Chapter 0

Introduction

Quantum Mechanics deserves the interest of mathematicians not only because it is a very important physical theory, which governs all microphysics, i.e. the physical phenomena at the microscopic scale of 10^{-8} cm., but also because it turned out to be at the root of important developments of modern mathematics.

The first branch of modern mathematics which was strongly influenced by Quantum Mechanics is the theory of *Algebras of Operators in Hilbert Spaces*, the so-called Von Neumann Algebras, whose foundations are due to Von Neumann also in connection with his interests in Quantum Mechanics. The theory of *Von Neumann Algebras*, as well as the related theory of C^* *Algebras* is now a well developed branch of mathematics [Dixmier (1981), (1977); Kadison and Ringrose (1983); Stratila and Zsido' (1979); Takesaki (1979); Jorgensen and Muhly (1987); AMS Proc. (1982)].

A strictly related topic is the study of the representations of the *Weyl algebra*, equivalently of the unitary representations of the *Heisenberg group*, which is at the basis of the canonical formulation of Quantum Mechanics of systems with a finite number of degrees of freedom [Putnam (1967); Bratteli and Robinson (1981); Folland (1989); Garding and Wightman (1954)].

The Schroedinger equation, which governs the time evolution of quantum systems, has motivated the theory of the *Schroedinger Operators* and the *Theory of Scattering*, which are now robust chapters of the theory of partial differential equations [Cycon et al. (1987); Eastham and Kalf (1982); Graffi (1984); Holten and Jensen (1989); Hislop and Sigal (1996); Amrein et al. (1977); Lax and Philips (1989), (1976); La Vita and Marchand (1974); Pearson (1988); Perry (1983); Petkov (1989); Sigal (1983); Yafaev (1992)].

Finally, it is worthwhile to mention that, under general stability conditions, the quantum mechanical time evolution allows an analytic continuation to purely imaginary time and the so-obtained theories (uniquely determined by the real time theories) correspond to stochastic processes.

Such a deep relation between Quantum Mechanics and *Stochastic Processes* has been at the origin of important developments in the theory of stochastic processes like the Feynman-Kac formula, the theory of functional (or path) integration etc. [Blanchard et al. (1987); Chung and Williams (1982); Glimm and Jaffe (1987); Kac (1980); Roepstorff(1993); Simon (1979)].

Along the lines of the deep philosophical changes, which led from Classical Mechanics to Quantum Mechanics, quite recently new steps were taken in the frontier developments of classical analysis and geometry, giving rise to the corresponding non-commutative (or quantum) extensions. These new developments were given the names of *Quantum Calculus, Non-Commutative Integration, Non-Commutative Geometry, Non-Commutative Harmonic Analysis, Quantum Probability, etc.* Even the discovery of *Quantum Groups*, a rapidly growing theory, is due to the influence of Quantum Mechanics [Araki (1993); Biane (1995); Connes (1992), (1994), (1995); Kassel (1995); Kirillov (1995); Madore (1995); Manin (1988); Meyer (1992); Parthasarathy (1992)].

In conclusion, Quantum Mechanics, as a very important physical theory, was not only a source of concrete and special mathematical problems arising in the solution of particular physical problems, but also provided a body of general mathematical structures which strongly influenced the development of modern mathematics. To better appreciate this role, it may be worthwhile to recall the strict relation between the development of Mechanics and Mathematics. Indeed, the origin of Calculus or classical Analysis can be traced back to Newton and Leibnitz, who discovered a mathematical language for the foundations of mechanics. The underlying idea is that physical quantities are described by functions of space and time (and possibly of additional variables) and therefore the mathematical description of observable quantities is related to the *theory of functions* and classical analysis.

When in the XIX century a major problem of theoretical physics was the description of complex systems, with 10^{23} degrees of freedom, as required for the foundations of thermodynamics and statistical mechanics, it became clear that new mathematical ideas were needed; one could not reasonably think to consider a Cauchy problem for 10^{23} initial data. This led to abandon the idea that a physical state is described by a point in phase space and to rather describe a state as a probability measure on the phase space. In this way *probability theory and random variables* entered in a crucial and philosophically important way into the framework of theoretical physics, at the basis of Classical Statistical Mechanics.

The quantum mechanical revolution, which took place in the twenties and early thirties, realized that at the microscopic level it is no longer correct to pretend that the physical observable quantities are described by an abelian algebra of functions or of random variables. The Heisenberg analysis of physically realizable experiments on microscopic systems indicated that the measurement of an observable in general limits the precision by which another observable can be subsequently measured. The mathematical abstraction of this deep physical fact is the realization that the algebra of observables is not described by an algebra of functions, but rather by an *algebra of operators in a Hilbert space.* As mentioned before, the passage from the commutative structure of classical mechanics and/or of classical statistical mechanics to the non-commutative structure of quantum mechanics is the deep and crucial feature shared by the modern non-commutative extension of calculus, probability, geometry etc.

Last but not least, quantum mechanics had a dramatic impact on the development of mathematical logic, giving rise to the so-called *Quantum Logic* : whereas the lattice of propositions of classical logic has the structure of a Boolean algebra (equivalently that of a lattice of commutative projections), the lattice of quantum propositions is non-boolean and it corresponds to a lattice of non-commutative projections [Birkhoff and Von Neumann (1936); Beltrametti and Cassinelli (1981); Cohen (1989); Garden (1984); Hooker (1975); Pitowski (1989); Rèdei (1998)].

The aim of these lectures is to provide at least the flavor of the philosophical revolution induced by quantum mechanics concerning the mathematical description of physical systems. The lectures are primarily addressed to people interested in questions of principle and in the mathematical foundations of physical theories, also in view of the fertile mutual influence between theoretical physics and mathematics.

In order to make the ideas at the basis of quantum mechanics understandable also to people with a mathematical education but with no great familiarity with physics, we will reduce the detailed description of the many experimental facts which led to the crisis of classical mechanics to the minimum and will rather extract and emphasize the overall simple and profound message for the mathematical description of quantum systems.

Once the Heisenberg revolutionary discovery has been accepted, namely that there are intrinsic limitations to the precise measurements of physical quantities (Heisenberg's uncertainty relations) leading to the non abelianity of the algebra of observables, the whole mathematical structure of Quantum Mechanics follows as a theorem (Gelfand-Naimark): the states of a physical system are described by vectors of a Hilbert space and the observables by Hilbert space operators. Also the Schroedinger formulation of Quantum Mechanics in terms of wave functions follows from Von Neumann uniqueness theorem on the regular (irreducible) representations of the Weyl algebra.

Those who will hopefully find the subject sufficiently interesting and stimulating are warmly referred to standard textbooks to deepen the mathematical and logical structure of quantum mechanics and to appreciate its impact on the description of the physical world [Dirac (1958); Feynman et al.(1963); Heisenberg (1930); Jauch (1968); Von Neumann (1955); Mackey (1963); Piron (1976); Segal (1963)].

GENERAL REFERENCES

J. Dixmier, *Von Neumann Algebras*, North-Holland 1981

J. Dixmier, *C^*Algebras*, North-Holland 1977

R.V. Kadison and J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras*, Vol.I-IV, Academic Press 1983

S. Stratila and L. Zsido', *Lectures on Von Neumann Algebras*, Abacus Press 1979

M. Takesaki, *Theory of Operator Algebras*, Vol.I, Springer 1979

P.E.T. Jorgensen and P.S. Muhly eds., *Contemporary Mathematics Vol. 62, Operator Algebras and Mathematical Physics*, Am. Math. Soc. 1987

Proc. Symposia in Pure Mathematics, *Operator Algebras and Applications*, Vol.I,II, Am. Math. Soc. 1982

O. Bratteli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol.I, Springer 1979; Vol.II, Springer 1981

G.B. Folland, *Harmonic Analysis in Phase Space*, Princeton Univ. Press 1989

L. Garding and A.S. Wightman, Proc. Natl. Acad. Sci. USA, **40**, 617 (1954)

C.R. Putnam, *Commutation Properties of Hilbert Space Operators and Related Topics*, Springer 1967

H.L. Cycon et al., *Schroedinger Operators*, Springer 1987

M.S.P. Eastham and H. Kalf, *Schroedinger type operators with continuum spectra*, Pitman 1982

S. Graffi, *Schroedinger Operators*, CIME course 1984, Lect. Notes Math. 1159, Springer 1985

P.D. Hislop and I.M. Sigal, *Introduction to Spectral Theory with Applications to Schroedinger Operators*, Springer 1996

M. Schechter, *Operator Methods in Quantum Mechanivs*, North-Holland 1981

W.O. Amrein et al., *Scattering Theory in Quantum Mechanics. Physical Principles and Mathematical Methods*, W.A. Benjamin 1977

J.A. La Vita and J.P. Marchand eds., *Scattering Theory in Mathematical Physics*, D. Reidel 1974

P.D. Lax and R.S. Phillips, *Scattering Theory*, Academic Press 1989

P.D. Lax and R.S. Phillips, *Scattering Theory for Automorphic Functions*, Princeton Univ. Press (Ann. Math. Studies) 1976

D.B. Pearson, *Quantum Scattering and Spectral Theory*, Academic Press 1988

P.A. Perry, *Scattering Theory by Enss Method*, Chur (Math. Reports) 1983

I.M. Sigal, *Scattering Theory for Quantum Mechanical Systems: Rigorous Results*, Lect. Notes Math. 1011, Springer 1983

D.R. Yafaev, *Mathematical Scattering Theory*, Am. Math. Soc. 1992

Ph. Blanchard et al., *Mathematical and Physical Aspects of Stochastic Mechanics*, Springer 1987

K.L. Chung and R.J. Williams, *Introduction to Stochastic Integration*, Birkhauser 1982

J. Glimm and A. Jaffe, *Quantum Physics. A Functional Integral Point of View*, Springer 1987

M. Kac, *Integration in Function Spaces and Some of Its Applications*, Lezioni Fermiane, Scuola Normale Superiore Pisa 1980

G. Roepstorff, *Path Integral Approach to Quantum Physics. An Introduction*, Springer 1993

B. Simon, *Functional Integration and Quantum Physics*, Academic Press 1979

H. Araki, *Quantum and Non-Commutative Analysis: Past, Present and Future Perspectives*, Kluwer 1995

A. Connes, *Géométrie non commutative*, InterEdition Paris 1990; *Non-Commutative Geometry*, Academic Press 1994

A. Connes, *Mathématiques Quantiques. Geometrie Non-Commutative et Physique Quantique*, Soc. Math. France 1992

A. Connes, *Quantized Calculus and Applications*, in Proc. XIth Int. Cong. of Mathematical Physics, D. Jagolnitzer ed., International Press 1995, p.15; Jour. Math. Phys. **36**, 6194 (1995)

A. Connes, *Non-Commutative Geometry and Physics*, Les Houches Lectures 1992, in *Gravitation and Quantizations*, B. Julia and J. Zinn-Justin eds. North-Holland 1995

C. Kassel, *Quantum Groups*, Springer 1995

A.A. Kirillov, *Representation Theory and Non-Commutative Harmonic Analysis*, Encyclopedia Math. Sci. N. 59, Springer 1995

J. Madore, *Non-Commutative Differential Geometry and Its Physical Applications*, London Math. Soc. Lect. Notes 206, Cambridge Univ. Press 1995

Y.I. Manin, *Quantum Groups and Non-Commutative Geometry*, CRM 1988

P.-A. Meyer, *Quantum Probability for Probabilists*, Springer 1992

K.R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*, Birkhäuser 1992

P. Biane, in *Lectures on Probability Theory*, Lect. Notes Math. 1608, Springer 1995

G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936)

E.G. Beltrametti and G. Cassinelli, *The Logic of Quantum Mechanics*, Addison Wesley 1981

D.W. Cohen, *An Introduction to Hilbert Space and Quantum Logic*, Springer 1989

R.W. Garden, *Modern Logic and Quantum Mechanics*, A. Hilger 1984

C.A. Hooker ed., *The Logico-Algebraic Approach to Quantum Mechanics*, D. Reidel 1975

I. Pitowsky, *Quantum Probability and Quantum Logic*, Lect. Notes Phys. 321, Springer 1989

M. Rédei, *Quantum Logic in Algebraic Approach*, Kluwer 1998

P.A.M. Dirac, *The Principles of Quantum Mechanics*, Oxford Claredon Press 1958

R.P. Feynman et al., *The Feynman Lectures on Physics*, Vol.III, Addison Wesley 1963

W. Heisenberg, *The Physical Principles of the Quantum Theory*, Dover 1930

J.M. Jauch, *Foundations of Quantum Mechanics*, Addison Wesley 1968

J. Von Neumann, *Mathematical Foundations of Quantum mechanics*, Princeton University Press 1955

G.W. Mackey, *Mathematical Foundations of Quantum Mechanics*, W.A. Benjamin 1963

C. Piron, *Foundations of Quantum Physics*, W.A. Benjamin 1976

I.E. Segal, Ann. Math. (2), **48**, 930 (1947); *Mathematical Problems of Relativistic Physics*, Am. Math. Soc. 1963

Chapter 1

Mathematical description of a physical system

1.1 Atomic physics and the crisis of classical mechanics

Quantum mechanics was invented on the basis of very cogent physical reasons. The large body of physical motivations and experimental facts, which are usually regarded as convincing enough by physicists, may appear not sufficiently forcing to mathematicians, also in view of the fact that the philosophical change involved is rather dramatic. To make the message more direct we will only briefly recall the basic experimental facts, which led to the crisis of classical mechanics for the description of atomic physics; we will rather dwell on the logical consequences of Heisenberg uncertainty relations and on the new mathematical structures which follow from them.

We list some of the most important phenomena in conflict with classical physics:

1) **Atomic physics**. There is a host of experimental evidence that an atom consists of a nucleus, made of neutrons and protons, and of electrons bound to it in a sort of planetary system, with the Coulomb potential playing the role of the gravitational potential. For example, the hydrogen atom has a nucleus made of a proton, of positive charge e, and an electron (of negative charge -e); the mass of the proton is about 1800 times the mass of the electron $m_e \simeq 10^{-27}$ gr. Such a planetary picture, which on one side is strongly supported by experimental data (typically Rutherford's experiment with α particles) leads to the following classical paradoxes:
i) all atoms have approximately the same dimensions $\simeq 10^{-8}$ cm., whereas classically the dimension of the orbit of a planet varies with the energy

(which can be arbitrary)
ii) the atoms are stable and therefore so must be the electron orbits, incompatibly with the laws of electrodynamics, according to which an accelerated charge emits electromagnetic radiation with inevitable energy loss. The electron should therefore collapse on the nucleus and a simple calculation shows that correspondingly the lifetime of an orbit of dimension 10^{-8} cm. should be of 10^{-10} sec.; in this case the dimensions of an atom would rapidly become those of its nucleus, namely of the order of 10^{-13} cm.
iii) the spectrum of the radiation absorbed or emitted by an atom, under the influence of external forces, consists of a discrete series of frequencies, contrary to the laws of classical physics, according to which the frequency ν of a planetary motion and therefore the radiation spectrum varies continuously with the dimensions of the orbit, for example for a circular orbit of radius r

$$\nu = (2\pi)^{-1}\sqrt{\frac{e^2}{m}}\, r^{-\frac{3}{2}}. \tag{1.1.1}$$

All this suggests that only a discrete set of orbits is allowed, equivalently that only a discrete set of frequencies of the electron (periodic) motion are allowed (*quantization of the electron periodic motion*).

2) **Photoelectric effect.** If light of short wavelength is sent on a metallic surface, electrons are emitted (roughly they gain enough energy to escape from being bounded inside the metal). Classically, one would expect that the phenomenon is crucially governed by the energy carried by the electromagnetic radiation, i.e. by the light intensity. On the contrary, what happens is that the crucial quantity is the frequency ν of the electromagnetic wave; indeed
i) the electron emission occurs only if $\nu > \nu_0 =$ *frequency threshold* (which depends on the metal),
ii) the maximum kinetic energy T of the emitted electrons is a function of the frequency,

$$T = a(\nu - \nu_0),$$

a a positive constant, rather than a function of the intensity of the radiation,
iii) the effect of the intensity (for fixed frequency ν) is only that of increasing the number of emitted electrons (not their energy !).

As argued by Einstein, this suggests that at the microscopic level the electromagnetic radiation of frequency ν does not carry energy in a continuous way, proportional to the radiation intensity, but rather in discrete fractions called *quanta* or *photons* each with energy

$$E = h\,\nu, \tag{1.1.2}$$

where $h = 6.6\ 10^{-27}$ erg sec is the Planck constant (and with momentum $p = h\nu/c$), and that the probability of more than one photon absorption by one electron is depressed. For fixed frequency, the intensity of the radiation is proportional to the number of photons carried by it and not to the energy of each photon (*quantum character of the electromagnetic radiation*). In this case, this can be interpreted as the evidence for the *particle or corpuscular behaviour* of *light* at the microscopic level.

3) **Particle-Wave duality of matter**. The above corpuscular behaviour of light (related to the way energy is carried by e.m. radiation) does not require a corpuscular localization of photons. In fact, in general they are not strictly localized in space; they are described by "wave functions" and this explains why interference and diffraction phenomena characterize the electromagnetic waves (*particle-wave duality of photons*). For example, in Bragg's experiment, if one sends a beam of light of wavelength λ on the plane surface of a crystal made of lattice planes at distance d, then, if θ is the incidence angle of the light beam on the crystal surface, the optical paths of two rays reflected by two lattice planes at distance d differ by the amount $2d\sin\theta$. Therefore, one has constructive interference for the reflected rays if $2d\sin\theta = n\,\lambda$, $n \in \mathbf{N}$.

A similar experiment was done by Davisson and Germer replacing the light beam by a well focused beam of electrons, all of (approximately) the same energy E. The result was a constructive interference if $2d\sin\theta = nh/p$, where $p = \sqrt{2mE}$ is the electron momentum. This indicates a wave behaviour of matter with wave length λ given by the De Broglie relation

$$\lambda = h/p = h/\sqrt{2mE} \tag{1.1.3}$$

(*particle-wave duality of matter*).

There are other important experimental facts which played a relevant role in the birth of Quantum Mechanics, like the *black-body radiation* and the temperature dependence of the *specific heats of gases*, but their discussion would lead us too far.[1]

The above sketchy account of the experimental facts, which led to the crisis of Classical Mechanics, may not provide a convincing evidence for the need of radical changes, especially if one is not familiar with the sharp constraints of classical physics. The implications of the above experiments at the level of general strategies may not appear logically inevitable, but a critical analysis would actually show that there is no alternative to changing the roots of classical physics.[2] Regrettably, we have to omit a detailed

[1] A discussion of the experiments at the basis of quantum mechanics can be found in many textbooks, see e.g. A. Messiah, *Quantum Mechanics*, North-Holland 1961 Vol. I, Chaps. I-III; S-I. Tomonaga, *Quantum Mechanics*, North-Holland 1962, Vol. I, Chap. 1,2; M. Born, *Atomic Physics*, Blakie 1958, Chap. VIII, Sects. 1-3.

[2] See e.g. J.A. Wheeler and W.H. Zurek, *Quantum Theory and Measurement*, Princeton University Press, 1983; A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer 1993.

discussion of these points, also because it would rely on a non-superficial mastering of physical arguments. Rather, we will follow the simpler logic of showing that the foundations of quantum mechanics can be deduced essentially by a single crucial fact, namely the *Heisenberg's realization of the uncertainty relations*, which affect the measurement of physical quantities at the microscopic level.

To fully appreciate the strength of Heisenberg intuition, we will start by a general revisitation of the mathematical structures and ideas underlying the foundations of classical mechanics.

1.2 Mathematical description of classical Hamiltonian systems

In order to realize the roots of the conflict between atomic physics and classical physics, we will isolate the basic structure underlying the mathematical description of a classical mechanical system.

Kinematics. The *configuration* or the *state* of a *classical Hamiltonian system* is (assumed to be) described by a set of canonical variables $\{q, p\}$, $q = (q_1, ..., q_n)$, $p = (p_1, ..., p_n)$, briefly by a point $P = \{q, p\} \in \Gamma \equiv$ phase space manifold. For simplicity, in the following, we will confine our discussion to the case in which Γ is compact. This is, e.g., the case in which the system is confined in a bounded region of space and the energy is bounded.

The physical quantities or *observables* of the system, clearly include the q's and p's and their polynomials and therefore, without loss of generality, we can consider as observables their sup-norm closure, i.e. (real) continuous functions $f(q,p) \in C_{\mathbf{R}}(\Gamma)$, (for a further extension see the remark after eq. (1.2.3)).

Every state P determines the values of the observables on that state and conversely, by the Stone-Weierstrass and Urysohn theorems, any $P \in \Gamma$ is uniquely determined by the values of all the observables on it (duality relation between states and observables).

Dynamics. The relation between the measurement of an observable f at an initial time t_0 and at any subsequent time t is given by the time evolution of the canonical variables

$$q \to q_t = q(t, q, p), \quad p \to p_t = p(t, q, p), \qquad f_t(q, p) \equiv f(q_t, p_t). \tag{1.2.1}$$

The time evolution of the canonical variables is given by the Hamilton equations

$$\dot{q} = \frac{\partial H}{\partial p}, \; \dot{p} = -\frac{\partial H}{\partial q}, \tag{1.2.2}$$

where $H = H(q,p)$ is the Hamiltonian. Under general conditions (typically if grad$H$ is Lipschitz continuous), for any initial data, the above system of equations has a unique solution local in time, which can be extended to all times under general conditions, e.g. if the surfaces of constant energy are compact [3].

The mathematical structures involved are the theory of functions (on phase space manifolds) and the theory of first order differential equations, defined by Hamiltonian flows on phase space manifolds.

From the above picture of elementary Hamiltonian mechanics one can extract the following algebraic structure.

I. **Algebraic properties of the classical observables**. The observable quantities associated to a classical system generate an abelian algebra $\mathcal{A}$ of real or more generally complex [4] continuous functions on the (compact) phase space (the product being given by the pointwise composition of functions $(fg)(x) = f(x)g(x)$ etc.). This algebra has an identity $\mathbf{1}$ given by the function $f = 1$ and a natural involution or $\star$ operation is defined by the ordinary complex conjugation, $f^\star(x) = \bar{f}(x)$, so that $\mathcal{A}$ is a *$\star$-algebra with identity.* To each element $f \in \mathcal{A}$ one can assign a norm, $\| f \|$, given by the sup-norm

$$\| f \| = \sup_{x \in \Gamma} | f(x) |, \tag{1.2.3}$$

so that $\mathcal{A}$ is a Banach space with respect to this norm. The product is continuous with respect to the norm topology since

$$\| fg \| \leq \| f \| \, \| g \| \tag{1.2.4}$$

and therefore $\mathcal{A}$ is a *Banach $\star$-algebra.* Finally, the following property holds (*C^*-condition*)

$$\| f^\star f \| = \| f \|^2. \tag{1.2.5}$$

Technically, an algebra with the above properties is called an *abelian C^*-algebra.*[5]

II. **States as linear functionals**. From an operational point of view, the identification of the states of a classical system with points of the phase space Γ relies on the unrealistic idealization, according to which the configuration of the system is sharply defined by measuring the canonical variables (typically positions and velocities) with infinite precision. Clearly, from a

[3] For a discussion of the existence theorems see V. Arnold, *Ordinary Differential Equations*, Springer 1992, and V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer 1989

[4] Such extension is both natural and convenient and in any case completely determined by the real subalgebra.

[5] For a beautiful account of the theory of abelian C^*-algebras see I.M. Gelfand, D.A. Raikov and G.E. Shilov, *Commutative Normed Rings*, Chelsea 1964; see also R.S. Doran and V.A. Belfi, *Characterization of C^*-Algebras*, Dekker 1986, esp. Chap. 2.

physical point of view it is more sensible to admit that in the preparation or detection of a state of a physical system a certain undeterminacy is unavoidable so that the configuration of the system at the initial time t_0 is known within a certain error, which inevitably propagates in time.

Since a state of the system is characterized by the measurements of the observables in that state, it is convenient to recall the operational meaning of such a procedure.

As it is well known, measurements with infinite precision are not possible and therefore the standard experimental way of associating a value of an observable f to a state ω is to perform replicated measurements of f, $m_1^{(\omega)}(f)$, $m_2^{(\omega)}(f), ..., m_n^{(\omega)}(f)$, on the system in the given state ω or more generally on replicas of it and to compute the average

$$< f >_n^{(\omega)} \equiv [m_1^{(\omega)}(f) + m_2^{(\omega)}(f) + m_n^{(\omega)}(f)]/n.$$

The limit $n \to \infty$ (whose existence is part of the foundations of experimental physics) defines the *expectation of f* on the state ω

$$\omega(f) \equiv \lim_{n\to\infty} \; < f >_n^{(\omega)} \tag{1.2.6}$$

as *average of the results of measurements of f* in the state ω.

The coarseness affecting the measurements of f is given by

$$(\Delta_\omega f)^2 \equiv \omega((f - \omega(f))^2), \tag{1.2.7}$$

since it indicates how much the results of measurements of f in the given state ω depart from the average $\omega(f)$; it is also called the mean square deviation or the variance of f (relative to ω). More generally, all experimental information on the measurement of an observable f in the state ω are recorded in the expectations of the polynomials of f. [6]

This is the way the experimental results are recorded and the operational identification of a state of a physical system is therefore given by the set of expectation values of its observables. Since the expectation $\omega(f)$ of an observable f has the interpretation (and actually corresponds to the operational definition) of the average of the results of the measurements of f in the given state ω, it follows that such expectations are linear, i.e.

$$\omega(\lambda f_1 + \mu f_2) = \lambda \omega(f_1) + \mu \omega(f_2), \;\; \forall f_1, f_2 \in \mathcal{A}, \;\; \lambda, \mu \in \mathbf{C} \tag{1.2.8}$$

and satisfy the *positivity condition*, namely

$$\omega(f^* f) \geq 0, \;\; \forall f \in \mathcal{A}. \tag{1.2.9}$$

[6] Such characterization of the measurements of f is somewhat related to the moment problem, for which a bound on the expectations of the polynomials of f is provided by the scale bound of the experimental apparatus associated to the measurements of f (such strict relations between observables and apparatuses yielding their measurements will be further discussed and exploited in the next section).

The positivity condition ($\omega((A+B)^*(A+B)) \geq 0$) implies the validity of Cauchy-Schwarz' inequality

$$|\omega(A^*B)| \leq \omega(A^*A)^{1/2}\,\omega(B^*B)^{1/2}, \quad \forall A, B \in \mathcal{A}, \tag{1.2.10}$$

and therefore $\omega(\mathbf{1}) > 0$ unless ω is the trivial state ($\omega = 0$ on $\mathcal{A}$). Thus, without loss of generality, given a (non-trivial) state ω one may always normalize it: $\omega \to \omega_{norm} = \omega(\mathbf{1})^{-1}\omega$, so that $\omega_{norm}(\mathbf{1}) = 1$.

Thus, in conclusion and quite generally, a classical system is defined by the abelian C^*-algebra $\mathcal{A}$ of its observables and a *state* of a classical system is a *normalized positive linear functional* ω on $\mathcal{A}$. A state ω on a C^*-algebra of continuous functions $C(X)$ on a compact (Hausdorff) space X is automatically continuous and therefore by the Riesz-Markov representation theorem [7] it defines a unique (regular Borel) measure μ_ω on X such that

$$\omega(f) = \int_X f\, d\mu_\omega, \quad \mu_\omega(X) = \omega(\mathbf{1}) = 1, \tag{1.2.11}$$

so that the expectations have a probabilistic interpretation [8]. Thus, the operational characterization of a state of a physical system leads to its description by a probability distribution rather than by a point of the phase space and the observables get the meaning of *random variables.*

The above considerations support the description of observables by continuous functions and justify the possible extension of the concept of observable to the pointwise limits of continuous functions, almost everywhere with respect to μ_ω.

Clearly, the above general concept of state contains as a very special case the definition of state in elementary mechanics; in fact, if μ_{P_0} is the probability measure concentrated on the point $P_0 = \{q_0, p_0\}$, i.e. for any measurable set S, $\mu_{P_0}(S) = 1$ if $P_0 \in S$ and $= 0$ otherwise, (namely μ_{P_0} is a Dirac δ function), then the corresponding state ω_{P_0} is

$$\omega_{P_0}(f) = \int_\Gamma f\, d\mu_{P_0} = f(P_0), \tag{1.2.12}$$

i.e. it is described by the point P_0. Such states are also called *pure states* since they cannot be written as convex linear combinations of other states (see also the next Sections). Clearly, for a pure state ω the variance vanishes, i.e. f takes a sharp value f_ω in the state ω, in the sense that all the measurements yield the same result, $f_\omega = \omega(f)$, i.e. there is no dispersion (such states are also called *dispersion free states*).

[7] A simple discussion is in M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Academic Press, Vol. I (Functional Analysis), Chap. IV, Sect. 4.

[8] For an introduction to the theory of probability and in particular to the concept of random variable see, e.g. J. Lamperti, *Probability*, W.A.Benjamin 1966, esp. Chap. 1, and H.G. Tucker, *A Graduate Course in Probability*, Academic Press 1967, esp. Chaps. 1,2.

For the general states defined above the time evolution can be defined by duality in terms of the time evolution of the observables

$$\omega_t(f) \equiv \omega(f_t). \tag{1.2.13}$$

The above mathematical description of a state of a classical system is not only strongly suggested by operational arguments, concerning the measurement of physical quantities, but it is absolutely necessary for the mathematical and physical description of complex systems, i.e. when the number of degrees of freedom become very large, typically 10^{23} for thermodynamical systems. In this case, it is technically impossible to control an initial value problem for such a large number of variables and it is also physically unreasonable to measure all of them. Moreover, such idealistic description of a complex system is not what is required on physical grounds to account for the time evolution of physically realizable measurements. Thus, the mathematical and physical description of a complex system inevitably requires new mathematical ideas and structures with respect to those of classical analysis, namely the *theory of random variables*. This was indeed the revolutionary step taken by Boltzmann in laying the foundations of Classical Statistical Mechanics and in deriving classical thermodynamics from the mechanical properties of complex systems.

III. **Algebraic Dynamics.** Under general regularity conditions, in the concrete case of the canonical realization of a classical system, the time evolution $\{q, p\} \to \{q_t, p_t\}$ is continuous in time $t \in \mathbf{R}$, with a continuous dependence on the initial data at time t_0 and therefore it defines a one-parameter family of continuous invertible mappings $\alpha_{t_0,t}$ of $C(\Gamma)$ into itself,[9] which preserves all the algebraic relations, including the $*$-operation (by eq. (1.2.2) $\alpha_t(f\,g) = \alpha_t(f)\,\alpha_t(g)$, $\alpha_t(f^*) = (\alpha_t(f))^*$). A linear invertible mapping of a C^*-algebra into itself, which preserves the algebraic relations is called a $*$-automorphism (it follows from a general result that it necessarily preserves the norm, see e.g. Proposition 2.2.3 in the next chapter).

Quite generally, given an abelian C^*-algebra $\mathcal{A}$ of observables a time-translation invariant (reversible) dynamics, (i.e. one which depends only on the difference $t - t_0$), can be algebraically defined as a one-parameter group of *-automorphisms α_t of $\mathcal{A}$, $t \in \mathbf{R}$ and by duality one can define a one-parameter group of transformations α_t^* of states into states given by

$$\omega_t(A) \equiv (\alpha_t^* \omega)(A) \equiv \omega(\alpha_t(A)), \tag{1.2.14}$$

[9]In the case of non regular dynamics, i.e. when the algebra of continuous functions on Γ is not stable under time evolution, one has to identify the observables with a larger C^*-algebra, than that of the continuous functions on Γ, in order to guarantee stability under time evolution (a necessary requirement for a reasonable physical interpretation).

(*time evolution of the states*). The abstract version of the continuity in time is that for any state ω, $\omega(\alpha_t(A))$, is continuous in time $\forall A \in \mathcal{A}$; technically α_t is said to be *weakly continuous.*

The recognition of the above mathematical structure at the basis of the standard description of classical systems suggests an abstract characterization of a classical (Hamiltonian) system, with no a priori reference to the explicit realization in terms of canonical variables, phase space, continuous functions on the phase space, etc. In this perspective, since a physical system is described in terms of measurements of its observables, one may take the point of view that a classical system is *defined* by its physical properties, i.e. by the algebraic structure of the set of its measurable quantities or observables, which generate an abstract abelian C^*-algebra $\mathcal{A}$ with identity. The states of the system being fully characterized by the expectations of the observables are described by normalized positive linear functionals on $\mathcal{A}$ and the dynamics is a one-parameter group of $*$-automorphisms of $\mathcal{A}$.

It is important to mention that quite generally, by the Gelfand-Naimark representation theorem [10], an (abstract) abelian C^*-algebra $\mathcal{A}$ (with identity) is isometrically isomorphic to the algebra of complex continuous functions $C(X)$ on a compact Hausdorff topological space X, where X is intrinsically defined as the Gelfand spectrum of $\mathcal{A}$.

From the point of view of general philosophy, the picture emerging from the Gelfand theory of abelian C^*-algebras has far reaching consequences and it leads to a rather drastic change of perspective. In the standard description of a physical system the geometry comes first: one first specifies the coordinate space, (more generally a manifold or a Hausdorff topological space), which yields the geometrical description of the system, and *then* one considers the abelian algebra of continuous functions on that space. By the Gelfand theory the relation can be completely reversed: one may start from the abstract abelian C^*-algebra, which in the physical applications may be the abstract characterization of the observables, in the sense that it encodes the relations between the physical quantities of the system, and then one reconstructs the Hausdorff space such that the given C^*-algebra can be seen as the C^*-algebra of continuous functions on it. In this perspective, one may say that the algebra comes first, the geometry comes later. The total equivalence between the two possible points of view indicates a purely algebraic approach to geometry: compact Hausdorff spaces are described by abelian C^*-algebras with identity, whereas if the algebra does not have an identity one has a locally compact Hausdorff space.

Non-commutative geometry is the structure emerging when the algebra is non-commutative. [11]

---

[10] For the convenience of the reader a brief outline of the Gelfand-Naimark theory is given in Appendix B.

[11] A. Connes, *Non Commutative Geometry*, Academic Press 1994.

1.3 General mathematical description of a physical system

In this section we argue that the structure of C^*-algebra of observables and states is the suitable language for the mathematical description of a physical system in general (including the atomic systems), with no reference to classical mechanics and its standard paradigms. [12]

I. **Observables.** From an operational point of view, a physical system is defined by its physical properties, i.e. by the set $\mathcal{O}$ of the physical quantities (briefly called *observables*) which can be measured on it and by the relations between them. Each observable has to be understood as characterized by a concrete physical apparatus yielding its measurements.[13]

For any $A \in \mathcal{O}$ and $\lambda \in \mathbf{R}$, one can define the observable λA as the observable measured by rescaling the apparatus by λ. By similar considerations one justifies the existence of elementary functions of an observable like the powers (with the standard elementary properties): if $A \in \mathcal{O}$, A^2 may be interpreted as the observable associated with squaring the apparatus scale (equivalently by squaring the results of measurements). Similarly, one defines the powers A^m, and their products $A^m A^n \equiv A^{m+n}$. It follows from this definition that A^0 is the observable whose results of measurements always take the value 1, independently of the state on which the measurement is done.

In the same way, one defines a polynomial of A as the observable obtained by taking as the new apparatus scale the given polynomial function of the scale for A. An element $A \in \mathcal{O}$ is said to be *positive* if all the *results of measurements* of A are positive numbers. By the operational definition of elementary functions of A, this implies that (and it is actually equivalent to) A is of the form $A = B^2$, $B \in \mathcal{O}$.

II. **States.** A *state* ω *of a physical system* is characterized by the results of the measurements of the observables in the sense that the *average over the results of measurements* of an observable A, when the system is in a state ω, defines the *expectation* $\omega(A)$ and the state ω is completely characterized by

[12]The C^*-algebraic approach to classical and quantum physics has been pioneered by I. Segal, Ann. Math.(2) **48**, 930 (1947); *Mathematical Problems of Relativistic Physics*, Am. Math. Soc. 1963; see also P. Jordan, J. Von Neumann and E.P. Wigner, Ann. Math. **35**, 29 (1934) for an early proposal and G.G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*, Wiley-Interscience 1972, esp. Chap. 2, for a historical review and a systematic treatment.

[13]The possibility that two distinct experimental apparatuses effectively define the same observable will be discussed below. It is convenient to deal with dimensionless observables, whose measurements are defined as ratios with respect to a set of reference measurements (e.g. of length, mass etc.); such a choice of scale is implicit in each physical apparatus.

all its expectations $\omega(A)$, when A varies over $\mathcal{O}$; thus ω is a (real) *functional* on $\mathcal{O}$. By the operational definition of $\omega(A)$ it easily follows that ω is a homogeneous functional

$$\omega(\lambda A) = \lambda\,\omega(A)\ , \quad \forall\lambda \in \mathbf{R}, \tag{1.3.1}$$

and that $\omega(A^n + A^m) = \omega(A^n) + \omega(A^m)$.

The realization that the only operational way of characterizing a state is in terms of its expectations of the observables, requires that two states yielding the same expectations must be identified, i.e.

$$\omega_1(A) = \omega_2(A), \quad \forall A \in \mathcal{O}\ , \tag{1.3.2}$$

must imply $\omega_1 = \omega_2$, (briefly the observables separate the states).

On the other hand, if we put at the basis of the mathematical description of a physical system the fact that the experimental way of identifying an observable is in terms of its expectations on the states, then two observables A and B having the same expectations on all the states, $\omega(A) = \omega(B)$, $\forall\omega$, cannot be distinguished. Such property of the states, of completely characterizing the observables and their relations, can be viewed as a *completeness of the states* with respect to the observables. This means that the states define an equivalence relation, denoted by $\sim$, between the elements of $\mathcal{O}$: $A \sim B$, if $\omega(A) = \omega(B)$ for all the physical states ω, (for example two distinct experimental apparatuses may effectively define the same observable). Two equivalent observables must therefore be identified and in the following the set $\mathcal{O}$ of observables will always denote the corresponding set of equivalence classes.

The definition of the zeroth power of an observable A implies that for any physical state $\omega(A^0) = 1$, since all the results of measurements take the value 1 and therefore so does their average. Thus, all the zeroth powers of observables fall in the same equivalence class which will be called the identity and denoted by $\mathbf{1}$. Clearly, for any state

$$\omega(\mathbf{1}) = 1, \tag{1.3.3}$$

i.e. a state of a physical system is a *normalized functional* on $\mathcal{O}$.

From the existence of products of powers of an observable follows the existence of a product of $\mathbf{1}$ with any observable:

$$A\,\mathbf{1} = A\,A^0 = A = \mathbf{1}A.$$

As we have also seen in the previous section, a crucial *positivity* property must be satisfied by the states. Since the expectation $\omega(A)$ is the *average* over the results of measurements in the state ω, it follows that for any state ω, if A is positive, then

$$\omega(A) = \omega(B^2) \geq 0, \tag{1.3.4}$$

i.e. ω is a *normalized positive functional* on $\mathcal{O}$.

By the completeness property of the states in identifying the observables, all the properties of an observable A have to be encoded in its expectation values and therefore, in particular, the positivity of an observable A has to be equivalently described by the positivity of all its expectations, $\omega(A) \geq 0$ for all ω.

III. **C^*-algebraic structure.** Since, as discussed above, an observable A is defined in terms of a concrete experimental apparatus, which yields the numerical results of measurements in any state, and since each concrete experimental apparatus has inevitable limitations implying a scale bound independent of the state on which the measurement is performed, the results of measurements of A in the various states is a bounded set of numbers, with bound related to the scale bound of the associated experimental apparatus. To each observable A it is then natural to associate the finite bound

$$||A|| \equiv \sup_{\omega} |\omega(A)| < \infty. \tag{1.3.5}$$

Clearly, by the homogeneity of the states

$$||\lambda A|| = |\lambda| \, ||A||, \quad \forall \lambda \in \mathbf{R}. \tag{1.3.6}$$

Moreover, since the states separate the observables, $||A|| = 0$ implies $A = 0$.

From the definition of $||A||$ it follows that

$$||A^2|| = ||A||^2. \tag{1.3.7}$$

In fact, by definition, for any physical state ω, $\omega(||A||\mathbf{1} \pm A) \geq 0$, so that $||A||\mathbf{1} \pm A$ are both positive; then $(||A||\mathbf{1} - A)\,(||A||\mathbf{1} + A)$ is a positive polynomial of A and

$$||A||^2 - \omega\,(A^2) = \omega\,((||A||\mathbf{1} - A)(||A||\mathbf{1} + A)) \geq 0, \forall \omega \tag{1.3.8}$$

which implies $||A||^2 \geq ||A^2||$. On the other side, the positivity of

$$(||A||\mathbf{1} \pm A)^2 = ||A||^2 + A^2 \pm 2\,||A||A$$

implies that for any state ω

$$2\,||A||\,|\omega\,(A)| \leq ||A||^2 + \omega\,(A^2) \leq ||A||^2 + ||A^2|| \tag{1.3.9}$$

and therefore $||A||^2 \leq ||A^2||$.

The duality relation between observables and states allows to display and define linear structures in $\mathcal{O}$. We have already argued that the sum of polynomials of one observable has a well defined operational meaning;

actually for a larger class of pairs A, B, (e.g. the kinetic and the potential energy) the sum can be defined in the sense that there is an observable C such that

$$\omega(C) = \omega(A) + \omega(B) \equiv \omega(A+B), \quad \forall \omega \tag{1.3.10}$$

and (by the duality relation between states and observables) one may write $C = A+B \in \mathcal{O}$. For arbitrary pairs, A, B, the sum defined by the expectations (1.3.10) may not correspond to an element of $\mathcal{O}$ and therefore it leads to an extension of $\mathcal{O}$, (on which the states are positive linear functionals), for which the definition of the powers $(A+B)^n$ may not have a direct operational meaning. The possibility, adopted in the sequel, of introducing the powers of $A+B$ for *any* pair A, B, with the same algebraic properties of the powers of the generating observables and the extension of the states to them as *positive linear functionals* is therefore a non-trivial extrapolation over the strict physically motivated structure. The so-obtained extension of $\mathcal{O}$ will still be denoted by $\mathcal{O}$.

Clearly, from eqs. (1.3.5), (1.3.10), $||A+B||$ is well defined and

$$||A+B|| \leq ||A|| + ||B||. \tag{1.3.11}$$

Thus $||\ ||$ is a norm on $\mathcal{O}$, which becomes a pre-Banach space. Technically it is convenient to consider the norm completion of $\mathcal{O}$, so that by a standard procedure one gets a real Banach space.

By definition of the norm, any state of the physical system satisfies

$$|\omega(A)| \leq ||A||, \tag{1.3.12}$$

i.e. any state is continuous with respect to the topology defined by the norm (*norm topology*) and therefore it has a unique continuous extension to the norm completion of $\mathcal{O}$, hereafter still denoted by $\mathcal{O}$.

The powers of the sum $A+B$ allow to define the following symmetric product

$$A \circ B \equiv \frac{1}{2}\left((A+B)^2 - A^2 - B^2\right) = B \circ A, \tag{1.3.13}$$

which however is not guaranteed to be distributive and associative.

The so obtained structure is close to that advocated by Jordan, the so-called *Jordan Algebra*, [14] for which no topological structure is assumed, but

[14] P. Jordan, Zeit f. Phys. **80**, 285 (1933); L.J. Page, Jordan Algebras, in *Studies in Modern Algebra*, A.A. Albert ed., Prentice Hall 1963; N. Jacobson, *Structure and Representations of Jordan Algebras*, Am. Math. Soc. 1968; H. Upmeier, *Jordan Algebras in Analysis, Operator theory and Quantum Mechanics*, AMS 1980; H. Hanche-Olsen and E. Størmer, *Jordan Operator Algebras*, Pitman 1984.
A Jordan algebra is said to be special if the symmetric product arises from an associative product: $A \circ B = \frac{1}{2}(AB + BA)$: otherwise it is said to be exceptional. The analysis of exceptional Jordan algebras does not seem to have led anywhere for possible physical applications (for a general review see A.S. Wightman, Hilbert Sixth Problem: Mathematical Treatment of the Axioms of Physics, in *Proceedings of Symposia in Pure Mathematics*, Vol. 28, Am. Math. Soc. 1976).

the symmetric product is assumed to be distributive and weakly associative, i.e.

$$A^2 \circ (B \circ A) = (A^2 \circ B) \circ A.$$

The structure discussed above is rather close to that advocated by Segal for the description of quantum systems (*Segal system*), [15] for which additional continuity properties of the powers are assumed:
i) The square is continuous in the norm topology, i.e. $A_n \to A$ implies $A_n^2 \to A^2$,
ii) $||A^2 - B^2|| \leq \max\,(||A||^2, ||B||^2)$.
The Segal structure is recovered from the one discussed above under the (mild looking) assumption that the symmetric product is homogeneous, i.e.

$$A \circ (\lambda B) = \lambda (A \circ B), \quad \lambda \in \mathbf{R} \tag{1.3.14}$$

and then, by symmetry, also $(\lambda A) \circ B = \lambda(A \circ B)$. Such a property is certainly satisfied when A and B are polynomial functions of the same observable C, since then a (distributive and associative) product is defined and $A \circ B = \frac{1}{2}(AB + BA)$; the extension to the general case looks like a reasonable assumption.[16]

Now, homogeneity of the above product implies distributivity, (a property which is not assumed by Segal). In fact, eq. (1.3.13) gives

$$(A + B)^2 = A^2 + B^2 + 2A \circ B, \tag{1.3.15}$$

$$(A - B)^2 = A^2 + B^2 + 2A \circ (-B) = A^2 + B^2 - 2A \circ B$$

and therefore $A \circ B = \frac{1}{4}((A + B)^2 - (A - B)^2)$, as in Segal, and

$$A^2 + B^2 = \frac{1}{2}((A + B)^2 + (A - B)^2). \tag{1.3.16}$$

Then, by eq. (1.3.15)

$$2(A + B) \circ C - 2A \circ C - 2B \circ C$$

$$= [(A + B + C)^2 + A^2] + [B^2 + C^2] - [(A + B)^2 + (A + C)^2] - (B + C)^2,$$

and eq. (1.3.16) applied to the three sums of squares in square brackets gives the vanishing of the r.h.s.

Distributivity of the symmetric product and positivity of the states imply

$$0 \leq \omega\,((A + \lambda B)^2) = \omega\,((A + \lambda B) \circ (A + \lambda B))$$

$$= \omega\,(A^2) + \lambda^2 \omega\,(B^2) + 2\lambda\omega\,(A \circ B), \quad \forall \lambda \in \mathbf{R},$$

[15] I.Segal, Ann. Math.(2), **48**, 930 (1947).
[16] G.G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*, Wiley-Interscience 1972, pp. 44-47.

so that

$$|\omega\,(A \circ B)| \leq \omega\,(A^2)^{1/2}\,\omega\,(B^2)^{1/2} \tag{1.3.17}$$

and

$$||A \circ B|| \leq ||A||\ ||B||. \tag{1.3.18}$$

Now, the continuity properties i) assumed by Segal follow easily. In fact, from distributivity one has $A_n^2 - A^2 = (A_n + A) \circ (A_n - A)$ and therefore

$$||A_n^2 - A^2|| \leq ||A_n - A||\,(||A_n - A|| + ||2A||),$$

which implies i). Equation ii) follows from the positivity of the squares and of the states, from the definition of the norm and from eq. (1.3.7), since $\forall a, b \in \mathbf{R},\ |a^2 - b^2| \leq \max\,(a^2, b^2)$.

As shown by Segal, the above structure allows to recover most of the mathematical basis for the description of quantum systems, like the concept of compatible observables, the joint probability distribution for compatible observables etc. However, the mathematical language becomes easier if one makes the technical assumption that the so obtained Segal system $\mathcal{O}$ can be embedded in a complex extension $\mathcal{A}$ generated by complex linear combinations of elements of $\mathcal{O}$, such that
1) the symmetric product arises from an associative (but not necessarily commutative) product in $\mathcal{A}$, i.e. $\forall A, B \in \mathcal{O}$

$$A \circ B = \frac{1}{2}\,(AB + BA),$$

2) a $*$ operation is defined on $\mathcal{A}$ with the properties that $\forall A, B \in \mathcal{O}$, $\lambda, \mu \in \mathbf{C}$

$$(\lambda A + \mu B)^* = \bar{\lambda} A + \bar{\mu} B,$$

$$(AB)^* = BA,$$

3) $\forall A \in \mathcal{A}$, $A^* A$ is positive and the states can be extended from $\mathcal{O}$ to $\mathcal{A}$ by linearity as linear functionals, with the natural extension of positivity

$$\omega\,(A^* A) \geq 0, \quad \forall A \in \mathcal{A}\ ,$$

and with the properties

$$||AB|| = \sup_\omega |\omega(AB)| \leq ||A||\,||B||, \quad ||A^* A|| = ||A^*||\,||A||. \tag{1.3.19}$$

Positivity, $\omega\,((\lambda A + \mathbf{1}\,)^* (\lambda A + \mathbf{1}\,)) \geq 0$, implies

$$\omega\,(A^*) = \overline{\omega(A)}\ , \qquad ||A^*|| = ||A||, \quad \forall A \in \mathcal{A}\ . \tag{1.3.20}$$

The so-obtained extension $\mathcal{A}$ has the properties of a C^*-algebra with identity $\mathbf{1}$, $\mathcal{O}$ is the subset of $*$-invariant (also called *self-adjoint*) elements and $\mathcal{A}$ is generated by $\mathcal{O}$.[17]

Necessary and sufficient conditions have been given for the existence of such an extension, [18] but their physical interpretation is not transparent. Honestly, this should be regarded as a technically motivated assumption. A Segal system which allows such an extension is called special and exceptional otherwise; so far no one seems to have succeeded in giving an interesting physical application of exceptional Segal systems, which are actually very difficult to construct. In the following discussion, we shall therefore assume that the observables generate a special Segal system.

The arguments discussed in this section do not pretend to prove as a mathematical theorem that the general physical requirements on the set of observables necessarily lead to a C^*-algebraic structure, but they should provide sufficient motivations in favor of it. In any case, the above mathematical structure is by far more general than the concrete structure discussed in Sect.2 for classical systems. As we shall see, the mathematical setting of quantum mechanics can be derived with a very strict logic solely from the C^*-algebraic structure of the observables and the *operational information of non-commutativity codified by the Heisenberg uncertainty relations* (Sect. 2.1). In this way one has a (in our opinion better motivated) alternative to the Dirac-Von Neumann axiomatic setting, which can actually be derived through the GNS theorem 2.2.4, the Gelfand-Naimark theorem 2.3.1 and Von Neumann theorem 3.2.2. For these reasons we adopt the following mathematical framework:

1. A physical system is *defined* by its *C^*-algebra $\mathcal{A}$ of observables* (with identity).

2. The states of the given physical system are identified by the measurements of the observables, i.e. a *state* is a *normalized positive linear functional on $\mathcal{A}$*. The set $\mathcal{S}$ of physical states separates the observables, technically one says that $\mathcal{S}$ is *full*, and conversely the observables separate the states.

In the mathematical literature, given a C^*-algebra $\mathcal{A}$, any normalized positive linear functional on it is by definition a state; here we allow the

[17]For an introduction to C^*-algebras see e.g. M. Takesaki, *Theory of Operator Algebras*, Vol. I, Springer 1979, Chap. I; R.V. Kadison and J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras*, Vol. I, Academic Press 1983, Chap. 4; O. Bratteli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol. I, Springer 1987, Sects. 2.1- 2.3; the basic textbook is J. Dixmier, *C^*-algebras* , North-Holland 1977. For the convenience of the reader a few basic notions about C^*-algebras are presented in the Appendices.

[18]D. Lowdenslager, Proc. Amer. Math. Soc. **8**, 88 (1957). For physically motivated conditions see E.M. Alfsen and F.W. Schultz, *Geometry of State Spaces of Operator Algbras*, Birkhäuser 2003.

possibility that the set $\mathcal{S}$ of states with physical interpretation (briefly called physical states) is full but smaller than the set of all the normalized positive linear functionals on $\mathcal{A}$.

Quite generally, given an abstract C^*-algebra $\mathcal{A}$, with identity $\mathbf{1}$, a positive linear functional ω on $\mathcal{A}$ is necessarily continuous with respect to the topology of the preassigned norm which makes $\mathcal{A}$ a C^*-algebra (see Appendix C, Proposition 1.6.3):

$$|\omega(A)\ | \leq ||A||\, \omega\,(\mathbf{1}).$$

Also, $\omega(A) \geq 0, \forall \omega$ implies $A = B^*B$ (Proposition 1.6.2). In the above presentation, these properties were obtained on the basis of the operational definition of states and observables.

The spectrum $\sigma(A)$ of an element $A \in \mathcal{A}$ is the set of all λ such that $\lambda \mathbf{1} - A$ does not have a two-sided inverse in $\mathcal{A}$. This is the purely algebraic version of the standard definition of spectrum for operators in a Hilbert space. An element A is said to be *normal* if it commutes with its adjoint A^*. If A is normal and $\lambda \in \sigma(A)$, then there exists at least one positive linear functional ω such that $\omega(A) = \lambda$ (see Appendix C). Thus the spectrum of a normal element A is a set of possible expectations of A. This implies that the set of all positive linear functionals on $\mathcal{A}$ separate the normal elements of $\mathcal{A}$ and therefore all the elements of $\mathcal{A}$, since any A can be written as a complex linear combination of normal elements $A = (A + A^*)/2 - i(iA - iA^*)/2$.

1.4 Appendix A : C^*-algebras

For the convenience of the reader, in this and in the following Appendices we recall a few basic notions about C^*-algebras.

A **C^*-algebra** $\mathcal{A}$ is
i) a linear associative algebra over the field $\mathbf{C}$ of complex numbers, i.e. a vector space over $\mathbf{C}$ with an associative product linear in both factors,
ii) a normed space, i.e. a norm $|| \, ||$ is defined on $\mathcal{A}$:

$$||A|| \geq 0, \quad ||A|| = 0 \quad \Leftrightarrow \quad A = 0, \quad \forall A \in \mathcal{A} \ ,$$

$$||\lambda A|| = |\lambda| \, ||A||, \quad \forall \lambda \in \mathbf{C},$$

$$||A + B|| \leq ||A|| + ||B||, \quad \forall A, B \in \mathcal{A} \ ,$$

with respect to which the product is continuous:

$$||AB|| \leq ||A|| \, ||B||, \tag{1.4.1}$$

and $\mathcal{A}$ is a complete space with respect to the topology defined by the norm (thus $\mathcal{A}$ is a Banach algebra),
iii) a *-(Banach) algebra, i.e. there is an involution $* : \mathcal{A} \to \mathcal{A}$,

$$(A + B)^* = A^* + B^*, \quad (\lambda A)^* = \bar{\lambda} A^*, \quad (AB)^* = B^* A^*, \quad (A^*)^* = A,$$

iv) with the property (*C^*-condition*)

$$||A^* A|| = ||A||^2. \tag{1.4.2}$$

The C^*-condition implies that

$$||A^*|| = ||A||. \tag{1.4.3}$$

In fact,

$$||A||^2 = ||A^* A|| \leq ||A^*|| \, ||A||,$$

i.e. $||A|| \leq ||A^*||$; on the other side, since $A = (A^*)^*$, by the same argument $||A^*|| \leq ||A||$.

An element $A \in \mathcal{A}$ is said to be *normal* if it commutes with its adjoint A^*; in this case the C^*-condition implies

$$||A^2|| = ||A||^2. \tag{1.4.4}$$

In fact,

$$||A^2||^2 = ||(A^*)^2 A^2|| = ||(A^* A)^* A^* A|| = ||A^* A||^2 = ||A||^4,$$

where in the second equality the normality of A has been used.

The above equation further implies that for a normal element A

$$||A^m|| = ||A||^m, \quad \forall m \in \mathbf{N}. \tag{1.4.5}$$

In fact, by iteration of eq. (1.4.4) one gets

$$||A^{2^k}|| = ||A||^{2^k}, \quad \forall k \in \mathbf{N}$$

and given $m \in \mathbf{N}$, one can find an n such that $m + n = 2^k$, so that

$$||A||^m \, ||A||^n = ||A||^{m+n} = ||A^{m+n}|| \leq ||A^m|| \, ||A^n|| \leq ||A||^m \, ||A||^n.$$

This means that the inequalities are actually equalities and the above equation is proved.

Given an element A of an algebra $\mathcal{A}$ with identity $\mathbf{1}$, as always assumed, its *spectrum* $\sigma(A)$ is defined as the set of all complex numbers λ such that $A - \lambda\mathbf{1}$ does not have a (two sided) inverse in $\mathcal{A}$.

Proposition 1.4.1 (Spectral radius formula) *Let $\mathcal{A}$ be a Banach algebra and $A \in \mathcal{A}$, then $\sigma(A)$ is a compact not empty set and*

$$\sup_{\lambda \in \sigma(A)} |\lambda| = \lim_{n \to \infty} ||A^n||^{1/n} \leq ||A||. \tag{1.4.6}$$

If $\mathcal{A}$ is a C^*-algebra and A is normal, then the above inequality becomes an equality.

Proof. We start by showing that the limit on the r.h.s. exists. For this purpose, let

$$r \equiv \inf_{n \in \mathbf{N}} ||A^n||^{1/n}.$$

Clearly, $r \leq ||A^n||^{1/n} \leq ||A||$, $\forall n \in \mathbf{N}$, and therefore

$$r \leq \lim_{n \to \infty} \inf ||A^n||^{1/n}.$$

Now, let $\varepsilon > 0$ and choose m such that $||A^m||^{1/m} < r + \varepsilon$. For any $n \in \mathbf{N}, \exists k_n \in \mathbf{N}$ such that $n = k_n m + l_n$, $l_n \in \mathbf{N}$, $0 \leq l_n \leq m$; then

$$||A^n||^{1/n} = ||A^{k_n m} A^{l_n}||^{1/n} \leq ||A^m||^{k_n/n} ||A||^{l_n/n} \leq (r + \varepsilon)^{m k_n/n} ||A||^{l_n/n}.$$

By construction $\lim_{n \to \infty} m k_n/n = 1$, $\lim_{n \to \infty} l_n/n = 0$, so that

$$\lim_{n \to \infty} \sup ||A^n||^{1/n} \leq r + \varepsilon.$$

Since ε was arbitrary, one has

$$\lim_{n \to \infty} \sup ||A^n||^{1/n} = \lim_{n \to \infty} \inf ||A^n||^{1/n},$$

i.e. the limit exists. The equality for normal elements of a C^*-algebra follows trivially from eq. (1.4.5).

To conclude the proof, we note that the existence of a norm allows an extension of the standard analytic calculus to Banach algebras (see e.g. the quoted book by Gelfand, Raikov and Shilov); in particular, by the extension of the Cauchy-Hadamard theorem of elementary analytic calculus, r^{-1} is the radius of convergence of the series $\mathbf{1} + zA + z^2A^2 + ..., z \in \mathbf{C}, A \in \mathcal{A}$, which converges to $(\mathbf{1} - zA)^{-1}$ for $|z| < r^{-1}$ and has a singularity for $|z| = r^{-1}$. Thus, $(\mu\mathbf{1} - A)^{-1}$ exists if $|\mu| > r$, and $r = \sup_{\lambda\in\sigma(A)} |\lambda|$; moreover $\sigma(A)$ is closed because the analyticity domain $\mathbf{C}/\sigma(A)$ of $(A - z\mathbf{1})^{-1}$ is open.

By a similar argument, $\sigma(A)$ cannot be empty, otherwise $(\lambda\mathbf{1} - A)^{-1}$ would be an entire function in the whole complex λ plane, vanishing for $|\lambda| \to \infty$ and therefore zero everywhere, contrary to the existence of A^{-1} (implied by $\sigma(A) = \emptyset$).

The above Proposition implies that if all elements, except 0, of a Banach algebra $\mathcal{A}$ are invertible, then $\mathcal{A}$ is isomorphic to the complex numbers; in fact, if $\lambda \in \sigma(A) \neq \emptyset$, then $\lambda\mathbf{1} - A$ is not invertible and therefore it must be 0, i.e. $A = \lambda\mathbf{1}$ (*Gelfand-Mazur theorem*).

A family $\mathcal{F} = \{A_\alpha, \alpha \in I\}$ is said to *generate* a (normed) algebra $\mathcal{A}$ if the polynomials of $\mathcal{F}$ are dense in $\mathcal{A}$.

1.5 Appendix B: Abelian C^*-algebras

A C^*-algebra $\mathcal{A}$ is called *abelian* or *commutative* if the product is commutative.

Definition 1.5.1 *A* **multiplicative linear functional** *m on a commutative Banach algebra $\mathcal{A}$ is a homomorphism of $\mathcal{A}$ into* $\mathbf{C}$, *i.e. a mapping which preserves all the algebraic properties:*

$$m(AB) = m(A)m(B), \quad m(A + B) = m(A) + m(B). \tag{1.5.1}$$

Clearly $m(\mathbf{1}) = 1$, if $m \not\equiv 0$.

Definition 1.5.2 *A linear subspace I of an algebra $\mathcal{A}$ is called a* **left** *(respectively* **right***)* **ideal** *if it is stable under left (resp. right) multiplication by elements of $\mathcal{A}$. I is* **proper** *if it is properly contained in $\mathcal{A}$ ($I \neq \mathcal{A}$), and it is* **maximal** *if it is not properly contained in a proper ideal.*

For commutative algebras left and right ideals coincide and are simply called ideals. Clearly, if $\mathcal{A}$ has an identity $\mathbf{1}$, as always assumed, an ideal

I is proper iff $\mathbf{1} \notin I$, ($\mathbf{1} \in I \Rightarrow \mathcal{A}\ \mathbf{1} \subseteq I$, i.e. $I = \mathcal{A}$). Hence, for Banach algebras, the closure of a proper ideal I is a proper ideal (if $\mathbf{1} \in \bar{I}$, $\exists x \in I$, with $||\mathbf{1} - x|| < 1$, $x = \mathbf{1} - (\mathbf{1} - x)$ is invertible and $x^{-1}x = \mathbf{1} \in I$) and therefore maximal ideals are closed and each proper ideal is contained in a proper maximal ideal.

Proposition 1.5.3 *For a commutative Banach algebra $\mathcal{A}$ there is a one to one correspondence between the set $\Sigma(\mathcal{A})$ of multiplicative linear functionals and proper maximal ideals of $\mathcal{A}$.*

Furthermore, given $A \in \mathcal{A}$, $\lambda \in \sigma(A)$ iff there exists a multiplicative linear functional $m \in \Sigma(\mathcal{A})$ such that $m(A) = \lambda$.

Proof. Each $m \in \Sigma(\mathcal{A})$ defines a proper ideal $K \equiv \ker(m)$. Since $\forall [A], [B] \in \mathcal{A}/K$, $m([A]) = m([B])$ implies $[A] = [B]$, it follows that $\mathcal{A}/K$ is isomorphic to $\mathbf{C}$; it is therefore a field and then it cannot contain any proper ideal, since an invertible element A of a commutative Banach algebra cannot belong to any proper ideal I (otherwise $\mathbf{1} = AA^{-1} \in I$). This excludes the existence of a maximal proper ideal I properly containing K, because otherwise I/K would be a proper ideal of $\mathcal{A}/K$; thus K is maximal.

Conversely, given a maximal proper ideal K of $\mathcal{A}$, $\mathcal{A}/K$ is a Banach space (with $||[A]|| = \inf_{k \in K} ||A + k||$), since K is closed, and actually a Banach algebra, since K is an ideal. Furthermore, since K is maximal, $\mathcal{A}/K$ cannot contain a proper ideal, otherwise its inverse image in $\mathcal{A}$ would be a proper ideal which properly contains K. Thus, all elements of $\mathcal{A}/K$, except 0, are invertible, since a non-invertible element A of a commutative Banach algebra $\mathcal{A}$ belongs to the ideal $A\mathcal{A}$, which does not contain $\mathbf{1}$ and therefore is proper. Hence, by the Gelfand-Mazur theorem, $\mathcal{A}/K$ is isomorphic to $\mathbf{C}$, and the homomorphism $m : \mathcal{A} \to \mathcal{A}/K \to \mathbf{C}$ defines a unique multiplicative linear functional with $\ker(m) = K$.

For the second part of the Proposition, if $\lambda \in \sigma(A)$ then $\lambda\mathbf{1} - A$ is not invertible and therefore it belongs to the proper ideal $I \equiv (\lambda\mathbf{1} - A)\mathcal{A}$. Let J be a maximal ideal containing I and m_J the corresponding multiplicative linear functional, then $\ker(m_J) = J \supseteq I$ so that $m_J(\lambda\mathbf{1} - A) = 0$, i.e. $m_J(A) = \lambda$.

Conversely, if $\exists m$, with $m(A) = \lambda$, then $\lambda\mathbf{1} - A$ is not invertible i.e. $\lambda \in \sigma(A)$, since otherwise

$$1 = m(\mathbf{1}) = m((\lambda\mathbf{1} - A)(\lambda\mathbf{1} - A)^{-1}) = m(\lambda\mathbf{1} - A)\, m((\lambda\mathbf{1} - A)^{-1}) = 0.$$

Because of the above relation between $\Sigma(\mathcal{A})$ and the points of the spectra of the elements of $\mathcal{A}$, $\Sigma(\mathcal{A})$ is called *the (Gelfand) spectrum of $\mathcal{A}$*. Indeed, if $\mathcal{A}$ is generated by a single element A, i.e. the linear span of the powers of A is dense in $\mathcal{A}$, then $\Sigma(\mathcal{A}) = \sigma(A)$; in fact, the above Proposition, establishes a correspondence between $\sigma(A)$ and $\Sigma(\mathcal{A})$, which is actually one to one since if $m_1(A) = m_2(A)$, then m_1 and m_2 coincide on the polynomials of A and, since the latter are dense, on the whole of $\mathcal{A}$.

By the spectral radius formula, the above relation between $\Sigma(\mathcal{A})$ and $\sigma(A)$ implies that multiplicative linear functionals are continuous:

$$|m(A)| \leq \sup_{\lambda \in \sigma(A)} |\lambda| \leq ||A||,$$

with an equality on the right if $\mathcal{A}$ is a (abelian) C^*-algebra.

Proposition 1.5.4 *Let $\mathcal{A}$ be a C^*-algebra (with identity), then any bounded linear functional m on $\mathcal{A}$, with $m(\mathbf{1}) = 1 = ||m||$, satisfies*

$$m(A^*) = \overline{m(A)}. \tag{1.5.2}$$

Proof. First we prove that if $A = A^*$, then $m(A)$ is real. Indeed, putting $m(A) = a + i\,b,\ a, b \in \mathbf{R}$, we have $\forall\, c \in \mathbf{R}$

$$b^2 + c^2 + 2bc = |b + c|^2 \leq |a + i(b + c)|^2 = |m(A + ic\mathbf{1}\,)|^2$$

$$\leq ||A + ic\mathbf{1}\,||^2 = ||A^*A + c^2|| \leq ||A||^2 + c^2,$$

where the C^* condition has been used. The above inequality requires $b = 0$. Now, a generic $A \in \mathcal{A}$ can be written as a linear combination of self-adjoint elements: $A = (A + A^*)/2 - i(iA - iA^*)/2$, so that the result follows by linearity.

If $\mathcal{A}$ is a commutative C^*-algebra is generated by A and A^*, then by exploiting the above property of multiplicative linear functionals and the general argument for abelian Banach algebras, one has that $\Sigma(\mathcal{A}) = \sigma(A)$. More generally, if $\mathcal{A}$ is a commutative C^*-algebra generated by (algebraically independent) $A_1, A_2, ..., A_n, A^*{}_1, ..., A^*{}_n$, then

$$\Sigma(\mathcal{A}\ ) = \times_i\ \sigma(A_i).$$

Theorem 1.5.5 *(Gelfand-Naimark characterization of abelian C^*-algebras) An abelian C^*-algebra $\mathcal{A}$ (with identity) is isometrically isomorphic to the C^*-algebra of continuous functions on a compact Hausdorff topological space, which is the Gelfand spectrum of $\mathcal{A}$ with the topology induced by the weak * topology.*

Proof. By duality, each $A \in \mathcal{A}$ defines a function $\tilde{A}$ on $\Sigma(\mathcal{A})$, called the Gelfand transform of A, by $\tilde{A}(m) \equiv m(A)$ and clearly

$$(\tilde{A} + \tilde{B})(m) = \tilde{A}(m) + \tilde{B}(m), \quad \mu\tilde{A}(m) = \widetilde{\mu A}(m), \ \ \forall \mu \in \mathbf{C}, \tag{1.5.3}$$

$$(\tilde{A}\tilde{B})(m) = \widetilde{AB}(m) = \tilde{A}(m)\,\tilde{B}(m), \quad \overline{\tilde{A}(m)} = (\tilde{A^*})(m) \equiv (\tilde{A})^*(m). \tag{1.5.4}$$

Thus, the functions $\tilde{A}$, for $A \in \mathcal{A}$, form an abelian *-algebra $\tilde{\mathcal{A}}$. By the above Proposition, for each $m \in \Sigma(\mathcal{A})$, $m(A)$ is a point of $\sigma(A)$, which is a closed set, and

$$|\tilde{A}(m)| = |m(A)| \leq \sup_{\lambda \in \sigma(A)} |\lambda| = ||A||,$$

$$||\tilde{A}||_\infty \equiv \sup_{m \in \Sigma(\mathcal{A})} |\tilde{A}(m)| = \sup_{\lambda \in \sigma(A)} |\lambda| = ||A||. \tag{1.5.5}$$

Thus $\mathcal{A}$ is isometrically isomorphic to $\tilde{\mathcal{A}}$.

We shall now show that $\Sigma(\mathcal{A})$ is a compact topological space and that $\tilde{\mathcal{A}} = C(\Sigma(\mathcal{A}))$. For this purpose, we note that $\Sigma(\mathcal{A})$ is a closed subset of the closed unit ball $\mathcal{B}$ of the set $\mathcal{A}^*$ of continuous linear functionals on $\mathcal{A}$. Indeed

i) each $m \in \Sigma(\mathcal{A})$ is a continuous functional since $|m(A)| \leq ||A||$, and, since $\mathbf{1} \in \mathcal{A}$ and $m(\mathbf{1}) = 1$,

$$||m||_\infty = \sup_{A \in \mathcal{A}} \; |m(A)|/||A|| = 1,$$

i.e. $m \in \mathcal{B}$.

ii) the weak * topology on $\mathcal{A}^*$ is defined by the following basic neighborhoods: given $\varepsilon > 0, A_1, ..., A_n \in \mathcal{A}$, the neighborhood of $\bar{l}$ is

$$U_{A_1,...,A_n}(\bar{l}; \varepsilon) = \{l \in \mathcal{A}^* : |l(A_i) - \bar{l}(A_i)| < \varepsilon, \quad i = 1, ..., n\}.$$

With respect to such a topology $\mathcal{A}^*$ is a Hausdorff topological space and the unit ball $\mathcal{B} \subset \mathcal{A}^*$ is compact, by the Alaoglu-Banach theorem[19]. The topology induced on $\Sigma(\mathcal{A})$ by the weak * topology is called the Gelfand topology and it is the weakest topology under which all the functions $\tilde{A}(m)$ are continuous. Clearly, $\Sigma(\mathcal{A})$ is a Hausdorff space because the weak * topology is Hausdorff.

iii) it remains to prove that $\Sigma(\mathcal{A})$ is a weak * closed set of $\mathcal{B}$. As a matter of fact, if $m_\alpha \in \Sigma(\mathcal{A})$ and $m_\alpha \to l$ in $\mathcal{B}$, then $\forall A, B \in \mathcal{A}$

$$m_\alpha(AB) = m_\alpha(A)\, m_\alpha(B) \quad \Rightarrow \quad l(AB) = l(A)\, l(B),$$

i.e. l is multiplicative.

Finally, since by definition $\tilde{\mathcal{A}}$ separates the points of $\Sigma(\mathcal{A})$, and it is closed by eq. (1.5.5), by the Stone-Weierstrass theorem it is the whole $C(\Sigma(\mathcal{A}))$.

Examples. To better grasp the properties of C^*-algebras it is instructive to work out the following Exercises.

1. Let X be a compact Hausdorff topological space and $C(X)$ the C^*-algebra of the continuous functions on X.

[19]See e.g. N. Dunford and J.T. Schwartz, *Linear Operators. Part I: General Theory*, Interscience 1958; M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Academic Press, Vol. 1, p. 115.

a. Determine the spectrum $\sigma(f)$ for $f \in C(X)$ and show that $\forall F \in C(\mathbf{C})$, $\sigma(F(f)) = F(\sigma(f))$.
b. Verify explicitly the validity of eq. (1.4.6).
c. Determine independently the proper maximal ideals of $C(X)$ and its multiplicative linear functionals; verify Proposition 1.5.3 and that $\Sigma(C(X)) = X$. [Hints: If, given a proper maximal ideal I, $\forall x \in X$, there is a $f_x \in I$ such that $f_x(x) \neq 0$, then, by exploiting the compactness of X, one could construct a never vanishing $h \in I$ and $h\,h^{-1} = \mathbf{1} \in I$, so that I is not proper. If the support K of a multiplicative measure μ on X (i.e. the smallest compact set such that $\mu(f) = 0$, if $\operatorname{supp} f \cap K = \emptyset$) contains two disjoint open sets K_1, K_2, then $\exists g \in C(X)$, $\operatorname{supp} g \subseteq K_2$, such that $\mu(g) \neq 0$ and $\forall f \in C(X)$, $\operatorname{supp} f \subseteq K_1$, $0 = \mu(fg) = \mu(f)\mu(g)$, i.e. $\mu(f) = 0$.]
2. Let $\mathcal{M}$ be the set of diagonal $n \times n$ matrices. Verify that $\mathcal{M}$ is a C^*-algebra. Determine the spectrum of $M \in \mathcal{M}$ and the Gelfand spectrum of $\mathcal{M}$.

1.6 Appendix C: Spectra and states

We discuss general properties of the states of a C^*-algebra $\mathcal{A}$ and their relation with the spectra of the normal elements of $\mathcal{A}$.

Proposition 1.6.1 *Let $\mathcal{A}, \mathcal{B}$ be C^*-algebras and $\mathcal{A} \subset \mathcal{B}$, then for any $A \in \mathcal{A}$, the set $\sigma_{\mathcal{A}}(A)$ of $\lambda \in \mathbf{C}$ such that $\lambda\mathbf{1} - A$ is not invertible in $\mathcal{A}$ coincides with*

$$\sigma(A) = \{\lambda \in \mathbf{C} : \lambda\mathbf{1} - A \;\; \textit{is not invertible in } \mathcal{B}\}.$$

Proof. In fact, if $(\lambda\mathbf{1} - A)^{-1}$ exists in $\mathcal{B}$, it can be expressed as a convergent power series, i.e. it is the norm limit of partial sums each belonging to $\mathcal{A}$, so that also $(\lambda\mathbf{1} - A)^{-1}$ belongs to $\mathcal{A}$.

Proposition 1.6.2 *Given a normal element A of a C^*-algebra $\mathcal{A}$, then any continuous function $F = F(A, A^*)$ of A, A^* defines an element of $\mathcal{A}$ and $\sigma(F) = \tilde{F}(\sigma(A))$, where $\tilde{}$ denotes the Gelfand transform.*

Proof. Let $\mathcal{A}_A$ be the abelian algebra generated by A, A^* and $\mathbf{1}$; by the Gelfand-Naimark theorem, F defines an element of $\mathcal{A}_A \subset \mathcal{A}$ and, by the preceding Proposition, $\sigma(F) = \sigma_{\mathcal{A}_A}(F) = \sigma(\tilde{F}) = \mathit{Range}\,\tilde{F} = \tilde{F}(\mathit{Range}\,\tilde{A}) = \tilde{F}(\sigma(A))$.

By Proposition 1.6.2 the functional calculus for normal A follows from the calculus on functions; e.g. $\sigma(A) \subseteq \mathbf{R}_+$ iff $\tilde{A} \geq 0$ and in this case $\tilde{A}^{1/2} \geq 0$ defines $A^{1/2}$. Similarly, the decomposition $\tilde{A} = \tilde{A}_+ - \tilde{A}_-$, with $\tilde{A}_\pm \geq 0$, $\tilde{A}_+\tilde{A}_- = 0$ yields $A = A_+ - A_-$, $\sigma(A_\pm) \geq 0$, $A_+ A_- = 0$.

The following properties are equivalent and define the set $\mathcal{A}_+$ of the *positive* elements, [20]

$$1)\,\sigma(A) \subseteq [0, ||A||], \;\; 2)\; A = B^2, \;\; B = B^*, \;\; 3)\; ||\mathbf{1} \, - A/||A||\; || \leq 1,$$
$$4)\;\; A = C^*\,C.$$

$\mathcal{A}_+$ is a closed convex cone, since 3) is stable under closure and multiplication of A by positive numbers, and $\forall A$, B, with $||A|| = ||B|| = 1$, one has $||\mathbf{1} \, - (A+B)/2|| \leq \frac{1}{2}\,(||\mathbf{1} \, - A|| + ||\mathbf{1} \, - B||) \leq 1$.

Proposition 1.6.3 *A linear functional* ω *on a* C^**-algebra* $\mathcal{A}$*, with identity, is positive iff: 1)* ω *is bounded, 2)* $||\omega\,|| = \omega\,(\mathbf{1})$.

Proof. Let ω be positive, then by the Cauchy-Schwarz' inequality

$$|\omega\,(A^*B)|^2 \leq \omega\,(A^*A)\,\omega\,(B^*B),$$

which implies

$$|\omega(A)\;\;|^2 \leq \omega\,(\mathbf{1})\,\omega\,(A^*A).$$

Thus, to get continuity it suffices to prove that

$$\omega\,(A^*A) \leq ||A||^2\,\omega(\mathbf{1}),$$

i.e. that for any positive B, $\omega\,(B) \leq ||B||\,\omega(\mathbf{1})$. This follows easily from the discussion after Proposition 1.6.2, in particular from the equivalent definitions of positive elements. In fact, one has that

$$||B||\mathbf{1} \, - B \geq 0$$

and therefore, by the positivity of ω,

$$\omega(B) \leq ||B||\,\omega(\mathbf{1}).$$

Conversely, let ω be bounded and $||\omega\,|| = \omega\,(\mathbf{1})$; it suffices to take $\omega(\mathbf{1}\,) = 1$ and consider A with $||A|| = 1$. By Proposition (1.5.4) $\omega(A^*\,A)$ is real and one has

$$|1 - \omega\,(A^*A)| = |\omega\,(\mathbf{1} \, - A^*A)| \leq ||\mathbf{1} \, - A^*A|| \leq 1,$$

where the first inequality follows from the continuity of ω and the last from the positivity of A^*A (see the equivalent property 3 above). This requires $\omega\,(A^*A) \geq 0$.

[20] Clearly, 1) $\Leftrightarrow$ 2), by the existence of $A^{1/2}$. 1) $\Rightarrow$ 3), since $\sigma(\mathbf{1} \, - A/||A||) = 1 - \sigma(A)/||A|| \subseteq [0,\,1]$ and conversely, $||\mathbf{1} - A/||A||\,|| \leq 1$ implies $\sigma(\mathbf{1} - A/||A||) \subseteq [-1, 1]$, i.e. $\sigma(A) \subseteq [0, 2||A||]$. For the equivalence 4) $\Leftrightarrow$ 1), one notes that $A = a_+^2 - a_-^2$, $a_\pm = A_\pm^{1/2}$ and $(Ca_-)^*\,(Ca_-) = a_-(a_+^2\, - a_-^2\,)\,a_- \, = -a_-^4 \, \in -\mathcal{A}_+$. On the other hand, by writing $D \equiv Ca_- \, = a_1 + ia_2$, $a_i \, = a_i^*$, one has $DD^* \, = -D^*D + 2a_1^2 + 2a_2^2 \in \mathcal{A}_+$. Now, if $\lambda \notin \sigma(D^*D) \cup \{0\}$, $\exists E = (D^*D - \lambda\mathbf{1})^{-1}$ and the identities $(DD^* \, - \lambda\mathbf{1})(DED^* \, - \mathbf{1}) = \lambda = (DED^* - \mathbf{1})(DD^* - \lambda\mathbf{1})$ imply $\lambda \notin \sigma(DD^*) \cup \{0\}$ and by symmetry $\sigma(D^*D) \cup \{0\} = \sigma(DD^*) \cup \{0\}$; then, one gets a contradiction, unless $a_- = 0$.

Proposition 1.6.4 *Let A be a normal element of a C^*-algebra $\mathcal{A}$, then $\lambda \in \sigma(A)$ iff there exists a positive linear functional ω on $\mathcal{A}$ such that, for any polynomial $\mathcal{P}(A, A^*)$, one has $\omega(\mathcal{P}(A, A^*)) = \mathcal{P}(\lambda, \bar{\lambda})$.*

Proof. Let $\mathcal{A}_A$ be the abelian algebra generated by A, A^* and the identity. By Proposition 1.6.1, $\lambda \in \sigma_{\mathcal{A}}(A)$ and by Proposition 1.5.3 there exists a multiplicative linear functional ω_A on $\mathcal{A}_A$ (and therefore $\omega_A(\mathbf{1}) = 1$), such that $\omega_A(A) = \lambda$, $\omega_A(A^*) = \bar{\lambda}$ and therefore $\omega_A(\mathcal{P}(A, A^*)) = \mathcal{P}(\lambda, \bar{\lambda})$. Furthermore, ω_A is positive on $\mathcal{A}_A$ and therefore, by Proposition 1.6.3, is a bounded linear functional with $||\omega_A|| = \omega_A(\mathbf{1}) = 1$. Since $\mathcal{A}_A$ is a closed subalgebra of $\mathcal{A}$, by the Hahn-Banach theorem ω_A has an extension ω to $\mathcal{A}$ with $||\omega|| = ||\omega_A|| = \omega_A(\mathbf{1}) = 1 = \omega(\mathbf{1})$. By Proposition 1.6.3, ω is a positive linear functional on $\mathcal{A}$ and coincides with ω_A on $\mathcal{A}_A$.
Conversely, if $\omega(\mathcal{P}(A, A^*)) = \mathcal{P}(\lambda, \bar{\lambda})$, then ω is a multiplicative linear functional on $\mathcal{A}_A$ and $\omega(A) = \lambda$. Then, by Proposition (1.5.3), $\lambda \in \sigma(A)$.

As an immediate consequence of the above Proposition we have

Proposition 1.6.5 *The positive linear functionals on a C^*-algebra $\mathcal{A}$ separate the elements of $\mathcal{A}$.*

Proof. Let $A, B \in \mathcal{A}, A \neq B$. Then, $||A - B|| \neq 0$ and, putting $A - B = A_1 + iA_2$, with $A_i = A_i^*$, $i = 1, 2$, one has $||A_1|| + ||A_2|| > 0$. Without loss of generality we consider the first case $||A_1|| \neq 0$. Then $\lambda = ||A_1|| \in \sigma(A_1)$ and by the preceding Proposition there exists a state ω such that $0 \neq \lambda = \omega(A_1) = \mathrm{Re}\, \omega(A - B)$.

Proposition 1.6.6 *An element A of a C^*-algebra $\mathcal{A}$ is positive iff $\omega(A) \geq 0$, for all positive linear functionals ω.*

Proof. In fact $\omega(A) \geq 0$, $\forall \omega$, implies $\omega(A - A^*) = 0, \forall \omega$, i.e. $A = A^*$, and therefore if $\lambda \in \sigma(A)$, $\exists \omega$ such that $\omega(A) = \lambda$, (see the proof of Proposition 1.6.4 above) i.e. $\lambda \geq 0$ and $\sigma(A) \subseteq [0, ||A||]$.

Positivity allows to introduce a natural ordering of linear functionals: given two positive linear functional $\omega_1, \omega_2, \omega_1$ is said to majorize ω_2, briefly $\omega_1 \geq \omega_2$, if $\omega_1 - \omega_2$ is a positive linear functional.

The functional calculus developed above allows to prove the spectral representation of a bounded self-adjoint operator A on a Hilbert space $\mathcal{H}$, (the so-called *spectral theorem*), which generalizes the standard representation of an $n \times n$ hermitian matrix M

$$M = \sum_{i=1}^{n} \lambda_i P_i, \tag{1.6.1}$$

where λ_i are the eigenvalues of M and P_i are the projections on the corresponding eigenvectors (if $\lambda_i = \lambda_j$, P_i and P_j have to be chosen as two independent projections on the corresponding two dimensional space).

Theorem 1.6.7 *For a bounded self-adjoint operator A one can write the following spectral representation*

$$A = \int_{\sigma(A)} \lambda dP(\lambda),$$

with $dP(\lambda)$ a projection valued measure defined on the spectrum of A.

A simple proof of the spectral theorem exploits the construction of projection operators $P(\Delta)$ which correspond to the characteristic functions of intervals $\Delta = [a, b) \subseteq \sigma(A)$ (see below for their explicit construction). They are characterized by the property of projecting on the subspace $\mathcal{H}_\Delta \subseteq \mathcal{H}$, such that the corresponding expectations of A are in Δ; thus

$$a\,P(\Delta) \leq A\,P(\Delta) \leq b\,P(\Delta). \tag{1.6.2}$$

Clearly, $AP(\Delta) = 0$ if $\Delta \cap \sigma(A) = \emptyset$, and if $\cup_i \Delta_i = \sigma(A)$, $\Delta_i \cap \Delta_j = \emptyset$, for $i \neq j$, then

$$P(\cup \Delta_i) = \sum P(\Delta_i) = \mathbf{1}. \tag{1.6.3}$$

Now, if $\nu_k \in [\lambda_k,\, \mu_k) \equiv \Delta_k$, $\delta \equiv \max|\mu_k - \lambda_k|$, and $\cup \Delta_k = \sigma(A)$ one has from eq. (1.6.2)

$$\sum_k (\lambda_k - \nu_k) P(\Delta_k) \leq A - \sum_k \nu_k P(\Delta_k) \leq \sum_k (\mu_k - \nu_k) P(\Delta_k)$$

and

$$\delta \leq A - \sum_k \nu_k P(\Delta_k) \leq \delta,$$

i.e. the Riemann sums $\sum \nu_k P(\Delta_k)$ converge to A as $\delta \to 0$. Thus, as in the ordinary case, one may introduce the operator valued integral

$$A = \lim_{\delta \to 0} \sum \nu_k P(\Delta_k) = \int_{\sigma(A)} \lambda\, dP(\lambda), \tag{1.6.4}$$

which gives the spectral representation of A in terms of the projection valued measure $dP(\lambda)$.

The existence of the projections $P(\Delta)$ can be argued by explicit construction. Since $A^2 \geq 0$, by the remark after Proposition 1.6.2, one may define the positive square root $\sqrt{A^2}$ and $A - \sqrt{A^2} \leq 0$. The projection P_+ on the subspace $\mathcal{H}_+ = \{x \in \mathcal{H};\ (A - \sqrt{A^2})x = 0\}$ has the meaning of the projection on the subspace on which the expectations of A are positive, i.e. $AP_+ \equiv A_+ \geq 0$. Clearly $\mathbf{1} - P_+$ projects on the subspace on which the expectations of A are negative. By the same reasoning applied to the operator $A_\lambda = A - \lambda\mathbf{1}$ one defines the projection $P_+(\lambda)$, (corresponding to the subspace on which $A \geq \lambda$), and finally for $\Delta = [\lambda,\, \mu)$, $P(\Delta) \equiv P_+(\lambda) - P_+(\mu)$ is the required projection.

A more elegant and compact proof of the spectral theorem can be obtained by exploiting the functional calculus and Proposition 1.6.2.

First, if A is a self-adjoint (more generally a normal) element of a C^*-algebra $\mathcal{A}$, a possible realization of the Gelfand isomorphism $\mathcal{A}_A \to C(\Sigma(\mathcal{A}_A)) = C(\sigma(A))$ is given by $\hat{A}(\lambda) = \lambda$, $\hat{F}(A) = F(\lambda)$, for any continuous function F. In fact, in this way one realizes an isometric isomorphism between the polynomial *-algebra generated by A and the *-algebra of polynomials on $C(\sigma(A))$ and by continuity such an isomorphism extends to the continuous functions on $C(\sigma(A))$.

The second step is to note that a vector $x \in \mathcal{H}$ defines a positive linear functional $(x,\, F(A)\, x)$ on the the algebra of continuous functions F of A, and therefore a positive linear functional on $C(\sigma(A))$. Then, by the Riesz-Markov theorem there exists a (unique) regular Borel measure μ_x such that

$$(x,\, F(A)\, x) = \int_{\sigma(A)} F(\lambda)\, d\mu_x(\lambda). \tag{1.6.5}$$

By the polarization identity

$$\begin{aligned}(x,\, Ay) = \frac{1}{4}[&(x+y,\, A(x+y)) - (x-y,\, A(x-y)) \\ &-i(x+iy,\, A(x+iy)) + i(x-iy,\, A(x-iy))]\end{aligned} \tag{1.6.6}$$

also $(x,\, F(A)y)$ is a continuous linear functional on $C(\sigma(A))$ and therefore expressible in terms of a complex measure μ_{xy} on $C(\sigma(A))$

$$(x,\, F(A)y) = \int_{\sigma(A)} F(\lambda)\, d\mu_{xy}(\lambda). \tag{1.6.7}$$

Such a spectral representation allows to extend the Gelfand Naimark isomorphism $F(A) \to F(\lambda)$ from the continuous functions of A to the bounded Borel functions B, by putting

$$(x,\, B(A)y) \equiv \int_{\sigma(A)} B(\lambda)\, d\mu_{xy}(\lambda). \tag{1.6.8}$$

In fact, the r.h.s. is a continuous sesquilinear form on $\mathcal{H}$ and by the Riesz lemma identifies a unique operator B.

In particular, if $\chi(\Delta)$ denotes the characteristic function of the Borel set Δ, the corresponding operator $P(\Delta)$ (called a *spectral projection*) satisfies

i) (*projection*) $P(\Delta) = P(\Delta)^*$, $\;P(\Delta_k)\, P(\Delta_j) = P(\Delta_k \cap \Delta_j)$

ii) $P(\emptyset) = 0$, $\;P(\sigma(A)) = \mathbf{1}$

iii) (σ-*additivity*) If $\Delta_j \cap \Delta_k = \emptyset$, $\forall\, j \neq k$

$$P(\cup \Delta_k) = s - \lim_{N\to\infty} \sum_{k=1}^{N} P(\Delta_k).$$

Such a system of projections is called a partition of unity, since if $\cup\Delta_k = \sigma(A)$, $\Delta_j \cap \Delta_k = \emptyset$, σ-additivity gives

$$\mathbf{1} = \sum_k P(\Delta_k).$$

Now, for any fixed $x \in \mathcal{H}$, the expectation $\mu_x(\Delta) \equiv (x, P(\Delta)x)$ defines a measure on the Borel sets of $\sigma(A)$ and by a standard measure theoretical argument

$$(x,\, F(A)x) = \int_{\sigma(A)} F(\lambda)\, d(x, P(\lambda)x), \tag{1.6.9}$$

where $d(x, P(\lambda)x)$ denotes the integration with respect to μ_x. Riesz lemma then gives the spectral representation of $F(A)$ and, in particular,

$$A = \int_{\sigma(A)} \lambda\, dP(\lambda), \tag{1.6.10}$$

where $dP(\lambda)$ is the projection valued measure defined by the family of projections $P(\Delta)$.

As a simple application we have obtain Stone's theorem

Theorem 1.6.8 *A bounded self-adjoint operator A defines a one-parameter group of strongly continuous unitary operators $U(t)$, $t \in \mathbf{R}$, of which A is the generator in the sense that*

$$\text{strong-}\lim_{t\to 0} t^{-1}\,(U(t) - \mathbf{1}) = i\,A. \tag{1.6.11}$$

Proof. The unitary group $U(t)$ is defined by the spectral integral

$$U(t) = e^{i\,A\,t} = \int_{\sigma(A)} e^{i\,\lambda\,t}\, dP(\lambda), \tag{1.6.12}$$

with $dP(\lambda)$ the spectral measure defined by A. The group law and unitarity follow from the Gelfand isomorphism and the spectral representation.

The strong continuity and eq. (1.6.11) follows from the dominated convergence theorem respectively applied the the sequence $f_t(\lambda) \equiv -1 + \exp(i\lambda\, t)$ which converges pointwise to zero, for $t \to 0$, and it is dominated by 2 and to the sequence $\Delta f_t \equiv t^{-1}(\exp(i\lambda\, t) - 1)$ which converges pointwise to $i\,\lambda$ and it is dominated by $|\lambda|$, for $t \to 0$.