

$$\forall x \in D(\Delta), \quad \exists \eta \in X^*, \quad \eta(x) = \|\eta\|\|x\|, \qquad \mathrm{Re}(\eta(\Delta(x))) \leq 0 \tag{18}$$

(i.e. (18) holds for each element of a face). ◇

We shall prove later that if (18) holds for an element of a face, then it holds for any other element of that face.

The notation *dissipative* originates form the fact that in the commutative setting the operator $\frac{d^2}{dx^2}$, which is the generator of the heat semigroup, is dissipative (the process dissipates energy).

Example The following example may be useful to better appreciate the meaning and the role of the definitions we will give in the general case.

Let $f \in C_0^2(R)$ and X be the Banach space of continuous functions that vanish at infinity, with norm $\|f\| = \sup_{x \in R} |f(x)|$. Let $x_0 \in R$ be a maximum point of f. Then

$$\left.\frac{d^2 f}{dx^2}\right|_{x=x_0} \leq 0. \tag{19}$$

Let δ_{x_0} be the functional defined by $\delta_{x_0}(g) \equiv g(x_0)$. Then (19) reads $\delta_{x_0}(\frac{d^2 f}{dx^2}) \leq 0$. By the maximum principle, (19) is satisfied if $\delta_{x_0}(f) = \sup_{x \in R} f(x)$. In particular if $f \geq 0$ one has $\|f\| = \sup_{x \in R} f(x)$ and therefore $\delta_{x_0}(f) = \|f\|$. ♣

A useful relation between contraction semigroups and dissipative operators is given by

Lemma 3 *If T_t is a contraction semigroup, then its generator a defined by*

$$A\,x = \lim_{t\to 0} t^{-1}(T_t x - x), \qquad x \in D(A)$$

is a dissipative operator. ◇

Proof If $\eta(x) = \|\eta\|\|x\|$ one has

$$\eta(Ax) = \lim_{t\to 0} t^{-1}\left(\eta(T_t x) - \|\eta\|\|x\|\right).$$

Since $|\eta(T_t x)| \leq \|\eta\|\|x\|$, by taking the limit $t \to 0$ one has $\mathrm{Re}(\eta(Ax)) \leq 0$. ♡

One has moreover

Theorem 4 *The operator A on the Banach space X is dissipative if and only if*

$$\|(A - \lambda I)x\| \geq \lambda\|x\|, \quad \forall\lambda > 0, \quad \forall x \in D(A).$$ ◇

Proof ($\Rightarrow$) If $\eta \in F_x$, $\mathrm{Re}\ \eta(Ax) \leq 0$, one has

$$\|\eta\|\,\|(\lambda I - A)x\| \geq \mathrm{Re}\ \eta((\lambda I - A)x) \geq \lambda\|x\|^2.$$

If $\eta \in F_x$ then $\|\eta\| = \|x\|$. Therefore $\|(\lambda I - A)x\| \geq \|x\|$.

($\Leftarrow$) Let $\|(\lambda I - A)x\| \geq \|x\|, \forall\lambda > 0, \forall x \in D(A)$. For $x \in D(A)$ let

$$\eta_\lambda \in F_{(\lambda I - A)x}, \quad \xi_\lambda = \eta_\lambda\|\eta_\lambda\|^{-1}$$

Then

$$\lambda\|x\| \leq \|(\lambda I - A)x\| = \frac{\eta_\lambda((\lambda I - A)x)}{\|\eta_\lambda\|} = \lambda\mathrm{Re}\ (\xi_\lambda(x)) - \mathrm{Re}(\xi_\lambda(Ax)). \quad (20)$$

By assumption $\mathrm{Re}\ \xi_\lambda(x) \leq \|x\|$ and therefore $\mathrm{Re}(\xi_\lambda(Ax)) \leq 0$. Dividing (20) by λ and taking the limit $\lambda \to \infty$, one has

$$\mathrm{Re}\ \xi(x) = \|x\| \quad (21)$$

where ξ is the limit point of a sequence ξ_λ (the Banach-Alaoglu theorem assures compactness and therefore existence of a limit). Because ξ is a weak limit and one has $\|\xi_\lambda\| = 1$ it follows that $\|\xi\| \leq 1$. From (21) it follows then

$$\|\xi\| = 1, \qquad \xi(x) = \|x\|. \quad (22)$$

If $\|\xi\| = 1$ the convergence ξ_λ is in the strong sense. Therefore ξ is the unique limit point. Equations (20) and (22) achieve the proof of the implication ($\Leftarrow$). ♡

Definition 6 (*maximal monotone*) An operator a on the Banach space X is *maximal monotone* if it is monotone, $(\lambda I - A)^{-1}$ is bounded in X if $\lambda > 0$ and, for some $\lambda > 0$ (and therefore for all $\lambda > 0$) the range of $\lambda I - A$ is X. ◇

A dissipative operator *is maximal monotone if and only if it is the generator of a contraction semigroup*. This is the content of next theorem of Lumer and Philips.

Theorem 5 (*Lumer-Philips*) *Let A be an operator on the Banach space X.*

1. *If A is dissipative and there exists $\lambda > 0$ such that* $\mathrm{Ran}(\lambda I - A) = X$, *then the operator A generates a contraction semigroup.*
2. *Conversely, if A is the generator of a contraction semigroup, then A is dissipative and* $\mathrm{Ran}(\lambda I - A) = X$, $\forall \lambda > 0$. *Moreover for every $\eta \in F_x$,* $\mathrm{Re}\ \eta(Ax) \leq 0$.

◇

Proof

1. By assumption $(\lambda I - A)^{-1}$ is closed and bounded, therefore A is closed. In order to use the Hille-Yosida theorem it suffices to prove

$$\mathrm{Ran}(\lambda I - A) = X, \qquad \forall \lambda \in R^+. \tag{23}$$

   Indeed if this is the case, setting $(A - \lambda I)x = f$, one has $\|f\| \geq \lambda \|(A - \lambda I)^{-1} f\|, \forall f \in X$. To prove (23) remark that $\{\lambda \in R^+, \mathrm{Ran}(\lambda I - A) = X\}$ is open (as intersection of the resolvent set with R^+) and closed because A is a closed operator. Since it is non empty it coincides with R^+.
2. From the Hille-Yosida theorem $\mathrm{Ran}(\lambda I - A) = X, \forall \lambda \in R^+$. Moreover

$$\eta(T_t x) \leq \|T_t x\| \|\eta\| \leq \|x\|^2$$

   $\forall \eta \in F_x$. It follows that $\mathrm{Re}(\eta(T_t x - x)) \leq 0$. Dividing by t and then taking the limit $t \to 0$ one verifies that the operator is maximal monotone. ♡

A special, important case of the Lumer-Philips theorem occurs when also the dual space X^* is a Banach space and also the operator A^* is dissipative.

Recall that the adjoint is defined by $D(A^*) = \{f \in X^*, f(Ax)$ is continuous in $X\}$ and if $f \in D(A^*)$ then $A^* f = f(Ax)$. Recall also that A^* is closed in the topology of X^*.

A dissipative operator cannot have positive eigenvalues. Indeed, if $Ax = \lambda x, \lambda > 0$ choose $\eta \in F_x$ so that $\eta(x) = |x|$. By assumption $\eta(x) = \lambda^{-1} \eta(Ax)$ therefore $\eta(Ax) = \lambda \|x\|$, a contradiction with $\mathrm{Re}\ \eta(Ax) \leq 0$.

Proposition 6 (corollary of Lumer and Philips theorem) *If A is closed and densely defined on the Banach space X, if X^* is a Banach space and A and A^* are both dissipative, then A is the generator of a continuous contraction semigroup.* ◇

Proof Under these assumptions $\mathrm{Ran}(I - A) = X$. In fact, if this were not true, since X^* separates X, there would exist $f \in X^*$ that has $\mathrm{Ran}(I - A)$ is in its kernel. Therefore $f(x) - f(Ax) = 0, \forall x \in D(A)$. From the density of $D(A)$ follows $f = A^* f$ and f is an eigenvalue of A^* to the eigenvalue 1; this contradicts the assumption that A be dissipative. ♡

This corollary is particularly useful if X is reflexive i.e. $(X^*)^* = X$ (as is the case if $X = L^p(R^n)$ $X^* = L^q(R^n), 0 < p, q < \infty, \frac{1}{p} + \frac{1}{q} = 1$) and if the operator A is defined on a dense domain (e.g. $C_0^1(R^n)$) and is symmetric under the duality between X and X^*:

$$(f, Ag) = (Af, g), \quad g \in C_0^1 \cap L^p, \quad f \in C_0^1 \cap L^q. \tag{24}$$

A special case occurs when $p = q = 2$. If (24) holds, then $A = A^*$.

Remark that a dissipative operator is closable. Indeed, suppose that A is dissipative. From the Hille-Yoshida theorem one has, setting $\mu \equiv -\lambda^{-1}$,

$$\|(I + \mu A)x\| \geq \|x\|, \qquad \forall x \in D(A), \quad \mu > 0.$$

Let $x_n \to 0$, $Ax_n \to y$ (norm convergence). Then

$$\left\|(I + \mu A)(x + \mu x')\right\| \geq \left\|xn + \mu x'\right\|, \qquad \forall x' \in D(A).$$

Taking the limit $n \to \infty$ one has $\|\mu y + \mu x' + \mu^2 A x'\| \geq \mu \|x'\|$. Dividing by μ and then taking the limit $\mu \to 0$

$$\left\|y + x'\right\| \geq \|x'\|, \quad \forall x' \in D(A). \tag{25}$$

Since $D(A)$ is dense, $y = 0$ follows from (25).

Nelson's theorem on derivations that we have discussed in "Lecture 9: Elements of C^*-algebras, GNS Representation, Automorphisms and Dynamical Systems" extends to dissipations. We state here the extension without proof.

Theorem 7 (Nelson) *Let A be closed and densely defined. The following statements are equivalent:*

1. *A is the generator of a one-parameter semigroup of contractions*
2. *A, A^* are both dissipative and have a common dense set of analytic vectors.*

◇

In the application to problems taken from Physics dissipations are often obtained as limit of bounded operators. In these cases the following theorem and its corollary are useful. We state them here without proof.

Theorem 8 [6] *Let A_n be a sequence of generators of contraction semigroups on a Banach space X. Denote by Γ_n^λ the graph of $I - \lambda A_n$, $\lambda > 0$, and denote with Γ^λ the limit graph. The following statements are equivalent:*

1. *There exists a strongly continuous contraction semigroup T_t such that*

$$\lim_{n\to\infty} \left\| e^{tA_n}x - T_t x \right\| = 0, \qquad \forall x \in X, \quad t \in R^+;$$

2. *For some value $\lambda > 0$ the domain and range of $I - \lambda A_n$ are dense in X.*

If these conditions are satisfied, the limit Γ^λ is the graph of $I - \lambda A$ where A is the generator of T_t. ◇

Corollary of Theorem 8 *Let A_n be generators of contraction semigroups. Assume that there exists a domain D dense in X and a limit operator A such that*

$$\lim_{n\to\infty} A_n x = Ax, \qquad x \in D(A) \tag{26}$$

If the range of $I + \lambda A$ is dense in X for at least one value of $\lambda > 0$, then A is the generator of a contraction semigroup and

$$e^{tA}x = \lim_{n\to\infty} e^{tA_n}x, \qquad \forall x \in X, \quad \forall t \geq 0,$$

where the convergence is uniform over compact set in t. ◇

The following theorem is used frequently.

Theorem 9 *Let A be dissipative and let* $\mathrm{Ran}(I - A) = X$. *If X is reflexive, then $D(A)$ is dense in X.* ◇

Proof It is sufficient to prove that if $f \in X^*$, $f(x) = 0\ \forall x \in D(A)$ then $f = 0$. As in the proof of the Lumer-Philips theorem one verifies that under the assumptions of the theorem $\mathrm{Ran}(\lambda I - A) = X$. Therefore for every $n \in N$ and $x \in D(A)$, there exists x_n defined by

$$x = x_n - \frac{1}{n}Ax_n \tag{27}$$

such that $x_n \in D(A^2)$. Multiplying by A one has

$$Ax = Ax_n - \frac{1}{n}A^2 x_n \to Ax_n = \left(I - \frac{1}{n}A\right)^{-1} Ax.$$

Since A is dissipative $\|Ax_n\| \leq \|Ax\|$. Since X is reflexive the balls in X are compact in the topology of X^* and therefore there exists a converging sequence, still denoted with $\{Ax_n\}$. Let y be the limit point. By assumption $f(x) = 0, \forall x \in D(A)$ and then $f(Ax) = nf(x_n - x) = 0$. Taking the limit $n \to \infty$ one concludes that $f(y) = 0$ since f is continuous. Since A is a closed operator $y = Ax$. Therefore $f(Ax) = 0$ and hence $f((I-A)x) = 0, \forall x \in D(A)$. By assumption $\mathrm{Ran}(I-A) = X$. Therefore $f(z) = 0, \forall z \in X$. ♡

3 Markov Approximation in Quantum Mechanics

In Quantum Mechanics an open system $\mathcal{S}_1$ can be regarded as a subsystem of a larger closed system $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$. Given a dynamics on $\mathcal{S}$ the dynamics of $\mathcal{S}_1$ can be seen as the trace on $\mathcal{S}_1$ of the dynamics in $\mathcal{S}$.

In general this dynamics in $\mathcal{S}_1$ *is not described by a semigroup* It is generically described by an integral equation with delay and the solution $u(t)$ at time t depends on the entire history $u(s)$, $0 \leq s < t$.

The evolution of the open quantum system $\mathcal{S}_1$ can have a semigroup description in some limit situations, that we shall call *weak coupling limit*. We shall outline now this approximation; for further details and exemplifications see [2–4].

Consider a quantum system S_1 which interacts with a quantum system S_2. Assume that S_1 has a finite number of degrees of freedom and is described in a finite dimensional Hilbert space $\mathcal{H}_1$. The system S_2 is described in a Hilbert space $\mathcal{H}_2$ which may be infinite dimensional.

Denote by P_1 the orthogonal projection on $\mathcal{H}_1$ of the total Hilbert space $\mathcal{H} \equiv \mathcal{H}_1 \otimes \mathcal{H}_2$. Let $P_2 = I - P_1$. Suppose the the evolution of the uncoupled systems is represented by a continuous one-parameter group of unitary operators U_t which leave $\mathcal{H}_1, \mathcal{H}_2$ separately invariant. Denote by K the generator of this unitary group.

Suppose that the interaction between the two systems is described by a bounded operator A. Define $A_{i,j} \equiv P_i A P_j$, $i, j = 1, 2$ and suppose that $A_{1,1} = 0$ and $A_{2,2}$ is selfadjoint. Denote by V_t^λ the unitary group generated by $K + \lambda A$ and by U_t^λ the one parameter group with generator $K + \lambda A_{2,2}$. Notice that U_t^λ it leaves $\mathcal{H}_1$ invariant. Duhamel's formula gives

$$V_t^\lambda = U_t^\lambda + \lambda \int_0^t U_{t-s}^\lambda (A_{1,2} + A_{2,1}) V_s^\lambda ds \tag{28}$$

Iterating (28) once and setting $W_t^\lambda \equiv P_1 V_t^\lambda P_1$, one obtains the following equation in $\mathcal{B}(\mathcal{H}_1)$ (note that we have assumed $A_{1,1} = 0$)

$$W_t^\lambda = U_t P_1 + \lambda^2 \int_{s=0}^t \int_{s'=0}^s U_{t-s}^\lambda A_{1,2} U_{s-s'}^\lambda A_{2,1} W_{s'}^\lambda ds ds'. \tag{29}$$

The evolution operator W_t^λ represents the trace left on on S_1 of the evolution of the entire system. The double integral in (29) is a sign of memory effects and it prevents W_t^λ to be a semigroup. If we assume that λ is very small, one can expect that the integral is bounded for any t of order λ^{-2} by Gronwall's lemma.

Define

$$Y_t^\lambda \equiv U_t^* W_t^\lambda \tag{30}$$

Introducing a new integration variable $\tau = \lambda^{-2}s$ equation (29) can be rewritten as

$$Y_t^\lambda = P_1 + \int_0^t H(\lambda, t, \tau) Y_\tau^\lambda d\tau \tag{31}$$

where

$$H(\lambda, t, \tau) = U^*_{\lambda^{-2}\tau} K(\lambda, t) U_{\lambda^{-2}\tau}, \quad K(\lambda, t) \equiv \int_0^{\lambda^{-2}t} U_\tau^* A_{1,2} U_\tau A_{2,1} d\tau \in \mathcal{B}(\mathcal{H}_1)$$

The integral in (32) does not converge in general when $\lambda \to 0$. If $\mathcal{H}_1$ is finite-dimensional (or more generally if $P_1 U_t$ has discrete spectrum of finite multiplicity), we can try to replicate the adiabatic approach.

If $\{Q_m\}$ are the spectral projection of K and α_m its eigenvalues ($\alpha_m \neq \alpha_n$ if $n \neq m$), one has

$$P_1 U_t = \sum_\alpha Q_\alpha e^{i\omega_\alpha t}. \tag{32}$$

For every bounded operator $B \in \mathcal{B}(\mathcal{H})$ define the *Cesaro mean*

$$\tilde{B} = \lim_{t\to\infty} \frac{1}{2t} \int_{-t}^t U^0_{-s} B U^0_s ds \tag{33}$$

where we have denoted by U^0_s the restriction of U_s to $\mathcal{H}_1$.

Theorem 10 *Let $\mathcal{H}_1$ be finite-dimensional and suppose that there is a constant c_{τ_1} such that $\|K(\lambda, \tau)\| \leq c_{\tau_1}$ for $|\lambda| \leq 1$ and $0 \leq \tau \leq \tau_1$. If there exists $K \in \mathcal{H}_1$ such that for $0 \leq \tau_0 \leq \tau_1 < \infty$ one has*

$$\text{s}-\lim_{\lambda\to 0} (K(\lambda, \tau) - K) = 0 \tag{34}$$

uniformly in $0 \leq \tau \leq \tau_1$, then for every $\phi \in \mathcal{H}_1$

$$\lim_{\lambda\to 0} \left\| (Y_\tau^\lambda - e^{iK\tau})\phi \right\|_2 = 0 \tag{35}$$

and moreover as bounded operators on $\mathcal{H}_1$ one has $K\ U_t = U_t\ K$ uniformly in $0 \leq t \leq \tau_1$. ◇

Proof Denote by Ξ the Banach space of continuous function with values in $\mathcal{B}(\mathcal{H}_1)$. Define the maps $\Xi \to \Xi$

$$(H_\lambda f)(\tau) \equiv \int_0^\tau H(\lambda, \tau - \sigma, \tau) f(\sigma) d\sigma,$$
$$(\tilde{H}_\lambda f)(\tau) \equiv \int_0^\tau U^*_{\lambda^{-2}\sigma} K U_{\lambda^{-2}\sigma} f(\tau) d\sigma = \sum_{\alpha,\beta} Q_\alpha Q_\beta \int_0^\tau e^{i(\omega_\alpha - \omega_\beta)\lambda^{-2}\sigma} f(\sigma) d\sigma. \tag{36}$$

with Q_α defined in (32).

Since $\alpha_m \neq \alpha_n$ if $m \neq n$ one has, uniformly in $0 \leq \tau \leq \tau_1$,

$$\lim_{\lambda \to 0} \tilde{H}_\lambda = \int_0^\tau \tilde{K} f(\tau) d\tau. \tag{37}$$

Therefore the function $H_\lambda(f)$ converges strongly to

$$Kf(\tau) = \int_0^t \tilde{K} f(\sigma) d\sigma. \tag{38}$$

Recall that for any $A \in \mathcal{B}(\mathcal{H}_1)$ the function $f_\lambda \equiv Y_\tau^\lambda$ solves

$$f_\lambda = g + H_\lambda f_\lambda, \qquad g(\tau) = A, \quad 0 \leq \tau \leq \tau_1$$

and $f_t = e^{Kt} A$ is a solution of the equation

$$f = g + Kf, \qquad g(\tau) = A, \quad 0 \leq \tau \leq \tau_1 \tag{39}$$

By an iteration of the above Volterra integral equations, one obtains for any $\phi \in \mathcal{H}_1$

$$\|(f_\lambda - f)\phi\|_2 \leq \sum_0^\infty \left\|(K_\lambda^n - K^n) g\phi\right\|_2 \leq 2 \|g\phi\|_2 \, c^n \frac{\tau_1^n}{n!} \tag{40}$$

and therefore f_λ converges strongly to f as $\lambda \to 0$. ♡

We now give a simple condition for the existence of the limit operator K.

Lemma 11 *If the perturbation satisfies* $A_{2,2} = 0$ *and*

$$\int_0^\infty \|A_{1,2} \, U_t \, A_{2,1}\| dt < \infty, \tag{41}$$

then the conditions for the validity of Theorem 10 are satisfied with

$$K = \int_0^\infty U_{-t} \, A_{1,2} \, U_t \, A_{2,1} dt. \tag{42}$$

◇

Proof The simple proof relies on the fact that under these assumptions the operator U_τ^λ is independent of λ. ♡

In the case $A_{1,1} \equiv B \neq 0$ we introduce *temporal ordered products*

$$B_n = U_{t_n}^* \, B \, U_{t_n} \tag{43}$$

The following theorem introduces assumptions that are often satisfied by physical systems.

Theorem 12 *Suppose that*

$$\int \| P_1 \ B_0 \ B_t \ P_1 \| dt < \infty. \tag{44}$$

Let

$$b_n(t) \equiv \int_0^t \cdots \int_0^{t_{n-1}} P_1 B_0 P_2 B_1 \cdots P_2 B_n P_2 B P_1 dt_n \cdots dt_1 \tag{45}$$

and assume that for $n \geq 1$ *the following estimate holds*

$$\| b_n(t) \| \leq c_n \ t^{\frac{n}{2}} \tag{46}$$

where the constants c_n *are such that the series* $\sum_0^\infty c_n \ z^n$ *has infinite radius of convergence. Assume that for* some $\epsilon > 0$ *and* some *sequence* $\{d_n\}$, $|d_n| < D, \quad \forall n$, *one has for all* $t \geq 0$

$$\| b_n(t) \| \leq d_n \ t^{\frac{n}{2} - \epsilon}. \tag{47}$$

Then the conditions of Theorem 10 are satisfied if one chooses

$$K = \int_0^\infty P_1 \ A_0 \ A_t \ P_1 \ dt. \tag{48}$$

◇

Proof The expansion in power series in λ of U_t^λ gives

$$\| K(\lambda, \tau) - K \| \leq \int_{\lambda^{-2}\tau} \| P_0 \ B_0 \ B_t \ P_0 \| dt + \sum_1^\infty \lambda^n \ \| a_n(t) \| \tag{49}$$

From (45) one proves that when $\lambda \to 0$ the integral converges to zero uniformly in τ in any open subset of R^+. Due to (46) the series is dominated when $0 \leq \tau \leq \tau_0$ by the convergent series $\sum_1^\infty c_n \ \tau_0^n$.

From (49) one sees that the nth term of the series is bounded by $\lambda^{2\epsilon} d_n \tau_0^{\frac{n}{2}}$ and converges therefore to zero when $\lambda \to 0$. Therefore the series converges to zero uniformly in $0 \leq \tau \leq \tau_0$. ♡

Remark that in general, without further assumptions, the convergence is not uniform in time. Therefore *without further assumptions* one cannot exchange the limits $\lambda \to 0$ and $t \to \infty$.

4 Quantum Dynamical Semigroups I

We have seen that in Quantum Mechanics the description of the evolution of an open system by a semigroup requires strong assumptions on the interaction with the environment and even in this case it holds only asymptotically in time.

An alternative "axiomatic" approach consists in *assuming* that action of the environment modifies the dynamics of the system by adding a bounded linear map on its state space. This leads to the theory of *quantum dynamical semigroups* [4].

Let H be an (unbounded) self-adjoint operator on the Hilbert space $\mathcal{H}$. For any trace class positive matrix ρ the formula

$$U_t(\rho) = e^{-itH}\rho e^{itH} \tag{50}$$

defines a strongly continuous group of isometries on the space of states. We denote by $-i\ ad_H$ the infinitesimal generator. The domain of the operator ad_H is the set of all trace class operators ρ on $\mathcal{H}$ such that

$$\rho D(H) \subset D(H) \tag{51}$$

and such that the operator $H\rho - \rho H$ is norm bounded and can be extended to a trace class operator $\sigma \in \mathcal{H}$.

Theorem 13 *Let H be an unbounded self-adjoint operator on $\mathcal{H}$ and Ξ a bounded positive linear map on the state space Σ. Then the operator Z defined by*

$$Z(\rho) = -ad_H(\rho) + \Xi(\rho) - \frac{1}{2}[\Xi^*(1)\rho + \rho\Xi(1)] \tag{52}$$

with domain equal to that of ad_H, is a generator of a semigroup on Σ. We call this semigroup a dynamical semigroup. ◇

Proof Since Ξ is a bounded perturbation of ad_H it is the generator of a semigroup T_t on Σ. If $\rho \in D(Z)$ then $T_t\rho \in D(Z)$ for all $t \geq 0$ and

$$\frac{d}{dt}tr(T_t\rho) = tr[Z(T_t\rho)] = 0 \qquad \Rightarrow tr(T_t\rho) = tr\rho \tag{53}$$

This relation holds then for all $\rho \in \Sigma$ since $D(Z)$ is dense in Σ.

Define W by

$$W(\rho) = -iad_H(\rho) - \frac{1}{2}[\Xi^*(I)\rho + \rho\Xi^*(I)] \tag{54}$$

Notice that by definition $\Xi^*(I)\rho = \rho$. It is easy to see that W is the generator of a semigroup S_t on Σ defined by

$$S_t(\sigma) = B_t\sigma B_t^* \tag{55}$$

where B_t is the one-parameter contraction semigroup on $\mathcal{H}$ with generator $-H - \frac{1}{2}\Xi^*(I)$.

The operator Z has the structure $Z = W + \Xi$ where Ξ is bounded, so that

$$T_t = lim_{n\to\infty}(S_{\frac{t}{n}}(exp(\Xi_{\frac{t}{n}}))^n\rho \tag{56}$$

is defined for all positive ρ (Trotter-Kato formula). The limit is taken in the trace norm. ♡

Since both S_t and $e^{t\Xi}$ are positive, also T_t is positive for all t.

We shall return in the next Lecture to the theory of Quantum Dynamical Semigroups. We discuss here briefly a converse procedure, the dilation.

5 Dilation of Contraction Semigroups

We have seen that the conditioning of a unitary group to a subalgebra does not lead in general to a semigroup; a semigroup is obtained under suitable assumptions on the interaction and only in an adiabatic limit. It is instead *always possible* the inverse process, *the unitary dilation*, from a one-parameter semigroup acting on the representation of a C^*-algebra $\pi(\mathcal{A})$ on a Hilbert space $\mathcal{H}$ to a one-parameter group acting on a representation of the same algebra on a larger Hilbert space $\mathcal{K}$. Of course one can always regard $\pi(\mathcal{A})$ as a subalgebra of an algebra $\mathcal{B}$ represented in $\mathcal{K}$.

We need first a definition

Definition 7 (*positive definite*) Let G be a group, $\mathcal{H}$ a Hilbert space and $T\colon\ G \to \mathcal{L}(H)$ an operator-valued function on G. W say that T is *positive definite* if for all N and all $g_1, \dots g_n \in G$ and all $\phi_1, \dots \phi_n \in \mathcal{H}$ one has

$$\sum_{m,n=1}^{N} (T(g_m^{-1}g_n)\phi_n, \phi_m) \geq 0 \tag{57}$$

◇

By taking $N = 1, N = 2$ this implies $T(e) \geq 0,\ T(g^{-1}) = T(g)^*$.

Theorem 14 [11] *Let U_g be a unitary representation of the group G on a Hilbert space $\mathcal{K}$ and P is an orthogonal projection of $\mathcal{K}$ onto a subspace $\mathcal{H}$. The restriction T_g of PU_g to $\mathcal{H}$ is a positive definite $\mathcal{L}$-valued function on G with $T_e = I$. If U_g is strongly continuous, also T_g is strongly continuous.*

Conversely given a positive definite $\mathcal{L}(\mathcal{H})$-valued function T_g on G with $T_e = I$ then there exists a unitary representation U_g of G on a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ such that if P is the projection of $\mathcal{K}$ onto $\mathcal{H}$ the T_g is the restriction of of PU_g to $\mathcal{H}$.

The representation is unique up to isomorphism if

$$\mathcal{K} = \{U_g \mathcal{H} \, ; \, g \in G\}^- \tag{58}$$

In this case it is called minimal unitary dilation *of T_g. If T_g is weakly continuous function of g, the the representation U_g can be chosen to be strongly continuous.* $\diamond$

Proof This is a standard proof in the theory of unitary representations of continuous groups.

To prove the first part, given U_g and P let $g_1 \dots g_n \in G$ and $\phi_1 \dots \phi_n \in \mathcal{H}$. Then T_g is positive definite because

$$\sum_{m,n=1}^{N} (T_{g_m^{-1} g_n} \phi_n, \phi_n) = \sum_{m,n=1}^{N} (U(g_n)\phi_n, U(g_m)\phi_m) = \| \sum_{n=1}^{N} U(g_n)\phi_n \|^2 \geq 0 \tag{59}$$

To prove the converse, given the positive definite operator valued function T_g let $\mathcal{M}$ be the vector space of all function $f \, : \, G \to \mathcal{H}$ of finite support (i.e. functions that are zero except for a finite set). The formula

$$(f_1, \, f_2) =\equiv \sum_{f,g \in G} (T_{g^{-1}h} f(h), \, f_2(g)) \tag{60}$$

defines a structure of pre-Hilbert product. If $\mathcal{N}$ is the subset for which $(f, \, f) = 0$ let $\mathcal{K}$ be the Hilbert space completion of $\mathcal{M}/\mathcal{N}$. Let J be the isometric embedding of $\mathcal{H}$ into $\mathcal{K}$ obtained by associating for each $\phi \in \mathcal{H}$ the function $f \in \mathcal{M}$ defined by $f(g) = 0$ if $g \neq e$ and $f(e) = \phi$.

If $g \in G$ and $f \in \mathcal{M}$ define

$$(V_g f)(x) \equiv f(g^{-1} x) \tag{61}$$

Since $V_g \, : \, \mathcal{M} \to \mathcal{M}$ is isometric and linear it induces a unitary map U_g on $\mathcal{M}/\mathcal{N}$ (and therefore on $\mathcal{K}$) which is a representation of G. One has, for $\phi, \, \psi \in \mathcal{H}$

$$P U_g(J\phi, (J\psi)) = (V_g \delta_e \phi, \delta_e \psi)(\delta_g \phi, \delta_e \psi) = (T_g \phi, \psi) \tag{62}$$

Therefore T_g is the restriction of $P U_g$ to $\mathcal{H}$.

Since $\mathcal{M}$ is generated by $\{\delta_g \phi, \; g \in G, \; \phi \in \mathcal{H}\}$ it follows that $\mathcal{K}$ is generated by $\{U_g(J\phi), \; g \in G, \; \phi \in \mathcal{H}\}$ and (58) holds for the space generated as above.

To prove uniqueness, let $\{\mathcal{K}', U'_g\}$ is another unitary dilation which satisfies (58). Define

$$Sf \equiv \sum_{g \in G} U'_g f(g) \tag{63}$$

Then

$$\|Sf\|^2 = \sum_{g,h\in G} (U'_g f(g), U'_h f(h)) = \sum_{g,h\in G} (T_{h^{-1}g} f(g), f(h)) = \|f\|^2 \quad (64)$$

and therefore S induces an isometric linear map $\mathcal{M}/\mathcal{N}$ and then a unitary map, denoted again by $S : \mathcal{K} \to \mathcal{K}'$. Moreover if $\phi \in \mathcal{H}$ one has $S(U_g)J\phi = U'_g S(J\phi)$ and therefore the two representations are unitary equivalent.

We finally prove that if T_g is a weakly continuous function of g then the representation U_g on $\mathcal{H}$ is weakly continuous (and therefore strongly continuous because it is unitary). One has

$$(V_g f_1, f_2) = \sum_{g_1,g_2\in G} (T_{g_2^{-1}g_1} f_1(g_1), f_2(g_2)) \quad (65)$$

and therefore t is continuous function of g for all functions f_1, f_2 in the dense subspace $\mathcal{M}/\mathcal{N}$ of $\mathcal{K}$ and therefore for all $f_1, f_2 \in \mathcal{K}$. ♡

We want to apply Theorem 14 to find a unitary one-parameter group of dilations of a contraction semigroup on a Hilbert space. In the next Lecture we attack the same problem in the setting of C^*-algebras. In order to apply the theorem we have to verify the technical hypothesis of the theorem.

We state the verification in a sequence of two theorems.

Theorem 15 *Let C be a contraction on a Hilbert space $\mathcal{H}$ and define for all $n \in Z$*

$$C(-n)^* = C(n) = C^n \quad (66)$$

then $C(n)$ is a positive definite operator-valued function on Z. ◇

Proof Let $\mathcal{K} \equiv l^2(Z^+, \mathcal{H})$ and B be any operator on $\mathcal{H}$ such that

$$C^*C + B^*B = I \quad (67)$$

Let $D : \mathcal{K} \to \mathcal{K}$ be the isometric linear operator defined by

$$D\phi_0 = C\phi_0, \quad D\phi_1 = B\phi_0 \quad D\phi_n = \phi_{n-1}, \ n \leq 2 \quad (68)$$

If $\psi : Z \to \mathcal{H}$ is a function with finite support define

$$(T\psi_m)_n = (1 - D^n)\psi_m \quad (69)$$

Then

$$\sum_{m,r\in Z} (C^{m-r}\psi_m, \psi_r) = \sum_{m,r\in Z} (D^{m-r} T\phi_m, T\phi_r) \quad (70)$$

where

$$(D^{-m})^* = D^m \tag{71}$$

Therefore it is sufficient to prove that D^m is a positive definite operator valued function on $\mathcal{K}$.

Suppose that $\psi'_n = 0$ for $m < a \in Z$. Then we can divide the sum in (70) in two parts: $m \geq r$ ad $r > m$ and then use the fact that D is an isometry. In this way one sees that the sum is equal to

$$\sum_{m,r\geq 0} (D^m \phi'_{m-a}, D^r \phi''_{r-a}) \geq 0 \tag{72}$$

♡

Theorem 16 *Let B_t be a strongly continuous one-parameter contraction semigroup on $\mathcal{H}$ and suppose $B_{-t} = B_t^*$ for all $t \geq 0$. Then $\{B_t, \}$, $t \in R$ is a positive-valued definite operator-valued function on R.* ◇

Proof Suppose $t_n \in R$ and ϕ_n are given for $1 \geq n \geq N$. Let $s_r(n)$, $r > 0$ be the integer that minimizes $|nt - s_r(n)|$. Let $C_r(s) \equiv B(\frac{s}{r})$. Therefore for all $s \in Z^+$ one has

$$C_r(-s)^* = C_r(s) = C_r(1)^s$$

Therefore by Theorem 15 for all N

$$\sum_{m,n=1}^{N} (B(t_n - t_m)\phi_n, \phi_n) = lim_{r\to\infty} \sum_{m,n=1}^{N} (C_r(s_r(n) - s_r(m))\phi_n, \phi_n) \geq 0$$

♡

Theorems 13 and 16 provide a strongly continuous group of unitary operators on the bigger space $\mathcal{K}$ that is a dilation of the strongly continuous semigroup of contraction we started with. Typically the generator of this group has spectrum that covers the entire real line. Indeed one has

Theorem 17 (Sinai) [10] *Let $U_t = e^{itH}$ be the minimal unitary dilation on $\mathcal{K}$ of some strongly continuous strict one parameter semigroup. Then there exist an auxiliary Hilbert space $\mathcal{N}$ and an isomorphism $\mathcal{K} \simeq L^2(R, \mathcal{N})$ under which U_t becomes the operator of translation by a distance t*

$$(U_t f)(\tau) = f(t + \tau)$$

As a consequence the spectrum of H is absolutely continuous and equals the real line. ◇

References

1. Brezis, H. (2010). *Functional analysis and applications*. New York: Springer Universitext.
2. Davies, E. (1974). *Communications in Mathematical Physics*, *39*, 91–130.
3. Davies, B. (1973). *Communications in Mathematical Physics*, *33*, 171–186.
4. Davies, B. (1976). *Quantum theory of open systems*. London: Academic Press.
5. Evans, L. C. (1989). Partial differential equations. *Transactions of the American Mathematical*, Providence.
6. Kuntz, T. (1970). *Transactions of the American Mathematical*, *148*, 205–240.
7. Meyer, P. (1993). *Quantum probability for probabilists*. Berlin: Springer.
8. Partasarathy, K. R. (1992). *An introduction to quantum stochastic calculus*. Basel: Birkhäuser.
9. Pazis, B. (1983). *Semigroups of linear operators*. Berlin: Springer.
10. Sinai, Y. (1964). *Translation of American Mathematical Society*, *39*, 83–110.
11. Sg-Nagy, B., & Foias, C. (1970). *Harmonic analysis of operator on Hilbert spaces*. Amsterdam: North-Holland.
12. Yosida, K. (1965). *Functional analysis*. New York: Springer.

Lecture 12: Positivity Preserving Contraction Semigroups on C^*-algebras. Conditional Expectations. Complete Dissipations

1 Complete Positivity. Dissipations

So far we have assumed that X is a Banach space. The semigroups have further properties when X is a C^*-algebra $\mathcal{A}$.

In the commutative case in the study of the contraction semigroups on $L^\infty(R^n)$ an important role is given to semigroups that *preserve positivity*. For them it is possible to define transition probabilities and stochastic processes, as we will see further in these Lectures.

Conversely if $\phi(t, \omega), \omega \in \Omega$, is the solution of a stochastic differential equation the semigroup

$$L^\infty(\Omega) \ni f \mapsto f_t, \qquad f_t(\omega) = f(\phi(-t, \omega))$$

is positivity preserving.

Positivity has also a role in the case in which that the Banach space X is a non-commutative C^*-algebra. We shall study in particular the subclass of contraction semigroups on C^*-algebra which *preserve positivity* (i.e. leave invariant *the cone of positive elements*).

Most (but not all) of the results we present can be generalized to the case in which in the C^*-algebra there exists an invariant cone V, with $V \cap (-V) = \emptyset$, $\overline{V \cup (-V)} = X$.

Theorem 1 *Let $\mathcal{A}$ be a C^*-algebra and α_t is a semigroup such that for each $a \in \mathcal{A}$, $t \in R^+$,*

$$\alpha_t(a^*) = (\alpha_t(a))^*, \quad \alpha_t(a^*a) \geq \alpha_t(a^*)\alpha_t(a). \tag{1}$$

where a is a generic element of $\mathcal{A}$. Let $\Delta \equiv lim_{t\to 0} \frac{\alpha_t(a)-a}{t}$ be the generator. *Then if $a^*a,\ a^*,\ a \in D(\Delta)$ the operator Δ satisfies*

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_12

$$\Delta(a^*a) \geq a^*\Delta(a) + \Delta(a^*), \tag{2}$$

◇

Proof From (1)

$$\alpha_t(a^*a) - a^*a \geq \alpha_t^*(a)\alpha_t(a) - a^*a = \alpha_t^*(a)(\alpha_t(a)\, a - a) + (\alpha_t^*(a) - a^*)\, a.$$

Dividing by t and taking the limit $t \to 0$ one obtains (2) under the assumption that all elements are in $D(\Delta)$. ♡

Definition 1 (*Dissipations*) If (2) holds, the operator Δ acting on $\mathcal{A}$ is called *dissipative*. ◇

Example Let $\mathcal{A} \equiv L^\infty(R)$ and $\Delta \equiv \frac{d^2}{dx^2}$. Then

$$\begin{aligned}\frac{d^2}{dx^2}(\bar{f}\ f) &= \frac{d^2\bar{f}}{dx^2}\ f + \bar{f}\ \frac{d^2 f}{dx^2} + 2\frac{d\bar{f}}{dx}\frac{df}{dx} \\ &\geq \frac{d^2\bar{f}}{dx^2}\ f + \bar{f}\ \frac{d^2 f}{dx^2}.\end{aligned} \tag{3}$$

♣

Remark that every derivation is a dissipation. It follows from (2) that the *the square of a derivation is a dissipation*. Moreover, if Δ is a dissipation, then for every integer N and for every choice of $a_1, \ldots, a_N \in \mathcal{A}$ one has

$$[[\Delta^2(a_i^* a_j)]]_N \geq [[\Delta^2(a_i^*)\ a_j + a_i^*\Delta^2 a_j]]_N \tag{4}$$

where the symbol $[[B]]_N$ denotes the $N \times N$ matrix with elements $b_{i,j}$.

The inequality holds in the C^*-algebra $\mathcal{A}\otimes\mathcal{B}_N$ where $\mathcal{B}_N$ is the algebra of complex valued matrices of rank N.

It is convenient to introduce a subclass of dissipations.

Definition 2 (*complete positivity*) A dissipation is called *N-positive* if for every choice of $A_k,\ k = 1, \ldots, N$, one has

$$[[\Delta(a_i^* a_j)]]_N \geq [[\Delta(a_i^*)\ a_j + a_i^*\ \Delta(a_j)]]_N \tag{5}$$

and is called *completely positive* if (5) holds for every integer N. ◇

In the case of groups of diffeomorphisms the corresponding conditions are

$$[[\alpha_t(a_i^* a_j)]]_N \geq [[\alpha_t(a_i^*)\alpha_t(a_j)]]_N.$$

As we have done in the case of the derivations we shall begin treating bounded dissipations.

Theorem 2 [8] *Let Δ be a closed linear map of a C^*-algebra $\mathcal{A}$ and define*

$$e^{t\Delta} \equiv \sum_0^\infty \frac{t^n}{n!}\Delta^n, \qquad t \in R^+. \tag{6}$$

The following conditions are equivalent:

(i)

$$\Delta(a^*a) \geq \Delta(a^*)\, a + a^*\, \Delta(a), \qquad a \in \mathcal{A};$$

(ii)

$$e^{t\Delta}(a^*a) \geq e^{t\Delta}(a^*)\, a + a^*\, e^{t\Delta}a, \qquad t \in R^+. \qquad \diamondsuit$$

Proof Condition (ii) follows from (i) differentiating with respect to t at the origin. To prove (i) $\rightarrow$ (ii) notice that

$$e^{t\Delta}(a^*a) - e^{t\Delta}(a^*)e^{t\Delta}(a) = \int_0^t ds\, \frac{d}{ds}\left[e^{s\Delta}\left(e^{(t-s)\Delta}(a^*)e^{(t-s)\Delta}(a)\right)\right]$$

$$= \int_0^t ds\, e^{s\Delta}\left\{\Delta\left[e^{(t-s)\Delta}(a^*)e^{(t-s)\Delta}(a)\right] - \Delta\left[e^{(t-s)\Delta}(a^*)\right]e^{(t-s)\Delta}(a)\right.$$

$$\left. - e^{(t-s)\Delta}(a^*)\Delta[e^{(t-s)\Delta}(a)]\right\} = \int_0^t e^{s\Delta}\Xi(t-s, a^*, a)ds$$

where $\Xi(t-s, a^*, a)$ is positive by assumption. Therefore it suffices to prove that the operator $e^{s\Delta}$ is positivity preserving. This is the content of the next theorem, which has an independent interest.

Theorem 3 *If Δ is bounded, point (ii) in Theorem 2 implies that $e^{t\Delta}$ is positivity preserving.* $\diamondsuit$

Proof Recall that for every state ω one has

$$a \in \mathcal{A}, \quad a \geq 0, \quad \omega(a) = 0 \Rightarrow \omega(\Delta(a)) \geq 0. \tag{7}$$

To see this, setting $a = b^2$ one has $\Delta(a) \geq \Delta(b)\, b + b\, \Delta(b)$ since Δ is a dissipation. From Schwarz inequality $|\omega(\Delta(b)b^*)|^2 \leq \omega(\Delta(b)^*\Delta(b))\omega(b^*b) \leq C\omega(a) = 0$.

If $\omega(a) \neq 0$ but there exists $y \in \mathcal{A}$ such that $a \geq 0, ay = 0$, then

$$y^*\Delta(a)y \geq 0. \tag{8}$$

Indeed by assumption $\omega(y^*ay) = 0, \forall\omega$ and therefore (7) implies $\omega(y^*\Delta(a)y) \geq 0$. From this (8) follows. $\heartsuit$

We remark now that $e^{t\Delta} = \lim_{n\to\infty}(I - \frac{t}{n}\Delta)^{-n}$. Therefore to prove that $e^{t\Delta}$ preserves positivity it suffices to prove it for $(I - \lambda\Delta)^{-1}$, when λ is sufficiently small.

It is sufficient to give the proof in a representation of $\mathcal{A}$. In what follows we choose the representation in which the elements a of $\mathcal{A}$ are bounded closed operators on the Hilbert space $\mathcal{H}$. By density, it suffices to prove that, if $a^* = a$ and $(I - \lambda\Delta)a \geq 0$, then $a \geq 0$.

Let $a = a_+ - a_-$ be the decomposition of a in its positive and negative parts. Recall that for every bounded closed symmetric operator a there exists an invertible isometric map $U : \mathcal{H} \to L^2(X, \mu)$ such that $U^{-1}\, a\, U$ is a bounded measurable function on the measure space (X, μ).

Denote by $f_\pm$ the positive (resp. negative) part of the function f. The operators satisfy $a_\pm = U^{-1} f_\pm U$. By construction $a_+ a_- = 0$. Therefore by (8) $a_-\Delta(a_+)a_- \geq 0$. Moreover

$$0 \leq a_-(I - \lambda\Delta)(a)a_- = -a_-^3 - \lambda a_-\, \Delta(a_+)\, a_- + \lambda\, a_-\, \Delta(a_-)\, a_-. \qquad (9)$$

Therefore $0 \leq -a_-^3 + \lambda\, a_-\, \Delta(a_-)\, a_-$ and hence $\|a_-\|^3 \leq \lambda\, \|a_-\|^3\, \|\Delta\|$. Choosing $\lambda < \|\Delta\|^{-1}$ one derives $a_- = 0$. This ends the proof of Theorem 3 and therefore of the implication (ii) $\to$ (i) in Theorem 2. ♡

Theorem 2 provides *sufficient* conditions for a semigroup to be positivity preserving. It is important to find *necessary and sufficient conditions*. These are given by the following important theorem which we will state without proof:

Theorem 4 (Evans and Hanche-Olsen) [8] *Let $\mathcal{A}$ be a C^*-algebra with unit e and let ω be a state. A bounded closed operator Δ is the generator of (norm-continuous) positivity preserving semigroup if and only if one of the following three* equivalent *conditions is satisfied:*

(i)
$$a \in \mathcal{A}, \quad a > 0, \quad \omega(a) = 0 \Rightarrow \omega(\Delta(a)) \geq 0;$$

(ii)
$$\Delta(a^2) - a\Delta(e) \geq \Delta(a)\, a + a\, \Delta(a),\ \forall a = a^* \in \mathcal{A};$$

(iii)
$$\Delta(e) - u^*\Delta(e)u \geq \Delta(u^*)u + u^*\Delta(u),\ \forall u \in \mathcal{A}, \quad u^*u = uu^* = e.$$

◇

A partial extension of this result to the case Δ is not bounded (and the semigroup is strongly continuous in a given representation) can be found in [3]. A very detailed analysis of order preserving semigroups on ordered Banach spaces can be found in [2].

At the beginning of this Lecture we have studied the dissipative operators on a Banach space X. We will now study the conditions under which a generator Δ is dissipative on a C^*-algebra $\mathcal{A}$. A simple result which gives a *sufficient* condition is the following theorem.

Theorem 5 *The generator Δ is dissipative if the following conditions are satisfied:*

(i) $D(\Delta)$ *is a dense subalgebra;*
(ii) $e \in D(\Delta)$*;*
(iii) $a \geq 0, a \in D(\Delta) \Rightarrow \sqrt{a} \in D(\Delta)$. ◇

Notice that if Δ is closed and bounded these conditions are satisfied. Moreover in this case every element $a \in \mathcal{A}$ is analytic and therefore Δ is maximal monotone and is the generator of a contraction semigroup.

Proof of Theorem 5 Recall that the face F_a of a is the collection of all elements l in dual of $\mathcal{A}$ such that $|l| = \|a\| \;\; l(a) = \|a\|^2$.

Let $a \in D(\Delta), \eta \in F_{a^*a}, \|\eta\| = 1$. Define $\eta_a(y) \equiv \eta(a^*y)$. Then

$$\eta_a(a) = \eta(a^*a) = \|a\|^2, \quad \frac{\eta_a}{\|\eta_a\|} \in F_a. \tag{10}$$

Since η is a state and Δ dissipative

$$2\text{Re } \eta_a(\delta(a)) = \eta(a^*\Delta(a) + \eta(\Delta(a)\, a) \leq \eta(\Delta(a^*a)). \tag{11}$$

Define

$$y \equiv \sqrt{(\|a\|^2 e - a^*a)}, \quad y^2 = \|a\|^2 e - a^*a \tag{12}$$

so that $\eta(y^2) = 0$.

Then (iii) implies $y \in D(\Delta)$ and from (12)

$$2\text{Re } \eta_a(\Delta(a)) \leq \|a\|^2 \eta(\Delta(e)) - \eta(\Delta(y^2)).$$

One has $\Delta(e) = \Delta(e^2) \geq 2e\ \Delta(e)$ and therefore $\Delta(e) \leq 0$. Moreover

$$\eta(\Delta(y^2)) \geq \eta(y\ \Delta(y)) + \eta(\Delta(y)\ y), \tag{13}$$

which follows from (13) and Schwarz inequality. Hence Re $\eta_a(\Delta(a)) \leq 0$ and therefore Δ is dissipative on $\mathcal{A}$. ♡

2 Completely Positive Semigroups. Conditional Expectations

We have seen that in the commutative case derivations and their squares are completely positive dissipations. In the commutative case every dissipation is completely positive; this is not true in the non-commutative case.

A special class of completely positive maps are the *conditional expectations* that are used to study subsystems of physical systems.

Definition 3 (*conditional expectation*) Let $\mathcal{A}$ be a C^*-algebra with unit, $\mathcal{B}$ a von Neumann algebra, $\mathcal{C}$ a sub-C^*-algebra of $\mathcal{A}$. Let $\Phi : \mathcal{A} \to \mathcal{B}$ be normalized (i.e. $\Phi(e) = e$), positive and such that the weak closure of $\Phi(\mathcal{C})$ coincides with $\mathcal{B}$.

Suppose moreover that $\mathcal{C} \to \Phi(\mathcal{C})$ be a homeomorphism. Then the map Φ is called *conditional expectation*. The map Φ is also called *projection of* $\mathcal{A}$ *on* $\mathcal{B}$ *relative to* $\mathcal{C}$. ◇

In the commutative setting the most common formulation is as follows. Let $X, \mathcal{B}, \mu$ be a measure space, $\mathcal{B}$ the σ-algebra of measurable sets, $\mathcal{B}_\infty$ a sub-σ-algebra of $\mathcal{B}$ and let f be a bounded function which is $\mathcal{B}$ measurable. The conditioning of f with respect to $\mathcal{B}_\infty$ is the map $f \to f_1$ where f_1 is the only $\mathcal{B}_\infty$-measurable function such that for every $g \in L^1(\mathcal{B}_\infty, \mu)$ one has

$$\int g f d\mu = \int g f_1 d\mu$$

Definition 4 (*N-positive maps, completely positive maps*) A linear map Φ of a C^*-algebra $\mathcal{A}$ in a C^*-algebra $\mathcal{B}$ is said to be *N-positive* ($\Phi \in P_N$) if

$$\Phi \otimes I_N : \qquad \mathcal{A} \otimes M_N \to \Phi(\mathcal{A}) \otimes M_N \tag{14}$$

is positive. It is *strongly N-positive* ($\Phi \in P_N^s$) if

$$[[\Phi(a_i^* a_j)]]_N \geq [[\Phi(a_i^*)\Phi(a_j)]]_N, \qquad \forall a_j \in \mathcal{A}, \quad j = 1, \ldots, N. \tag{15}$$

It is *completely positive* (CP) if (15) holds for every integer N. ◇

From the definitions the following inclusions are evident

$$CP \to P_N^s \to P_N \tag{16}$$

The following theorem, which we state without proof, states that these inclusion are strict unless $\mathcal{A}$ is a matrix algebra.

Theorem 6 [7] *The inclusions in (15)* are both strict *unless* $\mathcal{A} = M_N$. ◇

From Theorem 6 one derives as a corollary the following proposition.

Proposition 7 *Let* Δ *be a bounded derivation of a* C^**-algebra* $\mathcal{A}$ *and define*

$$e^{t\Delta} \equiv \sum \frac{t^n}{n!} \Delta^n, \qquad t \in R^+.$$

The following two statements are equivalent:

(i) Δ *is a completely positive dissipation;*
(ii) $e^{t\Delta}$ *is a completely positive contraction semigroup.* ◇

For a detailed study of completely positive semigroups one can consult e.g. [5, 6, 13]. Notice that in the commutative case every completely positive map is a conditional expectation. A similar result *does not hold* in the non-commutative case. The converse is however true: *every conditional expectation is a completely positive map*.

Theorem 8 [14] *Every conditional expectation* $\Phi : \mathcal{A} \to \mathcal{B}$ *is a completely positive map and for every* $a \in \mathcal{A}$ *and* $b_1, b_2 \in \mathcal{B}$ *one has*

$$\Phi(b_1 a b_2) = b_1 \Phi(a) b_2 \qquad \diamond$$

It is convenient to stress the reason why completely positive semigroups have a special role in Quantum Mechanics. Every one-parameter automorphisms group of a C^*-algebra $\mathcal{A}$ is a semigroup (if restricted to $t \in R^+$) which is positive and has a completely positive extension

$$a \otimes b \to \alpha_t(a) \otimes b, \quad b \in M_N, \tag{17}$$

which corresponds to considering N independent systems each described by the C^*-algebra $\mathcal{A}$.

One can consider equivalently a given system S in presence, *but without interactions*, of a system S' described by rank N matrices. If an interaction takes place between these two systems one can expect that at least some information on the dynamics of the systems S can be obtained by considering the evolution the entire system $S \cup S'$.

In Quantum Mechanics we do not expect complete information because, as we have seen in "Lecture 3: Axioms, States, Observables, Measurement, Difficulties", the interaction produces entanglement and for entangled states the knowledge of the whole does not imply knowledge of the parts.

To be more explicit denote by $\mathcal{A}$ the algebra of the observables of the entire system $S \cup S'$ and by α_t the one-parameter group of automorphisms that describes the dynamics of $\mathcal{A}$. Denote by $\mathcal{B}$ the sub-C^*-algebra that describes the sub-system S and denote by ι the identification of $\mathcal{B}$ as a subalgebra of $\mathcal{A}$. Let be γ be a *conditioning* $\mathcal{A} \to \mathcal{B}$ (a completely positive map which is the identity on $\mathcal{B}$).

Both γ and α_t are completely positive maps and composition preserves this property. Therefore if one measures only observables in $\mathcal{B}$ the evolution will be described at time t by the completely positive map $\mathcal{B} \to \mathcal{B}$ given by

$$T_t \equiv \gamma \alpha_t \iota, \qquad t \in R^+. \tag{18}$$

In general the maps $\{T_t\}$ *do not form a semigroup*, but, as we have seen at the end of the previous lecture, they may acquire a semigroup structure in a suitable adiabatic

limit when $\mathcal{A} = \mathcal{B} \otimes \mathcal{D}$ if the interaction is chosen suitably and the time is suitably rescaled.

It is interesting to notice that a converse is always true: every positivity preserving contraction semigroup of a C^*-algebra $\mathcal{A}$ can be written as conditioning of a one-parameter group of automorphisms of a bigger C^*-algebra $\mathcal{B}$. Of course this group of automorphism cannot in general represent the interaction between $\mathcal{A}$ and an algebra $\mathcal{C}$.

Theorem 9 (Steinspring) [15] *Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a C^*-algebra with unity and let Φ be a completely positive map $\mathcal{A} \to \mathcal{A}$.*

There exists a Hilbert space $\mathcal{K}$ and a bounded map $V : \mathcal{H} \to \mathcal{K}$ such that

$$\Phi(a) = V^* \pi_S(a) V \tag{19}$$

where π_S is a representation of $\mathcal{A}$ in $\mathcal{B}(\mathcal{K})$. ◇

This is to be compared with Theorem 14 of chapter "Lecture 11: Semigroups and Dissipations. Markov Approximation. Quantum Dynamical Semigroups I" (unitary dilations). Notice that in Theorem 9 is *bounded* but not unitary in general.

The notation $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ in Theorem 9 may seem superfluous: indeed every C^*-algebra has a faithful representation a algebra of operators in some Hilbert space. Remark however that the representation space enters in the construction of the Steinspring map and different constructions lead to different *concrete* Hilbert spaces. The representation given by Steinspring is *unique modulo partial isometries* between the "concrete" Hilbert spaces which are constructed in the dilation process.

Proof of Theorem 9 Denote by ξ_k and η_i two orthonormal bases in $\mathcal{H}$ and by $a_k,\ b_i$ two sequences of elements $\mathcal{A}$ dense in $\mathcal{A}$. Consider the space $\mathcal{A} \otimes \mathcal{H}$ endowed with the sesquilinear form

$$Q(a, \xi; b\,\eta) \equiv \langle \sum_i a_i \otimes \xi_i, \sum_k b_k \otimes \eta_k \rangle \equiv \sum_{i,k} \left(\Phi(b_k^* a_i) \xi_i, \eta_k \right) \tag{20}$$

where $a, \xi \equiv \{a_k\}, \{\xi_k\}$ and the sum runs over a *finite* set of indices. This quadratic form is positive because Φ is a completely positive map. It is easy to verify that the form Q is closable and its closure defines a scalar product.

Define a map Ψ from $\mathcal{A}$ to the linear maps on $\mathcal{A} \otimes \mathcal{H}$ through

$$\Psi(a) \sum_i b_i \otimes \eta_i = \sum_i (a\ b_i) \otimes \eta_i \tag{21}$$

For every $\xi \in \mathcal{A} \otimes \mathcal{H}$ one has $(\Psi(a)\xi, \Psi(a)\xi) \leq \|A\|^2 (\xi, \xi)$. Let $\mathcal{N}$ the kernel of the scalar product defined in (20). Call $\mathcal{K}$ the Hilbert space completion of $(\mathcal{A} \otimes \mathcal{H})/\mathcal{N}$ with respect to the scalar product (20). Define V by $V\xi \equiv I \otimes \xi + \mathcal{N}$. By construction it is an isometry of $\mathcal{H}$ in $\mathcal{K}$.

Let π_S be the representation of $\mathcal{A}$ in $\mathcal{K}$ given by

$$\pi_S(a)\left(\sum b_k \otimes \eta_k\right) = \sum ab_k \otimes \eta_k. \tag{22}$$

By direct computation one verifies $\Phi = V^*\pi_S V$. ♡

3 Steinspring Representation. Bures Distance

Stienspring's theorem can be reformulated as follows: every completely positive map T between two C^*-algebras $\mathcal{A}$, $\mathcal{A}_0$ *can be described as composition of a* **-homeomorphism (dilation)* π_S of $\mathcal{A}$ in $\mathcal{B}(\mathcal{K})$ followed by a *projection (conditioning)* to the C^*-algebra $\mathcal{A}_0$

$$T(a) = V^*\pi_S(a)V \tag{23}$$

We shall call the triple $\{\pi_S, V, \mathcal{K}\}$ *Steinspring representation*.

The Steinspring representation is unique modulo partial isometries between the space which are constructed using dilations. To see this, notice that by the construction if $\{\pi_S, V, \mathcal{K}\}$ and $\{\pi'_S, V', \mathcal{K}'\}$ are two Stinespring representations there exists a partial isometry $U : \mathcal{K} \to \mathcal{K}'$ such that

$$U\ V = V', \quad U^*V' = V, \quad U\pi_S(a) = \pi'_S(a)\ U. \tag{24}$$

The representation is called *minimal* if $\{\pi_S(a)\ V\phi,\ a \in \mathcal{A},\ \phi \in \mathcal{H}\}$ is dense in $\mathcal{K}$. Two minimal representations are unitary equivalent.

The following theorem of Arveson can be considered a corollary of Steinspring's theorem. It is a generalization of the Hahn-Banach theorem for states to the case of completely positive maps. Notice that if the Hilbert space $\mathcal{H}$ has complex dimension one, then Φ is a state.

Completely positive maps of subalgebras of $\mathcal{A}$ can be extended to completely positive maps of the full algebra.

Theorem 10 (Arveson) [1] *Let $\mathcal{A}$ a C^*-subalgebra with unit of a C^*-algebra $\mathcal{B}$. Then any completely positive map $\Phi : \mathcal{A} \to B(\mathcal{H})$ can be extended to a completely positive map $\Phi_1 : \mathcal{B} \to B(\mathcal{K})$ such that $\Phi_1/\mathcal{A} = \Phi$.* ◇

The states of a C^*-algebra are completely positive maps $\mathcal{A} \to C$ It is therefore natural to try and extend to completely positive maps the *Bures distance function* for states of a W^*-algebra, which generalizes to the non commutative setting the distance function between probability measures.

Definition 5 (*Bures distance*) [4] Let $\mathcal{A}$ be a W^*-algebra, Σ the collection of its normal states, π a representation in $\mathcal{B}(\mathcal{H})$. For every $\rho \in \Sigma$ define the following set $S_\pi(\rho)$

$$S_\pi(\rho) \equiv \{x \in \mathcal{H} : (x, \pi(a)x) = \rho(a), \quad \forall a \in \mathcal{A}\}. \tag{25}$$

The vector $x \in \mathcal{H}$ *represents* the state ρ in the representation π. The *Bures distance* $d(\rho_1, \rho_2)$ between two states ρ and τ is by definition

$$d(\rho, \tau) = \inf_{\pi \in \Lambda} d_\pi(\rho, \tau), \tag{26}$$

where if $S_\pi(\rho)$ or $S_\pi(\tau)$ are empty $d_\pi(\rho, \tau) \equiv [\rho(I)\tau(I)]^{\frac{1}{2}}$ and otherwise

$$d_\pi(\rho, \tau) = \inf\{\|x - y\|, \quad x \in S_\pi(\rho), y \in S_\pi(\tau)\}. \tag{27}$$

Defining

$$\xi_\pi(\rho, \tau) = \sup\{|(x, y)|, \quad s \in S_\pi(\rho), y \in S_\pi(\tau)\}, \quad \xi(\rho, \tau) = \inf_{\pi \in \Lambda} \xi_\pi(\rho, \tau) \tag{28}$$

one has

$$d(\rho, \tau)^2 = \rho(I) + \tau(I) - 2\xi(\rho, \tau). \tag{29}$$

♣

The Bures distance has interesting properties [1, 4]:

(1) Let $\{W_i, \ i \in \mathcal{I}\}$ be a collection of semi-finite W^*-algebras and for each of them let ρ_i be a normal state of W_i. Define $\mathcal{A} = \otimes_{i \in \mathcal{I}}(\mathcal{A}_i, \rho_i)$ and define on $\mathcal{A}$ the following normal state

$$\rho(a) = \otimes_{i \in \mathcal{F}} \rho_i(a_i), \quad a \in \mathcal{A}. \tag{30}$$

Suppose that for each value of the index i there exists an isomorphism ϕ_i of the algebra $\mathcal{A}_i$ with the algebra $\mathcal{B}_i$. Then necessary and sufficient condition for the existence of an isomorphism between $\otimes(\mathcal{A}_1, \rho_i)$ and $\otimes(\mathcal{B}_1, \tau_i)$ is that

$$\sum [d_{\phi_i}(\rho_i, \tau_i)]^2 < +\infty. \tag{31}$$

(2) If each algebra $\mathcal{A}_i$ is a finite factor with normalized trace τ_i then $\otimes(\mathcal{A}_i, \rho_i)$ is finite if and only if $\sum_i [d(\rho_i, \tau_i)]^2 < +\infty$.

Recall that for a linear map T from a normed space X to a normed space Y the *norm* $\|T\|$ of the map is defined by $\sup_{\|x\| \leq 1} \|Tx\|$. For completely positive maps it is convenient to introduce a norm for the product $T \otimes I_n$, $n = 1, \ldots,$ as $\|T\|_{c.l.} = \sup_{n \in N} \|T \otimes I_n\|$, where I_n is the identity map rank N matrices.

We shall call *completely bounded* the maps for which $\|T\|_{c.l.}$ is finite. Every completely positive map is completely bounded and $\|T\|_{c.l.} = \|T\| = \|T(I_{\mathcal{A}})\| =$

$\|V^*\ V\| = \|V^2\|$ where V is Steinspring's dilation for T. Of course one has $\|T\| \leq \|T\|_{c.l.}$ with equality sign if the algebras are abelian.

Stinespring representation is unique up to partial isometries in the dilation space. Given two representations $(\pi_1, V_1, \mathcal{K}_1)$ and $(\pi_2, V_2, \mathcal{K}_2)$ for a completely positive map $T : \mathcal{A} \to \mathcal{B}(\mathcal{K}_i)$ there exists a partial isometry $U : \mathcal{K}_1 \to \mathcal{K}_2$ such that $U\pi_1(a) = \pi_2(a)U$.

A representation is called minimal if $\bigcup_{a\in\mathcal{A},\ \psi\in\mathcal{H}} \pi(a)V\psi$ is dense in $\mathcal{K}$. Two minimal representations are unitarily equivalent. In the representation π the *Bures distance* between the completely positive maps T_1 and T_2 is defined as

$$\beta_\pi(T_1, T_2) \equiv inf\{\|V_1 - V_2\|\ ,\ V_i \in S(T_i, \pi)\} \tag{32}$$

where the fiber $S(T, \pi)$ is defined as the set of all operators $V : \mathcal{H} \to \mathcal{K}$ such that $(\pi, V, \mathcal{K})$ dilates T.

If one of the fibers is empty, we set $\beta_\pi(T_1, T_2) = 2$.

Definition 6 The Bures distance β between T_1 and T_2 is -the smallest π-distance

$$\beta(T_1, T_2) \equiv inf_\pi \beta_\pi(T_1, T_2) \tag{33}$$

◇

With these definition one has the following result.

Theorem 11 [12] *Let $\mathcal{A}$ be a C^*-algebra, and let T_1, T_2 be completely positive maps such that at least one is non-trivial.*

The following inequality holds:

$$\frac{\|T_1 - T_2\|_{cb}}{\sqrt{\|T_1\|_c} + \sqrt{\|T_2\|_c}} \leq \beta(T_1, T_2) \leq \sqrt{\|T_1 - T_2\|_c} \tag{34}$$

Moreover there exists a common representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{K})$ for T_1 and T_2 and two corresponding Steinspring representations $V_i : \mathcal{H} \to \mathcal{K}$ such that

$$\|V_1 - V_2\| = \beta_\pi(T_1, T_2) = \beta(T_1, T_2)$$ ◇

For C^*-algebras $\mathcal{A}$ other than $\mathcal{B}(\mathcal{H})$ it may seem natural to use the same definition. This poses a problem since the definition may depend on the representation of $\mathcal{A}$. The following definition does not refer to a representation and reduces to the previous definition when $\mathcal{A} = \mathcal{B}(\mathcal{H})$.

Definition 7 Given two C^*-algebras $\mathcal{A}$ and $\mathcal{B}$ and given two completely positive maps $T_i : \mathcal{A} \to \mathcal{B}$, $i = 1, 2$, the Bures distance is defined as

$$\beta(T_1, T_2) \equiv inf_{\hat{T}} \|\hat{T}_{1,1}(I_\mathcal{A}) + \hat{T}_{1,2}(I_\mathcal{A}) + \hat{T}_{2,1}(I_\mathcal{A}) + \hat{T}_{2,2}(I_\mathcal{A})\|^{\frac{1}{2}} \tag{35}$$

The infimum is taken over all positive extensions $\hat{T}\ \ \mathcal{A} \to \mathcal{B} \otimes \mathcal{B}(C^2)$ with completely bounded maps $\hat{T}_{i,j}\ :\ \mathcal{A} \to \mathcal{B}$ with the condition that $\hat{T}_{i,i} = T_i$ (whereas $\hat{T}_{i,j}$, $i \neq j$, is arbitrary). ◇

With this definition the Theorem holds for hyperfinite algebras; recall that a C^*-algebra is called hyperfinite if it is the closure of the union of an ascending sequence of finite dimensional subalgebras. The C^*-algebras considered in Quantum Statistical Mechanics are hyperfinite.

Further details and results on completely positive maps can be found in [14].

4 Properties of Dissipations

When discussing derivations of a C^*-algebra we have seen some classification results, in particular conditions under which they are inner. The results are not so complete in the case of dissipations. One has only partial results; we give some of them.

Theorem 12 [5] *Let Δ be a *-linear and bounded dissipation on the C^*-algebra $\mathcal{A}$ and let $e^{t\Delta}$ be defined by the power series expansion. Let $\mathcal{A}$ be concretely realized as closed subalgebra of $\mathcal{B}(\mathcal{H})$ and let $\mathcal{A}^-$ be its weak closure. The following statements are equivalent:*

(i) $e^{t\Delta}$ is completely positive for $t \in R^+$;

(ii) There exists a completely positive map $\Phi : \mathcal{A} \to \mathcal{A}^-$ and an operator $K \in \mathcal{A}^-$ such that

$$\Delta(a) = Ka + aK^* - \Phi(a). \tag{36}$$

In particular Δ is a complete positive dissipation if and only if it admits the decomposition Eq. 36 with

$$K + K^* - \Phi(e) \leq 0. \tag{37}$$

◇

Both conditions are satisfied if $\Delta = \delta^2$ where δ is a *-derivation.

To verify (i) notice that for each integer N and every choice of $a_1, \ldots, a_N$ one has

$$\delta^2(a_i^* a_j) = \delta^2(a_i^*)\ a_j + a_i^* \delta^2(a_j) + 2\delta(a_i^*)\delta(a_j). \tag{38}$$

If the rank N matrix A has elements $\delta(a_i^*)\ \delta(a_j)$, then A is positive. To see this notice that for each choice of $\eta \equiv \{\eta_k\}$ one has

$$\sum_{i,j} \bar{\eta}_i \delta(a_i^*)\delta(a_j)\eta_j = \left| \sum_i \delta(a_i)\eta_i \right|^2 \geq 0 \tag{39}$$

To verify (ii) notice that if δ^2 is closed and bounded also δ has these properties and therefore there exists $h = h^*$, $h \in \mathcal{A}^-$ such that $\delta(a) = [h, a]$. But then $\delta^2(a) = [h[h, a]] = h^2a + ah^2 - 2hah$.

Define $K = h^2$; (37) is verified with $\Phi(a) = 2hah$. Moreover

$$K + K^* - \Phi(e) = 2h^2 - 2h^2 = 0. \tag{40}$$

The condition $\Phi(e) = 0$ is necessary *but not sufficient* for Δ to be the square of a derivation.

Proof of Theorem 12 (ii) $\rightarrow$ (i): The map Φ extends by continuity to a completely positive map $\mathcal{A}^- \rightarrow \mathcal{A}^-$. Set

$$\Delta = \Delta_1 + \Delta_2 \quad \Delta_1(a) = Ka + aK^* \quad \Delta_2(a) = \Phi(a) \tag{41}$$

Notice that $e^{t\Delta_1}(a) = e^{tK}a(e^{tK})^*$ is completely positive and

$$e^{t\Delta_2}(a) = \sum_0^\infty \frac{t^n}{n!}\Phi^n(a), \quad t \geq 0 \tag{42}$$

is a completely positive map being uniformly convergent series of completely positive maps.

On the other hand $e^{t\Delta}(a) = \lim_{n\rightarrow\infty}(e^{\frac{t}{n}\Delta_1}e^{\frac{t}{n}\Delta_2})^n(a)$ and this map is completely positive being the uniform limit of such maps.

(i) $\rightarrow$(ii): To prove this point we will make use of the following theorem due to Christensen and Evans, a generalization of Kadison's theorem. We will state it without proof.

Theorem 13 [5] *Let $\mathcal{A}$ be a C^*-algebra, $\mathcal{A} \in B(\mathcal{H})$, and let π be a representation of $\mathcal{A}$ on a Hilbert space $\mathcal{K}$. Let Δ be a linear map $\mathcal{A} \rightarrow L(\mathcal{H}, \mathcal{K})$ such that*

$$\Delta(a)^*\Delta(b) \in \mathcal{A}^-, \qquad \Delta(ab) = \Pi(a)\Delta(b) + \Delta(a)\pi(b).$$

Then there exists $h \in \{\Delta(a)b,\ a, b \in \mathcal{A}^-\}$ for which

$$\Delta(a) - ha - \pi(a)h = 0, \quad \forall a \in \mathcal{A}. \tag{43}$$

$\diamond$

With the aid of this theorem we now prove (i) $\rightarrow$ (ii) in Theorem 12.

Remark that (i) implies that there exist a Hilbert space $\mathcal{K}$, a representation π of $\mathcal{A}$ in $B(\mathcal{K})$ and a linear map $\mathcal{A} \rightarrow L(\mathcal{H}, \mathcal{K})$ for which

$$(V(a^*))^*\pi(b)V(c) = \Delta(ab)c + a\Delta(bc) - a\Delta(b)c - \Delta(abc),$$

$$V(ab) = \pi(a)V(b), \qquad (\pi(\mathcal{A})V(a)\mathcal{H})^- = \mathcal{K}. \tag{44}$$

Indeed setting $D(a,b,c) \equiv \Delta(ab)c + a\Delta(bc) - a\Delta(b)c - \Delta(abc)$, one has that the map

$$(a_1,a_2)\times(b_1,b_2) \in (\mathcal{A}\times\mathcal{A})\times(\mathcal{A}\times\mathcal{A}) \mapsto D(a_1^*, a_2^* b_2, b_1) \tag{45}$$

is positive (since Δ is completely positive). By Steinspring's theorem there exist $\{\mathcal{K}, \Pi, V\}$ such that

$$D(a,b,c) = (V(a^*))^*\pi(b)V(c) \tag{46}$$

From (46) it follows that the identity $I \in B(\mathcal{K})$ belongs to the weak closure of $\pi(\mathcal{A})$. Let $\{b_n\}$ be a sequence for which $\Pi(b_n) \to I$. Then $V(a^*)^*V(b) \in \mathcal{A}^-$. It follows from the theorem of Christensen-Evans that there exists $h \in \mathcal{A}^-$ for which

$$V(a) = ha - \pi(a)h, \qquad \forall a \in \mathcal{A}.$$

Define the map $\Phi(a) \equiv h^*\pi(a)h$. From (46)

$$D(a,b,c) = (ah^* - h^*\pi(a))\pi(b)(hc - \pi(c)h) = \Phi(abc)$$

$$= a\Phi(b)c - a\Phi(bc) - \Phi(ab)c. \tag{47}$$

Define now $p \equiv \phi(e)$, $q \equiv D(e)$ (e is the identity of the algebra). Set $s \equiv \frac{p+q}{2}$ and $\{s,b\} \equiv sb + bs$; then

$$\Delta(ab) + \phi(ab) - a[\Delta(b) + \Phi(b)] - [\Delta(a) + \Phi(a)]b - aqb - apb$$

$$= -2sab = \{s, ab\} - a\{s,b\} - \{s,a\}b. \tag{48}$$

It follows that

$$\delta(a) \equiv \Delta(a) + \Phi(a) - \{s,a\} \tag{49}$$

is a *-derivation. By the Kadison-Sakai theorem there exists $b \in \mathcal{A}^-$, $b = b^*$, such that $\delta(a) = i[b,a]$. We have proved that (ii) holds with $C = a + ib$. This completes the proof of Theorem 12. ♡

Definition 8 A representation π is *normal* if for any filter $\{a_\beta\}$ one has $sup\pi(\{a_\beta\}) = \pi(\sup\{a_\beta\})$. ◇

In the case of *normal* representations one has stronger results. Let CP be the completely positive maps of a W^*-algebra $\mathcal{A}$ and let CP_n be the subset of normal ones.

It is convenient to recall that if π is a *normal* (*standard*) representation of $\mathcal{B}(\mathcal{H})$ on the Hilbert space $\mathcal{K}$ then there exists a direct sum decomposition

$$\mathcal{K} = \oplus_n \mathcal{H}_n$$

where the subspaces $\mathcal{H}_n$ are invariant under the action of $\mathcal{B}(\mathcal{H})$.

Theorem 14 (Kraus) [9] *A map T is $CP_n(\mathcal{H})$ if and only if*

$$T(a) = \sum_i V_i a V_i \quad i \in \mathcal{N} \tag{50}$$

where V_i and $V_i^* V_i \in \mathcal{B}(\mathcal{H})$. ◇

Proof Assume that T is completely positive; the projection P_n of $\mathcal{H}$ onto $\mathcal{K}_n$ commute with the representation ρ and therefore

$$T(b) = \sum_n V^* P_n \rho P_n V = \sum_n A_n^* X A_n, \quad A_n \equiv U N^* P_n V \tag{51}$$

where U_n are the unitary maps form $\mathcal{H}$ to $\mathcal{K}$.

If ψ_m is a countable dense set in $\mathcal{H}$ then

$$\infty \geq (T(|\phi_n)(\phi_n|)\phi_r, \phi_r) = \sum_r |(A_n \phi_r, \phi_m)|^2 \tag{52}$$

so that the set

$$\mathcal{N}_{r,m} \equiv \{n \ : (A_n \phi_r, \phi_m) \neq 0\} \tag{53}$$

is countable. Moreover the set

$$\mathcal{N} \equiv \cup_{r,m} \mathcal{N}_{r,m} \tag{54}$$

is countable.

It is easy to verify that the operators A_n satisfy $\sum_n A_n^* A_n \leq C I$ for some constant $C > 0$. ♡

The representation of T given above is not unique.

The maps CP_n can be lifted to product structures.

Lemma 15 *If $\Phi_1, \Phi_2 \in CP_n$ there exists $\Phi \in CP_n$ such that*

$$\Phi(a \otimes b) = \Phi_1(a) \otimes \Phi_2(b), \quad a \in B(\mathcal{H}_1), b \in B(\mathcal{H}_2). \tag{55}$$

◇

Notice also that setting $b \equiv e$ one derives that the map $D(a, b, c)$ is a measure of the dissipation of Δ. More precisely setting

$$\partial(\Delta) = \Delta(a^*a) - a^*\Delta(a) - \Delta(a^*)\, a, \tag{56}$$

it follows form the previous theorems that $\partial(\Delta)$ defines Δ modulo a derivation. Indeed

$$\partial(\Delta) = 0 \Rightarrow \Delta(a) = i[h, a], \qquad h = h^*, \qquad h \in \mathcal{A}^- \tag{57}$$

Some examples of totally positive bounded dissipations are

$$\Delta(a) \equiv bab - \frac{1}{2}(b^2a + ab^2) \tag{58}$$

and

$$\Delta(a) = U^*aU - a, \qquad U^*U = I = UU^* \tag{59}$$

(remark that if the algebra is abelian in both cases one has $\Delta(a) = 0$; this is in line with the fact that there is no bounded dissipation of an abelian algebra). One can also notice that in both examples above, the action of the generators takes the form

$$L(a) = \sum_1^n V_i^* a V_i - \frac{1}{2}\sum_i^n [V_i^*V_i\ a\ -\ aV_i^*V_i] \tag{60}$$

(in both cases $h = 0$).

We recall Steinspring's theorem which is valid even if $\mathcal{H}_2$ is infinite dimensional. We present it in a slightly different form; it corresponds to lifting the map to a larger space.

Theorem 16 (Steinspring) *If $\Phi \in CP(\mathcal{H}_1)$, $\Phi(I) = I$ there is a Hilbert space $\mathcal{H}_2$ and a unitary operator U in $\mathcal{H}_1 \otimes \mathcal{H}_2$ such that for all positive trace class operators $\sigma \in \mathcal{B}(\mathcal{H}_2)$ with trace 1 the following holds*

$$\Phi(a) = Tr_2(\sigma U^*(a \otimes I)U). \tag{61}$$

◇

Proof We have seen that

$$\Phi(a) = \sum_i V_i^* a V_i \tag{62}$$

Choose unitary operators U_i in $\mathcal{H}_2$ such that $U_i^*U_j = \delta_{i,j}$. Define U by $V = \sum_i V_i \otimes U_i$. then

$$U^*(a \otimes I)U = \sum V_i^* a V_i \otimes U_i^* \Phi(a) U_i \tag{63}$$

♡

From this result one derives

Lemma 17 (Stormer) [16] *If $\mathcal{A}$ is a C^*-algebra and $\Phi \in CP(\mathcal{A})$ with $\Phi(e) = e$ then for all $a \in \mathcal{A}$*

$$\Phi(a^*)\Phi(a) \leq \Phi(a^*a). \tag{64}$$

♡

5 Complete Dissipations

Let $\mathcal{A}$ be a W^*-algebra and let $\Phi(t) = e^{tL}$ be a norm continuous semigroup of CP maps of $\mathcal{A}$ with $\Phi(t)I = I$. Extend the generator L to a matrix algebra $M_N(\mathcal{A})$ setting $L_N = L \otimes I$.

Differentiating (64) with $\Phi \to \Phi(t)$ one obtains

$$L_n(a^*a) - L_n(a^*) \cdot a - a^* \cdot L_n(a) \geq 0 \tag{65}$$

From $\Phi_N(I) = I$ and from the fact that $\Phi(t)$ is a $*$-map it follows $L_N(I) = 0$, $L_N(a^*)(L_N(a))^*$.

Define the *dissipation function* by

$$D(L; a, b) = L(a^*b) - L(a^*)b - a^*L(b) \qquad a, b \in \mathcal{A} \tag{66}$$

Definition 9 (*completely dissipative map*) A bounded map $L : \mathcal{A} \to \mathcal{A}$ is called *completely dissipative* if it satisfies

$$L(I) = 0 \quad L(a_N^*) = L^*(a_N) \quad D(L_N, a_N, a_N) \geq 0, \forall N, \forall a_N \in M_N(\mathcal{A}) \tag{67}$$

We denote with $CD(\mathcal{A})$ a completely dissipative map and write $L \in CD(\mathcal{A})$ and denote with $CD(\mathcal{A})_\sigma$ the set of weak* continuous elements in CD.

Remark that $CD(\mathcal{A})$ is a convex cone. ◇

The following Theorem plays a special role.

Theorem 18 [13] *$D(L; a, b)$ determines L up to a Hamiltonian (a derivation).* ◇

Proof

(1) If $D(L; a, b) = 0$ $\forall a, b \in \mathcal{A}$ then it is a derivation. All derivations are weakly inner and therefore $\exists b \in \mathcal{A}$ such that $L(a) = [b, a]$. From $L(b^*) - (L(b)) = 0$ it follows that $b = ih$ with h self-adjoint. Then $\Phi(t)$ can be continued to a one-parameter group of automorphisms $\Phi(t)(a) = e^{iht} a e^{-ith}$.

(2) If L generates of group of CP maps the L and $-L$ satisfy (65). Hence $D(L; a, a)$ is a derivation. ◇

Theorem 19 [13] *Let L $\mathcal{A} \to \mathcal{A}$ be a bounded *-map and put $\Phi_t = e^{tL}$. Then the following are equivalent*

(1)

$$\Phi_t(a)^*\Phi_t(a) \leq \Phi_t(a^*a) \tag{68}$$

(2)

$$D(L; a, a) \geq 0 \quad \forall a \in \mathcal{A} \quad L(I) = 0 \tag{69}$$

◇

Proof We already proved $1 \to 2$. To prove the converse we proceed as in the general case of Banach spaces (see "Lecture 11: Semigroups and Dissipations. Markov Approximation. Quantum Dynamical Semigroups I").

Recall that in a Banach space L generates a norm continuous semigroup iff

$$\Xi(L) \equiv lim_{t \to 0_+}(\|I + tL\| - 1) \leq 0 \tag{70}$$

In our case since $\mathcal{A}$ is a C^*-algebra, we have

$$\|(I - tL)\| = \sup_U \|U + tL(U)\| \tag{71}$$

where the supremum is taken over the unitaries in $\mathcal{A}$. It is easy to show that (71) implies (70).

From (67) one concludes (2) $\to$ (1) by a standard argument using the Yosida approximants. ♡

Applying Theorem 19 to the matrix extensions L_N and Φ_N to the matrix algebras $M_N(\mathcal{A})$ one obtains

Theorem 20 *Let $\mathcal{A} \to \mathcal{A}$ be bounded *-map and let $\Phi_t = e^{tL}$. Then $\Phi_t \in CP(\mathcal{A})$, $\Phi_t(I) = I$ iff $L \in CD(M_n(\mathcal{A}) \equiv CD_N(\mathcal{A})$.*

It follows that Φ_t is a norm continuous dynamical semigroup on $\mathcal{A}$ iff $\Phi_t = e^{tL}$ with $L \in CD_n(\mathcal{A})$. ◇

It is easy to extend this theorem to CP maps that do not satisfy $\Phi_t(I) = I$. Differentiating (64) for Φ_m at $t = 0$ gives

$$L_N(A^*A) + A^*L_N(I)A - L_N(A^*)A - A^*L_N(A) \geq 0 \qquad \forall A \in M_N(\mathcal{A}) \tag{72}$$

It is easy to see that the map $L^1(A) = L(A) - \frac{1}{2}[L(I), A]$ belongs to $CD(\mathcal{A})$. On the other hand the map $\tilde{L}_K(X) = [K, A]$ generates the semigroup

$$e^{t\tilde{L}}(A) = e^{tK}Ae^{-tK} \tag{73}$$

the semigroup Φ_t generated by $K' + \tilde{K}$ is obtained by the Lie-Trotter formula

$$\Phi_t = lim_{n\to\infty}[e^{t\frac{L^1}{n}} e^{t\frac{\tilde{L}}{n}}]^n \tag{74}$$

and therefore $\Phi \in CP(\mathcal{A})$.

6 General Form of Completely Dissipative Generators

Assume that $\mathcal{A}$ is a hyperfinite factor in a separable Hilbert space. This definition includes the case $\mathcal{A} = \mathcal{B}(\mathcal{H})$ where $\mathcal{H}$ is a separable Hilbert space but also more general cases e.g. lattice of spins and Fermi fields on a lattice. It can be used to study the validity of the Markov approximation for the case of a quantum particle interaction with a lattice of spins or with a gas of fermions in an equilibrium state.

The factor $\mathcal{A}$ s therefore generated by a sequence of factors $\mathcal{A}_n$ of type I_n such that $\mathcal{A} = (\cup\mathcal{A}_n)''$ $\mathcal{A}_n \subset \mathcal{A}_{n+1}$. The collection U of unitary elements in $\mathcal{A}$ form a compact group with an invariant *mean* defined by the Haar measure.

One has then [12]

Theorem 21 *If $L \in CD(\mathcal{A})$ then there is a map $\Phi \in CP(\mathcal{A})$ and a self-adjoint element in $\mathcal{A}$ such for all $a \in \mathcal{A}$*

$$L(a) = \Phi(a) - \frac{1}{2}\{\Phi(I), a\} + i[H, a]. \tag{75}$$

$\diamond$

Proof Define $K_n \in \mathcal{A}$ by

$$K_n \equiv \mathcal{M}_n[L(U^*)U] \qquad \mathcal{M}_n = \mathcal{A}_n'' \tag{76}$$

One has $\|K_n\| \leq \|L\|$. If V is unitary in $\mathcal{M}_n$ then

$$\mathcal{M}_n[L(VU^*)U] = \mathcal{M}_n[L(U^*)UV] = K_nV \tag{77}$$

hence for all $X \in \mathcal{A}_n$

$$\mathcal{M}_n[L(XU^*)U] = K_nX \tag{78}$$

It follows that for all $X,\ Y \in \mathcal{A}_n$

$$\mathcal{M}_n[D(L, UX, UY)] = L(X^*Y) - K_nX^*Y - X^*YK_n^* \tag{79}$$

Define a linear map $\Phi_N :\ \mathcal{A} \to \mathcal{A}$ by

$$\Phi_n(X) = L(X) - K_n X - X K_n^* \tag{80}$$

Then, since $L \in CD(\mathcal{A})$ one has

$$\Phi_n(X^* X) = \mathcal{M}_n[D(L; UX)] \geq 0 \tag{81}$$

i.e. Φ_n restricted to $\mathcal{A}$ is positive for all $n \in Z$.

Applying the same procedure to $\mathcal{M}_N(\mathcal{A}_m) \equiv M_N(C) \otimes \mathcal{A}_m$ one proves that

$$\mathcal{M}_{m,N}[L_n(U^*)U] = K_m \otimes I_N \tag{82}$$

where $\mathcal{M}_{m,N}$ is the invariant mean over the group of unitary elements in $\mathcal{M}_N(\mathcal{A}_m)$. From this one derives

$$\Phi_m(\mathcal{A}_m) \in CP(\mathcal{A}_m, \mathcal{A}). \tag{83}$$

Define

$$\Phi_K(X) = L(X) - KX - X^* K$$

$$\Gamma_n = \{K \in \mathcal{A} \ : \ \Phi_K | \mathcal{A}_n \in CP(\mathcal{A}_n, \mathcal{A})\} \quad \|K\| \leq \|L\| \tag{84}$$

Γ_n is not empty, $\Gamma_n \subset \Gamma_{n+1}$ and Γ_n is weakly closed. Since the unit ball of $\mathcal{A}$ is compact in the weak operator topology it follows that Γ_n are weakly compact and therefore $\Gamma = \cap \Gamma_n$ is not empty.

Choose $K \in \Gamma$. Then $\Phi \equiv \Phi_K$ is CP on $\cup_n \mathcal{A}_n$. Since L is continuous in the ultraweak topology it follows that Φ is a CP map on $(\cup_n \mathcal{A}_n)''$. One has $\Phi(I) = -K - K^*$.

The thesis of Theorem 21 follows now putting $H = \frac{i}{2}(K^* - K)$. ♡

Notice that H is not uniquely defined by this procedure since we were free to choose $K \in \Gamma$.

Let now $\mathcal{A}$ be a C^*-algebra.

Theorem 22 *If $\Phi \in CP(\mathcal{A})$ and H is self-adjoint in $\mathcal{A}$ then L in Theorem 20 is in $CD(\mathcal{A})$.* ◇

Proof One has $L(I) = 0$ and $L(a^*) = (L(a))^*$. From Theorem 12 we know that the representation π can be chosen so that $\pi(I) = I$. Then

$$\Phi(a^* a) + a^* \Phi(I) a - \Phi(a^*) a - a^* \Phi(a) = (\pi(a)V - V\pi(a))^* (\pi(a)V - V\pi(a))^* \geq 0 \tag{85}$$

♡

The same argument applies to the extensions Φ_N and L_N; this shows that $L \in CD(\mathcal{A})$.

From Theorems 21 and 22 and Kraus lemma one derives

Theorem 23 [10] *The operator L belongs to $CD(\mathcal{H})_\sigma$ if and only if it is the form*

$$L(a) = \sum_j (V_j^* a V_j) - \frac{1}{2} V_j^* V_j, a^* + i[H, a] \tag{86}$$

with V_j, $V_j^* V_j, \in \mathcal{B}(\mathcal{H})$, and H self-adjoint in $\mathcal{B}(\mathcal{H})$.

The dual version of this theorem in state space is

$$L(\rho) = \frac{1}{2} \sum_j ([V_j \rho, V_j^*] + [V_j, \rho V_j^*]) - i[H, \rho] \tag{87}$$

◇

References

1. Arveson (1969). *Acta Mathematica*, *123*, 141–224.
2. Batty (1984). *Acta Applicandae Mathematicae*, *2*, 221–296.
3. Bratteli & Robinson (1981). *Mathematica Scandinavia*, *49*, 259–274.
4. Bures (1969). *Transactions of the American Mathematical Society*, *135*, 199–212.
5. Christensen-Evans (1979). *Journal of London Mathematical Society*, *20*, 358–368.
6. Davies, B. (1974). *Communications in Mathematical Physics*, *35*, 91–130.
7. Evans (1976). *Communications in Mathematical Physics*, *48*, 15–22.
8. Evans & Hanche-Olesen (1979). *Journal of Functional Analysis*, *32*, 207–212.
9. Kraus, K. (1970). *Annals of Physics*, *64*, 311–324.
10. Kraus, K. (1983). *Lecture notes in physics*. New York: Springer.
11. Kurtz (1973). *Transactions of the American Mathematical Society*, *186*, 259–272.
12. Kretschmann, D., Schlingermann, D., & Werner, R. (2008). *Journal of Functional Analysis*, *255*, 1889–1904.
13. Lindblad (1976). *Communications in Mathematical Physics*, *48*, 119–130.
14. Paulsen (2002). *Completely bounded maps and operator algebras*. Cambridge: Cambridge University Press.
15. Steinspring (1955). *Proceedings of the Royal Mathematical Society*, *6*, 211–216.
16. Störmer (1974). *Springer Lecture Notes in Physics*, *29*, 85–106.

Lecture 13: Weyl System, Weyl Algebra, Lifting Symplectic Maps. Magnetic Weyl Algebra

1 Canonical Commutation Relations

We have seen in the preceding Lectures that in the formulation of Quantum Mechanics for system with N degrees of freedom given by Heisenberg and Born an important role is played by the possibility of finding $2N$ self-adjoint operators, denoted here by $Q_k,\ P_k, k = 1, \ldots, N$, which on a suitable dense domain satisfy the (commutation) relations

$$[Q_k, Q_h] = [P_k, P_h] = 0, \qquad [Q_k, P_h] = i\delta_{k,h} \quad k, h = 1, \ldots, N \tag{1}$$

We shall call (1) *Canonical Commutation Relations (CCR)*. In the construction of a quantum dynamics operators that satisfy CCR play the same role as the coordinates in phase space, upon substitution of Poisson brackets with commutators.

We have already remarked in "Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories" that these two structures have the same algebraic properties. In this setting, if the Hamiltonian $H_{\text{class}} = p^2 + V(q)$ is the generator of classical dynamics, then the operator $H_{\text{quant}} = P^2 + V(Q)$ will be the generator (via Stone's theorem) of quantum dynamics.

The imaginary unit is needed in (1) because the commutator of two hermitian operators is anti-hermitian. The corresponding feature in Hamiltonian Mechanics is the fact that the symplectic structure of R^{2n} can be regarded as complex structure on C^n. The canonical symplectic matrix J with eigenvalues $\pm i$ is substituted in Quantum Mechanics by multiplication by the imaginary unit i.

A first difficulty one encounters in the analysis of the CCR comes from the fact these relations cannot be satisfied by bounded operators. This can be seen as follows.

If for given index k the operators P_k and Q_k were both bounded, from (1) it would follow for all integers n

$$Q_k^n - Q_k^{n-1} P_k = Q_k^{n-1}[Q_k, P_k] = i\, Q_k^{n-1} \tag{2}$$

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_13

and by iteration

$$in\,Q_k^{n-1} = Q_k^n P_k - P_k Q_k^n \quad \forall n$$

Therefore

$$n\|Q_k\|^{n-1} \leq 2\|Q_k\|^n \|P_k\|$$

If $\|Q_k\| \neq 0$ then $n < 2\|Q_k\|\,\|P_k\|$ for any natural number n, a contradiction.

On the other hand, if for each value of the index at least one of the operators Q_k, P_k is unbounded, (1) must be written

$$Q_k P_h - P_h Q_k \subset i\delta_{h,k} I \tag{3}$$

i.e. *on its domain of definition* the operator on the left of (3) it coincides with multiplication by $i\delta_{h,k}$. In this weaker sense the solution of (3) *is not unique* (up to unitary equivalence).

We consider the case $N = 1$ and give three *inequivalent* solutions of (1).

Solution (A) The (concrete) Hilbert space of the representation is $\mathcal{H} = L^2(R, dx)$.

The representation is given by the operators

$$\begin{array}{ll} (Qf)(x) \equiv xf(x), & D(Q) = \{f | f \in L^2, \quad xf(x) \in L^2\} \\ (Pf)(x) \equiv i\frac{df}{dx}, & D(P) = \{f | f \in L^2, \frac{df}{dx} \in L^2\} \end{array} \tag{4}$$

Remark that both Q and P are essentially self-adjoint on the dense subspace space $\mathcal{S} \in L^2(R^1)$ (indeed Q is multiplication by x, P is multiplication by p in the Fourier transform representation) and they have $\mathcal{S}$ as common *invariant* dense domain.

This representation is irreducible (no proper subspace is left invariant). By Stone's theorem $U_a \equiv exp\{iaQ\}$ and $V_b \equiv exp\{ibP\}$ with $a, b \in R$ are one-parameter strongly continuous groups of unitary operators.

It is easy to verify that the following identity holds

$$U_a V_b U_a^* = V_b exp\{-iab\} \tag{5}$$

Remark that (1) is the differential form of (5).

Solution (B) The Hilbert space is now $\mathcal{H} = L^2(S^1) \equiv L^2([0, 2\pi], d\theta)$.

Q is a bounded operator defined on all of $\mathcal{H}$ by $(Qf)(\theta) = \theta f(\theta)$.

The domain $D(P)$ of the operator P is made of all *periodic functions* which are square integrable with square integrable derivative. On this domain P acts as $-i\frac{d}{d\theta}$.

Also in this case both Q and P are self-adjoint and Q is bounded. However Q *doesn't leave the domain of* P *invariant* since Qf does not belong to the domain of P if $f(2\pi) \neq f(0)$.

Therefore the product PQ is defined only on functions for which $f(0) = f(2\pi)$. This set is dense in $L^2[0, 2\pi]$ but the restriction of P to this set does not define uniquely a self-adjoint operator (see "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method"; to define a self-adjoint operator one needs to impose two boundary conditions).

It can be verified by a direct computation that equation (5) *is not satisfied*: the operator $U_a V_b U_a^* V_b^*$ *is not a multiple of the identity*.

Solution (C) Consider the *characters* ξ_λ of the group R (i.e. $\xi_\lambda \equiv e^{i\lambda x}$, $\lambda \in R$). Denote by K the L^∞*-closure* in the topology of the continuous functions of finite linear combinations $\sum_1^N c_i \xi_{\lambda_i}$. K is the space of *quasi periodic functions*. The quadratic form

$$(\phi, \psi) \equiv \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \bar{\phi}(t)\ \psi(t) dt$$

is well defined on K and defines a pre-scalar product.

Denote by $\mathcal{K}$ the Hilbert space completion of K in the topology of this scalar product. The characters form a non-denumerable *orthonormal basis* in $\mathcal{K}$ and therefore the Hilbert space $\mathcal{K}$ is *non separable*. On $\mathcal{K}$ consider the two families of unitary operators U_a and V_b, $a,\ b \in R$ defined by

$$V_a \xi_\lambda = \xi_{\lambda - a}, \qquad U_b\ \xi_k = e^{ixb} \xi_k \tag{6}$$

This operators satisfy (5) but the map $a \to V_a$ is not continuous in the strong topology of $\mathcal{K}$ and not even Lebesgue-measurable (i.e. $(\phi, V_a \phi)$, $\phi \in \mathcal{K}$, is not measurable as functions of a).

To verify this, notice that weak and strong measurability coincide for unitary maps, and that

$$V_0 \equiv I, \qquad |(V_a - I)\xi_\lambda|_2 = \sqrt{2},\ a \neq 0$$

The representation (C) is irreducible (an element of $B(\mathcal{K})$ which commutes with U_λ and V_a is a multiple of the identity). It is not equivalent to (A) because the map $a \to V_a$ is not Lebesgue-measurable (whereas in solution A it is continuous).

Remark that one can consider in a representation of type (C) the C^*-subalgebra generated by V_a and U_b *for a given pair* $(a, b) \in R^2$. Equivalently one can consider the C^*-algebra generated by elements $U,\ V$ such that $UV = e^{2i\pi\Theta} VU$. This algebra is called *rotation algebra* and plays a relevant role in the analysis of the mathematical structure of aperiodic crystals. If $\Theta \notin \mathcal{Q}$ the rotation algebras has a unique (up to unitary equivalence) *faithful* irreducible representation in a *separable* Hilbert space. We shall come back to these issues in the second part of this Lectures.

2 Weyl System

Since there are many inequivalent representation we will not put (1) (the Heisenberg commutation relations) at the basis of Quantum Mechanics. Following Weyl [5, 10, 12] we will privilege the relation (5) among unitary operators, requiring also that the maps $a \to U_a, b \to V_b$ be *Lebesgue-measurable*.

If the system has N degrees of freedom we define the Weyl system by the following relations

$$U(a)V(b)U^*(a)V^*(b) = exp\{-i(a,b)\}, \qquad a,\ b \in R^N \tag{7}$$

with $a,\ b \in R^N$, where (a, b) denotes the scalar product in R^N and

$$U(a) = e^{i(a,Q)} \qquad V(b) = e^{i(b,P)}, \quad Q = Q_1, \ldots, Q_N, \qquad P = P_1, \ldots, P_N.$$

We shall prove that solution (A) given above is the *unique* (modulo unitary equivalence) irreducible representation of (7) for which the maps $a, b \in R^N \mapsto U(a),\ V(b) \in U(\mathcal{H})$ are Lebesgue-measurable. If one does not require irreducibility any representation of (7) decomposes as a direct sum of identical copies of this representation.

This result, due to Weyl and von Neumann, [12] implies the equivalence of Schrödinger's and of Born-Heisenberg's representations of the CCR.

We shall put (7) in an equivalent form. Let $z = a + ib$, $z \in C^N$, $a, b \in R^N$, and define

$$W(z) = exp\{-i\frac{(a,b)}{2}\}V(b)U(a) \tag{8}$$

Then (7) is equivalent to

$$W(z)W(z') = exp\{-\frac{i}{2}Im\langle z, z'\rangle\}W(z+z'), \qquad W^*(z) = W(-z), \qquad z \in C^N \tag{9}$$

We have denoted by $\langle z, z'\rangle$ the scalar product in C^N (antilinear in the first element). The relation (9) shows that $z \to W(z)$ is a *projective unitary representation* of the additive group C^N. It is strongly continuous by construction.

Definition 1 (*Weyl system*) We shall call *Weyl system* on the Hilbert space $\mathcal{H}$ a collection of operators which satisfy (9) and are continuous in z in the strong operator topology.

Notice that $exp\{-\frac{i}{2}Im\langle z, z'\rangle\}$ is a phase factor. The semidirect product the Weyl system with the group S^1, defines a group: the *Heisenberg group*.

As one sees from (8) in the definition of Weyl system the operators Q_k and P_k are not treated symmetrically. One could have defined the Weyl operators $W(z)$

inverting the position of $U(a)$ and $V(b)$ in the definition. It is easy to see that this would correspond to an anti-unitary transformation.

Conversely from Stone's theorem applied to the subalgebras corresponding to real and imaginary values of z (each of these subalgebras is commutative) it follows that there exist self-adjoint operators Q_k and P_k which generate the corresponding N-parameter subgroups and which satisfy (1) on a common dense invariant domain.

Notice that setting $x + iy = z \equiv \{x, y\}$ one has

$$Im\langle z, z'\rangle = \omega(z, z')$$

where ω is the standard symplectic 2-form: recall that the complex structure of C^N is isomorphic to the symplectic structure of $T^*(R^N) \equiv R^{2N}$. One can therefore write (9) as

$$W(z)W(z') = e^{-\frac{i}{2}\omega(z,z')}W(z+z') \quad z, z' \in C^N \tag{10}$$

3 Weyl Algebra. Moyal Product

To prove uniqueness of the representation of the Weyl system it is convenient to study first a more abstract algebraic structure, in analogy with what is done in the analysis of the representations of Lie groups through their group algebras.

Let $z \to W(z)$ be a Weyl system, Lebesgue-measurable in the weak sense. For each function $f \in L^1(C^N)$ with norm $\|f\|_1$ we define an operator W_f as follows:

$$W_f \equiv \int dz f(z) W(z) \tag{11}$$

where the integral is understood in the Bochner sense (as a Lebesgue integral for expectation values for vectors in $\mathcal{H}$). It is easy to see that W_f is a bounded operator with

$$\|W_f\| \leq \|f\|_1 \tag{12}$$

Therefore the linear map $f \to W_f$ is norm-continuous.

The following identities are easy to verify

$$W_f + W_g = W_{f+g}, \qquad W_f^* = W_{\tilde{f}}$$

$$W_f\ W_g = W_{f\times g}, \quad (f \times g)(z) \equiv \int dz' f(z - z')g(z')e^{\frac{i}{2}\omega(z,z')} \tag{13}$$

Definition 2 (*Moyal product*) [8] The product $f \times g$ defined in (13) is called *Moyal product*. It is *non-commutative*.

This product is sometimes called Weyl-Groenewold product [1].

Lemma 1 The map $f \mapsto W_f$ is injective.

Proof If $W_f = 0$, then for every $\psi, \phi \in \mathcal{H}$ one has

$$\int dz f(z)(\psi, W(z)\phi) = 0$$

Setting $\phi' = W(z_0)\phi$, $\psi' = W(z_0)\psi$, from

$$W(-z_0)W(z)W(z_0) = W(z)e^{\frac{i}{2}\omega(z,z_0)}$$

one derives that for every pair ψ', ϕ'

$$0 = \int dz f(z)(\psi', W(z)\phi')e^{\frac{i}{2}\omega(z,z_0)}$$

One concludes that the Fourier transform of $f(z)(\psi', W(z)\phi')$ vanishes. For each value of z one can choose $\phi' = W(z)\psi'$ and therefore

$$\langle \phi', W(z)\psi' \rangle = \langle \phi', \phi' \rangle = 1 \tag{14}$$

We conclude that $f = 0$. ♡

We can regard the W_f, with operator norm $\|W_f\|$, product law $W_f W_g = W_{f\times g}$ and conjugation given by $W_f^* = W_{\bar{f}}$, as a C^*-algebra, without reference to the Weyl system. Notice that (12) implies that W is norm dense in L_∞ and $\|W_f\| \leq \|\hat{f}\|_\infty$. Therefore

$$\|W_f\| = \|\hat{f}\|_\infty \tag{15}$$

since both terms are C^*-algebra norms.

Definition 3 (*Weyl algebra*) We will call *Weyl algebra* the algebra generated by the W_f and we will denote it by the symbol $\mathcal{W}$. We not indicate the number N of degrees of freedom, with the convention that $N < +\infty$ unless otherwise stated explicitly.

Notice that so far we have not taken into account that position and momentum have dimensions and that the product position times momentum has the dimension of an action. This suggests to choose units such that $[Q_k^\hbar, P_h^\hbar] = i\hbar\delta_{k,h}$. With this convention the elements of Weyl's system are now

$$W^\hbar(z) = e^{-i\frac{(a,b)}{2\hbar}} V(b)U(a) \tag{16}$$

and the product that defines Weyl's algebra is

$$W^{\hbar}(z)W^{\hbar}(z') = W^{\hbar}(z+z')e^{-\frac{i}{2\hbar}\omega(z,z')} \tag{17}$$

From the point of view of mathematics it is natural to define as *semiclassical limit* of this algebraic structure the limit $\hbar \to 0$. It is clear from (17) that the semiclassical limit is singular: the relation (17) contains a factor which has fast oscillations when $\hbar \to 0$ and therefore the limit $\hbar \to 0$ in (17) is *a singular limit*.

We will discuss this limit in "Lecture 16: Semiclassical Approximation for Fast Oscillating Phases. Stationary Phase. W.K.B. Method. Semiclassical Quantization Rules" and we shall analyse it further in the second part of these Lecture Notes. We will describe there the relation, in the semiclassical limit, of the algebraic structure of the generators of one-parameter groups in the Weyl quantization with the Poisson algebra of the generators in Hamiltonian Mechanics. The properties of oscillatory integrals play a crucial role in the formulation of the semi-classical limit in the Schrödinger representation.

From the point of view of Physics $\hbar$ is a physical constant and its value is not at our disposal. The *mathematical limit* $\hbar \to 0$ gives informations about those quantities that have the dimension of the action and take values large as compared to Planck's constant.

4 Weyl Quantization

In each representation of the Weyl system the map $f \to W_f^{\hbar}$ provides a correspondence between functions in classical phase space and operators on a Hilbert space. In "Lecture 6: Operators on Hilbert Spaces I; Basic Elements" we have called *quantization* any procedure that associates to a function (in a suitable class) on phase space an operator on a Hilbert space $\mathcal{H}$. We shall come back at the end of this Chapter to the problem of quantization.

We shall call *Weyl quantization* the quantization performed making use of the Weyl algebra. Notice that if we substitute the Weyl algebra with the (abelian) algebra of the characters of the multiplicative group R^{2N} we obtain the correspondence $f \mapsto \mathcal{F}(f) = \hat{f}$ (Fourier transform) and the product structure is mapped into convolution. Therefore one may regard the correspondence $f \to W_f^{\hbar}$ as a *twisted convolution* or a *symplectic Fourier transform*.

Definition 4 (*Weyl quantization*) *Weyl quantization* is the map $f \mapsto W_f^{\hbar}$ defined by

$$f \mapsto W_f^{\hbar} = \int f(z)W^{\hbar}(z)dz \qquad f \in L^1 \tag{18}$$

When we will came back to Weyl quantization in the second volume of these Lecture Notes, we will use the (more commonly used) notation

$$W_f^{\hbar} \equiv Op_{\hbar}^{W}(f) \tag{19}$$

Weyl quantization, originally defined for continuous functions, can be extended to a large class of functions. We shall discuss this in some detail in the second part of these Lecture Notes. It is also clear that a generalization of Weyl quantization can be obtained by substituting to ω in (17) other symplectic structure.

At the end of this Lecture we shall see an example: the *magnetic Weyl algebra*.

Coming back to the Weyl algebra we recall that every C^*-algebra has a faithful representation as operator algebra on a Hilbert space $\mathcal{K}$. If the C^*-algebra is separable, the space $\mathcal{K}$ can be taken to be separable.

Given a representation π of $\mathcal{W}$ in $\mathcal{K}$ one can ask what are the conditions under which one can *reconstruct a Weyl's system*. To answer this question consider a sequence $\{f_n \in L^1(C^N)\}$ which converges *in distributional sense* to the measure concentrated at the point $z_0 \in C^N$.

Consider the sequence of operators $\pi(W_{f_n})$ on $\mathcal{K}$. If this sequence converges *weakly*, denote by $\pi(W(z_0))$ the limit operator. If the limit exists for each subsequence, it is not hard to prove that the operators $\pi(W(z_0))$ determine a Weyl system.

We shall call *regular* those representations of the Weyl algebra that induce as above a Weyl system. From the uniqueness theorem of representations of Weyl algebra we shall then derive the uniqueness theorem for the Weyl system.

A crucial feature of the Weyl algebra is *that it contains projections*. Indeed, setting

$$f_0(z) = (2\pi)^{-N} e^{-\frac{|z|^2}{4}} \tag{20}$$

and using Weyl's relation one obtains

$$W_{f_0} = W_{f_0}^*, \qquad W_{f_0} W_{f_0} = W_{f_0} \tag{21}$$

Therefore W_{f_0} is a projection operator.

Since representation are homeomorphism, for every representation π, $\pi(W_{f_0})$ is a projection operator. It is not difficult to prove that *in the Schrödinger representation* it is the orthogonal projection on the vector $\phi_0 = Ce^{-\frac{x^2}{2}}$ where C is a normalization constant. Moreover from (20) one derives

$$W_{f_0} W_f W_{f_0} = W_f W_{f_0} \qquad \forall f \tag{22}$$

5 Construction of the Representations

We now construct the representations of the Weyl algebra. As we have seen in "Lecture 8: Properties of Free Motion, Anholonomy, Geometric Phase", every state ρ determines a representation and *every representation is obtained in this way*. Let us briefly recall the GNS construction.

A state ρ of a C^*-algebra $\mathcal{A}$ is by definition a linear positive functional continuous in the topology of $\mathcal{A}$. Every state induces a pre-Hilbert structure on $\mathcal{A}$ as follows

$$\langle a, b\rangle \equiv \rho(a^* b) \tag{23}$$

Quotient out the kernel of ρ and denote by $\mathcal{H}$ the Hilbert space obtained by completion. Denote by $\tilde{a}$ the equivalence class of a. If $b \in \mathcal{A}$ define an operator $\hat{b}$ on $\mathcal{H}$ as follows

$$\hat{b} \cdot \tilde{a} = \widetilde{ba}$$

The operator $\hat{b}$ is well defined since $\rho(a * b * ba) \leq \|b\|\rho(a * a)$. Moreover $\hat{b}$ is closable and bounded and extends to a bounded operator on the entire space $\mathcal{H}$; we shall denote it by the same name. From (23) follows $\hat{b}\hat{c} = \widehat{bc}$. The correspondence $a \mapsto \hat{a}$ provides therefore a representation of $\mathcal{A}$ by means of bounded operators on the Hilbert space $\mathcal{H}_\rho$.

We shall denote by π_ρ the representation induced by ρ and with $P \equiv \pi_\rho(W_{f_0})$ the representative of the projection operator W_{f_0} in this representation. For a generic C^*-algebra and a generic state the GNS representation is not faithful, but for the Weyl algebra *every representation obtained by this procedure is faithful*. This is due to the fact that the product of any two elements W_f and W_g of the Weyl algebra is an element of the Weyl algebra, apart from multiplication by a unitary element in the center.

Since the Weyl algebra is separable also the space $\mathcal{H}_\rho$ is separable. Let ω_j, $j = 1, \ldots, d_\rho$, be an orthonormal basis of $P\mathcal{H}_\rho$. Denote by $\mathcal{K}_i$ the subspace of $\mathcal{H}_\rho$ generated by the action of $\pi_\rho(W_f)$ applied to ω_i (for the sake of simplicity we omit the index ρ).

If we prove

$$\bigoplus_i \mathcal{K}_i = \mathcal{H}_\rho \tag{24}$$

it follows that the representation π_ρ decomposes in the direct sum of faithful irreducible representations each of which has ω_j, $j = 1, \ldots, d_\rho$ as cyclic vector. Indeed from

$$P\pi(W_f)P = \int dz f(z) exp\{-|z|^2/4\} P \tag{25}$$

follows that $\pi_\rho(W_f)$ vanishes if and only if $f = 0$. From

$$\langle \pi_\rho(W_f)\phi_i, \pi_\rho(W_g)\phi_j \rangle = \delta_{i,j} \int f(z)\bar{g}(z')exp\{i/2Im(z,\bar{z}')\}exp\{-1/2|z-z'|^2\}dzdz \qquad (26)$$

one derives that the representation π_ρ is the direct sum of irreducible representations each in the Hilbert space generated by the action of $\pi_\rho(W_f)$ on ϕ_i. They are all equivalent since the scalar product in (26) depends on the functions f and g *but not on the representation*. We have proved.

Theorem 2 All *irreducible* representations of the Weyl algebra are unitarily equivalent, and therefore they *are all equivalent to the Schrödinger representation.* As a consequence they are all regular, the operators $\pi(W(z))$ exist in every representation and define the same Weyl system. The map

$$z \mapsto \pi_\rho(W(z)) \qquad (27)$$

is a strongly continuous map of C^N in the unitary operators of $\mathcal{H}_\rho$.

The notation commonly used for the elements of the Weyl algebra is

$$W_f \equiv Op^W(f)$$

to stress that each element is associated to an L^1 function in phase space.

One can extend this notation also to the coordinate and write $Q = Op^W(q), \quad P = Op^W(p)$.

In the Schrödinger representation the elements of Weyl's algebra are compact operators; this is easy to verify because their integral kernels are known. It also follows from the fact the compact operators form the smallest C^*-algebra of operators that contains finite-dimensional projections and is closed in the uniform topology.

Remark that the construction of the Weyl system holds for any even-dimensional real vector space and for any symplectic form ω.

Notice however that *in the proof of uniqueness* of the irreducible representation we have used the Weyl algebra, and this has required the use of Lebesgue measure (to introduce L^1 functions). Lebesgue measure does non exist in R^∞ nor in this space exists a σ-continuous measure that is quasi-invariant (invariant modulo translations). Therefore the uniqueness theorem *does not hold* for a Weyl structure in a system with infinitely many degrees of freedom (e.g. in Relativistic Quantum Field Theory or in Quantum Statistical Mechanics).

6 Lifting Symplectic Maps. Second Quantization

In the Weyl system we shall call $\mathcal{K}$ *the base space* and shall call *representation space* the Hilbert space $\mathcal{H}$ on which $W(z)$ acts. In the previous Lecture we have remarked that the symplectic structure $w(z, z')$ is invariant for unitary maps in $\mathcal{K}$. From the

uniqueness theorem it follows therefore that there exists a correspondence Γ between the group of unitary operators $U(\mathcal{K})$ on the base space $\mathcal{K}$ and the group $U(\mathcal{H})$ of unitary operators on the representation space $\mathcal{H}$.

It is easy to see that this correspondence preserves weak continuity, and therefore it induces a correspondence $\partial\Gamma$ between generators, i.e. between self-adjoint operators on $\mathcal{K}$ and self-adjoint operators on $\mathcal{H}$. The application $\partial\Gamma$ extends by linearity to an application which we shall denote by the same name $\partial\Gamma : B(\mathcal{K}) \to B(\mathcal{H})$.

If one constructs Quantum Mechanics for one particle in the base space $\mathcal{K}$ the map Γ constructs Fock space for a system with an arbitrary number of identical particles; in this case the map Γ is called *second quantization*. It has functorial property.

In the next Lecture we will see if one requires that time evolution be described in $\mathcal{K}$ by a one-parameter group of linear maps *with positive generator*, the generator is uniquely determined.

In general the lift $\partial\Gamma$ of a representation of group does not lead to a representation unless the group is semi-simple. We have already seen this in the case of group of linear symplectic transformations, which under $\partial\Gamma$ is mapped into the Weyl group. If G is a group of linear symplectic transformations on $\mathcal{K}$ and $U(g) = \Gamma(U(g))$, in general the operators $\Gamma(U(g))$ do not provide a representation of G. This is due to the fact that $\Gamma(1)$ is in general not a multiple of the unit operators and therefore in general it does not commute with $\Gamma U(g)$ for an arbitrary element g. For the same reason $\Gamma(U(g)e^{i\phi})$ does not in general differ by a phase from $\Gamma(U(g))$ (but always differs by multiplication with a unitary operator).

If G is semi-simple one can choose unitary operators in $\mathcal{B}(\mathcal{K})$ in such a way that the resulting $\Gamma(U)$ give a representation of the group G on $\Gamma(\mathcal{K})$. We have seen an example in Wigner's theorem.

A simple example in which the lift of a representation of G on $\mathcal{K}$ to operators on $\Gamma(\mathcal{K})$ does not provide a representation of G is the Galilei group, which has time-translations as abelian invariant subgroup and therefore is not semi-simple. In this case the obstruction is an abelian algebra (a phase factor for irreducible representations) and one can obtain only a *projective representation*.

We treat first the case of the group of rigid motions in R^3 to establish notation. In this case the space $\mathcal{K}$ is $R^3 \oplus R^3$ with the natural symplectic structure and $\Gamma(\mathcal{K})$ is $L^2(R^3)$.

Denote by $j_k,\ k = 1, 2, 3$ the generators of the rotations around the axes. We use the Schrödinger representation and the notation

$$J = \partial\Gamma(j) = i\partial\Gamma(q) \wedge \partial\Gamma(p)$$

The analytic vectors of $N \equiv \partial\Gamma(J)$ are a set of analytic vectors for the J_k and in this domain the following relations are satisfied

$$[J_k, J_h] = i\epsilon_{k,m,m} J_m \tag{28}$$

since the rotation group is semisimple.

From (28) follows that each J_k commutes with $J^2 = \sum_k (J_k)^2$ and therefore for each value of k one can diagonalize simultaneously J^2 and J_k. With the notation

$$L_\pm \equiv J_1 \pm i J_2, \quad L_3 \equiv J_3 \quad L^2 \equiv J^2 \tag{29}$$

one has, on the analytic vectors of J^2

$$[L_3, L_\pm] = \pm L_\pm, \qquad L^2 = L_3^2 + L_3 + L_- L_+ \tag{30}$$

In the Schrödinger representation (and therefore in any other representation of the Weyl system) one has $e^{2\pi i L_3} = I$ (rotations around any one of the axis by an angle multiple of 2π is represented by the identity operator). Therefore the eigenvalues of L_3 must be a subset of the integers.

We shall denote by the symbol m the eigenvalues of L_3, with $g(l) \in N$ the eigenvalues of L^2 and with $|l, m\rangle$ the corresponding common eigenvalues. From (30) one derives

$$L_3 L_\pm |l, m\rangle = (m \pm 1 | l, m\rangle, \qquad L^2 L_\pm |l, m\rangle = g(l) L_\pm |l, m\rangle \tag{31}$$

and from (31)

$$L_\pm |l, m\rangle = \sqrt{g(l) - m(m+1)} |l, m \pm 1\rangle \tag{32}$$

From (27) noticing that $L_-.L_+$ is a positive operator (since $L_- = L_+^*$) one derives that $g(l)$ must have the form $g(l) = l(l+1)$ and that the joint eigenvalues of L_3 and L^2 are l, m with $m \in \{-l, \ldots l\}$.

In order to have more explicit formulas it is convenient to refer to the Schroedinger representation. On the domain of the harmonic oscillator one has

$$J_k f(x) = i \sum_{h,l} \epsilon_{k,h,l} x_h \frac{\partial f}{\partial x_l} \tag{33}$$

It is convenient to use the isomorphism $L^2(R^3) \simeq L^2(R^+) \times L^2(S^3, d\mu)$ (description in spherical coordinates) where μ is the invariant measure on the sphere of radius one. In these new coordinates the operators J_k take the form

$$\hat{J}_k = I \otimes J_k$$

where $\hat{J}_k$ has the same expression as J_k but now as an operator on $L^2(S^3, d\mu)$. Using spherical coordinates the common eigenvalues of J_3, J^2 to the eigenvalues m, l take the form $|l, m\rangle \equiv Y_l(\theta) e^{im\phi}$ where $Y_l(\theta)$ are the spherical harmonics.

We now treat briefly the Galilei group. This is a ten parameter Lie group; its defining representation, in the phase space for a material point of mass m is

1.
$$x \to x + a, \quad p \to p, \qquad x,\ p \in R^3,$$

2.
$$x \to x + vt, \quad p \to p + mv \qquad v \in R^3$$

3.
$$x \to Rx, \quad p \to Rp \qquad R \in O(3)$$

4.
$$x \to x, \quad p \to p \quad t \to t + \tau \quad \tau \in R$$

In this notation x is the cartesian coordinate, p is the momentum. The elements of the abelian subgroup (2) are called "boosts".

We shall denote with

$$P_k \quad B_k \quad J \quad T$$

the corresponding generators.

The subgroup (4) is the subgroup of time translations: it is an abelian non-compact invariant subgroup. Therefore the Galilei group is not semisimple and its representations by means of unitary operators are in general *projective representations*.

For each value of t the functions

$$k_i(x, p, t) \equiv p_i t - m x_i \tag{34}$$

generates the symplectic subgroup (2) with parameter $v \in R^3$. The operators k_m satisfy

$$\{k_i, p_j\} = m\delta_{i,j}, \qquad \{k_i, x_j\} = t\delta_{i,j} \qquad \{x_i, p_j\} = \delta_{i,j}$$

and together with

$$\{x_i, t\} = \{t, p_j\} = \{t, k_j\} = 0$$

define the structure of Galilei group as a Lie group.

Weyl quantization substitutes the functions q_i, p_i, k_i with the operators Q_i, P_i, K_i which satisfy the commutation relations (t and m are parameters)

$$[K_i, P_j] = im\delta_{i,j}, \qquad [K_i, x_j] = it\delta_{i,j} \qquad [x_m, P_j] = i\delta_{m,j} \tag{35}$$

The space $\mathcal{K}$ is now $R^3 \oplus R^1 \oplus R^3$ with the singular symplectic structure that is zero on R^1. The Euclidian space $R^3 \oplus R^1$ represents space-time. In the Schrödinger representation the representation space is the space of continuous functions of time

with values in $L^2(R^3)$. We prove that on this space we can only find a *projective representation* of the Galilei group.

Using (35) it is not difficult to give the explicit expression of the unitary operators $U_t(v,m)$ implementing the maps $x \to x + vt$, $p \to p + mv$. One has, for any function $g \in L^2(R^3)$

$$(U_t(v,m)g)(x) = exp\{-it\frac{mv^2}{2}\}exp\{-i(x.mv)\}exp\{i(p.v)\}g(x) \tag{36}$$

where $p_k \equiv -i\frac{\partial}{\partial x_k}$. From (36) one can verify that indeed this is a projective representation of the Galilei group. Notice that

$$2\sigma((0,mv),(tv,0)) = mt|v^2|.$$

7 The Magnetic Weyl Algebra

We describe now the *magnetic Weyl algebra*, a modification of the Weyl algebra useful for the description of particles in a magnetic field [4, 6, 7].

We have seen that the Weyl algebra is a structure adapted to *quantize* Hamiltonian Mechanics with phase space R^{2N} using the standard symplectic form. In classical Mechanics when treating charged particles in a magnetic field it may be convenient, instead of modifying the Hamiltonian through a redefinition of momentum (minimal coupling), to leave the Hamiltonian invariant and modify the Poisson brackets into *magnetic Poisson brackets*.

In the same way, in Quantum Mechanics, in the treatment of non relativistic particles interacting with an electromagnetic field, it may be convenient to make use of a modified form of Weyl algebra, the *magnetic Weyl algebra*. This permits often to clarify topological effects which are due to the presence of the magnetic field.

Let us briefly recall the Poisson structure associated to a symplectic manifold. In particular consider the configuration space $\mathcal{M}$ of a particle which we identify with R^N, $N \geq 2$. Each fiber of the tangent space is a copy of R^N and each fiber of the cotangent space can be also identified with R^N.

We shall denote by $q \equiv \{q_1, \ldots, q_N\}$ a system of coordinates relative to orthogonal axes and by $\{p_1, \ldots, p_n\}$ coordinates relative to orthogonal axes in the fibers of the cotangent space. A symplectic structure on $\mathcal{M}$ is a closed non-degenerate 2-form $\Sigma \in \Omega^2(\mathcal{M})$ (the space of 2-forms on $\mathcal{M}$). We suppose always that Σ has C^∞ coefficients with respect to the standard 2-form.

Remark that, being non-degenerate, Σ uniquely defines an isomorphism $\beta : \Omega^1 \to \Xi(\mathcal{M})$ (the fibered space of the vector fields on $\mathcal{M}$). Defining

$$\{f,g\}_\Sigma = \Sigma(\beta(df)\beta(dg)) \tag{37}$$

the symplectic manifold acquires a Poisson structure.

In the case $\mathcal{M} = R^N$ and without magnetic field the symplectic structure most commonly used is

$$\sigma : \Xi \times \Xi \to R, \quad \sigma[(q, p), (q', p')] = q'.p - q.p' \tag{38}$$

where $q,\ q' \in R^N$ are orthogonal coordinates in R^N and $p,\ p' \in R^N$ are coordinates relative to the axes parallel to $dq_1, \ldots, dq_N$.

We treat only the case of one particle in R^3 subject to an external magnetic field B; the case of several particles is analyzed in a similar way. The dynamics of a non-relativistic particle of mass m in R^3 in field of scalar potential V and subject to a magnetic field $B(q)$ is given by the Hamiltonian

$$H(p, q) = \frac{1}{2m}(p + eA(q))^2 + V(q), \qquad rot\ A = B \tag{39}$$

together with the symplectic form (37). Notice that the magnetic field defines a 2-form $\hat{B}$. In local coordinates $\hat{B}_{i,j}(q) = \frac{1}{2}\epsilon_{i,j,k}B_k(q)$ where $\epsilon_{i,j,k}$ is the totally antisymmetric Ricci symbol. This is an instance of Hodge duality in a manifold of dimension three.

The same dynamics can be equivalently described with the Hamiltonian $H'(p, q) = \frac{1}{2m}p^2 + V(q)$ but then the symplectic 2-form must be modified adding the closed 2-form $\hat{B}$. The resulting form is closed and non-degenerate and therefore defines a new Poisson structure.

When written using vector potentials this description has an intrinsic ambiguity: two vector potentials $A(q),\ A'(q) \in R^3, q \in R^3$, give rise locally to the same 2-form $\hat{B}(q)$ if $A'(q) - A(q) = \nabla\phi(q)$ where $\phi(q)$ is a C^1 function (local gauge invariance). The corresponding transformations form a group; it is represented by the addition of the functions $\phi(q)$. We shall call *gauge group* this group of transformations.

In local coordinates the magnetic field $B(q)$ is represented by the antisymmetric tensor $rotA$ and therefore corresponds to the antisymmetric 2-form $(\partial_i\ A_k - \partial_k\ A_i)dq_k \wedge dq_i$ which is by construction *invariant under local gauge transformations*.

This formalism can be extended to the case of generic smooth manifolds [2, 3] $\mathcal{M}$ and one can consider cases in which the magnetic field is represented by a 2-form which is closed but not exact. For example one can consider the case of the magnetic field of an infinite rectilinear wire with constant electric current. The corresponding 2-form is closed but not exact, and originates topological effects (Bohm-Aharanov effect).

In the phase space $R^3 \times R^3$ with natural coordinates $q_k,\ p_h, h, k = 1, 2, 3$, the equations of motion

$$\dot{q}_k = \frac{1}{m}p_k, \quad \dot{p}_k = \frac{e}{m}\sum_{h=i}^{3} B_{k,h}(q)\ p_h \tag{40}$$

are associated to the Hamiltonian (39) through the symplectic form $\sum_k dp_k \wedge dq_k$; they can also be associated to the Hamiltonian

$$H_0(p,q) = \frac{1}{2m}p^2 + V(q) \tag{41}$$

through the symplectic form

$$\sum_k dp_k \wedge dq_k + \frac{e}{m}\sum_{h,k=1}^{3} B_{i,k}\, dp_h \wedge dp_k \tag{42}$$

The corresponding Poisson brackets are

$$\{f,g\}_B = \sum_{h,k=1}^{3}\left(\frac{\partial f}{\partial p_k}\frac{\partial f}{\partial q_h} - \frac{\partial f}{\partial q_k}\frac{\partial f}{\partial p_h} + \frac{e}{m}B_{k,h}\frac{\partial f}{\partial p_k}\frac{\partial g}{\partial p_k}\right) \tag{43}$$

Notice that the equation of motion and the Poisson brackets (37) *are gauge invariant* (they only depend on the magnetic field) while the Hamiltonian under gauge transformations varies through the addition of a total derivative.

8 Magnetic Translations in the Magnetic Weyl Algebra

It seems preferable to introduce a quantization map invariant under gauge transformations. This leads to the introduction of the *magnetic translations* and to the *magnetic Weyl algebra* [13].

We want to find a symplectic transformations α that extends to phase space as symplectic transformations the maps $x \to x + q$, $x \in R^3$:

$$\alpha_x\{q,\ p\} = \{q + x,\ p + \tau_x(q,p)\} \tag{44}$$

The group property implies $\tau_{x+y}(q,p) = \tau_x(q,p) + \tau_y(q+x, p+\tau_x(q.p))$. The condition to be symplectic is

$$\alpha_x^*(\sigma_B) - \sigma_B = 0 \tag{45}$$

It is easy to show that this identity reads

$$T_{-x;i,j}(q,p)dq_i \wedge dq_j + S_{-x;i,j}(q - px)dq_j \wedge dp_i = 0 \tag{46}$$

where

$$T_{x;i,j}(q,p) = \frac{\partial}{\partial q_i}\tau(x,k) - \frac{\partial}{\partial q_k}\tau(x,j) + eB_{j,k}(q) - eB_{j,k}(q+x) \quad S_x = \frac{\partial}{\partial p_j}\tau_{x,i}(q,p) \tag{47}$$

From this one derives $\frac{\partial \tau(q,p)}{\partial p_k} = 0$, $k = 1, 2, 3$, and moreover

$$\frac{\partial}{\partial q_j}(\tau_{q,p})_k - \frac{\partial}{\partial q_k}(\tau_x(q,\, p))_j + eB_{j,k}(q) - eB_{j,k}(q+x) = 0 \tag{48}$$

To determine τ we must invert the differential relation (48) which can be written, taking into account that the magnetic field is (at least locally) an exact differential form

$$\frac{\partial}{\partial t}\alpha_{-tx}(q,\, p)_{t=0} = -(x, edA(q).x) \tag{49}$$

where the 1-form A satisfies $dA = B$ (in coordinates $(rotA)_{i,j} = B_{i,j}$).

The choice of A among the forms which satisfy $dA = B$ is the origin of the *gauge ambiguity*. Given A the solution of (49) is

$$\tau_x(q) = eA(q+x) - A(q) \qquad rotA \;= B \tag{50}$$

and is defined modulo the addition of the gradient of a function (gauge ambiguity). It follows that magnetic translations in the Schrödinger representation are *defined modulo a gauge transformation*.

Recall that in Hamiltonian Mechanics the *momentum* μ_Ξ associated by a symplectic form to a vector field Ξ (and therefore to an infinitesimal transformation in configuration space) is by definition the contraction of the symplectic form with the field Ξ. The *momentum map* associated to the magnetic translation is $\mu_A(q,\, p) = p - eA(q)$.

Gauss theorem implies that the integral of the 1-form A along a closed path is equal to the flux of the magnetic field across a surface that has the given path as boundary. The magnetic translations along a closed path may therefore generate a non trivial homotopy if the form which represents the magnetic field is closed but not exact. This will play a role in the quantization.

Weyl quantization in its algebraic structure does not pose serious problems since the symplectic form that defines the Weyl system is simply substituted by the magnetic symplectic form. Its *description in the Schrödinger representation* requires a choice of gauge, in accordance with the fact that the classical Hamiltonian (and therefore its quantum counterpart) depends on the choice of gauge. The representation of the state *will depend on this choice*, but all the expectation values of the observables are *gauge independent*.

The choice of a gauge in which to describe the Schrödinger equation corresponds in the geometric quantization scheme to a choice of a local Lagrangian manifold. In the presence of a magnetic field $B(x)$ the Moyal product takes the following form in dimension N [6]

$$(f \;*_B^\hbar\; g)(\xi) = (\frac{2}{\hbar})^{2N} \int d\eta \int d\zeta e^{-\frac{2i}{\hbar}\sigma(\eta,\zeta) - \frac{i}{\hbar}\int_{\mathcal{T}(q,y,z)} B(x;\eta,\zeta)d\xi} f(\xi-\eta)g(\xi-\zeta) \tag{51}$$

In (51) we have indicated with x, y, z the components in configuration space of the elements (ξ, η, ζ). Recall that for each value of x, $B(x)$ is a 2-form; we denote by $B(x; \eta, \zeta)$ its evaluation on elements η, ζ in the tangent space at x.

In (51) we have denoted by $\mathcal{T}(q, y, z)$ the projection on configuration space of the triangle $\mathcal{T}(\xi, \eta, \zeta)$ with vertices the vectors ξ, η, ζ i.e. the symplectic area of the triangle. This reference to the *symplectic area* is a feature common to Geometric Quantization [7, 11].

Only the projection on configuration space enters because the 2-form is modified only in that space; the particular form (51) is most naturally derived in its infinitesimal form and extended by parallel transport.

The product described by (51) is usually called *magnetic Weyl-Moyal product*. It is associative, non-commutative and satisfies

$$(f \ *_B^\hbar \ g)^* = (g \ *_B^\hbar \ f) \tag{52}$$

We can give (51) a more convenient form. Using Fourier transform it can be seen that

$$(f \ *^\hbar \ g)(\xi) = (\frac{2}{\hbar})^{2N} \int_\Xi d\eta \int_\Xi d\zeta e^{-\frac{i}{\hbar}\sigma(\xi-\eta,\xi-\zeta)} f(\eta) g(\zeta)$$

where Ξ is phase space and $\xi, \ \eta, \ \zeta \in T(\Xi) \equiv \Xi$. The *Weyl-Moyal magnetic product* is now

$$(f \ *_B^\hbar \ g)(\xi) = (\frac{2}{\hbar})^{2N} \int_\Xi d\eta \int_\Xi d\zeta e^{-\frac{i}{\hbar}(\sigma+eB)(\xi-\eta,\xi-\zeta)} f(\eta) g(\zeta) \tag{53}$$

If the magnetic field is of class C^∞ the Weyl-Moyal magnetic product is a map $\mathcal{S}(R^N) \times \mathcal{S}(R^N)$ to $\mathcal{S}(R^N)$, and can be extended by duality to a continuous map $\mathcal{S}(R^N) \times \mathcal{S}'(R^N)$ in $\mathcal{S}'(R^N)$ through

$$(F *_B^\hbar f, g) = (F, f *_B^\hbar f g), \quad (f *_B^\hbar F, g) * = (F, f *_B^\hbar g), \quad F \in \mathcal{S}'(R^N), \ \ f \in \mathcal{S}(R^N) \tag{54}$$

and, again by duality and with a limit procedure, to a continuous map $\mathcal{S}'(R^n) \times \mathcal{S}'(R^N)$ in $\mathcal{S}'(R^N)$ which satisfies

$$(F *_B^\hbar G, f)) = (G, F *_B^\hbar f), \quad F, \ G \in \mathcal{S}', \quad f \in \mathcal{S} \tag{55}$$

The resulting extended magnetic Weyl algebra is useful to compose quantum observables in a gauge invariant way. In the second part of the Lectures we will analyze the Wigner transform and there we will see that it is sometimes convenient to use a representation in term of functions on $\mathcal{M} \times \mathcal{M}$ rather then on $\mathcal{M} \times \mathcal{M}^*$ where $\mathcal{M}$ is configuration space.

In the case $\mathcal{M} = R^3$ that we are considering one obtains this more convenient form by taking the Fourier transform with respect to the second variables, keeping in mind that $p_k = i\hbar \frac{\partial}{\partial q_k}$. With the notation $\phi = (I \otimes \mathcal{F}) f$ where $\mathcal{F}$ is Fourier

transform, the multiplication law for the magnetic Weyl algebra becomes

$$(\phi *_B^{\hbar} \psi)(q,x) = \int_{R^3} dy \phi(q - \frac{\hbar}{2}(x-y), y))\ \psi(q + \frac{\hbar}{2}y, x-y) e^{-\frac{i}{\hbar}\Phi_B(q;x;y)} \tag{56}$$

where $\Phi_B(q;x;y)$ is the magnetic flux through the triangle with vertices in the points

$$\{q - \frac{\hbar}{2}x\ ,\ q - \frac{\hbar}{2}x + \hbar y\ ,\ q + \frac{\hbar}{2}x\} \tag{57}$$

This product *depends only on the magnetic field.*

If the 2-form representing the magnetic field is exact in any representation on the Hilbert $L^2(R^3)$ it is convenient to make use of a 1-form A related to the 2-form B by $B = dA$. In local coordinates the 1-form a is represented by a *vector potential* $A(x)$ related to B through $B(x) = rot\ A(x)$. This relation *has not a unique inverse* and the solutions differ from each other by a gradient (at least locally)

$$A'(x) = A(x) + \nabla\phi \tag{58}$$

where $\phi(x)$ is a scalar field. We shall see that the representations which correspond to different choices of A are all unitarily equivalent and we will give the unitary operator which implements the equivalence.

The need to introduce the vector potential has its origin in the fact that in the Schrödinger representation there exist unitary operators which commute with the elements of the magnetic Weyl algebra. This is the quantum counterpart of classical gauge invariance. The analysis of this problem could be done in the general context of projective representations of algebras defined by a *twisted product* (as are the Weyl algebra and the magnetic Weyl algebra).

In the following we consider only the special case of the Schrödinger representation of the magnetic Weyl algebra on the configuration space R^3. Let $A = A_k(x)dx_k$ be a 1-form in R^3, $dA = B$. In coordinates

$$B_{i,j} = h(x,y)_{k,j} dx_i \wedge x_j \qquad h(x,y)_{k,j} = \frac{\partial A_k(x)}{\partial x_j} - \frac{\partial A_j(x)}{\partial x_k} \tag{59}$$

If $x, y \in R^3$ define $\Gamma_A[x,y]$ to be the integral of the 1-form A on the segment

$$[a,b] \equiv \cup_{s\in[0,1]}[sx + (1-s)y] \tag{60}$$

Due to Stokes theorem, if we denote by $\Omega_B(q_1,q_2,q_3)$ the flux of the 2-form B across the triangle defined by the points q_1, q_2, q_3 one has

$$\begin{aligned}\Omega_B(q, q+\hbar x, q+\hbar x+\hbar y) \ &= \Gamma_A([q, q+\hbar x])\ \Gamma_A[q+\hbar x, q+\hbar x+\hbar y] \\ &\times(\Gamma_A[q+\hbar x+\hbar y])^{-1}\end{aligned} \tag{61}$$

Setting

$$\omega_b^\hbar(q,x,y) = e^{-\frac{i}{\hbar}\Omega^\hbar(q,x,y)} \quad \lambda_A^\hbar(q,x) = e^{-\frac{i}{\hbar}\Gamma_A([q,q+\hbar x])} \tag{62}$$

one obtains

$$\omega^\hbar(q,x,y) = \lambda_A^\hbar(q,x)\lambda_A^\hbar(q+\hbar x,y)\lambda_A^\hbar(q,x+y)^{-1} \tag{63}$$

Defining

$$[U_A^\hbar(x)\phi](q) = \lambda_A^\hbar(q,x)\phi(q+\hbar x), \quad V^\hbar(p)\hat{\phi}(k) = \phi(k+p) \tag{64}$$

one verifies that these unitary operators generate a representation, determined by A, of the magnetic Weyl algebra associated to the magnetic 2-form B.

Using (64) it is easy to verify that if one modifies the 1-form A in $A' = A + d\Phi$ where Φ is a sufficiently regular scalar field one obtains

$$e^{\frac{i}{\hbar}\Phi(q)}[U_A^\hbar(x)\phi](q) = [U_{A'}^\hbar(x)\phi](q)e^{\frac{i}{\hbar}\Phi(q)} \tag{65}$$

Notice that the magnetic Weyl algebra is described by the following relations

$$[\hat{Q}_j, \hat{Q}_k] = 0 \quad [\hat{Q}_k, \Pi_{A,j}^\hbar] = i\hbar\delta_{i,j} \quad [\Pi_{A,j}^\hbar, \Pi_{A,k}^\hbar] = i\hbar B_{k,j} \tag{66}$$

Denoting by $e^{iQ.p}$ the group that induces translations in Fourier transform and with

$$U_A^\hbar(x) \equiv e^{ix\Pi_A^\hbar} \equiv e^{-\frac{i}{\hbar}\Gamma_A([Q,Q+\hbar x])}e^{ix.P} \tag{67}$$

the group of magnetic translations in the configuration space one has

$$U_A^\hbar(x)U_A^\hbar(x') = \omega_B^\hbar(Q;x,x')U_A^\hbar(x+x') \tag{68}$$

and a unitary representation of the magnetic Weyl group is given by the unitary operators

$$W_A^\hbar(q,p) \equiv e^{-i\sigma(q,p;Q,\Pi_A^\hbar)} = e^{-\frac{i}{2}q.p}e^{-iQ.p}U_A^\hbar(x) \tag{69}$$

In a planar system with constant perpendicular magnetic field the relation (66) takes the form

$$[Q_1, Q_2] = 0 \quad [\Pi_1, \Pi_2] = \hbar B \quad [Q_k, \Pi_j] = \delta_{k,j} \tag{70}$$

Setting $K_j = P_j - \frac{1}{\hbar B}\epsilon_{j,k}Q_k$ one has

$$[K_j, K_m] = \frac{1}{\hbar B}\epsilon_{k,j}, \quad [\Pi_j, \Pi_k] = \hbar B\epsilon_{j,k} \quad [K_j, \Pi_k] = 0 \tag{71}$$

It is to be noted that the pair K_1, K_2 generates a Heisenberg algebra, and the same is true for the pair Π_1, Π_2. These two algebras commute. They can be represented in the space $L^2(R^2, dx_1 dx_2)$ by

$$K_1 \equiv \cdot x_1, \quad \Pi_2 \equiv \cdot x_2 \quad K_2 \equiv \hbar B\frac{\partial}{\partial x_1}, \quad \Pi_1 \equiv \frac{1}{\hbar B}\frac{\partial}{\partial x_2} \tag{72}$$

From (72) one can trace back [9] the symplectic transformation between this representation and the one in the space $L^2(R^2, dQ_1 dQ_2)$.

References

1. Gronewold, W. (1946). *Physica*, *12*, 405–460.
2. Hamilton, M. (2010). Locally toric manifolds. *American Mathematical Society Memoirs*, *207*(971).
3. Iwatsuka, A. (1983). *Journal of Mathematics Kyoto University*, *23*, 475–480.
4. Landsmann, J. (2002). *Contemporary Mathematics*, *315*, 8.
5. Mackey, G. (1998). *Contemporary Mathematics*, *214*, 91–109.
6. Mantoiu, M., & Purice, R. (2004). *Journal of Mathematical Physics*, *45*, 1394–1417.
7. Melgaard, M., & Rozenblum, G. (2003). *Communications in Partial Differential Equations*, *28*, 697–736.
8. Moyal, L. (1949). *Mathematical Proceedings of the Royal Philosophical Society*, *99*, 99–120.
9. Sakamoto, M., & Tanimura, S. (2003). *Journal of Mathematical Physics*, *44*, 5042–5069.
10. von Neumann, J. (1992). *Mathematical foundation of quantum mechanics*. Princeton: Princeton University Press.
11. Weinstein, A. (1994). *Contemporary Mathematics*, *179*, 261–270.
12. Weyl, A. (1932). *The theory of groups and quantum mechanics*. New York: Dutton.
13. Zak, J. (1968). *Physical Review*, *168*, 686–695.

Lecture 14: A Theorem of Segal. Representations of Bargmann, Segal, Fock. Second Quantization. Other Quantizations (Deformation, Geometric)

The lifting of positive generators of linear symplectic transformation by the Weyl algebra is uniquely determined; this is the content of the following theorem of Segal.

This poses one of the main problems in the construction of a relativistic Quantum Field Theory. In this theory the free field is defined as a Weyl system on a space of functions on in which is defined an infinite-dimensional linear representation of the inhomogeneous Lorentz group. Segal's theorem states that in a representation of the free field as Weyl's algebra in which the energy-momentum spectrum lies in the positive light cone, there is a unique choice of the generators of time-space translations (energy and momentum).

This is one of the main obstacles for the construction of a relativistic theory in which there are interactions and still one can speak of free fields at different times. One should then use different representation of the Weyl algebra at different times, and for a relativistically invariant interaction these representations are inequivalent. Therefore there is in general no unitary group that describes evolution in time.

Theorem 1 (Segal) *Let $\mathcal{K}$ a complex Hilbert space, W a Weyl system on $\mathcal{K}$. Let $\mathcal{H}$ the representation space W. Let A be an operator $\mathcal{K}$ with $A > 0$ and $Az \neq 0, \forall z \in \mathcal{K}$.*

Let ω be a cyclic vector for W. Suppose that there exists a one parameter group $\Gamma'(t)$ of unitary operators on $\mathcal{H}$ such that the following is true:

(a) $\Gamma'(t)W(z)(\Gamma'(-t)) = W(e^{iAt}z)$
(b) $\Gamma'(t)\omega = \omega \quad \forall t \in R$
(c) $\Gamma'(t) = e^{itH}, \quad H \geq 0$

Then there exist unique *a correspondence (second quantization) $\Gamma : U(\mathcal{K}) \to U(\mathcal{H})$ such that*

$$\Gamma(e^{iAt}) = \Gamma'(t) \tag{1}$$

Moreover for any *positive operator B on $\mathcal{H}$ one has $\partial\Gamma(B) > 0$ (recall that $\partial\Gamma(B)$ is the generator of the group $\Gamma(e^{iBt})$).*

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_14

Proof Let

$$f(u) = \langle e^{-uH} W(z)\omega, W(z)\omega\rangle \qquad u = s + it \qquad z \in K \tag{2}$$

The function f is bounded, holomorphic in $s > 0$, continuous in $s \geq 0$. Denote by Φ this space of functions and notice that they form an algebra. Weyl relations give

$$f(it) = e^{i/2Im(z_t,z)}\langle W(z_t - z)\omega, \omega\rangle, \qquad z_t = e^{-iAt}z$$

The function $g(u) = e^{-1/2(e^{-uA}z,z)}$ belongs to Φ therefore also $gf \in \Phi$ and one has

$$(fg)(it) = \langle W(z_t - z)\omega, \omega\rangle e^{-\frac{1}{2}Re(z_t,z)} \tag{3}$$

Substituting z with $-z$ one sees that also the function $\langle W(-z_t + z)\omega, \omega\rangle e^{-\frac{1}{2}Re(z_t,z)}$ belongs to Φ and its boundary value is $\bar{f}\bar{g}$. By taking adjoints, one can construct a function which is bounded, holomorphic in $s < 0$, continuous for $s \leq 0$ and has boundary value fg at $s = 0$.

We conclude that fg can be continued in the entire complex plane as an analytic function, and is therefore constant as a function of t. Evaluating this function at zero its value is seen to be $e^{-(z,z)/2}$; therefore

$$\langle W(z_t - z)\omega, \omega\rangle = e^{-|z_t - z|^2/4}$$

Since the kernel of A is the null vector, when z and t vary the vectors $z_t - z$ span a dense set in $\mathcal{K}$. Therefore for every z one has $\langle W(z)\omega, \omega\rangle = e^{-|z|^2/4}$. It follows that for every unitary $U \in \mathcal{B}(K)$ the map

$$\sum a_i W(z_i)\omega \mapsto \sum a_i W(Uz_i)\omega$$

is well-defined and isometric on $D \equiv \bigcup_z \{W(z)\omega\}$. By density the map extends to a unitary operator $\Gamma(U)$. By construction

$$\Gamma(U)W(z)\Gamma^*(U) = W(Uz), \qquad \Gamma(U)\omega = \omega$$

and therefore $U \mapsto \Gamma(U)$ is a representation of $U(K)$, continuous because

$$\langle \Gamma(U)W(z)\omega, W(z)\omega\rangle = e^{\frac{i}{2}Im(Uz,z)} e^{-|U(z)-z|^2/4} \tag{4}$$

In particular choosing $U = e^{-itA}$ one has for all z

$$\Gamma(e^{-iAt}W(z)\Gamma(e^{iAt}) = \Gamma'(t)W(z)\Gamma'(-t)$$

and this proves (1). To prove that if $B > 0$ as an operator on $\mathcal{K}$ then $\partial\Gamma(B) > 0$ it is sufficient to prove that if $B > 0$ then

$$\int \langle \Gamma(e^{-iBt})w, w'\rangle g(t)dt = 0, \qquad \forall w, w' \in \mathcal{K} \tag{5}$$

if $g \in L^2(R)$, $\hat{g}(p) = 0$, $p < 0$. By density it is sufficient to give the proof for $w = W(z)\omega$, $w' = W(z')\omega$. Notice that

$$\langle \Gamma(e^{-iBt})W(z)\omega, W(z')\omega\rangle = e^{-\frac{1}{4}(|z|^2+|z'|^2)+2(e^{-iBt}z,z')} \tag{6}$$

and the exponential map preserves positivity. Uniqueness follow from the cyclicity of ω. ♡

Segal's theorem can be extended, with a more complicated proof, to the case in which the self-adjoint operator $A \geq 0$ and zero is a simple eigenvalue.

1 Fock Space

We introduce now a representation of the Weyl algebra which is much used due to its simple structure. Most of the research paper in Quantum Field Theory make use of this formalism.

We shall formulate it in the Schrödinger representation for a system with a finite number N of degrees of freedom.

Let us consider in the Schrödinger representation in $L^2(R^N)$ the positive self-adjoint operator

$$\mathcal{N} = \sum_{k=1}^{N} \mathcal{N}_k, \quad \mathcal{N}_k = 1/2(P_k^2 + Q_k^2 - 1) = -\frac{1}{2}\Delta_k + \frac{1}{2}x_k^2 - \frac{1}{2} \tag{7}$$

where the operators P_k and Q_k satisfy the canonical commutation relations.

The operators $\mathcal{N}_k$ satisfy on the domain of $\mathcal{N}$ (a domain dense in $L^2(R^N)$) the following commutation relations with the operators Q_k, P_h

$$[\mathcal{N}_h, P_k] = \delta_{k,h} Q_k \quad [\mathcal{N}, Q_k] = -\delta_{h,k} P_k \tag{8}$$

The point spectrum of each of the operators $\mathcal{N}_k$ is non-degenerate and consists of the non-negative integers. For this reason it is often called *number operator* for the kth degree of freedom.

The eigenvector to the eigenvalue zero is $\frac{1}{2\pi}e^{-\frac{1}{2}x_k^2}$ and the eigenvector associated to the eigenvalue n is the nth Hermite polynomial.

All continuous and bounded functions of $\mathcal{N}$ belong to the Weyl algebra.
It is convenient to introduce the operators

$$\begin{aligned} a_k &= \tfrac{1}{\sqrt{2}}(\hat{Q}_k - i\hat{P}_k) \equiv \tfrac{1}{\sqrt{2}}(x_k + \tfrac{\partial}{\partial x_k})\ k = 1, \ldots, N \\ a_k^* &= \tfrac{1}{\sqrt{2}}(\hat{Q}_k + i P_k) \equiv \tfrac{1}{\sqrt{2}}(x_k - \tfrac{\partial}{\partial x_k})\ k = 1, \ldots, N \end{aligned} \tag{9}$$

These operators are densely defined and satisfy in the domain $D(\mathcal{N})$ the relations (that we shall still call *canonical commutation relation*)

$$[a_k, a_h] = [a_k^*, a_h^*] = 0, \qquad [a_h, a_k^*] = \delta_{h,k} \tag{10}$$

The operators a_k and a_k^* have a dense common domain of definition, are adjoints one of the other and on $D(\mathcal{N})$ satisfy

$$\mathcal{N} = \sum_{k=1}^{d} \mathcal{N}_k, \quad \mathcal{N}_k = a_k^* a_k, \quad [\mathcal{N}_k, a_h] = -a_h \delta_{h,k} \tag{11}$$

The spectrum of the operators a_k is the real axis, while a_k^* has empty spectrum. All have a dense set of analytic vectors (in particular the analytic vectors of $\sqrt{\mathcal{N}}$).

Definition 1 (*Number operator*) The operator $\mathcal{N}$ in (11) is called *number operator*; in the Schrödinger representation it coincides with the Hamiltonian of the harmonic oscillator in R^N.

The operators $\mathcal{N}_k$ are a complete system of commuting operators: any operator that commutes with all of them is a function of the $\mathcal{N}_k$. Therefore there exists a *natural isomorphism* of $\mathcal{H}$ with $(\ell^2)^{\otimes N}$ in which a complete orthonormal basis is given by the sequences of N non-negative integers.

To the sequence $\{n_1, \ldots, n_N\}$ corresponds an eigenvalue of $\mathcal{N}$ whose value is $\sum_1^N n_k$. To the non-degenerate eigenvalue 0 of $\mathcal{N}$ corresponds the sequence $\{0, \ldots, 0\}$, and this vector coincides with the cyclic vector ω (the product of the ground states of all one-dimensional harmonic oscillators).

From (9) one derives

$$\begin{aligned} a_k\{n_1, \ldots, n_N\} &= \sqrt{n_k}\{n_1, \ldots, n_k - 1, \ldots, n_N\}, \\ a_k^*\{n_1, \ldots, n_N\} &= \sqrt{n_k + 1}\{n_1, \ldots, n_k + 1, \ldots, n_N\} \end{aligned} \tag{12}$$

In view of (12) it is natural to call the operators a_k^* *creation operators* and the operators a_k *annichilation operators*.

The operators $(\mathcal{N}_k + I)^{-1/2} a_k$ and $(\mathcal{N}_k + I)^{-1/2} a_k^*$ are bounded operators of norm 1. In the Schrödinger representation for a system with N degrees of freedom in $L^2(R^N)$ one has

$$\{n_1, \ldots, n_N\} \to h_{n_1}(x_1) \cdots h_{n_N}(x_N) e^{-\frac{1}{2}|x|^2}, \quad |x|^2 = \sum_{k=1}^{N} |x_k|^2 \tag{13}$$

where h_i is the ith Hermite polynomial.

Definition 2 (*Fock representation*) The representation of the Weyl algebra in $\ell^2(Z^N)$ is called *Fock representation*. The space in which this representation is realized is called *Fock space*.

The functor Γ (second quantization) takes a particularly interesting form in the Fock representation.

If A is a complex-valued matrix of rank N we have

$$\Gamma(A) \equiv \{0, A, A \otimes I + I \otimes A, A \otimes I \otimes I + I \otimes A \otimes I + I \otimes I \otimes A, \ldots\} \tag{14}$$

In particular

$$\Gamma(0) = 1 \qquad \Gamma(I) = \mathcal{N} \qquad \Gamma(e^{-t}) = e^{-t\mathcal{N}}, \quad \lim_{t \to \infty} \Gamma(e^{-t})\phi = (\omega, \phi)\omega \tag{15}$$

Remark that the *explicit form* that Fock space takes in the Schrödinger representation depends on the choice of the basis. In the finite dimensional case a change of the basis is realized by a unitary transformation. In the infinite dimensional case it may lead to an inequivalent representation. We shall see later the condition on the change of basis in order to have equivalence.

Fock's representation is used in Quantum Mechanics in the context of Quantum Optics and in all cases where one wants to describe simultaneously states with different number of particles, although the number of particles is constant in time. A typical example is the treatment in Quantum Statistical Mechanics of the Grand Canonical Ensemble and of the Bose-Einstein Condensate. Fock space may be a useful instrument in describing the behavior of correlations among M particles when the number N of particles increases, in particular when $N \to \infty$.

In the case of a finite fixed number of particles Schrödinger's representation is in general more useful since it allows the use of techniques of Classical Functional Analysis (Sobolev inequalities, Lebesgue's dominated convergence, positivity preserving semigroups…) that do not have a *natural* counterpart in the Fock representation in its description through creation and destruction operators.

Fock's representation is mostly used when $\mathcal{K}$ is infinite dimensional and when the interaction does not preserve the number of particles; in this case difficulties may arise because the representation will depend in general from the basis chosen. We shall

see that a change of base associated to a linear map T cannot be realized by a unitary transformation (connects *inequivalent* Fock representations) unless $T = I + K$ with K an Hilbert-Schmidt operator.

In Relativistic Quantum Field Theory field theory if the space has dimension d one identifies usually $\mathcal{K}$ with the space $\mathcal{S}'(R^d)$ or with a Sobolev space $H^{-m(d)}$ where the value $m(d)$ depends on Sobolev embeddings.

The observables are associated to function in the dual space $\mathcal{S}(R^d)$) and can be *formally* be expressed as *smeared-out fields* $\phi(f)$. One has $\phi(f) = \int_{R^d} f(x)\phi(x)dx$ where $f \in \mathcal{S}(R^d)$ and $\phi(x)$ is an operator with values in $H^{-m(d)}$. The representation space is $L^2(H^{-m(d)}dG)$ where dG is a suitable Gauss measure. Gauss measures are defined as probability measures also in the infinite dimensional case. We will give some details when we discuss the real Bargmann-Segal representation.

It should be stressed that in infinite dimensions there are representations of the Weyl algebra *which are not of Fock type*.

2 Complex Bargmann-Segal Representation

For systems with a finite number of degrees of freedom Schrödinger's representation is based on Lebesgue measure. The Weyl algebra admits also representations in Hilbert spaces that are defined using Gauss measures; Gauss measures are probability measures and can be extended as σ-continuous measures in the infinite dimensional case. Since in R^∞ there are Gauss measures which are inequivalent to each other, we will have inequivalent representations of the Weyl algebra.

We limit ourselves here to the finite dimensional case and we now describe some of them because of their relevance also in the infinite dimensional case and in the study of the semiclassical limit of Quantum Mechanics.

Gaussian measures share with Lebesgue's the property of being cylindrical; they can be presented as inductive limit of gaussian measures on finite-dimensional spaces (conditional measures). This property allows for induction procedures.

Gaussian measures are characterized by their mean and covariance. In the finite dimensional in which we shall work, it is sufficient to consider *centered* gaussian measures; they have the form

$$d\mu_A = Ce^{-\frac{(x,Ax)}{2}} \quad x \in R^N \qquad A > 0 \tag{16}$$

where A is strictly positive $N \times N$ matrix and C is a numerical factor that provides normalization (the function identically equal to one has integral one). The matrix A is the *covariance* of the gaussian measure and has the property

$$\int x_i x_j d\mu_A = A^{-1}_{i,j}. \tag{17}$$

Notice that in the infinite dimensional case (Field Theory) (17) reads

$$\mu_A(\phi(f)\phi(g)) = (f, A^{-1}g) \qquad \mu_A(I) = 1 \tag{18}$$

We shall describe only briefly this representation because its relevance is more evident in the infinite dimensional setting where the Gaussian measure is called *weak normal distribution* on $\mathcal{S}'$ (in the terminology introduced by I. Segal).

The complex Segal-Bargmann-Segal representation diagonalizes the annihilation operators that we have introduced in the context of the Fock representation. It is related to the description of Quantum Optics through coherent states (eigenstates of the destruction operators).

The origin of the Bargmann-Segal representation can be traced to a remark by Fock who noticed that the commutation relations among a_k^* and a_h are satisfied by the operators z_k and $\frac{d}{dz_h}$ acting on a space $\mathcal{F}$ of entire functions. We shall come back to this point later; Fock's point of view the point is also related to the *Berezin-Toepliz quantization*.

To describe the complex Segal-Bargmann representation and to verify its equivalence with the Schrödinger representation, consider the isomorphism F of Hilbert spaces

$$L^2(R^N, dx) \to \mathcal{B} \equiv L^2(C_+^N, dG)_{an} \qquad \phi(x) \mapsto \psi(z) \tag{19}$$

where the target space is the Bargmann space $\mathcal{B}$ of functions analytic *in the sector* $z_n \geq 0, n = 1, \ldots, N$, which are square-integrable for the Gauss measure

$$(\frac{1}{\pi})^N e^{-\frac{|z|^2}{2}} \prod_{n=1}^{N} dx_n dy_n \tag{20}$$

The correspondence is given by

$$\mathcal{F}: \quad \phi(x) \mapsto \psi(z) \qquad \psi(z) = \mathcal{F}(\phi)(z) = (\frac{1}{\pi})^N \int_{R^N} e^{-1/2(z^2+|x^2|)+\sqrt{2}(z.x)} \phi(x) dx \tag{21}$$

In general, if Fock space is defined with the gaussian $C_\gamma e^{-\gamma|z^2|}$ the map is defined as [6]

$$\phi(z) \to \psi(z) = (\frac{\alpha^2\gamma}{2\pi^3})^{\frac{N}{4}} \int_{R^N} e^{\alpha x.z - \frac{\alpha^2}{4\gamma}x^2 - \frac{\gamma}{2}z^2} \phi(z) dz$$

The advantage of using $\mathcal{B}$ instead of the larger space $L^2(C^N, dG)_{an}$ (as in the Berezin quantization) is that in doing so one introduces a duality with the space $\mathcal{B}'$ obtained by complex conjugation and which is composed of functions analytic in the opposite sector. This anti-unitary involution J leads to a dual representation of the Weyl

algebra. The two representation are defined on two disjoint domains and can be considered as two commuting representations, or also the representation of an algebra and its commutant together with an anti-linear conjugation J. This is a modular setting. It allows to consider theories in which there is an anti-linear isometric map e.g. charge conjugation and is therefore adapted to set up in the infinite dimensional case a particle-antiparticle theory.

On the other hand, taking boundary values (functions analytic in a sector and continuous at the boundary are defined uniquely by their boundary values) one has a unitary map between $\mathcal{B}$ and the space $\mathcal{B}_{real}$ of functions on R^N which are square integrable with respect to the Gauss measure. This is the space on which we will define the real Bargmann-Segal representation; by uniqueness there is a unitary map between the complex and real Bargmann-Segal representations of the Weyl algebra.

Notice that we have not specified the mean and covariance that define the Gauss measure G. In the finite-dimensional case all Gaussian measures (including Lebesgue measure) are equivalent (in the sense that they all are related to Lebesgue measure by a Radon-Nykodim derivative). Therefore a specification is irrelevant and all representations of the Weyl algebra are equivalent. This is not the case in the infinite dimensional case; Gaussian measures can be inequivalent for different covariances and so are the corresponding representations.

In the complex Bargmann-Segal representation on a dense domain the operator $Fa_k^* F^{-1}$ is the operator of multiplication by z and $Fa_k F^{-1}$ is the operator $\frac{\partial}{\partial z}$.

Therefore in the complex Bargmann-Segal representation the operator *multiplication by z* corresponds in the Fock representation to a creation operator and the operator $\frac{\partial}{\partial z}$ corresponds to a destruction operator. The vacuum state corresponds therefore to the constant 1. The states with N particles correspond to homogeneous polynomials of order N in the variables $z_1, \dots, z_N$. They form a complete orthonormal basis in $\mathcal{K}_N \equiv L^2_{an}(C^N_+, dz^N)$.

In the complex Bargmann-Segal representation one sees that a creation operator cannot have eigenvalues because the equation $z\phi(z) = \lambda\phi(z)$ cannot be solved with $\phi(z)$ analytic. The spectrum of the destruction operators is the entire complex plane; in fact for every complex λ the equation $\frac{d\phi_\lambda}{dz} = \lambda\phi_\lambda(z)$ has the solution $\phi_\lambda(z) = e^{\lambda z}$, which is analytic and square integrable with respect to Gauss measure.

For $\lambda = 0$ the solution $\phi_0(z) = C$ corresponds to the vacuum in Fock space (and to the ground state of the harmonic oscillator in the Schrödinger representation). It is easy to see that the vectors

$$\frac{z_1^{m_1} \cdots z_n^{m_n}}{\sqrt{m_1! \cdots m_n!}} \qquad n = 1, \dots, N,\ m_n = 0, 1, 2, \dots \tag{22}$$

form a complete orthonormal basis in $\mathcal{K} = L^2_{an}(C^{\mathcal{N}}{}_+, dz^{\mathcal{N}})$.

Definition 3 (*coherent states*) The eigenvectors ψ_a of the destruction operators are called *coherent states* and play a major role in Quantum Optics. The vectors ψ_a are

not orthogonal (the operators a_k are not self-adjoint) but provide *an over-complete system* in the sense that any vector ϕ can be expressed as an integral over coherent states

$$\phi(z) = \int e^{z\bar{w}} d\mu(w) \qquad d\mu(w) = (\frac{1}{\pi})^n \psi(w) e^{-\frac{|z|^2}{2}} \prod dz_k \tag{23}$$

One sees from (23) that

$$(\frac{1}{\pi})^n \phi(w) e^{-\frac{|z|^2}{2}} \tag{24}$$

is a *reproducing kernel* in this representation.

The inverse transformation F^{-1} is given as follows

$$(F^{-1}g)(z) = \lim_{M\to\infty} \int_{|z|<M} \bar{A}(x,z) g(z) d\nu_n(z) \tag{25}$$

with $A(x,z) = e^{-\frac{1}{2}(z^2+x^2)+\sqrt{2}(z.x)}$. Remark that for each $x \in R^n$ one has $A(x,z) \in \mathcal{K}_n$ but *only for a dense subset of* $\mathcal{K}_n$ the integral in (25) is absolutely convergent for $|x| \to \infty$. For a generic vector in $\mathcal{K}_n$ convergence in (25) is understood in a weak (Cesaro) sense.

3 Berezin-Fock Representation

A (different) complex representation can be constructed on the space $\mathcal{F}(C^N, d\mu)$ of entire functions on C^N, square integrable with respect to the Gaussian measure $d\mu = e^{-\frac{1}{2}|z|^2} dz$. Again multiplication by z corresponds to creation operator and $\frac{d}{dz}$ corresponds to destruction operators.

This representation is often called *Berezin-Fock representation*. As already mentioned, Fock was the first to notice in the early 30s that a suitable complex-linear combination of the pair z, $\frac{d}{dz}$ satisfies the CCR; later Berezin described the representation by unitary operators, i.e. in a form similar to the Weyl system, and proved equivalence to the Schrödinger representation in the case of a finite number of degrees of freedom.

We shall come back at length to the Weyl and Berezin-Fock representations in the second part of these Lecture Notes. In the case of an infinite number of degrees of freedom, in the context of Relativistic Field Theory the Berezin-Fock representation is also called *Wick representation*.

In the Berezin-Fock representation the function

$$K(z,w) = (\frac{1}{2\pi})^N e^{\frac{1}{2}(z,w)} \tag{26}$$

is a *reproducing kernel* i.e. for $f \in \mathcal{F}(C^N, d\mu)$

$$f(z) = \int_{C^N} K(z, w) f(w) d\mu \tag{27}$$

At the same time the function $K(z, w)$ is the integral kernel of the orthogonal projection P of $L^2(C^N, d\mu)$ onto $\mathcal{F}(C^N)$ (recall that the latter is composed of *entire functions*).

4 Toeplitz Operators

The Berezin-Fock representation allows the introduction of *Toeplitz operators* defined in this particular case by

$$T_f : \mathcal{F} \to \mathcal{F} \quad T_f(g) = P(fg) \qquad f \in L^2(C^N, d\mu) \tag{28}$$

where f is entire and P is the projection to the entire part of the function fg.

These operators (see e.g. [1, 2]) have a great relevance in Operator Theory and in Quantum Mechanics; their relation with the Weyl operators in the Bargmann-Segal representation is given by

$$T_f = W_{\Theta f} \qquad \Theta f(z) = (\frac{1}{\pi})^N \int_{C^N} e^{-|z-w|^2} dw \tag{29}$$

The difference between the operators T_f and W_f comes from the difference in the space of representation; the map Θ is an isometry between the Berezin-Fock space $\mathcal{F}$ and the Bargmann-Segal space $\mathcal{B}$.

The Berezin-Fock representation plays an important role in the Berezin-Wick quantization; we will discuss it brefly in the second volume of these Lecture Notes. In this quantization the role of the *creation operators* is taken by Toeplitz operators. One should remark that the correspondence between $L^2(R^N, dx)$ and the space of function over C^N used in the Berezin-Fock representations can also be seen in terms of the symplectic structures:

$$\omega_N \equiv \sum_{k=1}^{N} dq_k \wedge dp_k \qquad i\mu_k \equiv \sum dz_k \wedge \bar{z}_k \tag{30}$$

defined respectively on $R^{2N} \equiv C^N$ and on C^{2N}. Notice that $D_N \equiv \{\{z, w\} \in C^{2N}, w = \bar{z}\}$ is a Lagrangian sub-variety with respect to the 2-form $\sum dz_k \wedge dw_k$. The corresponding symplectic reduction $C^{2N} \to C^N$ is given by

$$\{z, \bar{z}\} \mapsto z \tag{31}$$

The linear symplectic transformation $(T^*R^N, \omega_n) \to (C^N, \mu_N)$ has as generating function $\Phi(x, z) = -i\ log(x, z)$. One can indeed verify the following identities

$$p_k = -\frac{\partial \psi}{\partial x_k} \qquad w_k = \frac{\partial \psi}{\partial z_k} \qquad k = 1, \ldots, N \tag{32}$$

This symplectic transformation maps $\sum_{k=1}^{N}(p_k^2 + q_k^2)$ to $\sum_{k=1}^{N} z_k \bar{z}_k$ which is the classic counterpart of the map

$$\sum_{k=1}^{N} -(\frac{\partial}{\partial x_k})^2 + x_k^2 \mapsto \sum_k z_k \frac{\partial}{\partial z_k} \tag{33}$$

5 Landau Hamiltonian Constant Magnetic Field in R^3

We give a simple example of Berezin-Fock quantization.

Consider a particle in R^3 subjected to a magnetic field $B(x)$ oriented along a fixed axis which we take to be $\hat{3}$ [6, 11].

We have already introduced this system in "Lecture 13: Weyl System, Weyl Algebra, Lifting Symplectic Maps. Magnetic Weyl Algebra" when we described the magnetic Weyl algebra. Here we follow the approach of Berezin.

We shall start with the case in which the magnetic field has constant strength B_0 and is oriented along the third axis of a chosen reference frame.

The Hamiltonian, in suitable units, can be presented as

$$H_0 = -(\nabla + iA_0)^2, \quad A_0 = \frac{B_0}{2}\{-x_1, x_2, 0\} \quad B_0 > 0 \tag{34}$$

We have chosen a suitable gauge; one has $\nabla A_0 = B_0\hat{3}$. The notation H_0 is used to distinguish the case of constant magnetic field. This Hamiltonian corresponds to free motion along the axis $\hat{3}$; therefore we will consider only the motion in the $\{\hat{1}, \hat{2}\}$ plane. It is convenient to introduce the complex notation

$$z = x_1 + ix_2, \qquad \partial = \partial_z = \frac{1}{2}(\frac{\partial}{\partial x_1} - i\frac{\partial}{\partial x_2}) \quad \bar{\partial} = \partial_{\bar{z}} \tag{35}$$

It is also convenient also to introduce the operators $P_0^{\pm} = H_0 \pm B$. They describe in $L^2(R^3) \otimes C^2$ the dynamics a particle with magnetic moment 1 and spin $\frac{1}{2}$ under the influence of the magnetic field (often called Pauli system).

Define the operators

$$\mathcal{P} = -2i\bar{\partial}_z - A_0 \qquad \mathcal{P}' = -2\partial_z \tag{36}$$

These operators can be written, with $\psi = \frac{B_0}{4}|z|^2$ as

$$\mathcal{P} = -2ie^{-\psi_0}\bar{\partial}e^{\psi_0}, \qquad \mathcal{P}' = -2ie^{\psi_0}\partial e^{-\psi_0} \tag{37}$$

Notice that ψ_0 solves $\Delta\psi_0 = B_0$ i.e. is a potential for B_0. One verifies that

$$[\mathcal{P}, \mathcal{P}'] = 2B_0 I \tag{38}$$

i.e. that the pair $\mathcal{P}, \mathcal{P}'$ satisfy canonical commutation relations. One verifies also that the operators

$$P_+ = \mathcal{P}\mathcal{P}', \qquad P_- = \mathcal{P}'\,\mathcal{P} \tag{39}$$

are orthogonal projection operators with $P_+ + P_- = I$ and that the hamiltonian H_0 is now written

$$H_0 = \mathcal{P}\mathcal{P}' - BI = \mathcal{P}'\mathcal{P} + BI \tag{40}$$

(Landau hamiltonian). The vectors in the subspace $\mathcal{H}_-$ on which P_+ projects satisfy

$$\mathcal{P}u = 0 \Rightarrow e^{-\psi_0}\bar{\partial}(e^{\psi_0}u) = 0 \tag{41}$$

This implies that the function $f \equiv e^{\psi_0}u$ is entire analytic in C.

On the other hand the function $e^{-\psi_0}f$ belongs to $L^2(R^2)$. The vectors in the representation space are therefore in one-to-one correspondence with the elements of the space

$$\mathcal{F} \equiv \{f(z) : \ (\partial f(z))e^{-\frac{B_0|z|^2}{4}} \in L^2\} \tag{42}$$

which is Fock space.

From (38) one derives also that the spectrum of the operator H_0 is

$$Sp(H_0) = \{(2q+1)B_0, \quad q \in \mathbf{N}\} \tag{43}$$

and the multiplicity of each eigenvalue is infinite (because P_+ projects on a space of infinite dimension). One verifies the the projection operator P^k on the kth spectral subspace H_0 has kernel

$$K_k(z,w) = e^{\frac{1}{4}(wz-|z|^2-|w|^2)}\prod_{j=1}^{k}\mathcal{L}_k(\frac{1}{2}(z_j - w_j)^2) \tag{44}$$

where $\mathcal{L}_k$ is the Laguerre polynomial of order k

$$\mathcal{L}_k(y) = \frac{e^y}{n!}\frac{d^n}{dy^n}(e^{-y}y^n) \tag{45}$$

6 Non-constant Magnetic Field

We consider now briefly the case in which the magnetic field is not constant, but still directed along $\hat{3}$ [8, 12–14, 16].

The motion along $\hat{3}$ is still free motion, and we study only the motion on the $\{\hat{1}, \hat{2}\}$ plane. Define as before $z = x_1 + ix_2$, $\partial = \partial_1 - i\partial_2$ and introduce as before the potential ϕ_B solution of

$$\Delta\phi_B = B_0, \quad \Delta = -\bar{\partial}\partial \tag{46}$$

We consider $B(x)$ as a small perturbation of B_0 and write $B(x) = B_0 + \lambda b(x)$ with $b(x)$ of compact support and λ a small parameter. We write the solution of (46) as

$$\phi_B(z) = \frac{B_0}{4}|z|^2 + \phi_\lambda(z) \tag{47}$$

and the vector potential $A(x)$ as

$$A(x) = A_0(x) + \lambda a(x), \quad A_0(x) = \{-x_2 - \lambda a_2(x),\ x_1\} \tag{48}$$

Notice that if the flux Ψ of $b(x)$ (i.e. $\int b(x)dx$) is not zero, $a(x)$ cannot decay wit the distance R more than $\frac{1}{R}$ since $\int_0^{2\pi} a(R\hat{n}) \cdot n d\theta = \Psi$.

The Hamiltonian is now

$$H_\lambda = (i\nabla + A)^2, \quad A = A_0 + \lambda a \tag{49}$$

The parameter λ is small, and we can try to expand H in powers of λ. To first order there is a term $A_0 \cdot a$ which does not vanish at infinity if the total magnetic flux Ψ of $b(x)$ is not zero. Recall that $a(x)$ in that case cannot decay faster than R^{-1} since $\int_{S^2} a(x)d\omega = \Psi$ where we have denoted by $d\omega$ Lebesgue measure on the unit sphere.

One can prove that $H_\lambda - H_0$ is compact relative to H_0 and therefore by Weyl theorem the essential spectrum of H_λ is the same as that of H_0 i.e. $\{(2n+1)B_0\}$.

Also in this case one can define

$$\mathcal{P}_a = -2i\bar{\partial} - A = -2ie^{-\phi}\bar{\partial}e^{\phi} \qquad \mathcal{P}' = -2i\partial - \bar{A} = -2ie^{\phi}\partial e^{-\phi} \tag{50}$$

One has

$$[\mathcal{P}, \mathcal{P}'] = 2B_0 + 2b(x) \quad x \in R^2 \tag{51}$$

Notice that the algebra generated by $\mathcal{P}$ and $\mathcal{P}'$ is no longer the Heisenberg algebra. One can still define the operators

$$Z^+ = \mathcal{P}\mathcal{P}' \qquad Z^- = \mathcal{P}'\mathcal{P} \tag{52}$$

but they are no longer projection operators. Still one has

$$Z^+ = 2B_0 + 2b(x) \qquad H = \mathcal{P}\mathcal{P}' - (B_0 + b(x) = \mathcal{P}'\mathcal{P} + B_0 + b(x) \tag{53}$$

The spectrum of H is still contained in $[B, \infty)$ and the eigenspace for the lowest eigenvalue is the set of functions u for which $\mathcal{P}'u = 0$. This implies

$$\bar{\partial} e^{\psi} u = 0 \Rightarrow u = f e^{-\psi}.$$

Therefore the eigenspace to the lowest eigenvalue is made of entire functions which belong to $L^2(C)$ when multiplied by $e^{-\psi}$.

This is a Fock space relative to the (non-Gaussian measure) $e^{-\psi(z)}$ where $\psi(z)$ is the *potential* of $B_0 + b(z)$.

It is not easy now to construct a complete orthonormal set. In general the nth eigenfunction in this Fock representation is given by an entire function $\xi_n(z)$ (not a polynomial). And it is no longer easy to find the remaining part of the spectrum of H.

Since the perturbation is relatively compact, the points $\{(2n+1)B_0\}$ belong to the essential spectrum of H but in general they are no longer eigenvalues. The eigenvalues $\lambda_{n,k}$ have $(2n + 1)B_0$ as limit point when $k \to \infty$, and under suitable conditions $\lambda_{n,k}$ converge super-exponentially to $(2n + 1)B_0$ as $k \to \infty$ [14]. There is so far no complete theory to determine the location of the eigenvalues of H.

The corresponding Landau problem on a torus presents further difficulties since the requirement that the eigenfunctions be single-valued restricts B_0 to have quantized values of the flux across the torus.

The same problem occurs when one adds to the Landau hamiltonian a potential (scalar or vector) which is periodic. Let $\mathcal{C}$ a corresponding cell. The requirement that the eigenvector be single-valued requires also here that the flux of B_0 across $\mathcal{C}$ be quantized. In this case the fact that the $U(1)$ bundle over $\mathcal{C}$ (corresponding to the fact that on each point the phase of the wave function can be changed) can be non trivial leads to interesting topological problems.

7 Real Bargmann-Segal Representation

The real Bargmann-Segal representation is obtained using the isomorphism of Hilbert spaces (that depends on the positive matrix B)

$$L^2(R^N, dx) \to L^2(R^N, e^{-\frac{(x,Bx)}{2}} dx), \quad B > 0, \qquad \phi(x) \mapsto \psi(x) = C_B e^{\frac{1}{2}(x,Bx)^2} \phi(x) \tag{54}$$

where B is a strictly positive $N \times N$ matrix and C_B is a normalization constant that is such that the measure is a probability measure. This isomorphism induces on the canonical operators the map $X_k \mapsto X_k$, $P_k \mapsto P_k - i B_{k,h} x_h$.

The real Bargman-Segal representation has been used by I. Segal in the infinite dimensional case to represent the quantum fields as linear functions on spaces of distributions (in the same way as, in the finite dimensional case, the coordinates x_k are linear functions on R^N).

To give a historical perspective it is worth noting that the real Bargmann-Segal representation has been used by I. Segal for the quantization of the Klein-Gordon equation

$$\frac{\partial^2 u(t,x)}{\partial t^2} = \Delta u(t,x) - m u(t,x) \quad x \in R^d, \quad m \geq 0 \tag{55}$$

(for $m = 0$ this is the wave equation). This hyperbolic equation admits a unique real solution (in suitable function spaces) if one chooses as initial data at time $t = 0$ the (real) function $u(0,x)$ and its gradient $\nabla u(0,x) \equiv v(0,x)$. Notice that the solution is unique also in the space of distributions.

The initial data are pairs of real-valued functions that belong to H^1. On this space there exists a natural (and singular) symplectic (Poisson) structure defined by

$$\omega(f,g) = \frac{1}{2} \int [f(x) \nabla g(x) - \nabla f(x)\, g(x)] dx, \quad x \in R^d \tag{56}$$

This structure is defined for pairs $\{f, g\} \in H^1(R^d) \otimes H^1(R^d)$ and can be extended to pairs in $H^{\frac{1}{2}}$ or to couples $f \in H^1$, $g \in L^2$.

It is invariant for the hamiltonian flow defined by Eq. (55). The Hamilton function is the energy of the classical field. In this formulation the space K is the Hilbert space of pairs of functions

$$\{f(x),\ g(x)\} \qquad f(x) \in H^1 \quad g(x) \in L^2 \tag{57}$$

Second quantization in this case corresponds (see e.g. [15]) to the Schrödinger representation but has as configuration space, instead of R^N, the space of the real distribution-valued solutions of the Eq. (55). A Gaussian measure gives to this space a Hilbert space structure.

The connection with the Fock representation is seen by the following *formal argument*: chosen two bases $f_n, g_m \in \mathcal{S}$ orthonormal with respect to the L^2 scalar product, define the linear functions q_n, p_n on $\mathcal{S}'$ as coordinates on $\mathcal{S}'$ by

$$Q_n = \int f_n(x)\phi(x)dx, \qquad P_m = \int g_m(x)\pi(x)dx \tag{58}$$

where $\phi(x)$ and $\pi(x)$ are (tempered) operator-valued distributions. Notice that the Fourier transform is an isometry of $\mathcal{S}'$.

In the real wave representation $\{Q_n\}, \{P_m\}$ are (self-adjoint) operators which satisfy, on a suitable domain, the relations

$$[Q_n, P_m] = i\delta_{n,m}, \quad [Q_n, Q_m] = [P_n, P_m] = 0, \quad n, m = 1, 2, \ldots \tag{59}$$

The representation space is the space of tempered distributions, or one of its subspaces, e.g. the Sobolev space H^{-m}; the allowed values for m depend on the dimension of the space-time in which the Klein-Gordon equation is written (for a detailed analysis see [4]). There is a lower bound for this choice but this lower bound is not a minimum; there does not exists a minimal choice of space.

With this choice the measure space is a space of trajectories of Brownian motion with values in a space of distributions (rather then in R^n as usual Brownian motion). We shall see an example of the non-uniqueness of the measure space in the second part of these Lecture Notes, when we shall discuss Brownian motion and the Ohrstein-Uhlembeck process (that can be considered as a Quantum Field Theory in zero-dimension space). With a proper choice of Gaussian measure one gives measure 1 to the set of continuous functions.

The symplectic (Poisson) structure on the pair of functions f, g is given in (57).

The initial data of the solution of the Klein-Gordon equation determine a *Lagrangian manifold* (a polarization) in the space of distributions [10]; this manifold evolves in time according to the equation of the classical field. In this respect the real Bargmann-Segal representation is of Lagrangian rather than Hamiltonian origin.

The same remarks hold for the complex Bargamann-Segal representation for a system with an infinite number of degrees of freedom.

The representation space has as coordinates (linear functionals) $\xi(z)$, $z = \{f \in \mathcal{S}, g \in \mathcal{S}\}$ [4]. Remark that a Gaussian measure is completely characterized by the mean and the variance; but even if the configuration space has finite dimension *there does not exists* a privileged choice of measure space.

The problem of choosing a measure space with suitable properties is present in the theory of Relativistic Quantized Fields if one wants to introduce polynomials and the exponential functions which can be *informally written* as $\phi(x)^p$ and $e^{i\phi(x)}$. These expression are informal since "the coordinate $\phi(x)$" is a distribution.

One may e.g. be interested in the evaluation of the coordinates on regular functions, i.e. informally to give meaning to

$$\int \phi(x)f(x)dx, \int \pi(x)g(x)dx \quad \phi,\ \pi \in \mathcal{S}', \quad f, g \in \mathcal{S} \tag{60}$$

One can ask whether it is possible to choose $\phi(t,x)$, $\pi(t,x)$ as elements of a more regular function space, for example a space of function continuous in time (and distributions-valued in space) in such a way that it is possible to define fields at a fixed time (notice that the delta function is in the dual of continuous functions).

In Relativistic Field Theory this depends on the dimension of the space, since it is linked to the condition under which the immersion of H^k in L^2 is a Hilbert-Schmidt operator, where

$$H^k = \{\phi(x) \in L^2(R^d), (-\Delta + |x|^2 + 1)^k \phi \in L^2(R^d)\} \tag{61}$$

In the case of one space dimension one can introduce *quantum fields at a fixed time*. Notice that the gradient of a delta function is not in the dual of the continuous functions and therefore *time derivative of quantum fields at a fixed time* can be defined only as distribution. Therefore the free Hamitonian can be defined only in a weak sense; still the one-parameter group it generates can be properly defined as group of unitary operators in Fock space.

8 Conditions for Equivalence of Representations Under Linear Maps

We shall discuss now briefly the condition under which a *linear* symplectic map in the *base space* $\mathcal{K}$ of a second quantization map is realized by a unitary transformation *in the representation space* $\mathcal{H}$.

Consider the family of unitary transformations $z \to e^{iB}z$, $B = B^* \in \mathcal{B}(\mathcal{K})$. If $\mathcal{K}$ is finite dimensional for each value of t the Gauss measure μ_A with covariance A is transformed into an equivalent one μ_D.

We consider only the case in which $D \equiv I$. The other cases are treated in the same way. Using in $\mathcal{K}$ the base of the eigenvectors of A (recall that A is a positive definite matrix) it easy to see that in the finite dimensional case the Radon-Nykodym derivative of μ_A relative to μ_I is

$$C(det(A^*A)^{-1})exp\{-Tr[(A^*A)^{-1} - I]\} \tag{62}$$

where C is a normalization constant.

This considerations are valid also in the case $\mathcal{K}$ has infinite dimension, *under the condition that* (62) *be well defined*.

As a consequence the two measures are certainly equivalent (and therefore the map is implemented by a unitary operator and the representations are equivalent) if the operator $(A^*A)^{-1} - I$ is trace class.

However this condition *is not necessary*. A necessary and sufficient condition is that $((A^*A)^{-1} - I)$ be of class Hilbert-Schmidt (therefore $A = I + B$ where B is a Hilbert-Schmidt operator). This can be seen as follows. Notice that any H.-S.

operator B can be written as the limit, in the H.-S. topology (a Hilbert space topology) of operators $B_N = \sum_{n=1}^{N} b_n \Pi_n$ where b_n are eigenvalues of B and π_n are one-dimensional projection operators.

It is easy to verify that the limit

$$\lim_{N\to\infty} e^{-Tr B_N} det B_N \equiv e^{Tr_1 B} \tag{63}$$

is finite.

Taking $B = (A^*A)^{-1} - I$ proves the statement.

In the real Bargmann-Segal representation this result has the following interpretation: two gaussian measures with densities formally written as

$$D_\lambda = C_\lambda e^{-\sum_n \lambda_n x_n^2}, \qquad D_\mu = C_\mu e^{-\sum_n \mu_n x_n^2} \tag{64}$$

where C is a *normalization* factor and λ_n (respectively μ_n) are the eigenvalues of the operators L (resp. M) are equivalent if and only if $L - M$ is a Hilbert-Schmidt operator.

If $L - M$ is not trace-class, the series $e^{\sum_k \lambda_k x_k^2 - \sum_k \mu_k x_k^2}$ does not converge in general on the support of the Gaussian measure with density measure D_λ; however there exists a sequence of real numbers c_n (connected to the normalization constants C_λ and C_μ) such that the series

$$\sum_n ((\lambda_n - \mu_n) x_n^2 - c_n) \tag{65}$$

converges almost surely with respect to the Gaussian measure with density D_μ. The sequence of positive numbers $c_1, c_2, \ldots$ provides a *renormalization* of the measure to make it into a probability measure.

9 Second Quantization

If the space $\mathcal{K}$ is infinite-dimensional, in particular if it is a Banach space of functions on which U acts linearly, the map Γ takes a special form. Suppose K has a nuclear structure, so it can be seen as a denumerable union of one dimensional complex Hilbert spaces $\mathcal{H}_k$ with a suitable topology to make the union a Banach space.

In this case each $\gamma(\mathcal{H}_k)$ has the structure of a Fock space for one degree of freedom and one can define a number operator N_k. Therefore $\Gamma(\mathcal{H}_k)$ has the structure

$$\Gamma(\mathcal{H}_k) = \Omega \oplus \mathcal{H}_k' \otimes (\mathcal{H}_k' \oplus \mathcal{H}_k')_{symm} + \ldots$$

and

$$\Gamma(\mathcal{H}') = \otimes \Gamma(\mathcal{H}_k'), \qquad \Omega \otimes \Omega \equiv \Omega$$

The spaces $\mathcal{H}_k'$ are dual to $\mathcal{H}_k$ in a suitable topology that allows to consider $\Gamma(\mathcal{H}')$ as a measure space in which the elements of $\mathcal{H}$ are measurable. Typically the measure spaces considered are Gaussian since then the product of elements of $\mathcal{H}_k$ is measurable.

The lift $\mathcal{K} \to \mathcal{H} \equiv \Gamma(\mathcal{K})$ is called *second quantization*, $\oplus_k \mathcal{H}_k$ are the *one-particle states* and $\partial\Gamma(\mathcal{K})$ are generators of linear transformation in $\mathcal{H}$ which are lifts of (linear) unitary transformations in $\mathcal{K}$.

For the quantization of the Klein-Gordon or Dirac fields one can choose as orthonormal basis $\{\phi_k, \ \psi_k\}$ in $\mathcal{K} \oplus \mathcal{K}$, e.g. initial data at time $t = 0$ of the solutions of the *linear* hyperbolic differential equation (Klein-Gordon or Dirac) and perform the Weyl quantization on each of the two-dimensional symplectic spaces corresponding to each value of the index. The group of linear symplectic transformations in $\mathcal{H}$ is generated by the *linear* equation and the generators are the quadratic energy-momentum operators.

The generators of the one parameter family of unitary transformations in the representation space $\mathcal{H}$ are then hamiltonians at most quadratic in the fields (free hamiltonians).

Hamiltonians that are not quadratic in the fields can be introduced later (again with same care) but they *are not obtained* by second quantization of generators of classical non-linear hyperbolic equation (such as the non-linear Klein-Gordon equation).

We have remarked that the one-parameter groups of symplectic linear maps on $\mathcal{K}$ are realized by one-parameter groups of unitary operators on $\mathcal{H}$. Suppose that the group is determined by the solution of the liner homogeneous equation

$$\dot{z} = JBz \tag{66}$$

where B is a symmetric matrix. Making use of the Fock representation it is easy to see that $\partial\Gamma(B)$ (the corresponding generator of the unitary group on $\mathcal{H}$) is the self-adjoint operator

$$\partial\Gamma(B) \equiv \sum_{k,j=1}^{2N} b_k^* B_{k,j} b_j \tag{67}$$

where

$$b_j = a_j \quad b_{j+N} = a_j^* \qquad j = 1 \dots N \tag{68}$$

The operator $\partial\Gamma(B)(\mathcal{N} + I)^{-1}$ is bounded and therefore the analytic vectors for $\mathcal{N}$ are analytic vectors for $\partial\Gamma(B)$. It is easy to verify that the following relation holds

$$[\partial\Gamma(B_1), \partial\Gamma(B_2)] = \partial\Gamma([B_1, B_2]) + Tr(B_1 B_2) \tag{69}$$

One should notice that the last term at the right of (69) is not present when one considers the generators of a group of linear symplectic transformations. The pres-

ence of this term in Quantum Mechanics is relevant in the quantization of linear symplectic transformations and is connected with the *Maslov index* that we will encounter in "Lecture 15: Semiclassical Limit; Coherent States; Metaplectic Group" (semiclassical analysis).

10 The Formalism of Quantization

Generally speaking, a *quantization* is a linear map that associates to a function f of a suitable class on a function space X (phase space) an operator Q_f on a Hilbert space $\mathcal{H}$. If the function is real-valued it is required that Q_f be selfadjoint.

Only *some* of the functional relations are preserved under quantizaton. We will see that there is *no quantization that preserves all functional relations*.

We have defined as *Weyl quantization* the map that to each functon f (of a suitable class) on phase space $R^d \times R^d$ associates the operator $Q^W(f)$ acting on $L^2(R^d)$. From now on the we use standard notation is $Op^W(f)$.

Other quantizations have been constructed (e.g. the Berezin-Wick and Toepliz quantizations) for which the operators act on a Hilbert space of analytic functions.

11 Poisson Algebras

Definition 4 (*Poisson algebra*) A *Poisson algebra* (or *Poisson structure*) is a triple $(X, \times, \{.,.\})$ where X is a real vector space, $\times$ is a bilinear associative and commutative map $X \times X \to X$ (called product) and $\{.,.\}$ is an antisymmetric map from $X \times X$ to X that is for every $f \in X$ a derivation both with respect to $\times$ and with respect to $\{.,.\}$. Therefore for each three elements $f,\ g,\ h\ \in X$ one has

(i) $f \times g = g \times f, \quad (f \times g) \times h = f \times (g \times h)$
(ii) $\{f, g\} = -\{g, f\}$
(iii) $\{f, g \times h\} = \{f, g\} \times h + f \times \{g, h\}$ (Leibnitz rule)
(iv) $\{f, \{g, h\}\} = \{\{f, g\}h\} + \{g, \{f, h\}\}$ (Jacobi identity)

Notice that (iii) and (iv) are the requirement that the Poisson bracket act *as a derivation with respect to both product structures*.

A *Poisson manifold* is a smooth manifold $\mathcal{M}$ that admits on $C^\infty_{loc}(\mathcal{M})$ a Poisson structure in which the product is the the standard product of functions on $\mathcal{M}$.

12 Quantization of a Poisson Algebra

Often X is the space of all functions on a symplectic manifold $\mathcal{M}$; in this case one applies the quantization procedure only to a subalgebra $X_0 \subset X$ (for example to the algebra of functions of class C^∞_0 on $\mathcal{M}$).

Definition 5 (*strict quantization of a Poisson algebra*) Let X be a Poisson Algebra densely contained in the self-adjoint part of an abelian algebra $\mathcal{A}_0$. If I is a subset of R^+ which has zero as only accumulation point, a *strict quantization* of the Poisson algebra $(X, \times, \{.,.\})$ is a family of maps $\mathcal{Q}^\hbar$, $\hbar \in I$, from $\mathcal{A}_0$ to the real elements of a family $\mathcal{A}^\hbar$ of C^*-algebras, with norm $\|.\|_\hbar$, which satisfies the following conditions

(a) *linearity*: $\mathcal{Q}^\hbar$ is linear for each value of $\hbar$ and $\mathcal{Q}^0$ is the inclusion.
(b) *Rieffel condition*: if $a \in \mathcal{A}_0$ the map $I \ni \hbar \to \mathcal{Q}^\hbar(a) \in R^+$ is continuous.
(c) *von Neumann condition*: for $a, b \in \mathcal{A}_0$ one has

$$\lim_{\hbar \to 0} \|\mathcal{Q}^\hbar(a) \otimes_J \mathcal{Q}^\hbar(b) - \mathcal{Q}^\hbar(a \times b)\|_\hbar = 0 \tag{70}$$

where the suffix J denotes the Jordan product

$$\mathcal{Q}^\hbar(a) \otimes_J \mathcal{Q}^\hbar(b) \equiv \frac{1}{2}[\mathcal{Q}^\hbar(a)\ \mathcal{Q}^\hbar(b) + \mathcal{Q}^\hbar(b)\ \mathcal{Q}^\hbar(a)] \tag{71}$$

(d) *Dirac condition*: for $a, b \in \mathcal{A}_0$ one has

$$\lim_{\hbar \to 0} \|\frac{1}{2\hbar}[\mathcal{Q}^\hbar(a)\ \mathcal{Q}^\hbar(b) - \mathcal{Q}^\hbar(b)\ \mathcal{Q}^\hbar(a)] - \mathcal{Q}^\hbar(\{a, b\})\| = 0 \tag{72}$$

(e) *completeness condition*: $\mathcal{Q}^\hbar(\mathcal{A}_0)$ is dense in $\mathcal{A}^\hbar_{real}$ for $\hbar \in I$.

The notation *strict quantization* is introduced to distinguish it from the *formal quantization* obtained by formal power expansion in the parameter $\hbar$.

Definition 6 (*complete quantization of a Poisson algebra*) A *complete quantization of a Poisson algebra* $\mathcal{M}$ is a choice of a subalgebra $\mathcal{A}_0$ of $C^\infty(\mathcal{M}_r)$ and a strict quantization of this subalgebra.

Under favorable circumstances the liner maps $\mathcal{Q}^\hbar$ are morphisms for each value of $\hbar$ and define for each value of $\hbar$ a structure of modified product.

Consider the important special case in which the Poisson structure is realized in a space of functions (e.g. C^∞) on the classical phase space $T^*(R^d)$ and the corresponding quantum structure is realized by means of self-adjoint operators on $\mathcal{H} = L^2(R^d)$. Analogous considerations can be done in the case $X = T^*(T^d)$.

Let $\{x_k\}$ be cartesian coordinates in R^d. To simplify notations we write $\mathcal{Q}^\hbar(A) \equiv \hat{A}$ neglecting the parameter $\hbar$. We want to find a correspondence between classical observables A (real-valued functions on phase space) and quantum observables $\hat{A}$ (self-adjoint operators in a Hilbert space) which satisfies

(a) $A \leftrightarrow \hat{A}$ is linear
(b) $x_k \leftrightarrow \hat{x}_k$, where $\hat{x}_k$ is multiplication by x_k
(c) $p_k \leftrightarrow \hat{p}_k \equiv -i\hbar \frac{\partial}{\partial x_k}$

(d) the correspondence $A \leftrightarrow \hat{A}$ is such that if f is continuous then $\hat{f}(x) = f(\hat{x})$ and $\hat{f}(p) = (\mathcal{F}f)(\hat{x})$ where $\mathcal{F}$ denotes Fourier transform
(e) $L_\zeta \leftrightarrow \hat{L}_\zeta$ with $\zeta = (\alpha, \beta)$, $\alpha,\ \beta \in R^d$

The function L_ζ is the symplectic generator of the translations in phase space in the direction ζ and $\hat{L}_\zeta$ is the generator of the one-parameter group of unitary operators $t \mapsto W(t\zeta) = W_\zeta(t)$ defined by

$$(W_\zeta(t)\phi)(x) = e^{\frac{i}{2}(t\alpha,x)+\frac{1}{2}t^2(\alpha,\beta)}\phi(x + t\ \beta) \tag{73}$$

Notice that this is the one-parameter group associated to the direction z by the Weyl algebra. It is worth noticing that, through suitable limit procedures, (a), (e) imply (b), (c).

Through the correspondence $A \leftrightarrow \hat{A}$ linear symplectic transformations are mapped to unitary transformations. This is not true in general for non linear symplectic transformations, except the ones that obtained as lift of tranformation of coordinates in R^d.

One can prove that conditions (a), …, (e) *determine completely* the correspondence $A \leftrightarrow \hat{A}$.

Theorem 2 (van Hove [17]) *Let $\mathcal{G}$ the class of C^∞ functions on phase space that generate global one-parameter groups. Denote by $\Phi_g(t)$ the one-parameter group generated by $g \in \mathcal{G}$.*

There does not exist *a map $g \leftrightarrow \hat{g}$, with $\hat{g}$ self-adjoint such that*

$$\begin{aligned} \hat{p}_k &= i\,\tfrac{\partial}{\partial x_k} \\ \hat{x}_k &= \cdot x_k \\ (a\ g + b\ h)\hat{} &= a\hat{g} + b\hat{h} \\ (\{g, h\})\hat{} &= i[\hat{g}, \hat{h}] \\ \Phi_f(t)\ \Phi_g(t)\ \Phi_f(-s)\ \Phi_g(-t) &= e^{is\hat{f}}\ e^{i\hat{g}}\ e^{-is\hat{f}} e^{-it\hat{g}} \end{aligned} \tag{74}$$

13 Deformation Quantization, $*$-product

Let (Ω, ω) be a symplectic manifold and $C^\infty(\Omega)$ the space of smooth complex-valued functions on Ω. Denote by $C^\infty(\Omega)_\hbar$ the ring of all power series in a parameter $\hbar$ with coefficients in $C^\infty(\Omega)$.

Definition 7 (*quantization by deformation of the product*) A *quantization by deformation* associates to each value of the parameter $0 < \hbar \leq \hbar_0$ and to each $f \in C^\infty(\Omega)$ an element $a(f)$ of a C^*-algebra $\mathcal{A}$ (quantum observables) in such a way that the algebraic structure of the quantum observables converges *in a suitable sense* when $\hbar \to 0$ to the product structure described by the Poisson brackets.

Definition 8 ($*$-*product*) A $*$-*product* on Ω is a bilinear mapping

$$* : C^\infty(\Omega)_\hbar \times C^\infty(\Omega)_\hbar \to C^\infty(\Omega)_\hbar \tag{75}$$

defined for $0 < \hbar < \hbar_0$ (or on a subset $\mathcal{I}$ which has zero as accumulation point) with the following properties:

(i) $*$ is associative

(ii) there exist bilinear operators C_j as bilinear maps $C^\infty(\Omega) \times C^\infty(\Omega) \to C^\infty(\Omega)$ such that, for every $f,\ g \in C^\infty(\Omega)$, the following holds

$$f * g = \sum_{j=0}^{\infty} C_j(f, g)\hbar^j \tag{76}$$

(iii) the operators C_j satisfy

$$C_0(f, g) = f\ g \qquad C_1(f, g) - C_1(g, f) = \frac{i}{\hbar}\{f, g\} \tag{77}$$

where $\{f, g\}$ is the Poisson Bracket with respect to the 2-form ω and

$$C_j(\mathcal{I}, f) = C_j(f, \mathcal{I}) = f \qquad \forall j \geq 1 \tag{78}$$

We have denoted by $\mathcal{I}$ the function identically equal to one on Ω. The requirement (78) indicates that $\mathcal{I}$ is the identity element for the $*$-product.

A $*$-product defines a quantization, i.e. a map $C^\infty(\Omega)_\hbar \to \mathcal{A}$ by the rule

$$a_\hbar(f)a_\hbar(g) = a_\hbar(f * g) \tag{79}$$

Notice that the quantization is *always singular* when $\hbar \to 0$.

The $*$ product is called *of differential type* if the coefficients C_j can be expressed in local coordinates as

$$C_j(f, g) = \sum_{\alpha,\beta} \frac{1}{\hbar^j} c_{j;\alpha,\beta}(D^\alpha f)(D^\beta g) \qquad D^\alpha \equiv \frac{\partial^{\alpha_1+\ldots\alpha_n}}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_n}} \tag{80}$$

with some coefficients $c_{j;\alpha,\beta}$ that belong to $C^\infty(\Omega)$.

Definition 9 (*strict quantization, weak quantization*) A quantization procedure is called *strict quantization* if it is defined as a correspondence between real-valued function in phase space and self-adjoint operators operators in a Hilbert space. The Weyl quantization and the Berezin-Wick quantization are *strict quantizations* of the product structure defined on phase space by the Poisson brackets.

A quantization procedure is called *weak (or formal) quantization* if the sums in (76) and (80) are only formal or asymptotic with no control on their convergence in some Banach space topology.

Two $*$-products $*$ and $*'$ are called *equivalent* if there exists a sequence of linear operators V_j on $C^\infty(\Omega)$, $j = 0, 1, \ldots$, with V_0 a multiple of the identity such that

$$V_j(u *' v) = (V_j u) * (V_j v) \qquad \forall u, v \in C^\infty(\Omega) \tag{81}$$

Weak $*$-products exist on any symplectic manifold and their classes up to equivalence are classified by formal power series with coefficients in $\mathcal{H}^2(\Omega, R)$ (the second de Rham cohomology group). The $*$-product is not unique, even if one requires that it be strict.

14 Strict Deformation Quantization

Definition 10 (*strict deformation quantization*) A strict quantization is called *strict deformation quantization* if $\mathcal{Q}^\hbar(\mathcal{A}_0)$ is for each value of $\hbar$ a subalgebra of $\mathcal{A}^\hbar$ and the map is injective. If this is the case, we can define the product $\mathcal{A}_0 * \mathcal{A}_0 \to \mathcal{A}_0$ in such a way that

$$\mathcal{Q}^\hbar(A * B) = \mathcal{Q}^\hbar(A)\ \mathcal{Q}^\hbar(B) \tag{82}$$

Weyl quantization is a strict deformation quantization (corresponding to the *deformation* of the product of two functions given by the Moyal product). Another strict deformation quantization is Berezin-Wick quantization.

Any strict quantization $A \in C^\infty(\Omega) \mapsto Q(A) \in \mathcal{B}(\mathcal{H})$ defines (and is defined) by a $*$-product by the identity

$$Q(A)Q(B) \equiv Q(A * B)$$

A weak $*$-product leads to a *formal quantization*.

15 Berezin-Toeplitz $*$-product

There are $*$-products that are especially interesting because they are related to geometrical or analytic properties of Ω. For Kähler manifolds one example is given by the Berezin-Toepliz $*$-product related to the Berezin representation of the Weyl algebra.

Notice that Kähler manifolds admit locally a real *potential* $\Psi(z)$ with the property that the symplectic form can be written at least locally as $\omega_{i,j}(z) = \frac{i}{2}\partial_i \bar{\partial}_j \Psi(z)$.

Suppose $\Psi(z)$ *exists globally*. This is certainly the case if Ω is cohomologically trivial; otherwise in what follows one has to replace *functions with sections of lines bundles*.

Let $L^2_\hbar(\Omega)$ be the L^2 space on Ω with respect to the measure $e^{-\frac{\Psi}{\hbar}}\mu$ where μ is the Riemann volume element of Ω and $\hbar$ is a positive parameter. Consider the subspace $\mathcal{A}_\hbar$ of all *holomorphic functions* in $L^2_\hbar(\Omega)$ and let $P_\hbar : L^2_\hbar(\Omega) \to \mathcal{A}_\hbar(\Omega)$ be the orthogonal projection on the holomorphic part.

For $f \in L^\infty(\Omega)$ define *the Toeplitz operator* $T^\hbar_f$ associated to f by

$$T^\hbar_f F \equiv P_\hbar(f\ F) \qquad F \in \mathcal{A}_\hbar \tag{83}$$

In several cases, including $\Omega = C^d$, with the Kähler structure derived from the Euclidian metric, one can verify that one has an asymptotic expansion

$$\|T^\hbar_f\ T^\hbar_g - \sum_{j=0}^{N} \hbar^j T^\hbar_{C_j(f,g)}\|_{\mathcal{A}_\hbar} = O(\hbar^{N+1}) \tag{84}$$

as $\hbar \to 0$ for any $f,\ g \in K(\Omega)$ ($K(\Omega)$ is the space of C^∞ functions in Ω of compact support). It follows from (80) that the operators C_j define a *differential $*$-product* on Ω. This is the *Berezin-Toeplitz $*$-product*.

16 "Dequantization"

An interesting problem in the theory of quantization is the following. Let β be a quantization, let $H(q, p)$ be a Hamiltonian on R^{2d} and let $\beta(H) \equiv \hat{H}$. Let H be self-adjoint and let $U(t)$ be the corresponding unitary group.

Let A be a classical observable and let A_t be the classical evolution according to evolution

$$\frac{dA}{dt} = \{H, A\}$$

where H is a classical Hamiltonian. Suppose that $A_t \in D(\beta(H)), \forall t$.

A natural question to ask is what is the relation between $\beta(A_t)$ and $U(t)\beta(A)$ $U(-t)$, where $U(t) = e^{it\hat{H}}$ i.e. what is the obstruction to the commutativity of the following diagram

$$A \xrightarrow{\Phi_{cl}} A_t \xrightarrow{\beta} \hat{A}_t \xrightarrow{\Phi_q} (\hat{A})_t \xrightarrow{\beta^{-1}} A \tag{85}$$

where Φ_{cl} and Φ_q are respectively the flux associated to the classic Hamiltonian H and the one associated to the quantum Hamiltonian $\hat{H}$.

One can try to estimate the difference

$$\|\beta(A)_t - \beta(A_t)\| \tag{86}$$

or the difference

$$|[\beta(A)_t - \beta(A_t)]\psi| \tag{87}$$

for a suitable dense set of vectors ψ.

A dual problem, that privileges the role of Quantum Mechanics, is the estimate of

$$|\tilde{A}_t - A_t|_\infty, \qquad |\tilde{A}_t - A_t|_{L^p} \tag{88}$$

for a suitable class of operators, where the function $\tilde{A}$, *if it exists* is defined by

$$\beta(\tilde{A}_t) = \beta(A)_t \tag{89}$$

To obtain an estimate one introduces a small parameter $\hbar$, which codifies the difference between the two formalisms (classical and quantum), and requires that the quantization be defined for any $0 < \hbar \leq \hbar_0$. In particular we may want that (86–88) be infinitesimal in $\hbar$ and that

$$[\beta_\hbar(A), \beta_\hbar(B)] = \frac{1}{\hbar}\beta(\{A, B\}) + 0(1) \tag{90}$$

(recall that in the definition of the Weyl algebra there is a phase $e^{\frac{ic}{\hbar}}$).

In view of the singularity of the relation (90) it is unlikely (89) can be satisfied for all observables.

A weaker requirement would be that this correspondence exists only for evolution of states, i.e. that for any given classical Hamiltonian H there exist a quantum Hamiltonian $\hat{H}$ such that the evolution of a Liouville density under H is approximated by the evolution of the quantum state under $\hat{H}$. Or conversely if one start from a quantum Hamiltonian.

This requirement is much weaker, since the singularity when $\hbar \to 0$ may appear only in the phase of the vector that represents the state and therefore not in the state itself. Recall that in the evolution according to the free Schrödinger equation the phase of the vector oscillates as $e^{i\frac{p^2}{\hbar}}$ but the modulus is a continuous function of $\hbar$.

In general the requirement is satisfied only for two special classes of states, namely the WKB states and the coherent semiclassical states. We shall consider them in these Lectures.

In the second part of these Lectures we shall introduce a formalism that of *Wigner functions*. that allows to connect states in Quantum Mechanics to L^1 functions in classical phase space, but at the expense of allowing for functions that are not positive definite, and therefore do not correspond to Liouville densities, when $\hbar \to 0$.

17 Geometric Quantization

The *quantizations* that we have discussed so far are related to the Heisenberg formulation of Quantum Mechanics. In this sense, the quantizations of Bargmann-Segal and of Berezin reflect the description of Hamiltonian dynamics as motion in phase space [10].

Another form of quantization, *Geometric Quantization*, is related to the Schrödinger formulation and therefore is more directly connected to the Lagrangian aspects of Hamiltonian dynamics, e.g. the propagation of the stationary points of the action functional (which are Lagrangian manifolds); on the analytic side, the connection is to the Hamilton-Jacobi equations and conservation laws.

Notice that all quantizations have their root in the semiclassical approximation which we will describe in the next two Lectures.

The Geometric Quantization approach, initiated by B. Kostant and B. Souriau in the early 70s, has had a remarkable development in the following years and is still the object of intense research. It has particular interest for the WKB method (described in "Lecture 16: Semiclassical Approximation for Fast Oscillating Phases. Stationary Phase. W.K.B. Method. Semiclassical Quantization Rules") and the quantization of equations of *hydrodynamic type* such as the non-linear wave equation and the KdV equation describing the propagation of waves in shallow waters. Their "Heisenberg counterparts" are the non-linear harmonic oscillator and the Toda lattice (exponential interaction on a lattice).

For an introduction to Geometric Quantization on can consult [3, 9]; for later developments one can see [7].

Geometric quantization aims at constructing, starting with phase space of a mechanical system, a Hilbert space in which a quantum mechanical theory can be formulated. In "Lecture 16: Semiclassical Approximation for Fast Oscillating Phases. Stationary Phase. W.K.B. Method. Semiclassical Quantization Rules" we will consider a simple example of Geometric Quantization in the framework of the analysis of the semiclassical limit through the WKB method.

As remarked, the purpose of Geometric Quantization is to associate a Hilbert space $\mathcal{H}(\mathcal{M})$ to a symplectic manifold $\{M, \omega\}$ where ω is a closed 2-form. The Hilbert space $\mathcal{H}(\mathcal{M})$ is constructed (at least locally) with Lagrange sections of $\mathcal{M}$ (phase space) viewed as a fibered manifold on configuration space $\mathcal{V}$ with complex-valued fibers. The connection structure (curvature) is described locally by the symplectic 2-form ω.

Recall that, denoting by $\Gamma(V)$ the collection of smooth sections on $\mathcal{V}$, a *connection* ∇ is a map

$$\nabla : \ \Gamma(V) \to \Omega^1(M) \otimes \Gamma(V) \tag{91}$$

(Ω^1 is the collection of 1-forms σ on $\mathcal{M}$) which satisfies for any smooth function f

$$\nabla(\sigma_1 + \sigma_2) = \nabla\sigma_1 + \nabla\sigma_2, \qquad \nabla(f\sigma) = df \otimes \sigma + f\nabla\sigma \tag{92}$$

We shall assume that the fibered manifold is *locally trivializable* (reducible to a product manifold) by a change of coordinates in a neighborhood of each point of $\mathcal{M}$. It is then possible to represent in each point of $\mathcal{M}$ the connection by a 1-form Θ. With this notation the *curvature* of the connection is given by $\Omega = d\Theta$. A connection is *flat* if $\Omega = 0$. It is easy to verify that this is independent of the trivialization chosen.

The collection of all sections turns out to be *too large a set* for the construction of a Hilbert space. For example, if $\mathcal{M} = X \otimes T^+X$, where X is the configuration space of a mechanical system, if the connection is flat and if one takes the collection of all smooth sections, the resulting Hilbert space is $L^2(\mathcal{M}, dl)$ where l is Lebesgue measure on phase space, whereas the Hilbert space of Quantum Mecanics is $L^2(X, dx)$ where the integral is a Lebesgue integral in configuration space; notice that $L^2(X, dx)$ *is not a subspace* of $L^2(\mathcal{M}, dl)$.

It is then necessary to consider only a subset of sections in $\Gamma(V)$. This choice goes under the name of *polarization* (roughly speaking, a choice of one variable in a pair of conjugate variables).

Various polarization have been considered. One possible choice is the *Kähler polarization* which is determined by a choice of complex structure for $\mathcal{M}$ (since $\mathcal{M}$ is locally a symplectic manifold, its symplectic structure is equivalent locally to a complex structure). The Kähler polarization corresponds to the choice of holomorphic leaves.

A Kähler manifold has a local *potential* Φ such that the symplectic form has the structure $\omega_{i,k} = \sum_{i,k} \partial_i \partial_k \Phi(z)$, $z \in \mathcal{M}$. This corresponds locally to a harmonic potential in configuration space. The curvature corresponds to the symplectic form (so the the Lagrange manifolds are locally flat).

Remark that the symplectic manifold on which the quantization is performed is locally isomorphic to a fibered manifold and the fibers are classified by a *momentum map*.

The Hilbert space for the quantization is constructed as follows. We recall the definition of α-density, $\alpha > 0$. A 2-form ν can be seen as a linear function on the space of frames $\mathcal{E}$ (a *frame* is a choice of local coordinates) with the property

$$\nu(\mathcal{E}\ A) = \nu(\mathcal{E})\ detA, \quad A \equiv (a_{i,j}) \tag{93}$$

where A is the matrix that defines the local change of frame. We define α-*density*, $\alpha \in R^+$, a linear function on the frames that transforms under a change of frame as

$$\lambda(\mathcal{E}\ A) = \lambda(\mathcal{E})|detA|^\alpha \tag{94}$$

The linear general group acts transitively on the bases, therefore an α-density is determined by its evaluation on a fixed basis.

We denote with $|\Lambda|^\alpha\mathcal{E}$ the vector space of α-densities on $\mathcal{E}$.

Remark that $|\Lambda|^\alpha\mathcal{E}$ is a one-dimensional complex vector space. If Ξ is fibered over $\mathcal{M}$ one can define $|\Lambda|^\alpha\Xi$.

Definition 11

(a) $|\Omega|^{\alpha}\Xi$ is the vector space of regular sections of $|\Lambda|^{\alpha}\Xi$.
(b) $|\Omega|^{\alpha}_{c}\Xi$ is the vector space of the regular sections of $|\Lambda|^{\alpha}\Xi$ with compact support.

If $\mathcal{M}$ is a manifold we shall denote by $|\Omega|^{\alpha}(\mathcal{M})$ the α-densities on $\mathcal{M}$.
One has then a map

$$|\Omega|^{1}(\mathcal{M}) \to C, \qquad \sigma \mapsto \int_{\mathcal{M}} \sigma \tag{95}$$

that defines an integral over $\mathcal{M}$ and also a map $|\Omega|^{1/2}(\mathcal{M}) \to C$ that gives a scalar product

$$\langle \sigma, \tau \rangle \equiv \int_{\mathcal{M}} \bar{\tau} \cdot \sigma \tag{96}$$

Definition 12 We call *intrinsic Hilbert space* of the manifold $\mathcal{M}$ the completion of $|\Omega|^{1/2}(\mathcal{M})$ with respect to this scalar product. Notice that the density depends only on the curvature Ω and not on the polarization chosen.

To complete the quantization procedure one defines on this Hilbert space a complete set of operators by using the coordinates of the frame and the moment maps as generators of translations on the frame.

18 Bohr-Sommerfeld Quantization

For special systems there are other natural choices of the polarization. For example, in the case of the hydrogen atom, consider the set Ω_{-} of points in phase space in which the energy is strictly negative. Provide Ω_{-} with the standard symplectic structure. In this case the base manifold is compact, the symplectic fibers are smooth (except the origin) and one can consider only the *Bohr-Sommerfeld fibers* i.e. the fibers for which a globally flat section can be defined (so that a complete set of action-angle variables can be defined).

One proves [3] that the set of Bohr-Sommerfeld fibers is discrete, and this leads to the *Bohr-Sommerfeld Quantization*.

From the point of view of semiclassical analysis this can be interpreted as a procedure that replaces, for the construction of the Hilbert space, the space of function over Ω_{-} with a space of functions defined over the collection of all smooth Bohr-Sommerfeld orbits. Each Bohr-Sommerfeld orbit can be taken as a point in configuration space and the Hilbert space of the representation is the direct sum of the L^2 spaces over the orbits. The definition of Bohr-Sommerfeld fibers can be extended to other systems ad this extends the definition of *Bohr-Sommerfeld Quantization*.

For further details about this interesting field of research on can consult [3, 9, 12].

References

1. Berger, C., & Coburn, L. (1986). *Journal of Functional Analysis*, *68*, 273–299.
2. Berger, C., & Coburn, L. (1987). Toeplitz operators on Segal-Bergmann space. *Transactions of the American Mathematical Society*, *301*, 813–829.
3. Guilliemin, V., & Sternbeg, S. (1984). *Symplectic techniques in physics*. Cambridge: Cambridge University Press.
4. Gross, L. (1967). Potential theory in Hilbert space. *Journal of Functional Analysis*, *1*, 123–181.
5. Fock, V. (1928). *Zeitschrift fur Physik*, *47*, 446–448.
6. Folland, G. (1989). Harmonic analysis in Fock space. Annals of Mathematics Studies (Vol. 122). Princeton: Princeton University Press.
7. Hamilton, M. (2010). Locally toric manifolds american. Mathematical Societe Memoirs (Vol. 207. 971).
8. Iwatsuka, A. (1983). *Journal Mathematics Kyoto University*, *23*, 475–480.
9. Kirillov, A. (2001). Geometric quantization. Dynamical systems IV encyclopaedia of mathematical sciences (Vol. 4). Springer.
10. Landsmann, J. (2002). *Contemporary Mathematics*, *315*, 8.
11. Landau, L., & Lifshitz, E. (1977). Quantum mechanics 3rd (ed.). Oxford: Oxford Pergamon Press.
12. Onofri, E. (2001). *International Journal of Theoretical Physics*, *40*, 537–549.
13. Rozenblum, G., & Tashchiyan, G. (2008). *Commission Partial Differential Equation*, *3*, 1048–1081.
14. Rozemblum, G., & Sobolev, A. arXiv:0711.2158
15. Segal, I. E., et al. (1992). *Introduction to algebraic and constructive field theory*. Princeton: Princeton University Press.
16. Sakamoto, M., & Tanimura, S. (2003). *Journal Mathematics Physics*, *44*, 5042–5069.
17. van Hove, L. (1951). *Memoires de l'Academie Royale de Belgique, 26*, N. 6.

Lecture 15: Semiclassical Limit; Coherent States; Metaplectic Group

According to Bohr's correspondence principle when the energy levels are very closely spaced the description of the emission and absorption processes in Quantum Mechanics should approach the one given in Classical Mechanics.

More generally, including also Schrödinger's theory, we can say that the correspondence principle is *roughly* the following rule: the predictions of Quantum Mechanics and of Classical Mechanics should *almost coincide* when the value of Planck's constant is *much smaller then the value of all other measured quantities that have the dimension of an action*.

The quantity $\hbar$ is a physical constant with dimension of an action. In standard cms units it value is approximately equal to 10^{-8}. In this *macroscopic units* Planck's constant can be regarded as a small parameter and it is meaningful to look for effective equations for macroscopic bodies. In this lecture we look instead for equations for systems for which the Action is of the order $\hbar$ and still we want to find in their dynamics analogies with classical Lagrangian or Hamiltonian dynamics.

We find these analogies for two classes of wave functions. The functions in the first class have smooth modulus and a phase which has fast oscillations at scale $\hbar$. The second class is made of functions which have support of linear size $\sqrt{\hbar}$, therefore very small in natural units but very large with respect to $\hbar$. The support of the Fourier transform of the function in this class is very large, of order $\hbar^{-\frac{1}{2}}$. It is therefore convenient to introduce the Quantum Fourier Transform.

The Quantum Fourier Transform is defined in R^N by

$$\mathcal{F}f(p) = (\frac{1}{2\pi\sqrt{\hbar}})^N \int e^{\frac{i}{\hbar}(x.p)} f(x)dx \quad x,\ p \in R^N \tag{1}$$

The Quantum Fourier Transform of the functions in this class has very small support, of order $\sqrt{\hbar}$. When expressed in quantum Fourier transform the minimal uncertainty principe is $\delta_x \phi \delta_p \phi \geq \hbar$.

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_15

It is also convenient to write explicitly the parameter $\hbar$ in the Schrödinger equation

$$i\frac{\partial\phi}{\partial t} = -\frac{1}{2}\hbar\Delta\phi + \hbar^{-1}\, V\phi \tag{2}$$

In (2) we have divided both sides of the equation by $\hbar$ (which is legitimate since $\hbar \neq 0$).

From the point of view of mathematics the *semiclassical limit* is the study of the behavior when the parameter $\hbar$ tends to zero of the solutions of (2) *for these two classes of initial data*. The aim is to find the equations that are satisfied in the limit when $\hbar \to 0$.

Notice that the limit (2) is *very singular*: the coefficient of the derivative of highest order vanishes and the zero order term diverges. We must therefore expect that the limit exists *only for a suitable class of initial data* and when it exists it corresponds to an equation different from (1). For the class of initial data that we will consider the limit equations are the equation of Hamiltonian Mechanics in Hamilton's or Lagrange's form.

We have remarked that the class of initial data we shall consider contains two sub-classes:

(A) Functions well localized with support of linear size of order $\sqrt{\hbar}$ *both in position and in momentum* (the latter obtained by Quantum Fourier Transform). These is the class of *generalized coherent states* [1, 4, 5]. We will see that for this class of functions the limit equations are Hamilton equations in the phase space.
For functions which belong to class A one has

$$\frac{1}{\hbar}|(\nabla V)(x_0).(x-x_0)\phi_\hbar(x)|_2 = O(\frac{1}{\sqrt{\hbar}}) \qquad \hbar|\Delta\phi_\hbar(x)|_2 = O(1) \tag{3}$$

if $V(x) \in C^2$ in a neighborhood of x_0.
For class A solutions we shall estimate the difference between the solution of the Schrödinger equation and the wave function which is obtained by "moving" the initial datum keeping its barycenter on a hamiltonian trajectory and deforming its shape with the tangent flow.

(B) Functions with smooth modulus and phase which has oscillations of order $\hbar^{-1}$

$$\phi_\hbar(x,t) = \rho(x,t;\hbar)e^{\frac{i}{\hbar}S(x;t)} \tag{4}$$

with $\rho(x,t;\hbar)$ real and $S(x;t)$ real and sufficiently regular [2–4].
We shall call this states *W.K.B.* states, named after Wentzell, Kramer and Brillouin. For this class of states the equations satisfied in the limit $\hbar \to 0$ are the equations of Hamilton-Jacobi type and the conservation of Liouville density.
For class B solutions we shall estimate the difference between the solution of the Schrödinger equation and the wave function which is obtained by moving

the density according to Lagrange-Liouville's equation and the phase according to the stationary phase Lagrange condition.
The presence of class B states is suggested by the fact that if $V = 0$ the function

$$f_\hbar(x,t) = \int \phi(p) e^{\frac{i}{\hbar}[(x,p) - \frac{p^2}{2\hbar}t]} dp \tag{5}$$

is the solution of the Schrödinger equation with initial datum $f_\hbar(x,0)$. One sees from (5) that the phase of $f_\hbar(x,t)$ is a fast oscillating function as $\hbar \to 0$ and therefore this solution is a type B function.

1 States Represented by Wave Functions of Class A

We begin with an analysis of the states of type A. We will assume that the configuration space of Quantum Mechanics is R^N and the phase space of the classical system is $R^N \times R^N$ with the standard symplectic structure.

In the case of well localized functions this assumption is not essential since the analysis makes use only of the *local structure*. It can therefore be extended to a quantum theory based on a manifold Σ provided one substitutes the Laplacian with the Laplace-Beltrami operator and defines properly the Fourier transform. The phase space of the corresponding classical theory is $T^*(\Sigma)$.

In this more general case one seeks a correspondence between a solution of the Schrödinger equation and a classical trajectory in the space $T^*(\Sigma)$. This correspondence reflects the structure of the Heisenberg algebra generated by the local coordinates and the generators of geodesic motion on Σ.

In case Σ has a topological structure different from R^N one must pay attention in Quantum Mechanics to define correctly the Laplace-Beltrami operator and/or the boundary conditions.

In the second part of these Lectures we will give a more complete analysis, (through the Wigner transform and pseudo-differential operators), of the many interesting mathematical problems that are encountered in the study of the semiclassical limit of type A states. In this Lecture we provide a more elementary approach.

Definition 1 (*localization in phase space*) Let the phase space be R^{2N} with coordinates $z_k \equiv q_k$, $z_{k+N} = p_k$, $k = 1, \dots, N$. We will say that a function $\phi(z)$ is *well-localized* in a neighborhood of the point $z^0 \equiv \{q^0,\ p^0\}$ of phase space if there exists positive constants $c_1,\ c_2$ and $0 < \alpha < 1$ such that

$$|\phi|_2 = 1 \quad \int_{R^N} |\phi(x)|^2 (x - q^0)^2 dx \le c_1 \hbar^{2\alpha} \quad \int_{R^N} |\hat{\phi}(p)|^2 (p - p^0)^2 dx \le c_2 \hbar^{2(1-\alpha)} \tag{6}$$

◇

States well localized in phase space are the natural instrument to study the semiclassical limit. These states remain well localized near a classical trajectory both in position and momentum for any finite time T, with an error term which vanishes as $\hbar \to 0$ and increases with T. We shall provide a proof of this property (Theorems 1 and 6).

Write the Schrödinger equation in the form (2) and denote by

$$\gamma(t) \equiv \{q(t), p(t)\} \qquad q, p \in R^N \forall t \quad q(0) = q^0, \ p(0) = p^0 \tag{7}$$

the solution of Hamilton's equation for the hamiltonian $H_{class} = \frac{p^2}{2} + V(q)$.

We assume for simplicity that the trajectory of the classical system

$$\gamma(t) = \{q(t), p_(t)\} \in R^{2N}, \quad t \in R^1$$

with initial data $\gamma(0) = \{q^0, p^0\}$ is defined for all time $t \in R$.

Let $\phi^0(x, 0), \ x \in R^N$ the initial datum localised in a neighborhood of $\{q^0, p^0\} \in \gamma$. Denote by $\{q(t), p(t)\} \equiv \gamma(t)$ the solution of Hamilton's equations with hamiltonian H_{class} and initial data $\{q^0, p^0\}$.

We want to prove that for any fixed T the solution of the Schrödinger equation (1) with an initial datum ϕ_0 which is well localized in phase space at $q^0, \ p^0$ remains for $t \in [0, T]$ well localized in phase space around the point $\gamma(t)$ and its variance changes according to the quantization of the tangent flow at γ.

We will consider only the case in which the classical hamiltonian is

$$H_{class} = \frac{1}{2} p^2 + V(q) \qquad q, \ p \in R^N \tag{8}$$

with $V(q)$ smooth.

The trajectory γ is solution of Hamilton's equations is given in local coordinates by the solution of

$$\dot{\pi} = \frac{\partial V}{\partial q_i \partial q_k} \xi, \qquad \dot{\xi} = \pi \tag{9}$$

and the tangent flow is generated by the quadratic hamiltonian

$$\frac{1}{2} \sum_{i=1}^{N} p_i^2 - \frac{1}{2} \sum_{i,k=1}^{N} v_{h,k} q_h q_k \tag{10}$$

where $v_{h,k}$ is the Hessian of the potential at the point $\gamma(t)$.

Under suitable regularity assumptions we must prove

1. For any finite T the class of well localized states *is left invariant* by the quantum evolution in $[0, T]$ with hamiltonian $H = -\frac{1}{2}\Delta + V(x)$ (but the degree of localization may depend on time).

2. The solution of the Schrödinger equation with hamiltonian H and with initial datum $\phi_0 \in L^2(R^N)$ well localized at $q_0,\ p_0$ remains, uniformly in $t \in [0, T]$, clase in L^2 norm to a function well localized at time t at $\gamma_{q(t),\ p(t)}$ where $\{q(t),\ p(t)\}$ represent the trajectory solution of Hamilton's equations with hamiltonian $H(p, q) = \frac{1}{2}p^2 + V(q)$ and initial datum $q_0,\ p_0$.
3. The best approximation to the solution of the Schrödinger equation with hamiltonian H and well localized initial state is obtained by means of a family of unitary transformations $U_0(t)$ which solve the (time dependent) equation $i\hbar\frac{dU(t)}{dt} = H_2(t)U(t)$. Here $H_2(t)$ are (quantum) quadratic Hamiltonians obtained by quantization of the classical (time-dependent) quadratic hamiltonian that generate the flow tangent to the classical trajectory γ.
4. The error that is made with this approximation is of order $\hbar^{\frac{1}{2}}$ over finite interval of time (the error may increase in time).

Call $f_0^\hbar(x)$ the function that has initial data $\phi_0(x)$ and let $f^\hbar(\gamma_t(x))$ its evolution under Lagrangian flow associated to the Hamiltonian H_{class}. We must prove that for any finite value of T

$$sup_{0\le t\le T}\|f^\hbar(\gamma_t)(x) - e^{i\Theta_\hbar(t)}\phi^\hbar(x;t)\|_2 \le c(T)\hbar^{\frac{1}{2}} \tag{11}$$

for a suitable function $\Theta_\hbar(t)$. Recall that a change in phase in the wave function does not alter the quantum state. The estimate (10) follows if we can prove that for $t \in [0, T]$ one has

$$sup_{0\le t\le \tau}\|\frac{d}{dt}f^\hbar(\gamma_t)(x) - \frac{d}{dt}e^{i\Theta_\hbar(t)}\phi^\hbar(x;t)\|_2 \le c\hbar^{\frac{1}{2}} \tag{12}$$

and therefore

$$\|f^\hbar_{\frac{d\gamma_t}{dt}}(x) - i\frac{d}{dt}\Theta_\hbar(t)\phi^\hbar(x,t) - \frac{i}{2}\hbar\Delta\phi^\hbar(x;t) - \hbar^{-1}V(x)\phi^\hbar(x;t)\|_2 \le c(T)\hbar^{\frac{1}{2}} \tag{13}$$

The time T for which the approximation is valid to order $\hbar^\delta$ will be in general of order $\hbar^{-\frac{1}{2}+\delta}$ for $\delta > 0$.

2 Qualitative Outline of the Proof of (1), (2), (3), (4)

We give first a qualitative idea of the proof. The strategy is simple.

For well-localized states the term $\frac{\hbar}{2}\Delta\phi(x, t)$ in (12) is of order $O(1)$ in the limit $\hbar \to 0$ (recall that the localization in x is of order $\hbar^{1/2}$).

We write $\hbar^{-1}V(x)\phi$ as the sum of three terms plus a remainder

$$\begin{aligned}\hbar^{-1}V(x)\phi(x,t) &= (A_1 + A_2 + A_3 + R)\phi(x,t)\\ A_1 &\equiv \hbar^{-1}V(q(t))\\ A_2 &\equiv \hbar^{-1}(q(t)-x)\nabla V(q(t))\\ A_3 &\equiv \frac{1}{2}\hbar^{-1}\frac{\partial V}{\partial x_i \partial x_k}(q(t)(x_i - q_i(t))(x_j - q_j)\\ R &\equiv \hbar^{-1}\phi - A_1 - A_2 - A_3 \end{aligned} \tag{14}$$

The term A_1, singular in $\hbar$ of order $\frac{1}{\hbar}$, is multiplication by a term that does not depend on x. We choose $\Theta_\hbar(t)$ to cancel this term.

The term A_2 has a singularity of order $\frac{1}{\sqrt{\hbar}}$ and must be canceled by a term which has the same degree of singularity in the formal expansion in powers of $\hbar$ of the function $f^\hbar[\frac{d}{dt}\gamma_t](x)$. Notice that the L^2 norm of $f^\hbar$ is 1, but its H^1 norm is of order $\frac{1}{\sqrt{\hbar}}$.

The term $-\frac{\hbar}{2}\Delta\phi(x,t) + A_3$ is not singular when $\hbar \to 0$ but it does not vanish in this limit. We must prove that this term differs from the zeroth-order term in $\frac{d}{dt}\gamma_t(x)$ by terms which are infinitesimal in $\hbar$.

The remainder R is *formally* of order $\sqrt{\hbar}$.

It is easy to show that the functions of class C^∞ and of fast decrease at infinity are a dense common domain for the the generators of the Hamiltonian flow and for the generators of the tangent flow. Moreover this set of functions is left invariant (as a set) by the flow of the Schrödinger equation if the potential is smooth enough. Therefore the formal manipulations that we shall perform can be justified.

3 Tangent Flow, Quadratic Hamiltonians

The equations of the classical tangent flow are linear homogeneous (with time-dependent coefficients that depend on the trajectory considered) and therefore their symplectic generators are quadratic homogeneous in the local coordinates.

The semiclassical approximation for states of class (A) leads therefore to study Hamiltonians *which are quadratic in the* $\hat{q}_k$, $\hat{p}_k$; they are generators of the quantum analog of the classical linear flow. These time-dependent quadratic Hamiltonians are obtained as quantization of the terms of order two (quadratic terms) in the formal expansion of the classical Hamiltonian in a neighborhood of a classical trajectory. As a consequence, in the case of quantum mechanical Hamiltonians which are polynomials of order at most two, the evolution in quantum mechanics is determined, modulo a (Berry) phase, by the classical evolution.

In the following a crucial role has the following property: the commutator of homogeneous polynomials of order at most two in the operators $\hat{q}_k$, $\hat{p}_k$ gives *essen-*

tially (modulo a central term) the same result as the quantization of the Poisson brackets of the classical counterpart.

Moreover, we have seen in "Lecture 10: Derivations and Generators. K.M.S. Condition. Elements of Modular Structure. Standard Form" that the unitary operators $U_k(b) \equiv e^{ib\hat{q}}$ and $V_k(a) \equiv e^{ia\hat{p}}$ give rise to translations of $\hat{p}_k,\ \hat{q}_k$:

$$U_j(b)\hat{p}_k U_j(b)^* = \hat{p}_k + b\delta_{j,k}, \quad V_j(a)\hat{q}_k V_j(a)^* = \hat{q}_k + a\delta_{j,k}$$

By choosing appropriately $a(t)$, $b(t)$ one can therefore *move* the barycenter of a well-localized states in such a way that it follow a prescribed classical path.

We will assume that the Hamiltonian flow is defined globally in time and we shall denote by $\gamma_{q^0,p^0}(t)$ the trajectory corresponding to the initial datum $q^0,\ p^0$.

4 Coherent States

We give the proof of (i) to (iv) for a special class of well-localized states, the *coherent states*, which are maximally localized in a sense that we will make precise. The wave functions of these states are Gaussians (recall that the Fourier transform of a Gaussian is again a Gaussian). In this case the semiclassical limit is very natural.

One considers next *generalized coherent state* [1, 7], constructed both in configuration space and in momentum space by considering special Hermite fuctions (product of polynomials with a Gaussian); notice that this operation has a natural conjugate relation in momentum space since differential operators act on a Gaussian as a multiplication by a polynomial.

The case of an arbitrary well-localized state follows then because generalized coherent states are dense (in the L^2 topology) in the set of well-localized states; notice that the L^2 norm is conserved under unitary evolution. In this Lecture we shall treat in detail the case of coherent states, and briefly quote the general result due to.

Definition 2 (*coherent states*) *Coherent states* are pure states characterized by a wave function that has *minimal total dispersion* relative to the observables $\hat{q},\ \hat{p}$.

Recall that the *dispersion* $\delta_\psi(A)$ of a self-adjoint operator $A \in \mathcal{B}(\mathcal{H})$ in the (pure) state described by the wave function $\psi \in D(A)$ is defined as

$$\delta_\psi(A) \equiv \sigma_\psi(A^2) - (\sigma_\psi(A))^2, \qquad \sigma_\psi(A) \equiv (\psi, A\psi) \tag{15}$$

For any pair of operators one has

$$C \equiv [\frac{A - \sigma_\psi(A)I}{\delta_\psi(A)} + i\frac{B - \sigma_\psi(B)I}{\delta_\psi(B)}][\frac{A - \sigma_\psi(A)I}{\delta_\psi(A)} - i\frac{B - \sigma_\psi(B)I}{\delta_\psi(B)}] \geq 0 \tag{16}$$

with equality only if $C = 0$. This can occur only if ψ is an eigenstate both of A and of B.

If $AB - BA = i\hbar I$ (as is the case for $A = x,\ B = \nabla_x$) one derives from (16) $\delta_\psi(A)\delta_\psi(B) \geq \hbar$.

In particular for $A = \hat{q},\ B = -i\frac{\partial}{\partial q}$, *and requiring symmetry of the dispersions* the minimum is achieved for wave functions that satisfy

$$[(\hat{q} - a) - (\hat{p} - b)][(\hat{q} - a) + (\hat{p} - b)]\psi = 0 \quad a = (\psi, \hat{x}\psi) \quad b = (\psi, \hat{p}\psi) \tag{17}$$

From (17) one concludes that in the Schrödinger representation the functions which have minimal *symmetric dispersion* with respect to the pair $\hat{q},\ \hat{p}$ are parametrized by $z \in C^N$

$$\psi_z(x) \equiv Ce^{-\frac{(x-a)^2}{\hbar} + i\frac{bx}{\hbar}}, \quad z = a + ib \tag{18}$$

Identity (18) defines an $\{a, b\}$-*coherent state*. It is an eigenstate of the annihilation operator $a \equiv \hat{q} + i\,\hat{p}$ relative to the eigenvalue $a + ib \in C$.

If the dispersion is minimal but not symmetric the solutions of (19) are called *squeezed coherent states*; the name "squeezed" is given because when $\delta_\psi(\hat{q}) > \sqrt{\hbar}$ and $\delta_\psi(\hat{q})\ \delta_\psi(\hat{p}) = \hbar$ these states are more localized in position space and consequently less localized in momentum space.

In dimension N coherent states are products of coherent states in the different variables. They can be written

$$\psi_{K,a} = C_{k,h}e^{-\frac{(x,Kx)}{2\hbar} + i\frac{(a,x)}{\hbar}}, \qquad x, a \in R^N \tag{19}$$

where K is a strictly positive matrix and $C_{k,h}$ a normalization constant. The quantum Fourier transform of $\psi_{K,a}(x)$ has the form (19) with K^{-1} replacing K (the Fourier transform would also replace $\hbar$ with $\hbar^{-1}$).

Coherent states *are not orthogonal to each other*, but they satisfy a *completeness relation* which in one dimension takes the form (with $K \equiv I$)

$$\forall \phi \in L^2 \qquad \phi(x) = C\int (\psi_a, \phi)\psi_a(x)da \tag{20}$$

where C is a normalization constant.

Since (20) must hold also when $\phi(x) = \psi_b(x)$ one has

$$\psi_b(x) = c\int_{R^N} (\psi_a, \psi_b)\psi_a(x)da = c\int [\int_a \bar{\psi}_a(y)\psi_a(x)da]\psi_b(y)dy$$

From this one sees that $c\int \bar{\psi}_a(x)\psi_a(y)da$ is a *reproducing kernel*.

These formulas generalize easily o the N-dimensional case.

5 Quadratic Hamiltonians. Metaplectic Algebra

If the selfadjoint operator H_2 is at most quadratic in the $\hat{p}$, $\hat{q}$, then the group of unitary operators $e^{it\hat{H}_2}$ leaves invariant the set of coherent states. This remark will be at the basis of the analysis of the semiclassical limit [1].

For linear homogeneous Hamiltonians the equations of motion for the canonical variables have the same form in Classical and Quantum Mechanics. Therefore under the flow of a linear Hamiltonian the variation in time of the canonical quantum variables is the same as that of that of the classical variables.

In particular the unitary operators $U_k(b) \equiv e^{ib\hat{q}_k}$ and $V_h(b) \equiv e^{ib\hat{p}_k}$ where $k = 1, \ldots, N$ generate translation for the operators $\hat{p}$ and $\hat{q}$:

$$U_h(a)\hat{p}_k U_h(a)^* = \hat{p}_k + a\ I\delta_{h,k} \qquad V_h(b)\hat{q}_k V_h(a)^* = \hat{q}_k + b\ I\ \delta_{h,k} \tag{21}$$

By a proper choice of the functions $a(t)$, $b(t)$ one can therefore *move by this unitary transformations* the barycenter of the coherent states on any chosen trajectory.

We shall denote by the symbol $\mathcal{Q}$ the algebra of homogeneous quadratic polynomials in the coordinates q_k, p_h (the symbol $\mathcal{Q}$ stands for "quadratic"). They form a Lie algebra under Poisson brackets, the *metaplectic Lie algebra* mp_{2N} (isomorphic to the symplectic algebra sp_{2N}).

The algebra mp_{2N} generates the linear *metaplectic group* Mp_{2N}, a double covering of the symplectic group Sp_{2N} (double covering because quadratic polynomials are invariant under the map $z \to -z$) [5–7].

The metaplectic Lie algebra acts by Poisson Brackets as derivation on the space of linear polynomials in the canonical variables q_h, p_k; this action *is the same as that of* sp_{2N}.

Notice that the corresponding quadratic polynomials in the quantum mechanical coordinates obtained by quantization of $\mathcal{Q}$, i.e. for $A \in \mathcal{Q}$

$$Op(A) = \frac{1}{2}\sum_{k,h=1}^{2N} \hat{\xi}_h A_{h,k}\hat{\xi}_k, \quad \hat{\xi}_k = \hat{q}_k \quad \hat{\xi}_{k+N} = \hat{p}_k \qquad 1 \le k \le N \tag{22}$$

do not form a Lie algebra under commutation. This is due to the fact that the quantum canonical variables *do not commute* and therefore the ideals defining the zero elements are different in the two cases.

One can verify that $[Op(A), Op(B)] - Op(\{A, B\})$ is a multiple of the identity and therefore belongs to the centre in any representation. The polynomials $Op(A)$, $A \in \mathcal{Q}$, under commutation form together with the identity an algebra (the Heisenberg algebra) which is the semi-direct product of mp_{2N} and a commutative algebra [1, 7].

The flow induced on quantum quadratic polynomials by a symmetric quantum quadratic operator differs only by a phase form the corresponding classic (tangent) flow. The algebra generated by $Op(A)$ acts by commutation on the linear space $\mathcal{L}$

of the quantum canonical coordinates, and the unitary group generated acts in the adjoint representation.

Further details on the metaplectic group can be found in [5].

6 Semiclassical Limit Through Coherent States: One-Dimensional Case

We return now to the analysis of the semiclassical limit using coherent states. We shall treat first the one-dimensional case. We shall give later the generalization to the case of N degrees of freedom, $N > 1$.

We shall discuss first the action of the metaplectic group in Quantum Mechanics. It is convenient to use the operator $a = \frac{1}{\sqrt{2}}(\hat{q} - i\,\hat{p})$ and its adjoint a^*.

Consider the operators

$$D_\hbar(\alpha) = e^{\frac{\alpha a^* - \bar{\alpha} a}{\sqrt{\hbar}}} \tag{23}$$

$$T(\beta) = e^{\frac{\gamma(\beta)(a^*)^2 - \bar{\gamma}(\beta) a^2}{2}} \qquad \gamma(\beta) = \beta|\beta|^{-1} \arctan|\beta| \tag{24}$$

This operators are defined initially on finite linear combinations of Hermite functions; they are isometric with isometric inverse, and can be extended to unitary operators on $L^2(R)$ which will be denoted by the same symbol. The operator $T(\beta)$ can be written as

$$T(\beta) = e^{\frac{i}{2} Im\gamma(\hat{q}^2 - \hat{p}^2) - \frac{i}{2} Re\gamma(\hat{q}\ \hat{p} + \hat{p}\ \hat{q})} \tag{25}$$

with inverse $T(-\beta)$.

Repeated application of the operator $D_\hbar(\alpha)$ to the ground state of the harmonic oscillator generates the coherent states with symmetric dispersion. Indeed

$$\psi_\alpha = D_\hbar(\alpha)\psi_0, \qquad \psi_0 = \frac{1}{\sqrt{\pi}} e^{-\frac{(x-a)^2}{2\hbar} + i\frac{(p-b)}{\hbar}} \qquad \alpha = a + ib \tag{26}$$

is a coherent state centered in $\alpha \equiv \{a,\ b\}$.

Repeated applications of the operator $T(\beta)$ generates coherent states with *asymmetric* dispersion. On a dense domain one has

$$T(\beta) a T(-\beta) = (1 - |\beta|^2)^{-1/2}(a - \beta a^*) \tag{27}$$

The function

$$\phi_{\alpha,\beta} \equiv D_\hbar(\alpha)T(\beta)\psi_0 \tag{28}$$

represents a coherent state centered in α with dispersions

$$\delta_q(\phi) = \frac{1}{\sqrt{\beta}} \qquad \delta_p(\phi) = \hbar(\delta_q(\phi))^{-1} \tag{29}$$

Often one introduces *generalized coherent states* defined by

$$\phi^n_{\alpha,\,\beta} = D_\hbar(\alpha)T(\beta)h_n \tag{30}$$

where h_n is the nth Hermite function. These states can be used to give a better semiclassical estimate (up to an arbitrary order in $\hbar$) of the solution of the Schrödinger equation with initial data well-localized in phase space.

7 Semiclassical Approximation Theorems

We can now state the semiclassical approximation theorem in the case of one degree of freedom [4].

Let $q(t),\ p(t)$ be a trajectory of the classical Hamiltonian system with Hamiltonian $H(q,\, p;\, t) = p^2 + V(q,\, t)$, $V \in C^3$. Let $S(t)$ be the integral of the classical action along the trajectory

$$S(t) = \int_0^t ds[\frac{1}{2}p(s)^2 - V(q(s), s)] \tag{31}$$

Let $\xi(t), \pi(t)$ solutions of the tangent flow (linearized equations at $q(t),\ p(t)$)

$$\dot{\xi}(t) = \pi(t), \qquad \dot{\pi}(t) = -V''(q(t),\, p(t);\, t)\xi(t) \tag{32}$$

(we choose $Re(\xi(0)\pi(0)) = 1$).

Define the (complex-valued) functions

$$\alpha(t) = \frac{1}{\sqrt{2}}(q(t) + ip(t)), \qquad \beta(t) = \frac{\xi(t) - \pi(t)}{\xi(t) + \pi(t)} \tag{33}$$

and the generalized coherent states

$$\Phi^n_{\alpha,\beta}(x,t) \equiv \hbar^{-1/4}\phi^n_{\alpha,\beta}((\hbar)^{-1/2}\, x, t) \tag{34}$$

This states are localized in a neighborhood of diameter $O(\sqrt{h})$ of $q(t)$ and their Fourier transform is localized in a neighborhood of diameter $O(\sqrt{\hbar})$ of $p(t)$. One has

Theorem 1 *Let $U_\hbar(t)$ be the propagator of the Hamiltonian $H_\hbar = -\frac{\hbar}{2}\Delta + \frac{V}{\hbar}$. There exists a positive constant C (which depends on the initial datum and on the potential V) such that for every integer n and real positive number T one has for $0 \le t \le T$*

$$\|U_\hbar(t)\Phi^n_{\alpha(0),\beta(0)}(x) - e^{-i\frac{\delta(t)}{\hbar}(n+1/2)[\gamma(t)-\gamma(0)]}\Phi^n_{\alpha(t),\beta(t)}\|_2 \le C(T)\sqrt{\hbar}t \quad 0 \le t \le T \tag{35}$$

where

$$\delta(t) = S(t) - [q(t)p(t) - q(0)p(0)] \qquad \gamma(t) = -arg(\xi(t) - i\dot{\xi}(t)) \tag{36}$$

Moreover $C = 0$ if the Hamiltonian is at most quadratic.

The proof of Theorem 1 makes use of the following Lemma.

Lemma 2 *Let $U^0(t,s)$ be the family of unitary operators that describe the quantum motion due to the Hamiltonian $H(t) = \frac{1}{2}\hat{p}^2 + f(t)\hat{q}^2$ where $f(t) \in C^1$. Then*

$$U^0(t,s) = T(\beta(t))e^{\frac{i}{2}(\gamma(t)-\gamma(s))(a^*a+aa^*)}T(\beta(-s)) \tag{37}$$

Proof Let $U_0(t,s)$ be the evolution operator in $L^2(R)$ associated to the equation

$$i\hbar\frac{\partial\phi(x,t)}{\partial t} = (-\frac{1}{2}\hbar^2\frac{d^2}{dx^2} + f(t)x^2)\phi(x,t) \tag{38}$$

Define

$$K_0 \equiv (a^*a + aa^*)/4, \qquad K_+ \equiv (a^*)^2/2, \qquad K_- \equiv a^2/2 \tag{39}$$

One has $[K_0, K_\pm] = \pm K_\pm$ and then

$$\begin{aligned} T(\beta)K_0T(-\beta) &= \frac{1+|\beta|^2}{1-|\beta|^2}K_0 - \frac{\beta}{1-|\beta|^2}K_+ - \frac{\bar{\beta}}{1-|\beta|^2}K_- \\ T(\beta)K_+T(-\beta) &= \frac{1}{1-|\beta|^2}K_+ + \frac{\bar{\beta}^2}{1-|\beta|^2}K_+ - 2\frac{\bar{\beta}}{1-|\beta|^2}K_0 \end{aligned} \tag{40}$$

One can verify (40) using a power series expansion for the exponential in $T(\beta)$, an operation which is legitimate on a dense domain left invariant by K. By construction

$$
\begin{aligned}
i\frac{d}{dt}T(\beta(t)) &= \lambda(t)K_+ + \bar{\lambda}(t)K_- + \mu(t)K_0,\\
\lambda &= i\dot{\beta}(1-|\beta|^2)^{-1},\\
\mu &= i((\dot{\beta})^* - \beta^*\dot{\beta})(1-|\beta|^2)^{-1}
\end{aligned}
\tag{41}
$$

Using these formulas one can verify the following identity

$$
\begin{aligned}
i\tfrac{d}{dt}(T(\beta(t))e^{2i\gamma(t)K_0}) = [(\mu - 2\dot{\gamma}\tfrac{1+|\beta|^2}{1-|\beta|^2})K_0 + (\lambda + \tfrac{2\beta\dot{\gamma}}{1-|\beta|^2})K_+\\
+ (\lambda + \tfrac{2\beta\dot{\gamma}}{1-|\beta|^2})^*K_-]T(\beta(t))e^{2i\gamma(t)K_0}
\end{aligned}
\tag{42}
$$

On the other hand, it follows form the definition of $\beta(t)$ that this function satisfies the differential equation

$$
\dot{\beta} = -i\beta(1+f) + i(1-f)\frac{1+|\beta|^2}{2} \tag{43}
$$

with $(1-f)Re\beta - (f+1) \equiv 2\dot{\gamma}$. In this notation $2H(t) = 2(f+1)K_0 + (f-1)[K_+ + K_-]$. Hence

$$
i\frac{d}{dt}(T(\beta(t))e^{2i\gamma(t)K_0}) = H_0(T(\beta(t))e^{2i\gamma(t)K_0})
$$

This ends the proof of Lemma 2. ♡

Outline of the proof of Theorem 1, one-dimensonal case We outline now the proof of Theorem 1 for the one-dimensional case. A complete proof can be found in [1].

The strategy is to write

$$
H = H_2(t) + H_r(t)
$$

where $H_2(t)$ is the Taylor expansion up to second order of the operator H at the point (in phase space) occupied by the trajectory at time t and H_r is the residual term.

We shall use then Lemma 2 and the fact that the term H_r is suitably small on states that are well-localized on the classical trajectory. Indeed one has

$$
H_2(t) = \frac{\hbar^2}{2}\frac{d^2}{dx^2} + V((q(t)) + (x-q(t))V'(q(t)) + \frac{1}{2}(x-q(t))^2V''(q(t)) \tag{44}
$$

and if potential $V \in C^3$, then H_r is the multiplication operator by

$$
H_r(t) = V'''(q(t))(x-q(t))^3(1+o(1)) \tag{45}
$$

Therefore $|(H_r\psi)(x)|_2$ is small on functions localized in a sufficiently small neighborhood of $q(t)$.

To prove the theorem one must substitute these heuristic arguments with precise estimates. We shall sketch these estimates making use of Lemma 2 and of the fundamental theorem of calculus.

If $H(t)$ and $H_0(t)$ are two families of self-adjoint operators and $U(t,s)$, $U^0(t,s)$ are the families of unitary operators which solve

$$i\frac{dU(t,s)}{dt} = H(t)U(t,s), \qquad U(s,s) = I$$
$$i\frac{dU^0(t,s)}{dt} = H_2(t)U^0(t,s), \quad U^0(s,s) = I$$

one has, from the fundamental theorem of calculus

$$(U^0)^*(t,s)U(t,s) = I + \int_s^t \frac{d}{d\tau}(U_2^*(\tau,s)U(\tau,s))d\tau \tag{46}$$

Setting $W(t) = H(t) - H_2(t)$ one has

$$U(t,s) = U^0(t,s) + \int_s^t U(t,\tau)W(\tau)U^0(\tau,s)d\tau$$
$$W(t) = \frac{1}{\hbar}[V'''(q(t))(x-q(t))^3 + o|x-q(t)|^3]B$$

where B is a bounded operator.

For each $\phi \in \mathcal{H}$ one has

$$|(U(t,s) - U_2(t,s))\phi| \leq \int_s^t |W(H_r(t))U_2(\tau,s)\phi|_2 d\tau \tag{47}$$

In our case $W(t) = H_r(t)$. To be able to bound this term by $Ct\sqrt{\hbar}$ for $t \in [0,T]$ it is therefore sufficient to find a dense set of elements $\psi \in \mathcal{H}$, invariant (as a set) under the action of $U^0(t,s)$ and such that for each of them

$$|W(t)\psi| \leq C\hbar^{1/2}, \forall t \in [0,T] \tag{48}$$

This is the set of coherent states; with this choice, uniformly in $0 \leq t \leq T$

$$\| \int_0^t W(\tau)U_0(\tau,s)\phi\| ds \leq \frac{1}{\hbar}\int_0^t [\| V'''(q(t))(x-q(t))^3(1+o(1))]\phi\| ds = O((\hbar))^{\frac{1}{2}}$$

where we have made use of the fact that $\phi(t,x)$ is well-localized to order $\sqrt{\hbar}$ in a neighborhood of the classical trajectory. This ends the sketch of the proof of Theorem 1. ♡

Remark that in the case we are studying H_r contains no term that depends on the products $p_k\, q_h$. The case in which one has a term that depends linearly on both p an q (as is is the case in presence of a magnetic field) requires a more detailed treatment.

8 N Degrees of Freedom. Bogolyubov Operators

We shall now give a brief account of the generalization to the case of any (finite) number N of degrees of freedom. The procedure we will follow is the same as in the case of one degree of freedom, the formal complications which arise are due entirely to operations with matrices.

The coherent states are also here functions with minimal joint dispersion $\delta_q(\phi)$ $\delta_p(\phi) = \hbar$ and are eigenfunction of linear combinations of creation and annichilation operators. In place of the operator D one has now

$$D_\hbar(\alpha) \equiv e^{\hbar^{-1/2}[(\alpha,a^*)-(\alpha^*,a)]}, \qquad \alpha \in R^N, \qquad a \equiv a_1, \ldots, a_N \tag{49}$$

The Bogolyubov operators provide a representation of the metaplectic group as linear tranformation of the algebra of the canonical commutation relations.

The definition of *Bogolyubov operator* is as follows.

Let B be a $N \times N$ hermitian matrix, $|B| < 1$, and let $B = |B|U$ be its polar decomposition. Define the Bogolyubov operator $T(B)$ as

$$T(B) = e^{\frac{1}{2}((a^*,\Gamma a^*)-(a,\Gamma a))} \qquad \Gamma_{i,j} \equiv Arg(\tanh(|B|U)_{i,j}) \tag{50}$$

The operators $T(B)$ form a subgroup of the metaplectic group. Their relation with symplectic linear transformations is explicitely given by

Lemma 3

$$T(B)aT^*(B) = (1 - BB^*)^{-1/2}(a - Ba^*) \tag{51}$$

Proof The proof can be given introducing the path $B(t) = tB$ and proving that the derivative of the right side coincides with that of the right hand side (for $t = 0$ the relation holds trivially). ♡

In the following we shall also make use of the following Lemma.

Lemma 4 *Let A, D complex-valued rank N matrices such that*

$$DA^{-1} = (DA^{-1})^*, \qquad A^*D + DA^* = I$$

Define $B = (I + DA^{-1})^{-1}(I - DA^{-1})$. Then $I - B^*B = 4(A^* + D^*)^{-1}$ $(A + D)^{-1} \geq 0$.

Proof From the definition of B

$$B^*B = (A^* + D^*)^{-1}(AA^* + DD^* - 2I)(A + D)^{-1}$$

and from the assumptions on A and D

$$I = (A^* + D^*)^{-1}(A^*A + D^*D + 2I)(A + D)^{-1}$$

From this the assertion follows. ♡

Making use of Lemmas 3 and 4 we can now introduce approximate evolution equations that we be useful to study the semiclassical limit.

Consider the quantum Hamiltonian

$$\hat{H}(t) = \frac{\hbar^2}{2}(\hat{p}, \hat{p}) + V(t, x) \qquad p = -i\nabla \quad x \in R^3 \tag{52}$$

and denote with H the corresponding classic Hamiltonian

$$H(t) = \frac{1}{2}(p, p) + V(t, q) \qquad q, p \in R^3 \tag{53}$$

Let $q(t),\ p(t)$ be a solution of the classical Hamiltonian equations and denote by $S(t)$ the action integral on this trajectory

$$S(t) = \int_0^t [\frac{1}{2}(p(s), p(s)) - V(s, q(s))]ds \tag{54}$$

The tangent flow is described by linear equation with regular time-dependent coefficients and therefore the solutions can be extended to the complex plane. This extension will be useful in comparing the expressions we will obtain with those obtained solving Schrödinger equation.

The equation for the tangent flow can be put in the form

$$\dot{A}(t) = iD(t), \qquad \dot{D}(t) = F(t)A(t), \quad F_{i,j}(t) = \frac{\partial^2}{\partial q_i \partial q_j} V(t, q)_{q=q(t)} \tag{55}$$

For each value of t the matrix $D(t)A^{-1}$ is symmetric and $\dot{A}(t)D(t) + D(t)\dot{A}(t) = 2I$. Define

$$I(t) \equiv -\frac{1}{\sqrt{2}}(D(t)x + A(t)\hat{p}) \qquad \dot{I}(t) = i[\hat{H}_0(t), L(t)] \tag{56}$$

where $\hat{H}_0(t) = \frac{\hbar^2}{2}(\hat{p}, \hat{p}) + (x, F(t)x)$. With these notations one has

Lemma 5 *Let*

$$\alpha(t) \equiv \frac{1}{\sqrt{2}}(q(t) + ip(t)), \qquad \lambda_0 \equiv \frac{1}{\sqrt{2}}(iA^*(0)p(0) - D_0^* q(0)) \tag{57}$$

Then

$$T(-B(t))D_1(-\alpha(t))I^*(t)D_1(\alpha(t))T(B(t)) = \lambda_0 + W(t)a^* \tag{58}$$

where $W(t)$ is the unitary matrix $W(t) \equiv -(A(t)+D(t))^{-1}[((A(t)+D(t))(A^*(t)+D^*(t)))]^{\frac{1}{2}}$.

Proof Using $D_1(-\alpha)aD_1(\alpha) = a + \alpha$ one obtains

$$D_1(-\alpha)I^*D_1(\alpha) \equiv I^* + \frac{1}{\sqrt{2}}[iA^*p - D^*q] = I^* + \lambda_0 \tag{59}$$

To prove Lemma 5 it is therefore sufficient to prove that the time derivative of $[iA^*p - D^*q]$ is zero. This follows from Lemma 4. ♡

We can now state the theorem that proves the existence of the semiclassical limit for generalized coherent states i.e. states represented for N degrees of freedom by

$$\phi^n_{\alpha,\beta} \equiv D_1(\alpha)T(\beta)\phi_n \tag{60}$$

The generalization to N degrees of freedom of Theorem 1 is the following

Theorem 6 (Hagedorn [4]) *If the potential $V(x)$ is of class C^3 and there exist positive constants c, c_1 such that*

$$|V(x)| \leq e^{c|x|^2}, \quad V(x) \geq -c_1 \tag{61}$$

then for every generalized coherent state, for every integer n and for every $T \geq 0$ there exists a positive constant C such that

$$\|U_\hbar(t,0)\Phi^n_{\alpha,\beta} - e^{-iS(t)}e^{i(n+\frac{1}{2})[\gamma(t)-\gamma(0)]}\Phi^n_{\alpha,\beta}\| \leq CT\sqrt{\hbar} \qquad |t| \leq T \tag{62}$$

where

$$(n+\frac{1}{2})\gamma(t) \equiv \sum_{k=1}^{N}(n_k + \frac{1}{2})$$

and $e^{i\gamma_k}$ is the kth eigenvalue of the matrix $A + i\ B$ in Lemma 4.

The vector $\Phi^n_{\alpha,\beta}$ is obtained rescaling $\phi^n_{\alpha,\beta}$:

$$\Phi^n_{\alpha,\beta}(x) = \hbar^{-\frac{1}{4}}\phi^n_{\alpha,\beta}(\sqrt{\hbar}^{-\frac{1}{2}}x) \tag{63}$$

We will not give the proof of Theorem 6, as it follows the same line as the proof of Theorem 1; the difficulties come only from the need of manipulating matrices.

9 Linear Maps and Metaplectic Group. Maslov Index

We return now to the analysis of the quadratic operators $Op(A)$ that were defined in (22).

Recall that if A is a hermitian $2N \times 2N$ matrix, the self-adjoint operators

$$Op(A) = \frac{1}{2^N}\sum_{h,k=1}^{2N} \hat{z}^*_h A_{h,k}\hat{z}_k, \qquad A = A^* \tag{64}$$

have a dense common domain of analytic vectors (e.g. the analytic vectors of $\sum_{k=1}^{2N}(\hat{z}^*_k\hat{z}_k)$).

On this domain one has, for two such matrices $A_1,\ A_2$

$$[Op(A_1), Op(A_2)] = Op(C) + Tr(A_1 J A_2)\cdot I \qquad C = A_1A_2 - A_2A_1 \tag{65}$$

where J is the standard symplectic matrix.

We have already remarked that a similar relation holds among the homogeneous polynomials of degree two in the symplectic coordinates z_k under Poisson Brackets, but in this case the trace term in (65) is absent.

Notice that for operators the map $Op(S^A)(t) = e^{itOp(A)}Op(S)e^{-itOp(A)}$ gives the same result for $Op(A)$ and $-Op(A)$ while the corresponding action on $L^2(R^3)$ distinguishes between $e^{itOp(A)}\phi$ and $e^{-itOp(A)}\phi$.

The group generated by the operators $Op(A)$ in the Schrödinger representation is a projective representation of the symplectic group. Denote by $\theta(S)$ the angle that describes the relative phase of the vectors $Op(M(S))\phi_z$ and ϕ_{Sz} (which represent the same state and therefore differ at most by a phase). One can associate an *index* i_M to a closed non contractible curve γ in the space of symplectic matrices; it does not depend on the base point z. The index i_M counts the variation of the angle θ (modulo 2π) along the closed curve γ.

On the other hand to every symplectic matrix S one can associate an angle Θ_S by the following procedure. Every symplectic matrix can be written as $S = TR$ where T is a positive matrix and R an orthogonal one. Since the space of positive matrices is contractible it follows that Sp_{2N} has the same homotopy as $U(N)$. But $U(N) = SU(N) \times S^1$ and the group $SU(N)$ is contractible.

Therefore the homotopy of Sp_{2N} is the same as that of S^1 and can be characterised by an angle. The value of the angle Θ_S attributed to each symplectic matrix depends on the specific construction adopted. The construction allows to define an index i_S for any closed non-contractible curve in the space of symplectic matrices. It can be shown that the index i_S does not depend on the construction chosen and is a topological invariant.

One of the possible construction is as follows. A rank $2N$ symplectic matrix S can be written as

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{66}$$

where A, B, C, D are rank N real matrices.

If $det(A + iB)$ is not zero, one can define the angle γ_S as

$$\gamma_S \equiv arg \ \det \ (A + iB) \tag{67}$$

and extend the definition by continuity.

Notice that S can be written as

$$S = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix} \left(\begin{bmatrix} \cos(\frac{\alpha}{N}) & \sin(\frac{\alpha}{N}) \\ -\sin(\frac{\alpha}{N}) & \cos(\frac{\alpha}{N}) \end{bmatrix} \otimes I \right) \begin{bmatrix} X_0 & Y_0 \\ -Y_0 & X_0 \end{bmatrix} \tag{68}$$

where T_1 and T_2 are positive $N \times N$ matrices, X_0 and Y_0 are real matrices, I is the identity matrix and $det(X_0 + iY_0) = 1$.

From the definition of the matrices A and B it follows

$$A + iB = e^{i\alpha \, N^{-1}} (T_1 + iT_2)(X_0 + iY_0) \tag{69}$$

and therefore

$$det(A + iB) = e^{i\alpha} \det T_1 \det(I + iT_2 T_1^{-1}) \tag{70}$$

The matrix T_1^{-1} is symmetric (S is a symplectic matrix) and therefore the eigenvalues of $I + iT_2 T_1^{-1}$ are complex numbers $1 + i\lambda$.

When S follows a closed non self-intersecting curve (a cycle) in the space of symplectic matrices this eigenvalues follow a cycle. Complete cycles of $\alpha(t)$ correspond to complete cycles of $\gamma = arg \ \det \ (A + iB)$. Therefore $\gamma_S(t) \in S^1$ determines the index of a closed curve in the space of symplectic matrices.

We have associated to each closed non contractible curve γ in the space of symplectic matrices two indices i_S and i_M. The first was constructed using quadratic operators that generate a representation of the metaplectic group, i.e. using the semi-classical limit of quantum mechanics. The second is constructed using symplectic dynamics in phase space.

Definition 3 (*Maslov index,* [5, 7]) The Maslov index $\mathcal{M}$ is related to the difference between the index obtained in the quantized theory and the index obtained in symplectic dynamics. It is defined by

$$\mathcal{M} = i_M - \frac{1}{2} i_S \tag{71}$$

The Maslov index has its origin in the fact that the stationary phase analysis for the solutions of the Schrödinger equation provides for each turning point of the coordinate of each degree of freedom a change in phase of an angle π. It can be shown that for closed curve in configuration space the result is independent of the choice of coordinates. The Maslov index has a relevant role in the discussion of the semi-classical quantization rules. For a very clear presentation see [7].

It is interesting to notice that *if the spin in quantum mechanics were due to angular motion in some hidden space* the value of the spin would be a Maslov index.

To compute the Maslov index it is convenient to have an explicit description of the operators $M(S)$ as integral kernels in the Schrödinger representation. For the symplectic matrix

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{72}$$

we must find the integral kernel of the operator $Op(M(S))$ for which

$$Op(M^*(S))\hat{q}\, Op(M(S)) = A\hat{q} + B\hat{p}, \qquad Op(M(S))^*\hat{p}\, Op(M(S)) = C\hat{q} + D\hat{p}$$

One obtains

$$Op(M(S))(x, x') = \frac{e^{i\beta}}{\sqrt{(2\pi\hbar)^n \det |B|}} e^{[\frac{i}{2\hbar}(x', B^{-1}Ax') - 2(x', B^{-1}x) + (x, DB^{-1}x)]} \tag{73}$$

We choose the determination of the logarithm in such a way that the function is continuous in C/R_-; the factor $e^{i\beta}$ is not determined.

Notice that *in the Schrödinger representation* the integral kernel of $Op(M(S))$ becomes singular when $\det B \to 0$, while the operator $Op(M(S))$ has a smooth limit as can be seen choosing a different representation of the canonical commutation relations e.g. that in which the momentum $\hat{p}$ is chosen diagonal. In this representation the role of the matrix B is taken by the matrix C.

On a trajectory $S(t)$ in the space of symplectic matrices the locus $\det(B) = 0$ is called *caustic* (caustic surface). In the Schrödinger representation at a caustic the integral kernel has a singularity

$$\delta(y_1(x) - y_1(x')) \dots \delta(y_k(x) - y_k(x'))G(x, x') \tag{74}$$

where $G(x, x')$ is a regular function (k is the molteplicity of the zero eigenvalue of B).

In particular if $B = 0$

$$Op(M(S))(x, x') = \pm \frac{1}{\sqrt{det A^{-1}}} \delta(x - Ax') e^{\frac{i}{2\hbar}(x', Ax)} \tag{75}$$

Using (75) one can compare the integral kernel of $Op(M(S))$, $S = (S_1, S_2)$, with that of the product $Op(M(S_1)M(S_2))$. We assume that B_1 and B_2 be invertible. Denoting by N_+ (N_-) the number of positive (negative) eigenvalues of the symmetric invertible matrix $B_1^{-1}\ B\ B_2^{-1}$ of rank N one obtains

$$\beta = \beta_1 + \beta_2 + (N_+ - N_-)\frac{\pi}{4} \tag{76}$$

Setting $\beta = \nu\pi$, $\nu = \pm 1$, Eq. (76) reads

$$4 + N\pi + 2\nu\pi = (4\beta + N\pi + 2\nu_1\pi) + (4\beta + N\pi + 2\nu_2\pi) + 2\pi(\nu - \nu_1 - \nu_2 - N_-) \tag{77}$$

If we define

$$\beta(S) \equiv \beta + \frac{N\pi}{4} + \frac{\nu\pi}{2} \tag{78}$$

we have

$$\beta(S_1 S_2) = \beta(S_1)\beta(S_2) + n\pi, \qquad n \in Z \tag{79}$$

This shows that that the metaplectic group is a double covering of the symplectic one. Two metaplectic operators $Op(M(S, \sigma))$, $\sigma = \pm 1$, correspond to each symplectic matrix S.

An important property of the metaplectic operators is the transformation law for the Weyl system

$$Op(M^*(S))W(z)Op(M(S)) = W(S^{-1}z) \tag{80}$$

Making use of this property it is possible to introduce another group of operators, the *inhomogeneous metaplectic group*. Its elements are the operators

$$Op(M(z, \gamma, S)) \equiv e^{i\gamma\hbar} W(z) Op(M(S)) \equiv e^{i\gamma\hbar} Op(M(S)) W(S^{-1}z) \tag{81}$$

with composition law

$$Op(M(z_1, \gamma_1, S_1))Op(M(z_2, \gamma_2, S_2)) = Op[M(z_1 + S_1 z_2), \gamma_1 + \gamma_2 + \frac{1}{2}\omega(z_1, Sz_2), S_1 S_2] \tag{15.85}$$

It is therefore the semidirect product of the metaplectic group with the Weyl group. This group is called *Heisenberg group*.

References

1. Birkhoff, B. (1933). *Bulletin of the American Mathematical Society*, *39*, 61–79.
2. Duistermatt, J. J. (1974). *Communications on Pure and Applied Mathematics*, *27*, 207–281.
3. Evans, L. (1998). *PDE. AMS: Graduate Studies in Mathematics*.
4. Hagedorn, G. (1985). *Annales de l'Institut Henri Poincaré*, *42*, 363–374.
5. Maslov, V. (1994). The complex WKB method, Progress in Physics No. 16, Birkhauser.
6. Morette, C. (1951). *Physical Review*, *81*, 848–852.
7. Souriau, J.-M. (1975). *Lecture Notes in Physics, 50*(9), 117–147.

Lecture 16: Semiclassical Approximation for Fast Oscillating Phases. Stationary Phase. W.K.B. Method. Semiclassical Quantization Rules

We make in this Lecture an analysis of the semiclassical approximation for states represented in configuration space by wave function with fast oscillating phase. The W.K.B. method (named after Wentzell, Kramers, Brillouin) is a constructive method that permits to find approximate solutions of the Schrödinger (stationary or non-stationary) equation when the parameter $\hbar$ is small. The solutions are given as power series in $\hbar$ and *are different* from the solution we have described in "Lecture 15: Semiclassical Limit; Coherent States; Metaplectic group".

The W.K.B. method has analogies with Geometrical Optics, where the small parameter is the wave length. The solutions will have fairly regular absolute value and a strongly oscillating phase. We will see that the W.K.B. method is linked to the Hamilton-Jacobi equation in Analytic Mechanics and shares with it advantages and disadvantages.

The solutions are constructed in simple way for a time $0 \leq t < T_c$; the critical time T_c is the first time in which the projection of the trajectory on configuration space becomes singular (the first appearance of *a caustic*). The procedure for the construction of a solution beyond T_c is somewhat laborious; as in lagrangian mechanics the solution is unique only if one imposes no time delay. In general the solution after T_c is the sum of two or more functions that show a difference in phase and this leads lead to interference effects. We shall not discuss here the continuation problem and refer the interested reader to [3, 5, 12].

The WKB method is part of a more general analytic study of asymptotic solutions of evolution equations for strongly oscillating initial data. A clear description and connection with other methods can be found in [10].

1 Free Schrödinger Equation

We shall begin by studying the solutions of the free Schrödinger equation; their counterpart in geometrical optics are the solutions of the wave equation in a homogeneous medium.

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_16

Consider the equation

$$i\frac{\partial \phi}{\partial t} = -\frac{1}{2}\epsilon \Delta\ \phi \tag{1}$$

where ϵ is a small parameter (it has the role of $\hbar$ in Quantum Mechanics).

The solution of (1) with initial datum $\phi_0 \in L^2(R^N)$ is

$$\phi_\epsilon(x,t) = \epsilon^{-N}\int_{R^N} e^{i\frac{p^2 t}{\epsilon}} e^{i\frac{p\cdot x}{\epsilon}} \hat{\phi}_0(p)dp \tag{2}$$

One sees from (2) that when ϵ is very small the solution is *strongly oscillating* in x.

Equation (2) provides an explicit solution for all times of the free Schroedinger equation. In this case the Hamilton-Jacobi equation describes a rectilinear motion, therefore there is no inversion of motion.

If $V \neq 0$ we look for solutions of

$$i\frac{\partial}{\partial t} = -\frac{1}{2}\epsilon\Delta + \frac{1}{\epsilon}V(x) \tag{3}$$

corresponding to initial data of the form

$$\phi_0(x) = \sqrt{n(x)}e^{\frac{i}{\epsilon}S(x)} \qquad n(x),\ S(x) \in C^\infty \tag{4}$$

We construct the asymptotic expansion of the solutions of (4) in power series of the parameter ϵ.

Remark that the integral in (4) is a particular case of integrals of the form

$$I(\epsilon) = \int_{R^N} e^{i\frac{\theta(y)}{\epsilon}} f(y)\ dy \tag{5}$$

that are called (for obvious reasons) *oscillating integrals* [3, 6]. The function $\theta(y)$ is the *phase*.

The study of the limit of oscillating integrals when $\epsilon \to 0$ is based on the following two theorems.

2 The Non-stationary Phase Theorem

Non-stationary phase theorem

Consider a domain $\Omega \in R^d$ and assume that f and θ are of class $C^\infty(\Omega)$, that $f \in L^1(\Omega)$ and that for every $y \in \Omega$ one has $\nabla\theta(y) \neq 0$. *Then for each integer N* there exists a positive constant $C_N(\theta, f)$ such that

$$|I_\Omega(\epsilon)| = |\int_\Omega e^{i\frac{\theta(y)}{\epsilon}} f(y)\, dy| \leq C_N \epsilon^N \tag{6}$$

◇

Proof Set $g = \epsilon^{-1}\theta$.

We use the formula, valid when f and g are sufficiently regular

$$f e^{ig} = -i\ div(hf e^{ig}) + i\ e^{ig}\ div(hf) \qquad h = \frac{\nabla g}{|\nabla g|^2} \tag{7}$$

The formula (7) is derived from $div(uv) = u\ div\ v + (\nabla u, v)$ which is valid pointwise if the scalar valued function u and the vector valued function v are continuously differentiable. If $\nabla g \neq 0$ in Ω integration by parts gives

$$D_\epsilon \equiv \int_\Omega f(x) e^{ig(x)} dx = i \int_{\partial\Omega} e^{ig(x)} \sum_{k=0}^{d} (L_h f)^k \frac{\partial h_k}{\partial n} d\sigma + \int_\Omega e^{ig} L_h^k f dx \tag{8}$$

where $\frac{\partial}{\partial n}$ is differentiation in the direction perpendicular to $\partial\Omega$ and we have set

$$L_h f \equiv -i div(h\ f\ e^{ig}) + i e^{ig} div(h\ f) \tag{9}$$

Choosing

$$g \equiv \frac{\theta}{\epsilon} \qquad \delta \equiv\ inf_{x\in\Omega} |\nabla\theta|$$

(by assumption $\delta > 0$) and iterating n times one has, for $\lambda \geq 0$ and for any integer $n \in Z$

$$|D_\epsilon I \leq C_n \| f \|_n \epsilon^n \tag{10}$$

where $\| f_n \| = sup_{0\leq k\leq n} |\nabla^k f|_1$ ♡

Remark that if $f \in C^{n_0}$ this procedure can only be iterated only $n_0 - 1$ times, and therefore in this case one has the estimate

$$|D_\epsilon I \leq C_{n_0} \epsilon^{m_0} \tag{11}$$

3 The Stationary Phase Theorem

Stationary phase theorem *In the integral*

$$I(\epsilon) = \int_{R^d} e^{i\frac{\theta(y)}{\epsilon}} f(y)\, dy \tag{12}$$

let the function f have a finite number N of critical points in its support and suppose that they are all non degenerate. Let $D^2\theta$ the hessian matrix of θ in the critical points which by assumption has no zero eigenvalues. Then the following asymptotic expansion for (12) *holds*

$$I(\epsilon) = \sqrt{2\pi\epsilon} \sum_{k=1}^{N} e^{i\epsilon^{-1}[\theta(y_k)+\frac{\sigma_k}{\pi}]} f(y_k) |det\, D^2\theta|(y_k) + o(\sqrt{\epsilon}) \tag{13}$$

where N is the number of critical points and σ_k is the index *of the kth critical point (the difference between the number of positive and negative eigenvalues of the Hessian).* ◇

Proof We will give the proof in the one-dimensional case. The proof in the general case follows the same lines.

As a consequence of the non-stationary phase theorem we can consider disjoint neighborhoods of the critical points and use the non-stationary phase theorem in the complementary domain. Assume therefore that in Ω there is only a critical point, non degenerate, that we take to be the origin.

According to Morse lemma, there exists a neighborhood of the origin, that we assume coincides with Ω, and a C^1-function ϕ defined in Ω, with $\phi(0) = 0$ $\frac{d\phi}{dy} \neq 0$ in Ω such that for $y \in \Omega$ one has $\theta(y) - \theta(0) = \frac{1}{2}\phi(y)^2$. Therefore

$$I(\epsilon) \simeq e^{i\frac{\theta(0)}{\epsilon}} \int_{\Omega} e^{-i\frac{1}{2\epsilon}\phi(y)^2} g(\phi) d\phi \tag{14}$$

where $g(\phi) \equiv f(y(\phi))|det[\frac{dy}{d\phi}(y(\phi)]$. Setting

$$u(y) = \frac{1}{\epsilon^{\frac{1}{2}}}\phi(y) \tag{15}$$

one has

$$I(\epsilon) = e^{\frac{i}{\epsilon}\theta(0)} \int e^{-\frac{i}{\epsilon}u^2} g(\sqrt{\epsilon}u) du \tag{16}$$

Making use of the continuity of $g(y)$ we obtain

$$I(\epsilon) = 2\pi\sqrt{\epsilon} e^{i\theta(0)} f(0) |\frac{\partial y}{\partial \phi}(0)| e^{i\frac{\pi}{4}} + o(\sqrt{\epsilon}) \tag{17}$$

♡

One can give a more explicit form to the correction terms that we have denoted by $O(\sqrt{\epsilon})$ [5, 6]. For this purpose it is convenient to consider separately in (15) the integration over $[-\frac{a}{\sqrt{\epsilon}}, 0)$ and over $[0, \frac{b}{\sqrt{\epsilon}}]$.

We will give an analysis of the integration over the second domain.

Set $\lambda = \frac{1}{\sqrt{\epsilon}}$ and consider the integral

$$F(\epsilon) \equiv \int_0^b e^{i\frac{x^2}{\epsilon}} f(x)dx \tag{18}$$

where the function f is in C^∞ and vanishes in $x = b$ together with its derivatives up to order k.

We make use of the identity (obtained by integration by parts in the variable x)

$$F(\epsilon) = [f(x)\int_{-\infty}^0 e^{i\frac{(x+i\xi)^2}{\epsilon}} d\xi]_0^b - \int_0^b \frac{df}{dx}(\int_{-\infty}^0 e^{i\frac{(x+i\xi)^2}{\epsilon}} d\xi)dx \tag{19}$$

Integrating by parts k times with respect to ξ (remark that by dominated convergence on can exchange the orders of integration) one obtains

$$F(\epsilon) = -\sum_0^{k-1} \phi_j(0,\epsilon) f^j(0) + R_k(\epsilon)$$

$$\phi_j(x.\epsilon) = (j-1)!\int_0^\infty \xi^{j-1} e^{i\frac{(x-i\xi)^2}{\epsilon}} d\xi \qquad R_k(\lambda) = \int_0^b \phi_k(x,\epsilon) f^{(k)}(x)dx \tag{20}$$

(if the function $f(x)$ is only N times differentiable the iteration can be performed only N times). Summing the term which one obtains with a similar analysis in the first integration domain in (19) and setting $\lambda = \epsilon^{-1} > 0$ one obtains for any integer $k > 1$ that if the function $g(x)$ is k-times differentiable

$$\int_{-a}^b e^{i\frac{g(x)}{\epsilon}} f(x)dx = e^{i\frac{g(x_0)}{\epsilon}} \sum_0^{k-1} a_j(f,g)\epsilon^{j+\frac{1}{2}} + R_k(\epsilon) \quad |R_k(\epsilon)| \le c_k \epsilon^k \|f\|_{C^{k+1}_{[-a,b]}} \tag{21}$$

The coefficients $a_j(f,g)$ are given explicitly by

$$a_j(f,g) = \frac{\Gamma(j+\frac{1}{2})}{(2j)!} e^{i\frac{\pi\sigma}{4}(2j+1)} [\frac{d}{dx}]^j [h(x)^{-j-\frac{1}{2}} f(x)]_{x=x_0}]$$

$$h(x) \equiv 2\sqrt{\sigma(g(x) - g(x_0))}(x - x_0)^{-1} \tag{22}$$

where $\sigma \equiv sign\ g''(x_0)$. From (21) to (22) one derives the asymptotic behavior of the integral when $\epsilon \to 0$.

When the dimension d of the space is greater than one, Morse's Lemma still holds *if the quadratic form is not degenerate*. One can then choose coordinates in which the quadratic form is diagonal and proceed by induction starting with the one-dimensional case.

One can also give rather explicit formulae [6] for the coefficients of the term proportional to $\epsilon^{\frac{1}{2}+k}$, with $k \geq 0$. In the case in which g has only one non-degenerate critical point $x_0 \in R^d$, denoting by $\Omega \in R^d$ a neighborhood of this critical point one has

$$F(\epsilon) = e^{i\frac{g(x_0)}{\epsilon}} \sum_0^{k-1} a_j(f,g)\epsilon^{\frac{j}{2}} + R_k(\epsilon), \quad |R_k(\epsilon)| < c_k(g)\|f\|_{C_k^\beta(\Omega)}\epsilon^{\frac{d}{2}+k} \tag{23}$$

where $\beta < \infty$ and $a_j(f,g)$ is a linear operator of order $2j$. The explicit expression of the $a_j(f,g)$ is complicated. We give here only the term of lowest order in ϵ.

$$F(\epsilon) = (2\pi\epsilon)^{\frac{d}{2}} |det\, H_h(x_0)|^{-\frac{1}{2}} e^{i\frac{g(x_0)}{\epsilon} + i\frac{\pi}{4}\sigma_g(x_0)} \qquad \sigma_g(x_0) = sign H_g(x_0) \tag{24}$$

where $H_g(x_0)$ is the quadratic form (hessian) that appears in Morse's lemma for $d > 1$.

For a rather complete treatment one can see [5, 9]. Remark that in cases of practical interest the phase depends on additional parameters and the critical points form a manifold $\mathcal{M}$ in the space of parameters. In this case, in order to apply the stationary phase method, it is necessary to verify that the critical points be non degenerate *uniformly on* $\mathcal{M}$ (the eigenvalues of the hessian are uniformly bounded away from zero).

4 An Example

As a example we estimate up to terms of order $0(\sqrt{\epsilon})$ the integral (2). We consider only the case $d = 1$.

The solution (20) can equivalently be written

$$\phi_\epsilon(x,t) = (2i\pi t)^{-1} \int_{R^d} e^{i\frac{|x-y|^2}{2\epsilon}} \phi_0(y) dy \tag{25}$$

Therefore in our case $\theta_x(y) = \frac{m(x-y)^2}{2t} + S(y)$ and

$$\frac{\partial\theta}{\partial y} = \frac{m(y-x)}{t} + S'(y) \tag{26}$$

If $n(x)$ has compact support ($x \in (-L, L)$) for t sufficiently small (how small may depend on L) the equation $\frac{\partial\theta}{\partial y} = 0$ has a unique solution $y_0(x)$. One has

$$\frac{\partial\theta^2}{\partial y^2} = -\frac{m}{t} + S''(y) \tag{27}$$

therefore for t small enough $\frac{\partial\theta^2}{\partial y^2} < 0$ (remark that for larger values of t the Hessian becomes degenerate and the method cannot be applied).

We conclude that, up to an error of order smaller than $\sqrt{\epsilon}$ the solution is

$$\phi(x,t) = 2\pi\ \sqrt{\epsilon}\ e^{it\frac{S'(x(t))^2}{2\hbar}}[\frac{m}{t} - S''(x(t))] \tag{28}$$

where we have denoted by $y(t,x)$ the solution (assumed unique) of $\frac{x^2}{2t} + S(y(t,x)) = 0$.

We generalize now this example. We shall discuss equations of the form

$$i\epsilon\partial_t\phi^\epsilon = Op^w(H)\phi^\epsilon, \qquad x \in R^d, \quad t \in R \quad \epsilon \in (0,\epsilon_0) \tag{29}$$

with initial data

$$\phi^\epsilon(x,0) = \sqrt{\eta_I(x)}e^{i\frac{S_I(x)}{\epsilon}} \tag{30}$$

In (29) we have defined

$$(Op^w(H)\phi)(x) = (\frac{1}{2\epsilon\pi})^d \int_{R^d\times R^d} H(p, \frac{x+y}{2})\phi(y)e^{\frac{i}{\epsilon}(x-y,p)}dp\ dy \tag{31}$$

In the second part of these Lectures this operator will be called *Weyl quantization* of the Hamiltonian $\frac{1}{2}H(p,q)$. The parameter ϵ has here the role of $\hbar$. To stress the analogy with geometric optics we shall call *energy intensity* the function

$$|\phi^\epsilon(x,t)|^2 \equiv n^\epsilon(x,t), \qquad n^\epsilon(x,0) = n_I(x) \tag{32}$$

$0 < \epsilon \leq \epsilon_0$

We shall use the following notation

$$H(x,\xi) \in S^\sigma(R^d) \Leftrightarrow \forall\alpha,\ \beta \in N,\ \ \exists c_{\alpha,\beta} > 0\ \ t.c.\ \ |\frac{\partial^{\alpha+\beta}}{\partial x^\alpha\ \partial\xi^\beta}H(x,\xi)| < c_{\alpha,\beta}(I + |\xi|^{\sigma-\beta}) \tag{33}$$

On the hamiltonian we will make the following assumptions

A_1: There exists $\sigma \in R$ such that $H \in S^\sigma(R^d)$ uniformly in $0 < \epsilon \leq \epsilon_0$.
A_2: The operator $Op^w(H)$ is self-adjoint.

On the initial data we assume

$$n_I \in L^1(R^d), \qquad n_I \geq 0, \qquad S_I \in W^{1,1}_{loc}(R^d) \tag{34}$$

($W^{1,1}_{loc}$ is the class of functions which are locally in L^1 with first order derivatives in L^1).

Under these assumptions a weak solution of (29) exists and moreover

$$|n^\epsilon(t)|_{L^1} = |n^\epsilon(0)|_{L^1} \qquad \forall t \tag{35}$$

Examples of equation of type (29) are

(a) The Schrödinger equation

$$i\epsilon\partial_t\phi^\epsilon = -\frac{\epsilon^2}{2}\Delta\phi^\epsilon + V(x)\phi^\epsilon \tag{36}$$

(b)

$$i\epsilon\partial_t\phi^\epsilon + |D_y||a(\frac{x-y}{2})\phi^\epsilon(y)|_{y=x} = 0 \quad |D| = \sqrt{-\Delta} \tag{37}$$

(if $a(x) = 1$ this equation is related to the wave equation); indeed if $u^\epsilon_{tt} = \Delta u^\epsilon$, then $\phi^\epsilon_\pm \equiv \partial_t u^\epsilon \pm i|D|u^\epsilon$ satisfies

$$i\partial_t\phi^\epsilon_\pm = |D|\phi^\epsilon_\pm \tag{38}$$

(c) The linearized KdV equation

$$\epsilon\partial_t\phi^\epsilon = \frac{\epsilon^3}{3}\phi^\epsilon_{xxx} \tag{39}$$

To study (30) we will make the following *Ansatz*:

(1) The solution of (1) can be written as

$$\phi^\epsilon(x,t) = A^\epsilon(x,t)e^{\frac{i}{\epsilon}S(x,t)} \tag{40}$$

(2) A^ϵ can be written as power series in the parameter ϵ

$$A^\epsilon = A + \epsilon A_1 + \epsilon^2 A_2 + \cdots \tag{41}$$

Notice that in (41) by a redefinition of the $A_k, k \geq 1$ we can allow the function $S(x,t)$ to be independent of ϵ.

The solution of the Eq. (29) using this Ansatz constitutes the W.K.B. method.

We shall study in detail only the case when the classical hamiltonian can be written as

$$H(x,\xi) = \sum_{|k|\leq m} a_h(x)\xi^k, \tag{42}$$

where k is a multi-index and $\xi^k \equiv \xi_{k_1}, \ldots, \xi_{k_d}$

Substituting (40) and (41) in (42) one obtains

$$Op^w(H)(A^\epsilon\, e^{\frac{iS}{\epsilon}}) = e^{\frac{iS}{\epsilon}} \sum_{|k|=0}^{m} (i\epsilon)^{|k|} Op^w(R_k(A^\epsilon)) \tag{43}$$

where

$$R_0(A^\epsilon) \equiv H(x, \nabla S)\ A^\epsilon$$

$$R_1(A^\epsilon) \equiv \sum_1^d \frac{\partial H(x, \nabla S)}{\partial \xi_k} \cdot \frac{\partial A^\epsilon}{\partial x_k}$$

$$+ \frac{1}{2} \sum_1^d \frac{\partial^2 S}{\partial x_j\, \partial x_k} \cdot \frac{\partial H(x, \nabla)}{\partial \xi_j\, \partial \xi_k} A^\epsilon + \frac{1}{2} \sum_1^d \frac{\partial^2 H(x, \nabla)}{\partial y_k\, \partial \xi_j} A^\epsilon \tag{44}$$

To order zero in ϵ these equations have the following form, with $\xi = \nabla S(x)$

$$\frac{\partial A}{\partial t} + \sum_1^d \frac{\partial H(x, \xi)}{\partial \xi_k} \cdot \frac{\partial A}{\partial x_k} + \frac{1}{2} \sum_k \frac{\partial^2 S}{\partial x_k \partial x_j} \cdot \frac{\partial^2 H(x, \xi)}{\partial x_k \partial \xi_j} \cdot A + \frac{1}{2} \sum_1^d \frac{\partial^2 H(x, \xi)}{\partial x_k \partial \xi_j} A = 0, \tag{45}$$

$$\frac{\partial S}{\partial t} + H(x, \nabla_x S) = 0 \tag{46}$$

5 Transport and Hamilton-Jacobi Equations

Definition 1 Equation (45), linear in A, is the *transport equation*. Equation (46) is the *Hamilton-Jacobi equation*. ◇

The equations obtained to any order greater than one in ϵ are *linear in-homogeneous* in the A_k, with order zero term which depends on the A_s for $s < k$. When written in term of the pair $n \equiv A^2$, S Eqs. (45) and (46) represent a conservation law for the function n and the Hamilton-Jacobi equation for the function S.

For the case $H = \frac{1}{2}\xi^2 + V(x)$ one has

$$\partial_t n + div\ (n.\nabla S) = 0, \qquad \partial_t S + \frac{1}{2}|\nabla S|^2 + V(x) = 0 \tag{47}$$

For the wave equation with refraction index $n(x)$

$$\partial_t n + div\ (\frac{n(x)\nabla S}{|\nabla S|}) = 0, \qquad \partial_t S + n(x)|\nabla S| = 0 \tag{48}$$

i.e. the eiconal equation of geometric optics.

In the one-dimensional case system (48) takes a simpler form

$$\partial_t u + \partial_x(n(x))sign[\frac{\partial S}{\partial x}]u = 0 \tag{49}$$

If $\frac{dS}{dx}$ does not change sign, the system (48) decouples and the $W.K.B.$ ansatz gives the exact solution (if $A_k(0) = 0, \quad k \geq 1$ then $A_k(t) = 0 \;\; k \geq 1 \;\; \forall t$).

Following this procedure one finds an approximation to some solutions of (29) to order ϵ *as long as the solution of the Hamilton-Jacobi is unique*.

Theorem 1 *Let $\psi^\epsilon(t)$ be a solution of (29) with initial datum*

$$\psi^\epsilon(x,0) = \sqrt{n(x)}e^{\frac{iS(x)}{\epsilon}} \tag{50}$$

Assume that $n(x) \in S(R^d)$ and that $S(x) \in C^2$, and assume existence for $|t| < T_c$ of the solutions $S(x,t)$ and $n(x,t)$ of (48) and (49). Define

$$\phi^\epsilon_{WKB}(x,t) \equiv \sqrt{n(x,t)}e^{\frac{i}{\epsilon}S(x,t)} \tag{51}$$

There is then a constant C such that

$$sup_{|t|<T_c}\|\phi^\epsilon_{WKB}(t) - \phi^\epsilon(t)\|_2 < C\epsilon \tag{52}$$

◇

Proof Under the hypothesis we have made on $Op^w(H), n, S$, formal derivation gives

$$i\epsilon\frac{\partial}{\partial t}\phi^\epsilon_{WKB}(x,t) - Op^w(H)\phi^\epsilon_{WKB}(x,t) = \epsilon^2 Q^\epsilon(x,t) \tag{53}$$

where Q^ϵ depends on the derivatives of A and S and is bounded $L^1((-T_c,+T_c), L^2(R^d))$ uniformly in $\epsilon < \epsilon_0$. It follows from that, setting

$$u^\epsilon(x) \equiv \psi^\epsilon(x) - \psi^\epsilon_{WKB}(x) \tag{54}$$

one has

$$\partial_t u^\epsilon + \frac{i}{\epsilon}Op^w(H)u^\epsilon - \epsilon Q^\epsilon = 0, \qquad u^\epsilon(x,0) = 0 \tag{55}$$

From (55) one derives $\partial_t|u_\epsilon|^2 = -\epsilon\,|u_\epsilon||Q_\epsilon|$ and then, by integration

$$|u^\epsilon(t)| \leq \epsilon\int_0^t |Q^\epsilon(s)|\,ds \tag{56}$$

This implies

$$sup_{|t|<T_c}|u^\epsilon(t)|_{L^2} < C\epsilon. \tag{57}$$

♡

In the case of the Schrödinger equation one has

$$2Q^\epsilon = e^{i\frac{S}{\epsilon}}\Delta|u| \tag{58}$$

Equation (46) can be solved with the *method of the characteristics* and have in general a unique solution *up to the time when two characteristics meet*.

Consider at time $t = 0$ a point $x_0 \in R^d$. Any point y in a sufficiently small neighborhood Ω of x_0 can be reached from x_0 in a time τ following an integral curve of the dynamical system

$$\frac{dX}{dt}(t,x) = \nabla_\xi H(X(t,x),\xi), \qquad \frac{d\xi}{dt} = -\nabla_x H(X(t,x),\xi) \tag{59}$$

$$\xi_0 = \nabla S(x_0) \tag{60}$$

This curve is given implicitly by

$$\xi(t,x) = \nabla S(x(t,x_0),t) \tag{61}$$

and one has $x(\tau,x_0) = y$.

Suppose (59) and (60) have a unique solution for $t < T_0$, as is the case when the hamiltonian can be written as

$$H = \sum_{i,j} a_{i,j}(x)p_i p_j + \sum_j B_j(x)p_j + V(x) \tag{62}$$

where a_{ij} is a strictly positive matrix and all symbols are sufficiently regular. T_0 is the time at which the map $x \to X(T_0,x)$ is no longer invertible (the time of the first occurrence of a caustic).

In this case the transport equation is

$$\frac{\partial A}{\partial t} + (\nabla S, \nabla A) + A\sum_{i,j} a_{i,j}\frac{d^2S}{\partial x_i \partial x_j} = 0 \tag{63}$$

Denote $J_x \equiv det(\frac{\partial X_j(t,x)}{\partial x_i})$ and with $X_t^{-1}(x)$ the inverse of the map $x \to X((t,x))$, $t < T_0$. Then the first order term in the expansion of A in powers of ϵ is

$$A_0(t,x) = \frac{1}{\sqrt{J_t(X^{-1}(x))}}A(0,X_t^{-1}(x))$$

For a complete proof one can see e.g. [4].

In general the map $\Omega \ni y \to x(t, x_0)$ is not one-to-one for all times (characteristic curves intersect).

From a solution of (63) one obtains a solution of (59) integrating (over the trajectory considered) the equation

$$\frac{dS(x,t)}{dt} = \xi.\nabla_\xi H(x,\xi) - H(x,\xi), \qquad S(x(0,x)) = S(x) \tag{64}$$

In general the solution constructed in this way is discontinuous at the intersection of two characteristic curves. Methods to construct solutions at later times are laborious (see e.g. [5]).

We will discuss in the second part of these Lectures a method, based on the use of *Wigner functions*; it has the advantage, compared to the W.K.B. method, that the functions are defined on phase space. Caustics are the points in which the *projection on configuration space* of the characteristic curves is singular.

6 The Stationary Case

We have seen so far the application of the W.K.B. to find approximate solutions of the Schrödinger equation for initial data which are strongly oscillating. This has lead us to the time-dependent Hamilton-Jacobi equation. We now show that the stationary Hamilton-Jacobi equation has a role in the determination of approximate eigenfunctions ψ_E and eigenvalues E of the *stationary* Schrödinger equation.

Since the solutions of the stationary Hamilton-Jacobi equation are generating functions for a family of symplectic transformations which are continuously connected to the identity, we expect that this analysis play a role in quantizations methods based on lagrangian manifolds.

We begin also here by discussing the one-dimensional case. Let the classical hamiltonian be

$$H = \frac{p^2}{2m} + V(q) \tag{65}$$

with corresponding quantum mechanical hamiltonian

$$Op(H) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \tag{66}$$

We seek, for $\hbar$ small, stationary solutions $\phi_E(x,t) = e^{-iEt}\phi_E(x)$.

The function $\phi_E(x)$ must satisfy the (eigenvalue) equation

$$(Op(H) - E)\phi_E = 0, \qquad E = \hbar\omega \tag{67}$$

If $V = 0$ the solution is explicitly given by

$$\phi_E(x) = e^{i\xi x}, \qquad \xi = \pm\frac{\sqrt{2mE}}{\hbar} \tag{68}$$

If $V(x)$ varies slowly on a scale adapted to $\hbar$ it is natural to assume also in the case $V \neq 0$ that the solution has the form

$$\phi_E(x) = e^{\frac{iS_E(x)}{\hbar}} a_E(x;\hbar) \tag{69}$$

and to try to determine the functions $S_E(x)$ and $a_E(x,\hbar)$ as power series (may be asymptotic) in $\hbar$.

We study first the case $a_E(x) \equiv 1$. We will obtain *non-normalizable solutions* that will approximate the generalized eigenfunctions belonging to the continuum spectrum. Equation (67) reads now

$$(Op(H) - E)\phi_E = 0 = [\frac{(S'_E(x))^2}{2m} + (V(x) - E) - i\frac{\hbar}{2m}S''_E(x)]e^{i\frac{S_E(x)}{\hbar}} \tag{70}$$

We shall solve this equation under the assumption that the solution can be written as a power series in $\hbar$. To order zero one has

$$\frac{(S'_E(x))^2}{2m} + V(x) = E \Rightarrow S'_E(x) = \pm\sqrt{2m}\sqrt{E - V(x)} \tag{71}$$

This solution is defined only for x such that $V(x) \leq E$. With this choice of ϕ one has

$$(Op(H) - E)\phi_E = O(\hbar) \tag{72}$$

Setting $dS_E = S'_E(x)dx$ one verifies that the range of $S'_E(x)$ is contained in the subset of phase space where $\{q, p|H(q, p) = E\}$.

This analysis generalizes without difficulty to the case of N degrees of freedom. The equation for S_E is still the stationary Hamilton-Jacobi equation

$$\frac{|\nabla S_E(x)|^2}{2m} + V(x) = E \tag{73}$$

Definition 2 (*admissible function*) We shall call *admissible* a function $S_E(x)$ if it is solution of (73). ◇

In case E belongs to the point spectrum of $Op(H)$ we seek a normalized solution. We make the *Ansatz*

$$\phi_E(x) = e^{i\frac{S_E(x)}{\hbar}} a_E(x), \qquad a_E(x) \in L^2 \tag{74}$$

where $S_E(x)$ is admissible.

Therefore

$$(Op(H) - E)\phi_E = \hbar[\Delta\ S_E + 2\frac{\partial a_E}{\partial x_k}.\frac{\partial S_E}{\partial x_k}] + O(\hbar^2) \tag{75}$$

We conclude that *if* S_E *is admissible and moreover*

$$\Delta\ S_E + 2\frac{\partial a_E}{\partial x_k}.\frac{\partial S_E}{\partial x_k} = 0 \quad k = 1, \dots N \tag{76}$$

then $|(Op(H) - E)\phi_E| = O(\hbar^2)$. Therefore we have obtained an approximation to order $\hbar$ of the eigenvector of the stationary Schrödinger equation to the eigenvalue E.

If $S_E(x)$ is admissible the Eq. (76) is the *transport equation* for $a_E(x)$.

In case $N = 1$ the transport equation takes the form $a_E(x)'' + 2a_E'(x)S_E'(x) = 0$ or equivalently $(a_E(x)^2 S_E')' = 0$. Taking into account (73) the solution is

$$a_E(x) = \pm\sqrt{\frac{C}{4(E - V(x))}} \tag{77}$$

Denote by Γ_E the one-dimensional manifold $(q, p) : H(q, p) = E$. The points in which Γ_E does not have a smooth projection on the q axis are those for which $a_E(x) \to \infty$ (caustics, or point of inversion of motion in configuration space).

One can obtain a better approximation: $|(\hat{H} - E)\phi_E^M|_2 = O(\hbar^{M+1})$ by the ansatz

$$\phi(x) = e^{i\frac{S_E(x)}{\hbar}}(a_0^E(x) + \hbar a_1^E(x) + \hbar^2 a_2^E(x) + \dots + \hbar^M a_M^E(x, \hbar)) \quad x \in R^d \tag{78}$$

where the coefficients $a_k^E(x),\ k = 0 \dots M - 1$ satisfy a suitable system of linear inhomogeneuos equations which can be derived making a (formal) expansion in power of ϵ of the transport equation and of its solution.

7 Geometric Intepretation

It is interesting to connect this analysis of the WKB method to the formalism of geometric quantization; [12, 14] this gives a different prospective on semiclassical theory.

From a geometric point of view if $S : R^N \to R$ is admissible (satisfies the Hamilton-Jacobi equations) the manifold $\mathcal{L} = Range\ (dS)$ has the following properties

(1) $\mathcal{L}$ is a N-dimensional lagrangian manifold immersed in $H^{-1}(E)$
(2) The pull-back to $\mathcal{L}$ of the form $\alpha = \sum_{i=1}^N dp_i \wedge dq_i$ is a closed two form.

(3) The restriction to $\mathcal{L}$ of the natural projection $T^*R^d \to R^N$ induces an isomorphism between $\mathcal{L}$ and R^N

We shall use the following lemma:

Lemma 2 (Hamilton-Jacobi) *The function $H : R^{2N} \to R$ is locally constant on a lagrangian manifold $\mathcal{L}$ iff the vector field X_H is tangent to $\mathcal{L}$.* $\diamondsuit$

Proof

$$X_H \equiv \sum (\frac{\partial H}{\partial p_k}\frac{\partial}{\partial q_k} - \frac{\partial H}{\partial q_k}\frac{\partial}{\partial p_k}) \tag{79}$$

and therefore

$$d\alpha.X_H = dH, \qquad \omega(X_H, v) = -dH(v) \quad v \in T\ R^N$$

$$T_z\ \mathcal{L} \in Ker(dH) \Rightarrow T_z\ \mathcal{L} \cup X_H(p) \subset Ker(d\alpha) \Rightarrow X_H(p) \in T_z(\mathcal{L}) \tag{80}$$

$\heartsuit$

We return now to the transport equation.

Remark that the orthogonal projection Π of the hamiltonian vector field X_H on the tangent in x to the manifold L is

$$X_H = \sum_{k=1}^{N} (\frac{\partial S}{\partial p_k}\frac{\partial}{\partial q_k} - \frac{\partial S}{\partial q_k}\frac{\partial}{\partial p_k}) \tag{81}$$

and therefore one can write

$$div(a^2 X_H^q) = 0, \qquad X_H^q = \{\frac{\partial S}{\partial q_i}\} \tag{82}$$

It follows that the Lie derivative of $a^2 dx$ with respect to the vector field X_H^x vanishes. This means that the volume element

$$\Pi^{-1}(a^2 dx) \tag{83}$$

is invariant for the hamiltonian flow generated by H.

Since $a^2(x)$ is a Liouville density it is natural to consider $a(x)$ as *half-density*.

We have seen that if S_E is an admissible function and $a_E(x)$ is a half-density which is invariant under the hamiltonian flow of H, then

$$a_E(x) e^{i\frac{S_E(x)}{\hbar}} \tag{84}$$

approximates to second order in $\hbar$ a solution of the stationary Schrödinger equation with energy E.

If N = 1 this approximate solution is

$$\phi_{appr.} = \sqrt{n_+} e^{\frac{i}{\pi}\int_{x_0}^{x}\sqrt{2\,m(E-V(x)}dy} + \sqrt{n_-} e^{\frac{-i}{\pi}\int_{x_0}^{x}\sqrt{2\,m(E-V(x)}dy} \tag{85}$$

where $n_\pm(x) = \frac{A_\pm}{\sqrt{2m(E-V(x)}}$.

Observe that the semiclassical W.K.B. state we have described is *regular* on the Lagrange manifold $\mathcal{L}$ but its projection on R^N may have singular points.

8 Semiclassical Quantization Rules

We describe briefly an application of the W.K.B. method to completely integrable systems, and more specifically the semi-classical description of quantum systems which correspond to classical systems with periodic orbits. Some Authors in this context refer to this method as J.W.K.B. method, adding the initials of Jeffreys who introduced a similar procedure in the year 1923.

We shall follow in part the review article [1] and the book [7]; we refer to them for a more complete analysis. A rigorous treatment for a restricted class of potentials can be found in [2].

We shall also describe the semi-classical origin of the Bohr-Sommerfeld quantization rules. A detailed description of this subject is outside the scope of this Lecture. For further analysis one can see [11]. We only mention that the Maslov index and the W.K.B. method are treated in [3] in the context of oscillating integrals and of lagrangian manifolds.

Consider first system with one degree of freedom. The stationary Schrödinger equation we consider is

$$\frac{d^2}{dx^2}\phi + \frac{2}{\hbar^2}(E - V(x))\phi = 0 \qquad x \in R \tag{86}$$

We look for a solution which can be written as

$$\phi_E(x) = e^{i\frac{S_E(x)}{\hbar}}$$

For the function S one obtains

$$\frac{dS}{dx}^2 - i\hbar S''(x) = 2m(E - V(x)) \tag{87}$$

When $\hbar$ is very small, in the intervals in which $E - V(x) > 0$ we can try to approximate the solution with a solution of the equation $\frac{dS}{dx}^2 = 2m(E - V(x))$, i.e. with

$$S = \pm \int_a^x \sqrt{2m(E - V(y))}dy = \pm \int_a^x p(y)dy, \qquad p(y) \equiv \frac{\partial S}{\partial y} \tag{88}$$

where the choice of a corresponds to the choice of an additive constant for S.

The approximation we have introduced is valid if for the state considered the following estimate holds

$$\hbar(\phi, |S''|\phi) = \hbar(\phi, \frac{\partial p}{\partial x}\phi) \ll (\phi, p^2\phi)$$

Therefore *for these states* the second term on the r.h.s. in (87) can be neglected. This condition determines the possibility to make use of the W.K.B. method.

Expanding formally the solution of (88) in power series of $\hbar$ one obtains to first order $S = S_0 + S_1$ with

$$S_0(x) = \pm \int_a^x \sqrt{2m(E - V(y)}dy, \qquad S_1(x) = ilog\sqrt{S_0'} + C \tag{89}$$

To this order of approximation therefore

$$\phi_E(x) = C(S_0'(x))^{-1} e^{\frac{\pm i}{\hbar}\int_a^x \sqrt{2m(E-V(y)}dy} \tag{90}$$

where C is a normalization factor. If $V(x) \in C^2$ with $V(0) = E$ and there are constants b and $\delta > 0$ such that $V(x) > E + \delta$ for $x > b$ and $V(x) < E - \delta$ for $x < -b$, then the W.K.B. approximation is justified $|x| > b$ and $b \gg \hbar$ [13, 14].

Under this conditions (90) gives for $x < -b$ two independent polynomially bounded solutions of (87) and for $x > b$ one square integrable solution. If b is comparable to $\hbar$ the W.K.B. condition *is not satisfied* in the interval $|x| < b$ and therefore the solutions cannot be approximated by (90). On the other hand for each value of E Eq. (90) admits two bounded solutions $-b < x < b$. To obtain a regular solution of the Schrödinger equation one must connect smoothly the solution outside $[-b,\ b]$ to a linear combination of the two solutions of (90) [14].

In the W.K.B. method this connection is determined through a study of the asymptotic behaviour for $\hbar \to 0$ of the solution of a suitable *comparison equation*. This leads to *connection formulas*. In general the search for an optimal comparison equation is done through a change of coordinates $x \to y = y(x)$ (the function $y(x)$ is known as *mapping function*).

Through the change of coordinates $x = \hbar y$ the stationary Schrödinger to energy E takes the form

$$\frac{d^2}{dy^2}\psi(y) + v(y)\psi(y) = 0, \qquad v(y) = E - V(\hbar y) \tag{91}$$

and one looks for solutions comparing this with a simpler equation

$$\frac{d^2}{dz^2}\phi(z) + w(z)\psi(z) = 0 \tag{92}$$

where the function w is chosen such that approximate solutions of Eq. (46) are easily constructed and the qualitative features of the solution are not modified substantially.

Through a further change of coordinates $y \to z$ the function $\psi(y)$ can be written (at least formally and locally) $\psi(y) = \frac{dz}{dy}^{-\frac{1}{2}}\phi(z(y))$. The *mapping function* is then the solution of the equation

$$z(y) = w(z)\frac{dz^2}{dy^2} - (\frac{dz}{dy})^{\frac{1}{2}}\frac{d^2}{dy^2}[(\frac{dz}{dy})^{-\frac{1}{2}}]. \tag{93}$$

8.1 One Point of Inversion

In the case of potentials for which the classical motion in configuration space with energy E admits only one inversion point (that we assume to be the origin) a good approximation is obtained by choosing $w(z) = -z$.

In this case the comparison equation is the Airy equation $\frac{d^2\psi}{dy^2} + y\psi(y) = 0$. The change of coordinates is given by

$$z^{\frac{3}{2}} = \frac{3}{2}\int_0^y (\frac{V(\hbar y') - E}{\hbar^2})^{\frac{1}{2}} dy' \tag{94}$$

and the approximate solution of the Schrödinger equation is

$$\phi(y) = [\frac{z^2(y)}{E - V(\hbar y)}]^{\frac{1}{4}}[\alpha Ai_1(z(y)) + \beta Ai_2(z(y))] \tag{95}$$

where α β are constant which are determined by the conditions one imposes at $\pm\infty$. We have denoted with Ai_1 and Ai_2 two independent Airy functions, solutions of Airy's equation

$$\frac{d^2\phi}{dy^2} + y\phi(y) = 0 \tag{96}$$

The boundary condition we impose is that the solution is polynomially bounded (we assume that in a neighborhood of E the spectral measure of $Op(H)$ is absolutely continuous and has multiplicity one).

The two Airy functions are different boundary values of a function $F(z)$ which is analytic in the complex plane; its coefficients of the power series expansion in z are known explicitly. This allows an explicit determination of the connection formulas [11].

Moreover the asymptotic behaviour of the Airy functions when $y \to \pm\infty$ is known. Both have an oscillatory behaviour with an asymptotic phase difference of $\frac{\pi}{4}$.

The function Ai_1 converges exponentially when $y \to \infty$ while Ai_2 diverges exponentially.

For every value of μ one has the asymptotic behaviour (denoting by $\to$ and $\leftarrow$ the limits $x \to \infty$ and $x \to -\infty$)

$$\frac{cos(|w| - \frac{1}{4}) + \mu}{(V-E)^{\frac{1}{4}}} \leftarrow cos\mu\ Ai_1(\sigma) + sen\mu\ Ai_2(\sigma) \to \frac{e^{-|w|}cos\mu + 2e^{|w|}sen\mu}{(V-E)^{\frac{1}{4}}} \tag{97}$$

where

$$w = \int_0^x [\frac{(E - V(y))}{\hbar^2}]^{\frac{1}{2}} dy, \qquad \sigma = (\frac{3}{2}w)^{\frac{2}{3}} \tag{98}$$

8.2 Two Points of Inversion

If the potential is such that in the classical motion has two inversion points in configuration space (as is the case for closed orbits) the W.K.B. approximation is less accurate.

The simplest comparison potential in this case is a harmonic one and the comparison equation is

$$\frac{d^2}{dz^2}\phi + (t - z^2)\phi = 0 \tag{99}$$

which contains a parameter t which must be chosen as a function of the energy E.

For this potential, denoting by $x_\pm$ the turning points of the classical motion and defining t trough

$$\frac{\tau\pi}{2} = \int_{x_-}^{x_+} \frac{\sqrt{E - V(x)}}{\hbar} dx \equiv S(x_-, x_+) \tag{100}$$

the transformation of coordinates $x \to y$ is given implicitly by

$$S(x_+, x) = \hbar \int_{-\sqrt{\tau}}^{y(x)} (t - z^2)^{\frac{1}{2}} dz \tag{101}$$

Two linearly independent solutions are the two parabolic functions $D_{t-\frac{1}{2}}(\pm y\sqrt{2})$ [14] with well known asymptotic behaviour given by convergent power series.

9 Approximation Through the Resolvent

An alternative procedure to obtain an approximation for the eigenvalues may come from the observation that the point spectrum of a operator is characterised by the poles of its resolvent [8].

One can therefore try to express its resolvent as an integral over classical trajectories of a function which has strong oscillations as $\hbar \to 0$, and use the stationary phase methods to prove that the main contribution to the integral comes from classical trajectories which are closed. Therefore the eigenvalues are approximately given by the energies corresponding to closed classical trajectories. This correspond to the Bohr-Sommerfeld quantization rules, with a Maslov index correction term due to the presence of caustics [11, 14].

This program faces severe difficulties. We have seen in "Lecture 5: Automorphisms; Quantum Dynamics; Theorems ofWigner, Kadison, Segal; Continuity and Generators" that the propagator $e^{it\Delta}(x, y)$ (fundamental solution) can be *formally* given as an oscillating integral over classical trajectories

In the free case there is exactly one critical point for the oscillating phase due to the fact that for each choice of t_0, x_0, y_0 there exist only one classical trajectory that joins x_0 and y_0 in time t_0. This trajectory corresponds to energy $E_0 = \frac{p_0^2}{2m}$ with $t_0 p_0 = x_0 - y_0$.

In the case of the harmonic operator in R^d with classical hamiltonian $H_{class} = \frac{1}{2}p^2 + \frac{1}{2}q^2$ and quantum hamiltonian

$$Op(H_{osc}) = -\frac{1}{2}\Delta + \frac{1}{2}x^2 - \frac{d}{2} \tag{102}$$

the propagator is (Mehler's formula)

$$e^{-itOp(H_{osc})}(x, y) = \frac{e^{-id(1+2m\pi}}{|2\pi sent|^{\frac{d}{2}}} e^{[\frac{i}{sent}((x^2+y^2)cos\frac{t}{2}-xy)]}, \qquad t \in (m.m+1) \;\; m \in Z \tag{103}$$

If $t = m\pi, m \in Z$ one has

$$e^{-iOp(H_{osc})(t,x,y)} = e^{-im\frac{d\pi}{2}}\delta(x - (-1)^m y) \tag{104}$$

This formulas are easy to verify if one notices that $Op(H_{osc}) = \sum_{k1}^{d} a_k^* a_k$ where the operators $a_k = \frac{1}{\sqrt{2}}(x + \frac{d}{dx})$.

The classical trajectories are periodic and it is easy to verify that also in this case the propagator $e^{-itH_{osc}}(t, x, y)$ can be expressed *formally* as the sum of the integrals of e^{iS_γ} (S is the action integral) over those classical trajectories that connect x to y in time t [8].

Since all the trajectories have inversion points, one must keep into account Maslov's index. Making use of the relation between propagator and resolvent one

can *formally* represent the kernel of the resolvent $(Op(H_{osc}) - z)^{-1}$, $Imz \neq 0$ as an integral over classical trajectories. The integral in this case is better defined because the factor e^{izt}, $Imz > 0$ provides convergence for positive times.

An interesting procedure aimed at constructing a semiclassical approximation for the resolvent and the propagator, and thereby obtain a *formal* justification of the W.K.B. quantisation rules can be found in [2]. These Authors make use of stationary phase results, modify the scale of times and exchange the order of the operations $N \to \infty$ and $\hbar \to 0$.

In this way one derives for the kernel of the resolvent $(Op(H) - E)^{-1}(x, y)$ the semi-classical formula

$$(Op(H) - E)^{-1}(x, y)_{semicl} = \frac{2\pi\hbar^{\frac{(d-1)}{2}}}{i^{\frac{d}{2}}} \sum_r |\pi \Delta_{x,y;r} S(x, y; E)|^{\frac{1}{2}} H_d^1(S_r(x, y; E)) - \frac{M_r \pi}{2}) \tag{105}$$

where the sum is over all closed classical orbits of energy E that connect x and y, M_r is the relative Maslov index, H_d^1 is the first Hankel function and $S_r(x, y, E)$ is the integral of the Action along the rth orbit of energy E.

The function $\Delta_{x,y;r}$ depends on the orbit and is given as

$$\Delta_{x,y;r} = -\Big(\frac{\partial^2 S_r(x, y; E)}{\partial E^2}\Big)^{1-\frac{1}{d}} \det\Big[\frac{\partial^2 S}{\partial E^2}\frac{\partial^2 S_r}{\partial x_i \partial y_j} - \frac{\partial^2 S_r}{\partial x_i \partial E}\frac{\partial^2 S_r}{\partial y_j \partial E}\Big] \tag{106}$$

This expression simplifies in one dimension for classical hamiltonians. In this case one can derive, with further approximations, the Bohr-Sommerfeld quantisation rules.

In the one-dimensional case $H(p, x) = \frac{p^2}{2m} + V(x)$ one has

$$S_{x_1,x_2}(E) = 2\int_{x_1}^{x_2} \sqrt{2m(E - V(x)}dx, \qquad \Delta = \frac{m}{2(E - V(x))} \tag{107}$$

and then

$$n(E) = \frac{1}{\pi\hbar} \sum_{r=-\infty}^{\infty} cos(\frac{rS(E)}{\hbar}) \int_{x_1}^{x_2} [\frac{m}{2(E - V(x))}]^{\frac{1}{2}} = \sum_0^{\infty} \delta(E - E_n) \tag{108}$$

with E_n such that $S(E_n) = (2n + 1)\pi\hbar$, in accordance with the Bohr-Sommerfeld quantization rule.

The semi-classical approximation is most often used to derive approximate formulae for the density of bound states E_n by the (Gutzwiller) trace-formula

$$\rho(E) = \sum_n \delta(E - E_n) = -\frac{1}{\pi} Im \int dx (H - E)_+^{-1}(x, x), \quad x \in R^d \tag{109}$$

were we have denoted by $(H-E)^{-1}_{+}(x,y)$ the limit when $Imz \to 0_{+}$ of the kernel of the operator $(H-z)^{-1}$.

The formula (109) is a semiclassical formula for *the trace* (the sum of the projections on the eigen-spaces of the Hamiltonian, assuming that it has a point spectrum).

References

1. Berry, M., & Mount, K. (1972). *Reports on Progress in Physics*, *35*, 315–397.
2. Combescure, M., & Robert, D. (2012). *Coherent states*. Dordrecht: Springer.
3. Duistermaat, J. J. (1974). *Communications on Pure and Applied Mathematics*, *27*, 207–281.
4. Evans, L. (1998). *PDE, Graduate studies in mathematics*. Providence: AMS.
5. Fedoryiuk, M. (1999). *P.D.E. Encyclopedia of mathematical sciences* (Vol. 34). New York: Springer.
6. Fujiwara, D. (1980). *Duke Mathematical Journal*, *47*, 559–600.
7. Frohman, N., & Frohman, P. O. (1965). *J.W.K.B. approximation*, Amsterdam: North-Holland.
8. Gutzwiller, M. (1967). *Journal of Mathematical Physics*, *8*, 1979–2000.
9. Heading, J. (1962). *An Introduction to Phase-space Methods*, London: Methuen.
10. Hormander, L. (1967). Pseudodifferential operators and hypoelliptic equations. *Proceedings of the Symposium Pure and Applied Mathematics*, *10*, 138–183.
11. Keller, J. (1958). *Annals of Physics*, *4*, 180–188.
12. Maslov, V. (1994). *The complex W.K.B. method, Progress in physics N.16*, Basil: Birkhauser.
13. Voros, A. (1989). *Physical Review A*, *49*, 6814–6825.
14. Wittaker, E., & Watson, G. (1946). *A course of modern analysis*. Cambridge: Cambridge University Press.

Lecture 17: Kato-Rellich Comparison Theorem. Rollnik and Stummel Classes. Essential Spectrum

We have seen in "Lecture 5: Automorphisms; Quantum Dynamics; Theorems of Wigner, Kadison, Segal; Continuity and Generators" that when V is bounded Duhamel's formula can be used to prove that the operator $-\Delta + V(x), x \in R^d$ is self-adjoint and therefore it is the generator of a unitary group of operators.

When V is unbounded, as is almost always the case for physically interesting problems, this method does not work, since generically the Duhamel series does not converge. We must take a more general approach.

1 Comparison Results

We review first some general *comparison* result.

Let A be a self-adjoint operator on a separable Hilbert space $\mathcal{H}$.

Definition 1 (*relative smallness*) An operator B is $A-$*small* (denoted $B \prec A$) if there exist positive constants $0 \leq a < 1, b \in R^+$ such that

$$|B\phi| \leq a|A\phi| + b|\phi| \quad \forall \phi \in D(A) \tag{1}$$

If B is closed this implies $D(A) \subset D(B)$. ◇

We will call *bound* of B relative to A (denoted by $a_{A,B}$) the infimum of the positive constants a for which (1) holds *for some value of* b.

Definition 2 (*infinitesimal*) The operator B is *infinitesimal* with respect to A (denoted $B \prec\prec A$) if $a_{A,B} = 0$ ◇

Theorem 1 (Kato-Rellich) [4] *If A is self-adjoint and B is symmetric and A-small, then $A + B$ defined on $D(A) \cap D(B)$ is essentially self-adjoint* on any core of A.

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_17

Moreover if $A \geq mI$ *one has* $A + B \geq m - min_{a,b}(\frac{b}{1-a}, a|m| + b)$ *where the minimum is taken over all pairs* a, b *which satisfy (1).* ◇

Proof It suffices to prove that for μ real and sufficiently large $Range\ (A+B+i\mu) \equiv \mathcal{H}$ On $D(A)$ the following identity holds for any $\mu \in R$

$$(1 + B(A + i\mu)^{-1})(A + i\mu) = A + B + i\mu \tag{2}$$

Since A is self-adjoint, $A + i\mu$ is invertible. We must prove that also $(1 + B(A + i\mu)^{-1})$ is invertible for μ sufficiently large; it suffices to prove $|B(A + i\mu)^{-1}| < 1$. By (1), since $(A - i\mu)^{-1}\phi \in D(A)$ one has

$$|B(A + i\mu)^{-1}\phi| \leq a|A(A + i\mu)^{-1}\phi| + |(A + i\mu)^{-1}\phi|$$

$$\leq a|A(A + i\mu)^{-1}\phi| + \frac{b}{\mu}|\phi| \tag{3}$$

By the spectral theorem $\|A(A + i\mu)^{-1}\| < 1$ for all $\mu \in R - \{0\}$. If μ large enough the right hand side of (3) can be made smaller than $(1 - \delta)|\phi|$ for any δ with $0 < |\delta| < 1$.

The estimate on the lower bound for the spectrum of $A + B$ is obtained considering $A + \mu I - MI$ and estimating the smallest value of μ such that $A + b + \mu$ is invertible. ♡

As an application of Theorem 1 to Schrödinger operators we prove the following corollary. We shall use the notation $V \in L^\infty(R^3) + L^2(R^3)$ to indicate that there are potentials $V_2(x) \in L^1(R^3)$ and $V_1 \in L^2(R^3)$ such that $V(x) = V_1(x) + V_2(x)$.

Corollary *Let the potential* V *satisfy*

$$V = V_1 + V_2, \quad V_1 \in L^\infty(R^3), \quad V_2 \in L^2(R^3) \tag{4}$$

Then $-\Delta + V$ *is self-adjoint with domain* $D(\Delta)$ *and is essentially self-adjoint on any core of* Δ *(e.g. on* $C_0^\infty(R^3)$*).* ◇

Proof We verify that V is infinitesimal with respect to $-\Delta$. For this we prove first that if $\phi \in L^2$ for any $a > 0$ there exists $b > 0$ such

$$|\phi|_\infty \leq a|\Delta\phi|_2 + b|\phi|_2 \tag{5}$$

This inequality (technically a Sobolev inclusion) is easy to prove using the Fourier transform.

One has $|\phi|_\infty \leq \int |\hat{\phi}(k)|d^3k$. Choose $\delta > 0$ and consider separately the contribution to the integral of the region $|k| \geq \delta$ and $|k| > \delta$. From Schwarz's inequality

$$\int_{|k|\geq\delta}|\hat{\phi}(k)|d^3k \leq [\int_{|k|\geq\delta}|k^2\hat{\phi}(k)|^2d^3k]^{1/2}[\int_{|k|\geq\delta}|k|^{-4}d^3k]^{1/2} \leq C(\delta)|\Delta\phi|_2 \tag{6}$$

with $lim_{\delta\to\infty}C(\delta) = 0$. On the other hand

$$\int_{|k|<\delta}|\hat{\phi}(k)|d^3k \leq [\int_{|k|<\delta}|\hat{\phi}(k)|^2d^3k]^{1/2}(\frac{4\pi}{3}\delta^3)^{1/2}$$

A proper choice of δ concludes the proof of (6).

Let now $\phi \in L^\infty(R^3)$ and

$$V = V_1 + V_2, \quad V_1 \in L^2(R^3), \quad V_2 \in L^\infty(R^3) \tag{7}$$

and let $\phi(x)$ be such that $V\phi \in L^2(R^3)$. Then

$$|V\phi|_2 \leq |V_1\phi|_2 + |V_2\phi|_2 \leq |\phi|_\infty|V_1|_2 + |\phi|_2|V_2|_\infty$$

$$\leq C(\delta)|\Delta\phi|_2 + [|V|_2| + b(\delta) + |V|_\infty]\phi_2| \tag{8}$$

where $b(\delta) < \infty$ for any finite value of δ. Therefore by (6)

$$|(-\Delta + V)\phi|_2 \leq (1 - C(\delta))(\Delta\phi|_2 + b_1|\phi|_2)$$

Since $C(\delta)$ can be made smaller that one by choosing suitably δ, the theorem is proved. ♡

Conditions (4) on V are satisfied e.g. if $V \in L^2_{loc}(R^3)$ and there exist C and γ such that $|V(x)| \leq C$ for $|x| > \gamma$. This holds e.g. if $V(x) = c|x|^{-\alpha}$, $0 < \alpha < 3/2$. The same is true for the operator

$$-\sum_i \Delta_i + \sum_{i\neq j} V_{i,j}(x_i - x_j)$$

under the same assumptions on each $V_{i,j}$.

The Kato-Rellich theorem can be generalized to make it applicable to the N-body problem with Hamiltonian

$$H_{N,Z} = -\sum_{j=1}^N \Delta_{x_j} + \sum_{j<k}^N |x_j - x_k|^{-1} - \sum_{j=1}^N Z|x_j|^{-1}, \quad x_I \in R^3 \tag{9}$$

This system describes the interaction of N electrons with a nucleus of charge Z placed at the origin.

The hamiltonian $H_{N,Z}$ does not satisfy the assumptions of the Kato-Rellich theorem; the following theorem by T. Kato gives a generalization that includes this case.

Theorem 2 [4] *Let*

$$V_{j,k}, \quad W_j \in L^2(R^3) + L^\infty(R^3), \qquad j, \quad k \quad = 1, \dots N \tag{10}$$

and consider the hamiltonian on $L^2(R^{3N})$

$$H = \sum_{j=1}^{N} (-\Delta_{x_j} + W(x_j)) + \sum_{j<k} V_{j,k}(x_j - x_k) \tag{11}$$

The operator H *is self-adjoint and* $D(H) \equiv H^2(R^{3N})$. ◇

Proof We have already seen that under the hypothesis we have made on the potential one has for every $\delta > 0$ and every $\phi \in H^2(R^{3N})$

$$|W_j\phi|_2 \le C(\delta)|\Delta_{x_j}\phi|_2 + b(\delta)|\phi|_2, \qquad |V_{j,k}\phi|_2 \le C(\delta)|\Delta_{x_j - x_k}\phi|_2 + b(\delta)|\phi|_2, \tag{12}$$

where $C(\delta) \to 0$ when $\delta \to \infty$.

One verifies easily (e.g. using Fourier transform) that

$$|\Delta_{x_j - x_k}\phi| \le 2(|\Delta_{x_j}\phi| + |\Delta_{x_k}\phi|)$$

for every choice $j, \ k$. Therefore

$$|(\sum_j W_j + \sum_{j<k} V_{j,k})\phi| \le (NC(\delta) + N(N-1)C(\delta))|\sum_j \Delta_{x_j}\phi| + \frac{N(N-1)}{2} b(\delta)|\phi|$$

$$\le N^2 C(\delta)|\sum \Delta_j \phi| + b_1(\delta)|\phi| \tag{13}$$

Choosing δ such that $N^2 C(\delta) < 1$ the thesis of the theorem follows from the Kato-Rellich theorem. ♡

The Kato-Rellich theorem has a partial generalization to the case $a = 1$.

Theorem (Wurst) *If* A *is self-adjoint and* B *is symmetric with bound 1 relative to* A

$$|B\phi| \le |A\phi| + b|\phi|, \qquad \forall \phi \in D(A), \quad b > 0 \tag{14}$$

then $A + B$, *defined at first on* $D(A) \cap D(B)$, *is* essentially *self-adjoint on the domain of* A. ◇

Proof We prove that if $\xi \in \mathcal{H}$, $(A + B + i)^*\xi = 0$, then $\xi = 0$. The proof is obtained by contradiction.

Suppose that there exist a vector $\xi \neq 0$ which satisfies $(A + B + i)^*\xi = 0$. For every $t \in [0, 1)$ the operator tB is A-small. Therefore $A + tB$ is self-adjoint

with domain $D(A)$. Therefore there is a unique $\phi_t \in D(A)$, $|\phi_t| \leq |\xi|$, such that $(A + tB + i)\phi_t = \xi$.
Define

$$\psi(t) = \xi - (t-1)B\phi(t) \tag{15}$$

Then $|A\psi_t| \leq |(A + tB)\phi_t| + t|A\phi_t| + tb|\phi_t|$ and therefore

$$|A\psi_t| \leq \frac{1}{1-t}|(A + tB)\phi_t| + \frac{tb}{1-t}|\phi_t| \tag{16}$$

It follows that $(1-t)|A\psi_t|$ remains bounded when $t \to 1$.
Define

$$\psi_t = \xi - (t-1)B\phi_t \tag{17}$$

Let $\eta \in D(A)$, $lim_{t\to 1}(\psi_t - \xi, \eta) = lim_{t\to 1}(1-t)(B\phi_t, \eta) = 0$. Since $D(A)$ is dense one has

$$\xi = weak - lim_{t\to 1}\psi_t \tag{18}$$

On the other hand

$$(\xi, \xi) = ((A + tB + i)\phi_t, \xi) = (t-1)(B\phi_t, \xi) = (\psi_t - \xi, \xi) \tag{19}$$

We have used $((A + B + i)\phi, \xi) = (\phi, (A - B + i)^*\xi) = 0$.

The left hand side does not depend on t, the right hand side converges to zero when $t \to 1$. Therefore $(\xi, \xi) = 0$ and then $\xi = 0$. ♡

As an application of Wurst's theorem we prove now that the hamiltonian of the anharmonic oscillator

$$H = -\frac{1}{2}\frac{d^2}{dx^2} - 1 + x^2 - 1 + ax^4, \quad a > 0 \tag{20}$$

is an *essentially self-adjoint operator* on $C_0^\infty(R)$. The same proof applies to the case the term ax^4 is substituted by a polynomial $P_{2M}(x)$ of degree $2M$ in which the term of highest degree is positive.

Define

$$N \equiv -\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2 - 1, \quad Y = x^4, \quad Z = cN^2 \tag{21}$$

where the constant c will be chosen suitably.

Remark that N has as eigenvalues the positive integers (N is the number operator in the Fock representation) and that $N + Z$ is a self-adjoint operator.
We shall prove

(a) If c is sufficiently large Y is small relative to $N + Z$.
(b) For a suitable choice of b one has $|Z\phi| \leq |(N + Y + Z)\phi| + b|\phi|$.

From (a) and (b) and Wurst's theorem it follows that $X + Y \equiv (X + Y + Z) - Z$ is essentially self-adjoint on $D(N + Z)$.

The introduction of the auxiliary operator Z serves the purpose of making Y small with respect to $X + Z$, an operator of which we know the analytic vectors (notice that Y is not X-small).

To prove (a) and (b) it is convenient to define

$$\sqrt{2}\, a = x + \frac{d}{dx}, \quad \sqrt{2}\, a^* = x - \frac{d}{dx}, \quad -\frac{d^2}{dx^2} = 2a^*a + 1 \tag{22}$$

Then $Y \equiv \frac{1}{4}(a + a^*)^4$ and it is ready seen that if chooses c sufficiently large one has the following inequalities

$$|aN^{-1/2}| \leq 1, \qquad |Y\phi| \leq 4|(N+1)^2\phi| \leq \frac{1}{2}|(N + \frac{Z}{c})\phi| \tag{23}$$

This proves point (a).

To prove (b) an easy computation gives

$$\begin{aligned}(N + Y + Z)^2 &= (N + Y)^2 + Z^2 + 2cN^3 + 2cNYZ \\ &\quad + 2c[N, [N, Y]] \geq Z^2 + 2cN^3 + 2c[N, [N, Y]]\end{aligned} \tag{24}$$

On the other hand $[N, Y] = [N, (a + a^*)^4] = 4N(a - a^*)^3$ and therefore $[N[N, Y]] = 12(a + a^*)^2$. Therefore if one chooses p sufficiently large

$$(N + Y + Z)^2 \geq Z^2 + 2cN^3 - 3c(N + 1/2I)^2 \geq Z^2 - pI \tag{25}$$

This proves (b) and concludes the proof that $-\frac{d^2}{dx^2} + x^2 + x^4$ is essentially self-adjoint on $D(N^2)$.

The Kato-Rellich theorem can be used to prove that some Schrödinger operators with vector potential is essentially self-adjoint on C_0^∞. One has in particular the following theorem of Kato.

Theorem 3 [4] *If*

$$\begin{aligned}&A_i \in L^4(R^3) + L^\infty(R^3), \quad \partial_i A_k \in L^4(R^3) + L^\infty(R^3), \\ &V \in L^2(R^3) + L^\infty(R^3), \quad i, k = 1, 2, 3\end{aligned}$$

then $H \equiv (\nabla + A)^2 + V$ *is essentially self-adjoint on* $C_0^\infty(R^3)$. ◇

Proof From the previous inequalities

$$V \prec\prec \Delta, \quad \nabla \cdot A \prec\prec \Delta, \quad A^2 \prec\prec \Delta.$$

We must show

$$A \cdot \nabla \prec\prec \Delta \tag{26}$$

From Schwartz inequality for every value of the parameter α

$$|A_k \frac{\partial}{\partial x_k}\phi|_2 \leq |A_k|_4 |\frac{\partial \phi}{\partial x_k}|^{4/3} \leq |A_k|_4 |p_k \hat{\phi}|_{4/3}$$

$$\leq |A_k|_4 \ |(1+|p|)^{-\alpha}|_4 \ | \ |(1+|p|)^{\alpha} p_k \hat{\phi}|_2 \tag{27}$$

The second factor is finite if $\alpha \geq 3/4$. For all $a > 0$, if $\alpha < 1$ and b is chosen large enough

$$|(1+|p|^{\alpha}) p_i|^2 \leq (a|p|^2 + b)^2$$

Therefore $|(I + |p|)^{\alpha} p\hat{\phi}(p)|_2 \leq a|\Delta\phi|_2 + b|\phi|_2$ and a can be chosen arbitrary small if one chooses b sufficiently large. ♡

2 Rollnik Class Potentials

An important class of potentials which are small with respect to the Laplacian are the potentials of *Rollnik class*.

Definition 3 (*Rollnik class*) A function $V(x), \ x \in R^3$ is said to be *of Rollnik class* (abbreviated $V \in \mathcal{R}$) if

$$|V|_R^2 \equiv \int \frac{|V(x)||V(y)|}{|x-y|^2} d^3x \ d^3y < \infty \tag{28}$$

◇

Rollnik class potentials form a Banach space $\mathcal{R}$ (but not a Hilbert space). Indeed

$$\int \frac{|V(x) + aW(x)||V(y) + aW(y)|}{|x-y|^2} d^3x \ d^3y < \infty$$

implies

$$\int \frac{V(x)W(y) + V(x)W(y)}{|x-y|^2} d^3x \ d^3y < |V|_R |W|_R$$

and therefore

$$|V + W|_R^2 \leq (|V|_R + |W|_R)^2 \tag{29}$$

Remark The bilinear form

$$\langle V, W \rangle \equiv \int \frac{V(x)W(y)}{|x-y|^2} d^3x d^3y \tag{30}$$

is not positive for a generic choice of W, V and therefore it does not define a Hilbert space. It is positive if $V(x)$ and $W(x)$ have constant sign and the sign is the same for both. *In that case* $\langle W, W\rangle = |W|_R^2$. ♣

The following natural inclusions are frequently used. In R^3

$$L^2 \cap L^1 \subseteq R, \qquad |V|_R \leq \sqrt{3}(2\pi)^{1/3}|V|_2^{2/3}|V|_1^{1/3} \tag{31}$$

To prove Eq. (31) notice that

$$\int_{|x-y|>r} \frac{|V(x)||V(y)}{|x-y|^2} d^3x\, d^3y \leq \frac{1}{r^2}\left(\int_{|x-y|>r} |V(x)|d^3x\right)^2 \leq \frac{|V|_1^2}{r^2})$$

$$\int_{|x-y|\leq r} \frac{|V(x)||V(y)}{|x-y|^2} d^3x\, d^3y \leq \int_{|x-y|\leq r} \frac{|V(x)|^2}{|x-y|^2} d^3x d^3y = 4\pi r|V|_2^2 \tag{32}$$

Therefore for any $r > 0$ one has

$$|V|_R \leq \frac{|V|_1^2}{r^2} + 4\pi r|V|_2^2 \tag{33}$$

From this one obtains (31) by taking the minimum with respect to the parameter r.

More generally if X is a measure space and $f \in L^p(X)$ set

$$f_>(x) = f(x)(1 - \xi_f(x)), \qquad f_<(x) = f(x)\xi_f(x) \tag{34}$$

where $\xi_f(x)$ is the indicator function of the set $\{x \ : f(x) < 1\}$.

It is easy to verify (we omit reference to the measure and to the measure space).

$$f \in L^p \Rightarrow f_> \in L^q \quad \forall q \geq p, \qquad f \in L^p \Rightarrow f_< \in L^q \quad \forall q \leq p \tag{35}$$

Potentials of Rollnik class play an important role in the proof of regularity of the solution of the Schrödinger equation and also to establish their asymptotic behavior for large times.

It is therefore important to find conditions under which a potential is of Rollnik class. One such condition is

$$\|V\|_{\mathcal{R}} \leq C_1|V|_{3/2}, \qquad \|V\|_{\mathcal{R}}^2 \leq C_2\|V\|_p\|V\|_q, \quad \frac{1}{p} + \frac{1}{q} + \frac{2}{3} = 2 \tag{36}$$

where we have denoted by $\|.\|_{\mathcal{R}}$ the Rollnik norm and C_1, C_2 are suitable constants. In general for a suitable constant C_3 (depending on p, q, n)

$$\int \frac{|f(x)||h(y)|}{|x-y|^\lambda} d^nx d^ny \leq C_3|f|_p|h|_q, \qquad \frac{1}{p} + \frac{1}{q} + \frac{\lambda}{n} = 2, \quad \lambda < n \tag{37}$$

To prove (37) one must use a *generalized* Hölder inequality. In fact the ordinary Hölder inequality is not applicable because $\frac{1}{|x|^s} \notin L^{1/s}(R)$. Still it is true that $\frac{1}{|x|^s}$ belongs to the space $L_w^{1/s}$ (*weak* $L^{1/s}$) defined as follows

$$f \in L_w^p \leftrightarrow \mu(|f(x)| \geq s) < \frac{C}{s^p} \tag{38}$$

where μ is Lebesgue measure and C is a suitable constant.

Correspondingly one defines the *weak Hölder inequalities*. The following statements are true for weak Hölder inequalities

$$f \in L^p, \quad g \in L^q \Rightarrow f * g \in L^\infty, \qquad |f * g|_\infty \leq |f|_p |g|_q \qquad \frac{1}{p} + \frac{1}{q} = 1 \tag{39}$$

(we have denoted by $*$ the convolution)

$$L^p * L^q \subset L^s, \qquad |f * g|_s \leq |f|_p |g|_q \qquad \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{s} \tag{40}$$

Remark that if $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$ and if $r > 1$ one has $\frac{1}{p} + \frac{1}{q} > 1$. It follows that $f * g \in L^{r^*}$ if $1 < q < q^* < \infty$ (recall that q^*, the exponent *conjugated to* q, is defined by $\frac{1}{q} + \frac{1}{q^*} = 2$).

Since $|x - y|^\lambda \in L_w^{\frac{1}{\lambda}}$ one derives

$$L^p \cap L^q \subset \mathcal{R} \qquad p \leq 3/2 \leq q$$

Notice that if $V \in L^1 \cap L^2$ then $\hat{V} \in L^\infty$, $V(x)V(y) \in L^2(R^6)$ and one can make use of Fourier transform to obtain

$$\int \frac{V(x)V(y)}{a^2 + |x - y|^2} d^3x d^3y = \frac{1}{4\pi} \int \frac{|\hat{V}(p)|^2 e^{-ap}}{|p|} d^3p \tag{41}$$

Therefore if $V \in L^1 \cap \mathcal{R}$ the function V satisfies

$$\frac{1}{4\pi} \int \frac{(\hat{V})^*(p)\hat{W}(p)}{|p|} d^3p < \infty \tag{42}$$

If $\alpha < 2$ then $|x|^{-\alpha} \in R + L^\infty$.

The interest in the Rollnik class of potentials is related, among other things, to the fact that if V belongs to the Rollnik class, then $(-\Delta + V + \lambda)^{-1} - (-\Delta + \lambda)^{-1}$ is a compact operator for λ positive and large enough. In scattering theory this property makes it possible to compare in the limit $t \to \pm\infty$ the free evolution with the time evolution of given by the Schroedinger operator $-\Delta + V$.

Theorem 4 [4, 6] *If* $V \in \mathcal{R} + L^\infty(R^3)$ *then* $V \prec\prec \Delta$. *Moreover*

$$(-\Delta + V + \lambda)^{-1} - (-\Delta + \lambda)^{-1}$$

is a compact operator when λ *is sufficiently large.* ◇

Proof Notice that

$$lim_{\lambda\to\infty}||V|^{1/2}(-\Delta + \lambda\, I)^{-1}|V|^{1/2}| = 0 \tag{43}$$

Indeed the kernel of this operator is

$$\frac{|V(x)|^{1/2}e^{-\sqrt{\lambda}|x-y|}|V(y)|^{1/2}}{|x-y|} \tag{44}$$

and therefore if $V \in \mathcal{R}$ the operator is of Hilbert-Schmidt type. Its nucleus tends point-wise to zero when $\lambda \to \infty$ if $x \neq y$ and therefore its Hilbert-Schmidt norm tends to zero in the limit. From this one derives

$$||V|^{1/2}(-\Delta + \lambda\, I)^{-1/2}| \to_{\lambda\to\infty} 0$$

$$|(-\Delta + \lambda\, I)^{-1/2}|V|(-\Delta + \lambda\, I)^{-1/2}| \to_{\lambda\to\infty} = 0 \tag{45}$$

and therefore uniformly in ϕ

$$(\phi, |V|\phi) \leq a(\lambda)(\phi, (-\Delta + \lambda\, I)\phi), \qquad lim_{\lambda\to\infty}a(\lambda) = 0 \tag{46}$$

The first part of Theorem 4 follows from these estimates.

For the second part one writes

$$(H+\lambda)^{-1} = (-\Delta+\lambda)^{-1/2}[1+(-\Delta+\lambda)^{-1/2}\ V\ (-\Delta+\lambda)^{-1/2}]^{-1}(-\Delta+\lambda)^{-1/2} \tag{47}$$

For λ sufficiently large $(-\Delta+\lambda)^{-1/2}\ V\ (-\Delta+\lambda)^{-1/2}| < 1$ and therefore one can iterate (47) to obtain the Born series

$$(H+\lambda)^{-1} = (-\Delta+\lambda)^{-1/2}(1+\sum_n(-1)^n[(-\Delta+\lambda)^{-1/2}\ V\ (-\Delta+\lambda)^{-1/2}]^n)(-\Delta+\lambda)^{-1/2} \tag{48}$$

which converges in norm for λ large enough.

Therefore for λ sufficiently large $(H+\lambda)^{-1} - (-\Delta+\lambda)^{-1}$ is a compact operator; in fact it is the norm limit of the partial sums which are Hilbert-Schmidt operators. ♡

Notice that the negative eigenvalues $E_n < 0$ of H coincide with the zero eigenvalues of the operator

$$I - (-\Delta + E_n)^{-\frac{1}{2}} V (-\Delta + E_n)^{-\frac{1}{2}} \equiv I - K_{E_n} \tag{49}$$

where K_{E_n} is for every E_n a compact operator that converges to zero when $E_n \to \infty$.

If V is of Rollnik class the negative part spectrum in pure point and bounded below.

Following the same steps of the proof of Theorem 4 one proves

Corollary *If* $V \in \mathcal{R} + L^{\infty}_{\epsilon}$ *then* $V \prec\prec \Delta$ *and the conclusion of Theorem 4 hold.* ◇

In the same way one proves

Theorem 5 *If* $V \in R$ *then* $|V|^{1/2} \prec\prec (-\Delta + 1)^{1/2}$. ◇

A further condition on the potential V that implies $V \prec\prec -\Delta$ is given by the following Proposition

Proposition 6 *If* $d \geq 4, \quad V \in L^p, \quad p > d/2$ *then* $V \prec\prec \Delta$. ◇

Proof The function $(1 + k^2)^{-1}$ is in $L^p(R^d)$ if $p > d/2$. Therefore if $u \in D(\Delta)$ one has

$$|\hat{u}|_q \leq [(1 + k^2)|\hat{u}(k)|]_2 |(1 + k^2)^{-1}|_p \qquad \frac{1}{q} + \frac{1}{p} = \frac{1}{2} \tag{50}$$

It follows that if $V \in L^p$ and $u \in D(\Delta)$ then $V\ u \in L^2$. Moreover for every $t > 0$

$$|V\ u|_2 \leq |V|_p |(1 + tk^2)^{-1}|_p |(1 + tk^2)\hat{u}|_2 \tag{51}$$

Taking t large enough one sees that $V \prec\prec -\Delta$ and one can obtain an explicit estimate of the norm of $|V\ u|_2$ relative to $|\Delta u|_2$. ♡

3 Stummel Class Potentials

The Kato-Rellich theorem is sufficient for many applications but it is not optimal because it does not fully exploit the *local* properties of the potential.

For these it is convenient to introduce a new class of potentials, the *Stummel class*.

Definition 4 (*Stummel class*) A Lebesgue-measurable function $V(x)$ on R^d is said to be *of Stummel class* (denoted by S_d) if the following *local* conditions are satisfied

(1) if $d = 1, 2, 3$

$$sup_{x \in R^d} \int_{|x-y| \leq 1} |V(y)|^2 dy < \infty \qquad \forall x$$

(2) if $d = 4$

$$lim_{\alpha \to 0} sup_{x \in R^4} \int_{|x-y| \leq \alpha} log(|x-y|)^{-1} |V(y)|^2 dy < \infty$$

(3) if $d \geq 5$

$$lim_{\alpha \to 0} sup_{x \in R^d} \int_{|x-y| \leq \alpha} \frac{|V(y)|^2}{|x-y|^{d-4}} dy < \infty$$

$\diamondsuit$

It is easy to prove that the potential V belongs to S_d if the following condition is satisfied

$V(x) \in L^2_{unif}$ if $d \leq 3$

$V(x) \in L^p_{unif}$ $p > d/2$ if $d \geq 4$.

We have introduced the notation $V \in L^p_{unif}$ to indicate that $V \in L^p_{loc}$ and $sup \int_{|x-y| \leq 1} |V(x)|^p dy < \infty$.

To study Rollnik-type potentials we shall use the following Lemma, which has an independent interest.

Lemma 7 *Let Ω, μ be a measure space and let K be a function on $\Omega \times \Omega$ such that*

$$K(x, y) = K_1(x, y)\ K_2(x, y), \quad x,\ y\ \in \Omega \tag{52}$$

and let

$$sup_{x \in \Omega} \int_\Omega |K_1(x, y)|^2 d\mu(y) \equiv C_1^2 < \infty \quad sup_{y \in \Omega} \int_\Omega |K_2(x, y)|^2 d\mu(x) \equiv C_2^2 < \infty \tag{53}$$

Then the operator K_{op} with kernel $K(x, y)$ is bounded in $L^2(\Omega, \mu)$ and the norm satisfies $\|K_{op}\| \leq C_1\ C_2$. $\diamondsuit$

Proof For $\phi,\ \psi\ \in L^2(\Omega, \mu)$ repeated use of Schwarz's inequality gives

$$|(\phi, K\psi)| \leq \int |K(x, y)||\phi(x)||\psi(y)| d\mu(x) d\mu(y)$$

$$\leq (\int |K_1(x, y)|^2 |\phi(x)|^2 d\mu(x) d\mu(y))^{1/2} (\int |K_2(x, y)|^2 |\psi(y)|^2 d\mu(x) d\mu(y))^{1/2}$$

$$\leq [(\int |\phi(x)|^2 d\mu(x) d\mu(y))\ sup_x \int (K_1(x, y)^2 d\mu(y))^{\frac{1}{2}}]\ [(\int |\psi(y)|^2 d\mu(x) d\mu(y))$$

$$sup_y \int |K_2(x, y)|^2 d\mu(x)]^{1/2} = C_1\ C_2\ |\phi|_2 |\psi|_2 \tag{54}$$

$\heartsuit$

Theorem 8 [4, 5] *If $V \in S_d$ then $V \prec\prec \Delta$ and therefore $H \equiv -\Delta + V$ is essentially self-adjoint on C_0^∞. Moreover $D(H) = D(-\Delta + 1) = H^2(R^d)$.* ◇

Proof To prove Theorem 8 we shall use the representation

$$(e^{t\Delta}\phi)(x) = (4\pi t)^{-d/2} \int_{R^d} e^{-\frac{|x-y|^2}{4t}} \phi(y) dy, \quad t > 0 \tag{55}$$

Denote with $R_0(z, x)$ the integral kernel of the resolvent

$$(H_0 - z)^{-1} = \int_0^\infty e^{-(H_0 - z)t} dt, \qquad Rez < 0$$

For $d = 1, 2, 3$ one has the explicit formulas for $z \in C^+$

$$d = 1 \qquad R_0(z, x) = i \frac{e^{i|x|z}}{2\sqrt{z}}, \quad x \in R$$

$$d = 2 \qquad R_0(z, x) = \frac{e^{i|x|\sqrt{z}}}{2\pi|x|} log|x| \quad x \in R^2$$

$$d = 3 \qquad R_0(z, x) = \frac{e^{i|x|\sqrt{z}}}{4\pi|x|}, \quad x \in R^3$$

For $d \geq 3$ one has $|R_0(z, x)| < C|x|^{-d-2}$. Moreover for any d

$$|x| \geq \delta, \quad Rez < -\delta \quad \Rightarrow |R_0(z, x)| < Ce^{(-1+\epsilon)\sqrt{|Rez|}|x|} \tag{56}$$

This inequalities follow from the quadratic inequality $\frac{|x|^2}{4t} + Et \geq |x| \sqrt{E}$ which implies for every $\epsilon > 0$

$$|R_0(z, x)| \leq e^{-(1-\epsilon)|x|\sqrt{E}} \int_0^\infty (4\pi t)^{-\frac{d}{2}} e^{-\epsilon(\frac{|x|^2}{4t} + Et} dt = Ce^{-(1-\epsilon)E|x|} \tag{57}$$

Notice that $R_0(z, .)$ defines by convolution a bounded operator and therefore it is sufficient to prove the inequalities for the dense set $C_0^\infty(R^d)$.

The integral kernel of $(-\Delta - z)^2, \quad Rez < 0$ is

$$Q_0(z, x) \equiv (2\pi)^{-\frac{d}{2}} \int_0^\infty t^{-\frac{d-2}{2}} e^{-\frac{|x|^2}{4t}} e^{zt} dt$$

and one has

$$for \quad d \leq 3 \qquad |Q_0(z, x)| < C$$

$$for \quad d = 4 \qquad |Q_0(z, x)| < C\, log(|x|^{-1} + 1)$$

$$for \quad d \geq 5 \qquad |Q_0(z,x)| \leq C\ |x|^{-d+4}$$

and moreover for any d

$$|x| \geq \delta, \quad Re\ z < -\delta \Rightarrow |Q_0(z,x)| \leq C\ e^{-(1-\epsilon)|x|\sqrt{Re\ z}}$$

The integral kernel of $V(-\Delta + E)^{-2}V = [V(-\Delta + E)]^{-1}[(-\Delta + E)V]^{-1}$ is

$$V(x)Q_0(-E, x-y)V(y) \equiv K_1^E(x,y)K_2^E(x,y) \tag{58}$$

where K_1^E and K_2^E are the integral kernels of $V\ (-\Delta + E)^{-1}$ and $(-\Delta + E)^{-1}\ V$.

To estimate $sup_{x\in\Omega} \int_\Omega |K_1(x,y)|^2 dy$ divide the integration domain in $|x-y| > \alpha$ and $|x-y| \leq \alpha$. Since $V \in S_d \subset L^2_{loc}$ the contribution of the first region is bounded by

$$C'_\alpha \int_{R^d} e^{-\sqrt{E}\ \frac{|x-y|^2}{2}} \leq C_\alpha e^{-\frac{d}{2}} \tag{59}$$

To estimate the contribution of the second region we use the estimates given for the resolvents. In particular for α sufficiently small this contribution is bounded by ϵ uniformly in x and $E > \delta$. Therefore

$$lim_{E\to\infty} sup_x \int |K_1^E(x,y)|^2 dy = 0 \tag{60}$$

With a similar argument one obtains

$$lim_{E\to\infty} sup_y \int |K_2^E(x,y)|^2 dx = 0 \tag{61}$$

On the other hand for all $\phi \in L^2$

$$|V\phi| \leq |V(-\Delta + E)^{-1}||\Delta\phi| + E|V(-\Delta + E)^{-1}||\phi| \tag{62}$$

This achieves the Proof of Theorem 8. ♡

4 Operators with Positivity Preserving Kernels

All examples given up to now, with the exception of the anharmonic oscillator, are such that $D(H) = D(\Delta)$. There are interesting cases in which $D(H) \neq D(\Delta)$ but still H is *essentially self-adjoint* on $C_0^\infty(R^d)$. In most of these cases one makes use of the property of integral kernel of $\frac{1}{-\Delta+c},\ c > 0$ *to be positivity preserving*. We shall use the following inequality due to T. Kato.

Lemma 9 [4] *Let* $u \in L^1_{loc}(R^d)$, $\Delta u \in L^1_{loc}$ *and define the function* $sign\ u$ *as follows : if* $u(x_0) \neq 0$ *then* $(sign\ u)(x_0) = \frac{u(x_0)}{|u(x_0)|}$. *If* $u(x_0) = 0$ *then* $(sign\ u)(x_0) = 0$.

Define $\epsilon(u) \equiv u\ sign\ u$ *in the distributional sense. Then in the distributional sense*

$$\Delta|u| \geq sign(u)\ \Delta\ u \tag{63}$$

◇

Proof Let $u \in C^\infty$ and define $u_\epsilon(x) \equiv \sqrt{|u(x)| + \epsilon^2}$. Then $u_\epsilon \in C^\infty$ and

$$2\bar{u}_\epsilon.\nabla u_\epsilon = \nabla|u_\epsilon|^2 = \nabla|u|^2 = 2Re\ \bar{u}\ \nabla u \tag{64}$$

Therefore $|\nabla u_\epsilon| \leq |\nabla u|$ and, differentiating (64) with respect to x one derives $\Delta u_\epsilon \geq Re(\frac{\bar{u}}{u_\epsilon}\ \Delta u)$.

Since $lim_{\epsilon\to 0} u_\epsilon|u|$ in L^1_{loc} one has $\Delta u_\epsilon \to \Delta|u|$ xin the distributional sense. Moreover $lim_{\epsilon\to 0}\frac{\bar{u}}{u_\epsilon}\Delta u = sign(u)\ \Delta u$ in L^1_{loc} and (63) is proved in this case.

In the general case consider a regularization of u defined by

$$U^\delta = (u * j_\delta)(x), \quad j_\delta(x) = \delta^{-d} j(\frac{x}{\delta}), \quad \delta > 0, \quad j_\delta \in C_0^\infty, \quad \int j_\delta(x)dx = 1 \tag{65}$$

Since $u * j_\delta$ is regular (63) holds. Taking the limit $\delta \to 0$ proves (64) if $u \in L^1_{loc}$, $\Delta u \in L^1_{loc}$. ♡

Theorem 10 [5] *Let*

$$V(x) = V_+(x) - V_-(x), \quad V_\pm > 0, \quad V_+ \in L^2_{loc}(R^d) \quad V_- \prec -\Delta \tag{66}$$

Then $-\Delta + V$ *is essentially self-adjoint on* C_0^∞. ◇

Proof We must prove that under the hypothesis of the theorem if for at least one positive value of the parameter μ one has $u \in L^2, \quad (-\Delta + V + \mu)u = 0$ then $u = 0$. From Lemma 9 we know

$$\Delta|u| \geq Re\ sign(\bar{u})(V + \mu)u = (V + \mu)|u| \geq (-V_- + \mu)|u| \geq 0$$

It follows $(-\Delta - V_- + \mu)|u| \leq 0$ and therefore, since $(-\Delta + \mu)$ preserves positivity, $(-\Delta + \mu)^{-1}(-\Delta - V_- + \mu)|u| \leq 0$. From this one derives $|u| \leq (-\Delta + \mu)^{-1}V_-|u|$; choosing μ large enough

$$\|(-\Delta + \mu)^{-1}V_-\| \leq a + \frac{b}{\mu}|u|_2 < |u|_2$$

and therefore

$$|u|_2 \leq |(-\Delta+\mu)^{-1}V_-|u||_2 \leq (a+\frac{b}{\mu})|u|_2 < |u|_2 \tag{67}$$

This implies $|u|_2 = 0$, $u = 0$. ♡

Notice that one can also derive the condition on V_- as follows.

If V is positive we know from theorem 10.10 that $-\Delta_V$ is essentially self-adjoint on C_0^∞. One the other hand, since V_+ is positive from $V_- \prec\prec -\Delta$ it follows $V_- \prec\prec -\Delta + V_+$ and therefore from the Kato-Rellich theorem one derives that $-\Delta + V_+ + V_-$ is essentially self-adjoint on C_0^∞.

In Lemma 20 an essential role is played by the fact that the operator $(-\Delta+\mu)^{-1}$ is positivity preserving together with the assumption that the part of the potential which is not Kato-small with respect to the Laplacian is a positive function. From an abstract point of view Theorem 10 in "Lecture 10: Standard Form" holds if there exist in $\mathcal{H}$ a *convex cone* which is left invariant by e^{-tH_0} and $V_1 \prec H_0$.

The positivity criteria that we will discuss in part two of these lectures can be put in this abstract form. We shall not do it here because we will also use other properties of positive functions. It is important to keep in mind the central role of the convex cones which are invariant under H_0 and V_1. Extensions to the non-commutative case (C^*-algebras $\mathcal{A}$) of the Dirichlet forms can be obtained because also in that context there exists a convex cone (the positive elements of the algebra $\mathcal{A}$). Also in this new context it will be important to study the maps that leave invariant the cone of positive elements.

In "Lecture 10: Derivations and Generators. K.M.S. Condition. Elements of Modular Structure. Standard Form" we have seen some properties of the positive and completely positive semigroups, but we have not discussed their applications to Schrödinger's theory. We shall come back to the problem in the second part of these Lectures.

In Theorem 10 we have used the fact that the kernel the resolvent $\frac{1}{-\Delta-cI}$, $c > 0$ preserves positivity. The same is true for the kernel of $e^{t\Delta}$. This leads to the next Theorem which has application in Quantum Field Theory.

Theorem 11 (Davies, Faris) [5] *Let μ a regular Borel measure on the Banach space X. Consider on $L^2(X, d\mu)$ a* positive *operator H_0 and a Potential $V(x)$, $x \in X$. Assume that e^{-tH_0} is positivity preserving, that $H = H_0 + V_+$ is essentially self-adjoint on $D(H_0) \cap D(V)$ and that V_- is H_0-bounded. Then V_- is H-bounded.* ◇

Proof Let $\|V_-\ (H_0+b)^{-1}\| \leq a$. We must prove that $\|V_-\ (H+b)^{-1}\| \leq a$. Remark that if A is bounded and positivity preserving then $|A\phi|(x) \leq A|\phi(x)|$.

If e^{-tH_0}, $t \geq 0$ preserves positivity the same in true for $(H_0+b)^{-1}$ (recall that $(H_0+b)^{-1} = \int e^{-tH_0-tb}dt$). It is therefore sufficient to prove $|V_-(H_0+b)^{-1}\phi|_2 \leq a|\phi|_2$ and this follows if locally for $\phi(x) \geq 0$

$$0 \leq ((H+b)^{-1}\phi)(x) \leq ((H_0+b)^{-1}\phi)(x) \tag{68}$$

To prove (68) remark that $\phi(x) \geq 0 \Rightarrow (e^{-sH_0}(1-e^{-sV_+})\phi)(x) \geq 0$ and therefore

$$0 \leq (e^{-sH_0}e^{-sV_+}\phi)(x) \leq (e^{sH_0}\phi)(x)$$

Hence $0 \leq ((e^{-\frac{s}{n}H_0}e^{-\frac{s}{n}V_+})^n\phi)(x) \leq (e^{-\frac{s}{n}H_0}\phi)(x)$ and from the Feynmann-Kac formula

$$0 \leq (e^{-tH}\phi)(x) \leq (e^{-tH_0}\phi)(x) \tag{69}$$

One derives (68) from (69) by means of the Laplace transform. ♡

Corollary (Davies, Faris, Kato-Rellich) *If* $V_+(x) \in L^2_{loc}(R^d)$ *and* $V_-(x) \in L^p(R^d) + L^\infty(R^d)$ *for* $p = 2$ *if* $d \leq 3$, $p > 2$ *if* $d = 4$ *and* $p = d/2$ *if* $d \geq 5$, *then* $-\Delta + V(x)$ *is essentially self-adjoint on* C_0^∞. ◇

5 Essential Spectrum and Weyl's Comparison Theorems

Definition 5 (*essential spectrum*) The *essential spectrum* of a self-adjoint operator A (denoted by $\sigma_{ess}(A)$) is *the closure* of the set $\sigma(A) - I(A)$ where $I(A)$ denotes the collection of those eigenvalues of A which have finite multiplicity. ◇

The role of the essential spectrum is particularly relevant in scattering theory as we will see in the second part of these Lectures: in particular asymptotic completeness requires that the essential spectrum be absolutely continuous.

In what follows we shall denote by H_0 a reference operator (typically $-\Delta$) and by V a perturbation (typically multiplication with a measurable function). We have seen that if V is H_0-small, then $H \equiv H_0 + V$ is self-adjoint on $D(H_0)$. We want now to compare the spectra of H and H_0.

In general we will be interested in comparing the spectra of two self-adjoint operators A and B and in particular to have criteria under which the essential spectra coincide.

Proposition 12 [4] *If two bounded operators* A *and* B do not have point spectrum *and the operator* $A - B$ *is compact then* $\sigma_{ess}(A) = \sigma_{ess}(B)$. ◇

Proof Let $K \equiv A - B$ be compact. Define

$$F(z) \equiv K(A-z)^{-1}, \qquad z \in C - \sigma(A) \tag{70}$$

The function $F(z)$ is analytic with values in compact operators and $lim_{|z|\to\infty}\|F(z)\| = 0$. In particular $(I - F(z))^{-1}$ exists when $|z|$ is sufficiently large. By the analytic Fredholm theorem $(I - F(z))^{-1}$ is bounded for $z \in C - (\sigma(A) \cup D)$ where D is discrete.

Moreover $(1-F(z))^{-1}$ is meromorphic in $C-\sigma(A)$ and it has finite-order residues in D. If $1-F(z)$ is invertible one has $(B-z)^{-1} = (A-z)^{-1}(1-F(z))^{-1}$ and then

$$\sigma(B) \subset \sigma(A) \cup D \tag{71}$$

Observe now that $(B-z)^{-1}$ has residues of finite order in D and therefore $D \in \sigma_{disc}(B)$. Then by assumption $D = \emptyset$ and $\sigma_{ess}(A) \subset \sigma_{ess}(B)$. Exchanging the roles of A and B one concludes the proof of Proposition 12. ♡

In the case of Schrödinger operators one would like to apply this proposition to $A = (H-z)^{-1}$ and $B = (H_0-z)^{-1}$; since H may have bound states we must relax the condition that the discrete spectrum be empty.

Theorem 13 (first form of Weyl's comparison theorem) *Let A, B self-adjoint operators bounded below. If there exists $z \in \rho(A) \cap \rho(B)$ such that the operator $\frac{1}{A-z} - \frac{1}{B-z}$ is compact, then $\sigma_{ess}(A) = \sigma_{ess}(B)$.* ♢

Proof Notice that, by the resolvent identity, if the assumption holds true $z_0 \in \rho(A) \cap \rho(B)$ then it holds true for every $z \in \rho(A) \cap \rho(B)$.

Let z_0 be real, $z_0 \in \rho(A) \cap \rho(B)$; recall that we have assumed that A and B be bounded below. Define $E \equiv (A-z_0)^{-1}$, $\quad F \equiv (B-z_0)^{-1}$. By assumption $E = F + K$ with K compact. Therefore the function $K(E-z)^{-1}$ is analytic with value in compact operators; it is meromorphic in $C - (\sigma_{ess}(E))$ with residues which are finite rank operators.

For $z \notin \sigma(E)$ the function $(F-z)^{-1}$ is defined iff $[1-K(E-z)^{-1}]^{-1}$ exists. On the other hand $(F-z)^{-1}$ is analytic for $Im\, z \neq 0$ and also at $z = z_0$. From meromorphic Fredholm theory it follows that $[1-K(E-z)^{-1}]^{-1}$ is defined on $C - \sigma_{ess}(E)$ with the possible exception of a finite number of points in which it has a finite rank residue. Therefore the same is true for $(F-z)^{-1}$. This implies $\sigma_{ess}(F) \subset \sigma_{ess}(E)$. Interchanging the role of E and F achieves the Proof of Theorem 13. ♡

Following the same steps one can prove

Theorem 14 *Let A be self-adjoint and B symmetric and closed. Assume that for $z_0 \in \rho(B) \cap \rho(A)$ (and therefore for any $z \in \rho(B) \cap \rho(A)$) the operator $(A-z_0)^{-1} - (B-z_0)^{-1}$ be compact. Assume moreover*

(a) $\sigma(A) \neq R$, $\rho(B) \neq \emptyset$
(b) $\rho(B)$ *intersects both the half plane* $Im\, z < 0$ *and the half-plane* $Im\, z > 0$.

Then $\sigma_{ess}(B) = \sigma_{ess}(A)$. ♢

For the proof see [RS,IV p. 112].

Notice that in Theorems 13 and 14 the condition that A be self-adjoint is necessary, as seen by the following example.

Example Let $\mathcal{H} \equiv l^2(-\infty, \infty)$, $(A\phi)_n = \phi_{n+1}$. Define C as $(C\phi)_n = \delta_{n,0}\phi_1$. Therefore C is a rank one perturbation.

Set $B \equiv A - C$. We shall prove

(i) $\sigma(A) = \sigma_{ess} = \{z \mid |z| = 1\}$
(ii) $\sigma(B) = \sigma_{ess}(B) = \{z \mid |z| \leq 1\}$.

To prove (i) consider the map U (inverse of the discrete Fourier transform)

$$U \ : \ \{\phi_n\} \to (\sqrt{2\pi})^{-1} \sum_n \phi_n e^{inx}, \qquad l^2(-\infty, \ \infty) \to L^2(0, 2\pi) \tag{72}$$

Then $UAU^* = e^{-ix}$ therefore $\sigma(A) = \{e^{-ix} \mid x \in R\} = \{z, \ |z| = 1\}$ and moreover $\sigma(A)$ is absolutely continuous.

(ii) To find $\sigma(B)$ notice that $(B - z)^{-1}$ exists iff $(1 - C(A - z)^{-1})^{-1}$ exists. This requires that $C(A - z)^{-1}$ does not have one as eigenvalue and therefore z must not be an eigenvalue of B. Let $(B - z)\phi = 0$ or equivalently

$$\phi_{n+1} - \delta_{n,0}\phi - z\phi_n = 0 \tag{73}$$

Let $z \neq 0$. For $n = 0$ one has $\phi_0 = 0$ and from Eq. (66) follows that $\phi_n = 0, \quad n < 0$. Moreover $\phi_n = z^{n-1}, \ n \geq 1$. Therefore if $|z| \geq 1$ the equation $(B - z)\phi = 0$ has no solution in l^2 (but it has solutions if $|z| < 1$). It follows that

$$\rho(B) \equiv \{z, \ \mid |z| \geq 1\}, \quad \sigma(B) = \sigma_{ess}(B) = \{z, |z| < 1\} \tag{74}$$

♡

The first form of Weyl's comparison theorem has the following important corollary.

Corollary of Theorem 14 *Let A be a symmetric operator with defect indices (m, m) where $m < \infty$. Let A_1 and A_2 be self-adjoint extensions of A. Then $\sigma_{ess}(A_1) = \sigma_{ess}(A_2)$.* ◇

Proof If $\psi \in Range \ (A^* - i)$ one has $\psi = (A^* + i)\phi, \quad \phi \in D(A)$ and therefore $((A_1 + i)^{-1} - (A_2 + i)^{-1})\phi = 0$. It follows that $(A_1 + i)^{-1} - (A + i)^{-1}$ has finite rank. ♡

Definition 5 (*relative compactness*) We say that the operator K is A-compact if $(A - z_0)^{-1}K$ is compact for all $z_0 \in \rho(A)$. ◇

Let K be A-compact, and set $B \equiv A + K$. It is not difficult to prove that B is closed on $D(A)$ and self-adjoint on $D(A)$ if K is symmetric. It follows from Theorem 13 that $\sigma_{ess}(B) = \sigma_{ess}(A)$.

Notice that the theorem *does not* state that the point spectra of A and of B coincide. Indeed it is possible to prove that for any choice of ϵ there exist operators $A, \ B$ with $C \equiv A - B$ of trace class, $tr|C| < \epsilon$, such that A has continuous spectrum and B has purely point spectrum.

Theorem 15 (second form of Weyl comparison theorem) *Let H_0 be self-adjoint, V symmetric $V \prec H_0$. Suppose that $V(H_0 - z)^{-1}$ is compact for at least one value $z_0 \in \rho(H_0)$. Then the essential spectra of $H = H_0 + V$ and of H_0 coincide.* ◇

For the proof we shall make use of the following lemma of Weierstrass. Its proof is based on the fact that a compact operator is norm limit of a sequence of finite rank operator, and on Weierstrass lemma for finite matrices. In turn Weierstrass lemma for matrices is proved remarking that $(I - F(z))$ is invertible iff $det(I - F(z)) \neq 0$ and using Weierstrass lemma on the zeroes of an analytic function.

Lemma 16 (Weierstrass) [1, 6] *Let Ω be an open subset of the complex plane C and for every $z \in \Omega$ let $F(z)$ be a compact operator. Suppose $F(z)$ is analytic in Ω in the uniform topology. Then one has the following alternative: either $I - F(z)$ is not invertible for any $z \in \Omega$ or it is invertible in $\Omega - J$ where J is a denumerable set with no limit points in Ω.* ◇

Using this lemma we shall prove Theorem 15.

Proof of Theorem 15 We prove first

$$\sigma_{ess}(H_0 + V) \subset \sigma_{ess}(H_0) \tag{75}$$

If $\sigma(H_0) = R$ the inclusion is obvious. Suppose that $C - \sigma_{ess}(H_0)$ is open. Let I be a finite interval with $I \cap \sigma_{ess}(H_0) = \emptyset$. We must prove $I \cap \sigma_{ess}(H_0 + V) = \emptyset$. We apply Weierstrass lemma to the function $V\ (H_0 - z)^{-1}$ on $\Omega \equiv C - \sigma(H_0)$. We must verify

(a) $V(H_0 - z)^{-1}$ is compact for every $z \in \Omega$.
(b) $I + V(H_0 - z)^{-1}$ is invertible for every $z \in \Omega$.

To prove (a) we use the identity

$$V(H_0 - z)^{-1} = V.(H_0 - \zeta)^{-1}\ (H_0 - \zeta)\ (H_0 - z)^{-1} \qquad \forall z,\ \zeta \in \Omega \tag{76}$$

By assumption there is a z_0 for which $V(H_0 - z_0)^{-1}$ is compact. For every choice of $z \in \Omega$ the operator $(H_0 - \zeta)(H_0 - z)^{-1}$ is bounded. This proves point (a).

For the proof of point (b) it suffices to show that for $\zeta = i\xi,\ \ \xi \in R^+$ one has

$$|V(H_0 - i\xi)^{-1}| < 1 \tag{77}$$

if ξ large enough. From $|V\phi| \leq a\ |H_0\phi| + b|\phi|, a < 1$ setting $\phi = (H_0 - i\xi)^{-1}\psi$ one has

$$\begin{aligned} |V(H_0 - i\xi)^{-1}\psi| &\leq a\ |H_0(H_0 - i\xi)^{-1}\psi| + b|(H_0 - i\xi)^{-1}\psi| \\ &\leq a|H_0(H_0 - i\xi)^{-1}\psi| + b|(H_0 - i\xi)^{-1}\psi| \end{aligned} \tag{78}$$

(76) follows for ξ large enough since $H_0(H_0 - i\xi)^{-1}$ and $\xi(H_0 - i\xi)^{-1}$ are both contractions.

From $D(V) \subset D(H_0)$ we conclude that $(H_0 + V - z)^{-1} = (I + V\ (H_0 - z)^{-1})(H_0 - z)$ on $D(H_0)$. Using Weierstrass lemma for the function $V(H_0 - z)^{-1}$ one derives that $H + V - z$ is invertible for $z \in C - (\sigma(H_0) \cup J')$ where J' is denumerable and without limit points in $C - \sigma(H_0)$.

Let now I be an interval of the real axis such that $I \cap \sigma_{ess}(H_0) = \emptyset$. Then $I_2 \equiv I \cap \sigma(H_0)$ and $J_1 \equiv I \cap J'$ are finite sets. Therefore $H_0 + V - \lambda$ is invertible for $\lambda \notin J_1 \cup J_2$, a finite set of points.

Let $\lambda \in J_1 \cup J_2$. We must prove that if λ is an eigenvalue of $H_0 + V$ its multiplicity is finite. This will imply $\lambda \notin \sigma_{ess}(H_0 + V)$ and therefore $\sigma_{ess}(H_0 + V) \cap I = \emptyset$. Let λ be an isolated eigenvalue of $H_0 + V$. The orthogonal projection on the corresponding eigen-space is

$$\Pi_\lambda \equiv (2\pi)^{-1} \int_\gamma (H_0 + V - z)^{-1} d\gamma \tag{79}$$

where γ is a circle of radius δ around λ. If δ is sufficiently small

$$\Pi_\lambda = (2\pi)^{-1} \int_\gamma (H_0 - z)^{-1}(1 - F(z)^{-1} d\gamma, \quad F(z) \equiv -V\ (H_0 - z)^{-1} \tag{80}$$

By assumption $(F(z) - 1)^{-1}$ has a Mac Laurin expansion in $z - \lambda$ and since also $(H_0 - z)^{-1}$ admits such expansion we conclude that $(H_0 - z)^{-1}(1 - F(z))^{-1}$ has a Mac Laurin expansion with coefficient that are finite-rank operators. Therefore Π_λ projects on a finite-dimensional subspace. This prove part (a).

To prove the converse one may think to invert the roles of $H_0 + V$ and of H_0. This is not possible, since $V \prec H_0$ *does not imply* $V \prec H_0 + V$ (consider for example $V = -\frac{2}{3} H_0$). We will instead prove that if $V \prec H_0$ and $V\ (H_0 - \lambda_0)^{-1}$ is compact for $\lambda_0 \notin \sigma(H_0)$ then

(a) For every $\lambda \notin \sigma(H_0 + V)$ the operator $V\ (H_0 + V - \lambda)^{-1}$ is compact.
(b) For at least one λ in each connected component of $C - \sigma(H_0 + V)$ the operator $1 - V\ (H_0 + V - \lambda)^{-1}$ is invertible.

To prove (a) it suffices to prove the statement for some $\lambda_0 \notin \sigma(H_0 + V)$. One has

$$V(H_0 + V - \lambda)^{-1} = V(H_0 - \lambda)^{-1}[1 + V(H_0 - \lambda)^{-1}] \tag{81}$$

and the thesis follows because $V\ (H_0 - \lambda)^{-1}$ is compact and $I + V\ (H_0 - \lambda)^{-1}$ is bounded.

To prove (b) it suffices to remark that if $\lambda \notin \sigma(H_0) \cup \sigma(H_0 + V)$ one has

$$1 + V(H_0 - \lambda)^{-1} = (H_0 + V - \lambda)\ (H_0 - \lambda)^{-1} \tag{82}$$

From (81) and (82) one derives that $I - V(H_0 + V - \lambda)^{-1}$ is invertible if $H_0 + V - \lambda(H_0 - \lambda)^{-1}$ is bounded. If $V \prec H_0$ this is true provided λ is large enough. This achieves the Proof of Theorem 18. ♡

Definition 6 (*space* L^{∞}_{ϵ}) The function f belongs to L^{∞}_{ϵ} if there exists a positive constant C such that $|f(x)| < \epsilon$ if $|x| > \frac{C}{\epsilon}$. ◇

As an application of the second Weyl comparison theorem we prove the following proposition.

Proposition 17 *If* $V(x) \in L^2(R^3) + L^{\infty}_{\epsilon}(R^3)$ *then*

$$\sigma_{ess}(-\Delta + V) = \sigma_{ess}(-\Delta) \ \ (\equiv R^+) \tag{83}$$

Moreover the set $\sigma(-\Delta + V) - \sigma(-\Delta)$ *is* at most denumerable *and if it is infinite the only limit point is the origin. In particular all strictly negative eigenvalues have finite multiplicity.* ◇

Proof We first of all prove that $V(-\Delta + \lambda)^{-1}$ is compact for $\lambda \notin R^+$.

Let $\{C_n\}$ be a decreasing sequence of positive numbers converging to 0. By assumption there exist two sequences V_1^n, V_2^n such that

$$V_1^n \in L^2 \quad |V^n| \in L^{\infty} \quad |lim_{n\to\infty} V^n|_{\infty} = 0, \quad V_1^n + V_2^n = V \tag{84}$$

Since $(-\Delta + \lambda)^{-1}$ is bounded for $\lambda > 0$ it follows

$$lim_{n\to\infty}(V_1^n - V_1)(-\Delta + \lambda)^{-1}| = 0 \quad \forall \lambda > 0 \tag{85}$$

It suffices therefore to prove that if $W \in L^2(R^3)$ then $W(-\Delta + \lambda)^{-1}$ is compact for $\lambda > 0$. Using the explicit form of the integral kernel of $(-\Delta + \lambda)^{-1}$ one verifies easily that it is a Hilbert-Schmidt operator.

Setting $H_0 \equiv -\Delta$ we prove now that for a suitable choice of λ the operator $I + V(-\Delta - \lambda)^{-1}$ is invertible. This is certainly true if $Im\ \lambda \neq 0$ and the inverse operator is $I - V(H - \lambda)^{-1}$. This operator is bounded if $\rho\ V(H - \lambda)^{-1}$ is bounded for any $\rho \in R$. Therefore $1 + V(H_0 + \lambda)^{-1}$ is invertible if $1 + \rho V(-\Delta + \lambda)^{-1}$ is invertible for some value of ρ.

Consider the operator $zV(H_0 - \lambda)^{-1}$, $\lambda \notin \sigma(H_0)$. For every λ this operator is compact and $V \prec H_0$ the function $F(z) \equiv zV(H_0 - \lambda)^{-1}$ is analytic for z per $|z| < 1$. Moreover $1 + F(0)$ is invertible. According to the Fredholm alternative the function $F(z)$ is invertible for $|z| < 1$ except for at most a denumerable number of points. Therefore there exists at least one ρ_0 for which $I + V(H_0\lambda)^{-1}$ is invertible. Since $W \in L^2 \ \Rightarrow W \prec\prec H_0$, for every $a < 1$ if $Re\lambda = 0$ one has

$$|\rho_0 V(H - \lambda)^{-1}\phi| \leq \rho_0 a[|H(H - \lambda)^{-1})\phi| + |V(H - \lambda)^{-1}\phi + \rho\lambda^{-1}\phi| \tag{86}$$

It follows $\rho_0(1 - a)|V(H_0 - \lambda)^{-1}\phi| \leq \rho_0(a + b\lambda^{-1})|\phi|$. Therefore for every finite interval I

$$I \cap \sigma_{ess}(H_0 + V) = I \cap \sigma_{ess} H_0 \tag{87}$$

On the other hand $-\infty$ is not a limit point of $\sigma(H_0 + V)$ since $V \prec H_0$ implies that $H_0 + V$ is bounded below.

This achieves the Proof of Proposition 17. ♡

Notice that $\frac{c}{|x|} \in L^2(R^3) + L^\infty_\epsilon(R^3)$. Therefore all negative eigenvalues of the hydrogen atom have finite multiplicity and zero is the only limit point.

A generalization of Theorem 15 is the following

Theorem 18 *Let $V = R + W, \quad W \in L^\infty_\epsilon$ with R of Rollnik class. Then $\sigma_{ess}(-\Delta + V) = [0, \infty)$.* ◇

Proof Under the assumptions of Theorem 15 in chapter "Lecture 19: Estimates of the Number of Bound States. The Feshbach Method" the operator $V(-\Delta + 1)^{-1}$ is compact. Let W_n be obtained setting $W = 0$ for $|x| \geq n$. Set $V_n = R + W_n$. Then V_n is of Rollnik class for every finite n. Define $v_n \equiv \frac{V_n}{|V_n|^{1/2}}$. Setting $H_n \equiv -\Delta + V_n$ one has

$$(H_n + \lambda)^{-1} = (H_0 + \lambda)^{-1}$$

$$+ \sum_{k=1}^{\infty} (H_0 + \lambda)^{-1} |v_n| [-v_n (H_0 + \lambda)^{-1} |v_n|]^k v_n (H_0 + \lambda)^{-1} \tag{88}$$

If λ is chosen large enough, for each value of n the series (88) converges in operator norm, and the partial sums converge when $n \to \infty$; moreover the terms in the sum are equi-bounded. Setting $A_n \equiv (-\Delta + V_n)^{-1} |v_n|$ it is easy to see that the kernels of $A_n^* \, A_n$ are positive and equi-bounded in $L^2(R^3 \times R^3)$. The dominated convergence theorem implies that the sequence $(H_n + 1)^{-1}$ converges in norm; since each element is compact also $(-\Delta + V + 1)^{-1}$ is compact. ♡

6 Sch'nol Theorem

We conclude this Lecture with Sch'nol theorem on the spectrum of the Schrödinger operator.

We will call *generalized eigenvalue* a real number E such that

$$(-\Delta + V) u_E = E u_E \tag{89}$$

has a solutions with sub-exponential growth at infinity ($lim_{x \to \infty} |u_E(x)| e^{-\gamma |x|} = 0 \ \ \forall \gamma > 0$).

Sch'nol theorem, a consequence of the first Weyl criterion, states that under very general assumptions on the potential the spectrum $\sigma(H)$ is the closure of the set of generalized eigenvalues. Notice that from this theorem one derives that if a solution of (78) does not belong to L^2 then $E \in \sigma_{ess}$.

We will prove Sch'nol theorem in case V is of Rollnik class. We will use several inequalities that are local estimates of the gradient of the solution and hold for a large class of potentials (in particular Rollnik class). The reader will find useful to verify these inequalities in case $V(x) = 0$ taking advantage of the explicit form of the solution. The proof can be adapted to more general cases using further inequalities that depend on the *local regularity properties* of $V(x)$.

The following inequalities, which compare several norms of the function, are useful to of determine the local regularity of a function or its asymptotics at ∞. We begin with a simple lemma; the proof is by integration by parts, for smooth functions first and then for a larger class by density.

Lemma 19 *If $u \in L^2_{loc}$ and $\Delta u \in L^1_{loc}$, then $\nabla u \in L^2_{loc}$ and for any $\phi \in C_0^\infty$*

$$\int \phi |\nabla u(x)|^2 d^n x = \frac{1}{2} \int \Delta\phi |u(x)|^2 d^n x - \int \phi\, u(x)\, \Delta\bar{u}(x) d^n x \tag{90}$$

Moreover if $H\ u = 0$ and $\phi \in C_0^\infty, \quad \phi \geq 0$ then

$$\int \phi\, |\nabla u|^2 d^n x \leq \frac{1}{2} \int \Delta\phi\, |u|^2 d^n x + \int V_-\, \phi\, |u|^2 d^n x \tag{91}$$

$\diamond$

Using Lemma 19 we prove

Lemma 20 (Poincaré) *If V is of Rollnik class, for each open $\Omega \subset R^n$ and each compact $K \subset \Omega$ there exists a constant C (depending on K, Ω, and on the local behavior of V) for which*

$$\int_K |\nabla u|^2 dx \leq C\, [\int_\Omega |u(x)| dx]^2 \tag{92}$$

$\diamond$

Proof If V is Rollnik potential, $(\Delta + E)u = V\ u$ implies $u \in L^\infty_{loc}$. This in turn implies $\Delta u \in L^1_{loc}$ (since by assumption $V \in L^1_{loc}$). Choosing $\phi \in C_0^\infty$ with support S strictly contained in Ω and such that $\phi(x) \equiv 1$ for $x \in K$ from (89) one derives that there exists c_1 such that

$$\int_K |\nabla u|^2 d^n x \leq c_1 sup_{x \in S} |u(x)|^2 \tag{93}$$

On the other hand, since u is solution of

$$(-\Delta\phi, u) + ((V - E)\phi, u) = 0, \qquad \Delta u \in L^1_{loc}, \quad V\ u \in L^1_{loc} \tag{94}$$

one can prove that if V is of Rollnik class one can still have an analogue of the estimate for sub-harmonic functions (which correspond to the case $V = 0$)

$$|u(x)| < c_3 \int_{|x-y|<r} |u(y)| d^n y \tag{95}$$

for a suitable constant c_3 and for r chosen so the ball of radius r centered in x is contained in Ω. This, together with (93) implies (92). ♡

We shall use also the following estimate of the L^2 norm of the gradient of a function F in cube C as a function of the norm of F in a larger cube.

Lemma 21 *Denote with C_r the hypercube defined by $C_r = \{x \in R^n, max_i |x_i| \leq r\}$.*

Let V be of Rollnik class Then for every solution of $(\Delta + V)\ u = 0$ and for any choice of r

$$\int_{C_{r+1}/C_r} |\nabla\ u|^2 d^n x\ < c \int_{C_{r+2}/C_{r-1}} |u|^2 d^n x \tag{96}$$

where c is independent from n (it depends only on V). ◇

Proof Let C a cube of side 1, C' a cube of side 3 centered on C. From (96) and Schwarz inequality one derives that there exist constants b and c fior which

$$\int_C |\nabla u|^2 d^n x \leq b[\int_C |u| d^n x]^2 \leq c \int_{C'} |u|^2 d^n x$$

The proof is achieved summing these estimates on a partition of C_{r+1}/C_r in unit cubes. ♡

Theorem 22 (Sch'nol) [5] *Let V be a Rollnik potential, let E be a real number and $u_E(x)$ a solution of $Hu_E = Eu_E$ with $H = -\Delta + V$ and suppose that $u_E(x)$ has a* sub-exponential growth *at infinity. Then $E \in \sigma(H)$.* ◇

Proof Without loss of generality take $E = 0$. If the solution is in L^2 one has $0 \in \sigma(H)$ ($u(x)$ is an eigenfunction). Hence we may assume $u \notin L^2$.

Define as in Lemma 9 the hypercubes C_r and choose $\eta_r \in C_0^\infty(R^n)$ with support in C_{r+1}, equal to one in C_r and such that $|\Delta\ \eta_r|_\infty \leq M, \quad |\nabla \eta_r|_\infty \leq M$. Let $w_r \equiv \frac{\eta_r\ u}{|\eta_r\ u|_2}$ so that $|w_r|_2 = 1\ \forall r$. If there exists a sequence $\{r_n\} \in N, \quad r_n \to \infty$ such that

$$|H\ w_{r_n}| \to 0 \tag{97}$$

this sequence satisfies the first Weyl comparison theorem and therefore $0 \in \sigma(H)$. To prove existence of such a sequence remark that

$$\Delta\ (\eta_r\ u) = (\Delta\ \eta_r)\ u + (2\nabla\ \eta_r \cdot \nabla\ u) + \eta_r\ \Delta\ u$$

From $H\ u = 0$ it follows $H(\eta_r u) = -(\Delta \eta_r)\ u + 2\nabla \eta_t \cdot\ \nabla u$. Since both $|\Delta \eta_r|_\infty$ and $|\nabla \eta_r|_\infty$ are uniformly bounded, by Lemma 21 one derives, for suitable constants c_1 and c_2

$$|H\ (\eta_r u)|^2 \leq c_1 \int_{C_{r+1}/C_r} (|u|^2 + |\nabla u|^2) d^n x \leq c_2 \int_{C_{r+2}/C_{r+1}} |u|^2 d^n x \qquad (98)$$

Define $N(r) \equiv \int_{c_r} |u|^2 d^n x$. From (96)

$$|H\ w_r|^2 \leq c \frac{N(r+2) - N(r-1)}{|\eta_r\ u|^2} \leq \frac{N(r+2) - N(r-1)}{N(r-1)}$$

To prove the theorem we must show that there exists a sequence $\{r_k\}$ such that

$$\frac{N(r_k + 2) - N(r_k - 1)}{N(r_k - 1)} \to 0, \qquad k \to \infty$$

We proceed by contradiction. Assume such sequence does not exist. Then there exist $R \in \mathcal{N}$ and $\alpha > 0$ such that

$$r > R \Rightarrow \frac{N(r+2) - N(r-1)}{N(r-1)} \geq \alpha > 0$$

This implies $N(r+3) \geq (1+\alpha)N(r)$ when $r > R+1$ and by induction $N(R + 3s) \geq (1+\alpha)^s N(r)$ for any s. But then $N(r)$ has exponential growth against the assumption. ♡

For periodic potentials in R^n the analogue of Sch'nol theorem (with the same proof) is known as *Bloch's theorem*: E belongs to the spectrum of $H = -\Delta + V$ with V periodic and of suitable class iff there is a *bounded solution* of $Hu = Eu$.

References

1. Akhiezer, N., & Glazman, I. (1963). *Theory of linear opertors in Hilbert space* (Vols. I, II). New York: Frederick Ungar Publisher Co. (F. Ungar ed., 1961).
2. Albeverio, A., Gesztesy, F., Hoegh-Kroen, R., & Holden, H. (1988). *Solvable models in quantum mechanics*, New York: Springer.
3. Faddaev, L. (1965). *Mathematical aspects of the three-body problem*. New York: Daniel Davey & Co.
4. Kato, T. (1980). *Perturbation theory for linear operators*. New York: Springer.
5. Reed, M., & Simon, B. (1975). *Method of mathematical physics* (Vols. I, II, IV). London: Academic Press.
6. Yoshida, K. (1971). *Functional analysis*. New York: Springer.

Lecture 18: Weyl's Criterium, Hydrogen and Helium Atoms

1 Weyl's Criterium

We begin this Lecture with a discussion of a criterion established by H. Weyl in order to verify whether or not a symmetric Schrödinger operator on R^+ is self-adjoint. This criterium is often used in the case of radially symmetric potentials, upon reduction to polar coordinates.

Notice that in this case for strictly positive values of the angular momentum the potential in the reduced system becomes singular at the origin and the analysis in terms of boundary conditions is difficult. Most of the statements in the Weyl's criterium are contained in von Neumann's theory of self-adjoint extensions, but Weyl's [3–5] criterium is older and moreover and it gives conditions under which the operator is *essentially self-adjoint* on a suitable domain.

Theorem 1 (Weyl criterium) *Let $V(x)$ a real valued function on the positive real axis, continuous for $x > 0$. The function may be singular for $x \to 0$ R^+.*

The operator

$$H = -\frac{d^2}{dx^2} + V(x) \tag{1}$$

is essentially self-adjoint *on $C_0^\infty(R^+)$ if and only if the equation*

$$-\frac{d^2\phi(x)}{dx^2} + V(x)\phi(x) = i\lambda\phi(x) \tag{2}$$

has for at least one *real value of* λ (*and therefore for* all *real values of* λ) only one solution *which is square-integrable at the origin and* only one solution *that is square-integrable at infinity and* these two solutions do not coincide. ◇

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_18

Remark that if the condition stated holds true for λ_0 it holds true for *all* other values of λ. Indeed, suppose ϕ_0 solves (2) when $\lambda = \lambda_0$; take $\lambda_1 > \lambda_0$. Then $\phi_1(x)$ defined by

$$\phi_1(x) \equiv e^{-\sqrt{i(|\lambda_1 - \lambda_0)x}}\phi_0(x) \tag{3}$$

solves (2) for $\lambda = \lambda_1$. Choosing conveniently the determination of the square root, $\phi_1(x)$ can be made square-integrable at the origin (resp. at infinity) together with ϕ_0. So we can choose arbitrarily the parameter λ. In the following we choose $\lambda = 1$.

Proof of Theorem 1
(a) Necessity. We prove first that the condition is necessary. Notice that the adjoint H^* is by definition the differential operator

$$H^* = -\frac{d^2}{dx^2} + V(x) \tag{4}$$

acting on the space of distributions on R^+. Therefore we must prove that under the hypothesis of Weyl theorem there are in $L^2(R^+)$ no solutions in the distributional sense of the equation. The equation

$$(-\frac{d^2}{dx^2} + V(x))u(x) = iu(x) \tag{5}$$

is an ordinary elliptic differential equation with regular coefficients, therefore every weak solution is a classic solution. We must prove that there is no square-integrable function on R^+ which satisfies (5).

From the general theory of ordinary differential equations it is known that when $V(x)$ is a continuous function there is always a solution $u_0(x)$ of (5) which is regular at the origin (and therefore square-integrable at the origin) and also a solution $v_0(x)$ which is square-integrable at infinity. If $v_0(x)$ were proportional to $u_0(x)$, the operator $A \equiv -\frac{d^2}{dx^2} + V(x)$ would not be self-adjoint because $(A - iI)\phi = 0$ would have a solution in L^2.

The condition $v_0(x) \neq c\, u_0(x)$ is therefore necessary.
(b) Sufficiency We prove now that the condition is sufficient.

Let $f, g \in D(H^*)$ and define

$$W_x(f, g) \equiv f'(x)g(x) - f(x)g'(x) \qquad x \in (0, +\infty) \tag{6}$$

The function W_x is called *wronskian* of the pair f, g at the point x. Remark that $f \in D(H^*)$ implies that f' is a continuous function. Moreover one has $a, b \in R^+$

$$W_b(f, g) - W_a(f, g) = \int_a^b [H^* \bar{f}(x)g(x) - \bar{f}(x)(Hg)(x)]dx \tag{7}$$

Taking the limit $a \to 0$ and $b \to \infty$

$$W_\infty(f,g) - W_0(f,g) = (H^* f, g) - (f, H^* g) \tag{8}$$

To prove that H^* is a symmetric operator we must show that for any pair of functions $f, g \in D(H^*)$ one has $W_\infty(f,g) = W_0(f,g)$.

Choose x_0 positive and define

$$A \equiv (-\frac{d^2}{dx^2} + V(x)) \qquad D(A) = \{\phi,\ \phi \in C^\infty(0, x_0), \phi(0) = \phi(x_0) = 0\} \tag{9}$$

The operator A is a regular perturbation of the self-adjoint operator of $-\frac{d^2}{dx^2}$ with Dirichlet boundary conditions and is therefore self-adjoint. By construction H^* is an extension of A to the real line. Another possible extension of A to all R^+ is obtained by setting the extension to be equal to zero on the functions with support in $(x_0, +\infty)$.

For any given function $f \in D(A)$ choose a function $f_1 \in C_0^\infty(0, \infty)$ such that

$$f_1 \in D(H^*) \qquad f_1(x_0) = f(x_0) \tag{10}$$

and set

$$f_2(x) \equiv f(x) - f_1(x_0) \tag{11}$$

Notice that it is still $f_2 \in D(A)$. It follows $f_2(x_0) = 0,\ f_2'(0) = f'(0)$.

Define in the same way

$$g_2(x) \equiv g(x) - g_1(x), \quad g_1 \in C_0^\infty, \quad g_1 \in D(H^*), \quad g_2(x_0) = 0, g_2'(0) = g'(0). \tag{12}$$

Then

$$W_0(f_2, g_2) = W_0(f, g), \qquad W_{x_0}(f_2, g_2) = 0 \tag{13}$$

We have seen that A extends H. Therefore $f, g \in D(A)$. Moreover

$$(Af_2, g_2) - (f_2, Ag_2) = W_{x_0}(f_2, g_2) = 0 \tag{14}$$

Since this can be done for any function in the domain of A it follows that f_2, g_2 are arbitrary functions in the domain of A. Hence A is symmetric. On the other hand it is not possible that $f_2(x) = f(x) + cx;\ g_2(x) = g(x) + cx\ \forall f, g$ because otherwise there would be two solutions which are integrable at zero to the assumption. Hence A is a proper restriction of H^* and therefore it is self-adjoint, and essentially self-adjoint on $C_0^\infty(0, \infty)$. This ends the proof of Theorem 1. ♡

In Weyl' criterion one often refers to the case in which no solution is integrable both at zero and infinity as *limit point case* whereas if there is such solution as the

limit circle case. This notation reflects the fact that in the second case one has a one-parameter family of self-adjoint extension that can be parametrized by S^1.

We shall now give two applications of Weyl's criterion.

Lemma 2 *If* $V(x) \geq (-\frac{3}{4} + \epsilon)|x|^{-2}$ *in a neighborhood of the origin,* $\epsilon > 0$ *then*

$$(-\frac{d^2}{dx^2} + V(x))u(x) = 0 \tag{15}$$

has only one solution which is square-integrable at the origin. Therefore if it has only one solution which is square-integrable at infinity (as is the case e.g. if $V(x) = 0$ *for* $x > R$*) the operator* $-\frac{d^2}{dx^2} + V(x)$ *is essentially self-adjoint on* $C_0^\infty(R^+)$. ◇

Proof of Lemma 2 Consider first the case $V = -\frac{C}{x^2}$, $C > 0$. In this case the statement is true since the equation

$$(-\frac{d^2}{dx^2} + \frac{c}{x^2})u(x) = 0 \tag{16}$$

has the solution

$$x^{\alpha_\pm}, \qquad \alpha_\pm = \frac{1 \pm \sqrt{1 + 4c}}{2} \tag{17}$$

Notice next that if $V(x) > V_1(x)$ for $0 < x < \delta$ and if

$$-\frac{d^2u(x)}{dx^2} + V_1(x)u(x) = 0 \tag{18}$$

has only one square-integrable solution at the origin, also

$$-\frac{d^2u(x)}{dx^2} + V(x)u(x) = 0 \tag{19}$$

has this property. Indeed if $u(x)$ solves (16), if $v(x)$ solves (15) and $v(\delta) = u(\delta), u'(\delta) = v'(\delta)$, then $u(x) \geq v(x), 0 < x < \delta$ and therefore

$$\int_\epsilon^\delta |u(x)|^2 dx \to 0 \Rightarrow \int_\epsilon^\delta |v(x)|^2 dx \to 0$$

This ends the proof of Lemma 2. ♡

As a further example of application of Weyl's criterion we study the operator

$$H \equiv -\Delta + V(|x|), \qquad x \in R^d \tag{20}$$

where $V(r)$ is singular at the origin and of class C^∞ outside the origin. In polar coordinates

$$L^2(R^d, dx) \simeq L^2(R^+, r^{d-1}\, dr) \otimes L^2(S^d, d\Omega) \tag{21}$$

(S^d is the ball of radius one in R^d centered at the origin) the operator has the form

$$H = -\frac{d^2}{dr^2} + \frac{d-1}{r}\frac{d}{dr} + \frac{J^2}{r^2} + V(r) \tag{22}$$

where J is the Laplace-Beltrami operator on S^d. The operator J commutes with H; denote with λ_k its eigenvalues. The corresponding subspaces have finite dimension.

Define the isometric invertible operator

$$U: \quad L^2(R^+, r^{d-1}dr) \to L^2(R^+, dr); \qquad U\phi(r) \to r^{\frac{d-1}{2}}\phi(r). \tag{23}$$

One has, on the subspace which correspond to the eigenvalue λ_l of J

$$UH_lU^{-1} = -\frac{d^2}{dr^2} + V(r) + [\frac{(d-1)(d-3)}{4} + \lambda_l]\frac{1}{r^2} \tag{24}$$

From the preceding lemma one derives that H_k is essentially self-adjoint if

$$V(r) + (\frac{(d-1)(d-3)}{4} + \lambda_l)\frac{1}{r^2} \geq -\frac{\frac{3}{4} - \epsilon}{r^2} \quad \epsilon > 0 \tag{25}$$

while it has a one-parameter family of self-adjoint extensions if

$$V(r) + (\frac{(d-1)(d-3)}{4} + \lambda_l)\frac{1}{r^2} < -\frac{c}{r^2} \qquad c = 3/4 \tag{26}$$

Notice that if $d \geq 4$, one has $\frac{(d-1)(d-3)}{4} \geq 3/4$. Since $\lambda_l \geq 0$ one concludes that for $d \geq 4$ if $V(r)$ is less singular then $\frac{1}{r^2}$ the operator $-\Delta + V(r)$ is essentially self-adjoint on $C_0^\infty(R^d)$.

If $d = 3$ the eigenvalues λ_l of J^2 have the form $l(l+1)$ for $l = 0, 1, 2, \ldots$

In this case $\frac{(d-1)(d-3)}{4} + \lambda_l = \lambda_l$ and the self-adjointness condition is satisfied for $l \geq 1$ (if $V(r)$ is less singular than $\frac{1}{r^2}$) while for $l = 0$ the potential must satisfy $V(r) \geq -\frac{\frac{3}{4}-\epsilon}{r^2}$ at the origin.

For $d = 2$ one has $\lambda_l = 0, 1, 2, \ldots$ and $\frac{(d-1)(d-3)}{4} + \lambda_l = -1/4 + \lambda_l$. Also here the condition $-\frac{1}{4} + \lambda_l \geq \frac{3}{4} - \epsilon$ is not satisfied only for $\lambda_l = 0$.

2 Coulomb-Like Potentials. Spectrum of the Self-adjoint Operator

We shall analyze the properties of the spectrum of the Schrödinger operator for Coulomb-like systems. Coulomb-like systems have hamiltonians

$$H_{V,\alpha} = H_\alpha + V(x) \qquad H_0 = -\frac{1}{2m}\Delta + \frac{\alpha}{|x|} \quad m > 0 \quad x \in R^3 \tag{27}$$

where $V(x)$ is relatively compact with respect to H_0.

We shall treat first the case $V = 0$.

In "Lecture 17: Kato-Rellich Comparison Theorem. Rollnik and Stummel Classes. Essential Spectrum" we have seen that in dimension three the operator $\frac{1}{|x|}$ is infinitesimal with respect to the Laplacian. Therefore the operator H_0 is self-adjoint, essentially self-adjoint on the intersection $D(-\Delta) \cap D(\frac{1}{|x|})$. Its essential spectrum coincides with that of $-\frac{1}{2}\Delta$ and coincides therefore with R^+.

We shall now study the *point spectrum* of H_α.

Theorem 3 *If $\alpha \geq 0$ the point spectrum is empty. If $\alpha < 0$ the spectrum is bounded below, the eigenvalues are all negative, they are infinite in number and have 0 as limit point.* ◇

Proof

(1) We first prove that $H_\alpha \phi_E = E\phi_E \quad \phi_E \in L^2(R^3, dx)$ implies $\alpha < 0, E < 0$. Consider the unitary one parameter dilations group

$$U(\beta)\, x\, U^*(\beta) = e^\beta\, x, \qquad U(\beta)\, \hat{p}\, U^*(\beta) = e^{-\beta}\, \hat{p} \tag{28}$$

Its generator is $\frac{1}{2}(x\ \hat{p} + \hat{p}\ x)$ and its action on $H\alpha$ is

$$U(\beta)\, H_\alpha\, U^*(\beta) = e^{-2\beta}\, H_0 + e^{-\beta}\frac{\alpha}{|x|}, \qquad H_0 \equiv \frac{|\hat{p}|^2}{2m} \tag{29}$$

Therefore if $(H - E)\phi_E = 0, \quad \phi_E \neq 0$

$$(e^{-2\beta}\, H_0 + e^{-\beta}\frac{\alpha}{|x|} - E)U(\beta)\phi_E = 0 \tag{30}$$

Taking the scalar product with $U(\beta)\phi$

$$(U(\beta)\phi, [(1 - e^{2\beta})\, H_0 + (1 - e^\beta)\frac{\alpha}{|x|}]U(\beta)\phi) = 0 \quad \forall \beta \neq 0$$

Dividing by β, taking the limit $\beta \to 0$ and recalling that $lim_{\beta\to 0} U(\beta)\phi_E = \phi_E$ one obtains

$$(\phi_E, H_0\phi_E) = -\frac{\alpha}{2}(\phi_E, \frac{1}{|x|}\phi_E) \tag{31}$$

From (31) one derives that $\alpha < 0$ (H_0 is a positive operator).

Denote now by ϕ_E the eigenvector of H_α to the eigenvalue E normalized to have norm equal to one. From (31) one derives

$$E = (\phi_E, H\phi_E) = (\phi_E, H_0\phi_E) + \alpha(\phi_E, \frac{1}{|x|}\phi_E) = \frac{\alpha}{2}(\phi_E, \frac{1}{|x|}\phi_E) < 0 \tag{32}$$

From (31) and (32) follows

$$(\phi_E, H_0\phi_E) = -\frac{\alpha}{2}(\phi_E, \frac{1}{|x|}\phi_E)$$

a relation which is known with the name *virial theorem*.

From the derivation given above it is clear that Theorem 3 holds for potentials of the form $-\frac{C}{|x|^\gamma}$, $0 < \gamma < 2 \quad C > 0$. For $\gamma \geq 2$ derives still *formally* $\alpha < 0$ and $E \geq 0$ but recall that if $\gamma \geq 2$ the potential is no longer a small perturbation of H_0 and the operator is only symmetric. For $\gamma = 2$ we have seen in the first part of this Lecture the condition on the coefficient C under which the operator is self-adjoint.
(2) We shall now prove that there are infinitely many negative eigenvalues. Since the spectrum is bounded below and the essential spectrum coincides with R^+ their accumulation point must be zero. To prove the statement we shall construct an infinite-dimensional subspace $\mathcal{K} \in \mathcal{H}$ such that

$$(\phi, H\phi) < 0 \qquad \forall \phi \in \mathcal{K} \tag{33}$$

We shall make use again of the dilation group. It follows from (31) that for each $\phi \in \mathcal{H}$

$$(U(\beta)\phi, HU(\beta)\phi) = e^{-2\beta}(\phi, \frac{\hat{p}^2}{2\mu}\phi) + \alpha e^{-\beta}(\phi, \frac{1}{|x|}\phi) \tag{34}$$

Since $\alpha < 0$ for each ϕ there exists $\beta_0(\phi) > 0$ such that for $\beta > \beta_0$ the right hand side of (34) is negative.

Fix $0 < R_1 < R_2$ and choose ϕ supported in $R_1 < r < R_2$. If the support of ϕ is S then $U(\beta)\phi$ has support in $e^\beta S_\phi$. It follows that there exists an infinite increasing sequence $\{N_i\}$ such that the supports of $U(N_i\beta)\phi$ are pair-wise disjoint (so that the vectors are pair-wise orthogonal) and

$$(U(N_i\beta)\phi, HU(N_i\beta)\phi) < 0 \quad \forall N_i \tag{35}$$

From the min-max principle one concludes that there are infinitely many negative eigenvalues.

(3) A similar analysis shows that for any choice of $\epsilon > 0$ there exist an infinite dimensional subspace $\mathcal{K}_\epsilon$ such that

$$-\epsilon < (U(N_i\beta)\phi, \frac{1}{|x|}U(N_i\beta)\phi) < 0 \quad \forall\phi \in \mathcal{K}_\epsilon, \quad |\phi| = 1 \tag{36}$$

It follows that there exists a sequence ϕ_n of bound states of the hydrogen atom that have essential support outside a ball of radius R arbitrary large, i.e.

$$lim_{n\to\infty} \int_{|x|<R} |\phi_n(x)|^2 dx = 0$$

In this sense the support of the bound states of the hydrogen atom *extends to infinity* and we must expect some difficulties in the formulation of scattering theory for the Coulomb potential in the Schrödinger representation. ♡

We have so far considered the case $V = 0$. Since $V(x)$ is bounded, it is a regular perturbation of H_0. Therefore the essential spectrum of H is the same as the essential spectrum of H_0. And also for $V \neq 0$ there are infinitely many bound states since we have assumed that V is relatively compact and therefore the difference in the number of bound states is finite. If $V(x)$ has compact support, by minimax one can prove that also in this case the support of the bound states extends to infinity.

3 The Hydrogen Atom. Group Theoretical Analysis

In classical mechanics for negative energies the Keplerian system has only periodic solutions and for positive energies has a rather elementary scattering theory. Both features are due to the presence of the Runge-Lenz vector, that is constant of motion and transforms as a vector under rotation.

For negative energies the Poisson brackets induce between the angular momentum (a vector) and the Runge-Lenz vector (properly normalized in each energy shell) the structure of the $so(4)$ Lie algebra. This reduces to one the number of independent action variables; one determines the motion by quadratures and shows that all orbits are closed.

For positive energies the induced algebra is $so(3, 1)$, the trajectories are open and unbounded and it is possible to determine algebraically their asymptotic behavior (Compton scattering).

As mentioned in "Lecture 4: Entanglement, Decoherence, Bell's Inequalities, Alternative Theories" this structure has been used by Pauli, exploiting the identity of Poisson Bracket of these constant of motion with the commutators of the corresponding generators in Quantum Mechanics. This is possible, in spite of the fact the the generators are bilinear in the canonical variables, due to the symmetry of the problem, which cancels the polynomials of degree two in the commutators of the angular momentum and the Runge-Lenz vector (both represented by second order polynomials in quantum canonical variables).

This identity allows to give a complete group-theoretical description of the quantum version of the Keplerian system (i.e. the hydrogen atom) and in particular of its spectrum. The analysis made by W. Pauli of the quantum structure of the hydrogen atom by means of the so(4) Lie algebra paved the way to the rapid acceptance of Quantum Mechanics by a large part of the Physics community.

The analysis of Pauli is in the Heisenberg representation. If transcribed in the Schrödiger representation it allows to determine the eigenfunctions of the bound states and to verify in this case the *correspondence principle* according to which for energy very close to zero (the ionization threshold) the essential support of the quantum eigenfunctions in units of $\hbar^{-1}$ is in a neighborhood of size $\sqrt{\hbar}$ of the classical orbit, in accordance with the Correspondence Principle.

We expand here on the analysis given by Pauli [2], including some mathematical details. I particular we shall use the results in "Lecture 17: Kato-Rellich Comparison Theorem. Rollnik and Stummel Classes. Essential Spectrum" to prove that the Schrödinger operator for the Hydrogen (and Helium) atom is self-adjoint [3]. This cannot be seen by algebraic techniques and is not true in the one and two-dimensional case. At the end of this Lecture we shall analyze briefly the one and two-dimensional case.

The Hamiltonian of the hydrogen atom is

$$H = -\frac{1}{2m_1}\Delta_1 - \frac{1}{2m_2}\Delta_2 + \frac{e_1\, e_2}{|x_1 - x_2|} \tag{37}$$

where e_1, e_2 are the charges of the electron and of the proton, $m_1 m_2$ their masses, $x_1\ x_2, \in R^3$ their coordinates. The operator H acts on the Hilbert space

$$\mathcal{H} \equiv L^2(R^3, dx_1) \otimes L^2(R^3, dx_2) \tag{38}$$

Making use of the coordinates of the center of mass X and of the relative coordinates x the hamiltonian takes the form

$$H = -\frac{1}{2M}\Delta_X - \frac{1}{2\mu}\Delta_x + \frac{e_1\, e_2}{|x|} \tag{39}$$

where $M = m_1 + m_2$, $\mu = \frac{m_1 m_2}{m_1 + m_2}$.

As in Classical Mechanics the first term in the hamiltonian describes free motion of the barycenter. We shall limit ourselves therefore to the study of relative motion and we shall denote again by H the corresponding hamiltonian. With the notation $e_1 e_2 = \alpha$ the hamiltonian of the reduced system is

$$H = -\frac{1}{2\mu}\Delta_x + \frac{\alpha}{|x|} \tag{40}$$

i.e. the hamiltonian of a Coulomb (Kepler) system.

Before analyzing the algebraic structure of the system described by the hamiltonian H we state and prove a technical lemma which we are going to use

Technical Lemma *Let the operator A be self-adjoint and let B be essentially self-adjoint on a domain $\mathcal{D} \subset D(B)$ that is left invariant by the unitary group generated by A. If*

$$B\, e^{-iAt}\phi = e^{-iAt}\, B\phi, \quad \forall \phi \in \mathcal{D}, \quad \forall t \in R \tag{41}$$

then any bounded function of B commutes with any bounded function of A. If B is not closed, (41) *holds for the closure $\bar{B}$ of B.* ◇

Proof Set $W \equiv (B - iI)\,(B + iI)^{-1}$. The operator W has an extension to a unitary operator and for every $\psi \in \mathcal{H}$ there exists ϕ such that $\phi = (B + iI)\psi$.

Therefore since $\phi \in D$ by (41)

$$W\, e^{iAt}\phi = (B - iI)e^{iAt}\psi = e^{iAt}(B - iI)\psi = e^{iAt}W\phi \tag{42}$$

By density $We^{iAt} = e^{iAt}W$. Therefore any bounded function of the operator A commutes with any bounded function of W. The thesis of the Theorem follows noticing that any bounded function of B is also a bounded function of W. ♡

We use this technical Lemma for the proof of

Theorem 4 *The vector valued operators vectors L (angular momentum) and R (Runge-Lenz vector) defined by*

$$L = x \wedge \hat{p} \qquad R = \frac{1}{2}[\hat{p} \wedge L + L \wedge \hat{p}] + \mu\alpha\frac{x}{|x|} \tag{43}$$

where $\hat{p}_k \equiv= -i\frac{\partial}{\partial x_k}$ have self-adjoint components which are essentially self-adjoint on $C_0^\infty \cap D(H)$ and commute with H in the sense of the Technical Lemma. Moreover on the common domain of essential self-adjointness they satisfy

$$L.R = R.L = 0, \quad [L_n, L_l] = \sum_{n,l=1}^{3} \epsilon_{n,l,s}L_s \quad n, l = 1 \ldots 3 \quad [L, H] = [R, H] = 0$$

$$[R_n, R_l] = -2i\mu H \sum_{k=1}^{3} \epsilon_{n,l,k}L_k, \quad [L_n, R_l] = \sum k = 1^3 \epsilon_{n,l,k}L_k \quad n.l = 1 \ldots 3$$

$$|R|^2 = 2\mu H(|L|^2 + 1) + \mu^2\,\alpha^2 \tag{44}$$

where $\epsilon_{n,l,k}$ is Ricci's symbol. ◇

Proof Once one notices that all the operators which appear on (44) are essentially self-adjoint in $C_0^\infty \cap D(H)$ the proof consist of explicit calculations, making use of

the definitions of L_k and R_h and of the canonical commutation relations of $\hat{x}_h$ and $\hat{p}_h$. This proof is the same as in the classical case.

The commutation relations between L_k and R_h are consequence of the vector property of R. The commutation relations between R_k and R_h are easy to verify if one takes into account the fact that $\hat{x}$ and $\hat{p}$ are vectors. ♡

Set

$$A_k \equiv \frac{1}{2}[L_k + R_k(-2\mu H)^{-\frac{1}{2}}]\Pi_- \tag{45}$$

$$B_k \equiv \frac{1}{2}[L_k - R_k(-2\mu H)^{-\frac{1}{2}}]\Pi_- \tag{46}$$

where we have denoted by Π_- the projection operator on the negative part of the spectrum of H. Using (44) it is possible to verify that on the common domain of essential self-adjointness the following relations are satisfied

$$[A_k, A_h] = i\ \epsilon_{k,h,l} A_l \qquad [B_k, B_h] = i\ \epsilon_{k,h,l} B_l \tag{47}$$

$$\sum_k A_k^2 = \sum_k B_k^2 = [\frac{1}{4} - \frac{\mu\alpha^2}{8H}]\ \Pi_- \tag{48}$$

(recall that due to Hardy's inequality $H > -\frac{m\alpha^2}{2} I$).

Since A_k and B_k are generators of a representation of SO_4 it follows that the operators $\sum_k A_k^2$ and $\sum_k B_k^2$ have eigenvalues $l(l+1)$ where l is an integer (possibly zero). From (46) and from (47) we derive

$$E_n = -\frac{\mu\alpha^2}{2n^2} \quad n = 1, 3, 5 \ldots \tag{49}$$

We can now construct the eigenvectors of H. Consider the common eigenvectors of H, A_3, B_3 (a maximal set of commuting operators in the sense of the Technical Lemma). The highest possible value for A_3 and B_3 for fixed n, l are obtained when

$$R_+\phi = L_+\phi = 0 \qquad R_\pm \equiv R_1 \pm i R_2 \qquad L_\pm \equiv L_1 \pm i L_2 \tag{50}$$

We determine the eigenfunctions corresponding to this eigenvalues. The other eigenfunctions are then obtained by repeated action of powers of A_- and B_-. It is easy to verify that R and R_+ can be written as

$$R = \frac{i}{2}[\hat{p}, L^2] + \frac{\mu\alpha x}{|x|} \qquad R_+ = \frac{i}{2}[\hat{p}_+, L^2] + \frac{\mu\alpha x_+}{|x|} \tag{51}$$

where

$$\hat{p}_\pm = \hat{p}_1 \pm i\hat{p}_2 \qquad x_\pm = x_1 \pm i x_2 \tag{52}$$

We conclude that if $R_+\phi = 0$ and $H\phi = E_n\phi$ then $L_+\phi = 0$ and therefore ϕ corresponds to the maximum value that L^2 can have in the subspace corresponding to the eigenvalue E_n of the hamiltonian. The corresponding property in the classical case is that the hamiltonian depends only from one of the action variables.

Choose now as maximal set of commuting observables

$$H \quad |L|^2 \equiv \sum L_k^2 \quad L_3 \quad M_3$$

with eigenvalues $E_n, l(l+1), m, \nu$. The state described by (50) corresponds to the eigenvalue m both of L_3 and of R_3.

We consider now the description in the Schrödinger representation. We use polar coordinates with $z \equiv x_3$ so that

$$\frac{\partial}{\partial z} = cos\ \theta \frac{\partial}{\partial r} - \frac{sen\ \theta}{r} \frac{\partial}{\partial \theta} \tag{53}$$

For the eigenfunction $\phi_{n,l,l}$ corresponding $L_3 = l$ one has

$$i\,\hat{p}_+\ \phi_{n,l,l} = (\frac{\partial}{\partial r} - \frac{l}{r})\phi_{n,l+1,l+1}, \qquad \frac{x_+}{|x_+|}\phi_{n,l,l} = \phi_{n,l+1,l+1} \tag{54}$$

Using (52)

$$R_+\phi_{n,n-1,n-1} \equiv [-n(\frac{\partial}{\partial r} - \frac{n-1}{r}) + \mu\alpha]\phi_{n,n-1,n-1} = 0 \tag{55}$$

and therefore

$$\phi_{n,n-1,n-1}(r, \theta, \psi) = cr^{n-1}e^{\frac{-\mu\alpha r}{2}} Y_{n-1}^{n-1}(\theta, \psi) \tag{56}$$

($Y_{n-1}^{n-1}(\theta, \psi)$ are spherical harmonics).

When $\alpha < 0$ this function is in $L^2(R^3, dx) \simeq L^2((0, \infty), r^2dr) \otimes L^2(S^3)$.

The remaining eigenfunctions $\psi_{n,l,m}$ corresponding to $H = E_n$, $L^2 = l(l+1)$, $L_3 = m$ are obtained as

$$\psi_{n,l,m} = L_-^{l-m}\ R_-^{n-1-l}\ \phi_{n,n-1,n-1} \tag{57}$$

One can verify that these functions are a complete orthogonal basis in the subspace corresponding to the eigenvalue E_n of the hamiltonian.

We remark that from (56) one can verify that $\phi_{n,n-1,n-1}(x)$ as a function of the radius r reaches its maximal value at

$$r = \frac{n(n+1)}{\mu|\alpha|} = n(n+1)r_B \tag{58}$$

where $r_B \equiv \frac{1}{\mu|\alpha|} \simeq 0.52910^{-8}$ cm is the *Bohr radius*.

A rough description of the support of the probability density $|\psi_{n,n-1,n-1}(x)|^2$ is given by its mean and variance. One obtains

$$\int |\psi_{n,n-1,n-1}(x)|^2 |x| dx = n(n+1/2) r_B$$

$$\int |\psi_{n,n-1,n-1}(x)|^2 |x|^2 dx = n(n+1/2)(n+1) r_B^2 \tag{59}$$

When n is very large the function tends to concentrate on a spherical shell of mean radius $n(n+1/2)r_B$ and thickness $\simeq n^{\frac{3}{2}}$ so that the ration between radius and thickness decreases as $n^{-\frac{1}{2}}$.

Analogous results are obtained for $\phi(n,l,m)$ when n very large.

4 Essential Spectrum

We study now the *essential spectrum* of H_α. We claim that the singular continuum spectrum is empty.

Let $U(\beta)$ be the representation of the one-parameter group of dilation. The family of operators $U(\beta)$ can be continued as a function of β in the complex plane and defines an analytic function $U(z)$ for $Imz \neq 0$.

The operators $U(z)$ are not bounded but there is a dense set $\mathcal{D}$ of analytic vectors for the generator of dilations on which they all commute (e.g. the elements of the the bounded spectral subset of the generator).

For $\phi, \psi \in \mathcal{D}$ one has

$$(\phi, (H_\alpha - z)^{-1}\psi) = (U(\beta)\phi, [e^{2\beta}\frac{|\hat{p}|^2}{2\mu} + e^{\beta}\frac{\alpha}{|x|} - z]^{-1} U(\beta)\psi) \quad z \notin R^+, \quad i\beta \notin R \tag{60}$$

Since $\frac{1}{|x|}$ is compact relative to $\hat{p}^2$, for each $\beta \notin iR$ the essential spectrum of $e^{2\beta}\hat{p}^2 + e^{\beta}\alpha\frac{1}{|x|}$ is $e^{2\beta}R^+$.

From (60) we know that if $\phi \in \mathcal{D}$ the matrix elements $(\phi, (H_\alpha - z)^{-1}\psi)$, $\phi \in \mathcal{D}$ can be continued to both complex half-planes $\Im z > 0$ and $\Im z < 0$. This defines two functions, denoted by $F^+_{\phi,H_\alpha}(z)$ and $F^-_{\phi,H_\alpha}(z)$, analytic outside $Imz = 0$, $Rez < 0$. Also their difference is therefore analytic outside R^-.

Denoting by Θ the indicator function of the positive semi-axis we have

$$(\phi, [\Theta(H\alpha - a) - \Theta(H_\alpha - b)]\psi) = lim_{\epsilon \to 0} \int_a^b \frac{1}{2i\pi}[F^+_{\phi,H_\alpha}(x + i\epsilon) - F^-_{\phi,H_\alpha}(x - i\epsilon)] \tag{61}$$

The integrand in (61) is analytic outside R_- and the limits are continuous functions one has

$$lim_{b\to a}(\phi, [\Theta(H_\alpha - a) - \Theta(H_\alpha - b)]\psi) = 0 \tag{62}$$

Since $\mathcal{D}$ is dense, we conclude that the essential spectrum does not contain any set of Lebesgue measure zero in R. Therefore $\sigma_{sing}(H) = \emptyset$ and

$$\sigma_{ess}(H_\alpha) \equiv \sigma_{a.c.}(H_\alpha) \tag{63}$$

The identity (63) is important because it allows to define for Coulomb potential the Wave Operators and the Scattering matrix. But one should be aware that the Coulomb potential has very long range, so that the generalized eigenfunction are distorted. We shall come back to this problem in the second part of these Lectures.

5 Pauli Exclusion Principle, Spin and Fermi-Dirac Statistics

Before the analysis of Helium-like atoms, it will be convenient to make a digression on the *spin*, the *Pauli exclusion principle* and the *Fermi-Dirac statistics*.

5.1 Spin

Recall that in Quantum Mechanics particles having spin degrees of freedom are described by functions in $L^2(R^3) \otimes C^n$ which transform under a rotation $\mathcal{R}$ according to $\psi_k(x) \to U_{k,m}(\mathcal{R})\psi_m(\mathcal{R}x)$ where U is an irreducible representation in C^n of the group $SU(2)$ (the covering group of $O(3)$).

A particle with integer spin is described by an irreducible representation of of the rotation group while a particle of half-integer spin is described by an irreducible representation of $SU(2)$ *which is not* a representation of $O(3)$.

A particle is said *to have spin s* (s an integer or half-integer) if the representation of $SU(2)$ has dimension $2s + 1$. The name *spin* derives from *spinning* and is related to the fact that a representation of $SU(2)$ is a representation up to a phase of the rotation group and phases are irrelevant in the definition of a pure state in Quantum Mechanics.

5.2 Statistics

In Quantum Mechanics the definition of *statistics of identical particles* is introduced in the following way.

If the system is composed by N *identical* particles and each of them is separately described by a normalized vector in $\phi \in L^2(R^3) \otimes C^n$, the total systems is represented by an element of a subspace of $(L^2(R^3) \otimes C^n)^{\otimes N}$ which carries a one-dimensional representation of the permutation group Π_N (acting on the spatial coordinates as well as on the "spin variables").

The only one dimensional representations have correspond to $U(g) = 1$ and $U(g) = (-1)^{p(g)}$ where $p(g)$ is the parity of the permutation. In the first case the system of particles satisfies the Bose-Einstein statistics (i.e. the particles are *bosons*), in the second case it satisfies the Fermi-Dirac statistics (the particles are *fermions*).

According to this definition two particles may have the same physical properties *without being identical*. Remark that the definition of identical particle in Quantum Mechanics *has no counterpart in Classical Mechanics*.

The notation *two identical particles* is in fact misleading since it may lead to believe that the particles retain an identity (although it is the same for both). On the contrary in the quantum mechanical description of a system of N identical particles the particles *lose their identity* (whereas in Classical Mechanics the particles keep their identity but, if they have the same physical properties, became *indistinguishable*).

In quantum mechanics this leads to a conceptual difficulty: one should use a wave functions which is antisymmetric in the coordinates (position and spin) *of all the electrons* including the electrons *behind the moon*. One resolves this difficulty by the remark that the observables of interest are represented by operators with kernels which are *almost* compactly supported. Therefore, *for all practical purposes*, is sufficient to anti-symmetrize with respect to the variables of the particles which are in the support of the observables we are going to measure.

5.3 Pauli Exclusion Principle

In Quantum Mechanics these definitions have an interesting consequence: there is no state of two spin- $\frac{1}{2}$ particle state which is described by

$$\psi(x_i, \sigma_k)\psi(x_2, \sigma_k), \quad k = 1, 2 \quad \psi \in L^2(R^3 \otimes C^2)$$

This is *Pauli exclusion principle* [2]; it is responsible for the stability of matter. On the contrary the particles of spin zero can exist in the same state, giving rise to Bose-Einstein condensation.

The experimental data are *compatible* with the description of half-integer spin particles as fermions and of integer spin particles as bosons. This has led to establish the *empirical* rule of connection between spin and statistics.

It is important to notice that in non-relativistic Quantum Mechanics this principle is based on *empirical observation*. In the Theory of Relativistic Quantized Fields the connection between spin and statistics is instead a consequence of the two following assumptions

(1) the energy spectrum is positive
(2) observables associated to space-like separated regions are represented by operators that commute.

Experimental data suggest that the electrons have spin $1/2$ and satisfy Fermi-Dirac statistics.

A pure state of N electrons will therefore be represented by a wave function

$$\phi_{k_1,..k_N}(x_1, ..x_N) \tag{64}$$

which is antisymmetric with respect to the exchange *of particle index* (and therefore for *simultaneous exchange* of spatial and spin coordinates). Since the spin "coordinate" k can take only two values (the representation of $SU(2)$ we use has dimension two) if $N = 2$ the function ϕ in (64) *can be symmetric* in the space variable *provided it is antisymmetric with respect to the spin variables*.

This corresponds to a representation of the rotation group (the product of the two-dimensional representations of $SU(2)$ decomposes in a one dimensional representation and in a three dimensional representations of $O(3)$).

6 Helium-Like Atoms

We shall now study a system composed of a nucleus of mass M and charge Ze and two electrons of mass m and charge e with $m << M$. When $Z = 2$ this represents the Helium atom. We call *helium-like* this system for any value of Z; for Z integer larger than two it represents ionized rare earths atoms.

It turns out to be convenient to use $\alpha \equiv Z^{-1}$ as a parameter. The system is analogous to the three-body system in Classical Mechanics but, while in the study of the structure of the orbits in this "classical" system one faces serious difficulties, these problems are not present in the quantum counterpart.

We can now proceed with the analysis of the helium-like atoms.

We shall only treat the case in which the mass of the nucleus can be regarded as being infinite. The hamiltonian is

$$H = \frac{1}{2m}(|\hat{p}_1|^2 + |\hat{p}_2|^2) - Z\, e^2(\frac{1}{|x_1|} + \frac{1}{|x_2|}) + \frac{e^2}{|x_1 - x_2|} \tag{65}$$

The potential term is infinitesimal in Kato sense with respect to the kinetic energy. The hamiltonian is invariant under rotations and under permutation of the indices 1 and 2. Since the electrons satisfy the Fermi-Dirac statistics the Hilbert space in which the system is described is

$$\mathcal{H}_2 \equiv (\mathcal{H}_1 \otimes \mathcal{H}_1)_a, \quad \mathcal{H}_1 \equiv L^2(R^3) \otimes C^2$$

where the suffix $_a$ indicates that we keep only the antisymmetric part of the product. It is easy to verify that

$$\mathcal{H}_2 = [(L^2(R^3) \otimes L^2(R^3))_s \otimes (C^2 \otimes C^2)_a] \oplus [(L^2(R^3) \otimes L^2(R^3))_a \otimes (C^2 \otimes C^2)_s] \tag{66}$$

The space $(C^2 \otimes C^2)_a$ has complex dimension one and carries the trivial representation of $SU(2)$. The space $(C^2 \otimes C^2)_s$ has complex dimension three and carries an irreducible three-dimensional representation of $SU(2)$ which is also a vector representation of $O(3)$.

From (66) one sees that, *for operators that do not depend on the spin variables*, one can analyze the two-body problem on

$$(L^2(R^3) \otimes L^2(R^3))_s \oplus (L^2(R^3) \otimes L^2(R^3))_a \equiv L^2(R^3) \otimes L^2(R^3)$$

without reference to spin degrees of freedom.

We shall therefore study the hamiltonian (65) as self-adjoint operator on $L^2(R^3) \otimes L^2(R^3)$. Introduce the dilation

$$U\hat{p}U^* = Z\ m\ e^2\hat{p}, \quad UxU^* = (Z\ m\ e^2)^{-1}x \tag{67}$$

and redefine the hamiltonian through a scaling

$$Z^{-2}e^4m^2H \equiv H_\alpha = \frac{1}{2}(|\hat{p}_1|^2 + |\hat{p}_2|^2) - \frac{1}{|x_1|} - \frac{1}{|x_2|} + \frac{\alpha}{|x_1 - x_2|}, \quad \alpha \equiv Z^{-1} \tag{68}$$

This scaling does not change the spectral structure; if λ is an eigenvalue of $H(\alpha)$ the corresponding eigenvalue of H is $\frac{Z^2\lambda}{e^4\,m^2}$.

We write H_α as

$$H_\alpha = H(0) + \alpha H', \quad H_0 = \frac{1}{2}(|\hat{p}_1|^2 + |\hat{p}_2|^2) - \frac{1}{|x_1|} - \frac{1}{|x_2|} \quad H' = \frac{1}{|x_1 - x_2|} \tag{69}$$

The *physical* values of the parameter α are $\alpha = 1, 1/2, 1/3 \ldots$ For example $\alpha = 1/2$ corresponds to the helium atom, $\alpha = 1/3$ corresponds to a simply ionized lithium atom.

For a mathematical treatment it is convenient to regard α as a positive parameter.

The hamiltonian H_0 describes a system composed by two non-interacting hydrogen atoms. We have seen that the spectrum of the hydrogen atom is composed of an absolutely continuous part coinciding with R^+ and of an infinite number of points on R^-

$$\epsilon_0 \equiv -\frac{1}{2} \quad < \epsilon_1 \quad \leq \epsilon_2 \quad \leq \epsilon_3 \ldots\ldots < 0 \qquad lim_{n\to\infty}\epsilon_n = 0$$

It follows that the spectrum of H_0 has an absolutely continuous part coinciding with $[-\frac{1}{2}, \infty)$ and a point part given by $\{\epsilon_{i,j} = \epsilon_i + \epsilon_j\}$.

Remark that for an infinite number of indices i and j one has $\epsilon_i + \epsilon_j \geq \frac{1}{2}$. It follows that there are elements of the point spectrum of H_α that are *immersed* in the continuum part of its spectrum.

Denote by ϕ_0 the eigenfunction corresponding to the eigenvalue $-\frac{1}{2}$ of the hydrogen atom (the explicit form is (67) for n = 1).

The lower bound (in fact, the infimum) of the spectrum of $H(0)$ is the eigenvalue -1 which is simple and corresponds to the eigenfunction

$$\Phi_0(x_1, x_2, i_i, i_2) = \phi_0(x_1)\phi_0(x_2)\xi_{i_1,i_2}$$

where ξ is an antisymmetric 2×2 matrix. Remark that if the electrons were bosons the ground state energy would have been $-\frac{3}{4}$ and the corresponding eigen-space would have dimension three.

We now prove a theorem about the essential spectrum of H_α.

Theorem 5 *The essential spectrum of H_α does not depend on the value of the parameter α.* ◇

Proof We cannot make use of Weyl's theorem because $H_\alpha, \alpha \neq 0$ *is not* a relatively compact perturbation of H_0. To prove the theorem it will be however sufficient by Weyl's theorem to prove that for any $E \in [-\frac{1}{2}, \infty)$ it is possible to find an orthonormal sequence of vectors ϕ_n for which $lim_{n\to\infty}(H_\alpha - E)\phi_n = 0$ in $L^2(R^3)$.

We shall construct these vectors as anti-symmetrized product of the ground state ϕ_0 of H_0 times vectors ξ_n with support contained in $(2^n,\ 2^{n+1})$ (thus they are pairwise orthogonal) and with a Fourier transform with support in a smaller and smaller neighborhood of $\{p\ : |p|^2 = E + \frac{1}{2}\}$.

Notice that the requirement on the support on configuration space is compatible with the requirement on momentum space since $\hat{p}$ is translation invariant and the diameter of the spherical shell increases without bound when $n \to \infty$.

The condition on the support plays an important role since form Schwartz inequality one has in $L^2(R^3 \otimes R^3, dx)$

$$lim_{n\to\infty} \int \frac{1}{|x_1 - x_2|}[\phi_0(x_1)\xi_n(x_2) - \xi_n(x_i)\phi_0(x_2)]^2 dx_1 dx_2 = 0$$

From $H_0\phi_0 = -\frac{1}{2}\phi_0$ and $|(|p|^2 - 2E - 1)\hat{\xi}_n| \to 0$ one derives

$$lim_{n\to\infty}|(H_\alpha - E)\phi_0\ \xi_n| = 0 \qquad \heartsuit$$

Notice that the continuum part of the spectrum of H_α contains the interval $[-1/2, \infty)$ (consider the case in which one electron is in the ground state and the other is far away and has very small momentum).

On the other hand if $\alpha = 0$ the system has in the interval $[-1, 0]$ an infinity of bound states which correspond to the possible energies $-\frac{1}{2m} - \frac{1}{2n}, mn; \in Z^+$.

7 Point Spectrum

We shall prove that H_α has point spectrum in $[-\frac{1}{2}, 0]$ when α is sufficiently small.

Theorem 6 *If $\alpha < 1$ the operator H_α has infinitely many eigenvalues in $[-1, -1/2)$.* ◇

We remark that the condition $\alpha < 1$ is equivalent to $N > 1$. For the helium atom ($\alpha = 2$) the energies of the bound states are contained in $[-\frac{4}{e^4 m^2}, -\frac{2}{e^4 m^2}]$.

From a physical point of view the condition $Z > 1$ can be interpreted in this way: these bound states correspond to configurations in which one electron forms a bound state with the nucleus. The charge distribution of this electron provides *a partial screening* of the charge of the nucleus and the resulting system has charge $Z - 1 > 0$; this "residual" charge is sufficient to bind the other electron.

Proof of Theorem 6 The proof will follow closely the lines of the qualitative considerations given in the previous remark. We know the continuum spectrum of H_α coincides with $[-\frac{1}{2}, \infty)$; it is therefore sufficient that the subspace

$$\mathcal{K} \equiv \{\phi \ : \ (\phi, H_\alpha \phi) < -\frac{1}{2}\|\phi\|^2 \tag{70}$$

is infinite-dimensional.

We shall use as trial functions the anti-symmetrized product of the ground state ϕ_0 of the hydrogen atom with a function $\xi(x)$ which is localized "very far". We choose f such that $\int |\hat{\xi}|(p)\ p^2 dp \leq \epsilon$ where ϵ is a parameter which will be chosen sufficiently small. Since this is a translation invariant condition we can choose ξ such that

$$|\int \bar{\phi}_0(x_1)\bar{\xi}(x_2)\frac{1}{|x_1 - x_2|}\phi_0(x_2)\xi(x_1)dx_1 dx_2| < \epsilon \tag{71}$$

We have remarked that we can choose the function ψ to be symmetric in the spacial variables and then antisymmetric in the spin degrees of freedoms. Set

$$\psi(x_1,\ x_2) = \frac{1}{\sqrt{2}}[\phi_0(x_1)\ \xi(x_2) + \phi_0(x_2)\ \xi(x_1)] \tag{72}$$

and then

$$\langle \psi, H_\alpha \psi \rangle = -\frac{1}{2}\|\psi\|^2 + \int \bar{\xi}(x)(\frac{\hat{p}^2}{2} - \frac{1}{|x|})\xi(x)d^3x$$

$$+ \alpha \int \bar{\phi}_0(x_1)\ \bar{\xi}(x_2)\frac{1}{|x_1 - x_2|}\phi_0(x_1)\ \xi(x_2)d^3x_1\ d^3x_2 + 2\epsilon \tag{73}$$

To estimate the last term in (73) we can use the explicit form of ϕ_0 to obtain

$$\int |\phi_0|^2 \frac{1}{|x - y|}d^3x = \frac{1}{|y|} - e^{-2y}(1 + \frac{1}{|y|})$$

and therefore

$$\langle \psi, H_\alpha \psi \rangle = -1/2 \|\psi\|^2 + \int \bar{\xi}(x)[\frac{\hat{p}^2}{2}]\xi(x) d^3x$$

$$- (1-\alpha) \int |\xi|(y) \frac{1}{|y|} d^3y - \alpha \int |\xi|^2(y)(1 + \frac{1}{|y|}) d^3y + 2\epsilon \tag{74}$$

Choosing ϵ sufficiently small we obtain $\langle \psi, H_\alpha \psi \rangle < \frac{1}{2}$.

We can now repeat this construction choosing a sequence of functions ξ_m with disjoint support (and therefore mutually orthogonal) which satisfy the conditions that we have imposed on the function ξ in (74). ♡

We study now the existence of point spectrum in $[-1/2, +\infty)$. We prove first

Theorem 7 *The operator H_α does not have point spectrum in* $[0, +\infty)$. ◇

Proof As in the case of the hydrogen atom we use the dilation group as in the preceding theorems to prove that if $\phi \in L^2(R^6)$ satisfies $H_\alpha \phi = E\phi$ then (virial theorem)

$$2\langle \phi, T\phi \rangle = \int |\phi(x_1, x_2)|^2 (\frac{1}{|x_1|} + \frac{1}{|x_2|} - \frac{1}{|x_1 - x_2|}) dx_1 dx_2 \tag{75}$$

Hence $E = -\langle \phi, T\phi \rangle < 0$. ♡

We treat next the problem of the existence of point spectrum in $[-1/2, 0]$. Since the continuum spectrum covers this interval the simultaneous presence of point spectrum can only be due to *super-selection rules* that do not permit writing the bound state as superposition of generalized eigenfunctions.

In the helium atom this super-selection rule is due do the *joint presence of rotational invariance, Fermi statistics and independence of the hamiltonian from the spin degrees of freedom*. Invariance under rotations implies conservation of the angular momentum (with respect to the origin; recall that we have made the approximation in which the nucleus is infinitely heavy and therefore the barycenter of the system coincides with the position of the nucleus that we have chosen as origin).

Define a *spatial symmetry* map P as

$$(P\ \phi_{h,k})(x_1, x_2) = \phi_{h,k}(x_2, x_1) \tag{76}$$

The map P acts trivially on the spin subspace; the corresponding operator commutes therefore with the part of the generator of rotations that acts on spin space. It commutes also with the hamiltonian and with the part of the generator of rotations that describes rotations in configuration space.Therefore the spectrum of H_α can be analyzed *separately* in the eigenspaces corresponding to the eigenvalue $l(l+1)$ of L^2 and to the eigenvalue $p = \pm 1$ of P.

It is convenient to introduce some notation. We call *subspace of natural parity* the subspace where $p = (-1)^l$ and *subspace of unnatural parity* the one in which

$p = -(-1)^l$. The elements on the non-natural space are the wave functions which are antisymmetric in the spatial variables and have even angular momentum and those which are symmetric wit respect to the space variables and have odd angular momentum.

From the definition the parity of a state of a single electron is $(-1)^l$ (l is the absolute value of the angular momentum). The parity of a product state is then $(-1)^{l_1+l_2}$. The eigenvalues of H_0 which correspond to the choice $n_1 = 0$ and $n_2 \geq 0$ (therefore $l_1 = 0$ or $l_2 = 0$) are natural parity states.

The lowest energy for a state of non-natural parity is obtained setting $l_1 = 1,\ l_2 = 1$ (therefore $n_1 = 2,\ n_2 = 2$). The corresponding wave functions are (the suffix i, j gives the structure in spin space)

$$\psi_{i,j}(x_1, x_2) = (x_i \wedge x_2) f(|x_1|, |x_2|)\delta_{i,j} \tag{77}$$

where the function f is symmetric in its variables. The state described by (77) has energy $-\frac{1}{4}$; every other state of non-natural parity has larger energy. This state is isolated from the continuum spectrum of the restriction of H_0 to the subspace of non-natural parity states and therefore belongs to the point spectrum of H_0. Remark that the continuum spectrum of the restriction of H_0 to the subspace of natural symmetry covers $[-\frac{1}{4}, \infty)$.

The reduction with respect to the parity holds also for H_α. Since the perturbation $\frac{1}{|x_1-x_2|}$ is H_0 -small and therefore the eigenvalues vary continuously with the parameter α (as far as they remain bound states) we deduce that for small values of α there exists (at least) an eigenstate of H_α which has non-natural parity and eigenvalue immersed in the continuum. Explicit computation shows that this happens for $\alpha < \frac{1}{2}$.

Therefore this state exists for $\alpha = \frac{1}{n}, n = 3, 4, \ldots$ (ionized Litium atom, ionized Berillium atom, …*but not for the helium atom* (which corresponds to $n = 2$).

We remark that strictly speaking this super-selection rule for H_0 is valid only if one neglects the motion of the nucleus (we have set $\frac{1}{M} = 0$ where M is the mass of the nucleus) and one neglects the interaction with an external electromagnetic field. If one considers the corresponding corrections one has now a *resonance* instead of a bound state.

Roughly speaking the function $(e^{itH_\alpha}\psi)(x)$ remains *essentially localized* for $0 \leq t \leq \tau$ where τ is "very large"; however for any $R > 0$ one has

$$lim_{t\to\infty} \int_{|x_1|^2+|x_2|^2\leq R^2} |(e^{itH_\alpha}\psi)(x)|^2 dx = 0 \tag{78}$$

One can prove that asymptotically in time one has

$$w - limit_{t\to\infty}[e^{itH_\alpha}\psi - \phi(t)] = 0$$

where $\phi(t)$ is the (antisymmetric) product of the wave function of a free electron and the wave function of the ground state of the hydrogen atom. This phenomenon is called *Auger effect*.

8 Two-Dimensional Hydrogen Atom

It is interesting to compare the case of the Hydrogen atom in three space dimension with that of the same problem in one and two dimensions.

In two dimensions the Coulomb potential is not Kato-small with respect to the Laplacian and therefore the Hamiltonian is defined on $D(\Delta) \cap D(\frac{1}{|x|})$. On this domain Δ may not be self-adjoint; we use Weyl's criterion. The initial hermitian operator is, setting $\frac{\hbar^2}{m} = 2$

$$H_0 = -\Delta - \frac{e}{|x|} \qquad D(H_0) = C_0^\infty(R^2/\{0\}) \tag{79}$$

Introducing polar coordinates one has, on the dense set of product functions

$$f(r)g(\theta), \qquad f \in C_0^\infty(0,\infty) \quad g \in C^2(S_1)$$

$$H_0 f(r)g(\theta) = [-\partial_r^2 + \frac{1}{r}\partial_r - \frac{e}{r}]f(r)g(\phi) - \frac{f(r)}{r^2}\partial_\phi^2 g(\phi) \tag{80}$$

Decomposing in eigenstates of the angular momentum and using the isometric map

$$U \ : \ L^2(0,\infty), rdr \to L^2(0,\infty), dr \quad U(\phi)(r) \equiv \sqrt{r}\ \phi(r) \tag{81}$$

one has, on the lth eigenspace of ∂_ϕ^2

$$h_0^l \equiv U h_0^l U_l^{-1} = -(\partial_r^2 + (\frac{1}{4} - l^2)\frac{1}{r^2} - \frac{e}{r} \tag{82}$$

The adjoint has the same action of h_0^l but with domain

$$D((h_0^l)^*) = \phi \in L^2(0,\infty),\ \phi,\ \phi' \in AC(0,\infty),\ h_0^l\phi \in L^2(0,\infty) \tag{83}$$

To verify self-adjointness we use Weyl's criterion. The equation

$$[-\partial_r^2 + (l^2 - \frac{1}{4})\frac{1}{r^2} - \frac{e}{r}]f(r) = if(r) \tag{84}$$

For each value of l this equation has two solutions in $(0,\infty)$ (given in terms of Whittaker functions).

Both solutions are square integrable at infinity but for $l > 0$ one of the solution is not square integrable at the origin. According to Weyl's criterium the operator $\hat{h}_l$ is essentially self-adjoint.

When $l = 0$ both solutions are square integrable, independently of the value of e. Therefore the operator H_0 has defect indices (1,1) and a one parameter family of self-adjoint extensions.

It is worth noting that this result is true independently of the charge e and reflects the fact that the Laplacian in R^2 with domain the functions in H^2 that vanish in a neighborhood of the origin has defect indices (1, 1) and the function in the defect spaces can be chosen to be invariant under rotation. The presence of the Coulomb potential does not alter this feature.

This is true also in three dimensions, but one can avoid considering extensions because the Coulomb potential is small with respect to the laplacian.

9 One-Dimensional Hydrogen Atom

The *one-dimensional* hydrogen atom is defined *formally* by the differential operator

$$-\frac{1}{2}\frac{d^2}{dx^2} - \frac{e}{|x|} \tag{85}$$

Also in this case to see the problem as a quantum hamiltonian problem we must find a dense domain in the Hilbert space $L^2(R)$ on which the formal differential operator (85) can be interpreted as a self-adjoint operator.

We start with the symmetric operator

$$H_0 \equiv -\frac{1}{2}\frac{d^2}{dx^2} - \frac{e}{|x|}, \quad D(H_0) = C_0^\infty(R - \{0\}) \tag{86}$$

The closure of H_0 has defect indices (2,2) as seen by solving (using Wittaker functions) the differential equation

$$-\frac{1}{2}\frac{d^2u(x)}{dx^2} - \frac{e}{|x|}u(x) = iu(x) \tag{87}$$

Independently of the value of the parameter e this equation has two solution in $L^2(R)$. It follows this symmetric operator has defect indices (2,2) and there exists a four parameter family H_A of self-adjoint extensions of $\bar{H}_0$ parametrized by a symmetric matrix

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{88}$$

$a, c \in R, b \in C$.

The corresponding self-adjoint extensions are

$$H_A = -\frac{1}{2}\frac{d^2}{dx^2} - \frac{e}{|x|} \quad D(H_A) = \{\phi,\ \phi' \in AC_{loc}(R-0),\ \frac{1}{2}\phi''(x) - \frac{e}{|x|} \in L^2(R)\}$$

$$c_\pm = A_{\pm,l} d_l \qquad c_\pm = lim_{x\to 0_\pm}\{2e\phi(x) ln(\pm|e|x \pm \phi'(x)\}, \quad d_\pm = lim_{x\pm 0}\phi(x) \tag{89}$$

We have indicated with $AC_{loc}(\Lambda)$ the set of complex-valued functions ϕ which are are absolutely continuous in every subset of Λ and the first and second derivatives are Lebesgue integrable almost everywhere. Notice that if $e \neq 0$, $\phi'(x)$ diverges when $x \to 0$.

The entries of the matrix A are allowed to take the value $+\infty$. For example Dirichlet b.c. correspond to $a = c = +\infty, b = 0$. We denote the correspondent hamiltonian by H_D. Remark that in the case $e = 0$ each H_A corresponds to a "point interaction" at the origin.

The parameter e enters only in the condition that $\frac{1}{2}\phi''(x) - \frac{e}{|x|}\phi(x) \in L^2$. The nucleus of the resolvent of H_D can be expressed explicitly [1] in terms of Gamma and Wittaker functions; the nuclei of the resolvent for any matrix A can then be found using Krein's formula. Since the defect indices are finite, all self-adjoint extensions have the same essential spectrum, i.e. $[0, \infty)$ and for the same reason they are all bounded below.

From the explicit for of the resolvent one can see that for any positive value of the charge e there are infinitely many bound states accumulating at zero. For $e < 0$ there may be one or two bound states depending on the matrix A. For $e > 0$ one can choose the matrix A in such a way that the lower bound of the spectrum of H_A is an arbitrary negative number.

An interesting quantity is the probability current density

$$j_\phi(x) = Im\phi^*(x)\phi'(x), \qquad \phi \in H_A \tag{90}$$

One can verify that the current $j(x)$ is continuous for all $x \in R$ and at the origin is given by

$$j_\phi(0) = eIm(lim_{x\to +}\phi^*(x)\phi(x)) \tag{91}$$

As a consequence the choice $b = 0$ (Dirichlet b.c.) implies $j_\phi(0) = 0$. If the parameter b is zero, for any value of the charge e the hamiltonian is the orthogonal sum of two Hamiltonians defined on the two half-lines $R^\pm$.

While the continuous spectrum is R^+ for all self-adjoint extensions, the (discrete) negative spectrum depends on the self-adjoint extension. For a given self-adjoint extension the negative spectrum coincides with the poles of the resolvent but in general it is difficult to find the explicit form of the resolvent. One notable exception is the Dirichlet b.c. since in this case one can consider separately the problem in R^+ and in R^-. Due to the symmetry of the problem each eigenvalue will be doubly

degenerate and the solutions come in pairs, one even and one odd. One can prove that in this case the eigenvalues are

$$E_n = -\frac{e^2 m}{2\hbar^2}\frac{1}{n^2}, \quad n = 1, 2, 3, \ldots \tag{92}$$

as in the three-dimensional case (but the multiplicity is different). This can be understood since if one describes the problem in three dimensions in polar coordinates and one takes into account only the eigenfunctions that are spherically symmetric one recovers the one dimensional problem in R^+ and after the map U the condition that the solution be L^2 at the origin is transformed into the condition that it vanishes at the origin.

10 Capacity

In dimension d = 1, 2, 3 the symmetric operator Δ (the Lapacian) defined on functions that vanish in a neighborhood of the origin is symmetric but not self-adjoint. Its defect spaces are connected with the existence of functions in R^d which are not twice differentiable, which are locally in L^1 and solve the equation

$$-\Delta\phi(x) = \delta(x)\phi(x) \tag{93}$$

In dimension one they span a two-dimensional space, and one can take the functions $\frac{x}{|x|} = signx$ and $xsign(x)$ as a basis. Together with the two regular solution of $\Delta\phi = 0$ i.e., $\phi(x) = c, \phi(x) = x$ they determine the defect space and classify the possible extensions.

In two and three dimension the space is one dimensional and one can take as basis the functions $log\ |x|$ and $\frac{1}{|x|}$ respectively. In this case the regular solution is $\phi(x) = c$.

The solutions of (33) are usually called Green's Functions and their existence is often is often referred to by saying that the origin has finite capacity. In three dimensions the coefficient C in $\frac{C}{|x|}$ is usually called *charge*.

Remark that (93) has a solution in a distributional sense for any dimension. The defect space can be regarded as the finite dimensional space of solutions of (93).

The presence of this defect spaces can be seen by perturbing $H_0 = -\Delta$ with a potential that is singular enough at the origin. We have seen that in dimension one a Coulomb singularity is sufficient to probe the defect space. In dimension two and three one must introduce the stronger singularity $-\frac{C}{x^2},\ C > 0$.

An easy computation shows that the condition on $C > 0$ in order to be in the limit circle case is in dimension two $C \geq \frac{1}{2}$ and in dimension three $C \geq \frac{3}{4}$. If this condition is satisfied the operator $-\Delta - \frac{C}{x^2}$ defined on $C_0^2(R^d - 0)$ is bounded below but not self-adjoint and there is a one parameter family of extensions.

References

1. de Olivera, C., & Verra, A. (2009). *Annals of Physics*, *324*, 251–266.
2. Pauli, W. (1926). *Zeitschrift fur Physik*, *36*, 336–363.
3. Reed, M., & Simon, B. D. (1992). *Methods of Mathematical Physics*. Vol. 4. New York: Springer.
4. Weidmann, K. (1987). Spectral theory of O.D.E. *Lecture Notes in Mathematics*, *1258*, Berlin: Springer Verlag.
5. Weyl, H. (1910). *Mathematische Annalen*, *68*, 220–269.

Lecture 19: Estimates of the Number of Bound States. The Feshbach Method

In this lecture we discuss the problem of determining, under suitable assumption on the potential $V(x)$, an upper or lower bound to the number of bound states of the operator $H = -\Delta + V$ acting in $L^2(R^d)$.

We will make always the assumption the potential vanishes at infinity fast enough so that the continuous spectrum of H covers $[0, +\infty)$. The aim will to estimate the number of eigenvalues that satisfy $E \leq 0$ as a function of some properties of V, e.g. its L^p norm for a given value of p.

We shall also give sufficient conditions on V for the existence of strictly negative eigenvalues and estimates of the number of negative eigenvalues of the operator $-\Delta + \lambda V(x), \quad x \in R^3$ when λ is very large. Setting $\lambda = \hbar^{-2}$ one sees that these estimates are somehow connected with the semiclassical limit.

1 Comparison Theorems

We begin with simple comparison theorems.

Proposition 1 *Let $A \geq 0$ and B be self-adjoint operators on the Hilbert space $\mathcal{H}$. Let $Q(A)$ be the domain of the quadratic form $q(A)$ associated to A and assume that $Q(A) \cap Q(B)$ is dense in $\mathcal{H}$ and that the negative part of B is infinitesimal with respect to A in the sense of quadratic forms. Suppose moreover that for any $\beta \geq 0$*

$$\sigma_{ess}(A + \beta B) = [0, +\infty) \tag{1}$$

Then every negative eigenvalue $\mu_n(\beta)$ of the operator $A + \beta B$ is a non increasing monotone function of β. ◇

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_19

Proof From the min-max principle

$$\mu_n(A+\beta_0 B) = max_{\phi_1,\dots\phi_{n-1}}\, min_{\psi\in Q(A)\cap Q(B),|\psi|=1,(\phi_i,\psi)=0} min\{0,(\psi,(A+\beta_0 B)\psi)\} \tag{2}$$

Since $A \geq 0$ from $(\psi,(A+\beta B)\psi) < 0$ and $\beta > 0$ one derives $(\psi, B\psi) < 0$. Therefore

$$\beta_1 < \beta_2 \rightarrow (\psi,(A+\beta_1 B)\psi) \leq (\psi,(A+\beta_2 B)\psi) \tag{3}$$

The thesis of Proposition 1 follows from (2) and (3). ♡

Example Let $V \in L^2 + (L^\infty)_\epsilon$. Recall that this notation means

$$V = V_1 + V_2, \quad V_1 \in L^2, \quad V_2 \in L^\infty, \quad |V_2|_\infty < \epsilon$$

Then $\sigma_{ess}(-\Delta + \beta V) = [0, +\infty)$. Therefore if $V(x)$ has a negative part the eigenvalues (if they exist) are not decreasing functions of β. ♣

An inequality that finds frequent applications is given in the following theorem; the easy proof is obtained through the min-max principle.

Theorem 2 (Raileigh-Ritz) *Let H be a self-adjoint operator bounded below on a Hilbert space $\mathcal{H}$. Denote with $\mu_n, n = 1, \dots N$ its eigenvalues that lie below the continuum spectrum, in increasing order. Let M_n, $n \leq N$ be a n-dimensional subspace of $\mathcal{H}$ contained in $D(H)$. Let P_n be the orthogonal projection on M_n.*

Let the operator $P_n H P_n$ be defined in the sense of quadratic forms and let $\hat{\mu}_n$ be the corresponding eigenvalues. Then for $m = 1, \dots n$ one has $\hat{\mu}_m \geq \mu_m$. ◇

When one considers an increasing sequence of subspaces which tend to cover $\mathcal{H}$ one has convergence *at least of the smallest eigenvalue*. This is the content of the following theorem.

Theorem 3 *Let $\{\eta_k\} \in D(H)$ be a complete ortho-normal basis of $\mathcal{H}$. Let $H \geq E_0 I$ and let E_0 be an eigenvalue of H. Suppose that*

$$liminf_{\psi\in\mathcal{H},\|\psi\|=1 N\rightarrow\infty}(P_N\psi, H P_N\psi) = E_0$$

where P_N is the orthogonal projection on the subspace spanned by $\eta_1, \dots \eta_N$. Let $\hat{\mu}_0^N$ be the smallest eigenvalue of $P_N H P_N$. Then

$$lim_{N\rightarrow\infty}\hat{\mu}_0^N = E_0 \tag{4}$$

◇

Proof According to Proposition 1 the sequence $\hat{\mu}_0^N$ is increasing and bounded above by zero. Denote by $\hat{\mu}_0$ its limit. Theorem 2 implies $\hat{\mu}_0 \geq \mu_0$. Suppose $\hat{\mu}_0 > \mu_0$. Considering a sequence ψ_k, $\|\psi\| = 1$ that is minimizing for $(\psi, H\psi)$ and choosing the sequence $P_k\psi_k$ one obtains a contradiction. Therefore $\hat{\mu}_0 = \mu_0$. ♡

This estimates of the lowest eigenvalue which one obtains by application of the min-max principle are always *estimates from above*. A useful estimate *from below* is contained in the following theorem

Theorem 4 (Temple) *Let H be self-adjoint with pure point spectrum and let $\mu_0 = inf\sigma(H)$ be an eigenvalue of H isolated and simple. Suppose that $\psi \in D(H)$, $|\psi| = 1$ satisfies $(\psi, H\psi) < \mu_1$ where μ_1 is the second eigenvalue of H_0. Then*

$$\mu_0 \geq (\psi, H\psi) - \frac{(\psi, H^2\psi) - (\psi, H\psi)^2}{\mu_1 - (\psi, H\psi)} \tag{5}$$

◇

Proof By assumption

$$(H - \mu_0 I)(H - \mu_1 I) \geq 0 \tag{6}$$

because (5) holds on the eigenvector to the eigenvalue μ_0 and is also valid on any vector in the spectral projection of H on $[\mu_1, +\infty)$. Therefore it holds on $\mathcal{H}$.

Taking the expectation value of the left hand side of (6) in a generic vector ψ, $\|\psi\| = 1$ one has

$$\mu_0\mu_1 - \mu_0(\psi, H\psi) \geq \mu_1(\psi, H\psi) - (\psi, H^2\psi) \tag{7}$$

If the vector ψ satisfies the condition of the theorem, adding and subtracting $(\psi, H\psi)^2$ on the right hand side and dividing by $\mu_1 - (\psi, H\psi)$ one obtains (5). ♡

Remark that to apply Temple's theorem is is sufficient to take as μ_1 the second eigenvalue of a operator K such that $K < H$ and choose as ψ the eigenvector of H to the first eigenvalue.

Example We give a simple example of the use of Temple's theorem. Let

$$H = A + \frac{1}{|x_1 - x_2|}, \quad A = -\Delta_1 - \Delta_2 - \frac{2}{|x_1|} - \frac{2}{|x_1|} \tag{8}$$

One has $H \geq A$ and therefore $\mu_1(H) \geq \mu_1(A) = -\frac{5}{4}$ ♣

A useful theorem to estimate the number of a bound states of a Schrödinger operator is the following.

Theorem 5 *For $x \in R^3$, suppose $\sigma_{ess}(-\Delta + V) = [0, +\infty)$ and that there exists positive constants R_0 ϵ such that for $|x| > R_0$ one has $V(x) < -\frac{a}{|x|^{2-\epsilon}}$, a , $\epsilon > 0$.*

Then $\Delta + V$ has an infinite number of negative eigenvalues. ◇

Proof The proof follows the lines of the proof that there are infinitely many eigenvalues of the hydrogen atom.

We must show that $\mu_n < 0 \ \ \forall n \in N$. Choose $\psi \in C^\infty, \ |\psi|_2 = 1$ with support contained in $1 < |x| < 2$. Define

$$\psi_R(x) \equiv R^{\frac{3}{2}} \psi(R^{-1}x)$$

If $|x| > R_0$ one has

$$\begin{aligned}(\psi_R, H\psi_R) \leq (\psi_R, -\Delta\psi_R) - a(\psi_R, |x|^{-2+\epsilon}\psi_R) = \\ = R^{-2}(\psi, -\Delta\psi) - aR^{-2+\epsilon}(\psi, |x|^{-2+\epsilon}\psi)\end{aligned} \tag{9}$$

But

$$1 = (\psi, \psi) \leq (\psi, |x|^{-2+\epsilon}\psi) \leq (\phi, 2^{-2+\epsilon}\phi) = 2^{-2+\epsilon} \tag{10}$$

Since $\epsilon > 0$ there exists R_1 such that the right hand side of (9) is negative for $|x| > R_1$. The vectors $\phi_n \equiv \psi_{2^n R}$ are orthogonal for different value of n and span therefore a n-dimensional space. And $\mu_n(H) \leq sup_{1 \leq m \leq n}(\phi_m, H\phi_m) < 0$. ♡

If the negative part of the potential decays *more rapidly at infinity* the number of negative eigenvalues can be finite.

Theorem 6 *Let* $V(x) \in \mathcal{R} + L^\infty_\epsilon, x \in R^3$ *and for* $|x| > R_0$ *let* $V(x) \geq -\frac{1}{4}b|x|^{-2}, b < 1$.

Then $-\Delta + V$ *has at most a finite number of negative eigenvalues.* ◇

Proof Set $W \equiv V + \frac{1}{4}b|x|^{-2}$. Since $(\phi, \Delta\phi) > \frac{1}{4}(\phi, |x|^{-2}\phi)$ one has

$$(\phi, -\Delta + V\phi) \geq -(1-b)(\phi, \Delta\phi) + (\phi, \tilde{W}\phi), \qquad \tilde{W}(x) = inf\{0, W(x)\} \tag{11}$$

Therefore

$$\mu_n(-\Delta + V) \geq (1-b)\mu_n(-\Delta + \frac{\tilde{W}}{1-b}) \tag{12}$$

If $V \in \mathcal{R} + (L^\infty)_\epsilon$ then $W \in \mathcal{R}$ (Rollnik class). It suffices therefore to prove that the operator $-\Delta + V(x) \ \ x \in R^3, V(x) \in \mathcal{R}$ has a finite number of distinct eigenvalues and their multiplicity is finite. The thesis of Theorem 6 follows then from the important theorem of Birman-Schwinger that we now state and prove [14]. ♡

Theorem 7 (Birman-Schwinger, weak form) *Let* $d = 3$ *and* $V \in \mathcal{R}$. *Then* $N_-(V)$ *(the number of isolated negative eigenvalues counted with their multiplicity) satisfies the bound*

$$N_-(V) \leq \frac{1}{(4\pi)^2} \int \frac{|V(x)||V(y)|}{|x-y|^2} d^3x d^3y < \infty \tag{13}$$

◇

We remark that it suffices to prove the theorem for $V \leq 0$ and assume $V \in C_0^\infty$; one obtains by a limiting procedure the proof in case V is more singular but belongs to $\mathcal{R}$.

Proof of Theorem 7 Let $N_E(V)$ the dimension of the range of $P_{-\infty,E}$ (the spectral projection on the indicated interval). Define $\mu_n(\lambda) \equiv \mu_n(-\Delta + \lambda V)$ ($N_E(V)$ is the number of indices n such that $\mu_n(1) < E$). Since μ_n is monotone and continuous in λ (V is infinitesimal with respect to $-\Delta$) and $\mu_n(0) = 0$ it follows $\mu_n(1) < E$ iff there exists $\lambda \in (0, 1)$ for which $\mu_n(\lambda) = E$. Therefore

$$N_E(V) \leq \sum_{\lambda:\mu_k(\lambda)=E,\ k=1,\ldots N_E(V)} \lambda^{-2} \leq \sum_{\lambda:\mu_k(\lambda)=E} \lambda^{-2} \tag{14}$$

Set $H_0 = -\Delta$. Then

$$(H_0 + \lambda V - E)\psi = 0 \Rightarrow \lambda\sqrt{|V|}(H_0 - E)^{-1}\sqrt{|V|}\sqrt{|V|}\psi = \sqrt{|V|}\psi \tag{15}$$

which implies

$$\lambda \int \frac{\sqrt{|V(x)}e^{-\sqrt{-E}||x-y|}\sqrt{|V(y)|}}{4\pi|x-y|}\phi(y)dy = \phi(x), \qquad \phi \equiv \sqrt{|V|}\psi \tag{16}$$

Since $V \in \mathcal{R}$ its integral kernel K is of Hilbert-Schmidt type and

$$\sum_k \lambda_k^{-2} = TrK^*K = \frac{1}{(4\pi)^2}\int e^{-\sqrt{-E}||x-y|}\frac{\sqrt{|V(x)|}\sqrt{|V(y)|}}{|x-y|^2}dxdy \tag{17}$$

The proof of the weak form of the Birman-Schwinger theorem (and therefore of Theorem 6) is now achieved by noticing that $N_-(V) = lim_{E\to 0}N_E(V)$. ♡

We remark that from the weak form of the Birman-Schwinger theorem one concludes that *if the space dimension is at least three* the operator $-\Lambda + V$ *has no negative eigenvalue* if $||V||_{\mathcal{R}} < 1$.

Remark that if the space dimension is one or two one has instead:

Theorem 8 *Let* $V \in C_0^\infty$, $V \leq 0$, $V \neq 0$, $d = 1, 2$.

Then for every $\beta > 0$ *the operator* $-\Delta + \beta V$ *has at least one negative eigenvalue.* ◇

Proof From the proof of the weak form of Birman-Schwinger theorem it it sufficient to prove that for each $\beta > 0$ there exists k such that the operator $\sqrt{-V}(-\Delta + k^2)^{-1}\sqrt{-V}$ has an eigenvalue greater than β.

In dimension one and two this operator is of Hilbert-Schmidt class and therefore the maximum eigenvalue coincides with the Hilbert-Schmidt norm. It is then sufficient to prove

$$lim_{k^2\to 0}||\sqrt{-V}(-\Delta+k^2)^{-1}\sqrt{-V}||_{H.S}=\infty \tag{18}$$

and for this it is sufficient to prove that there exists $\eta\in L^2$ such

$$lim_{k^2\to 0}(\eta,\sqrt{-V}(-\Delta+k^2)^{-1}\sqrt{-V}\eta)=\infty \tag{19}$$

Let $V<0$ and choose η such that $\int\sqrt{-V(x)}\eta(x)dx>0$.

Performing Fourier transform the left hand side in (18) is

$$\int|\hat{\phi}(p)|^2\frac{1}{p^2+k^2}dp,\qquad \phi(x)\equiv\eta\sqrt{-V}(x) \tag{20}$$

By construction $|\hat{\phi}(0)|>0$. Therefore $lim_{k^2\to 0}\int|\hat{\phi}(p)|^2\frac{1}{p^2+k^2}dp=+\infty$. ♡

Theorem 8 admits the following generalization.

Proposition 9 *Let $V\in L^d(R^d)+(L^\infty(R^d))_\epsilon$ $d=1,2$.*

Then the operator $H=\Delta+V$ has at least one negative eigenvalue if at least one of the following inequalities is satisfied

(i) $V\le 0,\qquad V\neq 0$
(ii) $\int|V(x)|dx<+\infty,\qquad \int V(x)dx<0$
(iii) $\int V_+(x)dx<\infty,\qquad \int V_-(x)=+\infty$

◇

Proof Adding a suitable constant one reduce cases (i) and (iii) to case (ii). We prove case (ii).

From the assumptions made the essential spectrum in $[0,+\infty)$. It is therefore sufficient to prove that there exists $\psi\in L^2$ such that $(\psi,H\psi)<0$. We shall separately the cases $d=1$ and $d=2$.

$d=1$. Set $\psi_a=e^{-a|x|}$, $a>0$. Then

$$\int|\nabla\psi_a|^2dx=a,\qquad lim_{a\to 0}\int V(x)|\psi(x)_a dx=\int V(x)dx<0$$

Therefore $(\psi_a,H\psi_a)<0$ for a sufficiently small.

The case d = 2 is reduced to the case $d=1$ by considering only functions that are invariant under rotation around the origin and making use of polar coordinates. Essential for this procedure is the fact that $\frac{1}{|x|}\in L^2_{loc}$.

This is the reason why Proposition 9 *does not hold if* $d\ge 3$. ♡

We remark that the difference between the case $d=1,\ 2$ and the case $d\le 3$ can be as connected to the following: brownian motion is *recursive* for $d=1,2$ (the paths return with probability one to an arbitrarily small neighborhood of the initial point) while if $d\ge 3$ it is *dispersive* (the paths with probability one exit asymptotically form any compact in R^d).

We will discuss in the second part of these Lectures the relation between brownian motion and the solutions of the Schrödinger equation: the laplacian (a negative operator) is the generator of brownian motion as a stochastic process. Bound states of the operator $-\Delta + V$ correspond to invariant measures of the resulting stochastic process. We shall prove that adding to the free hamiltonian a negative potential V is equivalent to adding an attractive vector field ξ_V to the generator of brownian motion [10].

In this context the probabilistic analogue of Proposition 9 is the following: if brownian motion is recursive then the addition of an attractive vector field leads to a Stochastic process that has at least one invariant measure. On the contrary if the brownian motion is dispersive in order to have an invariant measure one must add a sufficiently strong attractive vector field (this can be proved also within the theory of stochastic processes).

The counterpart of these statements, valid for stochastic differential equations, are in the theory of the Schrödinger equation the strong version of Birman theorem and Hardy's inequality.

To give a further application of the method we have used to estimate $N(V)$ we prove

Theorem 10 *If $E < 0$ then (denoting by $-V_-$ the negative part of V)*

$$N_E(V) \leq 2 \int_0^\infty Tr\, V_- [e^{-tH_0} - e^{-t(H_0+V_-)}] e^{Et} dt \tag{21}$$

◇

Proof By min-max $N_E(V) \leq N_E(V_-)$. If $(H_0 - \beta V_-)\phi = 0$ then $\psi \equiv V_- \phi$ satisfies

$$\sqrt{V_-}(H_0 + V_- + k^2)^{-1} \sqrt{V_-}\psi = (1+\lambda)^{-1}\psi, \qquad k^2 = -E$$

and therefore

$$W\psi = [\lambda^{-1} - (1+\lambda^{-1}]\psi, \qquad W \equiv \sqrt{V_-}[(H_0 + k^2)^{-1} - (H_0 + V_- + k^2)^{-1}]\sqrt{V_-} \tag{22}$$

It follows that $N_E(V_-)$ is less than the number of eigenvalues of the operator W which are less than $1/2$ (when $0 < \lambda < 1$ one has $\lambda^{-1} - (1+\lambda)^{-1} \geq 1/2$). One has

$$Tr\, W = \sum_i \lambda_i \geq \sum_{i\,:\,\lambda_i > 1/2} \lambda_i \geq \frac{1}{2} \sum n^0(|\lambda_i) \geq 1/2 \tag{23}$$

and we conclude that the number of eigenvalues of the operator W which are larger than $1/2$ is less than $2Tr\, W$.

Performing the t integration in (21) concludes the proof. ♡

We remark the choice of the operator W in (22) was made in order to connect the estimate the number of bound states to estimates of the difference of the resolvents

of $H_0 + V_-$ and of H_0 i.e. estimates of Birman-Schwinger type. Other estimates can be obtained choosing functions that take value larger or equal to one on the positive integers and are somewhat related to the difference of these resolvents.

2 Estimates Depending on Banach Norms

We now analyze estimates on the number of negative eigenvalues which depend on some Banach norm of V (e.g. the norm as element of L^p). We start analyzing the sum of some powers of the eigenvalues. From these estimates one can then derive estimates on their number.

Notice that in many problems of physical interest (e.g. in the study of the stability of matter and more generally in the study of a system of N identical fermions) the quantity of interest is the sum of the first N eigenvalues. Introducing the function

$$S_\gamma(V) \equiv \sum_{E_j \leq 0} |E_j|^\gamma, \quad S_0(V) = N^o\{E_j \ : \ E_j \leq 0\} \tag{24}$$

one can prove e.g.

Theorem 11 ([14]) *For every $\gamma \geq 0$ and for any operator $-\Delta + V$ on $L^2(R^d)$*

$$lim_{\lambda\to\infty} \frac{S_\gamma(V)}{\lambda^{d/2}} = C_{d,\gamma} \int dx (V_-(x))^{\frac{d}{2}+\gamma}, \qquad C_{d,\gamma} = \int_{|y|<1} |y|^{d-1}(1-y^2)^\gamma dy \tag{25}$$

for potentials such that the integral is defined. ◇

For the proof of this theorem we refer to [14] vol IV, cap. XIII. We return later to this theorem.

Estimates for $S_\gamma(V)$ can be obtained by similar estimates for $N_E(V)$, the number of bound states which have energy less than E. The function $N_E(V)$ sometimes denoted *counting function* is the spectral dimension of the projection operator $P_{(-\infty,E]}$ associated to $H = -\Delta + V$. The function $N_E(V)$ is monotone increasing and

$$S_\gamma(V) = \sum_{E_j<0} |E_j|^\gamma = \gamma \int_0^\infty |E|^{\gamma-1} N_E(V) d|E| \tag{26}$$

We point out two simple inequalities satisfied by N_E

(a)

$$V(x) \leq W(x) \ \forall x \ \Rightarrow N_E(V) \leq N_E(W), \qquad E \leq 0 \tag{27}$$

(b)

$$\forall \alpha \in (0,1], \ E \leq 0 \quad N_E(V) \leq N_{\alpha E}[(1-\alpha)E - V] \tag{28}$$

The proof of (a) is trivial.

For the proof of (b) notice that $(\psi, [-\Delta + V]\psi) \leq E(\psi, \psi)$ implies

$$(\psi, [-\Delta + V - (1-\alpha)E]\psi) \leq \alpha E|\psi|^2 \tag{29}$$

Since $E \leq 0$ one derives $V-(1-\alpha)E \geq V$. Inequality (b) follows from the min-max principle since the subspace $H \leq E$ contains the subspace in which, as quadratic form $-\Delta + V - (1-\alpha)E$ is smaller than αE.

An important instrument in the study of the number of bound states is another Birman-Schwinger theorem that states the equality between $N_E(V)$ and the number of eigenvalues larger than one of a suitable integral operator that depends on V.

Define the *Birman-Schwinger operator* $K_E(V)$

$$K_E \equiv \sqrt{|V|}(-\Delta + E)^{-1}\sqrt{|V|} \tag{30}$$

Theorem 12 (Birman-Schwinger, strong version) *If* $V \leq 0, \quad E < 0$ *then* $N_E(V)$ *coincides with the number of eigenvalues of* K_E *in* $(1, \infty)$. ◇

Proof From monotonicity and continuity of the eigenvalues $N_E(V)$ coincides with the number of $\lambda \in (0, 1)$ such that $\Delta + \lambda V$ has E as eigenvalue, and therefore with the number of $\lambda \in (0, 1)$ such that λ^{-1} is an eigenvalue of K_E. ♡

If the Birman-Schwinger operator K_E is of Hilbert-Schmidt class (as is the case in most cases of physical interest) the number of its eigenvalues in $[1, \infty)$ is bounded by its Hilbert-Schmidt norm. Indeed if we denote by $\{\lambda_i\}$ its eigenvalues (counting multiplicity) one has

$$Tr(K_E^* K_E) = \sum_i |\lambda_i|^2 \geq \sum_{i, \lambda_i \geq 1} |\lambda_i|^2 \geq \sum_{i, \lambda_i \geq 1} = N^o(\lambda_i \ : \ \lambda_i > 1) \tag{31}$$

From the Birman-Schwinger theorem

$$N_E(V) \leq Tr\ K_E^* K_E \tag{32}$$

Keeping into account (30) and taking Fourier transform equation (32) takes the form

$$N_E(V) \leq \frac{1}{(4\pi)^3} \int_{R^3} \frac{|\hat{V}(p)|^2}{|p|} arctg \frac{|p|}{2\sqrt{|E|}} dp \tag{33}$$

In particular in three dimensions $\frac{1}{(4\pi)^3} \int_{R^3} \frac{|\hat{V}(p)|^2}{|p|} dp < 1 \Rightarrow N_0(V) = 0$.

Refinements and extensions of these results can be obtained using Hoelder's and Sobolev's inequalities. For example, setting

$$K_{d,r} \equiv \omega_{d-1}^{-\frac{1}{2r}} (d-2)^{[\frac{1}{2r-1}\frac{r-1}{r}]} \frac{r-1}{2r} [\frac{\Gamma(2r)}{\Gamma(r+1)\Gamma(r)}]^{\frac{1}{2r}} \tag{34}$$

where ω is the area of the unit sphere in R^d one can prove the following theorem (see [14] vol IV).

Theorem 13 *If $r \geq 1$ then*

(a) If $r \geq \frac{d}{2}$ then

$$inf_{x\in R^d} \int |x-y|^{2r-d}|V_-(y)|^2 dy < \frac{1}{K_{d,r}^{2r}} \Rightarrow N_0(V) = 0 \tag{35}$$

(b) If $1 \leq r < \frac{d}{2}$ and $V(x) = v(|x|)$ then

$$\int |x|^{2r-d} V_-(|x|)dx < \frac{1}{K_{d,r}^{2r}} \Rightarrow N_0(V) = 0 \tag{36}$$

$\diamond$

We have sen that *in one and two dimensions* under rather general assumptions there is *at least one* bound state.

It follows from (34) that if $d = 3$ for every $\alpha \in [0, 1)$

$$N_E(V) \leq Tr K_E^* K_E = \frac{1}{4\pi^2} \int |V_\alpha(x)||V_\alpha(y)||x-y|^{-2} e^{-2\sqrt{\alpha(E)}|x-y|} dxdy \tag{37}$$

where $V_\alpha(E) = V(x) - (1-\alpha)E)$.

Therefore, using Young's inequality for $p = r = 2,\ \ q = 1$ (and then $p' = r' = 2,\ \ q' = \infty$) and performing the integration in y

$$N_E(V) \leq \frac{1}{8\pi\alpha|E|} \int |V_\alpha(x)|^2 dx \tag{38}$$

To estimate the number of negative eigenvalues define

$$S_1(V) \equiv \sum_{E_k<0} |E_k| = E_1 N_{E_1} + \sum_{n=1}^{\infty} (E_{n+1} - E_n) N_{E_{n+1}}$$

One can bound this sum by $\int N_E(V)dE$ (recall that N_E is the number of eigenvalues $\leq E$ and therefore N_E is a non-decreasing function).

Using this estimate and taking the infimum with respect to α (which is obtained for $\alpha = 1/2$) one has

$$S_1(V) \leq \frac{1}{4\pi} \frac{\Gamma(3)\Gamma(1/2)}{\Gamma(3/2)} \int |V_-(x)|^{5/2} dx \tag{39}$$

In general in d dimension this procedure leads to

$$\sum_{E_k<0} |E_k| \leq L_{d,\alpha} \int_{R^d} |V_-(x)|^{1+d/2} \tag{40}$$

for a suitable constant L_d.

Theorem 14 *Let* $2 < p < 2^* \quad d \geq 3$. *Let* $r \equiv \frac{p}{p-2}, \quad 0 < \gamma = \frac{p}{p-2} - \frac{d}{2}$. *Then* $N_E(V) = 0$ *if for* $E < 0$ *and one has*

$$g_p^{-r} \int_{R^d} (V_(x))^{\frac{d}{2}+\gamma} dx < |E|^\gamma \quad g_p = inf_u \frac{\|\nabla u\|_2^2 + \|u\|_2^2}{\|u\|_p^2} \tag{41}$$

◇

Proof By Hoelder inequality $(\psi, H\psi) \geq \|\nabla\psi\|_2^2 - \|V_-\|_r \, \|\psi\|_p^2$. Define

$$f_p \equiv inf_u F_p(u), \quad F_p(u) = \|\nabla u\|_2^\alpha \|u\|_2^{2-\alpha}, \quad \alpha = \frac{d}{2r} \tag{42}$$

Denote by $\hat{u}$ its minimum which is reached by convexity and compactness. Then

$$(\psi, H\psi) \geq \|\nabla\psi\|_2^2 - f_p^{-1} \|V_-\|_2 |\nabla\psi\|_2^{2\alpha} \|\psi\|_2^{2-2\alpha} \tag{43}$$

By rescaling

$$g_p(\hat{u}) = \alpha^{-\alpha}(1-\alpha)^{\alpha-1} f_p(\hat{u}) \tag{44}$$

From (43) and (44) one derives

$$H \geq -g_p^{-\frac{r}{\gamma}} \|V_-\|_2^{\frac{r}{\gamma}} \tag{45}$$

A comparison with (37) achieves the proof of Theorem 7. ♡

3 Estimates for Central Potentials

If the potential is central, more refined estimates are obtained noticing that in this case the eigenvalue problem can be formulated as a problem for an ordinary differential equation.

We will consider briefly the case of dimension 3. The hamiltonian commutes with angular momentum. In spherical coordinates a common basis of eigen-states has the form

$$\phi_l(|k|^2, r) Y_{l,m}(\omega), \ \omega \in S^3$$

where $(Y_{l,m}\omega)$ are the spherical harmonics. The functions ϕ_l are in $L^2(R^+, r^2dr)$.

Traditionally one sets $\phi_l(r) \equiv ru_l(r)$ with $u_l \in L^2(R^+, dr)$ paying attention to the conditions at the origin and at infinity. The eigenvalue equation reduces to the ordinary differential equation

$$u_l''(|k|^2, r) = [|k|^2 + V(r) + \frac{l(l+1)}{r^2}]u_l(|k|^2, r) \tag{46}$$

with the supplementary condition that the solution vanish at the origin and at infinity. The determination of the eigenvalues is then reduced to a Sturm-Liouville problem. The bounds are often obtained by variational techniques and are almost always bound from above.

Denoting by N_l the negative eigenvalues with angular momentum l, and by V^- the negative part of the potential it is possible to prove the bound [3]

$$N_l < \frac{1}{2l+1}\int_0^\infty [-V^-(r)rdr \tag{47}$$

Bounds have been obtained by Martin [12]

$$N_l < [\int_0^\infty V_{l,eff}^{(-)}(r)r^2dr \int_0^\infty V_{l,eff}^{(-)}(r)dr]^{\frac{1}{2}} \tag{48}$$

where $V_{l,eff}(r) = V(r) + \frac{l(l+1)}{r^2})$.

Other were given by Glaser et al. [8, 9]

$$\forall p > 1 \quad N_l < (2l+1)^{1-2p}C_p[\int_0^\infty [-r^2V^-(r)]^p]\frac{dr}{r}]^{\frac{1}{p}} \tag{49}$$

Of course better estimates and also bounds from below can be obtained if one makes specific requirements on the potentials.

4 Semiclassical Estimates

We discuss now *asymptotic estimates* for the distribution of the negative eigenvalues of the Schrödinger operator when the negative part of the potential *is very large*.

These estimates are called *semiclassical* because the number of negative eigenvalues of the operator $H = -\hbar^2\Delta + V$ coincides with the number of negative eigenvalues of $H_\hbar = -\Delta + \hbar^{-2}V$ and the strictly negative part $\hbar^{-1}V_-$ is very large in absolute value when $\hbar$ is very small.

To have a first rough estimate of the number of bound states it is natural to consider an auxiliary problem obtained by assuming that the support Ω_- of the negative part of V is compact. One covers then Ω_- with cubic cells of linear size $\hbar$ and consider

the modification of the operator obtained by adding Dirichlet boundary condition at the boundary of the cells. Imposing Dirichlet boundary conditions at $\partial\Omega_-$ decouples the region Ω_- and increases the lower bound of the spectrum.

Call $(-\hbar^{-2}\Delta + V_-)^D_\Omega$ the restriction to Ω with Dirichlet boundary conditions. This operator has pure point spectrum and fewer negative eigenvalues then $-\hbar^{-2}\Delta + V_-$ (Dirichlet boundary conditions increase the quadratic form). Imposing Dirichlet boundary conditions at the boundary of each cell decrease further the number of negative eigenvalues. But now an estimate of the error is possible by comparison with the case in which one imposes Neumann boundary conditions at the boundary of the cells.

The latter operation increases the number of negative eigenvalues so that the number of negative eigenvalues of $(-\hbar^{-2}\Delta + V_-)^D_\Omega$ can be estimated from above and from below. In each cell the number of negative eigenvalues is $O(|V|)$ and if the cell has size $\hbar$ the number of negative eigenvalues in each cell is at most one due to the uncertainity principle. The number of cells is $O(\hbar^{-d})$ and then the number of negative bound states of the operator $H_\hbar$ is $O(\hbar^{-d})$.

One can estimate the error made in introducing Dirichlet boundary conditions at $\partial\Omega$ by comparing the result with that one would obtain introducing Neumann b.c. The result is that the error is small relative to $\hbar^{-1}$.

One can easily see that this estimate of the number of negative eigenvalues which is done for $(-\hbar^{-2}\Delta + V_-)^D$ is not changed if one considers $-\hbar^{-2}\Delta + V_-$ (the number of cubes at the boundary is much small then the number of cubes in Ω).

From this *heuristic* analysis one can conclude that if V is sufficiently regular in the semiclassical limit the number of negative eigenvalues is roughly given by the number of hypercubes of side $\hbar$ that can be placed in the region where V is negative. If $\hbar$ is sufficiently small this number is

$$N_{class}(V) = (2\pi)^{-d} \int\int_{H(x,p)\leq 0} dx\, dp = (2\pi)^{-d} B_d \int_{R^d} |V_-(x)| dx \qquad (50)$$

where $V_-(x) \equiv sup(-\{V(x), 0\}$ is the negative part of V and B_d is the volume of the unit cube in R^d. One has $B_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(1+\frac{d}{2})}$ with $\Gamma(t) = \int_0^t s^{\alpha-1}e^{-s}ds$.

Denoting with $N^0_\hbar(V)$ the number of negative bound states (the trace of the spectral projection of $H = -\hbar^2\Delta + V$ on $-\infty, 0]$) one expects therefore that the following relation holds for $\hbar$ sufficiently small: there exist $\delta > 0$ such that

$$\frac{N^0_\hbar(V)}{\hbar^{\frac{d}{2}}} - (\hbar)^{-d} \int_{R^d} |V_(x)|^{\frac{d}{2}} dx \leq \hbar^\delta \qquad (51)$$

To estimate the asymptotic number of negative eigenvalues is known as *Weyl problem* [16]. For this problem useful references are [6, 14] vol IV.

The estimate (51) was obtained by Martin [12, 13] for potentials which are negative, continuous and compactly supported. It was extended in [1] to much more general cases

In general let $S_\gamma(V) = \sum_{j:E_j<0} |E_j|^\gamma$; we will look for estimates of the form

$$S_\gamma(V) \leq C_{\gamma,d} \frac{1}{(2\pi)^d} \int_{H(q,p)\leq 0} |H_{class.}(q,\, p)|^\gamma dq\ dp \tag{52}$$

Notice that $S_0(V) = N^0(V)$.

Estimates in terms of phase space volume have been developed by Lieb [11] using the representation in terms of Wiener integral of the kernel of the Birman-Schwinger operator.

Denote by $d\mu_{x,y;t}$ Wiener measure on continuous paths $\omega(t)$, $\omega(0) = x$ and $\omega(t) = y$. Recall that

$$\int d\mu_{x,y;t}(\omega) = (\frac{1}{4\pi t})^{\frac{d}{2}} e^{-\frac{|x-y|^2}{4t}} \tag{53}$$

Let $V \leq 0\ \ V \in L^p(R^d) + L^q(R^d)$ where $p < q < \infty$ and $p = \frac{d}{2}$ if $d \geq 3$, $p > 1$ if $d = 2$ and $p = 1$ if $d = 1$.

Let f be a non-negative function, lower semicontinuous with

$$f(0) = 0 \qquad lim_{r\to\infty} x^r f(x) = 0 \tag{54}$$

for some $r < \infty$. Define F_f as $F_f(x) = \int_0^\infty f(xy)\frac{e^{-y}}{y} dy$. Then

$$TrF_f(K_E) = \int_0^\infty \frac{e^{-|E|t}}{t} dt \int_{R^d} d\mu_{x,x;t} f(\int_0^t V(\omega(s))ds) \tag{55}$$

(K_E is the Birman-Schwinger operator defined in defined in (30)).

If f is convex we can make use of Jensen's inequality

$$f[\int_0^t V(\omega(s))ds] \leq \frac{1}{t} \int_0^t f(sV(\omega(s)))ds \tag{56}$$

and obtain, exchanging order of integration (allowed by Fubini's theorem)

$$TrF_f(K_E) \leq \int_0^\infty t^{-1} e^{-|E|t} dt d\mu_{(0,0;t)} \int_{R^d} f(V(\omega(s) + x))dx \tag{57}$$

By the invariance of Lebesgue measure under translations we can omit the dependence on ω and perform the integration over Wiener measure. One obtains the *Lieb's trace formula*.

Theorem 15 (Lieb's trace formula)

$$Tr F_f(K_E) \leq (4\pi)^{-\frac{d}{2}} \int_0^\infty dt\, t^{-1-\frac{d}{2}} e^{-|E|t} \int f(tV(x))dx \tag{58}$$

◇

Different choices of the function f in the class Γ of *convex, non-negative lower semicontinuous functions* give different information on

$$S_\gamma(V) = \sum_{E_k \leq 0} |E_k|^\gamma \tag{59}$$

and therefore on the distribution of the negative eigenvalues of H. In particular, setting $E = 0$ and for each value of x changing variables $tV(x) \to \tau$ one has

$$S_0(V) \equiv N_0(V) \leq \frac{\int_0^\infty s^{-\frac{d}{2}-1} f(s)ds}{\int_0^\infty s^{-1} e^{-s} f(s)ds} (4\pi)^{-\frac{d}{2}} \int_R^d |V_-(x)|^{\frac{d}{2}} dx \tag{60}$$

The best estimate is obtained taking the infimum over all functions f in the class Γ. A lower bound is obtained by minimizing over the functions f_b that satisfy $f(s) = 0, s \leq b$, $f(s) = s - b$ for $s > b$. In this way one obtains, if $d \geq 3$, the two bounds

$$N(V) \leq C_d \int_{R^d} |V_-(x)|^{\frac{d}{2}} dx \tag{61}$$

(Cwikel-Lieb-Rosemblum inequality).

$$\sum_{i:E_i<0} |E_i|^\gamma \leq C_{d,\gamma} \int_{R^d} |V_-(x)|^{\frac{d}{2}+\gamma} dx \tag{62}$$

(Lieb-Thirring inequality).

For the proof we refer to [4, 14].

Notice that in the case of the classical hamiltonian $H_{class} = p^2 + V(q)$ the bound

$$S_\gamma(V) \leq C_1 \int_{R^d} |V_-(x)|^{\frac{d}{2}+\gamma} \tag{63}$$

is equivalent (after integration on the coordinates p) to

$$S_\gamma(V) \leq C_2 \int \int_{\{H(p,q)\leq 0\}} |H(p,q)|^\gamma dqdp \tag{64}$$

For $\gamma = 0$ this estimate is

$$N_0(V) \leq C_3 \int \int_{\{H(p,q)\leq 0\}} dqdp \tag{65}$$

a formula [16] that suggests that for $\hbar$ sufficiently small the number of negative eigenvalues is very close to the number of hypercubes of side $\hbar$ contained in the region where $H(p, q) \leq 0$.

One arrives to this conclusion also with more refined methods of semiclassical analysis; [6] gives a detailed and very interesting description of methods that lead to estimate the number of negative eigenvalues of the hamiltonian $-\hbar^2\Delta + V(x)$ in the limit $\hbar \to 0$.

5 Feshbach Method

We will consider now briefly the problem of determining (some of) the eigenvalues of a self-adjoint operator H on a Hilbert space $\mathcal{H}$ through the solution of a *non-linear* equation in a *smaller* Hilbert space $\mathcal{K}$. In the application of this method the space $\mathcal{K}$ is often finite-dimensional and the solution is found by algebraic means.

This method is known in Physics (especially atomic physics) as *Feshbach Method*, or also as *Feshbach Map*. The method was introduced by Feshbach [5] in the study of the scattering of a particle by a nucleus.

In this case in the "final" state of the system there will be several *channels* corresponding to the variation of the state of the nucleus (inelastic scattering). If we are interested only in the elastic part of the scattering (the channel in which the nucleus remains in its ground state) we could try to use a method similar to Gauss's elimination method to "project out" the inelastic part.

5.1 The Physical Problem

For the sake of concreteness we briefly describe the problem that originated the method. We shall give later an abstract version.

The system under study is composed of nucleus A of atomic number N and of an incident particle. We denote by $x \in R^3$ the coordinates of the particle, and by $X \equiv \{x_k, k = 1, \ldots N\}$ the coordinates of the particles in the nucleus. The representation space is

$$L^2(R^{3N+3}) \equiv \mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2, \quad \mathcal{H}_1 = L^2(R^3), \quad \mathcal{H}_2 = L^2 R^{(3N)}$$

We want to solve the *time-independent* Schrödinger equation

$$H\Phi = E\Phi, \qquad H = H_A + T_0 + V(x, X) \qquad T_0 = -\frac{1}{2m}\Delta_x \tag{66}$$

where H_A is the hamiltonian of the atom and m is the mass of the electron. The solution provides the bound states of the composite system of E is in the discrete spectrum of the system, and the generalized scattering states otherwise.

Suppose H_A has point spectrum and let ψ_k be a basis of eigenfunctions. A generic element $\psi \in \mathcal{H}$ can be written in the form

$$\psi = \sum_k u_k(x)\psi_k(X), \qquad u_k(x) \in \mathcal{H}_1, \quad k = 0, 1, \ldots$$

Making use of the orthogonality properties of the $\psi_k(X)$ in $\mathcal{H}_2$ one can write (66) as a matrix equation

$$(T_0 + V_{k,k}(x) + \epsilon_k - E)u_k(x) = -\sum_{j \neq k} V_{j,k}(x)u_k(x) \qquad k = 1, \ldots \tag{67}$$

where $V_{j,k}(x) \equiv \int(\phi_j(X), V(x; X)\psi_k(X))dX$.

We want to extract from (67) an equation on $u_0(x)$.

Denoting by Φ the column vector composed of the u_k, $k \geq 1$ Eq. (67) can be rewritten as

$$(T_0 + V_{0,0}(x) - E)u_0 = -(W_0(x).\Phi), \qquad (H - E)\Phi = -W_0^t u_0 \tag{68}$$

with $W_0(x) \equiv (V_{0,1}(x), V_{0,2}(x) \ldots)$. Solving for Φ the second equation in (68) and substituting the solution in the first equation on obtains

$$[T_0 + V_{0,0} + \int V_0[(E - H)^{-1}V_0^* - E]u_0(x) = 0 \tag{69}$$

which is the equation we are looking for. This a *consistency* relation under which we can solve (68).

When the spectrum of the hamiltonian is at least partly continuous some refinements are needed to make this argument in a rigorous proof: see e.g. [2] where the bound states and resonances are studied for the case of the interaction of an atom with the quantized electromagnetic field.

5.2 Abstract Setting

We shall now give the abstract version of the method [15]. In the mathematics literature this abstract method is often called *Schur complement formula* or, in the applications to linear partial differential equations, *Grushin Problem*.

Let the Hilbert space be $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ and assume that the operator H can be decomposed accordingly

$$H = \{H_{i,j}\} \;\; i, j = 1, 2 \tag{70}$$

Assume that there is a domain $\Omega \in C$ such that $H_{2,2} - zI_{2,2}$ is invertible for $z \in \Omega$.

Define *resonance operator* (this notation comes from Nuclear Physics) the operator on H_1 defined by

$$G_1(z) \equiv zP_1 - P_1 H_{1,1} - H_{1,2}(zI_{2,2} - H_{2,2})^{-1} H_{2,1} P_1 \tag{71}$$

where P_1 is the orthogonal projection on the subspace $\mathcal{H}_1$. In the mathematical literature, this operator is often called *Schur complement*.

Notice that (71) can be formally obtained by projecting on the space $\mathcal{H}_1$ an expression obtained by suitably resuming the Born series for the resolvent.

By substitution it is easy to see that $z \in \sigma(H)$ (i.e. $H - zI$) is not invertible) implies $0 \in \sigma(G_1(z))$ with the same multiplicity.

In fact one has

$$tr(\int_{\gamma_z} (\zeta I - H)^{-1} d\zeta) = tr(\int_{\gamma_z} [\partial G_1(\zeta)] G^{-1}(\zeta) d\zeta) \tag{72}$$

where $\gamma_z = z + \epsilon e^{it}, \;\; 0 \leq t \leq 2\pi$ for ϵ small enough.

The Feshbach method is used to find the eigenvalues of H (solutions of a *linear* equation from the knowledge of the solutions of the *non-linear* equation $G_1(z) = 0$). The latter is an *algebraic problem* if H_1 is finite dimensional.

A version of Feshbach method that goes back to Schur is the following: if one writes the resolvent H^{-1} as a matrix $\{B_{i,j}\}$ then the matrix $B_{1,1}$ is invertible iff $H_{2,2}$ is invertible and one has

$$B_{1,1}^{-1} = H_{1,1} - H_{1,2} H_{2,2}^{-1} H_{2,1}, \qquad H_{2,2}^{-1} = -B_{1,2} + B_{1,1}^{-1} - B_{2,1} \tag{73}$$

The Feshbach method can be generalized in several ways [2, 7] always with the purpose of proving that $H - z$ is invertible with inverse bounded if and only if a suitably defined operator $F_{H,P}$ is boundedly invertible.

As a generalization a Feshbach operator $F_{H,P}$ can be constructed using the decomposition of the Hilbert space $\mathcal{H}$ made using commuting bounded operators $\mathcal{P}$ and $\mathcal{Q}$ which satisfy $\mathcal{P}^2 + \mathcal{Q}^2 = I$ *but are not in general projection operators*.

If the domain of H contains the range of P define

$$H_1 \equiv PHP, \qquad H_2 = (I - P)H(I - P) \tag{74}$$

We regard H_2 as operator on $(I - P)\mathcal{H} = \mathcal{H}_2$ and assume that P has been chosen so that the inverse H_2^{-1} is bounded on $\mathcal{H}_2$.

Assume moreover that $(I-P)H_2^{-1}(I-P)HP$ and $PH(I-P)(H_2)^{-1}(I-P)$ be bounded operators.

On the range of P the Feshbach map is given by

$$F_P = PHP - PH(I-P)H_2^{-1}(I-P)HP \tag{75}$$

and the resonance operator is

$$S_P = P - (I-P)(H_2)^{-1}(I-P)HP \tag{76}$$

One has the identity

$$P\phi = [I + (I-P)(H_2)^{-1}(I-P)HP]S_P\phi \quad \forall \phi \in \mathcal{H} \tag{77}$$

and then $Ker\, S_P = Ker\, P$.

The Feschbach map is *isospectral*: if $z \in R(H_2)$ (the resolvent set) then $\sigma_*(H - z) = \sigma_*(F_{P_1}(H - z))$ (σ_* can denote the continuous spectrum or also the point spectrum).

Moreover

$$P\ Ker(H-z) = Ker(F_P(H-z)) \qquad dim\ Ker(H-z) = dim\ Ker(F_P(H-z)) \tag{78}$$

and there is a one-to-one correspondence between the eigenfunctions of $H - z$ and those of $F_P(H - z)$ given by

$$Ker[(H-z)S_P] = Ker\, F_P(H-z) \qquad S_P(z) = P-(I-P)(H_2-z)^{-1}(I-P)HP \tag{79}$$

These conclusions are easy consequences of the relation

$$H = H_1 + H_2 + P_1H(I-P) + (I-P)HP \tag{80}$$

and of the resolvent identity

$$\frac{1}{H-z} - \frac{1}{K-z} = \frac{1}{H-z}(K-H)\frac{1}{K-z} \tag{81}$$

which holds whenever all terms are well defined.

The Feshbach method can be put in a more general context for systems that depend analytically from a parameter and provides analytic solutions. Often this generalization is called *Grushin problem* and does not require that the operators be self-adjoint.

The natural setting for Grushin's problem is the analysis of differential operators, e.g. the problem of finding the normal derivative at the boundary of the solution

of a homogeneous elliptic problem in a domain Ω with regular boundary $\partial\Omega$ with in-homogeneous boundary conditions.

With the notations in (79) Grushin's problem can be described as solvability of a linear system as follows: find the solution (if it exists) u, v of the problem

$$\begin{aligned} Pu + R_- w = v \qquad R_+ u = \zeta \quad u \in \mathcal{H}_1, \ \ \xi \in \mathcal{H}_2 \\ P \ : \ \mathcal{H}_1 \to \mathcal{H}_2, \qquad R_- \ : \ \mathcal{H}_- \to \mathcal{H}_2 \qquad R_+ \ : \ \mathcal{H}_1 \to \mathcal{H}_+ \end{aligned} \tag{82}$$

If the solution exists it is given by

$$u = H_{1,1} v + H_{1,2} \zeta \qquad w = H_{2,1} v + H_{2,2} \zeta \tag{83}$$

In this case the problem is to reduce the solution of a linear problem (e.g. the Schrödinger equation) to a non-linear equation for a subsystem the operator $H_{2,2}$ plays the role of *effective hamiltonian*.

The Born-Oppenheimer approximation for the analysis of a system containing *fast and slow* degrees of freedom can be considered as an instance of Grushin's method.

References

1. Aizenman, E. (1978). *Lieb Phys. Lett.*, *66A*, 427–429.
2. Bach, V., Chen, T., Froelich, J., & Segal, I. (2003). *J. Functl. Anal.*, *203*, 44–92.
3. Bargmann, V. (1952). *Proc. Natl. Acad. Sci. USA*, *38*, 961–966.
4. Cwikel, M. (1977). *Ann. Math.*, *106*, 93–100.
5. Feshbach, G. (1958). *Ann. Phys.*, *5*, 357–390.
6. Folland, G. B. (1978). *Fourier analysis in phase space*. Princeton: Princeton University Press.
7. Griesemer, M., & Hasler, D. (2008). *J. Functl. Anal.*, *254*, 2329–2335.
8. Glaser, V. et al. (1976). *Studies in mathematical physics*. Princeton: Princeton University Press, pp. 169–194.
9. Glaser, V., et al. (1978). *Commun. Math. Phys.*, *59*, 197–212.
10. Kac, M. (1978). *Topics in functional analysis*. In I. Godberg & K. M. Kac (Eds.), New York: Academic Press.
11. Lieb, E. (1976). *Bull. A.M.S.*, 82, 751–753.
12. Martin, A. (1977). *Comm. Math. Phys.*, *55*, 293–298.
13. Martin, A. (1972). *Helv. Phys. Acta*, 45, 140–148.
14. Reed, M., & Simon, B. *Methods on mathematical physics 1, 2* New York: Springer Verlag 72,75.
15. Syöstrand, J. & Zworski, M. (2007). *Annales de l'Institut Fourier*, *57*, 2095–2141.
16. Weyl, A. (1911). *Mathematische Annalen*, *71*, 441–469.

Lecture 20: Self-adjoint Extensions. Relation with Quadratic Forms. Laplacian on Metric Graphs. Boundary Triples. Point Interaction

We have seen that for an operator on a Hilbert space $\mathcal{H}$ the condition to be symmetric *is not sufficient* to be the generator of a continuous one-parameter group of unitary operators and thereby to define a dynamics. We will provide now the basic elements of a procedure by which, under suitable conditions, the domain of a symmetric operator can be extended to be the domain of a self-adjoint operator.

1 Self-adjoint Operators: Criteria and Extensions

Let A be a closed, densely defined symmetric operator on $\mathcal{H}$.

Define

$$\mathcal{K}_+ \equiv Ker(A^* - iI), \qquad \mathcal{K}_- \equiv Ker(A^* + iI) \tag{1}$$

$$m_+ = dim\ \mathcal{K}_+, \qquad m_- = dim\ \mathcal{K}_- \tag{2}$$

Definition 1 The integers $m_\pm$ are called *defect indices* of the operator A and $\mathcal{K}_\pm$ are its *defect subspaces*. We have already seen that $m_+ = m_- = 0$ is a necessary condition in order that the closure of A be a self-adjoint operator. ◇

The defect spaces could have been defined as the Kernels of the operators $A^* - z_0 I$ and $A^* - \bar{z}_0 I$ for *any* z_0 such that $Im z_0 \neq 0$. The dimension of these spaces does not change if z_0 varies in the upper (or lower) half-plane but may be different for the two half-planes. The choice $z_0 = \pm i$ is the most common.

Example 1 Let A be the symmetric operator $-i\frac{d}{dx}$ on $L^2(0, 1)$ with domain those twice differentiable functions in $(0, 1)$ which vanish in a neighborhood of the extreme points. From the definition of adjoint, integrating by parts it easy to verify that

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5_20

the domain of the (closed) operator A^* are the absolutely continuous functions in $[0, 1]$ and that on this domain A^* coincides (a part from the factor $-i$) with the distributional derivative.

Notice that in the definition of the domain of the adjoint there is no condition on the function at the boundary (in the integration by parts the boundary terms are absent by the hypothesis on the domain of A). Therefore the closed operator A^* extends A. The defect subspaces are

$$\mathcal{K}_+ \equiv \{\lambda e^x, \quad \lambda \in C\}, \quad \mathcal{K}_- \equiv \{\lambda e^{-x}, \quad \lambda \in C\} \quad m_+ = m_- = 1 \tag{3}$$

Since $A^* \neq \bar{A}$ the operator $\bar{A}$ is *not* self-adjoint.

Notice that A^* *is not* a symmetric operator. Indeed if $f, g \in D(A^*)$ one has

$$\begin{aligned}(A^* f, g) &= \int_0^1 -i \frac{d\bar{f}(x)}{dx} g(x) dx \\ &= \int_0^1 i \bar{f}(x) \frac{dg(x)}{dx} dx + f(1)g(1) - f(0)g(0) \neq (f, A^* g)\end{aligned} \tag{4}$$

Since the boundary terms are *not* continuous in the topology of $L^2[0, 1]$, the domain of $(A^*)^*$ are the absolutely continuous functions which vanish in 0 and in 1. It follows $(A^*)^*$ is the closure of A in the graph norm. One has $D((A^*)^*) \subset D(A^*)$ with strict inclusion. In this example the domains of A and A^* differ only by the boundary conditions. This is typical for differential operators defined in an open domain Ω with regular boundary $\partial\Omega$. ♣

Example 2 Let A be the symmetric operator $i \frac{d}{dx}$ with domain the absolutely continuous functions on R^+ with compact support and vanishing in a neighborhood of the origin. In this case one has

$$\mathcal{K}_+ = \{0\} \quad \mathcal{K}_- \equiv \{\lambda e^{-x} \quad \lambda \in C\}, \quad m_+ = 0, \quad m_- = 1 \tag{5}$$

Also here the adjoint operator A^* is the operator $-i \frac{d}{dx}$ defined on absolutely continuous functions on $L^2(R^+)$ without limitations on the value at the origin. Therefore A^* extends A. Integrating by parts, one sees that the boundary term are not continuous in the $L^2(R^+)$ topology and therefore the double adjoint $(A^*)^*$ is the operator $i \frac{d}{dx}$ defined on the functions on $L^2(R^+)$ which are absolutely continuous and zero at the origin. Also in this case $(A^*)^*$ coincides with the closure of A. ♣

Example 3 Let A be the operator $\frac{d^2}{dx^2}$ defined on the functions in $L^2(R^+)$, of class C^∞, with compact support and vanishing in a neighborhood of the origin. It is not difficult to verify that in this case the defect space are

$$\mathcal{K}_+ = \{\lambda e^{\frac{(i-1)x}{\sqrt{2}}} \qquad \lambda \in C\}, \qquad n_+ = 1$$

$$\mathcal{K}_- = \{\lambda e^{-\frac{(i+1)x}{\sqrt{2}}} \qquad \lambda \in C\}, \qquad n_- = 1 \tag{6}$$

Also in this case $D(A) \subset D(A^*)$ and $(A^*)^* = \bar{A}$. The domain of A^* are the functions on R^+ which are square integrable together with their first and second derivative (while the functions in the domain of the closure of A satisfy the further requirement that the functions and their first derivative vanish at the origin). ♣

The examples we have given suggest the strategy to determine the conditions under which a closed symmetric operator admits a self-adjoint extension, i.e. the conditions for the existence of a self-adjoint operator A_1 with domain $D(A_1) \supset D(A)$ and such that on $D(A)$ one has $A_1 = A$.

Since in any case $D(A_1) \subset D(A^*)$, the operator A_1 must be a restriction of A^*. If A_1 is self-adjoint one must have $D(A_1) = D(A_1^*)$. Therefore A_1 must be *maximal* among the extension of A (or equivalently a *minimal* restriction of A^*).

In this Lecture we shall give a classification of all self-adjoint extensions of a symmetric operator A with equal defect indices.

For this purpose it is convenient to introduce the space $\mathcal{H} \oplus D(A^*)$. Consider the graph of A^* i.e. the pair $\Gamma(A^*) \equiv \{\phi, A^*\phi\}, \ \phi \in D(A^*)$. Since A^* is closed (we have assumed that A be densely in $\mathcal{H}$), the graph A^* is a closed subset of $\mathcal{H} \oplus D(A^*)$.

The closure is understood in the sense of the graph norm, defined by

$$|\{\phi, A^*\phi\}|^2 \equiv |\phi|^2 + |A^*\phi|^2 \tag{7}$$

Define on $\Gamma(A^*)$ a pre-Hilbert structure through the scalar product

$$\langle \tilde{\phi}, \tilde{\psi} \rangle \equiv (\phi, \psi) + (A^*\phi, A^*\psi) \tag{8}$$

Denote by $\tilde{\mathcal{H}}_A$ the corresponding Hilbert and by $\tilde{\phi} \in \tilde{\mathcal{H}}_A$ the element $\tilde{\phi} \equiv \{\phi, A^*\phi\}$. Define in $\tilde{\mathcal{H}}_A$ the antisymmetric form

$$\sigma_A(\tilde{\phi}, \tilde{\psi}) \equiv (\phi, A^*\psi) - (A^*\phi, \psi) \tag{9}$$

The subspace $\mathcal{K} \subset D(A^*)$ is said to be *A-symmetric* if

$$\sigma_A(\tilde{\psi}, \tilde{\phi}) = 0 \quad \forall \phi, \quad \psi \in \mathcal{K}$$

(i.e. if A^* restricted to $\mathcal{K}$ is *a symmetric operator*). Denote by $\tilde{\mathcal{K}}_A$ the subspace of $\tilde{\mathcal{H}}_A$ defined by

$$\tilde{\mathcal{K}}_A \equiv \cup_{\phi \in \mathcal{K}} \{\phi, \ A^*\phi\} \tag{10}$$

The subspace $\tilde{\mathcal{K}}_A$ is a closed subspace of $\tilde{\mathcal{H}}_A$; the antisymmetric form σ_A is closed and continuous in the topology of $\tilde{\mathcal{H}}_A$ and is degenerate on this subspace since $\sigma_A(\tilde{\phi}, \tilde{\psi}) = 0$ if $\phi, \quad \psi \in D(A)$.

Let $\Gamma(A)^{\perp}$ be the orthogonal complement of $\Gamma(A)$ in $\tilde{\mathcal{H}}_A$. The antisymmetric form σ_A may be degenerate also in other subspaces of $\tilde{\mathcal{H}}_A$. We shall call *maximally A-symmetric* the subspaces of $\tilde{\mathcal{H}}_A$ on which the form σ_A is non degenerate.

In analogy with the structure of the cotangent bundle in Hamiltonian mechanics, if σ_A is regarded as a symplectic structure, a closed A-symmetric subspace of $\Gamma(A)^{\perp}$ can be interpretedas *isotropic manifold*. If it is maximal, it is a Lagrangian manifold. With these notation, the following theorem holds.

Theorem 1 (Riesz-Nagy) *The symmetric closed extensions of the operator A are in one-to-one correspondence with the closed A-symmetric subspaces of* $\Gamma(A)^{\perp}$. *In particular an extension of A is self-adjoint if and only if the corresponding subspace is maximally symmetric (i.e. it is a Lagrangian manifold).* ◇

We shall not prove this theorem (see e.g. Riesz-Nagy, *Functional Analysis*) but we will analyze the structure of the A-symmetric subspaces and we will give an useful characterization (and construction) of the self-adjoint extensions, if they exist. Let us remark that the Riesz-Nagy theorem implies that every self-adjoint extension A_1 defines a lagrangian manifold and a maximal A-symmetric subspace. Indeed, since A_1 is self-adjoint and $D(A_1) \subset D(A^*)$, the subspace $\Gamma(A_1)$ is A-symmetric and contained in $\Gamma(A^*)$.

If it is not maximal there exists $\xi \notin D(A_1)$ such that

$$(\phi, A^*\xi) = (A_1\phi, \xi) \qquad \forall \phi \in D(A_1) \tag{11}$$

Therefore $\xi \in D(A_1^*) = D(A_1)$, a contradiction.

2 von Neumann Theorem; Krein's Parametrization

To characterize the self-adjoint extension corresponding to a given maximal subspace notice the orthogonal decomposition, with respect to the scalar product (8)

$$\tilde{\mathcal{H}}_A = \Gamma(A) \oplus \Gamma_+ \oplus \Gamma_- \tag{12}$$

where

$$\Gamma_+ \equiv \{\phi, i\phi\}, \quad \phi \in \mathcal{K}_+ \qquad\qquad \Gamma_- \equiv \{\phi, -i\phi\}, \quad \phi \in \mathcal{K}_-$$

To prove the decomposition (12) we prove first that $\Gamma(A) \perp \Gamma_+$.
If $\tilde{\phi} = \{\phi, A\phi\} \in \Gamma(A)$ and $\tilde{\eta} = \{\eta, i\eta\} \in \Gamma_+$ one has

$$\begin{aligned} \langle \tilde{\phi}, \tilde{\eta} \rangle &= (\phi, \eta) + (A^*\phi, A^*\eta) = (\phi, \eta) + i(A^*\phi, \eta) \\ &= (\phi, \eta) + i(\phi, A^*\eta) = (\phi\eta) - (\phi, \eta) = 0 \end{aligned} \tag{13}$$

In the same way one proves $\Gamma(A) \perp \Gamma_-$. To complete the proof of (12) we must show that if $\tilde{\phi}$ is orthogonal (in $\tilde{\mathcal{H}}$) to $\Gamma(A)$ and also to $\Gamma_\pm$, then $\phi = 0$. Assume $\phi \in D(A)$ and $\langle \tilde{\phi}, \tilde{\psi} \rangle = 0$. Then

$$(\phi, \psi) + (A^*\phi, \psi) = 0 \tag{14}$$

It follows $A\phi \in D(A^*)$ and $(A^*)^*\phi = -\phi$. But then $(A^* + i)(A^* - i)\phi = 0$ and therefore $(A^* - i)\phi \in K_-$. On the other hand if $\xi \in K_-$ one has $i(\xi, (A^* - i)\phi) = (\xi, \phi) = 0$. Therefore $(A^* - i)\phi \in K_- \cap K_-^\perp = \{0\}$ and $\phi \in K_+$. By assumption ϕ does not belong to K_+. Therefore $\tilde{\phi} = 0$ and $\phi = 0$. The proof of the decomposition (12) is complete.

Let us also notice that if the subspace $\Gamma_+ \oplus \Gamma_-$ is A-symmetric, then $|\xi| = |\eta|$ Indeed from $A^*\xi = i\xi, \quad A^*\eta = -i\eta$

$$0 = (\xi + \eta, A^*(\xi + \eta) - (A^*(\xi + \eta), \xi + \eta) = 2(|\xi|^2 - |\eta|^2) \tag{15}$$

We conclude that the A-symmetric subspaces of $\hat{\mathcal{H}}_A$ have the structure

$$\{\phi |\ \phi = \psi + \xi + \eta, \ \ |\xi| = |\eta|, \ \ \xi \in K_+, \ \ \eta \in K_-, \ \ \psi \in D(\bar{A})\} \tag{16}$$

These subspaces cannot be maximal if $dim K^+ \neq dim K^-$. Consider the decomposition

$$\phi_k = \psi_k + \xi_k + \eta_k \quad k = 1, 2 \quad \psi \in D(A) \tag{17}$$

The antisymmetric two-form σ

$$\sigma_A(\tilde{\phi}_1, \tilde{\phi}_2) = i[(\xi_1, \eta_2) - (\xi_2, \eta_1)] \tag{18}$$

is not degenerate iff $dim\ \mathcal{K}_+ = dim\ \mathcal{K}_-$.

We have proved.

Theorem 2 (von Neumann) *Let A be closed and symmetric. Its closed symmetric* maximal *extensions are in one-to-one correspondence with the isometries between $\mathcal{K}_+$ and $\mathcal{K}_-$.* ◇

Let $m_+ \geq m_-$. Denote by $\mathcal{J}$ a partial isometry of $\mathcal{K}_+$ to $\mathcal{K}_-$. The domain of the corresponding extension, denoted by $A_\mathcal{J}$ is

$$D(A_\mathcal{J}) = \{\psi + \xi + \mathcal{J}\xi\}, \quad \psi \in D(A), \quad \xi \in D(\mathcal{J}) \tag{19}$$

where $D_\mathcal{J}$ is the domain of the partial isometry $\mathcal{J}$. On this domain one has

$$A_\mathcal{J}\phi = A\psi + i\xi - i\mathcal{J}\xi \tag{20}$$

If $m_- > m_+$ the role of $\mathcal{K}_+$ and of $\mathcal{K}_-$ is inverted.

The operator A admits self-adjoint extensions if and only if $m_+(A) = m_-(A)$. In this case the isometry $\mathcal{J}$ is invertible and therefore the self-adjoint extensions are in one-to-one correspondence with the unitary maps U from $\mathcal{K}_+$ to $\mathcal{K}_-$. The domain and action of the extension corresponding to $\mathcal{J}$ are given in (19) and (20).

The extension $A_{\mathcal{J}}$ is obtained by adding to the closure of the domain of A the subspace of dimension m_+ of the vectors $\xi + \mathcal{J}\xi$, with $\xi \in K_+$. The image of this vector under $A_{\mathcal{J}}$ is $i(\xi - \mathcal{J}\xi)$. On the other hand, if ϕ is in the domain of A one has $A_{\mathcal{J}}\phi = A\phi$.

If one defines $\mathcal{K}_{\pm}^z$ as the Kernels of $H^* \mp z,\ Imz > 0$ the previous analysis goes through without change. From (12) one can see that the space

$$\mathcal{K}_z \cup \mathcal{K}_{\bar{z}} \cup D(A)$$

is *independent* of the choice of $z,\ Imz > 0$.

Recall that all Hilbert spaces which have the same dimension are isomorphic. If one chooses orthonormal complete bases in K_- and in K_+ the isometry $\mathcal{J}$ is represented by an unitary matrix $U_{\mathcal{J}}$. Since the choice of the orthonormal base is arbitrary the matrix $U_{\mathcal{J}}$ is defined modulo multiplication to the right or to the left by a unitary matrix.

Denoting by I the identity in the space K_+ one has $A_{\mathcal{J}}.(I + U_{\mathcal{J}}) = i(I - U_{\mathcal{J}})$ or equivalently

$$A_{\mathcal{J}} = i(I - U_{\mathcal{J}})(I + U_{\mathcal{J}})^{-1} \tag{21}$$

This is called *Krein parametrization*.

The choice of basis is not legitimate (singular) if the operator $U_{\mathcal{J}}$ has -1 as eigenvalue; this property depends on the basis chosen.

There are no general criteria for the choice of the bases. In the case of second order differential operators *with real coefficients* if $f \in \mathcal{K}_+$ then $\bar{f} \in \mathcal{K}_-$. Therefore in this case the natural choice is a complete basis

$$\{\xi_k(x)\} \in \mathcal{K}_+, \qquad \{\bar{\xi}_k(x)\} \in \mathcal{K}_-.$$

Denote by W the anti-unitary operator that extends the correspondence between the bases to a correspondence between $\mathcal{K}_+$ and $\mathcal{K}_-$. The operator $U_{\mathcal{J}}$ can be cast in the form $U_{\mathcal{J}} = UW$ with U unitary in $\mathcal{K}_+$.

One can also use a basis $\mathcal{K}_+ \oplus \mathcal{K}_-$ real with respect to $\mathcal{J}$

$$W_k^+ = \frac{1}{2}(\xi_k + \mathcal{J}\xi_k), \qquad W_k^- = -\frac{i}{2}(\xi_k - \mathcal{J}^*\xi_k), \tag{22}$$

In this basis one has

$$A_{\mathcal{J}} W^{\pm} = W^{\mp}$$

Moreover

$$(W_h^{\pm}, W_k^{\pm}) = \pm((A_{\mathcal{J}})^{-1}\phi_h, \phi_k) + \frac{1}{2}\delta_{h,k} \pm \frac{1}{2}\delta_{h,k}$$
$$(W_h^{\pm}, W_k^{\mp}) = -((A_{\mathcal{J}} + 1)^{-1}A_{\mathcal{J}}\phi_h, \phi_k) \tag{23}$$

Since

$$u = \phi + \sum_{k=1}^{2m}(c_k^+(u)W_k^+ + c_k^-(u)W_k^-), \quad \phi \in D(A)$$

when $u,\ v \in D(A^*)$ one has

$$(A_{\mathcal{J}}u, v) - (u, A_{\mathcal{J}}v) = \sum_{h=1}^{m}(c_k^-(u)c_k^+(v) - c_k^+(u)c_k^-(v) \equiv (\xi(u), J\xi(v)) \tag{24}$$

where J is the symplectic matrix.

In this way the defect spaces assume the form of the standard symplectic spaces. All other self-adjoint extensions are obtained by linear symplectic transformations.

In several simple cases this standard symplectic structure appears in a natural way. For example for the operator

$$(-\frac{d^2}{dx^2} + V)u \equiv A_0 u \quad u \in C_0^{\infty} \quad x \in R^+ \quad V \in \mathcal{R} \tag{25}$$

one has

$$(A_0^*u, v) - (u, A_0^*v) = (u'(0), v(0)) - (u(0), v'(0)) \tag{26}$$

With this notation if one considers different extension through invertible isometric operators U_1, U_2 from $\mathcal{K}_+$ to $\mathcal{K}_-$ the relation between their resolvents is given for $Im\ \lambda \neq 0$, choosing the base adapted to $\mathcal{J}$ given by (22), by

$$(A_1 - \lambda)^{-1} - (A_2 - \lambda)^{-1} = (A_1 - \lambda I)^{-1}P[Q - P\frac{1 + \lambda A_2}{A_2 - \lambda}P]^{-1}P(A_1 - \lambda)^{-1} \tag{27}$$

where $A_k \equiv A_{U_k}$ $Q \equiv \sum_k \phi_k \Gamma_{k,h}(., \phi_h)$ and P is the orthogonal projection on the subspace relative to the eigenvalue $+i$ of A_1. The matrix Γ is defined as

$$\psi = \phi + \Gamma\xi_+ + \xi_- \tag{28}$$

The relation (27) between the resolvents of two self-adjoint extensions of a symmetric operator with equal defect spaces was first obtained in the case of defect

indices $\{1,\ 1\}$ in [11] and then extended by Krein himself to the case of equal finite defect indices [12]. The same formula, under the appropriate convergence assumptions, holds for the case in which the defect spaces are infinite-dimensional. One should compare (27) with the formula of Birman ans Schwinger for the difference of the resolvents of the operators $\Delta - V_1(x)$ and $\Delta - V_2(x),\ x \in R^3$ when the two potentials $V_1,\ V_2$ belong to suitable regularity classes.

The extensions do not depend on the choice of $z_1, z_2;\ Imz_1 > 0, Imz_2 < 0$ in the definition of defect spaces. The defect spaces for different values of z_j are different as subsets of $\mathcal{H}$; their difference belongs to the closure of the domain of the symmetric operator $-\frac{d^2}{dx^2} + V(x)$.

3 The Case of a Symmetric Operator Bounded Below

One may wonder what happens when $z_\pm \to x_0 \in R$. This is particularly interesting if the symmetric operator bounded below. We consider the case in which A is bounded below and the defect space is finite-dimensional. Without loss of generality we assume $A \geq 0$.

Choose $z = i\epsilon$, and denote by $\mathcal{K}^\epsilon_\pm$ the correspondent defect spaces (the solutions of $A^*u = \pm i\epsilon u$). When $\epsilon \to 0$ these spaces tend to coincide and one can choose bases such that the unitary operators U_ϵ converge to I. The self-adjoint extensions in this limit are characterized by the operator

$$lim_{\epsilon\to} \frac{U_\epsilon - I}{\epsilon} \equiv B \tag{29}$$

which acts on the space of solutions of $A^*u = 0$. We shall denote this subspace by $\mathcal{N}$. It is not difficult to see that the operator B is symmetric; in case the defect spaces are finite dimensional, B is represented by a symmetric matrix and therefore it is self-adjoint. More generally, if A^* is bounded below, one can prove B is self-adjoint even if $\mathcal{N}$ is infinite-dimensional.

We conclude that in this case each self-adjoint extensions of A *is characterized by a self adjoint operator* B acting $\mathcal{N}$. Recall that $\mathcal{N}$ is a closed subset of $\mathcal{H}$ such that $\mathcal{H} - \mathcal{N}$ is dense in $\mathcal{H}$. Denote by A_B the self-adjoint extension associated with B and recall that the defect space is orthogonal to the domain of A in the graph topology. It follows that if $\psi = \phi + \xi,\quad \phi \in D(A) \xi \in D(B)$ then

$$(\psi, A_B\psi) = (\phi, A\phi) + (\xi, B\xi) \tag{30}$$

In this way we have described the quadratic forms associated to the self-adjoint extensions of the operator A as *the sum* of the quadratic form associated to the operator A and the quadratic form of an auxiliary self-adjoint operator B defined on an auxiliary Hilbert space $\mathcal{N}$ (called *Krein space*). It can be presented as a subset (*not a subspace*) of the Hilbert space $\mathcal{H}$ closed in the graph-topology of A^*.

This presentation, due to Krein, of the self-adjoint extensions of a symmetric operator *bounded below* by reduction to a sum of quadratic forms is a fundamental step in the construction of all self-adjoint extensions. It must be emphasized however that this method applies only if the symmetric operator is semi-bounded (bounded above or below).

4 Relation with the Theory of Quadratic Forms

In view of Krein's result it is natural to introduce quadratic form q_A defined on the domain of A by $q_A(\phi) = (\phi, A\phi)$ and exploit the additivity property (30). This shifts the problem of finding all self-adjoint extensions of a symmetric operator to the problem of describing the collection of closed bilinear forms that are associated to the extensions.

The forms associated to self adjoint extension of a symmetric operator A are a partially ordered set. The following theorem of Kato links, for self-adjoint operators, the partial order of the quadratic forms to the partial order of the resolvents.

Theorem 3 (Kato) [10, 14] *Consider two self-adjoint operators A, B on a Hilbert space $\mathcal{H}$, bounded below and let $c \in R$ be chosen in such a way that both $A + cI$ and $B + cI$ are positive. Then $q_A > q_B$ iff $\frac{1}{A+cI} < \frac{1}{B+cI}$. As a consequence the infimum of the spectrum of B is greater than the infimum of the spectrum of A if $q_A < q_B$.* ◇

We do not give here the proof of this theorem.

In view of Kato's theorem one can ask whether among the forms associated to the self adjoint extension of a symmetric operator bounded below there is a *minimal* one and a *maximal* one. One proves that even if the defect spaces are infinite-dimensional such minimal and maximal forms always exist. If one denotes by $q_{min}(A)$ respectively $q_{max}(A)$ the extremal points, one defines in this way *minimal* and *maximal* extensions of a symmetric operator A bounded below.

In the case $A \equiv -\Delta + V(x),\ x \in R^3$ with V in a suitable regularity class, the *maximal extension* is the Friedrichs extension.

In the second part of these Lectures we will discuss in more detail the relation between closed quadratic defined on the Krein space and the Krein-Birman classification of the self-adjoint extensions. Here we limit ourselves to two simple examples.

Example 4 The symmetric operator $H^0 \equiv -\frac{d^2}{dx^2} + 1$ is defined on functions of compact support which vanish in a neighborhood of the origin.

Choose $x_0 > 0$ and define

$$\mathcal{K}^{\epsilon}_{+} = \{ce^{-\alpha_+(\epsilon)x}\}, \qquad \mathcal{K}^{\epsilon}_{+} = \{ce^{-\alpha_-(\epsilon)x}\} \qquad c \in C \tag{31}$$

where $\alpha_{\pm}(\epsilon)$ are the solution of $\alpha^2 = (x_0 \pm i\epsilon)$ which have a negative real part so that $e^{\alpha_{\pm}(\epsilon)x} \in L^2(0, \infty)$. We set $\alpha_{\pm}(\epsilon = -\alpha \pm i\beta,\ \alpha = \sqrt{x}_0 > 0)$.

Different choices of x_0 provide defect space which are equivalent since the difference belongs to the closure of the domain of A.

For $\epsilon \neq 0$ the self-adjoint extensions were defined by a unitary relation (a phase) defined as

$$\xi_+^\epsilon + e^{i\beta}\xi_-^\epsilon = i(\xi_+^\epsilon - e^{i\beta}\xi_-^\epsilon), \qquad |\xi_+^\epsilon| = |\xi_-^\epsilon| \quad \xi_\pm^\epsilon \in \mathcal{K}_\pm^\epsilon \tag{32}$$

The domain of the self-adjoint extension $H_{\alpha,\beta}$ are all the function which can be written as

$$\psi(x) = c[e^{\alpha x} + e^{i\beta}e^{-\alpha x}] + \phi(x), \qquad \phi \in D(H^0) \qquad \psi \in L^2(R_+) \tag{34}$$

By assumption $\phi(0) = \frac{d\phi}{dx}(0) = 0$ for all $\phi \in D(H^0)$. Therefore the domain of the extension is characterized by

$$\psi'(0) = \alpha \frac{1 - e^{i\beta}}{1 + e^{i\beta}}\psi(0) \tag{35}$$

In Krein's theory, the quadratic forms associated to some of these self-adjoint extensions are

$$q_\alpha(\phi) = \int (\frac{d\phi}{dx}, \frac{d\phi}{dx})^2 dx + |\phi|^2 + \alpha^2|\phi(0)|^2 \tag{36}$$

(notice that $|\frac{1-e^{i\beta}}{1+e^{i\beta}}|^2 = 1$).

In this example Krein's space is identified with a subspace of boundary values at the origin. In Krein's notations α^2 is a point in the spectrum of a positive self-adjoint operator B defined on Krein's space. The maximal extension corresponds to the point $\alpha^2 = +\infty$ (i.e. $\phi(0) = 0$, Dirichlet b.c.) and the minimal one corresponds to the point $\alpha^2 = 0$ (Neumann boundary conditions).

Remark that the quadratic form can be written as $\int (\frac{d\phi}{dx}, \frac{d\phi}{dx})^2 dx + |\phi|^2$ both for $\alpha = 0$ (Neumann case) and for $\alpha = \infty$ (Dirichlet case) with the convention that in the second case the form is defined function of class H^1 *that vanish at the origin*.

Example 5

$$A = -\frac{d^2}{dx^2} + 1 \quad D(A) = \{\phi : \frac{d\phi}{dx} \in AC(0,\pi), \quad \phi(0) = \phi(\pi) = 0 \quad \frac{d\phi}{dx}(0) = \frac{d\phi}{dx}(\pi) = 0\} \tag{37}$$

The defect subspaces $\mathcal{K}_+$ and $\mathcal{K}_-$ have both (complex) dimension two. The space $\mathcal{K}_+$ is spanned by the distributional solutions of

$$\frac{d^2}{dx^2}\psi(x) = i\psi(x) \tag{38}$$

in $L^2[0, \pi]$ i.e. by the functions

$$\phi_+^{\pm}(x) = e^{\frac{\pm i-1}{\sqrt{2}}x} \tag{39}$$

In the same way the space $\mathcal{K}_-$ is spanned by the functions

$$\phi_-^{\pm} = e^{\frac{\pm i+1}{\sqrt{2}}x}$$

The isometries between $\mathcal{K}_+$ and $\mathcal{K}_-$ are parametrized by 2×2 unitary matrices U. Proceeding as in Example 3 one shows that they correspond to two linear relations among the four quantities

$$a_1 \equiv \psi(0), \quad a_2 \equiv \psi(\pi) \quad b_1 \equiv \frac{d\psi}{dx}(0), \quad b_2 \equiv \frac{d\psi}{dx}(\pi) \tag{40}$$

The relation has the structure

$$Ma + Nb = 0 \quad a = \{a_1,\ a_2\} \quad b \equiv \{b_1,\ b_2\} \tag{41}$$

where M and N are 2×2 matrices such that the 4×2 matrix obtained by setting M and N side to side has maximal rank. These relations are called *local* if they refer separately to $\psi(0)$, $\frac{d\psi}{dx}(0)$ and to $\psi(\pi)$, $\frac{d\psi}{dx}(\pi)$.

If N is invertible one can write $Bb = N^{-1}Ma$ and in the formulation by Krein the quadratic forms corresponding to the self-adjoint extensions can be written using the boundary values at the origin and at π of the functions in the domain. One has

$$q_{M,N}(\phi) = \int_0^{\pi} |\nabla\phi(x)|^2 + \sum_{i,j=1}^{2} C_{M,N}^{i,j} q_i q_j, \quad q_1 = \phi(0), \quad q_2 = \phi(\pi) \tag{42}$$

where the (non-negative) two by two matrix $C_{M,N}$ are in one-to one correspondence with the boundary conditions (42).

For Dirichlet boundary conditions one has $C_{M,N} = \infty I$.

If N is not invertible the quadratic form in Krein's space depends also on the value of the derivative at the boundary. This is the case e.g. of periodic and Neumann boundary conditions.

It is easy to generalize this approach to the case of the operator $-\frac{d^2}{dx^2} + V(x)$ on $(0, \pi)$ with $V(x)$ continuous and bounded.

5 Special Cases: Periodic, Dirichlet and Neumann Boundary Conditions

We consider now more in detail the self-adjont extensions of the symmetric operator $\frac{d^2}{dx^2}$ defined on C^2 functions over the interval $(0; 2\pi)$ and which satisfy either one of the following conditions:

(i) (*Dirichlet boundary conditions*)

$$\phi(0) = \phi(2\pi) = 0 \tag{43}$$

(ii) (*Neumann boundary conditions*)

$$\phi'(0) = \phi'(2\pi) = 0 \tag{44}$$

(iii) (*Periodic boundary conditions*)

$$\phi(0) = \phi(2\pi), \qquad \phi'(0) = \phi'(2\pi) \tag{45}$$

We treat first the periodic case (45). This corresponds to a degenerate form of the Krein operator. We shall denote by $-\Delta_P$ the corresponding self-adjoint extension. Its eigenfunctions are the restriction to $[0, 2\pi)$ of the periodic solutions of $-\frac{d^2}{dx^2} f = \lambda f$. The periodicity condition implies $\lambda = 0, 1, 4, ..n^2...$ and the solutions are

$$\psi_P^n(x) = \frac{1}{\sqrt{2\pi}} e^{inx} \qquad n \in N, \qquad x \in [0, 2\pi] \tag{46}$$

These functions form a complete orthonormal basis in $L^2(0, 2\pi)$. It follows that the operator described by periodic b.c. is self-adjoint. Its spectrum is the set of squares of natural numbers including zero. The operator is the square of the self-adjoint operator defined by $i\frac{d}{dx}$ with periodic boundary conditions. It follows that the each non zero eigenvalue is two-fold degenerate.

Consider next the Dirichlet b.c. conditions (43). We denote the corresponding operator by $-\Delta_D$. A complete set of eigenfunctions is now

$$\psi_D^n(x) = \frac{1}{\sqrt{\pi}} sen(\frac{nx}{2}) \tag{47}$$

The corresponding eigenvalues are $\frac{n^2}{4}$, $n = 1, 2, 3, ...$ and all are non degenerate. This shows that $-\Delta_D$ *is not* the square of a self-adjoint operator. To the domain of $-\Delta_D$ belong all function that can be written as

$$\phi(x) = \sum_k c_k \psi_D^k \qquad \sum |c_k|^2 |k|^4 < \infty \tag{48}$$

From the theory of Fourier series one derives that the domain of Δ_D is made of the functions which are zero in 0 and 2π with second derivative $L^2(0, 2\pi)$.

Consider now the Neumann condition (44); we will denote the corresponding operator by $-\Delta_N$. A complete set of eigenfunctions is now

$$\psi_N^n(x) = b_n cos(\frac{nx}{2})$$

The eigenvalues are $\frac{n^2}{4}$, $n = 0, 1, 2, 3...$ and are non-degenerate for $n > 0$. To the domain of Δ_N belong the functions that can be written

$$\phi(x) = \sum_k c_k \psi_N^k \qquad \sum k^4 |c_k|^2 < \infty \tag{49}$$

The domain of Δ_N are the functions with square integrable second derivatives and with first derivative vanishing at the boundary. Remark that the eigenfunctions of Δ_D *do not belong to the domain* of Δ_N (and conversely) in spite of the fact that these operators have a dense common domain, e.g. the twice differentiable functions with support strictly contained in $(0, 2\pi)$.

The operators $-\Delta_N$ and $-\Delta_D$ are not squares of operators but there is a relation with the symmetric non-self-adjoint operator ∂ defined to be $-i\frac{d}{dx}$ with domain the absolutely continuous functions which vanish in a neighborhood of 0 and of 2π. Notice that the adjoint ∂^* is the operator $i\frac{d}{dx}$ with domain the absolutely continuous functions on $[0, 2\pi]$.

Consider the operator $\partial^*\partial$. A function f in its domain is twice differentiable and its first derivative vanishes in 0 and 2π since f must belong to the domain of ∂ and ∂f must belong to the domain of ∂^*. On these functions $\partial^*\partial$ acts as $-\frac{d^2}{dx^2}$. It follows that the closure of $\partial^*\partial$ is $-\Delta_D$. In the same way one verifies that the closure of $\partial\ \partial^*$ is $-\Delta_N$.

This is a particular case of the following statement: if A is a closed operator the operation A^*A defines an essentially self-adjoint operator.

6 Self-adjoint Extensions of the Laplacian on a Locally Finite Metric Graph

Example 4 can be generalized without difficulty to the operator $-\frac{d^2}{dx^2} + V$ defined on a metric graph (we require that it be metric in order to define the differential operator $\frac{d}{dx}$).

We will consider only graphs which are concretely realized in R^2 as a set of vertices connected by segments (edges). In general a finite (or locally finite) graph is a quadruple

$$\mathcal{G} = \{\mathcal{V}\ , \mathcal{I}\ , \mathcal{E}\ , \partial\} \tag{50}$$

where $\mathcal{V}$ is a finite (or locally finite) set of *vertices*, $\mathcal{I}$ is a (locally) finite set of *oriented internal edges* and $\mathcal{E}$ is a locally finite set of *external edges*.

The map ∂ assigns to each internal edge i^k an ordered pair of vertices denoted by $\partial(i^k) = \{v_1^k, v_2^k\},\ k = \pm$ (they need not be different). We shall call $v_1 \equiv \partial(i^+)$ *initial vertex* of the i segment and $v_2 \equiv \partial(i^-)$ *final vertex*. We shall denote by $\partial(e)$ the vertex of the semi-infinite edges e.

A graph is *compact* if it has no external edge. We shall always assume that the graph is connected. The *degree* of the vertex v is by definition the number of edges incident to the vertex; an edge that has the same vertex as final and initial point (a lace) counts twice. We consider *metric graphs* i.e. we associate to each edge an interval $(0, a) \in R^+, a > 0$, a Hilbert space $\mathcal{H} = L^2(0, a)$ and a symmetric operator $\frac{d^2}{dx^2}$. In the case of external edges one has $a = \infty$.

We will consider only the case in which all edges are rectilinear. One can generalize without difficulty to edges which arbitrary C^1 curves (the Laplacian on the edge is then substituted by the Laplace-Beltrami operator) and one can add a potential and/or consider magnetic Laplacians.

Self-adjoint extensions are obtained by choosing *boundary conditions* at the vertices; the boundary conditions are *local* when they refer to a single vertex, non-local otherwise.

Given a metric graph, consider the Hilbert space

$$\mathcal{H}(\mathcal{G}) = \mathcal{H}_{\mathcal{E}} \oplus \mathcal{H}_{\mathcal{I}} \quad \mathcal{H}_{\mathcal{E}} = \oplus_{e\in\mathcal{E}}\mathcal{H}_e \quad \mathcal{H}_{\mathcal{I}} = \oplus_{i\in\mathcal{I}}\mathcal{H}_i \tag{51}$$

where $\mathcal{H}_j = L^2(I_j)$ and $I_j = [0, a_j]$ for a finite edge and $I_j = [0, \infty]$ for a semi-finite one.

Consider on the graph the operator Δ^0 defined for each edge as

$$(\Delta^0\phi)(x) = -\frac{d^2\phi(x)}{dx^2} + \phi(x), \quad x \in I_j \quad j \in \mathcal{I} \cup \mathcal{E} \tag{52}$$

with domain the functions which in each edge are twice differentiable with continuous second derivative and each vanishes in a neighborhood of the vertices. It easy to see that Δ^0 is a symmetric operator with deficiency index $m_\pm = 2|\mathcal{I}| + |\mathcal{E}|$.

To study its self-adjoint extensions we introduce the defect space $\mathcal{K}$

$$\mathcal{K} = \mathcal{K}_{\mathcal{E}} \oplus \mathcal{K}_{\mathcal{I}}^- \oplus \mathcal{K}_{\mathcal{I}}^+ \tag{53}$$

i.e. the space of the *boundary values* on each vertex of the functions and of their derivatives along the edges in the C^1 closure of the domain of Δ^0. The space $\mathcal{K}$ is finite-dimensional if the graph is finite.

We will have

$$\mathcal{K}_{\mathcal{E}} \equiv C^{|\mathcal{E}|} \quad \mathcal{K}_{\mathcal{I}}^{\pm} \equiv C^{|\mathcal{I}|} \tag{54}$$

The self-adjoint extensions are classified by linear maps on $\mathcal{K}$. It is also convenient to introduce the space $\mathcal{K} \oplus \mathcal{K}'$ in which these maps are associated to symplectic forms (see e.g. [10, 13, 16]).

Denote by $\mathcal{J} \subset \mathcal{E} \cup \mathcal{I}$ a subset of the edges and consider in the cartesian product $\otimes_{j\in\mathcal{J}} I_j$ a function ϕ

$$\phi(x) \equiv \{\phi_j(x_j),\ j \in \mathcal{J}\} \tag{55}$$

If $\phi \in D \equiv \oplus_{j\in\mathcal{E}\cup\mathcal{I}} D_j$ we pose

$$\psi \oplus \psi' \in \mathcal{K} \oplus \mathcal{K}' \tag{56}$$

where ψ is the vector

$$\psi \equiv \{\phi_{\mathcal{E}}(0),\ \phi_{\mathcal{I}}(0),\ \phi_{\mathcal{I}}(a_i)\} \tag{57}$$

and ψ' is the vector

$$\psi' \equiv \{\phi'_{\mathcal{E}}(0),\ \phi'_{\mathcal{I}}(0),\ -\phi'_{\mathcal{I}}(a_i)\} \tag{58}$$

(the apex indicates first order inner derivative at the vertex).

Let A and B be linear maps of $\mathcal{K}$. Denote by $\{A, B\}$ the linear map on $\mathcal{K} \oplus \mathcal{K}'$ defined by

$$\{A, B\}(\xi \oplus \xi') = A\xi + B\xi' \tag{59}$$

and let us define $\mathcal{M}(A, B) \equiv Ker\{A, B\}$.

Proceeding as in Examples 3, 4 one shows that the boundary conditions described by the kernel of the map $\{A, B\}$ provide a maximal symmetric (and therefore self-adjoint) extension of Δ_0. Indeed this is the condition under which the transposed map $\{A, B\}^t$ has maximal rank (equal to $|\mathcal{E}| + 2|\mathcal{I}|$).

It is easy to see that if we regard these boundary values as symplectic variables, the linear boundary condition are interpreted as determining a lagrangian manifold and different manifolds are connected by linear symplectic transformations.

The maps $\{A, B\}$ and $\{A', B'\}$ give rise to the same self-adjoint extension if and only if there exists an invertible map $C : \mathcal{K} \to \mathcal{K}$ such that $A' = CA,\ B' = CB$. In particular if the matrix B is invertible on can take $B' = I$ and $A' = B^{-1}A$. This equivalence is reflected in the fact that in Krein's formulation only the values of the functions at the vertices occur in the writing of the quadratic form. The values of the normal derivatives cannot occur because the quadratic form is defined in the space H^1 and in this space the derivative of a function need not be defined (this is also the reason for choosing the symmetric operator to be $-\frac{d^2}{d^2} + 1$ instead of $-\frac{d^2}{d^2}$).

If the graph is finite, the spectrum is pure point; if at lest one edge is infinite the continuous part of the spectrum is absolutely continuous.

Consider the special case of *star graph* i.e. a graph with one vertex and N infinite edges and place the vertex at the origin. In this case the boundary conditions determine, for a given self-adjoint extension, the relation at the vertex among the limits of the function and their derivatives along the edges.

The $N \times N$ matrix that describes a specific extension depends on the basis chosen; if one chooses as basis the value of the function and its derivative along the edge the choice $A = \infty,\ \ B = 0$ characterize Neumann boundary condition ($\phi'_n = 0,\ \ n = 1, \dots N$) and the choice $A = 0,\ \ B = \infty$ characterize Dirichlet b.c ($\phi_n(0),\ n = 1 \dots N$). Other boundary conditions frequently used the Kirchhoff b.c.

$$\sum_{n=1}^{N} \phi'_n(0) = 0, \qquad \phi_n(0) = \phi_m(0) \quad n, m = 1 \dots N \tag{60}$$

(this is the same law that in electrical circuits implies the conservation of the current) and

$$\sum_{n=1}^{N} \phi'_n(0) = b\phi_n 0,\ b \neq 0 \qquad \phi_n(0) = \phi_m(0) \quad n, m = 1 \dots N$$

The classification of the extensions in not changed if one adds a regular potential on the edges.

7 Point Interactions on the Real Line

Consider R as the $R = R^+ \cup R^-$ i.e. as a graph with two edges meeting a the origin. On can consider the self-adjoint extension corresponding to the matrices $A,\ B$ (which in this case are 2×2 matrices). On the real line one can also define

$$H_0 \equiv -\frac{d^2}{dx^2},\ x \in R \quad D(H_0) = H^2(R)$$

This operator is characterized among all possible extensions of the Laplacian on the graph $R^+ \cup R^-$ by having in its domain only functions that are continuous at the origin together with their derivative.

One can describe *some* of the other extensions as *perturbations* of H_0 *localized at the origin*. Such perturbations cannot be described by a potential; their precise description is within the theory of quadratic forms.

Consider the quadratic form

$$Q_\alpha \equiv \int_R |\frac{d\phi}{dx}|^2 + |\phi|^2 + \alpha|\phi(0)|^2 \qquad D(Q_\alpha = H^1 \tag{61}$$

Since functions in the Sobolev space $\mathcal{H}^1$ are absolutely continuous, the form Q_α is closed; it is obviously bounded below. It therefore corresponds to a self-adjoint operator bounded below. A formal integration by parts leads to interpret Q_α as the quadratic form associated to the operator

$$-\frac{d^2\phi(x)}{dx^2} + \phi(x) + \alpha\delta(x)\phi(x) \tag{62}$$

This expression is formal since multiplication by a δ function is an operation that is not Kato small with respect to the Laplacian.

The structure of Krein space indicates that Q_α is the quadratic form of an extension of the symmetric operator $\frac{d^2}{dx^2}$ defined on twice differentiable function which vanish in a neighborhood of the origin. This extensions are parametrized by $\alpha \in R$ and correspond to the requirement that a function $\phi(x)$ in the domain be continuous at the origin and satisfies

$$lim_{\epsilon\to 0_+}[\phi'(\epsilon) - \phi'(-\epsilon)] = \alpha\phi(0)$$

In the same way one can define the δ' interaction by means of a quadratic form by requiring continuity of the derivative of the functions and discontinuity of the functions the first derivative proportional to β time the value of the function at the origin.

The construction can be generalized to define point interactions on a finite number of points on the real line, or on any smooth curve by using local coordinates (the definition we have given for the Laplacian applies to any strictly elliptic operator and divergence form).

The formal expression (62) justifies the name *point interaction* that is given to a system described by the quadratic form (61). This interaction was introduced by Fermi [5] to describe the interaction of thermal neutrons with uranium nuclei.

If one applies perturbation theory to (62) regarding the delta function as a potential, first order perturbation theory gives the correct result; the terms of next orders are divergent.

It is interesting to notice that this extension of the Laplacian defined in $R - 0$ can be recovered as limit in the norm resolvent sense of the Laplacian on the real line plus a regular potential $V_\epsilon(x)$ in the limit when $\epsilon \to 0$ *provided* the operator $-\frac{d^2}{dx^2} + 1 + V(x)$, with $V(x)$ supported in $[-a, a]$, $a > 0$ has a zero-energy resonance [2].

Conventionally one says that this operator on the real line has a zero-energy resonance if the Sturm-Liouville problem on $[-2a, +2a]$ with Neumann boundary conditions at $-a$ and at a has an eigenvalue zero; equivalently one may ask that the resolvent have a singularity at zero momentum in Fourier space.

We shall return to this subject in the second part of this Lecture, where we shall discuss the corresponding problem in two and three dimensions.

One can consider also an infinite number of points on R or on a smooth curve with the condition that the infimum of the distance between any two points be strictly positive. If on the contrary the sites on which the point interactions are placed have an accumulation point, the extension one obtains is a symmetric operator which is in general not self-adjoint; its defect space contains functions that are singular at the accumulation point.

Of special interest is the case of the Laplacian with point interaction of strength α_n located in the points $x_n \in \Lambda$ where Λ is a periodic lattice in R.

If the strength does not depend on the point on the lattice one has the Krönig-Penney's model of a one dimensional crystal. The corresponding self-adjoint operator has absolutely continuous band spectrum.

If the strengths α_n are independent and identically distributed random variables (e.g. with possible values zero or one with equal probability) one has an examples of Schrödinger operator with random potential. This is one of the first *random potentials* studied in detail [1]. It has pure point spectrum.

8 Laplacians with Boundary Conditions at Smooth Boundaries in R^3

A natural generalization of the Laplacian on [0, 1] (i.e. a one dimensional manifold with boundary) is an elliptic operators defined in the interior of a regular domain $\Omega \in R^d$, $d \geq 2$ with regular boundary $\partial\Omega$. In electrostatics this is the theory of boundary potentials and of boundary charges.

In this case the defect subspaces have infinite dimension and the self-adjoint extensions are classified by the relation between two classes of functions defined on the boundary. One class represents the boundary value at $\partial\Omega$ of the functions in the domain of the operator, the other class represents, roughly speaking, the boundary values of the normal derivatives. This corresponds to single layer and double layer potentials in electrostatic.

It is known from electrostatic that the analysis is more difficult if the boundary is not smooth (boundary with sharp corners or with spikes). We limit ourselves to the case in which the boundary is a surface of class C^2.

If $d = 3$ from the properties of the Sobolev space $H^2(R^3)$ (the Laplacian is a second order elliptic operator) every function in the domain of the Laplacian Δ_0 (defined as the closure of $\sum \frac{\partial^2}{\partial x_i^2}$ on C_0^2) has a boundary value on $\partial\Omega$ which is a function of class $H^{\frac{3}{2}}(\partial\Omega)$, and its normal derivative is locally of class $H^{\frac{1}{2}}(\partial\Omega)$. Recall that the Sobolev spaces on $\partial\Omega$ are defined using the corresponding Laplace-Beltrami operator.

Dirichlet boundary condition corresponds to the vanishing of the function at the boundary, Neumann boundary conditions corresponds to the vanishing of the normal derivative. Relations of mixed type (Robin boundary conditions) correspond to linear relations between the boundary function f and its normal derivative g given by

$g(x) = \int_{\partial\Omega} K(x,y) f(y) dy, x \in \partial\Omega$ where $K(x,y)$ is a regular kernel. If the relation are between the function and its normal derivative at the same point, the boundary condition is called *local*; in this case the notation *boundary conditions* must be understood in a generalized sense since a function which belongs to $H^{\frac{1}{2}}(R^3)$ may not be defined pointwise.

We give here some detail of this classical approach. For a complete analysis of this problem one can see e.g. [G02], [LM72].

We want to classify the self-adjoint extensions of the operator defined as $\Delta_0\phi \equiv \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$ on twice differentiable functions with support strictly contained in a bounded closed domain $\Omega \subset R^3$ with regular boundary $\partial\Omega$ of class C^2.

Denote by $H^m(\Omega)$ the Sobolev space spanned by functions that are square integrable together with their (distributional) derivatives of order smaller or equal to m. Denote by $H^s(\partial\Omega),\ s \in R$ the completion of the functions C^∞ on $\partial\Omega$ with respect to the scalar product

$$(f,g)_{H^s(\partial\Omega)} \equiv (f, (-\Delta_{L.B.} + 1)^s g)_{L^2(\partial\Omega)}$$

where $\Delta_{L.B.}(\partial\Omega)$ is the Laplace-Beltrami on $\partial\Omega$ (seen as a Riemann surface).

We can use the fact [15] that $(-\Delta_{L.B.} + 1)^{\frac{s}{2}}$ defines for every choice of r and of s a map

$$(-\Delta_{L.B.} + 1)^{\frac{s}{2}} \ : \ H^r(\partial\Omega) \to H^{r-s}(\partial\Omega) \tag{63}$$

We denote by γ_j the linear surjective operator

$$\gamma_j \ : \ H^2(\Omega) \to H^{\frac{3}{2}-j}(\partial\Omega), \quad j = 0, 1 \tag{64}$$

defined as bounded extension of

$$D(\hat{\gamma}_j) = C^\infty(\partial\bar{\Omega}), \quad \hat{\gamma}_j \equiv \frac{\partial_j \phi}{\partial n} \tag{65}$$

where $n(x)$ is the normal to $\partial\Omega$ in $x \in \partial\Omega$ directed towards the interior of Ω.

9 The Trace Operator

In the construction of the self-adjoint extensions of the Laplacian in Ω will have a main role the operators

$$\rho(\phi) \equiv \gamma_0, \quad H^2(\Omega) \to H^{\frac{3}{2}}(\partial\Omega) \tag{66}$$

(evaluation operator or *boundary trace*) and the operator

$$\tau(\phi) \equiv \gamma_1 \ : \quad H^2(\Omega) \to H^{\frac{1}{2}}(\partial\Omega) \tag{67}$$

(*normal derivative*).

Denote by A the Laplacian defined on function of class $C^2(\Omega)$ vanishing in a neighborhood of $\partial\Omega$. Consider also the closed symmetric operator Δ_{min} which is the extension of A to the functions that vanish at the boundary together with their normal derivative

$$D(\Delta_{min}) \equiv \{\phi \in H^2, \ : \ \rho(\phi) = \tau(\phi) = 0\} \tag{68}$$

and its adjoint Δ_{max}

$$D(\Delta_{max}) \equiv \{\phi \in H^2, : \ \Delta\phi \in L^2\} \tag{69}$$

With these notation, the Laplacian with Dirichlet boundary condition at $\partial\Omega$, denoted by Δ_D, has domain

$$D(\Delta_D) \equiv \{\phi \in D(\Delta_{max}) \equiv \{\phi \in H^2, \ \ \hat{\rho}\,(\phi) = 0\} \tag{70}$$

while the Laplacian with Neumann b.c. Δ_N has domain

$$D(\Delta_D) \equiv \{\phi \in D(\Delta_{max}) \equiv \{\phi \in H^2, \ \ \hat{\tau}\,(\phi) = 0\} \tag{71}$$

where $\hat{\rho}$ and $\hat{\tau}$ are extensions of ρ and of τ to the domain of Δ_{max}.

The Green functions of these extensions are known, in the Physics Literature, especially in the text on Electrostatics, as *simple layer potentials* resp. *double layer potentials* (boundary charges).

Let us consider the operator Λ

$$\Lambda \equiv (-\Delta_{L.B.} + 1)^{\frac{1}{2}} \ : \ H^s(\partial\Omega) \to H^{s-1}(\partial\Omega) \tag{72}$$

and define G_0 as

$$\langle G_0 u, \phi\rangle_{L^2(\Omega} \ = \ -\langle \Lambda u, \tau(\Delta_D)^{-1}\phi\rangle_{L^2(\partial\Omega)} \quad \forall u \in H^1(\partial\Omega) \tag{73}$$

It follows $G_0 = K\Lambda$ where $K \ : \ H^{+\frac{1}{2}} \to D(\Delta_{max})$ is the (Poisson) operator that solves

$$\Delta(Ku) = 0, \quad \hat{\rho}(Ku) = u \quad \forall u \in H^{+\frac{1}{2}}(\partial\Omega) \tag{74}$$

Introduce also the operator $G_z = G_0 - zR_zG_0$ that solves

$$\Delta G_z u = zu, \quad \hat{\rho}(G_z u) = \Lambda u, \quad u \in H^{\frac{1}{2}}(\partial\Omega) \tag{75}$$

and the operator $P \equiv \hat{\tau} K$, a continuous linear operator from $H^{+\frac{1}{2}}(\partial\Omega)$ to $H^{-\frac{3}{2}}(\partial\Omega)$ also called *Neumann-Dirichlet map*. This operator maps the solution of the inhomogeneous Dirichlet problem to the solution of the corresponding inhomogeneous Neumann problem.

By means of these operators one can construct an operator Γ_z that can be used to determine all self-adjoint extensions of the operator A (the restriction of $-\Delta_{min}$ to the twice differentiable functions with support in Ω_0). In analogy with the one-dimensional (in which case H_1 is spanned by the functions defined on the border) the resolvents of all self-adjoint extension of A are classified by a pair $\{\Pi,\ \Theta\}$ where Π is an orthogonal projection in $\mathcal{K} \equiv H^{\frac{1}{2}}(\partial\Omega)$ and Θ is a self-adjoint operator in the Hilbert space $\Pi\mathcal{K}$.

Let us denote by $\Delta_{\Pi,\Theta}$ the corresponding extension, with domain

$$D(\Delta_{\Pi,\Theta}) = \{\phi \in D(\Delta_{max}),\ \Lambda^{-1}\hat{\rho}(\phi) \in D(\Theta)\ \ \Pi\hat{\tau}_0(\phi) = \Theta\Lambda^{-1}\hat{\rho}(\phi)\} \quad (76)$$

where $\tau_0(\phi) = \tau(\Delta_D^{-1}\Delta_{max}\phi.)$

The resolvent of $\Delta_{\Pi,\Theta}$ can then be written in the form (analogous to Krein's)

$$(-\Delta_{\Pi,\Theta} + z)^{-1} = (-\Delta_D + z)^{-1} + G_z[\Theta + \Pi\tau(G_0 - G_z)\Pi]^{-1}\Pi G(\bar{z})^* \quad (77)$$

Particular cases are the operators with Dirichlet or Neumann boundary conditions and the operators given by Robin boundary conditions defined by maps

$$B\ :\ H^{\frac{3}{2}} \to H^{\frac{1}{2}} \quad (78)$$

that connect the trace in $x \in \partial\Omega$ of the function with the trace in the same point x of its normal derivative. More general extensions are obtained through integral kernels; this extends the theory to boundary surfaces with singular points (e.g. sharp corners or spikes). A detailed analysis of this problem can be found e.g. in [8, 17].

10 Boundary Triples

If the symmetric operator is the Laplacian in the complement in R^3 of a set Σ of codimension two or three, e.g. a curve or a point, the theory described above is no longer applicable and one must generalize its setting. This is done with the theory of boundary triple [6].

Notice that if the ambient space is R^d and if the set Σ has co-dimension at least three its capacity is zero (with respect to the Laplacian). In this case the only self-adjoint extension of the Laplacian restricted to the complement of Σ is the Laplacian defined on $H^2(R^d)$.

In the second part of the Lectures we shall be mainly interested in the case $d = 3$ and the co-dimension of Σ is either two or three, i.e. Σ is a collection of smooth curves

or collection of points. In this lecture we will exemplify the theory of boundary triples discussing the case in which Σ is a collection of points in R^3 (point interaction). The theory of boundary triples has found applications also to the case of extensions of strictly elliptic operators defined in domains with rough boundaries (e.g. containing spikes).

We start by giving a general formulation of this method.

Definition 2 (*boundary triples*) [9, 17] Let A be a symmetric operator in a Hilbert space $\mathcal{H}$. A triple $\{\mathcal{K}, \beta_1, \beta_2\}$, where $\mathcal{K}$ is a Hilbert space with scalar product $\langle ., . \rangle$ (the Krein space) and

$$\beta_1 \ : \ D(A^*) \to \mathcal{K}, \qquad \beta_1 \ : \ D(A^*) \to \mathcal{K} \tag{79}$$

are two linear surjective maps is called *boundary triple* for A^* if for $\phi, \ \psi \in D(A^*)$ one has

$$(\psi, A^*\phi) - (A^*\phi, \psi) = \langle \beta_1(\psi), \beta_2(\phi) \rangle - \langle \beta_2(\psi), \beta_1(\phi) \rangle \tag{80}$$

(i.e.) A^* defines a symplectic form on $\mathcal{K}$. ◇

The notation *boundary triples* comes from the fact that in the case of partial differential operators defined on an open domain Ω with *regular* boundary, the Hilbert space $\mathcal{K}$ is made of functions on $\partial\Omega$ and the two maps are respectively the restriction on the boundary of a function and of its normal derivative.

Definition 3 (*symmetric relation*) A closed subspace $\Lambda \in \mathcal{K} \oplus \mathcal{K}$ is said to define a closed *symmetric relation* if

$$\langle \zeta_1, \xi_2 \rangle = \langle \zeta_2, \xi_1 \rangle \quad \forall \{(\zeta_1, \zeta_2), (\xi_1, \xi_2)\} \in \Lambda \oplus \Lambda \tag{81}$$

A relation is called *self-adjoint* (maximal symmetric) if it is not contained in any other closed symmetric relation. ◇

The main result in boundary triple theory is the following theorem [7] which generalizes Krein's classification of the self-adjoint extensions by means of the corresponding quadratic forms.

Theorem 4 *The self-adjoint extensions of A are parametrized by the self-adjoint relations in $\mathcal{K} \oplus \mathcal{K}$. Every self-adjoint extension is obtained by restricting A^* to the subspace*

$$\{\phi \in D(A^*) \ : \ \langle \beta_1(\phi), \beta_2(\phi) \rangle \in \Lambda\} \tag{82}$$

where Λ is a self-adjoint relation and $\{\mathcal{K}, \beta_1, \beta_2\}$ is a boundary triple for A^*. ◇

Remark that the graph of a self-adjoint extension of A is a particular case of a self-adjoint relation. Every self-adjoint extension of A defines (not uniquely) a self-adjoint relation.

One can prove that every self-adjoint relation in $\mathcal{K} \oplus \mathcal{K}$ has the form $\mathcal{G}(\Theta) \oplus \Pi_{\mathcal{K}_0^\perp}$ where $\mathcal{K}_0 \subset \mathcal{K}$ is a closed subspace, Θ is a self-adjoint operator in $\mathcal{K}$ and $\mathcal{G}(\Theta)$ is its graph. We shall denote by A_Θ the self-adjoint extension corresponding to the operator Θ. The boundary triples are a generalization of the defect subspaces in von Neumann's theory. Denote by $P_\pm$ the orthogonal projection on the defect subspace $\mathcal{K}_\pm$, with U an isometry of $\mathcal{K}_+$ in $\mathcal{K}_-$ and define

$$\beta_1 \equiv i P_+ - i U P_- \qquad \beta_2 \equiv P_2 + U P_1 \qquad \mathcal{K} = \mathcal{K}_+$$

Then $\{\mathcal{K},\ \beta_1,\ \beta_2\}$ is a boundary triple.

This allows to transpose to the theory of boundary triples the results of von Neumann's and Krein's theories on the relation between the resolvents of different self-adjoint extensions.

We shall not prove Theorem 4. We only remark that in the case of the Laplacian defined on a domain Ω with regular boundary $\partial\Omega$, the space $\mathcal{K}_0^\perp$ can be taken to be the space of functions vanishing in neighborhood $\Sigma \subset \partial\Omega$ and Θ an operator from $H^{\frac{3}{2}}$ to $H^{\frac{1}{2}}$ with integral kernel

$$\Theta(x, y) \quad x \in \partial\Omega,\ \ y \in \partial\Omega - \Sigma \tag{83}$$

11 Weyl Function

The relation between the theory of boundary triples and Krein's method can be clarified by the introduction in both procedures of the *Weyl function*.

Weyl's function was introduced by Weyl [19] as a tool in the study of the Sturm-Liouville problem. It was generalized in [4] (see also [3]) to an arbitrary symmetric operator with equal defect indices

Definition 4 (*Weyl function*) Given the boundary triple $\mathcal{K},\ \beta_1,\ \beta_2$ for the operator A^*, the *Weyl function* $\Gamma^A(z)$ of A is the map

$$\Gamma^A \ :\ \rho(A) \to \mathcal{B}(\mathcal{K}) \qquad z \in C^+ \tag{84}$$

uniquely defined by

$$\Gamma^A(z)\beta_2\phi = \beta_1\phi \qquad \forall\phi \in Ker(A^* - z) \equiv \mathcal{K}_z \tag{85}$$

◇

Let A_0 be the extension of A characterized by the self-adjoint relation $\mathcal{G} = \{0\} \times \mathcal{K}$ (in the case of the Laplacian defined in a regular domain this extension corresponds to Dirichlet boundary condition).

For every $z \in R(A_0)$ the maps β_1 and β_2 are bijections of $\mathcal{K}_z$ on $\mathcal{K}$ and therefore the map G_z can be defined through

$$\Gamma^A(z) = \beta_1 G_z$$

From this definition one derives

$$\Gamma^A(z) - \Gamma^A(\bar{w}) = (z - \bar{w}) G_w^* G_z \tag{86}$$

We have chosen the symbol G because in practical cases it is a Green function. In Krein's notations, the function Γ^A is called $\mathcal{Q}$ function. In this theory, as in Krein's, the Weyl function can be used to describe the spectral properties of the self-adjoint extensions.

One has e.g. the following theorem [4].

Theorem 5

$$\begin{gathered} z \in \rho(A_\Theta) \cap \rho(A_0) \Leftrightarrow 0 \in \rho(\Theta + \Gamma(z)) \\ \lambda \in \sigma_a(A_\Theta) \cap \rho(A_0) \Leftrightarrow 0 \in \sigma_i(\Theta + \Gamma(\lambda)), \quad a = p, c, r \end{gathered} \tag{87}$$

where the symbols p, c, r correspond respectively to the point, absolutely continuous and residual spectrum. Moreover the Krein-type formula *is valid*

$$(-A_\Theta + z)^{-1} - (-A_0 + z)^{-1} = G_z(\Theta + \Gamma(z)^{-1} G_{\bar{z}}^* \tag{88}$$

Through (88) *one can derive the form of the resolvent of any self-adjoint extension from the knowledge of a single one.* ◇

We remark that if the auxiliary space $\mathcal{K}$ has finite dimension, then all self-adjoint extensions have the same continuous and residual spectra.

The theory of boundary triples is useful [17, 18] to construct all self-adjoint extensions of the symmetric restriction S of a self-adjoint operator A to a suitable subset of its domain; typically the restriction to the interior of a domain which has rough boundaries or excludes a lower dimensional set.

For example A could be the self-adjoint operator that extends with Dirichlet b.c. a symmetric operator S defined on $R^N - \Sigma$ where Σ is a subset of measure zero of R^N (typically a set of points or a manifold of codimension one). In this case $\mathcal{K}$ is a Hilbert space of functions on Σ and one defines in a natural way a bounded operator τ which associates to a function in the domain of A^* its boundary value.

12 Interaction Localized in N Points

We exemplify the method of boundary triples by applying it in the case of point interaction in R^3 with N centers. In this case the Hilbert space $\mathcal{K}$ is C^n and the operator τ (*trace at the boundary*) is the operator

$$\tau \ : \ H^2(R^3) \to C^N \quad \tau(\psi) \equiv \{\psi(x_1), ..\psi(x_N)\} \tag{89}$$

The resolvent of the self-adjoint operator Δ is

$$(-\Delta + z)^{-1}(x, y) = \frac{e^{-\sqrt{z}|x-y|}}{4\pi|x-y|}, \quad Re\sqrt{z} > 0 \tag{90}$$

and therefore the function G_z is

$$G_z \ : \ C^N \to L^2(R^3), \quad [G_z\rho](x) = \sum_{k=1}^{N} \frac{e^{\sqrt{z}|x-x_k|}}{4\pi|x-x_k|}\rho_k \quad \rho = \{\rho_1, ..\rho_N\} \tag{91}$$

In this case ρ_k is the charge associated to the point x_k. From (91) one sees that the kth component of $(z-w)G^*_{\bar{w}}G_z\rho$ is

$$lim_{x\to x_k}[\frac{e^{-\sqrt{w}|x-x_k|} - e^{-\sqrt{z}|x-x_k|}}{4\pi|x-x_k|}\rho_k + \sum_{j\neq k}(\frac{e^{-\sqrt{w}|x-x_k|}}{4\pi|x_k-x_j|} - \frac{e^{-\sqrt{w}|x-x_k|}}{4\pi|x_k-x_j|})\rho_j]$$

Define

$$\Gamma_{k,k}(z) = \frac{\sqrt{z}}{4\pi}, \quad \Gamma_{k,j;k\neq j} = -\frac{e^{-\sqrt{z}|x_j-x_k|}}{4\pi|x_j-x_k|}$$

$$(\hat{\tau}\phi)_k = lim_{x\to x_k}[\phi(x) - \frac{\rho_k}{4\pi|x-x_k|}] \quad k = 1, ..N \tag{92}$$

(the regular part of the resolvent in x_k).

Introduce the orthogonal projection Π on the space $\mathcal{K} \equiv C^N$ and an operator $\Theta \in \mathcal{B}(\mathcal{K})$

$$\Pi\hat{\tau}\phi = \Theta\rho_\phi \tag{93}$$

The self-adjoint extensions S are then parametrized by Π, Θ according to

$$(-\Delta_{\Pi,\Theta}) = (-\Delta + z)^{-1} + G(z)\Pi[\Theta + \Pi\Gamma(z)\Pi]^{-1}\Pi G(\bar{z})^* \tag{94}$$

and their domain is characterized by the boundary condition (92).

With the same formalism one can treat the case in which the boundary is a curve γ in R^3. The defect space has infinite dimension and the functions in the domain of *some* of the self-adjoint extensions can be written as sum of a function in H^2 and of an integral on γ (with in general complex weight) of the Green functions. These

extensions are characterized by a linear relation between the *charge* on γ and the value taken by the function on γ. In this case the analysis is more cumbersome since the boundary singularities are logarithmic.

References

1. Aizenmann, M., & Molchanov, S. (1993). *Communications in Mathematical Physics*, *157*, 245–260.
2. Albeverio, S., & Hoegh-Krohn, R. (1981). *Journal Operator Theory*, *6*, 313–339.
3. Brashe, J., Malamud, M., & Neidhart, H. (2002). *Integral Equations and Operator Theory*, *43*, 264–289.
4. Derkad e, V. A., & Malamud, M. M. (1991). *Journal of Functional Analysis*, *95*, 1–95.
5. Fermi, E. (1936). *Ricerca Scientifica*, *7*, 13–52.
6. Gesztesy, F. (1980). One-dimensional Coulomb hamiltonian. *Journal of Physics A*, *13*, 867–875.
7. Gorbachuk, V. I., & Gorbachuk, M. L. (2003). *Boundary value problems for operator differential equations*. Dordrecht: Kluver.
8. Grubb, H. (1968). *Annali della Scuola Normale Superiore di Pisa*, *22*, 425–513.
9. Grubb, G. (2006). *Bulletin of the American Mathematical Society*, *43*, 227–230.
10. Kostrykin, V., & Schrader, R. (1999). Kirchoff's rules for Quantum Graphs. *Journal of Physics A*, *32*, 595–630.
11. Krein, M.G. (1946). *Doklady Akademii Nauk SSSR*, *43*, 339–342.
12. Krein, M.G. (1946). *Doklady Akademii Nauk SSSR*, *52*, 657–660.
13. Kuchement, P. (2004). *Waves and Random Media*, *14*, b107–128.
14. Kato, T. (1980). *Perturbation theory for linear operators*. Berlin: Springer.
15. Lions, J. L., & Magenes, E. (1972). *Non homogeneous boundary value problems*. Springer, London.
16. Novikov, S. (1999). Schroedinger operators on graphs and symplectic geometry. In: E. Bierstone (Ed.), *Proceedings of the Fields Institute Conference*.
17. Posilicano, A. (2008). *Operators and Matrices*, *2*, 109–147.
18. Posilicano, A. (2001). *Journal of Functional Analysis*, *183*, 109–147.
19. Weyl, H. (1910). *Matematische Annalen*, *68*, 220–269.

Index

G. Dell'Antonio, *Lectures on the Mathematics of Quantum Mechanics I*,
Atlantis Studies in Mathematical Physics: Theory and Applications 1,
DOI 10.2991/978-94-6239-118-5