Conceptual and mathematical aspects of quantum mechanics

Mario Paschke

January 12, 2009

Contents

Chapter 1

Introduction

1.1 A rickety pillar of physics

The apparent conceptual incompatibility of general relativity and quantum theory represents one of the most important challenges for contemporary theoretical physics. Many leading experts seem to agree that it's resolution will require a severe modification of the fundamental principles underlying these theories.

Most importantly, both theories do assume that energy (matter) can, in principle, be localized in an arbitrarily small region of spacetime. However all the prominent approaches to quantum gravity, in particular String theory and Loop Quantum Gravity, find that this is impossible within their model. Moreover several heuristic arguments indicate that one should indeed expect that a consistent theory of the interaction of the gravitational field with quantum fields will be nonlocal. However, none of these arguments is mathematically rigorous nor is it completely clear whether the assumptions they make about the sought-for theory of quantum gravity are realistic. Even though they appear very natural from our present understanding of the world: The real theory of quantum gravity might not.

On the one hand, quantum theory and general relativity are both verified experimentally to an impressive accuracy. Thus, if we believe that the principle of locality is violated at some energy-scale (or rather at sufficiently high energy densities), in other words if locality is not an indispensable fundamental principle, then it should be possible to formulate alternative theories that do not presuppose locality, but nevertheless are in accordance with all experimental facts.

Indeed, it is one of the aims of A. Connes' Noncommutative Geometry to generalize Einsteins theory of general relativity to noncommutative space-

times, i.e. spacetimes on which matter cannot be localized in arbitrarily small regions.

Likewise, quantum theories on noncommutative spacetimes have also been a subject of intense research over the last decade. Contrary to the original belief it turned out that certain theoretical arguments that such theories could not be consistent (for instance that they could not be unitary or renormalizable) are not justified. Yet there is as yet no model that is realistic enough to be confronted with experiment.

But what is actually known experimentally about locality ? Is there any direct consequence of locality that could be tested experimentally and used to rule out nonlocal theories in a model-independent way (i.e. without having to formulate a concrete model first)? Or is even the converse true: Do the experimental findings appear much more natural in such nonlocal theories ? In fact, there are many indications for the latter. We shall mention some of them later.

1.2 About these Lecture notes

In view of such questions, it seems to me that it is certainly worthwhile to present the conceptual and mathematical foundations of general relativity and quantum theory in a way that makes the role played by the principle of locality as explicit as possible. However, to the best of my knowledge none of the many excellent textbooks on these subjects considers this aspect. The following notes of a course that I have given for students of mathematics at the university of Opava attempts to fill this gap for the case of quantum mechanics. At least as far as it had been possible in the given time, which unfortunately is not very far. In particular only very few of the basic experiments and mathematical tools of quantum mechanics will be described, and in fact very briefly. Excellent introductions to these issues can be found in many textbooks. So there would have been no point to write another book on them here. Yet as I necessarily had to describe the conceptual foundations of quantum physics in much greater detail and much more carefully than can usually be done in a regular course, these notes could well serve as a supplement to the more conventional textbooks or a regular course in quantum mechanics. Yet, as these notes take a rather different vantage point, I hope that some readers may find them inspiring. More importantly I hope that they can help them to realize that some statements in many textbooks a bot the interpretation and experimental verification are not as clear as they are usually sold to be. Even more importantly, still, I intended to point out that in view of recent experimental and theoretical progress many statements

about the interpretation and experimental verification of quantum mechanics are not as unclear as they are sold to be.

The lecture started with some introduction to mathematical prerequisites: C^*-algebras, unbounded operators in Hilbert spaces and some basic differential geometry. Accordingly these notes start with a brief outline of the mathematical structures that are used in the later chapters. However, as all these topics are well covered in the literature, I shall be very sketchy here. In the course itself I tried to present the proofs of the stated theorems in full detail, while here they are usually omitted. Only when I felt that some details of the proof are important to understand the concepts that are essential for the remainder of the notes, I have indicated the respective part of the proof and sketched the part I did not cover. Exceptions are also made if the proof is less than half a page or if I was not aware of a good reference where the proof is stated.
Thus I recommend readers familiar with the topics covered in the first chapter to simply skip it. Those readers, who do not feel too comfortable with those subjects, might get a first impression, and a rather broad overview of them. But no more. Thus, some excellent and very comprehensive textbooks on these topics that are listed in the bibliography and recommended for further reading.

We then carefully analyzed which operation on the set of observables and the states of a physical system can safely be assumed. For instance, given two observables, position and velocity of a particle say, their sum can in general only be measured (if at all) by an experimental setup that is completely different from the two apparatuses used to measure each of them separately. It is then well conceivable that the result of a measurement of the sum of the observables will not be the sum of the results of measurements of each of the observables (in particular as any measurement has some inaccuracy and does in general change the state of the system). In fact, as Bell pointed out, in a hidden variable theory this linearity of states must fail. If this linearity holds, the system can always be described via operators (the observables) on some Hilbert space (whose elements represent the states).

Going on with our analysis we then find that the failure of classical mechanics to correctly describe certain experiments means that either there is an uncertainty relation for the simultaneous measurement of momentum and position, or the theory must contain hidden variables, and thus cannot be described by operators on a Hilbert space. The latter possibility is then seemingly ruled out by the experiments proposed by J.Bell and carried out later (by A.Aspect and others). However as we are going to see this is only

true if one additionally assumes that the hidden variable theory is a local theory. Following Bell we shall explicitly give a nonlocal hidden variable theory that does correctly describe the Aspect experiment. Thus, if do not want to exclude the possibility of nonlocality, then the experimental findings in the non relativistic regime do not seem to exclude a deterministic theory. Quite to the contrary.

We therefore also briefly considered relativistic quantum mechanics, only to find that a one particle theory would not have an operationally well defined interpretation (at least if one keeps the assumption of locality). Thus one would necessarily have to consider quantum field theory. Unfortunately neither time nor my abilities did allow to enter a discussion of locality in quantum field theory. Thus, admittedly, these notes are more than unsatisfactory as concerns the questions related to locality. Yet I still hope that they may serve as food for thought. Some readersmight e.g. not be aware of the posibility to derive the Schrödinger-equation from the uncertainty-relation, which implies that the only interactions compatible with these relations are gravitational and Yang-Mills interactions, i.e. precisely those that we do find in nature, but no others. In fact, what might render these notes special is that they necessarily give a rather broad overview of physics, including the basic equations, like the Yang-Mills, Einstein-, Schr"odinger- or the Dirac-equation. Furthermore they point out in how far these equations rely on the fundamental principles.

Chapter 2

Mathematical Preliminaries

Most physical systems have a natural description on some Hilbert space $\mathcal{H}$ where observables are represented as selfadjoint operators. This is particularly true for all quantum theories and all models of (quantum) statistical physics. Moreover, recently A.Connes' has shown that also Einstein's Theory of General Relativity can be (almost) equivalently reformulated in this language.
Thus, as this formalism seems to be rather universally applicable, it is certainly worthwhile to gain some familiarity with its basic concepts, definitions and results. In the following chapter we shall therefore give a broad overview of the most elementary aspects of operator theory.

2.1 Bounded Operators on Hilbert Spaces

In the sequel $\mathcal{H}$ will always denote a complex Hilbert space whose inner product shall be denoted by $\langle \cdot, \cdot \rangle$. We shall denote the norm of a vector $\xi \in \mathcal{H}$ by $\|\xi\| = \sqrt{\langle \xi, \xi \rangle}$.

2.1.1 The Algebra $\mathcal{L}(\mathcal{H})$

An **operator** (i.e. a linear transformation) A will be called **bounded** if there is a $C \geq 0$ such that for all $\xi \in \mathcal{H}$ one has $\|A\xi\| \leq C\|\xi\|$. The **operator norm** of a bounded operator A is by definition the smallest C with this property, i.e.

$$\|A\| := \sup_{\|\xi\|=1} \|A\xi\|.$$

One may check that this indeed defines a norm, i.e. for all bounded operators A, B and $\lambda \in \mathbb{C}$ one has

$$\|A+B\| \leq \|A\|+\|B\|, \qquad \|\lambda A\| = |\lambda|\|A\|, \qquad \|A\| = 0 \;\Leftrightarrow\; A = 0.$$

Moreover it is

$$\|AB\| \leq \|A\|\|B\|.$$

An operator is continuous if and only if it is bounded. We shall denote the set of all bounded operators on $\mathcal{H}$ by $\mathcal{L}(\mathcal{H})$.
In view of the triangle inequality and homogeneity of the norm, $\mathcal{L}(\mathcal{H})$ is obviously a vector space. Moreover it is also an algebra, as the product of two bounded operators is always bounded by the above stated submultiplicativity of the operator norm. Even more so it is also **complete**:
Let us call a sequence A_n of bounded operators a Cauchy sequence with respect to the norm topology if $\|A_n - A_m\|$ tends to zero as $n, m \to \infty$. It then easily follows that every such Cauchy sequence converges in $\mathcal{L}(\mathcal{H})$.
Furthermore each $A \in \mathcal{L}(\mathcal{H})$ has an **adjoint** in $\mathcal{L}(\mathcal{H})$, i.e. there exists a unique operator $A^* \in \mathcal{L}(\mathcal{H})$ such that

$$\langle \eta, A\xi \rangle = \langle A^*\eta, \xi \rangle \qquad \forall \eta, \xi \in \mathcal{H}.$$

Some obvious properties of taking adjoints include

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

One also easily checks that $\|A^*\| = \|A\|$.
Finally it follows immediately from the definition of the adjoint of A that its kernel equals the orthogonal complement of the image of A, i.e.

$$A^*\xi = 0 \qquad \Leftrightarrow \qquad \langle \xi, A\eta \rangle \qquad \forall \eta \in \mathcal{H}.$$

Note that while the kernel of a bounded operator A is always closed, its image need not be.
In view of the above properties $\mathcal{L}(\mathcal{H})$ is a **Banach-$*$-algebra**. In fact it is more, namely a **C^*-algebra**, i.e. one also has $\|A^*A\| = \|A\|^2$ for all $A \in \mathcal{L}(\mathcal{H})$:

$$\|A\|^2 = \sup_{\|\xi\|=1} \langle A\xi, A\xi \rangle = \sup_{\|\xi\|=1} \langle \xi, A^*A\xi \rangle \leq \sup_{\|\xi\|=1} (\|\xi\|\|A^*A\xi\|) = \|A^*A\|,$$

while, on the other hand, from the above properties it also holds that

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

$\mathcal{L}(\mathcal{H})$ is also **unital**, i.e. the identity operator 1 on $\mathcal{H}$ is a bounded operator.

Definition:
We shall call a $*$-subalgebra of $\mathcal{L}(\mathcal{H})$ that is complete with respect to the norm topology a **concrete C^*-algebra**.

The definition and properties of abstract C^*-algebras will be stated later.

Definition:
$A \in \mathcal{L}(\mathcal{H})$ will be called

- **normal** if $A^*A = AA^*$
- **selfadjoint** if $A^* = A$
- **projection** if $A = A^* = A^2$
- **unitary** if $A^*A = AA^* = 1$
- **isometry** if $A^*A = 1$

If P is a projection, we call the dimension of $P\mathcal{H}$ the **rank of** P. (If $P\mathcal{H}$ is infinite dimensional, we say that P **is of infinite rank**.)

Examples:

1. Let (X, μ) be a measure space, and $\mathcal{H} = L^2(X, \mu)$. Then every bounded measurable function f defines a bounded operator $\hat{f}$ on $\mathcal{H}$ by setting $\hat{f}\psi(x) = f(x)\psi(x)$ (which however only makes sense almost everywhere.) This operator is selfadjoint if and only if f is real. It is unitary if and only if $|f(x)| = 1$ almost everywhere.

2. Let now $\mathcal{H}$ be any separable Hilbert space and let $\{\psi_k\}_{k=1,2,\ldots}$ denote a complete orthonormal set. Let the operators P_n be defined by

$$P_n\left(\sum_k \alpha_k\psi_k\right) = \sum_{k=1}^{n} \alpha_k\psi_k.$$

   Then the P_n are projectors.
   Moreover define T_k by

$$T_k\psi_n = \psi_{k+n}.$$

   Then $T_k^*\psi_n = \psi_{n-k}$ if $n \geq k$ but $T_k\psi_n = 0$ otherwise. Hence the T_k are isometries, but not unitaries.

3. There is a one-to-one correspondence between projections and closed subspace of $\mathcal{H}$: If P is a projection, then the space $\mathcal{H}_P = P\mathcal{H}$ is necessarily closed (as it is the orthogonal complement of the kernel of P.)
   Vice versa, given a closed subspace $\mathcal{V}$ of $\mathcal{H}$, every vector $\xi \in \mathcal{H}$ can be uniquely decomposed as $\xi = \eta + \psi$ where $\eta \in \mathcal{V}$ and ψ is in the orthogonal complement of $\mathcal{V}$. The linear map defined via $P_{\mathcal{V}}\xi = \eta$ is then a projection.

2.1.2 The Weak and Strong Operator Topologies

There are other interesting Topologies on $\mathcal{L}(\mathcal{H})$:

So far we have defined the **norm topology** on $\mathcal{L}(\mathcal{H})$. To recall, we say that the sequence $A_n \in \mathcal{L}(\mathcal{H})$ converges in norm to $A \in \mathcal{L}(\mathcal{H})$ if

$$\|A_n - A\| \to 0.$$

This topology is also often called *the topology of uniform convergence.*
Next, we shall say that the sequence $A_n\mathcal{L}(\mathcal{H})$ converges in the **strong topology** (respectively the *topology of pointwise convergence*) to $A \in \mathcal{L}(\mathcal{H})$ if for all $\xi \in \mathcal{H}$ one has

$$\|(A_n - A)\xi\| \to 0$$

Finally we shall say the sequence converges in the **weak topology** if for all $\xi \in \mathcal{H}$ and all $\eta \in \mathcal{H}$ one has

$$\langle (A_n - A)\xi, \eta \rangle \to 0.$$

Obviously, norm convergence implies strong convergence which in turn implies weak convergence. The converse is not true:
Consider example 2 above. Then the sequence P_n converges strongly to the unit operator 1, because the component of each fixed vector ψ in the orthogonal complement of $P_n\mathcal{H}$ tends to zero. However it does not converge in the norm topology because $\|P_n - P_m\| = 1$ whenever $n \neq m$.
The sequence T_k converges weakly to zero, but not strongly.

$\mathcal{L}(\mathcal{H})$ is then complete with respect to any of these topologies.

2.1.3 The Spectrum of bounded operators

Clearly, an element of $\mathcal{L}(\mathcal{H})$ need not be invertible, and if the inverse exists, it need not be a bounded operator. Now suppose that, given a bounded

operator A, $\lambda \in \mathbb{C}$ is chosen such that $\|A\| < |\lambda|$. Then the *von Neumann series*

$$(\lambda \cdot 1 - A)^{-1} = \frac{1}{\lambda}\left(1 + \frac{A}{\lambda} + \frac{A^2}{\lambda^2} + \cdots\right)$$

will converge in norm and thus be bounded.

Definition:
Let A be a bounded operator. We call the set

$$r(a) := \{\lambda \in \mathbb{C} \,|\, (A - \lambda \cdot 1) \text{ is invertible in } \mathcal{L}(\mathcal{H})\}$$

the **resolvent set** of A.
Its complement

$$\sigma(A) = \mathbb{C} \setminus r(A)$$

will be called the **spectrum** of A. The number

$$\rho(A) = \sup\{|\lambda\,| \in \sigma(A)\}$$

is called the **spectral radius** of A.

Example: Consider example 1 above with X compact, and take e.g. a fixed *continuous* function f on X, represented as a bounded multiplication operator $\hat{f}$ on $L^2(X, \mu)$. Then $\hat{f} - \lambda$ is invertible if and only if there does not exist $x \in X$ such that $f(x) = \lambda$. Consequently $\sigma(\hat{f}) = f(X)$ and $\rho(\hat{f}) = \max_{x \in X} |f(x)| = \|\hat{f}\|$.

Elements of the spectrum of A can be viewed as generalized eigenvalues. Indeed, if λ is an eigenvalue of A, i.e if there exists a $\psi_\lambda \in \mathcal{H}$ such that $A\psi_\lambda = \lambda\psi_\lambda$ the obviously $\lambda \in \sigma(A)$. However the converse is not true. We shall have to say more about the distinction of the spectrum and the eigenvalues of an operator in the section on unbounded operators.

Note hat the resolvent set of A is always open. In fact, if $\lambda_0 \in r(A)$ then for any λ in the open ball $|\lambda - \lambda_0| < \|(\lambda_0 - A)^{-1}\|^{-1}$ one checks that the series

$$\sum_{m=0} (\lambda_0 - \lambda)^m (\lambda_0 - A)^{-m-1}$$

converges in norm to $(\lambda - A)$, i.e. $\lambda \in r(A)$. Thus $\sigma(A)$ is closed. Since from the discussion previous to the definition of $\rho(A) \leq \|A\|$, it is also bounded. Finally it is also nonempty. Otherwise, if $r(A) = \mathbb{C}$ then the function $(\lambda - A)^{-1}$ would be holomorphic (in fact entire) and bounded on $\mathbb{C}$.

Thus it would vanish for all $\lambda \in \mathbb{C}$, in contradiction to the existence of A^{-1} which is implied by $0 \in R(A)$. In conclusion:

Proposition:
For all $A \in \mathcal{L}(\mathcal{H})$ the spectrum $\sigma(A)$ is a nonempty, compact subset of $\mathbb{C}$.

Note that if λ is in the spectrum of A, because $(\lambda - A)^* = \bar{\lambda} - A^*$. Likewise:

Proposition:
Let A be bounded. Then

1. $\sigma(A^*) = \overline{\sigma(A)}$.
2. If A is invertible then $\sigma(A^{-1}) = (\sigma(A))^{-1}$.
3. If A is an isometry then $\rho(A) = 1$.
4. If A is unitary then $\sigma(A) \subset S^1 = \{\lambda \in \mathbb{C} \,|\, |\lambda| = 1\}$.
5. If A is selfadjoint then $\sigma(A) \subset [-\|A\|, \|A\|]$ and moreover $\sigma(A^2) \subset [0, \|A\|^2]$..
6. If $P(z)$ is any complex polynomial, then

$$\sigma(P(A)) = P(\sigma(A)) = \{P(\lambda) \,|\, \lambda \in \sigma(A)\}.$$

2.1.4 Spectral decomposition of selfadjoint and unitary bounded operators

Definition:
We say that a selfadjoint element $A \in \mathcal{L}(\mathcal{H})$ is **positive,** $A \geq 0$ if $\langle \xi, A\xi \rangle \geq 0$ for all $\xi \in \mathcal{H}$.

Note that if $A \in \mathcal{L}(\mathcal{H})$ then A^*A is positive. In particular, if A is selfadjoint then A^2 is positive.

Lemma:
Every positive element $A \in \mathcal{L}(\mathcal{H})$ has a positive square root in $\mathcal{L}(\mathcal{H})$.

We shall later see a quick and elegant argument for this fact. For the convenience of the reader we note here that a direct construction is also possible (though less elegant):
Consider the following sequence of functions on $[0, 1]$ defined recursively via :

$f_0(t) = 0$ and $f_{n+1}(t) = f_n(t) + \frac{t}{2} - \frac{(f_n(t))^2}{2}$. one then proves that this sequence converges uniformly to $\sqrt{t}$ on $[0, 1]$.
Suppose now $\|A\| \leq 1$ (which can always be achieved by multiplying A with a positive real number). It is then possible to prove that $f_n(A)$ (as the f_n are polynomials this definition needs no explanation) is an increasing sequence of bounded operators that converges in norm. Obviously the limit is then $\sqrt{A}$.

Theorem
Let $A \in \mathcal{L}(\mathcal{H})$ be selfadjoint. Then there exists a projection valued measure $\mathrm{d}P(\lambda)$ on the real line such that

$$A = \int_{-\infty}^{\infty} \lambda \, \mathrm{d}P(\lambda).$$

Here the integral is to be understood in the sense that the Riemann sum approximations to this integral converge in the norm topology.

A sketch of the proof:
A is selfadjoint and therefore A^2 is positive. Hence $|A| := \sqrt{A^2}$ is well defined. As the kernel of any bounded operator is closed, the orthogonal projection P_+ on the kernel of $A - |A|$ is uniquely and well defined. Intuitively it can be viewed as the projection on the space on which A is "positive". More precisely $A_+ = AP_+$ is positive, i.e. $A_+ \geq 0$. Likewise, considering for $\alpha \in \mathbb{R}$ the operator $A - \alpha$ we get projections P_α such that $(A - \alpha)P_\alpha \geq 0$. For any Borel-set $\Delta = [\alpha, \beta)$ one may therefore introduce the projections

$$P(\Delta) := P_\alpha - P_\beta.$$

(They are indeed projections, since $P_\alpha P_\beta = P_\beta P_\alpha$ by construction.) One immediately checks the following properties

1. $\alpha P(\Delta) \leq AP(\Delta) \leq \beta P(\Delta)$.
2. Since $\sigma(A) \subset [-\|A\|, \|A\|]$ it follows that $P(\Delta) = 0$ if $\Delta \cap \sigma(A) = \emptyset$.
3. If $\Delta_1 \cap \Delta_2 = \emptyset$ then the corresponding projections are orthogonal, i.e. $P(\Delta_1)P(\Delta_2) = P(\Delta_2)P(\Delta_1) = 0$.
4. Accordingly, if $\cup_i \Delta_i = \sigma(A)$ and the Δ_i are mutually disjoint, $\Delta_i \cap \Delta_j = \emptyset$ for all $i \neq j$, then
$$\sum_i P(\Delta_i) = 1.$$

The projections $P(\Delta)$ are called the **spectral projections** of A.
Let now $\sigma(A) = \cup_{k=1}^{N} \Delta_k^N$ be a decomposition of $\sigma(A)$ in N disjoint Borel subsets $\Delta_k^N = \left[\alpha_k^N, \beta_k^N\right)$. Let $\delta_N = \sup_k |\beta_k^N - \alpha_k^N|$. We shall assume that $\delta_N \to 0$ as $N \to \infty$. Moreover let $\nu_k^N \in \Delta_k^N$ be arbitrarily chosen. It then follows immediately from the above properties of the $P(\Delta_k^N)$ that

$$-\delta_N \leq A - \sum_{k=1}^{N} \nu_k P(\Delta_k) \leq \delta_N$$

(Just subtract for each k the term $\nu_k P(\Delta_k)$ and sum over k.) Thus for $N \to \infty$ the sequence $A_N := \sum_{k=1}^{N} \nu_k P(\Delta_k)$ converges in norm to A, and this defines the above integral.

Now, let A be any normal operator. Then, writing $A = S + iT$, where S, T are selfadjoint, it follows immediately from $AA^* = A^*A$ that $ST = TS$ and this continues to hold for the spectral projections $P_T(\Delta)$, $P_S(\tilde{\Delta})$ for T and S respectively. It is then possible to construct a spectral decomposition also for normal operators along similar lines as above. Finally, since unitary operators are normal, and since for U unitary one has $\sigma(U) \subset S^1 = \{e^{i\lambda} \,|\, \lambda \in \mathbb{R}\}$ one infers

Theorem:
Let U be any unitary operator. Then there exists a projection valued spectral measure $\mathrm{d}P(\lambda)$ on $[0, 2\pi)$ such that

$$U = \int e^{i\lambda} \,\mathrm{d}P(\lambda).$$

Finally the spectral decomposition also allows to define bounded (and in fact, by continuity also some unbounded) functions of a normal operator. For simplicity we shall state the theorem only for selfadjoint operators. The general statement will be obvious.

Theorem (functional calculus):
Let A be a bounded and selfadjoint operator. If f is a bounded Borel function on $\sigma(A)$ then the following uniquely defines a bounded operator:

$$f(A) := \int f(\lambda) \,\mathrm{d}P(\lambda).$$

1. $f(A)$ is selfadjoint if and only if f is real-valued

2. $f(A)$ is a projection if and only if f only takes the values 0 or 1, i.e. if f is a characteristic function,

3. $f(A)$ is unitary if and only if $|f(\lambda)| = 1$ for all $\lambda \in \sigma(A)$.

2.1.5 Commutants and von Neumann algebras

Let A, B be two bounded operators. We call the bounded operator

$$[A, B] := AB - BA$$

the **commutator** of A and B. Two bounded operators are said to *commute* with each other, if their commutator vanishes. Thus, a subalgebra of $\mathcal{L}(\mathcal{H})$ is commutative if and only if all its elements commute with one another.

Proposition:
Two normal bounded operators commute if and only if all their respective spectral projections do commute.

Definition:
Let $\mathcal{S} \subset \mathcal{L}(\mathcal{H})$ be a subalgebra. We call the subspace

$$\mathcal{S}' := \{T \in \mathcal{L}(\mathcal{H}) \mid [T, S] = 0 \quad \forall S \in \mathcal{S}\}$$

the **commutant** of S. Likewise the subspace $\mathcal{S}'' = (\mathcal{S}')'$ is called the **bi-commutant** of $\mathcal{S}$.

Obviously $\mathcal{S} \subset \mathcal{S}''$ and $\mathcal{S}$ is commutative if and only if $\mathcal{S} \subset \mathcal{S}'$. Moreover $\mathcal{S}_1 \subset \mathcal{S}_2$ then $\mathcal{S}_2' \subset \mathcal{S}_1'$. Furthermore $1 \in \mathcal{S}'$ and the commutant is always weakly closed (by continuity of the commutator of bounded operators with respect to the weak topology).

Let $\mathcal{A}$ be a concrete C^*-algebra. We say that $\mathcal{A}$ acts irreducibly on $\mathcal{H}$ if there is no proper subspace of $\mathcal{H}$ that is invariant under the action of $\mathcal{A}$. In other words, any $\psi \in \mathcal{H}$ is a **cyclic vector** for $\mathcal{A}$, i.e. the set $\mathcal{A}\psi$ is a dense subset of $\mathcal{H}$.

Proposition:
If a concrete C^*-algebra $\mathcal{A}$ acts irreducibly then $\mathcal{A}' = \mathbb{C} \cdot 1$

Proof:
If $\mathcal{A}$ acts not irreducibly, i.e. if there exists a proper invariant subspace, then the $\mathcal{A}$ commutes with the orthogonal projection on that subspace. Hence the

commutant not only consists of multiples of the identity.
Conversely, if the commutant contains a nontrivial element $T \neq \lambda \cdot 1$, then also T^* commutes with $\mathcal{A}$ (by using the definition of the adjoint). Thus we may assume that T is selfadjoint (or take e.g. $T+T^*$). But then the spectral projections of T will also commute with $\mathcal{A}$ and there would exist a proper invariant subspace (unless $C + C^*$ is a multiple of the identity).

We shall say that a subalgebra $\mathcal{S}$ of $\mathcal{L}(\mathcal{H})$ acts **nondegenerately** on $\mathcal{H}$ if $S\xi = 0$ for all $S \in \mathcal{S}$ implies $\xi = 0$. If $1 \in \mathcal{S}$ then $\mathcal{S}$ acts obviously nondegenerately.

Theorem (von Neumann bicommutant theorem):
Let $\mathcal{S} \subset \mathcal{L}(\mathcal{H})$ be a $*$-subalgebra that acts nondegenerately on $\mathcal{H}$. Then $\mathcal{S}$ is strongly dense in $\mathcal{S}''$.

A sketch of the proof:
We only show that given any $\xi \in \mathcal{H}$ and $S'' \in \mathcal{S}''$ there exists a sequence $S_n \in \mathcal{S}$ so that $S_n\xi$ converges to $S''\xi$ in norm. To see this, on considers the strong closure of the image of ξ under $\mathcal{S}$, i.e the closed subspace

$$\Xi = \overline{\{S\xi \,|\, S \in \mathcal{S}\}}.$$

The orthogonal projector P_Ξ on this space then necessarily commutes with $\mathcal{S}$. (As $\mathcal{S}$ leaves Ξ and its orthogonal complement invariant). Therefor we have for all $S \in \mathcal{S}$:

$$S(1 - P_\Xi)\xi = (1 - P_\Xi)S\xi = 0$$

since $S\xi \in \Xi$. But then it follows from the nondegeneracy of the action of $\mathcal{S}$ that $(1 - P_\Xi)\xi = 0$, i.e $\xi \in \Xi$.
On the other hand P_Ξ also commutes with the bicommutant of $\mathcal{S}$ and hence with S''. Thus S'' leaves Ξ invariant, and so $S''\xi \in \Xi$ and by definition of Ξ there must exist the required sequence S_n. One then continues to prove that there must also exist a sequence S_n in $\mathcal{S}$ such that $S_n\phi \to S''\phi$ for all $\phi \in \mathcal{H}$ (which is much less obvious). The full proof can be found in the references given in the bibliography.

Thus we have that $\mathcal{S}' = \mathcal{S}'''$. We shall call a $*$-subalgebra $\mathcal{M}$ of $\mathcal{L}(\mathcal{H})$ a **von Neumann** algebra if $\mathcal{M} = \mathcal{M}''$. In other words a von Neumann algebra is a $*$-algebra of bounded operators that is weakly closed. Note that it follows that $1 \in \mathcal{M}$ and that any von Neumann algebra is also a concrete C^*-algebra. But the converse is not true.

We shall say that a von Neumann algebra $\mathcal{M}$ separates the elements of $\mathcal{H}$ if from $\langle \xi, m\xi \rangle = \langle \eta, m\eta \rangle$ for all $m \in \mathcal{M}$ it follows that $\xi = \eta$.

Lemma:
The algebra $\mathcal{L}(\mathcal{H})$ separates the elements of $\mathcal{H}$.

Lemma:
Let $\mathcal{H}$ be separable. There exists a commutative von Neumann algebra that separates the elements of $\mathcal{H}$.

As for the proof, one may take any orthonormal basis ψ_n of $\mathcal{H}$. Then the projections $P_n \psi_k = \delta_{nk} \psi_n$ are mutually commuting and thus generate a commutative von Neumann-algebra.

Lemma:
If a von Neumann algebra $\mathcal{M}$ separates the elements of $\mathcal{H}$ then there exists a commutative von Neumann algebra $\mathcal{N} \subset \mathcal{M}$ that separates the elements of $\mathcal{H}$.

This follows from the spectral decomposition and functional calculus, as for any normal element in $\mathcal{M}$ also all its spectral projections belong to $\mathcal{M}$ (apply characteristic functions to the element). From the separating property one then concludes that there do exist sufficiently many mutually commuting spectral projections in $\mathcal{M}$.

2.2 Abstract C^*-algebras

As we shall see later, the operational structure of observables for physical systems naturally leads to consider C^*-algebras. It is only because every C^*-algebra has a faithful representation on a Hilbert space that the latter structure plays such a prominent role in physics. In this chapter we shall briefly describe the way from abstract to concrete C^*-algebras.

Definition: A (complex) **Banach-algebra** $\mathcal{A}$ is an associative algebra over $\mathbb{C}$ which is also a normed space with norm $\| \cdot \| : \mathcal{A} \to \mathbb{R}$ such that

- $\|a\| \geq 0$ for all $a \in \mathcal{A}$,
- $\|\lambda a\| = |\lambda| \|a\|$ for all $a \in \mathcal{A}$ and all $\lambda \in \mathbb{C}$,
- $\|ab\| \leq \|a\| \|b\|$ for all $a, b \in \mathcal{A}$.

- The normed vector space $(\mathcal{A}, \|\cdot\|)$ is complete, i.e every Cauchy-sequence with respect to $\|\cdot\|$ converges in $\mathcal{A}$.

A Banach-$*$-algebra is a Banach-algebra that is equipped with an involution $^* : \mathcal{A} \to \mathcal{A}$, i.e. a map such that for all $a, b \in \mathcal{A}$ and all $\lambda \in \mathbb{C}$:

- $(a+b)^* = a^* + b^*$,
- $(\lambda a)^* = \bar{\lambda} a^*$,
- $(ab)^* = b^* a^*$,
- $(a^*)^* = a$.

A Banach-$*$-algebra $\mathcal{A}$ is called **unital** if there exists a (necessarily unique) element $1 \in \mathcal{A}$ such that $1a = a1 = a$ for all $a \in \mathcal{A}$.

Definition:
A **C^*-algebra** $\mathcal{A}$ is a Banach-$*$-algebra such that for all $a \in \mathcal{A}$ it holds that (C^*-condition)

$$\|a^* a\| = \|a\|^2.$$

Remark:
Note that the C^*-condition implies

$$\|a^*\| = \|a\|$$

and thus that also

$$\|aa^*\| = \|a\|^2.$$

Indeed from $\|a\|^2 = \|a^* a\| \leq \|a^*\|\|a\|$ it follows that $\|a\| \leq \|a^*\|$. Applying the same reasoning to $\|a^*\|^2$ one gets the assertion.

Examples:

1. Obviously any concrete C^*-algebra (and thus in particular $\mathcal{L}(\mathcal{H})$ is a C^*-algebra. In fact, as we shall see, any C^*-algebra can be viewed as a concrete C^*-algebra. However, there are some advantages to consider abstract C^*-algebras. In particular as concerns the definition of states.

2. Let X be a (locally compact) topological Hausdorff space. We shall say that a continuous function $f : X \to \mathbb{C}$ "vanishes at infinity" if for all $\epsilon > 0$ there exists a compact set K_ϵ such that $|f(x)| < \epsilon$ for all $x \in X \setminus K_\epsilon$.
   Consider the algebra $C_0(X)$ of all continuous functions which vanish at

infinity with the involution given by complex conjugation, i.e. for all $f \in C_0(X)$ and all $x \in X$ we have that $f^*(x) = \overline{f(x)}$. Moreover set for all $f \in C_0(X)$:

$$\|f\| := \sup_{x \in X} |f(x)|.$$

Then $(C_0(X),^*, \|\cdot\|)$ is a C^*-algebra, as is easily checked.
Note that this C^*-algebra is unital if and only if X is compact. (In the latter case $C_0(X)$ is just the algebra of all continuous functions on X.) Most importantly this C^*-algebra, usually simply denoted by $C_0(X)$, is commutative. In fact, as we shall prove in the next section, every commutative C^*-algebra is of this form. This then also completely justifies the shorthand notation $C(X)$.

Just as for elements of $\mathcal{L}(\mathcal{H})$ one may define **selfadjoint** ($a = a^*$), **normal** ($a^*a = aa^*$) and **unitary** ($u^* = u^{-1}$) elements, **projections** ($p^* = p = p^2$) and **isometries** ($v^*v = 1$) for (unital) C^*-algebras.
Likewise one defines the resolvent set of an element a of a C^*-algebra $\mathcal{A}$ as the set $r_{\mathcal{A}}(a)$ of complex numbers z for which $a - z$ is invertible in $\mathcal{A}$. The spectrum of a is then the closed set $\sigma_{\mathcal{A}}(a) = \mathbb{C} \setminus r_{\mathcal{A}}(a)$. Finally the supremum of $|\lambda|$ over all $\lambda \in \sigma_{\mathcal{A}}(a)$ is called the spectral radius $\rho_{\mathcal{A}}(a)$ of a.

Theorem (Spectral radius formula):
Let $\mathcal{A}$ be a C^*-algebra and $a \in \mathcal{A}$. Then $\sigma_{\mathcal{A}}(a)$ is a nonempty compact set and

$$\rho_{\mathcal{A}}(a) = \lim_{n \to \infty} \|a^n\|^{\frac{1}{n}} \leq \|a\|.$$

Sketch of the Proof:
Consider the von Neumann series

$$(1 - za)^{-1} = \sum_k (za)^k.$$

One can show – by first using the Banach-Steinhaus-Theorem (i.e. considering $\rho((1 - za)^{-1})$ for all linear functionals ρ on $\mathcal{A}$) and then the Hadamard-Cauchy-Theorem – that the radius of convergence of this sequence is $\frac{1}{\lim_n \|a^n\|^{\frac{1}{n}}}$.
Thus for $z \neq 0$ the resolvent $(z - a)^-1 = \frac{1}{z}(1 - \frac{1}{z}a)$ will exist if and only if $z \leq \lim_n \|a^n\|^{\frac{1}{n}}$, which proves the asserted formula, provided the limit exists.
We repeat Strocchis argument to prove the latter:
Let $r = \inf \|a^n\|^{\frac{1}{n}}$. Thus $r \leq \liminf \|a^n\|^{\frac{1}{n}} \leq \|a\| < \infty$ and we only need to show that $r \geq \limsup \|a^n\|^{\frac{1}{n}}$.
Now let $\epsilon > 0$ and choose $m \in \mathbb{N}$ such that $\|a^m\|^{\frac{1}{m}} < r + \epsilon$. (If no such m

exists the assertion already follows.) Then by decomposing any $n \in \mathbb{N}$ as $n = k_n m + l_n$ where $k_n, l_n \in \mathbb{N}$ and $0 < l_n < m$ we have

$$\|a^n\|^{\frac{1}{n}} = \|a_{k_n m} a_{l_n}\|^{\frac{1}{n}} \leq \|a^m\|^{\frac{k_n}{n}} \|a\|^{\frac{l_n}{n}} \leq (r+\epsilon)^{\frac{m k_n}{n}} \|a\|^{\frac{l_n}{n}}.$$

Hence the assertion follows as ϵ has been arbitrary and $\frac{l_n}{n} \to 0$ as $n \to \infty$, while $\frac{m k_n}{n} \to 1$ and thus $\limsup \|a^n\|^{\frac{1}{n}} \geq r$.
The proof of the equality for normal elements essentially uses the fact that for normal elements $\|a^{2^k}\| = \|a\|^{2^k}$, as follows from

$$\|a^2\|^2 = \|(a^*)^2 (a)^2\| = \|(a^* a)^* (a^* a)\| = \|a^* a\|^2 = \|a\|^4.$$

The remaining statements of the proposition are proven as in the case of concrete C^*-algebras.

Theorem (Gelfand-Mazur):
If all elements, except 0, of a C^*-algebra $\mathcal{A}$ are invertible then $\mathcal{A} \cong \mathbb{C}$.

As for the **proof** one notes that for all $0 \neq \lambda \in \sigma(a)$ one would have that $\lambda - a$ is not invertible and thus it vanishes, i.e. $a = \lambda 1$. Otherwise, if the spectrum of a only contains 0, then $\|a\| = 0$ and thus $a = 0$.

In the following, a selfadjoint element $a \in \mathcal{A}$ will be called **positive** if there exists $b \in \mathcal{A}$ such that $a = b^2$. A linear functional ω on $\mathcal{A}$ will be called positive if $\omega(a) \geq 0$ for all positive elements $a \in \mathcal{A}$. It will be called normalized if $\omega(1) = 1$ (in the case that $\mathcal{A}$ is unital).

Definition:
A positive, normalized linear functional on a C^*-algebra on $\mathcal{A}$ is called a **state on $\mathcal{A}$**.

The set of states is then automatically convex, as with ω_1, ω_2 also the convex combination $\omega = t\omega_1 + (1-t)\omega_2$ is a state for all $t \in [0,1]$. A state ω which can not be written in this way for $t \in (0,1)$ and some ω_1, ω_2 is called a **pure state**.

Example:
Let us consider a concrete C^*-algebra $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H})$. Then any $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ defines a state ω_ξ on $\mathcal{A}$ via

$$\omega_\xi(a) = \langle \xi, a\xi \rangle \qquad \forall a \in \mathcal{A}.$$

We shall later see that this state is pure if and only if $\dashv$ acts irreducibly on $\mathcal{H}$. Moreover let ξ_n be any orthonormal basis of $\mathcal{H}$ and let w_n be a sequence of nonnegative real numbers such that $\sum_n w_n = 1$. Then the operator ρ defined via

$$\rho\psi = \sum_n w_n \langle \xi_n, \psi \rangle \xi_n$$

is obviously selfadjoint, positive,i.e. $\rho \geq 0$ and $\mathrm{tr}\rho = \sum_n w_n = 1$. Such an operator is called a **density matrix**. Every density matrix defines a state ω_ρ on $\mathcal{A}$ as

$$\omega_\rho(a) = \mathrm{tr}(\rho a) \qquad \forall a \in \mathcal{A}.$$

As they are linear functionals, there is a norm on the set of states as

$$\|\omega\| = \sup_{\|a\|=1} |\omega(a)|.$$

Obviously $\|\omega\| \geq \omega(1) = 1$ for all states ω. We shall later see that in fact $\|\omega\| = 1$.

Definition:
A left (right) **ideal** of an algebra $\mathcal{A}$ is a subspace $\mathcal{I}$ such that $ai \in \mathcal{I}$ (respectively $ia \in \mathcal{I}$) for all $a \in \mathcal{A}$ and all $i \in \mathcal{I}$. An ideal is called **two-sided** if it is a left and a right ideal. A **proper ideal** of $\mathcal{A}$ is an ideal that is not equal to $\mathcal{A}$, i.e. $\mathcal{I} \neq \mathcal{A}$. A proper ideal is called maximal if it is not contained in any other proper ideal.

Not that if $\mathcal{A}$ is unital then $1 \notin \mathcal{I}$ if and only if $\mathcal{I}$ is proper. Moreover if $\mathcal{A}$ is a Banach-algebra one easily sees that the closure of a proper ideal is a proper ideal:
Otherwise one would have that $1 \in \bar{\mathcal{I}}$. However that would imply that there exists $x \in \mathcal{I}$ so that $\|a - x\| < 1$ and from this is follows (via the von Neumann series) that $x = 1 - (1 - x)$ is invertible in $\mathcal{A}$. But then $1 = x^{-1}x \in \mathcal{I}$ in contradiction to $\mathcal{I}$ being proper.
Thus maximal ideals are closed. On the other hand, by using Zorn's Lemma, one can also show that each proper ideal is contained in a maximal proper ideal.

If $\mathcal{I}$ is a proper ideal, then the vector space $\mathcal{A}/\mathcal{I}$ is an algebra by setting $[a][b] = [ab]$ which is well-defined if $\mathcal{I}$ is an ideal.

Example: Consider $C_0(X)$ and let $Y \subset X$ be closed. Then the subspace of $C_0(X)$ of all functions that vanish on X is an ideal, and this ideal is

obviously proper unless $X = Y$. Moreover it is maximal if and only if $Y = \{y\}$ is a single point $y \in X$.

2.2.1 Commutative C^*-algebras and the Gelfand-Naimark-Theorem

An important tool for the study of abstract C^*-algebras whose value can not be overestimated is the Gelfand-Naimark Theorem. We shall later see that is also indispensable for the operational approach to quantum mechanics and the interpretation of the latter. In brief it says that an abstract C^*-algebra is commutative if and only if it is the algebra of continuous functions that vanish infinity on some topological Hausdorff space. Henceforth $\mathcal{A}$ will always denote a commutative and unital C^*algebra. (The unital case is somewhat easier to handle, though the statement is also true for nonunital algebras.)

Definition:
A **multiplicative functional** on $\mathcal{A}$ is a homomorphism $x : \mathcal{A} \to \mathbb{C}$, i.e. a $\mathbb{C}$-linear functional on $\mathcal{A}$ such that for all $a, b \in \mathcal{A}$ one has $x(ab) = x(a)x(b)$. Hence in particular $x(1) = 1$. We shall denote the set of all multiplicative functionals on $\mathcal{A}$ by $Spec(\mathcal{A})$, respectively call it the **Spectrum of** $\mathcal{A}$.

Note that multiplicative functionals will rarely exist for noncommutative C^*-algebras. It is important to observe that multiplicative functionals are states.

Proposition:
Let x be any multiplicative functional. Then

$$\|x\| = 1$$

and hence x is continuous.

Sketch of the Proof:
Due to the multiplicativity $ker(x)$ is a two-sided $*$-ideal in $\mathcal{A}$ and hence $\mathcal{A}/ker(x)$ is a $*$-algebra. One may write $x = \tilde{x} \circ \pi$, where π is the natural homomorphism on $\mathcal{A}/ker(x)$, while $\tilde{x}([a]) = x(a)$. Now, it obviously follows from $\tilde{x}([a]) = \tilde{x}([b])$ that $[a] = [b]$. Hence there is an isomorphism $i : \mathcal{A}/ker(x) \to \mathbb{C}$ that is explicitly given by $i([a]) = \tilde{x}([a])$. This implies that $\|\tilde{x}\| = 1$ and hence that $\|x\| \leq \|\tilde{x}\|\|\pi\| \leq 1$. (For a projection π one has $\|\pi(a)\| \leq \|a\|$.) On the other hand we already know that $\|x\| \geq |x(1)| = 1$. Hence the claim follows.

On the space of continuous linear functionals on $\mathcal{A}$ there are, besides the one inherited from the norm, several interesting topologies. among them is the *weak*-$*$-topology, which is also called quite appropriately the topology of pointwise convergence. The precise definition can be found in the references on C^*-algebras, but is of no importance for what follows. Indeed, in the examples we shall consider later, the physical interpretation will always give a clear picture of the topology of the corresponding spaces.
The only important properties of the weak-$*$-topology needed here are:

- it is weaker than the topology inherited by the norm.
- By the Banach-Alaoglu-Theorem the unit ball (of the space of continuous linear functionals on a Banach-space and thus in particular $\mathcal{A}'$) is a compact Hausdorff space in the weak-$*$ topology.

From the above proposition it is clear that $Spec(\mathcal{A})$ is a subset of this unit ball. Moreover this subset is closed, as the multiplicativity $x_n(ab) = x_n(a)x_n(b)$ is obviously a preserved property if the limit $n \to \infty$ exists. Thus we have the

Proposition:
In its relative weak-$*$-topology, $Spec(\mathcal{A})$ is a compact Hausdorff space.

Definition:
To any $a \in \mathcal{A}$ we assign the complex-valued function $\hat{a}$ on $Spec(\mathcal{A})$ via

$$\hat{a}(x) = x(a) \qquad \forall x \in Spec(\mathcal{A}).$$

The map $a \mapsto \hat{a}$ is called the **Gelfand transformation**.

Note that by definition of the weak-*-topology the Gelfand-transform $\hat{a}$ of any $a \in \mathcal{A}$ is continuous.

Theorem:
For every $a \in \mathcal{A}$ it is

$$\sigma_{\mathcal{A}} = \{\hat{a}(x) \,|\, x \in Spec(\mathcal{A})\}.$$

Proof:
It is clear that for any $a \in \mathcal{A}$ and $\lambda \in \mathbb{C}$ one has that the Gelfand transform of $a - \lambda$ is $\hat{a} - \lambda$ and that $\sigma(a - \lambda) = \sigma(a) - \lambda$. Hence it is sufficient to prove that: **an element $a \in \mathcal{A}$ is invertible if and only if $\hat{a}$ is a non vanishing function.** This is proven as follows:

"$\Leftarrow$": If a is invertible there exist a^{-1} and since $\hat{a}(x)\hat{a^{-1}}(x) = 1$ it follows that $\hat{x} \neq 0$ for all $x \in Spec(\mathcal{A})$..
Suppose now that a is not invertible. Then the two-sided ideal $a\mathcal{A} = \{ab \,|\, b \in \mathcal{A}\}$ (remember that $\mathcal{A}$ is commutative) is a proper ideal (as it does not contain 1). Hence $a\mathcal{A}$ is contained in a (closed) maximal proper ideal $\mathcal{I}$. Consider $\mathcal{B} = \mathcal{A}/\mathcal{I}$. Since $\mathcal{I}$ is maximal this algebra $\mathcal{B}$ must be simple, i.e cannot contain a proper ideal. Moreover it is commutative. Thus all elements b of $\mathcal{B}$ are invertible (because the ideal $b\mathcal{B} = \mathcal{B}$ contains 1). By the Gelfand-Mazur Theorem we then have that $\mathcal{B}$ is isomorphic to $\mathbb{C}$. Since $a \in \mathcal{B}$ we may indeed define a natural isomorphism $i : \mathcal{B} \to \mathbb{C}$ by setting $i([\lambda a]) = \lambda$. Then $\mathcal{I} = ker(i)$ and i can be extended to an element x of $Spec(\mathcal{A})$ as $x(b) = i([b])$. But then $\hat{a}(x) = x(a) = 0$ since $a \in a\mathcal{A} \subseteq \mathcal{I}$. Hence $\hat{a}$ vanishes at least in one point in $Spec(\mathcal{A})$.

Corollary:
The states on a (not necessarily commutative) C^*-algebra $\mathcal{A}$ separate the elements of $\mathcal{A}$.

Proof:
Any element a of $\mathcal{A}$ can be uniquely written as $a = a_1 + ia_2$ where a_1, a_2 are selfadjoint and thus normal. Suppose now that $a \neq b$. Thus at least one of the two selfadjoint operators $c_i = a_i - b_i$ $(i = 1, 2)$ does not vanish. Now any of the c_i does generate a commutative C^*-algebra and by the Gelfand transformation and the above Theorem, this algebra can be viewed as a subalgebra of $C(Spec(\mathcal{A}))$. But the points of $Spec(\mathcal{A})$ separate all the elements of $C(Spec(\mathcal{A}))$. Thus there exists a multiplicative functional x_i on the subalgebras generated by each of the c_i with the property that $x_i(c_i) \neq 0$. It is easy to see that x_i can be extended to a state on all of $\mathcal{A}$ by applying the Hahn-Banach-Theorem.

Theorem (Gelfand-Naimark):
The Gelfand transformation is an isometric $*$-isomorphism of $\mathcal{A}$ onto $\mathbb{C}$.

Proof:
It remains to prove that the Gelfand transformation is *onto* and *isometric*.
Let us first prove the latter assertion: The norm in $C(X)$ (for any topological space X) is given as $\|f\|_\infty = \sup_{x \in X} |f(x)|$. Hence, using the above theorem we have for the Gelfand transform $\hat{a}$ f any $a \in \mathcal{A}$:

$$\|\hat{a}\|_\infty = \sup_{x \in Spec(\mathcal{A})} |\hat{a}(x)| = \sup_{\lambda \in \sigma(a)} |\lambda| = \rho_{\mathcal{A}}(a) = \|a\|.$$

The last equality follows from the spectral radius formula because every element of a commutative C^*-algebra is normal.
Next, we show that for all $a \in \mathcal{A}$ on has $\hat{a^*} = \overline{\hat{a}}$, by considering the exponential series e^{ita} for selfadjoint a and $t \in \mathbb{R}$ (The existence of the exponential series is easy to see.) Since by multiplicativity and linearity for all $x \in Spec(\mathcal{A})$ one has $x(e^{ita}) = e^{it(\hat{a}(x))}$ the unitarity of e^{ita} implies that $\hat{a}(x)$ is real. Since every element of $\mathcal{A}$ can be written as a complex linear combination of two selfadjoint elements the assertion follows. Thus the Gelfand transform preserves the involution $*$.
Moreover, by definition the Gelfand transform separate the elements of $Spec(\mathcal{A})$. (Because two multiplicative functionals are only different if they differ on some $a \in \mathcal{A}$.) As the constant functions are in the image of the Gelfand transformation, we may apply the Stone-Weierstrass-Theorem (all suppositions thereof are met) which implies that the subset $\{\hat{a} \,|\, a \in \mathcal{A}\}$ is dense in $C(Spec(\mathcal{A}))$. As it is also a closed subset, the statement of the Theorem follows.

Corollary (Functional calculus):
Let a be a normal element of a (not necessarily commutative) C^*-algebra $\mathcal{A}$. Then any continuous function $f(a, a^*)$ of a and a^* defines an element of $\mathcal{A}$, and

$$\sigma_{\mathcal{A}}(f(a, a^*)) = \hat{f}(\sigma_{\mathcal{A}}(a))$$

Now from this Corollary we obviously get that $a = b^2$ for some $b = b^* \in \mathcal{A}$ if and only if $\sigma_{\mathcal{A}}(a) \subseteq [0, \|a\|]$. One may also show (using some easy algebra) that the latter property is equivalent to the existence of $c \in \mathcal{A}$ such that $a = c^*c$. In any case we have for such positive elements a that $\|a\| - a \geq 0$. Hence for every state ω by positivity of ω :

$$\omega(a) \leq \|a\|\omega(1)$$

But then it is not hard to prove that for all $c \in \mathcal{A}$:

$$|\omega(c)|^2 \leq \|c\|^2$$

by considering $a = c^*c$. Hence one gets:

Corollary: All states on a C^*-algebra are bounded.

2.2.2 The GNS-representation

Definition:
A **representation** of a C^*-algebra $\mathcal{A}$ is an (isometric) *-homomorphism

$\pi : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$.
One calls a representation **faithful** if $ker(\pi) = \{0\}$.
A representation is **irreducible** if no proper subspace is invariant under the action of $\pi(\mathcal{A})$.
A cyclic representation is a representation that admits a **cyclic vector** $\psi \in \mathcal{H}$, i.e. $\pi(\mathcal{A})\psi$ is dense in $\mathcal{H}$.

Note that in an irreducible representation every vector is cyclic. However, the existence of a cyclic vector does not imply that the representation is irreducible, nor that all vectors are cyclic. (Consider for instance commutative C^*-algebras.)

Theorem (Gelfand-Naimark-Segal):
Given a state ω on the unital C^*-algebra $\mathcal{A}$ there exists a Hilbert space $\mathcal{H}_\omega$ and a representation $\pi_\omega : \mathcal{A} \to \mathcal{L}(\mathcal{H}_\omega)$ such that

1. There exists a cyclic vector ψ_ω for π_ω.
2. $\omega(a) = \langle \psi_\omega, \pi_\omega \psi_\omega \rangle$ for all $a \in \mathcal{A}$.
3. Every other representation π on some Hilbert space $\mathcal{H}$ such that there exists a cyclic vector ψ with $\omega(a) = \langle \psi, \pi(a)\psi$ for all $a \in \mathcal{A}$ is **unitarily equivalent** to the above representation, i.e there exists a unitary operator $U : \mathcal{H} \to \mathcal{H}_\omega$ such that for all $a \in \mathcal{A}$:

$$U\pi(a)U^* = \pi_\omega(a) \qquad \psi_\omega = U\psi.$$

Sketch of the Proof:
Given a state ω one may define a positive semidefinite inner product on $\mathcal{A}$ by

$$(a, b) := \omega(a^* b).$$

However, $(\cdot, \cdot)$ will be degenerated in general. Now set

$$\mathcal{I} := \{a \in \mathcal{A} \,|\, \omega(b^* a) = 0 \qquad \forall b \in \mathcal{A}\},$$

which obviously is a left ideal in $\mathcal{A}$
Hence we can consider the vector space $\mathcal{A}/\mathcal{I}$ on which the restriction $\langle \cdot, \cdot \rangle$ of $(\cdot, \cdot)$ is obviously nondegenerate and positive definite. We may thus complete $\mathcal{A}/\mathcal{I}$ with respect to $\langle \cdot, \cdot \rangle$ to obtain a Hilbert space $\mathcal{H}$.
The representation π_ω of $\mathcal{A}$ on $\mathcal{H}$ is the defined as

$$\pi_\omega(a)[b] = [ab].$$

One easily checks that π_ω is well defined. Moreover from the inequality $\omega(a^*a) \leq \|a\|^2$ we proved above, one infers that $\|\pi_\omega(a)\| = \|a\|$.
Finally, $\psi_\omega := [1]$ is obviously a cyclic vector and

$$\langle \psi_\omega, \pi_\omega(a)\psi_\omega \rangle = \omega(1^*a) = \omega(a)$$

The stated unitary equivalence is given by

$$U^*\pi_\omega(a)\psi_\omega = \pi(a)\psi \qquad\qquad \forall a \in \mathcal{H}$$

which obviously defines U uniquely (and well) due to the cyclicity of the representations π and π_ω.

Theorem (Gelfand-Naimark:
Every C^*-algebra admits a faithful representation on some Hilbert space $\mathcal{H}$.

Sketch of the Proof:
As the states do separate the elements of a C^*-algebra there exists a family $\mathcal{F}$ of states which also does so. Then one may consider the sum of the GNS-representations for elements of $\mathcal{F}$, i.e. $\mathcal{H} = \bigoplus_{\omega\in\mathcal{F}} \mathcal{H}_\omega$ and $\pi = \bigoplus_{\omega\in\mathcal{F}} \pi_\omega$. This will obviously provide a faithful representation.

Proposition:
The GNS representation defined by a state is irreducible if and only if the state is pure.

Proof:
"$\Leftarrow$": If the GNS-representation is reducible, then there would exist a projector P on the invariant subspace. Setting $\lambda = \|P\psi_\omega\|^2$, and $\omega_1(a) = \langle P\psi_\omega, \pi_\omega(a)P\psi_\omega\rangle$ and $\omega_2(a) = \langle (1-P)\psi_\omega, \pi_\omega(a)(1-P)\psi_\omega\rangle$ one easily checks that $\langle \psi_\omega, \pi_\omega(a)\psi_\omega\rangle = \lambda\omega_1(a) + (1-\lambda)\omega_2(a)$.
"$\Rightarrow$": Let now be ω such that the resulting GNS-representation is irreducible. Suppose $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$ were not pure. Thus for all $a \in \mathcal{A}$ we have:

$$\lambda\omega_1(a^*a) \leq \omega(a^*a) \leq \|a\|^2.$$

But then from the cyclicity of the representation and the Riesz-representation theorem it follows that there exists a bounded operator T on $\mathcal{H}$ such that for all $a, b \in \mathcal{A}$ we have

$$\lambda\omega_1(a^*b) = \langle \psi_\omega, \pi_\omega(a^*)T\pi_\omega(b)\psi_\omega\rangle =: \tau(a^*Tb).$$

Then, for all $a, b, c \in \mathcal{A}$

$$\tau(a^*Tcb) = \lambda\omega_1(a^*cb) = \lambda\omega_1((c^*a)^*b) = \tau((c^*a)^*Tb) = \tau(a^*cTb).$$

Using again the cyclicity of the representation, we get from its irreducibility that $T \sim 1$ and thus $\omega = \omega_1$.

Corollary:
A state on a commutative C^*-algebra is commutative if and only if it is multiplicative.

This follows easily, because for a commutative algebra all irreducible representations are one-dimensional, while conversely the GNS-representation of a multiplicative state is easily seen to be one-dimensional.

2.3 Unbounded Operators on Hilbert spaces

In physics it is often necessary, or at least useful, to work with operators that are unbounded and are thus at best defined on dense subsets of $\mathcal{H}$. Typical examples of such Operators include Differential Operators, in particular the momentum- and Hamilton-Operators, the latter being the generators of time translations. In fact, as we shall see in this section, Stone's Theorem asserts that every strongly continuous group of unitaries (to be defined below) is generated by a possibly unbounded selfadjoint operator. For instance the operator $i\frac{\partial}{\partial x}$ defined e.g on appropriate sets of smooth functions $f(x)$ on the real line $\mathbb{R}$ generates translations, $f(x) \to f(x+\alpha)$, where $\alpha \in \mathbb{R}$ can be arbitrarily chosen. However, as it is local operator, on might also consider its action on smooth functions on some subset $X \in \mathbb{R}$. However then the above translations need not be well defined. According to Stone's Theorem the Operator $i\frac{\partial}{\partial x}$ can thus only be well-defined as a selfadjoint-operator if one finds a way to make the translations for all α well defined. As we shall see, this is indeed possible if X is open and also if X is compact – by fixing appropriate boundary conditions. In the latter case different choices of boundary conditions will lead to different operators, however. For other choices of X, for instance the non-negative half-line it is impossible to define $i\frac{\partial}{\partial x}$ as a selfadjoint operator.

2.3.1 Densely defined operators on Hilbert space.

As usual $\mathcal{H}$ shall denote a Hilbert space, while $\mathcal{D}$ will denote a subset of $\mathcal{H}$. Any complex linear map $A : \mathcal{D} \mapsto \mathcal{H}$ will be called a **partially defined operator with domain** $\mathcal{D}$. We shall write $\mathrm{dom}(A) = \mathcal{D}$. If the domain $\mathcal{D}$ of A is dense in $\mathcal{A}$, one says that A is densely defined. Note that for $\xi \in \mathcal{D}$ we have $\|A\xi\| < \infty$ by definition.

For any densely defined operator A there is a natural definition of its **adjoint**:
Let $D(A^*)$ denote the subset of all $\xi \in \mathcal{H}$ for which there exists a $\xi^* \in \mathcal{H}$ such that

$$\langle A\eta, \xi\rangle = \langle \eta, \xi^*\rangle \qquad\qquad \forall \eta \in \mathrm{dom}(A).$$

Note that for $\xi \in \mathcal{H}$ there is, by density of the domain of A, at most a unique ξ^* with the above property. Moreover $D(A^*)$ is dense in $\mathcal{H}$ by Riesz' Lemma. Thus, setting $A^*\xi = \xi^*$ densely defines an operator A^*, called the **adjoint of** A in $\mathcal{H}$, with $\mathrm{dom}(A^*) = D(A^*)$.
Note that from the definition of A^* it also immediately follows that the kernel of A^* is the orthogonal complement of the range of A.

Definition:
A densely defined operator A is called **symmetric** if

$$\langle \xi, A\eta\rangle = \langle A\xi, \eta\rangle \qquad\qquad \forall \xi, \eta \in \mathrm{dom}A.$$

Theorem (Hellinger-Toeplitz):
A symmetric operator A with $\mathrm{dom}(A) = \mathcal{H}$ is automatically bounded.

In other words an unbounded symmetric operator is necessarily only densely defined. As for the proof of the above theorem one notes that if an everywhere defined symmetric operator A would be unbounded, there would exist a sequence $\chi_n \in \mathrm{dom}(A) = \mathcal{H}$ so that $\|\chi_n\| = 1$ for all n and $\|A\chi_n\| \to \infty$. But then, for all $\xi \in \mathcal{H}$ we have

$$|\langle \xi, A\chi_n\rangle| = |\langle A\xi, \chi_n\rangle| \leq \|A\xi\|\|\chi_n\| = \|A\xi\| < \infty.$$

It is an easy exercise to conclude that then $\|A\chi_n\|$ must be bounded, in contradiction to the hypothesis.

Definition:
If A and B are densely defined operators we shall say that B is an **extension** of A if $\mathrm{dom}(A) \subset \mathrm{dom}(B)$ and $A = B$ on the common domain $\mathrm{dom}(A)$.

Let A be an unbounded operator, B bounded. One might be tempted to define the commutator of A and B as $[A, B] = AB - BA$, just as in the case of two bounded operators. In fact one does use the same notation. However, the meaning is necessarily slightly different:
Note that $AB = BA$ as operators would imply that the domains of the two operators are equal. Now suppose $B = 0$ is the trivial operator that maps

all $\xi \in \mathcal{H}$ to 0. Then the domain of AB is all of $\mathcal{H}$, while BA is only defined on the domain of BA. Hence $A0 \neq 0A$. One therefore defines:

Definition:
We say that an unbounded operator A **commutes** with the bounded operator B if AB is an extension of BA.
We shall call the subspace

$$\{A\}' := \{B \in \mathcal{L}(\mathcal{H}) \,|\, B \text{ commutes with } A\}$$

the commutant of A. Accordingly, given a collection $\mathcal{T}$ of unbounded operators we shall write $\mathcal{T}'$ for the intersection of the commutants of all $T \in \mathcal{T}$. If $\mathcal{A}$ is a concrete C^*-algebra we shall say that such a collection $\mathcal{T}$ is **affiliated** to $\mathcal{A}$ if

$$\mathcal{A}' \subset \mathcal{T}'.$$

Affiliated operators may be viewed as unbounded generators of subalgebras of $\mathcal{A}$. Thus, e.g. the position operator x on $L^2(\mathbb{R})$ is affiliated to $C_0(\mathbb{R})$. In general, if A and B are densely defined and if there is a common dense domain $\mathcal{D}$of the operators AB and BA, we shall write $[A, B]$ for the operator $AB - BA$ with domain $\mathcal{D}$. Whenever we write $[A, B]$ in the sequel this interpretation is implied. We shall later see in a concrete example that severe contradictions arise if the commutator of two unbounded operators is used carelessly.

Definition:
We shall call a densely defined operator A **closed** if for any sequence $\xi_n \in$dom(A) with

$$\xi_n \to \xi \qquad \text{and} \qquad A\xi_n \to \eta$$

one has that $\xi \in$dom(A) and $A\xi = \eta$.

Example:
A^* is automatically closed: If $\xi_n \in$dom(A^*) converges to ξ such that $A\xi_n$ converges to η one has for all $\zeta \in$dom(A):

$$\langle A\zeta, \xi\rangle = \lim\langle A\zeta, \xi_n\rangle = \lim\langle \zeta, A^*\xi_n\rangle = \langle \zeta, \eta\rangle$$

i.e. $\zeta \in$dom(A^*) and $A\zeta = \eta$.

Definition:
A densely defined operator is called **closable** if it has a closed extension.

The closure $\bar{A}$ is the smallest closed extension.

A densely defined operator A is called **selfadjoint** if $A = A^*$, i.e if it is symmetric and $\mathrm{dom}(A) = \mathrm{dom}A^*)$.
A densely defined operator A is called **essentially selfadjoint** if $\bar{A}$ is self-adjoint.

In other words, an essentially selfadjoint operator admits a unique **self-adjoint extension**.

Lemma:
A symmetric Operator A on $\mathcal{H}$ is selfadjoint if and only if

$$Ran(A \pm i) = \mathcal{H}$$

Proof:
"$\Leftarrow$": Let's first assume that $Ran(A \pm i) = \mathcal{H}$ i.e. that for any $\phi \in \mathcal{H}$ there exists $\eta \in D(A)$ such that $\phi = (A - i)\eta = (A^* - i)\eta$. (The last equality follows because $D(A) \subset D(A^*)$. This then holds in particular for every $\xi \in D(A^*)$. But then one would have $(A^* - i)(\xi - \eta) = 0$, which implies $\xi = \eta$ because the kernel of $A^* - i$ is the orthogonal complement of the image of $A + i$ and thus only contains 0 by assumption.
In conclusion $D(A^*) = D(A)$ and thus A is selfadjoint.
"$\Rightarrow$" Let's now assume that A is selfadjoint. Then $\langle \xi, A\xi$ is real for all $\xi \in D(A)$, and hence $\langle \xi, (A \pm i)\xi = 0$ implies $\xi = 0$ (by density of $D(A)$). Thus the kernel of $A \pm i$ is trivial. Moreover the Range of $A \pm i$ is closed:
First of all the reality of $\langle \xi A\xi \rangle$ for all ξ implies that for any sequence $\xi_n \in D(A)$:

$$\|(A \pm i)(\xi_n - \xi_m)\|^2 = \|A(\xi_n - \xi_m)\|^2 + \|\xi_n - \xi_m\|^2.$$

Hence if $(A \pm i)\xi_n$ converge in norm, so do $A\xi_n$ as well as ξ_n. But then since A is closed, this implies that $Ran(A \pm i)$ are closed. This completes the proof.

Similarly one shows that a symmetric Operator A on $\mathcal{H}$ is essentially selfadjoint if and only if

$$\overline{Ran(A \pm i)} = \mathcal{H}$$

(In fact, by simply applying the above theorem to $\bar{A}$.)

Definition:
The two numbers :

$$n_\pm = \mathrm{dim}ker(A^* \mp i)$$

are called the **deficiency indices** of A.

Theorem:
A symmetric operator A on $\mathcal{H}$ admits selfadjoint extensions if and only if $n_+ = n_-$. These selfadjoint extensions $\tilde{A}_U$ are then described as follows:
Pick a unitary operator $U : ker(A^* - i) \mapsto ker(A^* + i)$. Then the domain of $\tilde{A}_U$ is given as

$$D(\tilde{A}_U) = \{\eta = \eta_0 + \eta_+ + U\eta_+ \,|\, \eta_0 \in D(A),\ \ \eta_+ \in ker(A^* - i)\}.$$

$\tilde{A}_U$ acts as $\tilde{A}_U\eta = A\eta_0 + i\eta_+ - iU_A\eta_+$ on this domain.

Examples:

1. Let us – as a first example – consider the operator $i\frac{\partial}{\partial x}$ on $L^2([0,1])$. We shall try to define it as a selfadjoint and thus in particular symmetric operator. Now, for two differentiable functions ψ, φ on $[0,1]$ one easily computes :
$$\langle \psi, i\frac{\partial}{\partial x}\varphi\rangle = \langle i\frac{\partial}{\partial x}\psi, \varphi\rangle - i\,\overline{\psi}\varphi\Big|_0^1.$$
One might thus be tempted to take as a domain the space of all smooth functions of compact support on $(0,1)$. However this space is not dense in $L^2([0,1])$. (It's closure has codimension 1).
A family of dense domains is described as:
$$D_\theta = \{\psi \in L^2([0,1]) \,|\, \psi \text{ is absolutely continuos and } \psi(0) = e^{i\theta}\psi(1)\}$$
Indeed, absolutely continuous functions can be characterized as those who admit a derivative almost everywhere. Thus $i\frac{\partial}{\partial x}$ is closed and symmetric on D_θ. We shall denote the corresponding densely defined operator by $\dot{x}_\theta$. Moreover the spaces D_θ contain the functions $\psi_{n,\theta} = e^{i(2\pi n - \theta)x}$, $n \in \mathbb{Z}$ which form an orthogonal basis of $L^2([0,1])$ by Fourier decomposition.
In fact the $\psi_{n,\theta}$ are eigenfunctions of $\dot{x}_\theta$ to the eigenvalues $\lambda_{n,\theta} = 2\pi - \theta$. This immediately implies selfadjointness. In conclusion:
$i\frac{\partial}{\partial x}$ admits uncountably many different selfadjoint extensions on $L^2([0,1])$.
Note that the multiplication operator x does not leave any of the domains D_θ invariant, as it is not compatible with the boundary conditions. Hence, in this case the commutator $[x, \dot{x}_\theta]$ makes no sense at all. Indeed, otherwise we could conclude that $\langle\psi_n, [x, \dot{x}_\theta]\psi_n\rangle = \lambda_n\langle\psi_n, (x - x)\psi_n\rangle = 0$ for all n and thus on all of $L^2([0,1])$.

2. Next we consider again the operator $i\frac{\partial}{\partial x}$ but this time on $L^2([0,\infty))$. The differential equation

$$(i\frac{\partial}{\partial x} - i)\psi = 0$$

then has a unique square integrable solution, while the equation $(i\frac{\partial}{\partial x} + i)\psi = 0$ has none. Thus $n_+ \neq n_-$ and therefore there do not exist selfadjoint extensions in this case.

3. Similarly it is immediately clear that $i\frac{\partial}{\partial x}$ is essentially selfadjoint on $L^2(\mathbb{R})$, as $n_+ = n_- = 0$. A closed dense domain is given by the set of all L^2- functions whose derivative is also L^2.

4. Let's now return to $L^2([0,\infty))$. Yet this time we consider the operator $-\frac{\partial^2}{\partial x^2}$. One checks that $n_+ = n_- = 1$ in this case. Square integrable solutions being given by $e^{-\sqrt{\pm i}x}$ where any complex square root can be taken. Thus this operator does admit a unique selfadjoint extension. Note that $-\frac{\partial^2}{\partial x^2}$ is positive. Thus this result is a special case of

5. **Friedrichs' Theorem:**
Every semi-bounded symmetric operator admits semi-bounded selfadjoint extension.

Here an operator A is called semi-bounded if $\langle \xi, A\xi \rangle \geq C$ (respectively $\leq$) for some finite $C \in \mathbb{R}$ and all ξ in the domain of A.

2.3.2 Spectral decomposition of unbounded selfadjoint operators and Stone's Theorem

The spectral decomposition and functional calculus for selfadjoint operators can be generalized to unbounded selfadjoint operators, even though that may not be apparent at first sight. For instance it is not clear at this point that the operator $|A| = \sqrt{A^2}$ is well defined for selfadjoint A, and thus also the existence of the spectral projections is not obvious at all. Yet, there are various methods to construct these spectral projections. The quickest and perhaps most elegant one goes back to von Neumann, and is in fact also the oldest one:
Let A be a selfadjoint operator, and thus $Ran(A - i) = \mathcal{H}$. Then $A - i$ is invertible and one easily proves that the following operator is in fact *unitary*:

$$C(A) = \frac{A + i}{A - i}.$$

(Unitarity follows from $\|(A+i)\xi\| = \|(A-i)\xi\|$.) This operator is called the Cayley-Transform of A. Next it is readily seen that 1 is not (i.e. for no A) in the spectrum of $C(A)$. (Otherwise the operator $(A-i)(C(A)-1) = -2i$ would not be invertible.) Even more so, one has the identity

$$A = i\frac{C(A)+1}{C(A)-1}.$$

Finally it is also easy to check that the operators A and $C(A)$ do commute and in fact that $C(A)$ commutes with any bounded operator that commutes with A.
Thus it follows (with some more work) that, given a bounded Borel set $\Delta \subset \mathbb{R}$, setting $P_A(\Delta) := P_{C(A)}(C(\Delta))$ (with an obvious notation) will give the spectral projections of A.
We still need to define the integral over $\sigma(A)$ (which is no longer a compact set). Let Δ_k be an increasing sequence of finite Borel sets, such that $\cup\Delta_k = \mathbb{R}$ and let P_k be the corresponding spectral projection for A. The the set $\cup P_k\mathcal{H}$ is dense in $\mathcal{H}$. Now for each k one can define

$$A_k = \int_{\Delta_k} \lambda \, \mathrm{d}P(\lambda),$$

and obviously $A_k = A|_{P_k\mathcal{H}}$ and in fact (by letting $k \to \infty$) A is the closure of $A|_{\cup P_k\mathcal{H}}$. We may therefore define the integral as the closure of the limit of Riemann-Integrals over sequences of bounded Borel sets.

Theorem:
Let A be a selfadjoint operator in $\mathcal{H}$. Then there exists a projection valued measure $P(\Delta)$ on the Borel subsets of $\mathbb{R}$, consisting of projections that commute with A and all bounded operators that commute with A, such that

$$A = \int_{-\infty}^{\infty} \lambda \, \mathrm{d}P(\lambda).$$

Remark:
It is important to note at this point that the above derivation makes essential use of the selfadjointness of A. In fact, there is no spectral decomposition of symmetric operators, in general. This is, of course, also clear from the fact that the spectra of the different selfadjoint extensions (if they exist) may be very different.

Next by using the same procedure of choosing an increasing sequence of finite Borel-sets, using the bounded spectral decomposition theorem, and then defining the desired operator as the closure of the limiting operator, one also has

Theorem:
For any Borel measurable function $f : \mathbb{R} \mapsto \mathbb{C}$ the expression

$$f(A) = \int_{\infty}^{\infty} f(\lambda) \mathrm{d}P(\lambda)$$

densely defines an operator $f(A)$ (with a proper interpretation of the integral).

One says that a Borel set Δ has A-measure zero if $P(\Delta) = 0$. Accordingly we shall say for two Borel-measurable functions f, g that $f = g$ A-almost-everywhere if equallity holds up to a set of A-measure zero. In that case, obviously $f(A) = g(A)$. Moreover:

Proposition:
If A is a selfadjoint operator and f a Borel measurable function on $\mathbb{R}$.

- $f(A)$ is symmetric if and only if is real-valued A-almost-everywhere.
- $f(A)$ is bounded if and only if f is bounded A-almost-everywhere.
- $f(A)$ is positive if and only if f is positive A-almost-everywhere.
- $f(A)$ is unitary if and only if $|f(\lambda)| = 1$ A-almost-everywhere.

Note that we made no statement about the selfadjointness of $f(A)$, as the dependence of the domain of $f(A)$ on f is not clear in general.

It follows in particular that the expression

$$U_A(t) := e^{iAt} = \int_{-\infty}^{\infty} e^{i\lambda t} \, \mathrm{d}P(\lambda)$$

is a family of unitary operators for all selfadjoint A. In fact it is a **strongly continuous group of unitaries**,i.e.

$$U_A(t)U_A(s) = U_A(t + s)$$

for all $t, s \in \mathbb{R}$ and the function

$$t \to \|(U_A(t) - 1)\xi\|$$

is continuous.
Vice versa, given a strongly continuous group of unitaries $U(t)$, one may define the generator A via

$$A\psi = -i\frac{\partial}{\partial t}\left(U(t)\psi\right)\Big|_{t=0}$$

on an appropriate domain. In fact one has to show that this operator will be densely defined. However, one easily checks that for any $\xi \in \mathcal{H}$ and every compactly supported smooth function f the vector $\xi_f = \int(f(t)U(t)\xi)$ is well defined (due to the strong continuity of $U(t)$ which is essential here) and in the domain of A. The set of such vectors is then obviously dense in $\mathcal{H}$, as one may approximate the Dirac-δ-distribution by compactly supported smooth functions. One easily verifies from the unitarity of the $U(t)$ that A is symmetric and that its deficiency indices vanish. Indeed, to see the latter statement, if $\xi_\pm$ is in the kernel of $(A \pm i)$, then one would get that for all ϕ in the domain of A one has $\frac{\partial}{\partial t}\chi_\pm(t) = \chi_\pm(t)$, where $\chi_\pm(t) = \langle \xi_\pm, U(t)\phi\rangle$. Thus the unitarity of $U(t)$ implies $\chi_\pm(t) = 0$. (Otherwise $\chi_\pm(t) = \chi_\pm(0)e^{\pm t}$ would give a contradiction.)
Finally, one easily shows that if $U(t) = e^{iBt}$ for some selfadjoint operator B, then $A = B$. This sufficiently motivates:

Stone's Theorem:
If $U(t)$ is a strongly continuous group of unitaries then there exists a selfadjoint operator A such that $U(t) = e^{iAt}$.
Vice versa, if A is selfadjoint, then $U_A(t) = e^{iAt}$ defines a strongly continuous group of unitaries.
The assignment $t \to U(t)$ is norm continuous, i.e $\|U_A(t) - 1\|$ is a continuous function of t if and only if A is bounded.

2.3.3 The pure point and the continuous spectrum of selfadjoint operators

We shall now make the relation of the spectrum of an (un)bounded selfadjoint operator and its eigenvalues more precise. After all, the points in the spectrum aren't as bad as their name might indicate[1]. In the sequel A can be any selfadjoint operator, bounded or unbounded. We will continue to call the set of all $z \in \mathbb{C}$ such that $(A-z)^{-1}$ exists as a *bounded* operator the resolvent set $r(A)$ of A, even if A is unbounded. The spectrum $\sigma(A) = \mathbb{C} \setminus r(A)$ of A will then be a closed set.

[1]The name spectrum is derived from the Greek word for "ghosts".

Now, from the spectral decomposition and functional calculus theorems it is clear that the resolvent, if it exists, is given as

$$(A-z)^{-1} = \int_{-\infty}^{\infty} \frac{1}{\lambda - z}\, \mathrm{d}P(\lambda),$$

and this integral exists if and only if the function $\frac{1}{\lambda - z}$ is bounded on the support of the measure given by the spectral projections P_λ of A. Thus, in particular the resolvent will exist whenever z has a non vanishing imaginary part, i.e. $z \in \mathbb{C} \setminus \mathbb{R}$.
Recall that the spectral projections can be viewed as the orthogonal projections on the closed subspace on which $A \geq \lambda$, i.e. $A - \lambda$ is positive, while the measure is then, as usual defined as the limit $\mu \to \lambda$ of the operators $P(\Delta) = P_\lambda - P_\mu$ for Borel subsets $[\lambda, \mu)$. Thus $\lambda \in \mathbb{R}$ will not lie in the support of the spectral measure of A if and only if the projection valued function $\lambda \to P_\lambda$ is constant on some finite Borel set $[\mu_1, \mu_2)$ which contains λ, i.e. $\mu_1 \leq \lambda \leq \mu_2$.
In other words the spectrum of A consists of all $\lambda \in \mathbb{R}$ for which no finite Borel set Δ exists such that $\lambda \in \Delta$ and $P(\Delta) = 0$. In particular those real λ belong to $\sigma(A)$ at which the function P_μ "jumps", i.e. for which there exist $\mu_- < \mu_-$ so that P_μ is constant on $[\mu_-, \lambda)$ and $[\mu_0, \mu_+)$ for all $\mu_0 > \lambda$. One may then consider the projection

$$P(\lambda) = P_\lambda - P_{\mu_+} = P_{\mu_-} - P_\lambda$$

and the corresponding subspace $\mathcal{H}_\lambda = P(\lambda)\mathcal{H}$. Then A leaves $\mathcal{H}_\lambda$ invariant and

$$A\xi = \lambda\xi \qquad\qquad \forall \xi \in \mathcal{H}_\lambda.$$

Such values $\lambda \in \sigma(A)$ are called the **characteristic values** of A, but are sometimes also called the eigenvalues of A. Note however that the **characteristic subspaces** $\mathcal{H}_\lambda$ need not be finite dimensional.
Obviously, if $\mathcal{H}$ is separable A can have at most denumerable many different characteristic values, and characteristic subspaces $\mathcal{H}_\lambda$ and $\mathcal{H}_\mu$ corresponding to different characteristic values $\lambda \neq \mu$ are orthogonal.
A selfadjoint operator A on a separable Hilbert space need not have any characteristic values at all. In that case one says that A has **purely continuous spectrum**. In the other extreme case, when the characteristic subspaces of A add up to $\mathcal{H}$, and there thus exists an orthonormal basis ξ_n of $\mathcal{H}$ such that $A\xi_n = \lambda_n \xi_n$ for some $\lambda_n \in \mathbb{R}$ and all n, one says that A has **completely discrete spectrum**. Sometimes operators of the latter type are said to have

pure point spectrum.
Let $\lambda_1, \lambda_2, \ldots$ be a complete set of characteristic values of A and consider the projectors

$$P_d := \sum_k P_{\lambda_k} \qquad \text{and} \qquad P_c := 1 - P_d.$$

Note that P_d is well defined due to the orthogonality of the projectors P_λ for different values of λ. Obviously the spaces $\mathcal{H}_d = P_d\mathcal{H}$ and its orthogonal complement $\mathcal{H}_c$ are invariant under A. The assumed completeness of the set $\{\lambda_k\}$ then requires that the restriction of A to $\mathcal{H}_c$ has purely continuous spectrum, while the restriction of A to $\mathcal{H}_d$ has purely discrete part of the spectrum. The spectrum of the operator AP_d is called the **discrete part of the spectrum of** A, the spectrum of AP_c is the **continuous part of the spectrum of** A.

Among the characteristic values of A are those with finite multiplicity. The remainder of $\sigma(A)$ are called the limit points of $\sigma(A)$. In other words:

Definition:
Let A be a selfadjoint operator. We call all the points of the continuous spectrum, all characteristic values of infinite degeneracy and all limit points of the discrete spectrum of A the **limit points of** $\sigma(A)$. In particular, if A is not bounded from below (resp. above) then $-\infty$ (respectively $+\infty$) is a limit point of A.

Note that by definition, if $\lambda \in \sigma(A)$ is not a limit point then the corresponding projector $P(\lambda) = P_\lambda - P_{\mu_+}$ (see above) is of finite rank. Vice versa, suppose $\mu \in \sigma(A)$ is a limit point. Suppose that there exists a Borel set Δ with $\mu \in \Delta$ and such that $P(\Delta)$ is of finite rank. Then, obviously the projection valued function P_λ must jump at a finite number of points in Δ. Thus if μ is a limit point of $\sigma(A)$, then for all Δ with $\mu \in \Delta$ the corresponding projector $P(\Delta)$ must be of infinite rank.

Theorem (Weyl):
Let $\mu \in \mathbb{R}$ be finite and A be a selfadjoint operator on the separable Hilbert space $\mathcal{H}$. Then μ is a limit point of $\sigma(A)$ if and only if there exists a sequence $\eta_n \in \mathrm{dom}(A)$ such that

$$\|\eta_n\| = 1, \qquad \eta_n \to 0 \qquad \text{and} \quad (A - \mu)\eta_n \to 0.$$

Sketch of the Proof:
"$\Rightarrow$": If μ is a limit point then by the above discussion for all Borel sets Δ

with $\mu \in \Delta$ the spectral projection $P(\Delta)$ is of infinite rank. One may then choose a sequence Δ_n of nested Borel set , $\Delta_1 \subset \Delta_2 \subset \cdots$, which contract to μ, i.e $\mu = \cap_n \Delta_n$. Let $\mathcal{H}_n := P(\Delta_n)\mathcal{H}$ and choose an orthonormal sequence $\eta_n \in \mathcal{H}_n$, i.e $\langle \eta_n, \eta_m \rangle = 0$ for $n \neq m$. Such a sequence exists because each $\mathcal{H}_n$ is infinite dimensional. One easily checks that this sequence has the above properties.

"$\Leftarrow$": The proof of the converse statement, namely that the existence of a sequence η_n with the properties stated in the Theorem implies that μ is a limit point of $\sigma(A)$ essentially relies on the fact in a finite dimensional Hilbert space weak convergence implies strong convergence. Hence no subsequence that also weakly converges to 0 of the sequence η_n can lie entirely in a finite-dimensional subspace. However one easily proves that for any Borel set Δ that contains μ there must exist a subsequence of η_n that weakly converges to 0 and that lies entirely in $P(\Delta)\mathcal{H}$. Thus all these spaces must be infinite dimensional. That μ cannot lie in $r(A)$ is easy to see as well.

Remark:
We should note here that besides the purely continuous and pure point part of the spectrum there is a third possibility for operators that are *not normal*: Namely the residual (or singular) part of the spectrum. It is the set of all λ for which the von Neumann series for $(\lambda - a)^{-1}$ converges to a bounded operator, but the range of this operator is not dense in $\mathcal{H}$. In fact, operators that are not normal play an important role in physics. E.g. the Dirac-Operator on Lorentzian Spin-manifolds or Random Schrödinger-Operators.

Examples

1. Consider the selfadjoint operator $p = -i\frac{\partial}{\partial x}$ on $L^(\mathbb{R})$. Then p has purely continuous spectrum, $\sigma(p) = \mathbb{R}$.

2. Similarly the selfadjoint operator $\hat{x}$ on $L^2(\mathbb{R})$, densely defined by $(\hat{x}\psi)(x) = x\psi(x)$, has purely continuous spectrum $\sigma\hat{x} = \mathbb{R}$.

3. On the other hand, as we have seen above, the selfadjoint operators $\dot{x}_\theta$ i.e. the selfadjoint extensions of $i\frac{\partial}{\partial x}$ on $L^2([0,1])$ have completely discrete spectrum.

4. Similarly to the above example the Dirac- and Laplace-Beltrami-operators on *compact* manifolds without boundary have completely discrete spectrum. (The precise definition of these operators shall be given later.)

Note that in the first two examples above, there do also exist **generalized eigenfunctions**. Given a selfadjoint operator P that leaves the space Φ of test function space (e.g. the space $\mathcal{D}(\mathcal{M})$ of compactly supported smooth functions on some smooth manifold $\mathcal{M}$) invariant, such a generalized eigenfunction to the eigenvalue $\lambda \in \mathbb{R}$ is given as a *distribution* T_λ such that :

$$T_\lambda(P\varphi) = \lambda T_\lambda(\varphi) \qquad \forall \varphi \in \Phi.$$

Thus the (general) definition of such generalized eigenfunctions requires a **Gelfand triple** , which is given by a dense subspace Φ of $\mathcal{H}$ (or a space continuously embedded into $\mathcal{H}$ as a dense subspace) so that

$$\Phi \subseteq \mathcal{H} \subseteq \Phi'$$

where Φ' denotes the space of continuous linear functionals on Φ. Particularly useful examples arise if Φ is a **nuclear space**, which however need to be a possible choice. In fact, the difficulty with the use of Gelfand triples relies in the problem to find a suitable space Φ for a given operator P. In many important examples from physics (the existence of) a feasible choice of Φ is not known. On the other, there do exist concrete algorithms to find a nuclear space Φ in case such a choice is possible.
Note that, by density, it is clear that one will always find sequences $\tau_{n,\lambda} \in \Phi$ (and thus in particular in the domain of P) which converge weakly to T_λ, i.e.

$$\langle \tau_{n,\lambda}, \varphi \rangle \to T_\lambda(\varphi) \qquad \forall \varphi \in \Phi \subseteq \mathcal{H}.$$

As a concrete example, one may consider the sequence $\frac{\sin(n(x-x_0))}{\pi(x-x_0)}$ that weakly converges to the delta-distribution $\delta(x - x_0)$ centered at x_0, which is the generalized eigenfunction to the position operator on $L^2(\mathbb{R})$, viz

$$\delta(x - x_0)(\hat{x}\varphi) = x_0\varphi(x_0) = x_0\delta(x - x_0)(\varphi).$$

2.4 Differential Geometry

In this section we shall give a brief introduction to the background in differential geometry used in the main part of the lecture notes. From a differential geometric point of view most properties of manifolds are most conviently described in terms of the pre-C^*-algebra of smooth functions that vanish at infinfinity (together with all their derivatives). Moreover this algebraic framework appears very naturally in quantum theory.
Consequently, the following introduction will use the algebraic and thus global description of geometric data in terms of operations on (modules over)

smooth functions. An important advantage of this approach over the more "geometric", local approach in terms of local coordinates is that one does not need to worry about coordinate independence or global well-definiteness of the construction. Yet, for some local constructions the description in terms of smooth functions is not known, or too complicated to be of much use for practical computations. Thus for later use some constructions are also described in terms of local coordinates.

2.4.1 Manifolds and Vector bundles

Recall that by the Gelfand-Naimark Theorem a locally compact topological Hausdorff space $\mathcal{M}$ is uniquely described by the commutative C^*-algebra $C_0(\mathcal{M})$ of continuous functions that vanish at infinity. Vice versa any commutative C^*-algebra is of this form. For technical simplicity only, we shall in the sequel restrict ourselves to consider only *compact* manifolds when every continuous function belongs to $C_0(\mathcal{M}) = C(\mathcal{M})$ and there exists a finite covering. In particular $C(\mathcal{M})$ is unital.

Definition:
Let $\mathcal{A}$ be a unital, commutative C^*algebra. We shall call the Gelfand-spectrum $\mathcal{M} := Spec(\mathcal{A})$ a compact n**-dimensional manifold** if there exists a finite set I and for each $i \in I$ there exist n selfadjoint elements $a_i^1, a_i^2, \ldots a_i^n$ of $\mathcal{A}$ such that

- Let $U_i := \bigcap_{\alpha=1}^{n} \mathrm{supp}(a_i^\alpha)$. Then it is $\bigcup_{i\in I} U_i = \mathcal{M}$ i.e the U_i form an open covering of $\mathcal{M}$.

- For each $i \in I$ the map

$$\alpha_i = (a_i^1, \ldots a_i^n) : U_i \subset \mathcal{M} \to a_i(U_i) \subset \mathbb{R}^n$$

  is a homeomorphism, i.e. it is invertible, with a continuous inverse a_i^{-1}.

The functions $a_i^\alpha : U_i \to \mathbb{R}$ are called the *coordinates* for the *local chart* U_i. The collection $\{(U_i, a_i)\}_{i\in I}$ is called an *atlas*.

Definition:
Let $\mathcal{M}$ be a (compact) n-dimensional manifold. A *smooth atlas* of $\mathcal{M}$ is an atlas $\{(U_i, a_i)\}_{i\in I}$ such that

$$a_i \circ a_j^{-1} : a_j(U_i \cap U_j) \to a_i(U_i \cap U_j)$$

is a smooth map (between subsets of $\mathbb{R}^n$) whenever $U_i \cap U_j \neq \emptyset$.
A **smooth manifold** is a manifold that admits a maximal smooth atlas.

A given smooth atlas is called maximal if no other smooth atlas contains the same charts, except the given one. Every smooth atlas is contained in a unique maximal smooth atlas. Note that on a given manifold there may exist different smooth structures in the sense that there is no diffeomorphism from the smooth atlas corresponding to the one smooth structure to the other smooth atlas. Hence the smooth atlas is an important part of the data defining a smooth manifold.

Given a compact smooth manifold it is natural and convenient to work with the dense subalgebra of $C(\mathcal{M})$ of smooth functions, denoted by $C^\infty(\mathcal{M})$. From a C^*-algebraic point of view this is a **pre-C^*-algebra**, i.e. a dense subalgebra of a C^*-algebra that is closed under the *holomorphic functional calculus*, i.e. if f is any holomorphic function and a an element of the pre-C^*-algebra then $f(a)$ is also an element of the pre-C^*-algebra. In particular the elements a_i^α of the atlas are then taken from $C^\infty(\mathcal{M})$.
Let $F_i : \mathcal{M} \to \mathbb{C}$ be a smooth function with $\mathrm{supp}(F) \subset U_i$. We may then define the smooth function $f_i : a_i(U_i) \subset \mathbb{R}^n \to \mathbb{C}$ by $f = F \circ a_i^{-1}$. Thus, vice versa any such function F_i can be written as

$$F_i = f_i \circ a_i = f_i(a_i^1, \ldots a_i^n)$$

with a uniquely determined f_i. Note that if e.g. $\mathrm{supp}(F) \subset U_i \cap U_j$, then we may write $F = f_j \circ a_j = f_i \circ a_i$.
Let the family $\{\psi_i\}_{i\in I} \subset C^\infty(\mathcal{M})$ be a **partition of unity subordinate to the atlas** $\{(U_i, a_i)\}_{i\in I}$. That is to say the smooth functions $\psi_i : \mathcal{M} \to \mathbb{R}$ satisfy

$$\mathrm{supp}(\psi_i) \subset U_i \qquad \text{and} \qquad \sum_i \psi_i^2 = 1.$$

Then any smooth function $F \in C^\infty(\mathcal{M})$ can be written as

$$F = \sum_i (\psi_i^2 F) = \sum_i F_i = \sum_i (f_i \circ a_i).$$

Of particular importance is the set of all **first order differential operators on** $\mathcal{M}$, i.e. of $\mathbb{C}$-linear maps $P : C^\infty(\mathcal{M}) \to \mathcal{C}^\infty(\mathcal{M})$ obeying the *Leibniz-rule*:

$$P(ab) = P(a)b + aP(b) \qquad \forall a, b \in C^\infty(\mathcal{M}).$$

Obviously this set is a vector space over $\mathbb{C}$. Moreover it is a **left $C^\infty(\mathcal{M})$-modul** because if P is a first order differential-operator so is aP for all $a \in C^\infty(\mathcal{M})$.
Note that by the inverse function theorem, the invertibility of the map $a_i = (a_i^1, \ldots, a_i^n) : U_i \to a_i(U_i) \subset \mathbb{R}^n$ is equivalent to the existence of n linearly independent (over $C^\infty(\mathcal{M})$) first order differential operators χ_i^α such that the Jacobian

$$J_{\chi,i}^{\alpha\beta} := \chi_i^\alpha(a_i^\beta) : U_i \to M_n(\mathbb{R})$$

is invertible in every $p \in U_i$.

Let P be a first order differential operator. Then the action of P on any function $F_i \in C^\infty(\mathcal{M})$ with $\mathrm{supp}(F_i) \subset U_i$, i.e. $F_i = f_i \circ a_i$, can be written by using the chain rule as

$$P(F_i) = \sum_\alpha P(a_i^\alpha)(f_i^{(\alpha)} \circ a_i)$$

where $f_i^{(\alpha)}(x^1, \ldots, x^n) = \frac{\partial f_i}{\partial x^\alpha}(x^1, \ldots, x^n)$.
Thus, *locally*, in U_i the action of P is completely determined by the n functions $P(a_i^\alpha)$. Alternatively we may describe P in U_i by the n smooth functions ϕ_i^β

$$\phi_i^\beta = \sum_\alpha P(a_i^\alpha) \left(J_{i,\chi}^{-1}\right)^{\alpha\beta} \qquad \text{i.e.} \qquad P(a_i^\alpha) = \sum_\beta \phi_i^\beta \chi_i^\beta(a_i^\alpha).$$

Thus we may locally view P as a linear combination over $C^\infty(\mathcal{M})$ of the operators χ_i^β.
If $U_i \cap U_j \neq \emptyset$ this in particular applies to the χ_j^α when restricted to $U_i \cap U_j$. Hence there do exist real-valued smooth functions $q_{ij}^{\alpha\beta}$ with $\mathrm{supp}(q_{ij}^{\alpha\beta}) \subset U_i \cap U_j$ such that

$$\chi_j^\beta = \sum_\alpha q_{ij}^{\alpha\beta} \chi_i^\alpha.$$

For later convenience we shall also introduce the functions $q_{ii}^{\alpha\beta} = \delta_{\alpha\beta}$ and observe that it follows that for any i, j the $(n \times n)$-matrices q_{ij} with entries $q_{ij}^{\alpha\beta}$ fullfill the consistency relation

$$q_{ij} q_{jk} = q_{ik} \qquad \text{on} \quad U_i \cap U_j \cap U_k.$$

In particular $q_{ij} q_{ji} = q_{ij}^2 \equiv 1$ on $U_i \cap U_j$.
Note that we can write P on $U_i \cap U_j$ either as $P = \sum_\alpha \phi_i^\alpha \chi_i^\alpha$ or as $P =$

$\sum_\alpha \phi_j^\alpha \chi_j^\alpha$. Consistency then requires that for the respective restrictions of ϕ_i^α and ϕ_j^β to $U_i \cap U_j$ we have

$$\phi_i^\alpha|_{U_i\cap U_j} = \sum_\beta q_{ij}^{\alpha\beta}\ \phi_j^\beta\Big|_{U_i\cap U_j}$$

using the consistency relation for the matrices q_{ij}.

Let now ψ_i be a partition of unity subordinate to the atlas $\{(U_i, a_i)\}_{i\in I}$ and let $r = |I|$ be the cardinality of the finite set I.
As we saw above any first order differential operator is fully characterized by a collection of $n \cdot r$ functions ϕ_i^α subject to a consitency relation on the overlaps $U_i \cap U_j$.
Define

$$\varphi_i^\alpha = \psi_i \phi_i^\alpha$$

and the $(n \cdot r) \times (n \cdot r)$-matrix valued function p with entries

$$p_{ij}^{\alpha\beta} = \psi_i \psi_j\, q_{ij}^{\alpha\beta}.$$

From the consistency relation for the matrices q_{ij} it then follows immediately that p is a projector, i.e.

$$p^* = p = p^2.$$

Moreover the consistency relation for the functions ϕ_i^α gives

$$\begin{aligned}\varphi_i^\alpha &= \sum_j \psi_j^2 \psi_i\, \phi_i^\alpha = \sum_{j,\beta} \psi_j^2 \psi_i\, q_{ij}^{\alpha\beta} \phi_j^\beta \\ &= \sum_{j,\beta} \psi_j\, p_{ij}^{\alpha\beta} \phi_j^\beta = \sum_{j,\beta} p_{ij}^{\alpha\beta} \varphi_j^\beta.\end{aligned}$$

In other words the $(n \cdot r)$-tupel $\{\varphi_i^\alpha\}$ can be viewed as an element of the $C^\infty(\mathcal{M})$-module

$$(C^\infty(\mathcal{M}))^{n\cdot r} \cdot p,$$

i.e. the complex vector space of all $(n \cdot r)$-tuples $\mathbf{f} = (f_1, \dots f_{n\cdot r})$ of smooth functions f_k such that

$$\mathbf{f}p = \mathbf{f}.$$

Vice versa any such tupel will define a first order differential operator on $\mathcal{M}$.

Definition:
Let $\mathcal{A}$ be an unital, associative $*$-algebra. In particular $\mathcal{A}$ may be a unital

(pre)-C^*-algebra. A left $\mathcal{A}$-modul $\mathcal{E}$ will be called **finitely generated projective** if there exists an $N \in \mathbb{N}$ and an projector $p = p^2 = p^*$ in $M_N(\mathcal{A})$ such that $\mathcal{E} = \mathcal{A}^N p$.
If $\mathcal{A} = C^\infty(\mathcal{M})$ is the algebra of smooth functions on a (compact) smooth manifold $\mathcal{M}$ we shall call any finitely generated projective module $\mathcal{E}$ over $\mathcal{A}$ the space of **smooth sections of some vector bundle over $\mathcal{M}$**.

Thus the space of first order differential operators is a finitely generated projective module over $C^\infty(\mathcal{M})$. In the sequel we shall only consider finiteley generated projected modules over this pre-C^*-algebra.
Not that then, by smoothness of its entries, the *rank* of the projector p is constant on all of $\mathcal{M}$. For instance, in the example of first order differential operators the rank of p is $n =$dim$(\mathcal{M})$. Moreover in that example one can describe the elements of the module, i.e. the differential operators, *locally*, on each chart U_i by n functions ϕ_i^α. However, due to the consistency relations and the possibly nontrivial topology of $\mathcal{M}$ it may not be possible to characterize any differential operator by only n functions also globally.

Let E and $\mathcal{M}$ be smooth manifolds, $\mathcal{M}$ being compact, and $\pi : E \to \mathcal{M}$ be a surjective map. Suppose that for each $p \in \mathcal{M}$ the fibre $E_p = \pi^{-1}(p)$ carries the structure of an d-dimensional complex vector space. Moreover assume that there exists an open covering $\{U_i\}_{i\in I}$ of $\mathcal{M}$ such that for each $i \in I$ there exists a smooth map $\Phi_i : \pi^{-1}(U_i) \to U_i \times \mathbb{C}^d$ that is invertible and has a smooth inverse, i.e. it is a diffeomorphism, and obeys the equation

$$\pi = \pi_t \circ \Phi$$

where $\pi_t : U_i \times \mathbb{C}^d$ is the natural projection on the first factor. Thus Φ_i preserves the **fibers** $E_p = \pi^{-1}(p)$ and we shall require that this map $\pi_t \circ \Phi_i|_{E_p}$ is a vector space isomorphism.
We may then view E as a collection of d-dimensional complex vector spaces indexed by the points of M and glued together smoothly over each U_i. In fact due to requirement that the corresponding maps $\pi_t \circ \Phi_i$ are vector-space isomorphisms, this gluing is smooth over all of $\mathcal{M}$. Indeed, if $U_i \cap U_j \neq \emptyset$ the map $\Phi_i \circ \Phi_j^{-1}$ is a smooth map from $U_i \cap U_j$ to isomorphisms of d-dimensional complex vectorspaces.

The triple $(E, \mathcal{M}, \pi)$ is then called a smooth **vector bundle** of rank d. E is called the **total space**, $\mathcal{M}$ the **base** of the vector bundle. The collection of maps Φ_i is called a **local trivialization** of the bundle.

A **smooth section** in the vector bundle $(E, \mathcal{M}, \pi)$ is a smooth map $s : \mathcal{M} \to E$ such that

$$\pi \circ s = \mathrm{id}_{\mathcal{M}}.$$

Thus a section s assigns to each $p \in \mathcal{M}$ an element of the vector space E_p and this assignment depends smoothly on p. Thus the space of all smooth sections is obviously a (left) module over $C^\infty(\mathcal{M})$. From the above discussion of the example of first order differental operators it should also be clear (how to prove) that this module is finitely generated projective. In fact the converse statement, that each finitely generated projective module over $C^\infty(\mathcal{M})$ defines a smooth vector bundle, is also true:

Theorem (Serre-Swan):
There is a one-to-one correspondence of vector bundles over $\mathcal{M}$ and finitely generated projective modules over $C^\infty(\mathcal{M})$.

Examples:

1. A vector bundle that is globally of the form $\mathcal{M} \times \mathbb{C}^d$ is called a **trivial bundle**.

2. The bundle corresponding to the finitely generated projective module of first order differential operators is called the **tangent bundle of** $\mathcal{M}$. Its total space is usually denoted by $T\mathcal{M}$, the fibers by $T_p\mathcal{M}$ and the space of smooth sections by $\Gamma^\infty(T\mathcal{M})$. Sections $X \in \Gamma^\infty(T\mathcal{M})$ are called **vector fields**. The name tangent bundle (respectively tangent vectors) is motivated by the fact that elements of $T_p^*\mathcal{M}$ can be viewed as tangent vectors to curves $\mathbb{R} \to \mathcal{M}$ in the point p.
   Note that, as they are first order differential operators one may define the **commutator** $[\cdot, \cdot]$ **of vector fields** X, Y as

   $$[X, Y](a) = X(Y(a)) - Y(X(a))$$

   for all smooth functions a. One easily checks then that $[X, Y]$ obeys the Leibniz-rule, i.e. it is well-defined as a vector-field.

3. Most operations from linear algebra can be carried out on vector bundles by defining them *fiberwise*. Most importantly, if $(E, \mathcal{M}, \pi)$ is a vector bundle, then one may define its **dual vector bundle** $(E^*, \mathcal{M}, \pi)$, which is defined by setting for all $p \in \mathcal{M}$

   $$(E^*)_p := (E_p)^*.$$

Similarly one may define the direct sum and tensor product of two bundles $(E, \mathcal{M}, \pi)$ and $(E_2, \mathcal{M}, \pi_2)$ over $\mathcal{M}$ and, accordingly also the **exterior** (i.e. antisymmetric) and **symmetric** powers of $(E, \mathcal{M}, \pi)$ are defined.

4. According to the last example, there is a dual vector bundle to $T\mathcal{M}$, called the **cotangent bundle** $T^*\mathcal{M}$ of $\mathcal{M}$. Sections $\omega \in \Gamma^\infty(T^*M)$ are called **one-forms**. We shall also often write $\Omega^1(C^\infty(\mathcal{M}))$ for the space of one-forms. Note that there is a canonical map d: $C^\infty(\mathcal{M}) \to \Gamma^\infty(T^*\mathcal{M})$ defined via

$$\mathrm{d}f(X) = X(f) \qquad \forall f \in C^*(\mathcal{M}) \quad \text{and} \quad \forall X \in \Gamma^\infty(T\mathcal{M})$$

and called the **differential d**. With respect to the local coordinates a_i^α in some chart U_i we may then define a local basis of $\Gamma^\infty(T^*\mathcal{M})$ as

$$\mathrm{d}a_i^\alpha \qquad \alpha = 1, \ldots, \dim(\mathcal{M}).$$

The global picture of this module is then described similarly as for the tangent bundle. Before we proceed it is important to observe that $\mathrm{d}(fg) = (\mathrm{d}f)g + f\mathrm{d}g$ as follows immediately from the definition.

5. Next, one may define for $k \leq n$ the k-th exterior power of $T^*\mathcal{M}$ i.e the k-fold antisymmetric tensor product

$$\Lambda^k T^*\mathcal{M} := \underbrace{T^*\mathcal{M}\hat{\otimes}T^*\mathcal{M}\hat{\otimes}\cdots\hat{\otimes}T^*\mathcal{M}}_{k-\text{times}}.$$

The space of sections of this bundle shall be denoted by $\Omega^k(C^\infty(\mathcal{M}))$. Its elements are called **k-forms**. An explicit local basis is provided by

$$\mathrm{d}a_i^{\alpha_1} \wedge \mathrm{d}a_i^{\alpha_2} \wedge \cdots \wedge \mathrm{d}a_i^{\alpha_k}.$$

Note that for $k = n$ this bundle $\Lambda^n T^*\mathcal{M}$ has rank 1 with the single local basis $\mathrm{d}a_i^1 \wedge \cdots \wedge \mathrm{d}a_1^n$.

6. Let $p \in \mathcal{M}$. We denote by $|\Lambda\mathcal{M}|_p$ the set of all functions $v : \Lambda^n T_p^*\mathcal{M} \to \mathbb{R}$ such that for all $\lambda \in \mathbb{R}$ and all $\omega_p^n \in \Lambda^n T_p^*\mathcal{M}$ one has

$$v(\lambda\omega_p^n) = |\lambda| v(\omega_p^n).$$

Thus each $|\Lambda\mathcal{M}|_p$ is a one dimensional real vector space and together these vector spaces define a *real* vector bundle $|\Lambda\mathcal{M}|$, which turns out

to be trivial. Sections of this bundle are called **densities**. Their importance lies in the fact that they can be integrated. (Recall that we assume throughout that $\mathcal{M}$ is compact for simplicity). In a chart (U_i, a_i) we can define a local basis for densities as the unique density $|\mathrm{d}a_i|$ such that

$$|\mathrm{d}a_i|(\mathrm{d}a_i^1 \wedge \cdots \wedge \mathrm{d}a_i^n) = 1$$

Any section $s \in \Gamma^\infty(|\Lambda\mathcal{M}|)$ with support in U_i is then of the form $F_i|\mathrm{d}a_i|$ with F_i being a smooth function with support conatined in U_i. There is then a unique linear map $\int_M$ such that for all $i \in I$ and all such $F_i = f_i \circ a_i$ with $\mathrm{supp}(F_i) \subset U_i$ we have

$$\int_M (F_i|\mathrm{d}a_i|) = \int\limits_{a_i(U_i)} f_i \, \mathrm{d}^n x.$$

2.4.2 Metrics, connections and curvature

Since the algebra $C^\infty(\mathcal{M})$ is commutative all finitely generated modules $\mathcal{E}$ over this algebra are actually bimodules, i.e. elements of $\mathcal{E}$ (sections of some vector bundle) can be multiplied from the right and from the left with smooth functions.
Given two bimodules $\mathcal{E}, \mathcal{F}$ over an assiciative unital algebra $\mathcal{A}$ one may build their tensor product $\mathcal{E} \otimes_\mathcal{A} \mathcal{F}$ over $\mathcal{A}$. As a vector space it is defined as the quotient $(\mathcal{E} \otimes \mathcal{F})/\sim$ where the equivalence relation $\sim$ is given by

$$(ea) \otimes f \sim e \otimes (af) \qquad\qquad \forall a \in \mathcal{A}.$$

Obviously $\mathcal{E} \otimes_\mathcal{A} \mathcal{F}$ is again a bimodule over $\mathcal{A}$. As a warning, note that in general (for noncommutative algebras $\mathcal{A}$) the bimodules $\mathcal{E} \otimes_\mathcal{A} \mathcal{F}$ and $\mathcal{F} \otimes_\mathcal{A} \mathcal{E}$ need not be isomorphic.
We may thus consider $C^\infty(\mathcal{M})$-bilinear maps

$$g : \Gamma^\infty(T^*\mathcal{M}) \otimes_{C^\infty(\mathcal{M})} \Gamma^\infty(T^*\mathcal{M}) \to C^\infty(\mathcal{M})$$

i.e. we have for all $a, b \in C^\infty(\mathcal{M})$ and all one-forms ω_1, ω_2 on $\mathcal{M}$:

$$g(a\omega_1, b\omega_2) = ab\, g(\omega_1, \omega_2).$$

Moreover we shall assume that the map g is symmetric, i.e. $g(\omega_1, \omega_2) = g(\omega_2, \omega_1)$.
Thus, locally, in every chart (U_i, a_i) the map g is uniquely determined by the matrix valued function g with entries

$$g^{\alpha\beta} = g(\mathrm{d}a_i^\alpha, \mathrm{d}a_i^\beta).$$

For every point $p \in U_i$ the matrix $g^{\alpha\beta}(p)$ is then obviously syymetric. We shall say that the map g is nondegenerate if (in all charts) the matrix $g^{\alpha\beta}(p)$ is nondegenerate (i.e. has nonvanishing determinant) for all $p \in \mathcal{M}$.

Definition:
A symmetric, nondegenerate $C^\infty(\mathcal{M})$-bilinear map

$$g : \Gamma^\infty(T^*\mathcal{M}) \otimes_{C^\infty(\mathcal{M})} \Gamma^\infty(T^*\mathcal{M}) \to C^\infty(\mathcal{M})$$

is called a **metric** on $\mathcal{M}$. If the above defined matrices with components $g^{\alpha\beta}(p)$ are positive definite in all $p \in \mathcal{M}$ the metric g is called a **Riemannian metric**. It is called **Lorentzian metric** if the $g^{\alpha\beta}$ have signature

$$(-1, \underbrace{+1, \ldots, +1}_{\dim(\mathcal{M})-1}).$$

A metric g on $\mathcal{M}$ yields a nowhere vanishing smooth density, locally given as $\sqrt{|g|}|\mathrm{d}a_i|$ where g denotes the determinant of the matrix $g^{\alpha\beta}$. Thus there is a canonical way to integrate a function on a Riemannian or a Lorentzian manifold.
As in linear algebra we may then also define the metric on the sections of the dual bundle, i.e. the tangent bundle $T^{\mathcal{M}}$. The local corresponding matrices $g_{\alpha\beta}$ are then the inverse matrices to the $g^{\alpha\beta}$.

Definition:
Let $\mathcal{E}$ be a finitely generated projective left module over $C^\infty(\mathcal{M})$. A smooth **connection** on $\mathcal{E}$ is a $\mathbb{C}$-linear map

$$\nabla : \mathcal{E} \to \mathcal{E} \otimes_{C^\infty(\mathcal{M})} \Gamma^\infty(T^*M)$$

such that for all $\xi \in \mathcal{E}$ and all $a \in C^\infty(\mathcal{M})$ it is:

$$\nabla(a\xi) = \mathrm{d}a \otimes \xi + a\nabla(\xi).$$

Example:
Note that the differential d: $C^\infty(\mathcal{M}) \to \Gamma^\infty(T^*\mathcal{M})$ is a connection, and this connection obviuosly extends to trivial modules $(C^\infty(\mathcal{M}))^r$ of higher rank by componentwise application. Then, however, one easily checks that $p\mathrm{d}p$ also defines a connection on an arbitrary finitely generated projective module $(C^\infty(\mathcal{M}))^r p$. Thus every finitely generated projective module admits a connection.

Definition:
A **covariant derivative** on a finitely generated projective module $\mathcal{E}$ over $C^\infty(\mathcal{M})$ is a $\mathbb{R}$-linear map

$$\begin{aligned} \nabla : \Gamma^\infty(T\mathcal{M}) \times \mathcal{E} &\rightarrow \mathcal{E} \\ (X, s) &\mapsto \nabla_X(s) \end{aligned}$$

with the following properties:

1. The map ∇ is $C^\infty(\mathcal{M})$-linear in the first argument,i.e.

   $$\nabla_{aX}(s) = a\,\nabla_X(s)$$

   for all smooth functions a, all vector fields X and all $s \in \mathcal{E}$

2. ∇ is a derivation with respect to the second argument, i.e. it is $\mathbb{C}$-linear and

   $$\nabla_X(as) = X(a)e + a\,\nabla_X(s)$$

   holds for all smooth functions a, vector fields X and all $s \in \mathcal{E}$.

It follows from these properties that, given two covariant derivatives ∇, ∇', their difference defines a $C^(\mathcal{M})$-linear endomomorphism of $\mathcal{E}$ for any vector field X, viz

$$(\nabla_X - \nabla'_X) : \mathcal{E} \rightarrow \mathcal{E} \qquad\qquad (\nabla_X - \nabla'_X)(as) = a(\nabla_X - \nabla'_X)(s).$$

Hence it may be viewed as a matrix valued function. Moreover the assignement $(\nabla - \nabla') \rightarrow (\nabla_X - \nabla'_X)$ is $C^\infty(\mathcal{M})$-linear. Therefor there must exist a matrix valued one-form (namely a matrix valued element A of the dual space of $\Gamma^\infty(T^*\mathcal{M})$) such that

$$\nabla_X - \nabla'_X = A(X).$$

Deliberately fixing a "background" covariant derivative ∇^0 then any covariant derivative is given as $\nabla = \nabla^0 + A(\cdot)$. The one-form A with values in the $C^\infty(\mathcal{M})$-linear endomorphisms of $\mathcal{E}$ is called the **connection one-form** on $\mathcal{E}$.
The above nomenclature is well motivated by the fact that there is a one-to-one correspondence of covariant derivatives $\nabla.(\cdot)$ and connections ∇, given via

$$\nabla_X(s) = \nabla(X)(s).$$

Given a covariant derivative ∇ on $\mathcal{E}$ its **curvature** is the map

$$R : \Gamma^\infty(\mathcal{M}) \times \Gamma^\infty(\mathcal{M}) \times \mathcal{E} \to \mathcal{E}$$

defined as

$$R(X,Y)(s) = \nabla_X(\nabla_Y(s)) - \nabla_Y(\nabla_X(s)) - \nabla_{[X,Y]}(s)$$

It is straightforward to check that R can be viewed as an element of

$$\Omega^2(C^\infty(\mathcal{M})) \otimes End_{\mathbb{C}}(\mathcal{E})$$

i.e. for all smooth functions a, b, c and all vector fields X, Y and all $s \in \mathcal{E}$ it is

$$R(aX, bY)(cs) = abcR(X,Y)(s)$$

and R is antisymmetric in the first two arguments.

Definition:
A covariant derivative on $\Gamma^\infty(T^*\mathcal{M})$ is called **metric compatible** if for all one forms ω_1, ω_2 and all vector fields X it holds that

$$X(g(\omega_1, \omega_2)) = g(\nabla_X(\omega_1), \omega_2) + g(\omega_1, \nabla_X(\omega_2)).$$

Likewise we may define metric compatible connections on $\Gamma^\infty(T\mathcal{M})$ by using the inverse metric.

It turns out that there exists a unique metric compatible connection ∇^{LC} on $\Gamma^\infty(T\mathcal{M})$ that is also *torsion free*, i.e.:

$$\nabla^{LC}_X(Y) - \nabla^{LC}_Y(X) = [X,Y]$$

for all vector fields X, Y. This covariant derivative is called the **Levi-Civita**-connection. Its connection-one form defines in any local chart (U_i, a_i) and with respect to any local trivialization of the tangent bundle a one-form $\sum_\alpha \Gamma^\alpha \mathrm{d}a_i^\alpha$ whose components are $\dim(\mathcal{M}) \times \dim(\mathcal{M})$-matrices. Their matrix entries $\Gamma^\alpha_{\beta\gamma}$ are called the *Christoffel-symbols*. The curvature of ∇^{LC} is called the *Riemann-Tensor Riem*. It is an element of

$$\Omega^2(C^\infty(\mathcal{M})) \otimes End(\Gamma^\infty(T\mathcal{M})) \equiv \Omega^2(C^\infty(\mathcal{M})) \otimes \Gamma^\infty(T\mathcal{M}) \otimes \Gamma^\infty(T^*\mathcal{M}).$$

Let $e_1, \ldots, e_n$ (where n =dim($\mathcal{M}$) as before) be n vector fields that are orthogonormal with respect to the inverse metric g^{-1}, i.e.

$$g^{-1}(e_\alpha, e_\beta) = \epsilon_\alpha \delta_{\alpha\beta}$$

where δ_{ij} is the Kronecker-delta and the numbers $\epsilon_\alpha = \pm 1$ are to be chosen according to the signatre of the metric. One may then define the *Ricci-curvature* $Ric \in \Gamma^\infty(T^*\mathcal{M} \otimes T^*\mathcal{M})$ by

$$Ric(X,Y) = \sum_\alpha \epsilon_\alpha g^{-1}(Riem(X, e_\alpha)(e_\alpha), Y)$$

and the *scalar curvature* $R \in C^\infty(\mathcal{M})$ by

$$R = \sum_\beta \epsilon_\beta Ric(e_\beta, e_\beta).$$

To conclude this subsection we should also mention that the metric g allows to define to further operations:

1. On any Lorentzian or Riemannian manifold there is a canonical map

   $$\mathrm{grad} : C^\infty(\mathcal{M}) \to \Gamma^\infty(T\mathcal{M})$$

   called the **gradient** defined via

   $$\mathrm{grad}(a)(b) = g(\mathrm{d}a, \mathrm{d}b) \qquad \forall a, b \in C^\infty(\mathcal{M}).$$

2. Another natural operation is the **divergence** i.e. the map

   $$\mathrm{div} : \Gamma^\infty(T\mathcal{M}) \to C^\infty(\mathcal{M})$$

   evaluated on a vector field X as

   $$\mathrm{div}(X) = \sum_\alpha \epsilon_\alpha g^{-1}(e_\alpha, \nabla^{LC}_{e_\alpha}(X))$$

   where we have used the notation introduced above.

2.4.3 Laplacians and d'Alembertians

A generalized Laplace Operator P on a vector bundle $E \to M$ over a Lorentzian or Riemannian manifold (M, g) is a symmetric second order differential operator whose principal symbol is given by the metric g.
Thus P maps smooth sections of the bundle to smooth sections, $P : \Gamma^\infty(E) \to \Gamma^\infty(E)$, and with respect to a local trivialization of the bundle P is locally, in a point $x \in M$, of the form

$$P = \sum_{ij} g^{ij}(x) \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} + \sum_i A^i(x) \frac{\partial}{\partial x^i} + B(x).$$

Let ∇ be a connection on E, i.e. $\nabla : \Gamma^\infty(E) \to \Gamma^\infty(T * M) \otimes \Gamma^\infty(E)$. Using the Levi-Civita connection ∇^{LC} on $\Gamma^\infty(T^*M)$ we can thus build the operator

$$\Delta^\nabla = -tr \circ \left(\mathrm{id}_{\Gamma^\infty(T^*M)} \otimes \nabla + \nabla^{LC} \otimes \mathrm{id}_{\Gamma^\infty(E)}\right) \circ \nabla$$

where $tr(\omega_1 \otimes \omega_2 \otimes s) = g(\omega_1, \omega_2)s$ for one-forms $\omega_1, \omega_2 \in \Gamma^\infty(T^*M)$ and $s \in \Gamma^\infty(E)$.
This operator Δ^∇ is called the covariant Laplacian for the connection ∇. A lengthy calculation shows that it can be written locally as

$$\Delta^\nabla = -\sum_{ij} \frac{1}{g} \left(\frac{\partial}{\partial x^i} - A_i\right) g g^{ij} \left(\frac{\partial}{\partial x^j} - A_j\right)$$

where $g = \det(g^{ij})$ and A_i denote the local components of the connection one-form $A = \sum_i A_i \mathrm{d}x^i$ corresponding to ∇. Thus Δ^∇ is a generalized Laplacian. In fact, every generalized Laplacian is of this form:

Theorem:
Let P be a generalized Laplace Operator on a vector bundle $E \to M$. Then there exists a *unique* connection ∇ on E and a *unique* endomorphism B of E such that:

$$P = \Delta^\nabla + B.$$

As for the easy proof we refer to the book by Bär, Ginoux and Pfäffle or the one by Gracia Bondia, Figuero and Varilly, respectively the references given therein.

More explicit examples of Laplace-Operators are given by

- The Laplace-Beltrami-Operator on functions on M:

$$\begin{aligned} \Delta(f) &= -\mathrm{div}\ \mathrm{grad}(f) = -\sum_{ij} \frac{1}{g} \frac{\partial}{\partial x^i} \left(g\, g^{ij} \frac{\partial f}{\partial x^j}\right) \\ &= -\sum_{ij} g^{ij} \left(\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j} - \sum_k \Gamma^k_{ij} \frac{\partial}{\partial x^k}\right)(f) \end{aligned}$$

  where Γ^k_{ij} denote the Christoffel symbols. If the metric g has Lorentzian signature this operator is also called the d'Alembert Operator.

- The Hodge-de Rham Laplacian on differential forms, given by $-(*d*d + d*d*)$.

- The Dirac-Laplacian, i.e. the square D^2 of the Dirac-Operator (introduced later). By Lichnerowicz' formula it is related to Δ by

$$D^2 = \Delta + \frac{R}{4}$$

where, as usual, R denotes the scalar curvature.

Chapter 3

An operationalists road to quantum physics

In this chapter we shall derive the mathematical structure of physical systems from rather general considerations on the operational structure of observables. We shall also get a first, yet slightly superficial glimpse of its physical interpretation. As a result of this section we shall find that the failure of classical mechanics in comparison with experiment leads essentially to two alternatives: Quantum theory or theories with hidden variables.

3.1 General physical systems and C^*-algebras

Before we shall come to our actual definition of the systems under consideration, let us first motivate this definition by some intuitive ideas:

In the following we shall consider a rather general class of closed physical systems with finitely many degrees of freedom, even though most of our considerations can, of course, as well be applied to most other examples of dynamical systems. What we have in mind by a physical system is rather general: This can mean the molecules of a gas, the Santa Maria on its way to India, the solar system, an electric circuit and so on. The meaning of the terms "closed" and "finitely many degrees of freedom" will be explained later.

To such a system we can now associate a set **Obs** of observables. Each such observable should correspond to a concrete experimental setup by which we investigate the system. Thus each time we measure an observable O by way of performing the corresponding experiment we shall obtain a real num-

ber $o^i(\omega)$. Of course the system can possibly be in different states, which will be denoted by ω. Moreover we shall not assume that the same value is obtained each time we measure the same observable in the same state. First of all a realistic experimental apparatus has only a finite accuracy. Thus due to the statistical error caused by the inaccuracy of our experiment we might obtain different values each time we measure O in the same state. Then however it would not make sense to assume that we could assign a definite value $o(\omega)$ to the state ω. Our experimental data are simply not good enough to safely assume this to be possible. In fact, quantum theory teaches us that it is't.
Moreover, a realistic measurement necessarily involves an interaction of the apparatus with the physical system, and this will, in general, cause a change of the system during the process of measurement. As we can not assume that we know exactly all the physical laws governing these interaction (or that we can completely compute the change of the state of the system caused by them), we cannot assume at this point of the analyses that we can completely determine the state of the system after the measurement from the values we found in the measurement.

Note that by averaging over infinitely many measurements of O we can, however, define an "expectation value"

$$\omega(O) := \sum_i o^i(\omega)$$

of O in ω. In case we find the same value $o(\omega)$ each time the experiment O is performed we have $\omega(O) = o(\omega)$.
In view of the example of the molecules of a gas it is clear, however, that it might sometimes be wise only to consider observables which are already defined as mean values. It would obviously not make sense to try to mark one of the 10^{-26} molecules (i.e. by a red dot) and consider its velocity as an observable.

Certainly it will do no harm if we assume that with O also any scalar multiple λO for $\lambda \in \mathbb{R}$ is an observable , i.e. an element of **Obs**. (Since we can always rescale our instrument). Moreover, also any real polynomial $P(O)$ of O should be in **Obs**: If we measure the values $o^i(\omega)$ for O then we may e.g. simply declare that for $P(O)$ we have found the values $P(o^i(\omega))$. No change of our experimental setup is needed to do so. It follows immediately that for any state ω

$$\omega(\lambda O) = \lambda\omega(O).$$

Since, by our definition, the observable O^0 will only take the value 1 in any experiment, we shall define $O^0 = 1$ for all O. We then also have that

$$\omega(1) = 1$$

for all states ω.

We shall not assume at this point that the sum or the product of two different observables is an observable, nor that states are linear if the sum happens to exist. First of all, given two observables O_1, O_2, for which there thus does exist an experimental setup to measure them, it is not at all trivial that there is an experimental setup to measure $O_1 + O_2$, say. (For instance the sum of the velocity and the position of a particle a priori makes no sense as an observable.) If there does exist such an apparatus that measures the sum, it will in general be different from the experiments corresponding to the individual observables. It is therefore not clear at this point that the expectation value of the sum $O_1 + O_2$ equals the sum of the expectation values. As we shall see, this is actually the case for classical and quantum physics. However, this fact is experimentally confirmed. It should thus be viewed as an experimental finding, valid at the presently accessible energy scales, but not as a necessary presupposition.

A realistic experiment has generically only a finite range of values that it can measure. For instance the measurement of energy by some apparatus is typically bounded because the apparatus would be destroyed if it would come in contact with a higher energy. Likewise no known experiment could measure the distance to an object that is arbitrarily far away. Of course, the upper bounds for distances that can in principle be measured with astronomical methods are pushed forward continuously. However, this usually requires to find new experimental methods, i.e. to define a new observable. Still, the expectation value for this observable in any state does not exceed a certain bound. Thus we shall assume that

$$\|O\| := \sup_{\omega} \{|\omega(O)|\} < \infty.$$

Now, by definition of O^{2k}, the expectation value $\omega(P(O))$ will be positive for any positive polynomial P, i.e. a polynomial with the property that $P(x) \geq 0, \quad \forall x \in \mathbb{R}$. As is explained in the mathematical preliminaries, it can then be shown that $\|\cdot\|$ does actually define a C^*-norm on the algebra generated by a single observable O, i.e the complex valued polynomials of O: As the expectation values of real polynomials of O are real, it is possible to extend them to complex linear combinations of such polynomials.

Of, course, we shall only consider real (i.e. self-adjoint elements) of the so obtained complex algebra as observables, as it might, for instance, be hard to find an apparatus that measures an imaginary distance. However, provided we only assign an operational meaning to the self-adjoint elements the passage to a complex algebra is completely harmless, and, as we shall see, only a matter of convenience. As we have a C^*-norm on this algebra it also appears natural to complete it with respect to this norm. Note that any state then defines (via its expectation values) a normalized, positive and bounded linear functional on this algebra. Thus:

Any observable O generates a commutative C^*-algebra $C(O)$, of continuous functions of O. All self-adjoint elements of $C(O)$ are operationally well-defined observables of the system. Moreover any state of the system induces a C^*-state on $C(O)$ for all $O \in \mathbf{O}bs$. We can therefore view $\mathbf{O}bs$ as a collection of commutative C^*-algebras.

There would be no practical or theoretical gain to distinguish two observables (i.e. experiments) which would give the same expectation value for all states. Thus we can identify such observables. In other words we shall assume that the states of the system separate the observables, i.e $\omega(O_1) = \omega(O_2)$ for all states ω will imply that $O_1 = O_2$.

On the other hand it might be too restrictive to assume that the observables separate the states. It is conceivable – though anything but desirable – that some degrees of freedom of the system are not accessible to any experiment, as they do not interact with any (known) device. Such degrees of freedom are called *hidden variables*.
Let us now suppose that the observables do separate the states. We may then formally define the sum of any two observables $O_1 + O_2$ by its expectation values, i.e as the unique "observable" such that

$$\omega(O_1 + O_2) = \omega(O_1) + \omega(O_2).$$

(This will define a *unique* observable because the states do separate the observables by assumption.) Thus, in case the observables do separate the states, we may safely assume that the sum of any two observables is an observable, even it is not clear how to ascribe a concrete measurement device to this sum.
The measurement of (the expectation values of) O_1 and O_2 will in general require very different experimental setups. Thus the above sum $\omega(O_1) + \omega(O_2)$

of the expectation values is only operationally meaningful if we can be sure that the corresponding experiments are carried out in the same state. For this reason the assumption that the observables do separate the states, i.e. that we can uniquely determine the state of the system by experiments is crucial here. In other words:

For a system with hidden variables the sum of any two variables may not be well-defined. For a system without hidden variables it always is.

It is then tempting to try to define a product of observables along the same lines as well. It can be shown that under certain, rather mild, additional assumptions this can indeed be done. However, the physical interpretation of such additional assumptions is not known. Moreover we may gain some flexibility if we assume as less as possible about the system. Nevertheless we shall now restrict the class of physical systems that we shall consider severely: As we want to discuss the role of the principle of locality we shall assume that there exists a manifold X such that all real-valued continuous functions $f \in C(X)$, $f^* = f$ are observables. Note that, given some local coordinates x^k on a local chart U for X, the algebras of continuous functions of each x^i would automatically define observables. Thus we assume that we can build the tensor-product of the resulting algebras which would then (after an appropriate completion) would lead to the algebra $C(X)$. Of course it is not clear at all that this can always be done, in particular not as we assume that X can be given the structure of a manifold. In fact we shall later see that for realistic physical systems this assumption appears more and more problematic.
For the moment we shall not specify the physical interpretation of the manifold X. It may for instance be the phase space of a classical system, or its configuration space.

Definition: We say that the system has (at most) N degrees of freedom if there exists an N-dimensional manifold X such that

1. $C(X) \subset \mathbf{Obs}$

2. $C(X)$ separates the physical states

3. No proper C^*-subalgebra of $C(X)$ separates the physical states

Hence for a system with N degrees of freedom we have that

$$\omega_1(f) = \omega_2(f) \qquad\qquad \forall f \in C(X)$$

implies that $\omega_1 = \omega_2$. The assumptions that there are no proper C^*-subalgebras which do also separate the states means that X is chosen minimal in this sense. In fact, any proper C^*-subalgebra of $C(X)$ will be isomorphic to some algebra $C(Y)$ with $Y \subset X$ but $Y \neq X$. However, given X there might of course be a completely different manifold $\tilde{X}$, of lower dimension such that $C(Y)$ belongs to **Obs** and separates all the physical states. Therefore the above definition only defines an upper bound for the number of degrees of freedom. For simplicity we shall in the following assume that the systems under consideration do have a finite number of degrees of freedom and that the dimension of the manifold X is chosen minimal.

Now, as explained in the appendix, by the Gelfand-Naimark-Theorem and the Gelfand-Naimark-Segal-Theorem for any commutative C^*-algebra – and thus for $C(X) \subset$ **Obs** defined above – and every state ω we know that

1. There exists a representation π_ω of $C(X)$ on a Hilbert space $\mathcal{H}_\omega$ and a cyclic vector $\psi_\omega \in \mathcal{H}_\omega$ of norm 1 such that

   $$\omega(f) = \langle \psi_\omega, \pi_\omega(f)\psi_\omega \rangle$$

   for all $f \in C(X)$. This representation will, however, not be faithful in general.

2. There will also exist a faithful representation $(\pi, \mathcal{H}, \psi)$ of $C(X)$, admitting a cyclic vector $\psi \in \mathcal{H}$. Any element $\phi \in \mathcal{H}$ of norm one, $\langle \phi, \phi \rangle = 1$ does then define a state on $C(X)$.
   Moreover any state on ω on $C(X)$ can be approximated arbitrarily well by elements of $\mathcal{H}$. Thus, given ω there exists a sequence $\psi_n \in \mathcal{H}$, $\|\psi_n\| = 1$ such that for any $\epsilon > 0$ we have

   $$|\langle \psi_n, \pi(f)\psi_n \rangle - \omega(f)| \leq \epsilon$$

   for all $f \in C(X)$ and sufficiently large n.

To be more concrete, the above representation is, for a compact[1] manifold X, given by $\mathcal{H} = L^2(X)$ with functions $f \in C(X)$ acting as multiplication operators, i.e. $\pi(f)\psi(x) = f(x)\psi(x)$. The pure states on $C(X)$ are in one to one

[1]For technical simplicity we shall only consider this case in the sequel. However, all that follows also holds for non compact manifolds.

correspondence with the states of X. Thus if $y \in X$ the corresponding state ω_y acts as $\omega_y(f) = f(y)$. Any sequence $\psi_n^y(x)$ of functions, such that $|\psi_n^y(x)|^2$ converges to the Dirac-distribution $\delta(x-y)$ and $\int\limits_X |\psi_n^y(x)|^2 \, \mathrm{d}x = 1$, will then approximate the state ω_y in the above sense, i.e. $\lim\limits_{n\to\infty} \langle \psi_n^y, \pi(f)\psi_n^y \rangle = f(y)$. Any other state is a linear combination of pure states and can thus be approximated similarly (as any distribution can be approximated by a sequence of integrable functions).

Note that it will generally not be true that any state of the above representation is a physical state. Thus, in practice we may need to add some subsidiary conditions to specify which states are physical. In the examples we shall consider below, this will, however, not be necessary. Moreover,one may imagine at this point that it may happen that no element of $\mathcal{H}$ represents a physical state, but only limits of such vectors. In the next section, when we shall discuss the physical interpretation of the mathematical formalism developed here, we shall however see that such a situation cannot occur for realistic physical systems.
Of course, in general there will be more observables than the elements of $C(X)$ and this raises the question whether these observables can also be represented on the Hilbert space. In fact, the usual C^*-theoretic description assumes that **Obs** is actually given as the set of selfadjoint elements of a C^*-algebra $\mathcal{A}$. One then works with a representation of $\mathcal{A}$, i.e. of all observables. Here shall not assume that all observable are represented on $\mathcal{H}$ nor that any selfadjoint operator on $\mathcal{H}$ defines an observable, as the operational meaning would not be clear. However we have not yet discussed the dynamics, i.e. the time evolution of the system. As we shall see now, this will not only define a large class of additional observables that *are* represented on $\mathcal{H}$, but also the operational meaning of these observables will be clear.

The dynamical equations of a physical system will in general induce a time evolution on the set of physical states. We shall denote the latter by $\mathcal{S}$ from now on. We shall call the system *closed* if every state ω in $\mathcal{S}$ lies on a unique trajectory under this time evolution. In other words every state $\omega \in \mathcal{S}$ defines a unique curve $\{\omega_t\}_{t\in\mathbb{R}}$ such that $\omega = \omega_{t=0}$. This time evolution will automatically be translation invariant, i.e. if $t_1 - t_2 = t_3 - t_4$ then $\omega_{t_1-t_2} = \omega_{t_3-t_4}$. Now, as we assumed that states are separated by observables and vice versa, we can equivalently describe this time evolution on the observables. (Provided we assume that the experimental setups we use to define these observables do not depend on time.) Thus, given an

observable O we can uniquely define its trajectory O_t by

$$\omega(O_t) = \omega_t(O) \qquad \forall \omega \in \mathcal{S}.$$

Here $\omega = \omega_{t=0}$ and the states are considered not to depend on time once we consider observables time dependent.
Thus, in particular each element $f \in C(X)$ will define a unique trajectory f_t. We shall not assume that f_t is an element of $C(X)$ for all $t \in \mathbb{R}$. However, by the definition it is very natural to assume that the algebraic structure is invariant under this time evolution. Thus, for fixed t the set of all f_t will again form a C^*-algebra that is isomorphic to $C(X)$. We shall denote the family of C^*-algebras so obtained by $\{C(X)_t\}_{t\in\mathbb{R}}$.
Returning now to the representation $(\pi, \mathcal{H}, C(X), \psi)$ we first note that, as it admits a cyclic vector ψ every other representation $(\tilde{\pi}, \tilde{\mathcal{H}}, C(X), \tilde{\psi})$ that admits a cyclic vector $\tilde{\psi}$ such that

$$\langle \psi, \pi(f)\psi \rangle = \langle \tilde{\psi}, \tilde{\pi}(f)\tilde{\psi} \rangle \qquad \forall f \in C(X)$$

is unitarily equivalent to $(\pi, \mathcal{H}, C(X), \psi)$, i.e there exists a unitary operator $U : \tilde{\mathcal{H}} \to \mathcal{H}$ such that $\pi(\tilde{f}) = U^*\pi(f)U$. Assume that the cyclic vector ψ represents a physical state, ω say. As any vector in $\mathcal{H}$ is cyclic for $\pi(C(X))$ this will be possible whenever there exists at least one vector in $\mathcal{H}$ which is physical. Thus so far this is a very mild assumption.
We shall however also assume that then the full trajectory ω_t is represented by elements ψ_t of $\mathcal{H}$. In other words we shall assume that the representation is stable under the time evolution, at least on a sufficiently large set of states. Certainly this need not be the case. Yet it is a very natural assumption, and, in fact, it is valid for all known physical systems (that fall in the class considered here).
Then, as the vectors ψ_t are cyclic, each of them defines a new representation π_t via:

$$\pi_t(f)\psi = \pi(f)\psi_t$$

Namely, the set $\pi_t(C(X))$ will be dense in $\mathcal{H}$. However, then we can define the operators

$$\pi_t(f)\left(\pi_t(g)\psi\right) = \psi_t(fg)\psi.$$

Since the so defined operators are bounded they can be uniquely extended to all of $\mathcal{H}$.
Moreover, the representation $(\pi_t, \mathcal{H}, C(X), \psi_t$ will by the above remark be unitarily equivalent to the original one. Thus there will for all t exist a unitary operator $U(t)$ such that for all $f \in C(X)$ we have

$$\pi_t(f) = U(t)^*\pi(f)U(t).$$

One may then assume that the operators $U(t)$ define a strongly continuous group of unitaries. (The group properties actually follow from translation invariance.) Then there will exist a selfadjoint operator H on $\mathcal{H}$ such that

$$U(t) = e^{iHt}.$$

Note that for any (strongly) smooth family $A_t = U(t)AU^*(t)$ of operators it then follows that

$$\dot{A}_t := \frac{\partial}{\partial t}(U(t)AU^*(t)) = U(t)(i[H, A])U^*(t) = i[H, A_t].$$

Furthermore it is important to stress that the smoothness of the family A_t with respect to the strong topology does not imply that $\dot{A}_t$ is a bounded operator. It only requires that this operator is densely defined. In contrast the boundedness of $\dot{A}_t$ would be implied by smoothness with respect to the norm topology.

This then motivates the following

Definition 1:

A closed physical system with N degrees of freedom is given by $(\mathcal{H}, \pi, C(X), H)$, where:

- *$\mathcal{H}$ is a Hilbert space.*
- *$C(X)$ is the algebra of continuous functions on an N-dimensional manifold X.*
- *$\pi : C(X) \to \mathcal{L}(\mathcal{H})$ is a faithful representation.*
- *$\mathcal{H} = L^2(X, E)$ is the space of square integrable sections of some complex line bundle $E \to X$ over X.*
- *H is a selfadjoint operator on $\mathcal{H}$.*
- *There is a dense subalgebra $\mathcal{A}$ of $C(X)$ such that the family $\pi_t(a) = e^{iHt}\pi(a)e^{-iHt}$ is strongly smooth in t for all $a \in \mathcal{A}$.*

In addition there are sometimes given subsidiary conditions that specify which states on $C(X)$ are (un)physical. There might also exist other observables that are represented on $\mathcal{H}$ as bounded, selfadjoint operators. We shall not specify them here.

Remark:

Note that in the above definition we did not assume that there does exist a cyclic vector as this would be a somewhat unphysical restriction. We shall

see this later. There is no clear physical reason, why there should be such a vector. In fact, it is actually not used in calculations. What is used, however, is that the commutant $\pi(C(X))'$ is the weak closure of $\pi((C(X))$ itself. This does, of course, only make sense for a commutative algebra, and it it would easily follow from the existence of a cyclic vector. However, it is also true if $\mathcal{H}$ is isomorphic to the space of square integrable section of a complex line bundle see the discussion in the mathematical preliminaries). In fact, representations for topologically nontrivial bundles are realized in nature, as is proven experimentally via some topological effects like e.g. the Aharonov-Bohm effect.
Note that this commutant condition is operationally well motivated: It simply excludes that the Hilbert spaces is chosen too large. Otherwise, as is easily shown, there would exist a genuine subspace of unphysical states, invariant under the action of $C(X)$. We could then simply project on the orthogonal complement of this subspace.
Finally, as we shall see below, the requirement of the existence of a complex line bundle may be seen as a (nontrivial but natural) regularity condition, as it ensures that also tangent vectors, (which in classical physics would correspond to components of the velocity), are represented on $\mathcal{H}$, and that also the algebra of smooth functions of compact support on X is faithfully represented on $\mathcal{H}$.

Remark:
In order to gain some flexibility one could actually replace the algebra $C(X)$ in the above definition by an arbitrary C^*-algebra. In fact, in the usual C^*-theoretic description of physical systems this is done. However as will become clear in the next section, the physical interpretation will then still rely on the study of its commutative subalgebras which, by the Gelfand-Naimark-Theorem would be isomorphic to $C(\Sigma)$ for some topological space Σ.

3.2 Physical Interpretation

Before we come to the construction of concrete models it will be helpful to obtain a somewhat clearer picture of the physical interpretation of the mathematical formalism that has been motivated in the last section.

To start with, a physical system has been defined essentially as a family of representations π_t of representations of the C^*-algebra $C(X)$ that is generated by the unitaries e^{iHt}. For simplicity we assume that X is compact. Note that

we do *not* assume that the subalgebras $\pi_t(C(X))$ and $\pi_{t'}(C(X))$ of $\mathcal{L}(\mathcal{H})$ do commute for $t \neq t'$. Moreover there may be other selfadjoint operators on $\mathcal{H}$ that can be interpreted as operationally well-defined observables. We do also not assume that these operators commute with the π_t.
Every physical state ω induces for all t a state on $\pi_t(C(X))$, i.e. a positive, bounded, normalized linear functional. The Riesz-Markov-Theorem then asserts (shifting the time dependence back to the states for the moment) that it is given by a unique (regular,..) Borel measure μ^t_ω, viz

$$\omega(f) = \int\limits_X f(x)\, \mathrm{d}\mu^t_\omega(x).$$

Now, according to probability theory we may then interpret μ^t_ω as a *probability measure* (or probability distribution) on the probability space X. (If we would actually specify some σ-algebra of (Borel-)subsets of X.) The corresponding *random variables* would then be given by the elements of $C(X)$, i.e. by observables. Hence we directly obtain the probabilistic interpretation we intended. Note that it is also known that a probability measure is uniquely characterized by its "moments", i.e. by by the expectations of a countable subset of $C(X)$, whose linear range is dense. This then gives a probabilistic interpretation of the fact that the elements of $C(X)$ separate the states. Of course, we may take as $\mu(x)$ a Dirac distribution $\delta(x-y)$. In that case every measurement of every $f \in C(X)$ would give the same value $f(y)$. However, as pointed out above due to the inaccuracy of any measurement, it is wise to consider more general states, where we might find different values each time we measure f. We shall see a better reason below. The coarseness of these measurements is then in each state ω well described by the variance

$$\Delta_\omega(f)^2 = \omega((f - \omega(f))^2) = \omega(f^2) - (\omega(f))^2\,.$$

Analogously ω would produce a probability measure on the spectrum of any commutative C^*-algebra that is represented on $\mathcal{H}$. Note that it may happen that two such commutative algebras, that do also commute with each other, generate a larger commutative C^*-algebra of observables. For all these algebras we would obtain probability measures.

As we shall see below, the probability measure for different commutative subalgebras need, however, not be compatible in any reasonable sense. Before coming to this point, let us analyze the interpretation of actual measurements for the commutative subalgebras. For simplicity of notation, we shall consider only a situation in which there exists an observable A with purely discrete

spectrum, and such that no eigenvalue is degenerate, i.e all eigenspaces are one-dimensional. The distinct eigenvalues of A will be denoted by λ_n with $n \in \mathbb{N}$ and corresponding eigenvectors ψ_n. As the latter form an orthonormal basis of $\mathcal{H}$ (remember that A is selfadjoint and bounded), every $\phi \in \mathcal{H}$ can be decomposed as

$$\phi = \sum_n p_n \psi_n \qquad\qquad p_n \in \mathbb{C}.$$

In order to provide a state on the C^*-algebra generated by A, the vector ϕ has to be normalized, viz

$$\sum_n |p_n|^2 = 1.$$

In this case the expectation value of A in the state ϕ is given by

$$\omega_\phi = \langle \phi, A\phi \rangle = \sum_n |p_n|^2 \lambda_n.$$

This then strongly suggests to interpret:

In a measurement of A one will find one of the eigenvalues λ_n as possible outcome. The numbers $|p_n|^2$ are then interpreted as the probability to find the value λ_n in such a measurement. No value other than one of the eigenvalues can be found.

In the more general case the eigenspace corresponding to some eigenvalue λ_n is not one dimensional, the the probability to find λ_n in a measurement would be given by $\langle \phi, P_n \rangle$, where P_n denotes the orthogonal projection on eigenspace to the eigenvalue λ_n. Finally, we should also comment on the case of an observable with continuous spectrum. To be explicit we may consider the case where $\mathcal{H} = L^2(\mathbb{R}^3)$ and consider the characteristic function χ_K of a compact subset $K \subset \mathbb{R}^3$. Strictly speaking this is of course not a continuous function but it nevertheless defines a bounded, selfadjoint operator on $\mathcal{H}$ with eigenvalues 0 and 1. Thus it is a projector. The expectation value

$$\langle \phi, \chi_K \phi \rangle = \int_K |\phi(x)|^2 \mathrm{d}^3 x$$

can therefore be interpreted as the probability to find the particle (respectively the system) in K. Accordingly we interpret $|\phi(x)|^2$ as the corresponding probability density.
The above probabilistic interpretation of the measurement of an observable

A is, obviously, only consistent, if we assume that the system is in the corresponding eigenstate (-space) after the measurement. For then we know the state of the system as concerns A. Thus one also assumes:

As a result of a measurement of the eigenvalue λ of the observable A in the state ϕ the system will be in the state $P_\lambda\phi$, where P_λ is the orthogonal projection on the eigenspace corresponding to λ.

This axiom, usually also called the "*collapse of the wave function*", that appears rather natural at this point, is nevertheless the subject of an intense and controversial debate ever since the invention of quantum mechanics. We shall comment on this debate later, as it only becomes nontrivial in view of the noncommtativity of observables in quantum physics. If there were only mutually commuting observables (as is the case in classical mechanics) then this axiom is indeed completely harmless: Even if – in the example above – the support of the probability density $|\phi(x)|^2$ is all of $\mathbb{R}^3$ and we find the particle nevertheless in a small region K, this obviously causes no paradox. We would simply infer that the particle has been in K at the time of the measurement, although we assumed (for some reason) before the measurement that there was a finite probability to find it somewhere else. If any other observable commutes with χ_K, and thus leaves the subspace $\chi_K\mathcal{H}$ invariant, no measurement before or after that of χ_K will lead to a contradiction with this conclusion. Thus we may indeed consistently conclude that the result of our measurement has in fact been predetermined, and that our ignorance (or inability to obtain sufficient information) has prevented us from knowing the result in advance. (This will be discussed in more detail later.)

However there might also be (subalgebras of) observables that do not commute with each other. In particular up to now we did not assume that the algebras $\pi_t(C(X))$ and $\pi_{t'}(C(X))$ do commute with other, and we should try to find whether it is a physically realistic option if they don't.

Let A and B be any two selfadjoint operators on $\mathcal{H}$ and let ψ be any vector of norm 1 in their common domain and that of their products (so that the following expressions are all well-defined). We then set $\omega(A) = \langle\psi, A\psi\rangle$ and likewise for $\omega(B)$. (In other words ω will define a state on the C^*-algebra generated by A and B. We also define $\tilde{A} = A - \omega(A)$ and $\tilde{B} = B - \omega(B)$ so that we have $\Delta^2_\omega(\tilde{A}) = \omega(\tilde{A}^2)$ and $\Delta^2_\omega(\tilde{B}) = \omega(\tilde{B}^2)$. Note that $\omega(A^2)$ is real (and positive). Thus $\tilde{A}$ is selfadjoint, and so is $\tilde{B}$.
Now the operator $C = (\tilde(A) + i\lambda\tilde{B})(\tilde{A} - i\lambda B)$ will be positive for all $\lambda \in \mathbb{R}$.

Thus, as ω is a state, so is $\omega(C)$, viz

$$\omega(\tilde{A}^2) + \lambda^2\omega(\tilde{B}^2) + i\lambda\omega([\tilde{A}, \tilde{B}]) \geq 0 \qquad \forall\lambda \in \mathbb{R}.$$

We should remark that $i\lambda\omega([\tilde{A}, \tilde{B}])$ is real. Not that if $\omega(\tilde{B}^2) = 0$ the above inequality cannot hold for all λ if $\omega(\tilde{A}^2)$ is finite, unless $\omega([\tilde{A}, \tilde{B}]) = 0$. Otherwise we may set $\lambda = -\frac{\omega([\tilde{A},\tilde{B}])}{2\omega(\tilde{B}^2)}$ to obtain

$$4\omega(\tilde{A}^2)\omega(\tilde{B}^2) \geq |\omega([\tilde{A}, \tilde{B}]|^2.$$

Hence we obtain that for all selfadjoint operators and all states one has

$$\Delta_\omega(A)\Delta_\omega(B) \geq \frac{1}{2}|\omega([A, B])|.$$

Now suppose that for some $f \in C(X)$ we have that for some $t \neq t'$

$$[\pi_t(f), \pi_{t'}(f)] \neq 0.$$

Now, since the elements of $\mathcal{H}$ can, by construction, be identified with functions on X, one may think that the above commutator might only be nonzero when applied to elements of $\mathcal{H}$ that are "poorly localized", i.e whose support is rather large or all of X, and such that they are greater than a given $\epsilon > 0$ on rather large sets. However, as all the states can be given as linear combinations of vectors whose support is contained in an arbitrarily small compact set, one easily shows that generically there have to exist well localized states on which the above commutator does not vanish. We shall see this later in the examples. Thus, for such a state we get that there exists some bound $\kappa > 0$ such that $\Delta(\pi_t(f))\Delta(\pi_{t'}(f)) \geq \kappa$. Thus if such a state was well localized at time t, say, so that the corresponding variance is small,then it will be poorly localized at t'. (Though this conclusion is of course not strictly true. However we shall see in the concrete examples that it comes out so). In other words a probability distribution μ^t that is well localize ed at some time, will start to spread its support more and more as time goes by. The "smearing of wave functions" takes place.
From the point of view of a classical physicist, this situation is certainly unacceptable, as it seemingly contradicts determinism, i.e. that the complete knowledge of the state of a physical system at one time is sufficient to have complete knowledge of the state at any other time: The better one knows a state at one time the lesser it is known at later times and this solely due to the time evolution!

One might therefore be tempted to assume that the representations π_t have to commute at different times. We shall do so in the next section. Only to find that this (historically) lead to theories that do contradict experimental facts.

Before we come to this point, we should however briefly mention another problem that arises if one adopts a probabilistic interpretation also to non-commuting observables:
In view of the spectral decomposition theorem we can associate to each observable its family of spectral projections. (Strictly speaking these spectral projections do not (necessarily) belong to the C^*-algebra, but to some suitable completion of it). As it only has the eigenvalues 1 or 0 we may consider these projections also as genuine observables, namely the "Yes/No" experiments. Note that two operators commute if and only if their spectral projections do commute.
We can then consider these "Yes/No" experiments as the elementary propositions about the state of the system. (e.g. "The state is contained in a subregion of X (or not)"). One finds such elementary propositions also in any classical probability theory. Given a projection P we let $P\mathcal{H}$ denote its range, i.e. the subsets of states on which the proposition attains the value "true". It is then natural to define in analogy to classical logic the operations "AND", "OR" and "NOT" on these projections. Hence, given P_1, P_2 we shall denote by $P_1 \cap P_2$ (P_1 and P_2) the orthogonal projection on the intersection of $P_1\mathcal{H}$ and $P_2\mathcal{H}$. Likewise $P_1 \cup P_2$ (P_1 or P_2) will be the orthogonal projection on the union of the ranges of the two projections. Finally $\bar{P} = 1 - P$ (Not P), will correspond to the logical negation. Now, in analogy to classical logic, one might also expect that the distributional law

$$P_1 \cap (P_2 \cup P_3) = (P_1 \cap P_2) \ \cup \ (P_1 \cap P_3)$$

holds. Indeed if the three projections P_1, P_2, P_3 mutually commute, this is easily shown, using that then the set of eigenvectors of all three projections agree. However, if they do not commute this is not the case:

This is most easily seen via a simple example. We take the Hilbert space $\mathbb{C}^2$ (resp. a two-dimensional subspace). Then one may take as P_1 the orthogonal projection on the subspace spanned by the vector $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. The projector P_2 will be the orthogonal projection on the subspace spanned by $e_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Thus $\bar{P}_2 = 1 - P_2$ will project on $\bar{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. One easily checks that P_1 and P_2 do not commute. Moreover, obviously $P_1 \cap (P_2 \cup \bar{P}_2) = P_1 \cap 1 = P_1$. On the other hand, since the intersection of

the eigenspaces is obviously empty we have that $P_1 \cap P_2 = 0 = P_1 \cap \bar{P}_2$.

We may now return to the discussion of the "collapse of the wave function". We argued above, that if there were only observables that mutually commute, we could actually assume after the measurement of one of them that the system has actually ever been in the eigenstate we found in the measurement. However, this conclusion will no longer work if some observables do not commute with each other: For instance, if we would measure P_2 in the above example, and obtain the answer "Yes", we would according to this reasoning conclude that the system has always been in the state e_2. But what if we would have measured P_1 before P_2 ? Certainly we would have obtained some answer for the measurement of P_1, so for definiteness let's say "Yes" (obviously it doesn't matter which one). Thus we would have concluded that the system has always been in the state $e_1 = \frac{1}{\sqrt{2}}(e_2 + \bar{e}_2)$. If we thus immediately (that is to say quickly enough that the time evolution of the state e_2 can't have reached e_1 yet) measure P_1, then we obviously get an incompatible answer. The system can't be in the states e_2 and e_1 at the same time. Thus the outcome of a measurement can not have been predetermined in this case. After the measurement of P_1 there really is a probability of $\frac{1}{2}$ to find each eigenvalue of P_2. But then the assumed collapse of the wave function during the measurement of P_2 really does change the state of the system and it does so instantly. Even if the observables P_1 and P_2 would correspond to measurements in different regions of the universe. We shall discuss later some of the seemingly paradoxical conclusions that result from this fact but, more importantly, also its experimental significance.

3.3 Deterministic Systems

As we saw in the last section, if as a consequence of the time evolution of the system the representations π_t of $C(X)$ do not commute at different times, then this will lead to a delocalization of states. Thus our knowledge of the state of the system would diminish as time goes by. As this does, seemingly, contradict determinism, and certainly did not appeal to classical (Newtonian) physicists, the following two definitions (which follow the analysis of Bell) seem to provide natural additional assumptions:

Definition 2 :
Let O be any observable and let $O_t = e^{iHt}Oe^{-iHt}$ denote its time evolution. We say that O is a beable if $[O_t, O_{t'}] = 0$ for all t, t'.

The term beable here is to be read as be-able, in sharp contrast to observable. The point is that the property of the system that we measured by some observable that is not a beable will more and more be resolved during the time evolution. Thus we can only state that we have this property at some t, but no more than that. In contrast to that if the observable is in fact a beable this property would be preserved during the time evolution. Thus in this sense, beables refer to properties of the states that "are" (and remain), and not simply to mere observations.

Definition:
We shall call a system deterministic if all $f \in C(X)$ are beables.

Thus if a state of a deterministic system is a Dirac delta-distribution, then it will remain so during the time evolution. If we know the state of the system completely at one instance of time, then we shall know it at all times. Note that for the faithful representations of $C(X)$ that we consider here, we have that $\pi_t(C(X))' = \pi_t(C(X))''$. Hence we obtain that if the system is deterministic, then for any t, t' we have that$\pi_t(C(X))' \subset \pi_{t'}(C(X))''$. Hence, as by von Neumanns Theorem $\pi_{t'}(C(X))''$ is the weak closure of $\pi_{t'}(C(X))$, there is a sense subset of $\pi_t(C(X))$ that actually belongs to $\pi_{t'}(C(X))$. It is thus natural to assume that there is dense subalgebra of $C(X)$ which is stable under the time evolution. If one now assumes that this algbra contains a dense subset of smooth functions, the one gets that the time evolution $f \to f_t$ where $\pi(f_t) = \pi_t(f)$ is actually a diffeomorphism. The strong continuity then enforces the generator H of the time evolution be a vector field, i.e. a first order differential operator. We shall give a full account of this argument elsewhere. Here we restrict ourselves to a brief outline of the derivation. As we shall see in the next section, the system described here is not realistic anyway.
Now in order to get second order equations of motions (as is empirically the case for the basic equations of classical physics) we have to interpret the space X as phase space. Thus points in X, i.e the pure states, describe the position and the velocity of the particle, say, at a given time. Here this then means that we would have to assume that there are local coordinates $q_1, \dots q_n, p_1, \cdot_n$ such that the Hamiltonian locally has the form

$$H = \sum_i \left(p_i \frac{\partial}{\partial x_i} + F_i(x,p) \frac{\partial}{\partial p_i} \right) + K(x,p)$$

Hence one would obtain Newtons law $\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = F_i(x,p)$.

Remark:
In fact, as we shall see later, we are not completely forced to these conclusions. They require certain additional assumptions and it would not be worthwhile to explain them here, as the result is simply unphysical. We will, however, later meet a deterministic system that is much more realistic (though not completely), namely Dirac's equation. We should also note in passing that the above operator H is not bounded from below, and this creates severe problems concerning the dynamical stability of the system. We shall have to discuss them in connection with the Dirac equation.

3.4 The need for Quantum Mechanics

Thus, all seems to be fine so far. The assumptions of the previous sections have lead us – albeit on a seemingly unnecessary long and rickety road – to Newtonian mechanics. An elegant and clearcut theory that even got us to the moon. Yet some may object that without the discovery of quantum mechanics there would have been no live broadcast of Armstrongs first steps in Television. And they would be serious, I'm afraid. Unfortunately, they would even be right. Quantum mechanics did in fact bring so many dramatic changes in every days live[2] that that alone provides strong enough arguments to show that it describes nature much more accurately than classical mechanics. Unfortunately, for the majority of people this also seems to be the only kind of argument they find acceptable.

Yet, it is hard to believe that Max Planck introduced a modification of classical physics because he thought it might be useful in case someone makes it to the moon someday. In fact he did so for very good reasons, and fortunately they still can be found in many books on quantum mechanics. However, not in this one. I thought for a while that I really should explain some of the discrepancies of classical physics with experimental data that had arisen around 1900 and eventually lead to the development of quantum mechanics. Hopefully I made myself clear enough in the above (somewhat cynical) introduction to this section: To my point of view there have been extraordinary grave theoretical reasons which really forced such a dramatic change in the paradigms of physics. Yet, to fully appreciate the strength of the arguments requires a very deep and thorough understanding of classical

[2]E.g. the modern computer- and information technology would not have been possible without the discovery of the Transistor and the Laser. Both these technologies thus essentially rely on pure quantum effects.

physics, and much more time and space than I had during the lecture, respectively in this script. Thus I added some references in the bibliography that provide excellent reviews of these issues.
On the other hand, the historical path to a great discovery, is rarely the most easy to follow. Moreover, quantum theory, like all physical theories can not be rigorously derived from empirical facts, and most most probably sooner or later it will be overcome by another, better theory. Better in the sense that it explains more experimental findings than quantum theory and starting from fewer assumptions. Just as quantum theory explains far more than classical theory.
Thus after all, I came to conclude that one no longer needs to explain why quantum mechanics was invented. Its fully sufficient to explain that it really does describe nature much more accurate than classical physics, and that it does indeed start from fewer assumptions. In fact, as we shall see in the next section, almost all predictions of quantum mechanics are a direct consequence of one single assumption: Heisenberg's uncertainty relation.

Chapter 4

Nonrelativistic Quantum mechanics

As briefly mentioned in the last section, there has been overwhelming evidence around 1900 that classical (Newtonian) mechanics had to be replaced by a different theory. One may argue that the double slit experiment by Gelmer and Davies clearly shows that in this theory $[\pi_t, \pi_{t'}] \neq 0$ (unless there are some hidden variables. We shall postpone the discussion of the existence of hidden variables to the next section.)
Moreover,by carefully analyzing the possible experimental setups to measure position and momentum of a microscopic object of (small) mass m, W. Heisenberg then found that there did not seem to be a way to perform an experiment such that $\Delta(v)\Delta(x) < \frac{\hbar}{2m}$ where $\Delta(O)$ denotes the inaccuracy of the measurement of an observable O in a given state.
Thus in view of the relation

$$\Delta_\omega(A)\Delta_\omega(B) \geq \frac{1}{2}|\omega([A,B])|$$

it appears natural to assume that for all t one has

$$[x_i(t), \dot{x_j}(t)] = \delta_{ij}\frac{i\hbar}{m},$$

though one is of course not forced to this conclusion yet. However, that this assumption alone is sufficient to explain all the experimental finding that lead lead to dismiss classical mechanics, as well as being accordance with all the other experimental facts. Moreover it lead to many nontrivial predictions that were later (and still are) confirmed by experiments. Thus, at any rate, assuming the above commutator relation of position and velocity, leads to a theory that is to be preferred over classical mechanics.

However, working with coordinates x as is done above has, to my point of view, several aesthetic and practical disadvantages. First of all, a physical result should not depend on the coordinates chosen to calculate (or measure) it. Thus in dealing with coordinate expressions one always has to check that the result of a calculation does in fact not depend on the coordinates chosen to write it down. Secondly, if the configuration space of the system is topologically nontrivial, then there will not exist global coordinates. However, if one only uses local coordinates, one might not be able to derive global (topological) effects (which are observed in nature). Thus it appears preferable to work with a coordinate free expression.
To this end, we first observe that from the commutator $[x, \dot{x}] = i$ it follows that for all smooth functions $f(x)$ one has: $[f(x), \dot{x}] = i\frac{\partial f(x)}{\partial x}$. To see this one first proves this equality for polynomials, by using the Leibniz rule for commutators. It is then extended to all smooth functions by a density argument. Then, as for any function $g(x)$ we have that $\dot{g(x)} = \frac{\partial g(x)}{\partial x}\dot{x}$, and as functions do commute with each other, we obtain:

$$[f(x), \dot{g(x)}] = i\frac{\partial f(x)}{\partial x}\frac{\partial g(x)}{\partial x}.$$

This may still seem to depend on the coordinates – which it does not, in fact. However, for the following discussion it will be enough to observe that for all functions f, f the commutator $[f, \dot{g}]$ is itself a function (and thus a very special operator). It will turn out gratifying, not to limit the explicit expression for this commutator any further.

4.1 Quantum mechanics on generic configuration spaces

We shall now start to investigate the consequences of imposing the uncertainty relation in the above form. For simplicity we shall abbreviate

$$a_t := \pi_t(a) = e^{iHt}\pi(a)e^{-iHt} \qquad \forall a \in \mathcal{A}.$$

Moreover we shall henceforth assume that $\mathcal{A}$ can be chosen as the algebra of smooth functions of compact support on X, i.e. $\mathcal{A} = \mathcal{C}^\infty_\rfloor(\mathcal{X})$. This of course implies that X is equipped with with a smooth atlas, i.e. we shall assume that it is a smooth manifold. Hence we define:

Definition: *We shall call a closed physical system* $(\mathcal{H}, \pi, C^\infty_c(X), H)$ *with* N *degrees of freedom* **nonrelativistic scalar quantum mechanics** *if*

1.

$$[a_t, \dot{b}_t] \in \pi_t(\mathcal{A}) \qquad \forall \dashv, \lfloor \in \mathcal{A}.$$

2.

$$-i[a_t, \dot{a}_t] \geq 0 \qquad \forall a = a^* \in \mathcal{A}.$$

3. If for some selfadjoint $a \in \mathcal{A}$ one has $[a_t, \dot{a}_t] = 0$ for all t then it follows that $a = 0$.

Recall that by definition of a physical system we also know that

- $\mathcal{H} = L^2(X, E)$ for some complex line bundle $E \to M$.
- The commutant $\pi(\mathcal{A})'$ is the weak closure of $\pi(\mathcal{A})$.
- $\dot{a}_t = i[H, a_t]$ for all $a \in \mathcal{A}$.

Remark:
Some remarks are in order here:
The term "scalar" in the above definition refers to the restriction to line bundles $E \to X$. In principle it would also have been possible to take as $\mathcal{H}$ the space of square integrable sections of a bundle of higher rank. We shall later come back to this possibility when we discuss particles of higher Spin.
The third condition in the above definition essentially requires that there is an uncertainty relations for all directions in any local chart: If it would not hold for some $a \in \mathcal{A}$, then one may take a as a local coordinate function (at least on some open set in X on which it is a diffeomorphism to $\mathbb{R}$).
The physical relevance of the second condition will become apparent later. We should stress however, that the following arguments do not make use of this condition. It is only added to exclude some examples that are not in accordance with the empirical facts and would render the name "quantum mechanics" in the definition problematic. Thus in this sense it can be viewed as an experimental finding.

Lemma:
The following formula defines a smooth family of Riemannian metrics $\{g_t\}_{t\in\mathbb{R}}$ on X:

$$g_t(f^{(1)}\mathrm{d}f^{(2)}\,,\, h^{(1)}\mathrm{d}h^{(2)}) := -if_t^{(1)}h^{(1)}[f_t^{(2)}, \dot{h}_t^{(2)}]$$

where $f^{(1)}, f^{(2)}, h^{(1)}, h^{(2)}$ are all selfadjoint, i.e. real valued functions.

A Sketch of the Proof:
We first need to prove that the above metric is well-defined as a symmetric Tensor of rank two on the sections of the cotangent bundle, i.e. on one-forms:

- By the uncertainty relation, i.e. condition 1 in the above definition, $g_t(f^{(1)}\mathrm{d}f^{(2)}\,,\,h^{(1)}\mathrm{d}h^{(2)})$ is a function.

- Linearity with respect to $C(X)$ follows from the Leibniz-rule for commutators

$$[AB,C] = A[B,C] + [A,C]B \qquad\qquad \forall A,B,C$$

  (on appropriate domains), and the uncertainty relation.

- The Leibniz-rule for commutators and for the time derivative then also gives the compatibility of the definition of g_t with the Leibniz-rule for the exterior differential, viz

$$g_t(\mathrm{d}(f\tilde{f})\,,\,h^{(1)}\mathrm{d}h^{(2)}) = fg_t(\mathrm{d}\tilde{f}\,,\,h^{(1)}\mathrm{d}h^{(2)}) + \tilde{f}g_t(\mathrm{d}f\,,\,h^{(1)}\mathrm{d}h^{(2)})$$

- The symmetry of g_t follows from

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}[f_t,h_t] = [f_t,\dot{h}_t] - [h_t,\dot{f}_t]$$

- The positive definiteness and nondegeneracy of g_t then follows from conditions 2 and 3.

Finally, the stated smoothness follows from the smoothness of $U(t)$ and the fact that $[f,\dot{h}]$ is a bounded operator by the uncertainty relation.

As shown in the preliminaries, on any Riemannian manifold there exists to any smooth function f a canonical vector field $\mathrm{grad}(f)$, defined via :

$$\mathrm{grad}(f)(h) = g(\mathrm{d}f\,,\,\mathrm{d}h).$$

Moreover it is shown there that for any complex vector bundle $E \to X$ on a smooth manifold there exists a covariant derivative ∇ on the smooth sections of E. It is then obvious that ∇ extends to a densely defined symmetric operator on $L^2(X,E) = \mathcal{H}$. (We shall later comment on the selfadjoint extensions of this operator).
We shall abbreviate $\nabla_h = \nabla_{\mathrm{grad}(h)}$. Now since by definition of a covariant derivative

$$\nabla_h(f\psi) = \mathrm{grad}(h)(f)\psi + f\nabla_h\psi$$

one has

$$[f_t,\nabla_{h_t}] = [f_t,-\dot{h}_t]$$

and thus:

Proposition:
There exists for all t a complex linear map $\beta_t : \pi_t(\mathcal{A}) \mapsto \pi_t(\mathcal{A})'$ such that $\dot{h}_t = -\nabla_{h_t} + \beta_t(h_t)$.

Definition:
We shall call a nonrelativistic scalar Quantum mechanics **regular** *if the image of the above maps β_t is contained in $\pi_t(\mathcal{A})$, i.e. is always a smooth function.*

From now on we shall assume regularity in the above sense. One may then, of course, continue with the analysis of the properties of the map β. However there is a quicker route to the main result:
As also shown in the mathematical preliminaries there exists for all covariant derivatives ∇ and all metrics g on X a covariant Laplacian Δ_g^∇ on $\mathcal{H}$ such that:

- Δ_g^∇ has a dense domain and is symmetric.
- $$[\Delta_g^\nabla, f] = -\nabla_f + \frac{1}{2}\Delta_g(f)$$
  where $\Delta_g(f)$ is the Laplacian of f. Thus it is a smooth function.
- Thus
  $$[[\Delta_g^\nabla, f], h] = -ig(\mathrm{d}f\,,\,\mathrm{d}h) \qquad\qquad \forall f, g \in C_c^\infty(X)$$

Thus, taking into account that $\dot{h}_t = i[H, h]$ we see that for any covariant derivative ∇ we have that for all t

$$[[(\Delta_{g_t}^\nabla - H), h_t], f_t] = 0.$$

More explicitly one computes with the help of the above stated formulas that for all $h \in \mathcal{A}$ we have

$$[(\Delta_{g_t}^\nabla - H), h_t] = \frac{1}{2}\Delta_g(h_t) - \beta_t(h) \in \pi_t(\mathcal{A}).$$

The inclusion follows from regularity. Thus, commuting with the operator $\Delta_{g_t}^\nabla - H$ is a linear map from $C_c^\infty = \mathcal{A}$ to itself that admits the Leibniz-rule. Hence it is a vector field, γ say, on X. But then, by the same reasoning as

before, there exists a covariant derivative $\tilde{\nabla}$ and an element $\tilde{\beta}$ of $\pi_t(\mathcal{A})'$ such that

$$H = \Delta^{\nabla} + \tilde{\nabla} + \tilde{\beta}$$

But then, as explained in the preliminaries, any operator of the above form can be rewritten as the sum of a covariant Laplacian and an element of the commutant of $\pi_t(\mathcal{A})$ in a unique way. Finally we should also use the fact the, by assumption H did not depend on t.

Theorem:
Let $(L^2(X,E), \pi, C_c^\infty(X), H$ be a regular nonrelativistic scalar quantum mechanics. Then there exists a unique metric g, a unique covariant derivative ∇ and a unique measurable function V on X such that

$$H = \Delta_g^{\nabla} + V.$$

We should remark that for a trivial bundle E with ∇ being the Levi-Civita connection and $V = 0$ it is known by Chernovs Theorem that the Laplacian is selfadjoint if and only if g is geodesically complete. Thus we can also conclude the geodesical completeness here. Note however that the above Theorem only holds under several additional assumptions like regularity and strong smoothness of the time evolution. I'm not aware of any compelling physical argument why these suppositions should be expected to hold.

It is instructive but tedious to calculate the equations of motion for the position operator q_k referring to some local chart[1] in the above mentioned case of a trivial bundle, with Levi-Civita-connection and $V = 0$. As a result one obtains that q essentially obeys the geodesic equation. More precisely:

$$\ddot{q_k} + \langle \Gamma_k^{ij} \dot{q}_i \dot{q}_j \rangle = 0$$

where Γ_k^{ij} denote the Christoffel-symbols, summing over repeated indices is understood, and $\langle \cdot \rangle$ is the total symmetrization, i.e for a product of three operators, $\langle ABC \rangle = \frac{1}{6}(ABC + ACB + BAC + BCA + CAB + CBA)$.

4.2 Simple examples on $\mathbb{R}^n$

In order to gain some understanding of the physical significance of the uncertainty relation and the structure of the Operator H, that results from it, we

[1] In this case, local coordinates can be represented on $\mathcal{H}$ by using a partition of unity.

shall now discuss some simple, concrete and, most importantly realistic examples of nonrelativistic scalar quantum mechanics. In later chapters, more examples of quantum systems will follow.
We shall, however, only state the solution of the equations of motion for these systems, but not describe the derivation of these results which can be found in most textbooks.
Moreover, in order to reduce the complexity of expressions to a minimum, these solutions are only given in "natural" units. Thus we have consistently $\hbar = 1 = m$, while the unit of charge e of the particle is 1, 2 or $\frac{1}{2}$ depending on the circumstance. Whenever this appears necessary, the magnitude of the corresponding physical effect for macroscopic bodies, respectively for the electron is given, so that no misunderstanding can arise: The quantum mechanical effects described in this chapter do not play a role on the macroscopic scale, but are very important and easy to observe on the microscopic scale (e.g. for single atoms or electrons). Note however that this does, of course, not mean that there can be no quantum effects that are observable on the macroscopic scale. Indeed, for systems with a large number of particles some quantum effects can (and often do) accumulate: Many properties of materials can not be understood with classical (non quantum mechanical) models, even though that is often not apparent without a detailed analysis. Yet there are also many spectacular effects that obviously have no classical explanation, for instance the Quantum Hall effect, Superconductivity and -fluidity, or the quantization of the magnetic flux.
In these notes, however, we shall only be concerned with systems that contain only a few particles. Thus effects of this type will play no role. Moreover, on the atomic scale, spacetime can well be approximated as being flat: For instance for the hydrogen atom the masses of the participating particles, the proton and the electron are extremely small ($m_p \approx 10^{-24}g$, and $m_e \approx 10^{-27}g$) as compared to their distance $10^{-8}cm$. The curvature of spacetime, i.e. the gravitational field, only plays an important role on the macroscopic scale (compare e.g with the mass of the earth $m_{Earth} \approx 10^{+27}g$ and its distance to a satellite of approximately $10^{+9}cm$). However, gravityy should play a leading role if distances are extremely small, as must have been the case in the very early universe.
For the systems considered in this chapter we can therefore assume flat spacetime, i.e. $X = \mathbb{R}^N$, equipped with the euclidean metric. As $\mathbb{R}^N$ is parallelizable it is then clear that Δ^∇ is given in the global coordinates x_i as

$$\Delta^\nabla = -\frac{1}{2}\sum_i (\frac{\partial}{\partial x^i} - A_i)^2.$$

Here A_i are the (well defined) components of the one form characterizing ∇ with respect to the frame $\frac{\partial}{\partial x^i}$.
Let us now compute the equation of motion for the operators x_i : (we shall comment on domain questions later)

$$\ddot{x}_k = F_k^L(x, \dot{x}) + F_k^C(x)$$

where $F_k^C = -\frac{\partial V}{\partial x_k}$ is the force felt by a particle in the electric potential V, while $F_k^L = \frac{1}{2}\sum_{ij} \epsilon_{kij}\,(\dot{x}_i B_j(x) + B_j(x)\dot{x}_i)$ is the symmetrized Lorentz-Force, i.e the force felt by a particle moving with velocity $\dot{x}$ in a magnetic field $B_j(x) = \sum_{mn} \epsilon_{jmn}(\frac{\partial A_n}{\partial x_m})$. ($\epsilon_{ijk}$ denotes the completely antisymmetric symbol. Thus $\epsilon_{123} = 1$, and for all i, j, k one has $\epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{jik}$. Hence $\epsilon_{ikk} = \epsilon_{kki} = 0$.)
Note that in classical mechanics also other forms of the Force would be allowed. In particular it would be allowed to have terms of higher order in $\dot{x}$ or friction terms. Thus we may already conclude:

The only forces compatible with Heisenberg's uncertainty relation are the gravitational and the electromagnetic force.

Later we shall see that, if one allows bundles E of higher rank, also Yang-Mills interactions are compatible with the uncertainty relation. Thus, from the point of view of a quantum physicist it is no surprise that no other fundamental forces are found in nature.
In any case, we are now in a position to discuss some more concrete examples. For all these examples we shall actually not consider the equation of motion for the observables but rather will consider the time evolution of states and thus consider the observables as constant. In the physics literature this is usually called the "Schrödinger picture". (In contrast to the "Heisenberg picture" that we mainly considered so far.) Note that given the wave function $\psi(x)$ at some initial time, $t = 0$ say, it will be given at any time t by

$$\psi(x, t) = e^{iHt}\psi(x).$$

In particular it will obey the equation

$$-i\frac{\partial}{\partial t}\psi(x, t) = (-\frac{1}{2}\sum_i (\frac{\partial}{\partial x_i} - A_i(x))^2 + V(x))\psi(x, t)$$

which is called the Schrödinger equation.

4.2.1 The Schrödinger representation of the Weyl-algebra

In the sequel we shall only consider examples where $A_i = 0$. (Examples with $A_i \neq 0$ will be considered later.) One then obviously would get $\dot{x} = -i\frac{\partial}{\partial x}$ (in one dimension, say). As shown before, both the operators x and $\dot{x}$ are essentially selfadjoint on $L^2(\mathbb{R})$.
Let us now consider the C^*-algebra generated by the two strongly continuous families of unitaries

$$U(\alpha) = e^{i\alpha x}, \qquad V(\beta) = e^{i\beta\dot{x}} \qquad \alpha, \beta \in \mathbb{R}.$$

This algebra is called the Weyl-algebra. One readily verifies that for all $\psi(x) \in L^2(\mathbb{R})$

$$U(\alpha)\psi(x) = e^{i\alpha x}\psi(x) \qquad V(\beta)\psi(x) = \psi(x+\beta)$$

and from this one gets:

$$U(\alpha)V(\beta) = e^{-i\alpha\beta}V(\beta)U(\alpha)$$

Proposition:
The Weyl-algebra acts irreducibly on $L^2(\mathbb{R})$.

Proof: If not, there exists a subspace $\tilde{\mathcal{H}}$ invariant under the action of the Weyl algebra. Thus for all ϕ in the orthogonal complement of $\tilde{\mathcal{H}}$ and all nonvanishing $\psi \in \tilde{\mathcal{H}}$ we get from $\langle \phi, U(\alpha)V(\beta)\psi\rangle = 0$:

$$0 = \int e^{i\alpha x}\overline{\phi(x)}\psi(x+\beta)\,\mathrm{d}x.$$

The above integral is the Fourier transform of $\overline{\phi(x)}\psi(x+\beta)$ taken in the point α. Now, since the Fourier transform is an isomorphism from $L^2(\mathbb{R})$ into itself, its vanishing implies $\overline{\phi(x)}\psi(x+\beta) = 0$. Since $\beta \in \mathbb{R}$ can be chosen arbitrarily this requires $\phi(x) = 0$.

4.2.2 The free particle and the smearing of the wave function

As the first and simplest example we shall consider a particle in one dimension, $X = \mathbb{R}$, that is not subject to any force, i.e. $A_i = V = 0$.
For definiteness we shall consider the initial wave function $\psi_\sigma(x) = C_\sigma e^{-(\frac{x}{2\sigma})^2}$

at $t = 0$, i.e a Gaussian with variance σ, normalized to 1. It readily verified that ψ_σ lies in the domain of the operators H, x and $\dot{x} = -i\frac{\partial}{\partial x}$. Moreover

$$\langle \psi_\sigma, x\psi_\sigma \rangle = C_\sigma \int x e^{-\frac{x^2}{2\sigma^2}} \, \mathrm{d}x = \sigma^2 C_\sigma \int \frac{\partial (e^{-\frac{x^2}{2\sigma^2}})}{\partial x} \, \mathrm{d}x = 0$$

while

$$\langle \psi_\sigma, x^2\psi_\sigma \rangle = C_\sigma \int x^2 e^{-\frac{x^2}{2\sigma^2}} \, \mathrm{d}x = -2\sigma^2 C_\sigma \frac{\partial}{\partial \alpha} \left(\int e^{-\frac{\alpha x^2}{2\sigma^2}} \, \mathrm{d}x \right)_{\alpha=1} = -\sigma^2$$

so that $\Delta_{\psi_\sigma}(x) = \sigma$.
One then easily checks that

$$\psi(x,t) = \psi_{\sigma_t} e^{\frac{2i\sigma}{t}}$$

where

$$\sigma_t = \sigma \sqrt{1 + (\frac{\hbar t}{2m\sigma^2})^2}$$

is the corresponding solution of the Schrödinger equation. Thus, as expected from our previous discussion, the variance for the position operator increases with time t.

Note that for the convenience of the reader we reintroduced the constants $\hbar$ and m in the last formula above. The time to double the initial spread is of the order of 10^{-13} seconds for the electron if $\sigma = 10^{-8}cm$. Thus even for a very good initial localization the wave function delocalizes extremely fast in this case.
On the other hand, for a small macroscopic body with $m = 10^{-3}g$ and $\sigma = 10^{-3}cm$ it takes 10^{11} years till the initial spread doubled.

4.2.3 The Quantum Harmonic Oscillator, energy quantization and the Tunnel-Effect

Next, we shall add the potential $V(x) = \frac{x^2}{2}$ (still in one dimension), so that we get what is called the quantum harmonic oscillator. The equation of motion is given as

$$\ddot{x} = -x.$$

Let's consider the classical dynamical system corresponding to this equation. Recall that the energy of the system, here $E = \frac{\dot{x}^2}{2} + \frac{x^2}{2}$ is a conserved quantity, i.e. $\frac{\mathrm{d}E(x(t))}{\mathrm{d}t} = 2(\ddot{x} + x) = 0$ for any solution $x(t)$ of the equation of motion.

But then, since $\dot{x}^2$ is positive, we conclude that for any solution of energy E one has

$$x(t) \in [-\sqrt{2E}, +\sqrt{2E}] \qquad \forall t \in \mathbb{R}.$$

As we shall see this property is not true for the quantum mechanical harmonic oscillator.

Let's now return to the quantum case. The Hamilton-Operator is then given as

$$H = -\frac{1}{2}\left(\frac{\partial}{\partial x}\right)^2 + \frac{x^2}{2} = a^* a + \frac{1}{2}.$$

Here we have already introduced the two operators

$$a = \frac{1}{\sqrt{2}}(x + \frac{\partial}{\partial x}) \qquad a^* = \frac{1}{\sqrt{2}}(x - \frac{\partial}{\partial x}).$$

since x and $i\frac{\partial}{\partial x}$ do have a common dense domain and are essentially selfadjoint on this domain, it is clear that the above operators are defined on a common dense domain on which they are the adjoints of one another. Moreover it follows that H is densely defined and positive $H \geq 0$ on its domain and finally from Friedrich's Theorem that H is essentially selfadjoint.
One then checks that

$$[a, a^*] = 1, \qquad [H, a] = -a, \qquad [H, a^*] = a^*$$

and that there is a unique (up to a complex constant) vector $\psi_0 \in L^2(\mathbb{R})$ such that

$$a\psi_0 = 0 \qquad \Rightarrow \psi_0(x) = Ce^{-\frac{x^2}{2}}.$$

Obviously $H\psi_0 = a^* a\psi_0 + \frac{1}{2}\psi_0 = \frac{1}{2}\psi_0$, $\frac{1}{2}$ is an eigenvalue of H. Moreover the vectors $\psi_n = (a^*)^n\psi_0$ are eigenvectors of H since $[H, (a^*)^n] = n(a^*)^n$ (by the Leibniz rule) and thus:

$$H\psi_n = (a^*)^n H\psi_0 + n(a^*)^n\psi_0 = (n + \frac{1}{2})\psi_n.$$

One easily checks by direct computation that all ψ_n are square integrable and have a nonvanishing norm.

Proposition:
The vectors ψ_n are a complete orthogonal set. Thus the operator H has purely discrete spectrum

$$\sigma(H) = \{(n + \frac{1}{2}), \quad n = 0, 1, 2, \ldots\}$$

Proof:
Mutual orthogonality of the ψ_n is clear, since they belong to different eigenvalues.
Next one notes that the subspace they span is invariant under the action of a^* and a and thus also under the action of $x = \frac{1}{\sqrt{2}}(a + a^*)$ and $i\frac{\partial}{\partial x} = \frac{i}{\sqrt{2}}(a^* - a)$. Accordingly it is also invariant under the action of the Weyl-algebra these operators generate. By irreducibility of the Schrödinger-representation of the Weyl-algebra it is therefore dense in $L^2(\mathbb{R})$. This completes the proof.

Thus we get that, unlike in the classical case, the set of possible values for the measurement of the energy is discrete, i.e. there is a "**Quantization of the energy levels**". Moreover, while in the classical case solutions of a given energy E are bound to stay in a compact interval, this is not the case for the quantum harmonic oscillator. Even for the smallest value of the energy there is a finite (albeit) small probability to find the particle outside the interval $[-1, 1]$, i.e. in the region that is forbidden by energy conservation in the classical case. This is the so called **Tunnel-Effect**.
We should stress however that there is nevertheless no violation of the conservation of energy in the quantum case: H and x do not commute with each other. Thus if the particle is found in the region outside $[1, -1]$, i.e. after a measurement of the position, it is no longer in an eigenstate of the energy (if it was before the measurement).
Note that once we are given a complete set of eigenstates of H it is trivial to calculate the solution $\psi(t)$ for any given initial vector ψ: As there do exist complex numbers α_n such that

$$\psi = \sum_{n=0}^{\infty} \alpha_n \psi_n, \qquad \sum_n |\alpha_n|^2 < \infty$$

the corresponding solution is by linearity of H given as

$$\psi(t) = \sum_{n=0}^{\infty} e^{i(n+\frac{1}{2})} \alpha_n \psi_n.$$

4.2.4 Some comments on the Hydrogen atom

We shall finally consider the potential $V(x) = -\frac{2}{r}$ in 3 dimensions, where $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$ is the distance to the origin. In classical physics this potential describes e.g. the gravitational force between two massive bodies (in natural units), one of which is located in the origin. Thus one of the two bodies might e.g. be the sun, the other a planet. Thus obviously among the

solutions of the classical equations of motion there are also "bound states" where the planet moves along an ellipse in a fixed plane around the sun. As concerns these bound states, they all have negative energy (the potential energy corresponding to the above potential is negative) and their maximal radius depends on the energy: The larger the energy , the larger the radius, which can, in fact, become arbitrarily large. On the other hand, there are also solutions with arbitrarily small energy (i.e. $|E|$ is very large, as E is negative for these solutions. The maximal radius tends to zero as E approaches $-\infty$.
The plane in $\mathbb{R}$ in which the orbit of the planet stays depends on the angular momentum of the planet relative to the sun (which is also a conserved quantity since the above potential is symmetric under rotations.) Every plane in $\mathbb{R}^3$ is possible.

On the other hand the above potential also describes (in other natural units) the electric (Coulomb-) force between to particles of opposite charge. One would then expect that the solutions of the equations of motion look the same as for the gravitational case, essentially. Well, there is a difference, however: A charge that moves on a circular (or elliptical or any accelerated) orbit radiates and thus looses energy to the outside electromagnetic field. Accordingly, there are no stable solutions for the Coulomb-force: All particles fall in into the particle at the origin after some finite time.
By Rutherford's experiments it is also well known, that a hydrogen atom consists of a proton (sitting in the center) and an electron, bound to it, but being typically quite far away. As the electromagnetic force is by more than a factor 10^{34} stronger than the gravitational (force) for these two particles (the other forces play no role here and are even smaller than the gravitational one, if present at all), it is clear that the electron must be bound by the Coulomb-force.
According to classical physics one would thus expect that

- Any value of the energy of the electron is possible (i.e. is in the spectrum of the Hamilton-Operator)
- Depending on the energy, any radius of the atom is possible.
- To any value of the energy there exist infinitely many bound states
- The hydrogen atom is not stable. The electron typically drops into the proton after 10^{-10} seconds.

Needless to say, none of the above predictions of classical physics is realistic, or even close to reality. To start with, hydrogen atoms are quite stable

and typically exist since 10^{10} years. Moreover if a hydrogen atom is subjected to electromagnetic radiation it only absorbs certain discrete values of the energy, which are given by $E_0(\frac{1}{n^2} - \frac{1}{m^2})$ where n, m are any two strictly positive integers and $E_0 = 13.6eV$. Likewise the atom only admits radiation according to this formula. Finally all energy levels have a finite degeneracy and apprixametely the same radius of about $10^{-10}m$.

In fact, it turns out that the negative part of the Hamilton-Operator $\Delta - \frac{2}{r}$ is given as $-\frac{E_0}{n^2}$ for all $n = 1, 2 \ldots$ and each level is $(2n+1) - fold$ degenerate. In particular it is bound from below (unlike the classical energy) by $-E_0$.
This explains all the experimental findings (in contrast to classical physics, which explained none) apart from one: Actually I was a bit unfair to classical physics when mentioning that the classical solutions would be unstable: This is only true if one takes into account the coupling of the electron to the electromagnetic field. However this has not been done for the quantum hydrogen atom. Thus the discreteness of the spectrum of the quantized hydrogen atom that is not coupled to the electromagnetic field explains nothing. In fact, the coupled Hamilton-Operator has purely continuous spectrum. Nevertheless one can prove the stability of the ground state of the hydrogen atom within quantum electrodynamics (as has recently been done by Bach, Fröhlich and Sigal), but that's to be appreciated as the highly nontrivial result of an extremely difficult analysis.

4.3 Spin

As turned out in the Stern-Gerlach-experiment described below, the wave function of the electron is not correctly described as a (square integrable) section of a complex *line* bundle: A much more appropriate description is obtained by using sections of a certain bundle of rank 2. The precise nature of this bundle has only been understood later in the relativistic description in terms of the **Dirac equation**. In fact, in general a bundle of rank 4 has to be used. We shall describe the Dirac equation and *Spin bundles* later. It will then be shown that, in the nonrelativistic limit on $\mathbb{R}^3$, Pauli's approach to describe the electron by the section of a trivial bundle of rank 2 is fully justified as an excellent approximation. Further evidence for the existence of the Spin is given by the *Feinstructure* and the *Zeeman-effect*, but also by the Polarization effects in scattering of electrons. These effects will all be explained in the discussion of the Dirac equation.

To be more concrete, the Hilbert space is then given by

$$\mathcal{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^2.$$

Besides the representations π_t of $C_0^\infty(\mathbb{R}^3)$ there are then also additional observables of the form $\mathrm{id} \otimes m$ with m a two by two matrix acting on $\mathbb{C}^2$. Obviously these observables together separate the states. The time evolution will then again be governed by a Hamilton-Operator which a priori is now an operator on $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. In the next subsection we shall describe the form of this operator as derived from empirical facts. Later it shall then be obtained as the nonrelativistic limit of the Dirac-Hamiltonian.

A relevant basis of two by two matrices is given by the identity and the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The **Spin-Operators** $S_i = \frac{1}{2}\sigma_i$ then obey the commutation relations

$$[S_i, S_j] = i\epsilon_{ijk} S_k$$

of the Lie-algebra $\mathbf{su(2)}$. In fact, as we shall see the Spin can be interpreted as an "internal" angular momentum of the electron. The total Spin $S^2 = \sum_i S_i^2$ is an conserved quantity, i.e. it commutes with the Hamilton-operator for all choices of the external (gravitational, electromagnetic, etc) fields, For the electron one then has $S^2 = \frac{1}{2}(\frac{1}{2}+1) = \frac{3}{4}$. One says the electron has Spin $\frac{1}{2}$. There are also particles with Spin 0 (e.g. the π-meson), Spin 1 (e.g. photons, Gluons or the W- and Z- mesons) and higher Spins (e.g. the Ω-mesons with Spin $\frac{3}{2}$).

4.3.1 The Stern-Gerlach-Experiment

As stated above, the relation of the Spin to the internal angular momentum, and thus the dynamics of the Spin degrees of freedom, is only understood in the relativistic theory, and we shall thus postpone the discussion thereof. In the chapter on the Dirac-equation we shall discuss all the dynamical couplings of the Spin and the corresponding observations in atomic physics. For this and the next chapter it is fully sufficient to describe the coupling of the Spin to the magnetic field, which historically also lead to the discovery of the Spin.

The relevant term is given as

$$H_P = 2\mu_B \mathbf{S} \cdot \mathbf{B}(x) = \mu_B \sum_i \sigma_i B_i(x)$$

where the constant $\mu_B = \frac{e}{mc}$ is called the Bohr-Magneton. We should remark that this term resembles the energy $\mu_B \mathbf{L}\cdot\mathbf{B}$ of a magnetic dipole, i.e. a charge with angular momentum $\mathbf{L}$ in the magnetic field $\mathbf{B}$.
The total Hamiltonian is then given as

$$H = H_0 + H_P$$

where $H_0 = \Delta^\nabla + V$ denotes the Spin-independent term considered above. The resulting equations of motion in the Heisenberg picture are then given as

$$\ddot{x}_k = F_k^L(x, \dot{x}) + F_k^C(x) + 2\mu_B \sum_i \sigma_i \frac{\partial B_i}{\partial x_k}.$$

Thus, unless the magnetic field is constant, there is an additional Force that results from the coupling of the Spin to the magnetic field.

Let us now consider an electrically neutral atom. Then the Lorentz Force F_k^L and the Coulomb-Force F_k^C will vanish. Assuming then that the atom has an effective Spin $\frac{1}{2}$ only the term H_P will lead to a nonvanishing force. An example for such an atom is Ag (silver). To a very good approximation another example is, of course, provided by the Hydrogen atom: Note that (as for any atom) the nucleus has a much larger mass m_N than the electron, which is 2000 times lighter than the proton. Hence the coupling of the Spin of the nucleus (the proton has Spin $\frac{1}{2}$) given by $2\frac{e}{m_N}\sigma_P \cdot \mathbf{B}$ can to a good approximated be negelected as compared to the same term for the electron. Suppose further that the magnetic field depends only on one coordinate, x_3 say, i.e. $\frac{\partial B}{\partial x_1} = \frac{\partial B}{\partial x_2} = 0$. (Such a magnetic field is easily realized experimentally). Then, if a beam of electrons is subjected to B it will obviously split in two beams corresponding to the two different eigenvalues $\pm\frac{1}{2}$ of S_3. This is the (result of the) classical experiment performed in 1924 by Stern and Gerlach. Actually they intended to investigate whether the Ag-atom has a magnetic dipole moment, i.e. whether there is a nontrivial contribution of the form $\mathbf{L}\cdot\mathbf{B}$ to the *classical (unquantized)* Hamiltonian. One would then expect, classically, as this term can take a continuum of values, that the width of the beam in the x_3 direction will be enlarged. The outcome of the experiment therefor came as a complete surprise.

Note that if the experiment is performed with a magnetic field that has a gradient only in the x_3-direction, one may take one of the two beams, corresponding to the eigenvalue $+\frac{1}{2}$ say, and subject it to another magnetic field with a gradient only in a direction *perpendicular* to x_3. E.g. the x_2-

direction. As the two eigenspinors of the matrix σ_2 are given by

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

the beam will then split again in two beams that contain equally many atoms (as the probability to find each of the eigenvalues ± 1 of σ_2 is $\frac{1}{2}$ if the system is in the eigenstate

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

of σ_3.
It is an easy excercise to show that from these experimental observations and the Ansatz $2\mu_B \sigma \cdot \mathbf{B}$ the above form of the matrices σ_k and their commutation relations can be deduced.

Note that it then follows, and more importantly from the necessary invariance of the term $\sigma \cdot \mathbf{B}$, that the 3-tupel S_k transforms like a vector under the change of orthonormal frame x_i of $\mathbb{R}^3$. Thus if O is any orthogonal 3×3-matrix, then

$$S_i' = \sum_k O_{ki} S_i$$

will be the components of the Spin with respect to the rotated (by O) orthonormal frame.
The same statement of course also applies to the tuple of matrices σ_k. We may then introduce the component of the Spin with respect to any normalized vector $\mathbf{a} \in \mathbb{R}^3$, $\mathbf{a}^2 = \sum_k a_k^2 = 1$, as

$$S_{\mathbf{a}} = \frac{1}{2}\mathbf{a} \cdot \sigma = \frac{1}{2}\sum_k a_k \sigma_k.$$

The eigenvalues of $S_{\mathbf{a}}$ will then obviously correspond to measurements of the Spin along the $\mathbf{a}$-axis (e.g. by a Stern-Gerlach-type experiment).

4.3.2 Two particles with Spin $\frac{1}{2}$

It is worthwile to consider the case of two identical electrically neutral atoms (as described above) in more detail. The Hilbert space may then be described as $\mathcal{H} = L^2(\mathbb{R}^6) \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ with a typical wavefunction given as a linear combination of terms of the form $\psi(x^1, x^2) \otimes \xi^1 \otimes \xi^2$. Here $x^1, x^2 \in \mathbb{R}^3$ denote the position of the two particles. The Hamilton-Operator is then given as

$$H = H_0^1 + H_0^2 + \mu_B(\sigma_{\mathbf{1}} + \sigma_{\mathbf{2}}) \cdot \mathbf{B}(x^1, x^2) =: H_0^1 + H_0^2 + H_1$$

if we neglect the gravitational interaction of the two atoms (as they are neutral there is no elctromagnetic interaction).
We shall in the sequel only consider the Spin part $\mathbb{C}^2 \otimes \mathbb{C}^2$. Let then

$$\sigma_k^1 = 1 \otimes \sigma_k \qquad\qquad \sigma_k^2 = \sigma_k \otimes 1$$

denote the observables corresponding to the Spin of the first, respectively second particle. Note that H above commutes with the operator

$$\sigma_{total}^2 = \sum_k (\sigma_k^1 + \sigma_k^2)^2$$

and with the operator τ that interchanges the two particles, in particular

$$\tau(\xi^1 \otimes \xi^2) = \xi^2 \otimes \xi^1.$$

Thus we also have

$$[\tau, \sigma_{total}^2] = 0.$$

Hence every eigenspace to these operators is invariant under the time evolution. Note that the eigenspace to the eigenvalue -1 of the operator τ is one-dimensional the elements being of the form

$$\xi^1 \otimes \xi^2 - \xi^2 \otimes \xi^1,$$

while the eigenspace to the eigenvalue $+1$ is three-dimensional. If e_1, e_2 denotes a basis of $\mathbb{C}^2$ then a basis of the latter eigenspace is given as

$$e_1 \otimes e_1, \qquad\qquad e_2 \otimes e_2, \qquad\qquad e_1 \otimes e_2 + e_2 \otimes e_1.$$

One easily checks that

$$\sigma_{total}^2(\xi^1 \otimes \xi^2 - \xi^2 \otimes \xi^1) = 0$$

while e.g.

$$\sigma_{total}^2(e_1 \otimes e_1) = 2e_1 \otimes e_1$$

and in fact $\sigma_{total}^2 \xi = 2\xi$ for all $\xi \in \mathbb{C}^2 \otimes \mathbb{C}^2$ with $\tau\xi = \xi$.

Let ξ_s denote the (up to multiplication with a phase factor) unique normalized eigenvector to the eigenvalue -1 of τ in $\mathbb{C}^2 \otimes \mathbb{C}^2$. Then we have

$$e^{iHt}\sigma_{total}^2 e^{-iHt}\xi_s = \sigma_{total}^2 \xi_s = 0 \qquad\qquad \forall t \in \mathbb{R}.$$

Thus if the system is at some moment of time in Spin-state ξ_s it will remain in this Spin-state at all times.

Then, if we measure for some component of the Spin, e.g. the component S_3, say, for particle one the eigenvalue $\frac{1}{2}$ the measurement of the same component of the Spin for particle two will necessarilly give the value $-\frac{1}{2}$, as one easily checks that $\mathbf{b} \in \mathbb{R}^3$ one has

$$(\mathbf{b} \cdot \sigma^1)(\mathbf{b} \cdot \sigma^2)\xi_s = -\xi_s.$$

To see this one may use a basis β_+, β_- of $\mathbb{C}^2$ of eigenvectors of $\mathbf{b} \cdot \sigma$. Then it is clear that there exists a complex number z such that

$$\xi_s = z(\beta_+ \otimes \beta_- - \beta_- \otimes \beta_+).$$

Next let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ are two *orthogonal* normalized vectors, $\mathbf{a}^2 = \mathbf{b}^2 = 1$ and $\mathbf{a} \cdot \mathbf{b} = 0$. We shall denote the basis of eigenvectors of $\mathbf{a} \cdot \sigma$, respectively $\mathbf{b} \cdot \sigma$, by $\alpha_\pm$, respectively $\beta_\pm$. One then computes that there exist a complex number w such that

$$\alpha_\pm = w(\beta_+ \pm \beta_-)$$

and from this it follows that

$$(\mathbf{a} \cdot \sigma^2)(\mathbf{b} \cdot \sigma^1)\xi_s = wz(\beta_- \otimes \beta_- + \beta_+ \otimes \beta_+).$$

Hence the vectors ξ_s and $(\mathbf{a} \cdot \sigma^2)(\mathbf{b} \cdot \sigma^1)\xi_s$ are orthogonal. Therefore the expectaion value

$$\big(\xi_s, (\mathbf{a} \cdot \sigma^2)(\mathbf{b} \cdot \sigma^1)\xi_s\big) = 0$$

vanishes in this case. Here we have denoted the sscalar product in $\mathbb{C}^2 \otimes \mathbb{C}^2$ by $(\cdot, \cdot)$.
If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ are two arbitrary normalized vectors, we may decompose $\mathbf{a} = (\mathbf{a} \cdot \mathbf{b})\mathbf{b} + \tilde{\mathbf{a}}$ where $\tilde{\mathbf{a}} \cdot \mathbf{b} = 0$. We then obtain for the expectation value for the measurement of $(\mathbf{a} \cdot \sigma^2)(\mathbf{b} \cdot \sigma^1)$ in the state ξ_s:

$$\big(\xi_s, (\mathbf{a} \cdot \sigma^2)(\mathbf{b} \cdot \sigma^1)\xi_s\big) = -(\mathbf{a} \cdot \mathbf{b}).$$

Chapter 5

In how far can the Quantum description of reality be considered complete ?

The title of this chapter refers to the 1935 paper by Einstein, Podolski and Rosen, entitled "Can the Quantum Mechanical description of reality be considered complete." One may expect that the answer to such a simple YES/NO question may not be found. However, one may also expect that the question can be modified in such a way that the answer is within reach of an experimental test. For instance, one may ask, whether the quantum mechanical description of reality can be considered complete, if, in addition, we assume that any such description must necessarily be a *"local" theory*. In fact, this is the modification J.S. Bell proposed in 1964 (including a clear and rather general definition of locality). He then showed that this question can indeed be answered with the help of certain experiments he proposed. A.Aspect and others then carried out these experiments and obtained the answer "Yes". We shall comment on the interpretation of this outcome later. First we shall have a closer look on these questions, Bell's ingeniously simple idea and the experimental verification.

5.1 Hidden variables and the EPR paradoxon

In the paper mentioned in the introduction above, Einstein and his coauthors intended to show that quantum mechanics can't be a complete theory, whether or not one believes – as Einstein did – in "*local realism*".By that Einstein meant that any property of the system that can be measured with a sharp value in an experiment must be a real (existing) property of the

system. " The Moon is still there, even if nobody looks at it" as he said. We already saw in the discussion of beables and deterministic theories that this point of view is not in accordance with quantum theory. For example, the electron does not have a well-defined position: We can only measure such a sharp position of the electron at some instant of time, but at the next instance the electrons wave function will again be spread over some large region of space. It is worthwhile to discuss Einstein's objection and the way out that he suggested in some detail. We shall however not consider the systems used in the original paper respectively for the experimental test. Instead we shall use a simpler system, where the argument is most transparent.

Consider two electrically neutral atoms (e.g. hydrogen atoms) in the singlet Spin-state ξ_s, as defined in the last chapter. Let us now reintroduce the spatial part $\psi(x^1, x^2)$ of the wave function for this system. We may then prepare an initial state in which the two particles are most likely found to be in two distant points $x, y \in \mathbb{R}^3$ say. The separation of the particles might be achieved in such a way that the Spin part of the wave function is always the singlet state ξ_s (e.g. by using gravitational forces).
Choosing a normalized vector $\mathbf{a} \in \mathbb{R}^3$ we may then measure the component $\mathbf{a} \cdot \sigma^1$ of the Spin for the atom at the point x. Denoting as before the eigenvectors of $\mathbf{a} \cdot \sigma^1$ in $\mathbb{C}^2$ to the eigenvalues ± 1 by $\alpha\pm$ the system will then be in the state

$$\alpha_- \otimes \alpha_+$$

if we assume for definiteness that we have found the value -1 in the measurement. A subsequent measurement of $\mathbf{a} \cdot \sigma^2$ in the point y after some time t will thus give the value $+1$, irrespective of t. In particular one may choose t to be short enough, so that no signal (in particular light) that is emitted at x can reach y in the time t, i.e. $t < \frac{|x-y|}{c}$, where c denotes the speed of light. The result of the measurement then seemingly contradicts causality, as the atom at the point y cannot know that we performed a measurement of the spin of its partner at the point x. One may still object that it will take long enough to prepare the measurement, i.e. to adjust the magnets, and that the state of the atoms may be influenced by the arrangement of this setup. However, as pointed out by D.Bohm, one may of course measure another component $\mathbf{b} \cdot \sigma^2$ of the Spin at y, and it is easily seen that the apparatus (i.e. the magnet) can be rotated quickly enough from $\mathbf{a}$ to $\mathbf{b}$ so that no signal emitted at x can reach y during the time of the rotation and the subsequent measurement. Indeed, if the measurements is performed with an ensemble of such pairs of atoms, then one would find for the expectation

value of measurement of $\mathbf{a}\cdot\sigma^1$ followed by that of $\mathbf{b}\cdot\sigma^2$ in the state ξ_s:

$$\big(\xi_s,(\mathbf{b}\cdot\sigma^1)(\mathbf{a}\cdot\sigma^2)\xi_s\big)=-\mathbf{a}\cdot\mathbf{b}.$$

Thus, unless $\mathbf{a}\cdot\mathbf{b}=0$ the two eigenvalues of $\mathbf{b}\cdot\sigma^2$ are not equally distributed, once the measurement of $\mathbf{a}\cdot\sigma^1$ has been performed. Thus also in this general case the result of the measurement of $\mathbf{b}\cdot\sigma^2$ is predetermined to some extent by that of $\mathbf{a}\cdot\sigma^1$.

Einstein, Podolsky and Rosen argued that, as the atom at y cannot know about the result and the measurement at x the only way it can happen that the result of the measurement at y is predetermined must be that it has been predetermined all the time. However, then the question arises why in the measurement of $\mathbf{a}\cdot\sigma^1$ both eigenvalues are found with equal probability.
They then proposed that there do exist hidden parameters, i.e. degrees of freedom of the system which can (or simply are) not measured , and which determine the outcome of any measurement uniquely. We shall describe this idea in the next section. Before we come to that, however, a few remarks are in order.

Note that it is not possible for the two experimentalists at x and y to submit any information faster than the speed of light by this experimental setup. The point is that they cannot know about the result of each others experiment, unless they they communicated about it, which however will require some signal that can at most propagate with the speed of light. This, of course, has been clear to Einstein. However he did assume that the atoms also need to know about this measurement.

The latter however need not be the case. In fact, Einstein did mistakingly assign an individual existence to the spins of the two atoms. This, however, is, from the modern point of view, not possible for the "entangled" singlet spate ξ_s. In fact the state Einstein, Podolsky and Rosen constructed in their paper can be viewed as the first example of an entangled state. The fact that such states can be realized experimentally nowadays is expected to cause – e.g. via the development of quantum computation and quantum cryptography – another technical revolution. Einstein, in the above paper was the first to realize that the most important difference of quantum and classical mechanics is the existence of such states in quantum theory.

In fact, what make these states – first shown to exist by Einstein, Podolsky and Rosen – so remarkable is that they are highly nonlocal, but have to

be viewed as a single entity. The (Spins of) the two atoms behave just as one irrespective of how far the two atoms have been separated.
Schrödinger was the first to point out that Einsteins argument rather than constructing a contradiction in the Kopenhagen interpretation of quantum theory in fact clearly pointed out the nonlocal character of quantum mechanics as compared to classical mechanics. He also introduced the fashionable term "entangled states" for such nonlocal states.

5.2 Bell's inequalities

Einstein's idea to resolve the problem he thought to be raised by the Gedankenexperiment (which became a real experiment only later) considered above has been that there exist hidden parameters. Thus, to stay in our operational language, the set of all observables does not separate the states, but to any state ω and any observable O there is a unique value $o(\omega)$ that can be found in measurements of O in the state ω.

Recall that for a system that admits hidden parameters it is not necessarily true that the sum (respectively the product) of two different observables is an observable. Thus, when dealing with hidden parameters we have to refrain from the C^*-algebraic framework advocated so far. This has first been observed by D.Bohm and later been stressed by J.S.Bell. In fact, after the appearance of the paper by Einstein, Podolsky and Rosen several authors published proofs that a theory of hidden parameters can not be consistent with the already existing empirical facts. The first such proof is due to von Neumann. Later Jauch and Pirron and then Gleason pointed out some important refinements of von Neumann's argument. However all these proofs made essential use of the assumption that the sum of two observables is always an observable.
Yet, in 1951 D.Bohm gave a concrete counterexample that showed that there do exist operationally well defined hidden variable theories which do not fulfill this linearity assumption but are in accordance with most empirical facts.

Thus it had become clear that to dismiss Einstein's idea – if that is possible at all – additional assumptions that are physically well motivated would be needed. J.S.Bell then pointed out in 1962 that such an additional assumption is locality. In fact, Einstein was mainly concerned with the apparent lack of locality in quantum theory. The main purpose of the hidden parameters then has been to restore this locality. Thus Bell assumed that the theory with hidden parameters should be a local theory.

To become more explicit Bell made the following assumptions on the hypothetical theory with hidden parameters that describes the Gedankenexperiment by Einstein, Podolsky and Rosen:

1. Besides the vector ξ_s the state of the system is described as a point $\lambda \in X$ in some *probability space* X.

2. Given a normalized vector $\mathbf{a} \in \mathbb{R}^3$ the measurement of the observable $\mathbf{a} \cdot \sigma^1$ in the state corresponding to the state $\lambda \in X$ will give the unique answer $A(\mathbf{a}, \lambda) \in \{-1, +1\}$. Thus for any such normalized $\mathbf{a} \in \mathbb{R}^3$ there is a random variable

$$A(\mathbf{a}, \cdot) : X \to \{-1, +1\}$$

3. Likewise, to the measurement of the observable $\mathbf{b} \cdot \sigma^2$ with $\mathbf{b}^2 = 1$ there is a random variable

$$B(\mathbf{b}, \cdot) : X \to \{-1, +1\}.$$

4. Given a state the corresponding point $\lambda \in X$ can not be determined by experiments. However one may assume that for any given ensemble of N such systems there is a probability distribution ρ i.e.

$$\rho(\lambda) \geq 0 \qquad \int_X \rho(\lambda)\, \mathrm{d}\lambda = 1$$

   such that $\rho(\lambda)N$ is the number of systems in the state $\lambda \in X$.

5. The expectation value of the observable $\mathbf{a} \cdot \sigma^1$ for that ensemble is then given as

$$P^1_{\mathbf{a}} = \int_X A(\mathbf{a}, \lambda)\rho(\lambda)\, \mathrm{d}\lambda.$$

   Likewise we have the expectation value

$$P^2_{\mathbf{b}} = \int_X B(\mathbf{b}, \lambda)\rho(\lambda)\, \mathrm{d}\lambda.$$

6. **(Locality)** By choosing the time interval between the measurement of $\mathbf{a}\cdot\sigma^1$ and $\mathbf{b}\cdot\sigma^2$ short enough we may assume that $A(\mathbf{a},\cdot)$ and $B(\mathbf{b},\cdot)$ are *independent random variables*. Thus the expectation value for the measurement of $\mathbf{a}\cdot\sigma^1$ followed by that of $\mathbf{b}\cdot\sigma^2$ is given as

$$P(\mathbf{a},\mathbf{b}) = \int_X A(\mathbf{a},\lambda)B(\mathbf{b},\lambda)\rho(\lambda)\,\mathrm{d}\lambda$$

Theorem (Bell's inequality):
Under the above six assumptions one has for any three normalized vectors $\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{R}^3$:

$$1 + P(\mathbf{b},\mathbf{c}) \geq |P(\mathbf{a},\mathbf{b}) - P(\mathbf{a},\mathbf{c})|$$

Proof:
Note that in view of the properties 2,3 and 4 $P(\mathbf{a},\mathbf{b})$ can not be less than -1. Moreover it can be -1 at $\mathbf{a}=\mathbf{b}$ (which is the experimentally observed value) only if

$$A(\mathbf{a},\lambda) = -B(\mathbf{a},\lambda)$$

for almost all $\lambda\in X$,i.e. except at a set of points λ of zero probability. Thus we may write

$$P(\mathbf{a},\mathbf{b}) = -\int_X A(\mathbf{a},\lambda)A(\mathbf{b},\lambda)\rho(\lambda)\,\mathrm{d}\lambda.$$

Then we have

$$\begin{aligned} P(\mathbf{a},\mathbf{b}) - P(\mathbf{a},\mathbf{c}) &= -\int_X \left[A(\mathbf{a},\lambda)A(\mathbf{b},\lambda) - A(\mathbf{a},\lambda)A(\mathbf{c},\lambda)\right]\rho(\lambda)\,\mathrm{d}\lambda \\ &= \int_X A(\mathbf{a},\lambda)A(\mathbf{b},\lambda)\left[A(\mathbf{b},\lambda)A(\mathbf{c},\lambda) - 1\right]\rho(\lambda)\,\mathrm{d}\lambda \end{aligned}$$

where $(A(\cdot,\cdot))^2 = 1$ has been used. Thus from the triangle inequality we obtain

$$|P(\mathbf{a},\mathbf{b}) - P(\mathbf{a},\mathbf{c})| \leq -\int_X \left[1 - A(\mathbf{b},\lambda)A(\mathbf{c},\lambda)\right]\rho(\lambda)\,\mathrm{d}\lambda$$

The statement of the theorem now follows, since the second term on the right hand side is $P(\mathbf{b},\mathbf{c})$.

The remarkable fact about this seemingly innocent inequality is that it is not obeyed by the **quantum mechanical expectation value**

$$(\xi_s, (\mathbf{b}\cdot\sigma^2)(\mathbf{a}\cdot\sigma^1)\xi_s) = -\mathbf{a}\cdot\mathbf{b}$$

computed in the last chapter. To see this one may take the explicit example of normalized vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ chosen such that

$$\mathbf{a} \cdot \mathbf{c} = 0, \qquad \mathbf{a} \cdot \mathbf{b} = -\mathbf{b} \cdot \mathbf{c} = \frac{1}{\sqrt{2}}.$$

Thus, the above inequality gives a possibility to distinguish hidden variable theories from quantum mechanics experimentally. Such experiments require a very high accuracy (in particular because the set of vectors in $\mathbb{R}^3$ for which the quantum mechanical prediction for the expectation value contradicts the inequality is extremely small) and could only have been performed in the 1980ies. The result of these experiments clearly rules out hidden parameter theories that obey the above 6 assumptions.

However, Bell himself stressed that this result does not rule out the possibility that, after all, the fundamental theory is a completely deterministic theory admitting hidden variables. But it can , in view of these experimental findings, not be a *local* theory in the above sense. Bell and Bohm gave explicit examples of nonlocal hidden variable theories that do describe the above Gedankenexperiments correctly and do not obey Bell's inequalities. Thus, in view of the fact that most theoretical physicists no longer consider locality as an indespensable pillar of physics, hidden variable theories are not ruled out at all. Yet, as I should stress, they do not seem to appealing to me either, as they imply that there is an important part of reality that we will never know about.

Chapter 6

Bibliography

Here a selection of references will be given, ordered by subject and with some additional remarks.

Excellent introductions to **functional analysis** and **operator algebras** are

- **B.Blackaddar** *Operator Algebras*, Enzycl. of Mathematical Sciences Vol.122, Springer 2006
- **O.Bratteli and D.W.Robinson** *Operator Algebras and Quantum Statistical Mechanics* Vol I, Springer 1979; Vol II, Springer 1981
- **J.Dixmier** *C^*-algebras* North Holland 1977
- **M.Reed and B.Simon** *Methods of Modern Mathematical Physics* Vol I 1980, Vol II 1975
- **F.Riesz and B.Sz.Nagy** *Functional Analysis* Fred.Ungar Publ., New York 1955 (Reprint: Dover 1990)
- **S.Sakai** *C^*-algebras and W^*-algebras* Springer 1971 (reprint 1998)
- **R.V.Kadison and J.R.Ringrose** *Fundamentals of the Theory of Operator Algebras* Vol I.-IV Academic Press 1983

Some (of) many good introductions to **Differential geometry** that mainly consider aspects that are important for applications in physics (except the rather general introduction given by Spivak) include

- **C.Bär, N. Ginoux and F.Pfäffle** *Wave Equations on Lorentzian manifolds and Quantization* ESI Lectures, EMS 2007

- **H.Baum** *Spin-Strukturen und Dirac-Operatoren über pseudoriemannschen Mannigfaltigkeiten* Teubner, Leipzig 1981
- **J.Gracia Bondia, H.Figueroa and J.Varilly** *Elements of Noncommutative Geometry* Birkhäuser 2001
- **T.Frankel** *The Geometry of Physics* Cambridge Univ.Press second edition 2004
- **M.Spivak** *A Comprehensive Introduction to Differential Geometry* Publish or Perish, Houston, second edition 1979
- **R.Wald** *General Relativity* Univ. of Chicago Press, 1984

Introductions to **Quantum mechanics**:
The books by Bell, Peres, d'Espagnat and Wheeler/Zurek are mainly concerned with the interpretation of Quantum mechanics, while the other books decribe the mathematical formalism in much more detail. While the classic by Landau Lifschitz gives a more conventional overview, Segal and Strocchi advertises the C^*-algebraic framework used here. The book by Strocchi also contains a nice introduction to the theory of Schrödinger Operators (as does the books by Reed and Simon given above). Finally, the books by Galindo and Pascual give excellent and very modern accounts of both, the interpretation and the technical and mathematical aspects of Quantuma mechanics. Unfortunately it is no longer in Press, even though its treatment appears much more modern than that of most textbooks that appeared later (and still do). To my point of view, it is still the best textbook on quantum mechanics.

- **J.S.Bell** *Speakable and Unspeakable in Quantum Mechanics* Cambridge Univ.Press revised edition 2004
- **Galindo and Pascual** *Quantum mechanics (1 +2)*, Springer 1991
- **L.D.Landau and E.M. Lifshitz** *Quantum Mechanics* Pergamon Press, third editition 1977
- **B. d'Espagnat** *Conceptual Foundations of Quantum Mechanics*, Addison Weslsy 1989
- **A.Peres** *Quantum Theory: Concepts and methods* Kluwer Acad. Press 1995
- **I.Segal** Ann.Math. (2), 48, 1947, p.930

- **Strocchi** *An Introduction to the Mathematical Structure of Quantum Mechanics* Advanced Series in Math.Phys. Vol.27, Worl Scientific 2005
- **J.A.Wheeler and W.H.Zurek** *Quantum Theory and Measurement* Princeton Univ. Press 1983