

Chapter 3

Formalism and interpretation

Gott würfelt nicht mit dem Universum!
Albert Einstein

*Ich denke nicht, dass es unsere Aufgabe ist dem Herrgott
Vorschriften zu machen ...*
Niels Bohr

*The theory of quantum electrodynamics describes Nature as ab-
surd from the point of view of common sense. And it agrees fully
with the experiment. So I hope you can accept Nature as She is -
absurd.*

Richard P. Feynman, "QED"

The mathematical formalism of quantum theory, which we want to develop in this chapter, is based on the fact that the solutions of the Schrödinger equation form a Hilbert space, i.e. a vector space that is complete with respect to the norm defined by an inner product. All equations of the theory can be interpreted in terms of operators, i.e. linear maps on this space. This point of view is useful for theoretical as well as for practical reasons. As an example, we will solve the Schrödinger equation for the harmonic oscillator purely algebraically by introducing creation and annihilation operators. Along the way we will discuss the axioms and the interpretation of quantum mechanics, derive the general uncertainty relation, and develop new concepts and computational tools like the Heisenberg picture and density matrices.

3.1 Linear algebra and Dirac notation

The Schrödinger equation is a linear homogeneous differential equation. Its set of solutions therefore forms a vector space \mathcal{H} over the complex numbers, because linear combinations of solutions with complex coefficients are again solutions. But this vector space is, in general, infinite dimensional. We should hence also admit infinite linear combinations so that convergence properties of such infinite sums have to be considered. The notion of convergence is based on a measure $\|v\|$ for the length of a vector, where a sequence is called convergent if the distance between its members and its limit vector goes to 0. The length $\|v\|$ has to be positive and is required to satisfy the triangle inequality $\|v+w\| \leq \|v\| + \|w\|$. It is called a norm on a vector space if it scales linearly according to $\|\alpha v\| = |\alpha| \|v\|$, where $|\alpha|$ is the modulus of the complex number $\alpha \in \mathbb{C}$. A vector space with such a norm, a normed space, is called Banach space if it is complete (i.e. if it contains the limits for all Cauchy sequences, where a Cauchy sequence is a sequence for which the distances between its elements converge to 0).

Observables in quantum mechanics, like momentum or energy, are given by linear operators, i.e. by linear maps, which are the analogues of matrices in finite-dimensional spaces. Many of the concepts and tools of linear algebra can be extended to infinite-dimensional linear spaces. This is the subject of the mathematical discipline of functional analysis [Reed, Kreyszig].

Hilbert spaces: In quantum mechanics there is a natural norm, namely the square root of the integral of the probability density of a wave function $\psi(x)$ at some given time t ,

$$\|\psi\| = \sqrt{Q} \quad \text{with} \quad Q = \int_{\mathbb{R}^3} d^3x |\psi(x)|^2 \quad (3.1)$$

(as we have shown it is time-independent for solutions of the Schrödinger equation). This norm has the additional property that it can be defined in terms of an inner product (φ, ψ) by

$$\|\psi\| = \sqrt{(\psi, \psi)} \quad \text{with} \quad (\varphi, \psi) = \int_{\mathbb{R}^3} d^3x \varphi^*(x) \psi(x). \quad (3.2)$$

An inner product (φ, ψ) is semi-bilinear and symmetric up to complex conjugation,

$$(\varphi, \alpha\psi_1 + \beta\psi_2) = \alpha(\varphi, \psi_1) + \beta(\varphi, \psi_2), \quad (\varphi, \psi) = (\psi, \varphi)^*, \quad (3.3)$$

where semi-bilinear means linear in the second entry and anti-linear in the first,

$$(\alpha\varphi_1 + \beta\varphi_2, \psi) = \alpha^*(\varphi_1, \psi) + \beta^*(\varphi_2, \psi). \quad (3.4)$$

as implied by eq. (3.3).

Note that anti-linearity (i.e. the complex conjugation of scalar coefficients) for the first entry is necessary because strict bilinearity would be inconsistent with positivity of the norm

$\|\psi\|^2 = (\psi, \psi) \geq 0$. To see this compare $\|(i\psi)\|^2 = (i\psi, i\psi)$ with $\|\psi\|^2$. A Banach space whose norm is defined by (3.2) in terms of a positive definite inner product (φ, ψ) is called a Hilbert space. The standard Hilbert space \mathcal{H} of quantum mechanics is the space of complex-valued square-integrable functions $\psi(x) \in \mathcal{H} = L^2(\mathbb{R}^3)$, where the letter L stands for *Lebesgue* integration (which has to be used to make the space complete).¹ This is an ∞ -dimensional vector space.

Let us pretend for a while that our Hilbert space is a finite-dimensional complex vector space. We will introduce a number of concepts like commutators and exponentiation of linear operators. The definitions will be straightforward for (finite-dimensional) matrices, but the same calculus can then be used for linear operators in Hilbert spaces. Refinements that are needed for the infinite-dimensional situation will then be discussed later on.

In linear algebra each vector space V automatically provides us with another linear space, called the *dual space* V^{dual} , which consists of the linear maps $w \in V^{dual}$ from vectors $v \in V$ to numbers $w(v) \in \mathbb{C}$. The numbers $w(v)$ are real for real and complex for complex vector spaces, respectively. We can think of vectors $v \in V$ as column vectors and of dual vectors $w \in V^{dual}$ as line vectors, so that their product, the duality bracket $\langle w, v \rangle \equiv w(v)$ is a number. If we introduce a basis e_i of V we can write each vector v as a unique linear combination $v = v^i e_i$ and each co-vector $w = w_j e^j$ is a sum of the elements of the dual basis e^j , which has upper indices and is defined by $\langle e^j, e_i \rangle = \delta_i^j$. Evaluation of w on v by linearity thus implies the formula

$$\langle w, v \rangle \equiv w(v) = w_j \langle e^j, e_i \rangle v^i = w_j v^j, \quad \text{with} \quad w = w_j e^j, \quad v = v^i e_i, \quad \langle e^j, e_i \rangle = \delta_i^j. \quad (3.5)$$

If we now make a change of basis $\hat{e}_i = G_i^j e_j$ then the components of vectors transform with the inverse transposed matrix, and the same is true for the dual basis vectors \hat{e}^j :

$$v = v^i e_i = \hat{v}^i \hat{e}_i, \quad \hat{e}_i = G_i^j e_j \quad \Rightarrow \quad \hat{v}^i = v^k (G^{-1})_k^i = (G^{-1T})^i_k v^k, \quad (3.6)$$

$$\delta_i^j = \langle e^j, e_i \rangle = \langle \hat{e}^j, \hat{e}_i \rangle, \quad \hat{e}_i = G_i^j e_j \quad \Rightarrow \quad \hat{e}^j = e^l (G^{-1})_l^j = (G^{-1T})^j_l e^l. \quad (3.7)$$

Co-vectors $w \in V^{dual}$, on the other hand, transform in the same way as the elements e_i of the basis, $\hat{w}_j = G_j^l w_l$, and also have the same index position. They are therefore called *covariant* vectors. It might be tempting to identify *contravariant* vectors $v \in V$ (column vectors, with upper indices, transforming like the dual basis \hat{e}^j) and covariant vectors $w \in V^{dual}$ (line vectors, with lower indices, transforming like the original basis) by transposition. Indeed this is possible in Euclidean space if we restrict ourselves to use orthonormal bases, because then the matrix G for the change of basis has to be orthogonal $G = G^{-1T}$ so that upper and lower indices transform in the same way. In other situations, like in the Minkowski space of special relativity

¹ Riemann integration would define a *pre-Hilbert space* or *inner product space*, whose Cauchy sequences need not converge. Such spaces can always be completed to Hilbert spaces similarly to the completion of \mathbb{Q} to \mathbb{R} .

(where the metric is not positive definite) or in quantum mechanics, where the inner product is *semi*-bilinear, it is important to distinguish between the two kinds of vectors.²

Dirac notation: Dirac introduced a very elegant and efficient notation for the use of linear algebra in quantum mechanics that is also called bra-ket notation because products are written by a bracket $\langle \dots \rangle$ as in eq. (3.8). We introduce bra-vectors $\langle w|$ and ket-vectors $|v\rangle$, which are just the co- and contravariant vectors $\langle w| \equiv w \in V^{dual}$ and $|v\rangle \equiv v \in V$, respectively. Their duality pairing can be written as a *bra-ket* product,

$$\langle w, v \rangle = w_i v^i = \langle w| \cdot |v\rangle \equiv \langle w|v\rangle. \quad (3.8)$$

The Dirac notation is basis independent. Instead of using vector components v^i with respect to some predefined basis e_i we will rather identify a state vector by specifying its physical properties, i.e. by the quantum numbers of the state of the physical system which it describes. For the energy eigenfunctions of the harmonic oscillator we can write, for example,

$$u_n(x) \equiv |E = \hbar\omega_0(n + \frac{1}{2})\rangle \equiv |E_n\rangle \equiv |n\rangle, \quad (3.9)$$

where it is sufficient to characterize the state by the number $n = 0, 1, \dots$ if it is clear from the context what quantum number we are referring to. The bra-ket notation is sufficiently flexible to allow us to write as much (or as little) information as we need. Note, however, that even a complete set of quantum numbers, which by definition uniquely defines the *physical* state of the quantum system, fixes the wave function only up to an overall phase. Bra- and ket-vectors, accordingly, are determined by the quantum numbers only up to a phase $|n'\rangle = e^{i\rho}|n\rangle$ and $\langle n'| = e^{-i\rho}\langle n|$. It is important not to change the implicit choice of that phase during the course of a calculation! Observable quantities will then be independent of such choices.

The inner product allows us to define a natural map from V to its dual by inserting an element v into the first position of the inner product. For $|v\rangle \in V$ the Hermitian conjugate vector $\langle v| \in V^{dual}$ is defined by

$$|v\rangle^\dagger \equiv \langle v| \in V^{dual} \quad \text{such that} \quad \langle v|u\rangle = (v, u) \quad \text{for all } u \in V. \quad (3.10)$$

Since the inner product is positive definite this conjugation is a bijective map from V to V^{dual} (this is also true for infinite-dimensional Hilbert spaces), but it is an “anti-isomorphism” and not an isomorphism because it is “anti-linear”

$$\boxed{(\alpha|v\rangle + \beta|w\rangle)^\dagger = \alpha^*\langle v| + \beta^*\langle w|.} \quad (3.11)$$

Linearity can be achieved by an additional complex conjugation so that V^{dual} is isomorphic to the complex conjugate space V^* , while V can be identified with its Hermitian conjugate

² In solid state physics the same distinction has to be made between the lattice Λ of atoms in a crystal and the reciprocal lattice Λ^{dual} of wave vectors; if Λ becomes finer then the reciprocal lattice becomes coarser.

$V \cong V^\dagger \equiv (V^{dual})^*$. We will henceforth use these identifications and the antilinear map $|v\rangle \rightarrow |v\rangle^\dagger = \langle v| \in V^{dual}$, which corresponds to the equation $\langle v, u \rangle \equiv \langle v|u\rangle = (v, u)$. For column vectors $|v\rangle$ the Hermitian conjugate is the line vector $\langle v|$ with complex conjugate entries. For wave functions $|\psi\rangle = \psi(x, t)$ it is the complex conjugate function $\langle\psi| = |\psi\rangle^\dagger = \psi^*(x, t)$.

3.2 Operator calculus

The components v^i of a vector v in an arbitrary basis can be obtained by evaluation of the dual basis $v^i = e^i(v) = \langle e^i, v \rangle$ because $e^i(v) = e^i(v^j e_j) = v^j e^i(e_j) = v^j \delta_j^i = v^i$. For an orthonormal basis $(e_i, e_j) = \delta_{ij}$ we observe that $|e_i\rangle^\dagger = \langle e_i| = \langle e^i|$, i.e. the Hermitian conjugate vector $|e_i\rangle^\dagger$ coincide with the dual basis $\langle e^i|$ and

$$|v\rangle = v^i e_i = |e_i\rangle \langle e^i| v \rangle = \sum_i |e_i\rangle \langle e_i| v \rangle, \quad (3.12)$$

where we have chosen, for later convenience, to use Einstein's summation convention only for contractions of upper and lower indices. Since the identity (3.12) holds for all v we get a representation of the unit matrix, or identity operator

$$\mathbb{1} = \sum_i |e_i\rangle \langle e_i| = \sum_i P_i \quad \text{with} \quad P_i = |e_i\rangle \langle e_i|. \quad (3.13)$$

Orthonormal bases are thus characterized by the two equations

$$\langle e_i|e_j\rangle = \delta_{ij} \quad \text{orthonormality,} \quad (3.14)$$

$$\sum_i |e_i\rangle \langle e_i| = \mathbb{1} \quad \text{completeness.} \quad (3.15)$$

P_i is the (orthogonal) **projector** onto the direction of the basis vector $|e_i\rangle$. As an example we consider the standard basis of \mathbb{C}^3 ,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.16)$$

The orthogonality relation reads

$$\langle e_1|e_1\rangle = (1, 0, 0) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1, \quad \langle e_1|e_2\rangle = (1, 0, 0) \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0, \quad \dots \quad (3.17)$$

and the projectors

$$|e_1\rangle \langle e_1| = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot (1, 0, 0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad |e_2\rangle \langle e_2| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad |e_3\rangle \langle e_3| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.18)$$

add up to the completeness relation

$$\sum_i |e_i\rangle\langle e_i| = \mathbb{1}. \quad (3.19)$$

While the product $\langle v|w\rangle$ of a covector and a vector yields a complex number, the tensor product $|w\rangle\langle v|$ is a matrix of rank 1 that is sometimes called *dyadic product*.

For a linear transformation $v \rightarrow Av$ the components A^i_j of the matrix representation $v^i \rightarrow A^i_j v^j$ can be obtained by sandwiching the operator A between basis elements. For an orthonormal basis $\langle e_i|e_j\rangle = \delta_{ij}$ we can use the Kronecker- δ to pull all indices down so that the entries (elements) of the matrix $A^i_j = e^i(Ae_j)$ in Dirac notation become

$$A_{ij} = \langle e_i|A|e_j\rangle. \quad (3.20)$$

In quantum mechanics the numbers $\langle v|A|w\rangle$ are hence called **matrix elements** even for arbitrary bra- and ket-vectors v and w . The normalized diagonal term

$$\langle A\rangle_v = \frac{\langle v|A|v\rangle}{\langle v|v\rangle} \quad (3.21)$$

is called expectation value of the operator A in the state $|v\rangle$, where the denominator can obviously be omitted if and only if $|v\rangle$ is normalized $\langle v|v\rangle = 1$.

Hermitian conjugation. If we apply a linear transformation $v \rightarrow Av$ to a vector v and evaluate a covector w , i.e. multiply with w from the left, the resulting number is

$$\langle w, Av\rangle = w_i A^i_j v^j = \langle w| \cdot A|v\rangle. \quad (3.22)$$

But we might just as well first perform the sum over i in $w_i A^i_j$ and then multiply the resulting bra-vector $\langle w|A$, with the ket-vector $|v\rangle$ from the right. In the language of linear algebra this defines the *transposed* map A^T on the dual space V^{dual} , which can be written as a matrix multiplication $w_j \rightarrow (A^T)_j^i w_i$ with the transposed matrix A^T . Using the non-degenerate inner product we can define the *Hermitian conjugate* A^\dagger of the linear operator A by

$$\boxed{(A^\dagger v, w) \equiv (v, Aw) \quad \forall v, w \in V.} \quad (3.23)$$

Using $(\varphi, \psi) = (\psi, \varphi)^*$ we obtain the matrix elements

$$\langle v|A|w\rangle = \langle A^\dagger v|w\rangle = (\langle w|A^\dagger v\rangle)^* \quad \Rightarrow \quad \langle w|A^\dagger|v\rangle = \langle v|A|w\rangle^*. \quad (3.24)$$

For an orthonormal basis $|e_i\rangle$ the components become

$$\boxed{(A^\dagger)_{ij} = \langle e_i|A^\dagger|e_j\rangle = \langle e_j|A|e_i\rangle^* = A_{ji}^*,} \quad (3.25)$$

so that Hermitian conjugation is transposition combined with complex conjugation of the matrix elements. Like transposition, Hermitian conjugation reverses the order of a product of matrices $(AB)^\dagger = B^\dagger A^\dagger$ and

$$(\alpha \langle \varphi | A \dots B | \psi \rangle)^* = (\alpha \langle \varphi | A \dots B | \psi \rangle)^\dagger = \alpha^* \langle \psi | B^\dagger \dots A^\dagger | \varphi \rangle \quad (3.26)$$

because Hermitian conjugation of a number is just complex conjugation.

An operator is called *self-adjoint* or *symmetric* or *Hermitian*³ if $A^\dagger = A$. Consider a normalized eigenvector $|a_i\rangle$ for the eigenvalue a_i of a **self-adjoint** operator

$$A|a_i\rangle = a_i|a_i\rangle \Rightarrow \langle a_i|A^\dagger = \langle a_i|a_i^*, \quad a_i = \langle a_i| \cdot (A|a_i\rangle) = \langle a_i| \cdot (A^\dagger|a_i\rangle) = (\langle a_i|A^\dagger) \cdot |a_i\rangle = a_i^*, \quad (3.27)$$

i.e. all **eigenvalues are real**, and hence

$$0 = \langle a_i|(A^\dagger - A)|a_j\rangle = \langle a_i|A^\dagger \cdot |a_j\rangle - \langle a_i| \cdot A|a_j\rangle = (a_i - a_j)\langle a_i|a_j\rangle \quad (3.28)$$

so that **eigenvectors** for different eigenvalues $a_i \neq a_j$ are **orthogonal** $\langle a_i|a_j\rangle = 0$.

Self-adjoint operators and spectral representation. The importance of self-adjoint operators $A = A^\dagger$ in quantum mechanics comes from the fact that they are exactly the operators for which all expectation values are real,⁴

$$(\langle \varphi | A | \psi \rangle)^* = \langle \psi | A^\dagger | \varphi \rangle = \langle \psi | A | \varphi \rangle \Rightarrow \langle \psi | A | \psi \rangle \in \mathbb{R}, \quad (3.32)$$

as we require for observable quantities. Hermitian matrices can be diagonalized and have real eigenvalues. Eigenvectors for different eigenvalues are orthogonal. In case of **degenerate eigenvalues**, i.e. eigenvalues with multiplicity greater than 1, a basis of eigenvectors for the

³ For infinite-dimensional Hilbert spaces there is a subtle difference between the definitions of symmetric and self-adjoint operators, respectively, because due to convergence issues an operator may only be defined on a dense subset of \mathcal{H} (see below). *Hermitian* is used synonymical with *symmetric* by most authors.

⁴ To see that Hermiticity is also necessary for real expectation values we bring A to Jordan normal form and assume that there is a non-trivial block $A = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$ with basis vectors $|e_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|e_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, i.e.

$$A|e_1\rangle = a|e_1\rangle, \quad A|e_2\rangle = |e_1\rangle + a|e_2\rangle. \quad (3.29)$$

Reality of $(a_i, Aa_i) = a_i(a_i, a_i) = a_i||a_i||^2$ for eigenvectors $|a_i\rangle$ implies reality of all eigenvalues a_i . For $|\psi\rangle = \alpha|e_1\rangle + \beta|e_2\rangle$ we find

$$\begin{aligned} \langle \psi | A | \psi \rangle &= (\alpha^* \langle e_1| + \beta^* \langle e_2|) ((\alpha a + \beta)|e_1\rangle + \beta a|e_2\rangle) \\ &= a (|\alpha|^2 ||e_1||^2 + |\beta|^2 ||e_2||^2 + \alpha^* \beta (e_1, e_2) + \alpha \beta^* (e_2, e_1)) + |\beta|^2 (e_2, e_1) + \alpha^* \beta ||e_1||^2 \end{aligned} \quad (3.30)$$

which cannot be real for all α if $\beta \neq 0$. Real expectation values hence imply diagonalizability. It remains to show that eigenvectors for different eigenvalues are orthogonal. We consider $|\varphi\rangle = \alpha|a_i\rangle + \beta|a_j\rangle$ and compute

$$\langle \varphi | A | \varphi \rangle = (\alpha^* \langle a_i| + \beta^* \langle a_j|) (a_i \alpha |a_i\rangle + a_j \beta |a_j\rangle) = \text{real} + a_j (\alpha^* \beta (a_i, a_j)) + a_i (\alpha^* \beta (a_i, a_j))^*, \quad (3.31)$$

which cannot be real for $a_i \neq a_j$ and arbitrary α, β unless $(a_i, a_j) = 0$. We conclude that a matrix A with real expectation values is diagonalizable with real eigenvalues and orthogonal eigenspaces, and hence is Hermitian.

respective eigenvalue can be **orthonormalized** by the **Gram-Schmidt** algorithm and the resulting vectors have to be distinguished by *additional “quantum numbers”* l_i in $|a_i, l_i\rangle$ with $l_i = 1, \dots, N_i$. The l_i have to be summed over in the completeness relation. For Hermitian matrices we thus can construct an orthonormal basis of eigenvectors $A|a_i\rangle = a_i|a_i\rangle$ with $\langle a_i|a_j\rangle = \delta_{ij}$, or, more precisely,

$$A|a_i, l_i\rangle = a_i|a_i, l_i\rangle \quad \text{with} \quad \langle a_i, l_i|a_j, k_j\rangle = \delta_{ij}\delta_{l_i k_j} \quad (3.33)$$

in the degenerate case. Using the completeness relation this implies the *spectral representation*

$$A = \sum_{i, l_i} A|a_i, l_i\rangle\langle a_i, l_i| = \sum_{i, l_i} a_i|a_i, l_i\rangle\langle a_i, l_i| = \sum_i a_i P_i, \quad (3.34)$$

where

$$P_i = \sum_{l_i=1}^{N_i} |a_i, l_i\rangle\langle a_i, l_i| \quad (3.35)$$

is the orthogonal projector onto the eigenspace for the eigenvalue a_i .

Unitary, traces and projection operators. We have seen that Hermitian matrices provide us with orthonormal bases of eigenvectors. A matrix U is called unitary if $U^\dagger U = UU^\dagger = \mathbb{1}$ or $U^\dagger = U^{-1}$. Different orthonormal bases $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$ are related by a unitary transformation $U_{ij} = \langle a_i|b_j\rangle$ because

$$|b_j\rangle = \left(\sum_i |a_i\rangle\langle a_i|\right) |b_j\rangle = \sum_i |a_i\rangle U_{ij}, \quad \Rightarrow \quad U_{ij}(U^\dagger)_{jk} = \sum_j \langle a_i|b_j\rangle \cdot \langle b_j|a_k\rangle = \langle a_i|\mathbb{1}|a_k\rangle = \delta_{ik}, \quad (3.36)$$

where we used the completeness relation. In other words, the inverse change of basis is given by $\langle b_j|a_k\rangle = \langle a_k|b_j\rangle^* = (U_{kj})^* = (U^\dagger)_{jk} = U_{jk}^{-1}$.

Projection operators in quantum mechanics are always meant to be orthogonal projections and they are characterized by the two conditions

$$P = P^\dagger \quad \text{and} \quad P^2 = P. \quad (3.37)$$

It follows from our previous considerations that projectors satisfy these equations. In turn, Hermiticity $P = P^\dagger$ implies the existence of a spectral representation $P_i = \sum_i \lambda_i |\lambda_i\rangle\langle \lambda_i|$ and $P^2 = P$ implies $\lambda_i^2 = \lambda_i$ so that all eigenvalues are either 0 or 1. Hence $P_i = \sum' |\lambda_i\rangle\langle \lambda_i|$ is a sum of projectors $|\lambda_i\rangle\langle \lambda_i|$ onto one-dimensional subspaces spanned by $|\lambda_i\rangle$ where the sum \sum' extends over the subset of basis vectors with eigenvalue 1. While the eigenvectors for a degenerate eigenvalue a_i of a matrix A in the spectral representation (3.34) are only defined up to a unitary change of basis of the respective eigenspace, the projector $P_i = \sum_{l_i} |a_i, l_i\rangle\langle a_i, l_i|$ onto such an eigenspace is independent of the choice of the orthonormal eigenvectors $|a_i, l_i\rangle$.

The axioms of quantum mechanics imply that measurements of the observable corresponding to a self-adjoint operator A yield the eigenvalue a_i with probability $\mathcal{P}(a_i) = |\langle \psi|a_i\rangle|^2$ if the state

of the system is described by the normalized vector $|\psi\rangle \in \mathcal{H}$. The resulting expectation value, i.e. the mean value $\langle A \rangle = \sum a_i \mathcal{P}(a_i)$ of the measured values weighted by their probabilities, is in accord with the definition (3.21) because the spectral representation of A implies

$$\langle \psi | A | \psi \rangle = \langle \psi | \sum_i a_i | a_i \rangle \langle a_i | \psi \rangle = \sum_i a_i |\langle \psi | a_i \rangle|^2. \quad (3.38)$$

The *trace* of a matrix is the sum of its diagonal elements and can be written as

$$\text{tr } A = \sum_i A_{ii} = \sum_i \langle e_i | A | e_i \rangle \quad (3.39)$$

for any orthonormal basis $|e_i\rangle$. An important property of traces is their invariance under cyclic permutations,

$$\text{tr}(AB) = \text{tr}(BA) \quad \Rightarrow \quad \text{tr}(A_1 A_2 \dots A_{r-1} A_r) = \text{tr}(A_r A_1 A_2 \dots A_{r-1}). \quad (3.40)$$

Probabilities and expectation values can be written in terms of traces and projection operators, which often simplifies calculations. Insertion of the definition $P_i = |a_i\rangle\langle a_i|$ shows that

$$\langle a_i | A | a_i \rangle = \text{tr}(P_i A) = \text{tr}(A P_i), \quad \Rightarrow \quad \mathcal{P}(a_i) = \langle \psi | P_i | \psi \rangle = \text{tr}(P_i P_\psi), \quad (3.41)$$

where $P_\psi = |\psi\rangle\langle\psi|$ projects onto the one-dimensional space spanned by the normalized state vector $|\psi\rangle$. The second formula $\mathcal{P}(a_i) = \text{tr}(P_i P_\psi)$ also holds for the probability of the measurement of the degenerate eigenvalue a_i if we use the projector $P_i = \sum_{l_i} |a_i, l_i\rangle\langle a_i, l_i|$ onto the complete eigenspace.

Commutators and anti-commutators. The commutator $[A, B]$ of two operators is defined as the difference between the two compositions $AB \equiv A \circ B$ and $BA \equiv B \circ A$,

$$\boxed{[A, B] = AB - BA \quad \Rightarrow \quad [A, B] = -[B, A].} \quad (3.42)$$

In the finite dimensional case it is just the difference between the matrix products AB and BA . We will often be in the situation that we know the commutators among a basic set A, B, \dots of operators, like the position operator $X_i = x_i$ and the momentum operator $P_i = \frac{\hbar}{i} \frac{\partial}{\partial x_i}$

$$\boxed{[X_i, P_j] = i\hbar\delta_{ij}.} \quad (3.43)$$

This can be verified by application to an arbitrary wave function

$$[X_i, P_j]|\psi\rangle = (X_i P_j - P_j X_i)\psi(x) = \frac{\hbar}{i}(x_i \partial_j \psi(x) - \partial_j(x_i \psi(x))) = -\frac{\hbar}{i}(\partial_j x_i)\psi(x) = i\hbar\delta_{ij}|\psi\rangle. \quad (3.44)$$

If we want to compute commutators for composite operators like the Hamilton operator $H = \frac{1}{2m}P^2 + \dots$ one should then always use the identities

$$\boxed{[A, BC] = [A, B]C + B[A, C], \quad [AB, C] = [A, C]B + A[B, C]} \quad (3.45)$$

rather than inserting and evaluating all the terms on a wave function and trying to recombine the result to an operator expression. (3.45) is easily verified by expanding the definitions

$$[A, BC] = ABC - BCA, \quad [A, B]C + B[A, C] = (AB - BA)C + B(AC - CA) \quad (3.46)$$

and similarly for $[AB, C]$. These identities can be memorized as the Leibniz rule for the action of $[A, *]$ on a product BC and a similar product rule for the action of $[*, C]$ on the product AB “from the right”. This “Leibniz rule” also holds for the action of $[A, *]$ on a commutator $[B, C]$ and for the action of $[*, C]$ on $[A, B]$

$$[A, [B, C]] = [[A, B], C] + [B, [A, C]], \quad [[A, B], C] = [[A, C], B] + [A, [B, C]]. \quad (3.47)$$

Each of these equations is equivalent to the *Jacobi identity*

$$\boxed{[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0,} \quad (3.48)$$

which states the sum over the cyclic permutations of ABC in a double commutator is zero. This is again easily verified by expanding all terms

$$A(BC - CB) - (BC - CB)A + B(CA - AC) - (CA - AC)B + C(AB - BA) - (AB - BA)C = 0. \quad (3.49)$$

The equivalence of the “product rule” (3.47) with the Jacobi identity follows from the antisymmetry of the commutator $[A, B] = -[B, A]$.

Similarly to the commutator we can define the anti-commutator

$$\boxed{\{A, B\} = AB + BA \quad \Rightarrow \{A, B\} = \{B, A\}.} \quad (3.50)$$

For two Hermitian operators $A = A^\dagger$ and $B = B^\dagger$ the commutator is anti-Hermitian and the anti-commutator is Hermitian,

$$[A, B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A, B], \quad (3.51)$$

$$\{A, B\}^\dagger = (AB + BA)^\dagger = B^\dagger A^\dagger + A^\dagger B^\dagger = BA + AB = \{A, B\}. \quad (3.52)$$

Since iC is Hermitian if C is anti-Hermitian the decomposition

$$AB = \frac{1}{2}(AB + BA) + \frac{1}{2}(AB - BA) = \frac{1}{2}\{A, B\} + \frac{1}{2}[A, B] \quad (3.53)$$

of an operator product AB as a sum of a commutator and an anti-commutator corresponds to a decomposition into real and imaginary part for products of Hermitian operators.

Complete systems of commuting operators. We show that two self-adjoint operators A and B commute $AB = BA$ if and only if they can be diagonalized simultaneously. Since

diagonal matrices commute, it is clear that $[A, B] = 0$ if there exists a basis such that both operators are diagonal. In order to prove the “only if” we assume that $[A, B] = 0$ and that A has been diagonalized. Then B must be block-diagonal because

$$0 = \langle a_i | [A, B] | a_j \rangle = \langle a_i | AB - BA | a_j \rangle = (a_i - a_j) \langle a_i | B | a_j \rangle \quad (3.54)$$

so that all matrix elements of B between states with different eigenvalues of A vanish. B can now be diagonalized within each block, by a change of basis that does not mix eigenstates for different eigenvalues of A and hence does not spoil the diagonalization of A . It is clear from the proof that the proposition extends to an arbitrary number of mutually commuting operators. Moreover, we see that any set of mutually commuting operators can be extended to a complete set in the sense that the simultaneous diagonalization uniquely fixes the common normalized eigenvectors up to a phase (just add an operator that lifts the remaining degeneracies within the common eigenspaces of the original set). The set of all eigenvalues (a_i, b_j, c_k, \dots) of such a complete system A, B, C, \dots thus completely characterizes the state $|a_i, b_j, c_k, \dots\rangle$ of a quantum system.

Functions of operators. If we consider the position vector \vec{x} of a particle as a vector of operators \vec{X} then the potential $V(X) = V(x)$ can be a complicated function of operators X_i . If such a function is analytic $f(x) = \sum_{n=0}^{\infty} c_n x^n$ then the corresponding function of operators can be defined by the power series expansion

$$f(x) = \sum c_n x^n \quad \Rightarrow \quad f(A) = \sum c_n A^n. \quad (3.55)$$

For matrices the series always converges if the radius r of convergence of the Taylor series is infinite. If $0 < r < \infty$ then $f(\mathcal{O})$ can be defined by analytic continuation of its matrix elements. Of particular importance is the exponential function

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} A\right)^n, \quad (3.56)$$

which usually appears if we are interested in the finite form of infinitesimal transformations. For example, the infinitesimal time evolution of the wave function is given by the Schrödinger equation

$$\partial_t |\psi(x, t)\rangle = \frac{1}{i\hbar} H |\psi(x, t)\rangle \quad \Rightarrow \quad |\psi(x, t_0 + \delta t)\rangle = \left(1 + \frac{\delta t}{i\hbar} H + O(\delta t^2)\right) |\psi(x, t_0)\rangle \quad (3.57)$$

For a *time-independent Hamiltonian* H we obtain, after n infinitesimal time steps $\delta t = (t - t_0)/n$ with $n \rightarrow \infty$,

$$|\psi(x, t)\rangle = U(t - t_0) |\psi(x, t_0)\rangle, \quad U(t - t_0) = e^{-\frac{i}{\hbar}(t-t_0)H}. \quad (3.58)$$

$U(t)$ is called **time evolution operator**. It is, actually, a one-parameter family of operators satisfying $U(t_1)U(t_2) = U(t_1 + t_2)$ and $\partial_t U(t) = -\frac{i}{\hbar}HU(t)$ with $U(0) = \mathbb{1}$.

For operators A, B the product of exponentials is not the exponential of the sum if the operators do not commute. The correction terms are expressed by the **Baker–Campbell–Hausdorff** formula

$$\boxed{e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}([A,[A,B]]-[B,[A,B]])+\text{multiple commutators}}} \quad (3.59)$$

(for a proof consider example (1.21) in [Grau]). In many applications the double commutators $[A, [A, B]]$ and $[B, [A, B]]$ vanish or are proportional to $\mathbb{1}$ so that the series terminates after a few terms. In particular, since A and $-A$ commute, the exponential of an anti-Hermitian operator iA is unitary,

$$A = A^\dagger \quad \Rightarrow \quad (e^{iA})^\dagger = e^{-iA} = (e^{iA})^{-1}. \quad (3.60)$$

The Hamilton operator of a quantum system has to be self-adjoint because it corresponds to the energy, which is an observable.⁵ Time evolution is hence described by a unitary transformation $U(t) = U^\dagger(-t)$. We have already checked this in chapter 2 by showing that $\langle \psi | \psi \rangle$ is preserved under time evolution for a nonrelativistic electron in an electromagnetic field. But the present discussion is more general. Another important formula

$$\boxed{e^{\lambda A} B e^{-\lambda A} = B + \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} [A, B]_{(n)} = B + \lambda [A, B] + \frac{\lambda^2}{2} [A, [A, B]] + \dots} \quad (3.61)$$

with $[A, B]_1 = [A, B]$ and $[A, B]_{(n+1)} = [A, [A, B]_{(n)}]$ describes the “conjugation” UBU^{-1} of an operator B by the exponential $U = e^{\lambda A}$ of λA .⁶

Arbitrary functions of *Hermitian* operators can be defined via their spectral representation,

$$\boxed{A = A^\dagger = \sum a_i |a_i\rangle\langle a_i| \quad \Rightarrow \quad f(A) = \sum f(a_i) |a_i\rangle\langle a_i|} \quad (3.62)$$

For analytic functions f this coincides with the power series (3.55), as is easily checked in a basis where A is diagonal. The definition (3.55) only makes sense for analytic functions, but it has the advantage that it does not require diagonalizability. With (3.62), on the other hand, even the Heaviside step function $\theta(A)$ of an operator A makes sense.

Tensor products: If we have a quantum system that is composed of two subsystems, whose states are described by $|i\rangle \in V_1$ with $i = 1, \dots, I$ and $|m\rangle \in V_2$ with $1 \leq m \leq M$, then

⁵ In certain contexts, like the description of particle decay, it may nevertheless be useful to consider Hamilton operators with an imaginary part.

⁶ Conjugation of operators corresponds to a change of orthonormal bases $|e\rangle \rightarrow U|e\rangle$, for which the dual basis transforms as $\langle e| \rightarrow U^\dagger\langle e| = U^{-1}\langle e|$.

the states of the composite systems are superpositions of arbitrary combinations

$$|i, m\rangle \equiv |i\rangle \otimes |m\rangle, \quad 1 \leq i \leq I, \quad 1 \leq m \leq M \quad (3.63)$$

of the independent states in the subsystems. The vector space $V_1 \otimes V_2$ describing the composite system is called tensor product and it consists of linear combinations

$$|w\rangle = \sum_{i=1}^I \sum_{m=1}^M w_{im} |i, m\rangle \in V_1 \otimes V_2 \quad (3.64)$$

with an arbitrary matrix w_{im} of coefficients. Its dimension $\dim(V_1 \otimes V_2) = I \cdot M$ is the product of the dimensions of the factors V_1 and V_2 . The Dirac notation is particularly useful for such composite systems because we just combine the respective quantum numbers into a longer *ket*-vector. It is a simple fact of linear algebra that generic vectors in a tensor product cannot be written as a product

$$|w\rangle = \sum w_{im} |i, m\rangle \neq |v_1\rangle \otimes |v_2\rangle \quad (3.65)$$

for any $|v_1\rangle = \sum c_i |i\rangle$ and $|v_2\rangle = \sum d_m |m\rangle$ because this is only possible if the coefficient matrix factorizes as $w_{im} = c_i d_m$ and hence has rank 1. In quantum mechanics non-product states like (3.65) are often called *entangled states*. They play an important role in discussions about the interpretation of quantum mechanics like in the EPR paradoxon (see below).

The inner product on the tensor product space is defined by

$$\langle i, m | j, n \rangle = \langle i | j \rangle \cdot \langle m | n \rangle \quad (3.66)$$

for product states and extended by semi-bilinearity to $V_1 \otimes V_2$. In the product basis $|i, m\rangle$ operators on a tensor product space also have double-indices

$$|i, m\rangle \rightarrow \mathcal{O}_{i,m;j,n} |j, n\rangle \quad (3.67)$$

Such operators will often correspond to the combined action of some operator $\mathcal{O}^{(1)}$ on V_1 and $\mathcal{O}^{(2)}$ on V_2 , like for example the rotation of the position vector of the first particle and the simultaneous rotation of the position vector of the second particle for rotating the complete system. In that situation the trace of the product operator factorizes into a product of traces

$$\mathcal{O}_{i,m;j,n} = \mathcal{O}_{ij}^{(1)} \otimes \mathcal{O}_{mn}^{(2)} \Rightarrow \text{tr } \mathcal{O}_{i,m;j,n} = \sum_{im} \mathcal{O}_{i,m;i,m} = \sum_i \mathcal{O}_{ii}^{(1)} \sum_m \mathcal{O}_{mm}^{(2)} = \text{tr } \mathcal{O}^{(1)} \cdot \text{tr } \mathcal{O}^{(2)}. \quad (3.68)$$

As an example consider $\mathcal{O}^{(1)} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\mathcal{O}^{(2)} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ for $V_1 = V_2 = \mathbb{C}^2$. In the basis $e_1 = |11\rangle$, $e_2 = |12\rangle$, $e_3 = |21\rangle$ and $e_4 = |22\rangle$ of the product space the product operator corresponds to the insertion of the second matrix into the first,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a\mathcal{O}^{(2)} & b\mathcal{O}^{(2)} \\ c\mathcal{O}^{(2)} & d\mathcal{O}^{(2)} \end{pmatrix} = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}. \quad (3.69)$$

The Dirac notation is obviously more transparent than this. It is easy to verify (3.68) for (3.69).

3.3 Operators and Hilbert spaces

Recall that the normalizable solutions $\psi(x, t)$ of the Schrödinger equation form an inner product space, i.e. a vector space with a positive definite semi-bilinear product

$$\langle \psi | \varphi \rangle = \int d^3x \psi^*(x, t) \varphi(x, t). \quad (3.70)$$

Inner product spaces are also called pre-Hilbert spaces. Such a space is called Hilbert space if it is complete with respect to the norm

$$\|\psi\| = \sqrt{\langle \psi | \psi \rangle}, \quad (3.71)$$

i.e. if every Cauchy sequence converges. Cauchy sequences are sequences ψ_n with the property that for every positive number ε there exists an integer $N(\varepsilon)$ with

$$\|\psi_m - \psi_n\| < \varepsilon \quad \forall m, n > N(\varepsilon). \quad (3.72)$$

Pre-Hilbert spaces can be turned into Hilbert spaces by a standard procedure called completion, which amounts to adding the missing limits. The standard Hilbert space of quantum mechanics is the space of square integrable functions called

$$L^2(\mathbb{R}^3). \quad (3.73)$$

The letter L stands for *Lesbeques integration*, which has to be used because Riemann's definition of integration only works for a restricted class of square-integrable functions $\int |\psi|^2 < \infty$ that is not complete and the Lesbeques integral can be regarded as the result of the completion.⁷ A Hilbert space basis is a set of vectors $|e_i\rangle$ with some (possibly not countable) index set I such that every vector $|\psi\rangle \in \mathcal{H}$ can be written as a convergent infinite sum

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n |e_{i_n}\rangle \quad (3.74)$$

for a sequence c_n of coefficients and a sequence i_n of indices $i \in I$ and hence of basis vectors $|e_{i_n}\rangle$ taken from the complete set of basis elements $|e_i\rangle$. A Hilbert space is called separable if there exists a countable basis, i.e. if we can take the index set to be $I = \mathbb{N}$. All Hilbert spaces that we need in this lecture will be separable.

⁷ For example, the integral of the function $I_{\mathbb{Q}}(x)$ that is 1 for rational numbers and 0 for irrational numbers x is 0 for Lesbeques integration, because \mathbb{Q} is countable so that $I_{\mathbb{Q}}$ is the limit of a Cauchy sequence of function with only finitely many values different from 0. But the Riemann integral does not exist.

3.3.1 Inequalities

In this section we derive three inequalities that hold in any Hilbert space. Let us denote the vectors as $f, g, h, \dots \in \mathcal{H}$. The orthogonal projection of f onto g is the vector $|g\rangle \frac{\langle g|f\rangle}{\langle g|g\rangle}$ with the projection vector

$$|h\rangle = |f\rangle - |g\rangle \frac{\langle g|f\rangle}{\langle g|g\rangle} \quad (3.75)$$

orthogonal to $|g\rangle$ since $\langle g|h\rangle = 0$. Now we use the defining equation of $|h\rangle$ to obtain the Pythagorean theorem

$$\|f\|^2 = \langle f|f\rangle = \left(\langle h| + \langle g| \frac{\langle g|f\rangle}{\langle g|g\rangle} \right) \left(|h\rangle + |g\rangle \frac{\langle g|f\rangle}{\langle g|g\rangle} \right) = \|h\|^2 + \frac{|\langle g|f\rangle|^2}{\langle g|g\rangle} \quad (3.76)$$

Since $\|h\|^2 \geq 0$ we see that:

$$\|f\|^2 \geq \frac{|\langle g|f\rangle|^2}{\|g\|^2} \quad (3.77)$$

and we obtain the **Schwartz inequality**

$$\|f\| \|g\| \geq |\langle g|f\rangle| \quad (3.78)$$

which will later be used in the derivation of Heisenberg's uncertainty relation.

More generally, we can consider a set g_1, \dots, g_n of orthonormal vectors $\langle g_i, g_j\rangle = \delta_{ij}$ and write f as a sum of orthogonal projections $|g_i\rangle \langle g_i|f\rangle$ and the difference vector

$$|h\rangle = |f\rangle - \sum_{i=1}^n |g_i\rangle \langle g_i|f\rangle, \quad (3.79)$$

which is orthogonal to the linear subspace spanned by the $|g_i\rangle$. The Pythagorean theorem thus becomes

$$\|f\|^2 = \sum_{i=1}^n |\langle g_i|f\rangle|^2 + \|h\|^2 \quad (3.80)$$

and the **Bessel inequality**

$$\|f\|^2 \geq \sum_{i=1}^n |\langle g_i|f\rangle|^2 \quad (3.81)$$

follows from positivity of $\|h\|^2$. For a Hilbert space basis $g_i, i \in \mathbb{N}$ the norm of h thus has to converge to 0 monotonously from above for $n \rightarrow \infty$.

The norm of $|f\rangle + |g\rangle$ is

$$\|f + g\|^2 = \langle f + g|f + g\rangle = \|f\|^2 + \|g\|^2 + \langle f|g\rangle + \langle g|f\rangle \quad (3.82)$$

Since we can write the last two terms as

$$\langle f|g\rangle + \langle g|f\rangle = \langle f|g\rangle + (\langle f|g\rangle)^* = 2\text{Re}\langle f|g\rangle \leq 2|\text{Re}\langle f|g\rangle| \leq 2|\langle f|g\rangle| \quad (3.83)$$

we can use the Schwartz inequality in this relation and obtain

$$\|f + g\|^2 \leq \|f\|^2 + \|g\|^2 + 2\|f\|\|g\| \quad (3.84)$$

whose square root yields the **triangle inequality**

$$\|f + g\| \leq \|f\| + \|g\|, \quad (3.85)$$

which shows that the definition (3.71) of the norm in inner product spaces makes sense.

3.3.2 Position and momentum representations

As compared to matrices in finite-dimensional vector spaces we will encounter two kinds of complications for operators in Hilbert spaces. Consider, for example, the Hamilton operator for the potential well. For negative energies we obtained a discrete spectrum of bound states. But for free electrons there are no normalizable energy eigenstates and normalizable wave packets are superpositions of states with a continuum of energy values. Hence, the spectrum of self-adjoint operators will, in general, consist of a discrete part and a continuum without *normalizable* eigenstates. Moreover, the eigenvalues may not even be bounded, which leads to additional complications.

As an example we first consider the momentum operator $P = \frac{\hbar}{i} \frac{\partial}{\partial x}$. Working, for simplicity, in one dimension we define

$$|p_x\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p x}, \quad P|p_x\rangle = p|p_x\rangle, \quad (3.86)$$

where the argument x of the wave function is indicated as a subscript of the eigenvalue p if necessary. The normalization of the momentum eigenstates $|p\rangle$ have been chosen such that

$$\langle p'|p\rangle = \frac{1}{2\pi\hbar} \int dx e^{\frac{i}{\hbar}(p-p')x} = \delta(p' - p), \quad (3.87)$$

where we used $\int dx e^{ikx} = 2\pi\delta(k)$. In three dimensions $|\vec{p}\rangle = |p_1\rangle \otimes |p_2\rangle \otimes |p_3\rangle$ so that

$$|\vec{p}_x\rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i}{\hbar} \vec{p} \cdot \vec{x}} \quad \text{and} \quad \langle \vec{p}' | \vec{p} \rangle = \delta^3(\vec{p}' - \vec{p}). \quad (3.88)$$

The product

$$\langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dx e^{-\frac{i}{\hbar} p x} \psi(x) = \tilde{\psi}(p) \quad (3.89)$$

yields the Fourier transform⁸ of the wave function and the validity of the formula for the Fourier representation

$$\int dp |p_x\rangle \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dp e^{+\frac{i}{\hbar} p x} \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dp e^{+\frac{i}{\hbar} p x} \tilde{\psi}(p) = \psi(x), \quad (3.90)$$

⁸ The extra factor $1/\sqrt{\hbar}$ in the normalization, as compared to the conventions in section 2, is due the rescaled argument $p = \hbar k$ of the Fourier transform.

which holds for all $\psi \in L^2(\mathbb{R})$, implies the spectral representation

$$\int dp |p_x\rangle\langle p_{x'}| = \mathbb{1}_{x,x'} = \delta(x - x'), \quad (3.91)$$

but now with the sum over eigenvalues with normalizable eigenstates replaced by an integral over the continuum of eigenvalues with non-normalizable eigenfunctions. The spectral representation thus becomes

$$P = P\mathbb{1} = \int dp P|p\rangle\langle p| = \int dp p |p\rangle\langle p|. \quad (3.92)$$

For more general self-adjoint operators like the Hamilton operator of the potential well we hence anticipate a spectral representation that will combine a sum over bound state energies with an integral over continuum states.

Similarly to the momentum operator we can now introduce a basis of eigenstates for the position operator X , where we would like to have

$$X|x\rangle = x|x\rangle \quad \text{with} \quad \int dx |x\rangle\langle x| = \mathbb{1} \quad \text{and} \quad \langle x|x'\rangle = \delta(x - x'). \quad (3.93)$$

But what are the wave functions $\psi_x(x')$ corresponding to these states? Since $X|x\rangle = x|x\rangle$ the wave function $\psi_x(x')$ of the state $|x_{x'}\rangle$ should vanish for $x' \neq x$ and hence be proportional to $\delta(x' - x)$, i.e. $\psi_x(x') = c\delta(x' - x)$. This ansatz satisfies (3.93) if we choose the prefactor $c = 1$ so that $\psi_x(x') = \langle x|x'\rangle$. This should not come as a surprise if we recall from section 3.1 that we can obtain the components v^i of a vector $v = v^i e_i$ by evaluation of the dual basis vectors $v^i = e^i(v)$ and that the bra-vectors obtained by Hermitian conjugation of an orthonormal basis provide the dual basis. Hence, for an arbitrary state $|\psi\rangle \in \mathcal{H}$ the products

$$\psi(x) = \langle x|\psi\rangle, \quad \psi(p) = \langle p|\psi\rangle \quad (3.94)$$

are the **wave functions** $\psi(x)$ in **position space** and $\psi(p)$ in **momentum space**, respectively.

We hence **regard** $|\psi\rangle \in \mathcal{H}$ **as an abstract vector in Hilbert space** independently of any choice of basis and write $\langle x|\psi\rangle = \psi(x)$ for the position space wave function and $\langle p|\psi\rangle = \psi(p)$ for the wave function in the momentum space basis $|p\rangle$. The “unitary matrix” $\langle x|p\rangle$ for the change of basis from position space to momentum space $|p\rangle = \int dx |x\rangle\langle x|p\rangle$ and its inverse $\langle p|x\rangle$ are

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{ipx}{\hbar}}, \quad \langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{ipx}{\hbar}}. \quad (3.95)$$

Since the spectra of eigenvalues of P and X and the corresponding “matrix indices” p and x are continuous, matrix multiplication amounts to integration and the change of basis becomes

$$\psi(p) = \langle p|\psi\rangle = \langle p|\mathbb{1}|\psi\rangle = \langle p| \left(\int dx |x\rangle\langle x| \right) |\psi\rangle = \int dx \langle p|x\rangle \psi(x), \quad (3.96)$$

which is nothing but a Fourier transformation.

The basis independence of the integrated probability density $||\psi||^2 = \langle \psi | \psi \rangle$

$$\begin{aligned} \int dx |\psi(x)|^2 &= \int dx \langle \psi | x \rangle \langle x | \psi \rangle = \int dx \langle \psi | \int dp |p\rangle \langle p| x \rangle \langle x | \int dp' |p'\rangle \langle p'| \psi \rangle \\ &= \int dp \int dp' \langle \psi | p \rangle \int dx \langle p | x \rangle \langle x | p' \rangle \langle p' | \psi \rangle = \int dp \langle \psi | p \rangle \langle p | \psi \rangle = \int dp |\psi(p)|^2 \end{aligned} \quad (3.97)$$

expresses the unitarity $\int dp \langle p | x \rangle \langle x | p' \rangle = \delta(p - p')$ of the matrix $\langle x | p \rangle$ of the change of basis. In Fourier analysis (3.97) is called Parseval's equation.

The matrix elements of X and P are now easily evaluated in position space

$$\langle x' | X | x \rangle = x \delta(x - x'), \quad \langle x' | P | x \rangle = \frac{\hbar}{i} \frac{\partial}{\partial x} \delta(x - x'), \quad (3.98)$$

and in momentum space

$$\langle p' | P | p \rangle = p \delta(p - p'), \quad \langle p' | X | p \rangle = -\frac{\hbar}{i} \frac{\partial}{\partial p} \delta(p - p'), \quad (3.99)$$

which shows that $X = -\frac{\hbar}{i} \frac{\partial}{\partial p}$ and $P = p$ in momentum space. The generalizations of these formulas to three dimensions are obvious.

3.3.3 Convergence, norms and spectra of Hilbert space operators

Having gained some intuition about spectra and eigenbases of Hilbert space operators we are now turning to general definitions and results. Already for the case of a discrete spectrum, like in the Harmonic oscillator for which electrons are always bound, it is clear that the spectral decomposition of the identity

$$\hat{1} = \lim_{n \rightarrow \infty} \sum_{i=1}^n |e_i\rangle \langle e_i| \quad (3.100)$$

requires some notion of convergence for sequences of operators in order to be able to define infinite sums as limits of finite sums.

For sequences of Hilbert space *vectors* there are, in fact, two different notions of convergence: The obvious one, which we used for the definition of completeness, is called *strong convergence*:

$$|\psi_n\rangle \longrightarrow |\psi\rangle \quad \text{if} \quad \lim_{n \rightarrow \infty} ||\psi_n - \psi|| = 0 \quad \text{strong limit.} \quad (3.101)$$

A second notion of convergence, which is called *weak* because it is always implied by strong convergence (see section 4.8 of [Kreyszig]), only requires that all products with bra-vectors converge:

$$|\psi_n\rangle \xrightarrow{\text{weak}} |\psi\rangle \quad \text{if} \quad \lim_{n \rightarrow \infty} \langle \varphi | \psi_n \rangle = \langle \varphi | \psi \rangle \quad \forall \langle \varphi | \in \mathcal{H}^{\text{dual}} \quad \text{weak limit.} \quad (3.102)$$

An example of a sequence that weakly converges to 0 but that is divergent in the strong sense is the sequence $\{e_n\}$ of basis vectors of a Hilbert space basis: A sequence pointing into the infinitely many directions of a Hilbert space with constant length 1 does not converge (in the norm) because the distance $\|e_n - e_m\|$ between any two elements of such a sequence is always $\sqrt{2}$. But the scalar products $\langle \varphi | e_n \rangle$, which are the expansion coefficients of $\langle \varphi |$ in the basis $\{e_n\}$, have to converge to 0 because of Bessels inequality.

Convergence of operators: For us, Hilbert space operators are always meant to be linear

$$\mathcal{O}(\alpha|\varphi\rangle + \beta|\psi\rangle) = \alpha\mathcal{O}|\varphi\rangle + \beta\mathcal{O}|\psi\rangle \in \mathcal{H}. \quad (3.103)$$

These operators are important in quantum mechanics because they correspond to observables. We now have two options: Every real measurement has a bounded set of possible results. For example, we can never measure the position of a particle, say, behind the Andromeda nebula, because our particle detector has a finite size. We could hence simply restrict the concept of an observable to bounded operators, which are quite well-behaved. But, like for wave packets and plane waves, it is much more convenient to work with unbounded operators like the position operator X rather than with more realistic approximations of this operators.

We hence first define the concept of the norm of an operator, which we can think of as the modulus $|\lambda|$ of the largest eigenvalue λ :

$$\|\mathcal{O}\| = \sup_{\psi \neq 0} \left(\frac{\|\mathcal{O}\psi\|}{\|\psi\|} \right). \quad (3.104)$$

In this definition we have to use the supremum instead of the maximum because in the infinite-dimensional case the maximum may not exist and the supremum (which is the smallest upper bound) may be infinite $0 \leq \|\mathcal{O}\| \leq \infty$. Considering a sequence ψ_n of localized waves packets for electrons whose distance from the earth increases, say, linearly with n , it is clear that X is unbounded $\|X\| = \infty$, and similarly one can show that the momentum P is also unbounded.

An operator is called bounded if $\|\mathcal{O}\| < \infty$. Bounded operators $\mathcal{O} : V \rightarrow W$ can, in fact, be defined for any normed spaces V and W . For two such operators we can consider linear combinations defined by

$$(\alpha\mathcal{O}_1 + \beta\mathcal{O}_2)|\psi\rangle = \alpha\mathcal{O}_1|\psi\rangle + \beta\mathcal{O}_2|\psi\rangle \in W \text{ for } |\psi\rangle \in V, \quad (3.105)$$

so that the set of all bounded operators $\mathcal{B}(V, W)$ again forms a vector space. With the operator norm defined by (3.104) the normed space $\mathcal{B}(V, W)$ is complete and hence a Banach space. In this statement we refer to the strongest notion of convergence, which is called *uniform convergence* or *convergence in the norm*. For operators there are, however, even two different weaker notions of convergence: A **sequence of operators** $O_n : V \rightarrow W$ is said to be:

- **uniformly convergent** if (\mathcal{O}_n) converges in the norm of \mathcal{B} , i.e.

$$\lim_{n \rightarrow \infty} \|\mathcal{O}_n - \mathcal{O}\| = 0, \quad (3.106)$$

- **strongly convergent** if $(\mathcal{O}_n\psi)$ converges strongly in W for every $\psi \in V$, i.e.

$$\lim_{n \rightarrow \infty} \|\mathcal{O}_n\psi - \mathcal{O}\psi\| = 0 \quad \forall |\psi\rangle \in V \quad (3.107)$$

- **weakly convergent** if $(\mathcal{O}_n\psi)$ converges weakly in W for every $\psi \in V$, i.e.

$$\lim_{n \rightarrow \infty} |\langle \phi | \mathcal{O}_n\psi \rangle - \langle \phi | \mathcal{O}\psi \rangle| = 0 \quad \forall |\psi\rangle \in V \text{ and } \forall \langle \phi | \in W^{dual}, \quad (3.108)$$

where \mathcal{O} denotes the limit operator $\mathcal{O} : V \rightarrow W$. The notions of strong and weak operator convergence make perfect sense also for unbounded operators, and, moreover, $\mathcal{O}_n - \mathcal{O}$ may be bounded and uniformly convergent even if the operators \mathcal{O}_n and \mathcal{O} are unbounded.

Spectra and resolvents of operators. Naively we think of the spectrum of an operator as the set of eigenvalues λ of the matrix A , which coincides with the values λ for which

$$A_\lambda = A - \lambda \mathbb{1} \quad (3.109)$$

is not invertible so that $\det A_\lambda = 0$. In that case there exists an eigenvector $|a_\lambda\rangle$ with

$$A_\lambda |a_\lambda\rangle = 0 \quad \Leftrightarrow \quad A |a_\lambda\rangle = \lambda |a_\lambda\rangle. \quad (3.110)$$

The generalization to infinite dimensions is based on this fact and defines the spectrum as the set of complex numbers $\lambda \in \mathbb{C}$ for which A_λ is not invertible.

We have to take into account one further complication: For unbounded operators A it may happen that they are only defined on a subset of the Hilbert space vectors. As an example consider the position operator X and the wave function $\psi(x) = \theta(x)/\sqrt{1+x^2}$ where the step function $\theta(x)$ is 1 for $x > 0$ and 0 for $x < 0$. The integral $\int |\psi|^2 = \int_0^\infty \frac{dx}{(1+x^2)} = \frac{\pi}{2}$ exists, but $\langle \psi | X | \psi \rangle = \int_0^\infty \frac{x dx}{1+x^2}$ diverges. Hence $x\psi(x) \notin \mathcal{H} = L^2(\mathbb{R})$ and we have to restrict the domain of definition of X to a subset $\mathcal{D}_X \subset \mathcal{H}$ if we want X to be an operator with values in \mathcal{H} .

We hence consider operators $A : \mathcal{D}_A \rightarrow \mathcal{H}$ with domain of definition $\mathcal{D}_A \subseteq \mathcal{H}$ and assume that \mathcal{D}_A is dense in \mathcal{H} , which means that every vector $|\psi\rangle \in \mathcal{H}$ can be obtained as a limit of a sequence $\psi_n \in \mathcal{D}_A$ of vectors in the domain of definition.⁹ We now define the **resolvent** R_λ , if it exists, as the inverse of $A_\lambda = A - \lambda \mathbb{1}$, i.e.

$$R_\lambda = A_\lambda^{-1} \quad (3.111)$$

⁹ For the position operator X we can take the sequence $\psi_n(x) = \theta(n - |x|) \cdot \psi(x)$.

with $R_\lambda A_\lambda = \mathbb{1}$ on \mathcal{D}_A . The resolvent R_λ hence is a linear operator from the range of A_λ to the domain of A_λ . It does not exist if and only if there exists a vector $|\psi\rangle \in \mathcal{D}_A$ with $A_\lambda|\psi\rangle = 0$. In that case $|\psi\rangle$ is an eigenvector of A with eigenvalue λ .

A complex number $\lambda \in \mathbb{C}$ is called a **regular value** if the resolvent R_λ exists as a bounded operator and λ is called **spectral value** otherwise. The set of regular values is called **resolvent set** $\rho(A) \subset \mathbb{C}$ and its complement $\sigma(A) = \mathbb{C} - \rho(A)$ is called **spectrum** of the operator A . The spectrum $\sigma(A)$ consists of three disjoint parts:

- The **point spectrum** or **discrete spectrum** $\sigma_p(A)$ is the set of values λ such that R_λ does not exist. $\sigma_p(A)$ is the set of **eigenvalues** of A (with normalizable eigenstates; this corresponds to the bound state energies for the Hamilton operator).
- The **continuous spectrum** $\sigma_c(A)$ is the set of values λ such that R_λ exists and is defined on a set which is dense in \mathcal{H} , but is not bounded (for the Hamilton operator this corresponds to the energies of scattering states).
- The **residual spectrum** $\sigma_r(A)$ is the set of λ such that R_λ exists but the domain of definition is not dense in \mathcal{H} .

We thus obtain a decomposition of the complex plane as a disjoint union of four sets $\mathbb{C} = \rho(A) \cup \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$. From the definition it follows that the resolvent set is open and one can show that the resolvent R_λ is an (operator valued) holomorphic function on $\rho(A)$, so that methods of complex analysis can be used in spectral theory [Reed]. In finite dimensional cases the spectrum of a linear operator is a pure point spectrum, i.e. $\sigma_c(A) = \sigma_r(A) = \emptyset$. For self-adjoint operators it can be shown that the residual spectrum is empty $\sigma_r(A) = \emptyset$.

3.3.4 Self-adjoint operators and spectral representation

A densely defined Hilbert space operator A is called **symmetric** (or **Hermitian**) if its domain of definition is contained in the domain of definition of the adjoint operator¹⁰

$$\mathcal{D}_A \subseteq \mathcal{D}_{A^\dagger} \quad \text{and} \quad \langle A\varphi|\psi\rangle = \langle\varphi|A\psi\rangle \quad \forall\varphi, \psi \in \mathcal{D}_A. \quad (3.112)$$

A symmetric operator is called **self-adjoint** if $\mathcal{D}_A = \mathcal{D}_{A^\dagger}$.

The difference between symmetric and self-adjoint hence is based on the domain of definition. If the domain of definition \mathcal{D}_{A^\dagger} of the adjoint operator A^\dagger , which is defined by $\langle A^\dagger\varphi|\psi\rangle = \langle\varphi|A\psi\rangle$, is smaller than \mathcal{D}_A , then we first have to restrict the definition of A to

¹⁰It can be shown that \mathcal{D}_{A^\dagger} consists of all vectors $\varphi \in \mathcal{H}$ for which $(|\langle A\psi|\varphi\rangle|) / (|\psi|)$ is (uniformly) bounded for all $\psi \in \mathcal{D}_A$ with $|\psi\rangle \neq 0$; see e.g. section VIII.1 of [Reed].

a subset of \mathcal{D}_A , which will at the same time increase \mathcal{D}_{A^\dagger} . If A thus becomes (or already is) a symmetric operator then we can ask the question whether it is possible to extend \mathcal{D}_A such that A becomes self-adjoint. This question has been answered by a theorem first stated (for second order differential operators) by Weyl in 1910, and generalized by John von Neumann in 1929:

Self-adjoint extension of operators: The existence of a self-adjoint extension depends on the so-called deficiency indices n_\pm of A , which are the dimensions of the eigenspaces N_\pm of A^\dagger for some fixed positive and negative imaginary eigenvalues $\pm i\varepsilon$, respectively,

$$A^\dagger \psi = \pm i\varepsilon \psi, \quad \varepsilon > 0, \quad (3.113)$$

where one may set, for example, $\varepsilon = 1$. Depending on these indices there are three cases:

- If $n_+ = n_- = 0$ then the operator A is already self-adjoint.
- If $n_+ = n_- \geq 1$ then A is not self-adjoint but admits infinitely many self-adjoint extensions.
- If $n_+ \neq n_-$ then a self-adjoint extension of A does not even exist.

A detailed discussion of simple examples for these situations can be found in a paper by Bonneau, Faraut and Valent.¹¹

Spectral theorem. The content of the spectral theorem is that self-adjoint operators A are essentially multiplication operators in an appropriate eigenbasis, i.e. there exists a decomposition of unity as a sum of projection operators with the direction of the projections aligned with the eigenspaces of A ,

$$\mathbb{1} = \sum_i P_i, \quad A = \sum_i a_i P_i. \quad (3.114)$$

In the infinite-dimensional case of self-adjoint operators in Hilbert spaces the first complication is that the spectrum may consist of discrete and continuous parts, so that the sum has to be generalized to an operation that can at the same time describe sums and integrals. This is achieved by the Riemann-Stieltjes integral, which allows to assign different weights to different parts of the integration interval. Assume that $\mu(x)$ is a monotonously increasing function with only isolated discontinuities. Then we think of the mass density given by the derivative $d\mu = \mu' dx$ which has (positive) δ -function like concentrations at the discontinuities of μ and the **Riemann–Stieltjes integral** for smooth functions can be written as

$$\int_a^b f(x) d\mu(x) = \int_a^b f(x) \mu'(x) dx \quad (3.115)$$

¹¹ G. Bonneau, J. Faraut, G. Valent, *Self-adjoint extensions of operators and the teaching of quantum mechanics*, Am.J.Phys. 69 (2001) 322 [<http://arxiv.org/abs/quant-ph/0103153>]

where we include, by convention, the contribution of δ -functions located at the upper integration limit

$$\int_a^b f(x) d\mu(x) = \lim_{\varepsilon \rightarrow 0^+} \int_a^{b+\varepsilon} f(x) d\mu(x) \quad (3.116)$$

and accordingly drop point-like contributions at the lower limit to make the whole integral additive for intervals. The extension of the definition of the integral for non-smooth integrands $f(x)$ then proceeds like for the case of Riemann integration by taking limits of upper and lower bounds. Using the methods of measure theory this can be generalized to the (Lesbeques-) Stiltjes integral, allowing general *measurable functions* to be integrated.

The application of Stiltjes integration to spectral theory introduces the concept of a spectral family $\{\mathbb{E}_\lambda\}$ associated with an operator A , which is the one-parameter family of sums/integrals of the projectors for all spectral values up to a certain number $\lambda \in \mathbb{R}$. At first we assume that A is bounded, so that its spectrum is contained in an interval $\lambda \in [a, b]$. \mathbb{E}_λ grows monotonically in λ and is a family of projectors that is continuous from above, i.e. one can show [Kreyszig]

$$\forall \nu \geq \lambda \quad : \quad \mathbb{E}_\nu \geq \mathbb{E}_\lambda, \quad (3.117)$$

$$\forall \lambda < a \quad : \quad \mathbb{E}_\lambda = 0, \quad (3.118)$$

$$\forall \lambda > b \quad : \quad \mathbb{E}_\lambda = \mathbb{1}, \quad (3.119)$$

$$\lim_{\nu \rightarrow \lambda^+} \mathbb{E}_\nu = \mathbb{E}_\lambda. \quad (3.120)$$

The **theorem of Stone** then asserts that a **bounded self-adjoint operator** A has the spectral representation

$$A = \int_{a-}^b \lambda d\mathbb{E}_\lambda \quad (3.121)$$

for a spectral family \mathbb{E}_λ , where the Riemann–Stiltjes integral is uniformly convergent (with respect to the operator norm). The lower limit $a-$ indicates that we have to include the δ -function contribution at $\lambda = a$ if a is part of the discrete spectrum.

Unbounded and unitary operators.¹² The spectral theorem can now be extended to unbounded operators using the Cayley transformation to a unitary operator

$$U = (A - i\mathbb{1})(A + i\mathbb{1})^{-1} \quad (3.122)$$

where the resolvent $(A + i\mathbb{1})^{-1}$ exists as a bounded operator because the spectrum of A is real.

The spectral decomposition for unitary operators follows from the fact that we can decompose them into a commuting set of self-adjoint operators $V = \frac{1}{2}(U + U^\dagger) = V^\dagger$ and

¹² The most general family of operators for which a spectral decomposition exists are the normal operators, defined by the equation $NN^\dagger = N^\dagger N$, i.e. N commutes with its adjoint, which obviously covers both the self-adjoint and the unitary case. Normal operators can also be characterized by the fact that they are unitarily diagonalizable.

$W = \frac{1}{2i}(U - U^\dagger) = W^\dagger$ which commute because

$$U = V + iW, \quad UU^\dagger = U^\dagger U \quad \Rightarrow \quad V^2 - i(VW - WV) + W^2 = V^2 + i(VW - WV) + W^2 \quad (3.123)$$

so that $VW - WV = 0$. Hence they have a spectral decomposition with a common spectral family. Putting together real and imaginary part of the eigenvalues of U we find

$$U = \int_{-\pi}^{\pi} e^{i\theta} d\mathbb{E}_\theta. \quad (3.124)$$

The spectrum is located on the unit circle and the convergence of the integral is again uniform. For unbounded operators A the Cayley transform can now be inverted with the formula

$$A = i(\mathbb{1} + U)(\mathbb{1} - U)^{-1} \quad (3.125)$$

and we obtain

$$A = \int_{-\pi}^{\pi} \tan\left(\frac{\theta}{2}\right) d\mathbb{E}_\theta = \int_{-\infty}^{\infty} \lambda d\mathbb{F}_\lambda \quad (3.126)$$

with the appropriate change of measure density in the spectral family. Since the spectrum can be unbounded the Stiltjes integral is now defined in the sense of strong operator convergence.

3.4 Schrödinger, Heisenberg and interaction picture

We now return to the issue of time dependence in quantum mechanics, which we described so far by the time dependence of states

$$\psi(x, t) = \langle x | \psi(t) \rangle \in L^2(\mathbb{R}^3) \quad \text{for } t \geq t_0, \quad (3.127)$$

i.e. by time-dependent vectors in Hilbert space that are determined by some initial condition $\psi(x, t_0)$ at an initial time and by solving the Schrödinger equation for later times. Since the map $\psi(x, t_0) \rightarrow \psi(x, t)$ is linear it defines a linear operator $U(t, t_0)$

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle \quad (3.128)$$

called **time evolution operator**. More precisely $U(t, t_0)$ is a family of operators depending on two parameters, the initial time t_0 and the final time t , where we can also consider $t < t_0$ by solving the Schrödinger equation backwards in time. If we choose some orthonormal basis $|e_i(t_0)\rangle$ at an initial time then $|e_i(t)\rangle$ also forms an orthonormal basis at later times: The normalization $\langle e_i(t) | e_i(t) \rangle = \langle e_i(t_0) | e_i(t_0) \rangle = 1$ expresses the conservation of probability, and orthogonality at later times follows from the general fact that conservation of norms implies conservation of scalar products: Since

$$\begin{aligned} \|e_1 + e_2\|^2 &= \langle e_1 + e_2 | e_1 + e_2 \rangle = \|e_1\|^2 + \|e_2\|^2 + \langle e_1 | e_2 \rangle + \langle e_2 | e_1 \rangle \\ \|e_1 + ie_2\|^2 &= \langle e_1 + ie_2 | e_1 + ie_2 \rangle = \|e_1\|^2 + \|e_2\|^2 + i\langle e_1 | e_2 \rangle - i\langle e_2 | e_1 \rangle. \end{aligned} \quad (3.129)$$

the scalar product $\langle e_1|e_2\rangle$ can be reconstructed from the norm by solving the equations (3.129) for $\langle e_1|e_2\rangle$ and $\langle e_2|e_1\rangle$, which become complex conjugates of one another because their sum is real and their difference imaginary. We conclude that orthonormal bases stay orthonormal (and complete) during time evolution, so that the time evolution operator $U(t, t_0)$ amounts to a change of basis and hence is a unitary operator $U^\dagger = U^{-1}$.

Inserting the definition (3.128) into the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle \quad (3.130)$$

and using $\frac{\partial}{\partial t} |\psi(t_0)\rangle = 0$ we obtain

$$i\hbar \left(\frac{\partial}{\partial t} U \right) |\psi(t_0)\rangle = HU |\psi(t_0)\rangle. \quad (3.131)$$

Since this relation has to hold for all $|\psi(t_0)\rangle$ it implies the operator differential equation

$$i\hbar \frac{\partial U}{\partial t} = HU. \quad (3.132)$$

If the Hamiltonian does not explicitly depend on time this equation can be solved formally and we obtain

$$U(t, t_0) = U(t - t_0) = e^{-\frac{i}{\hbar} H(t-t_0)}, \quad (3.133)$$

which only depends on the time difference $t - t_0$.

The Heisenberg picture. With the time evolution operator we can now write expectation values of operators as

$$\begin{aligned} \langle A \rangle &= \langle \psi(t) | A | \psi(t) \rangle = \langle \psi(t_0) | U^\dagger A U | \psi(t_0) \rangle \\ &= \langle \psi(t_0) | A_H | \psi(t_0) \rangle \quad \text{with} \quad A_H(t) = U^\dagger(t) A U(t) \end{aligned} \quad (3.134)$$

where we assume that A does not explicitly depend on time.

So far we discussed quantum mechanics in terms of the so-called **Schrödinger picture**, in which the time dependence of the system is governed by the Schrödinger equation for a time dependent wave function and operators are time independent, at least if the apparatus is not moved and if there are no other external sources of time dependence,

$$|\psi_S(t)\rangle \equiv |\psi(t)\rangle, \quad A_S \equiv A. \quad (3.135)$$

But since all observable quantities in quantum mechanics can be expressed in terms of expectation values, eq. (3.134) shows that we can take a different point of view and interpret the time evolution as acting on the operators according to

$$A_H(t) = U^\dagger(t, t_0) A(t_0) U(t, t_0) \quad (3.136)$$

while the states do not change

$$|\psi_H(t)\rangle = |\psi(t_0)\rangle. \quad (3.137)$$

The description of quantum mechanics in terms of $A_H(t)$ and $|\psi_H\rangle$ is called **Heisenberg picture**, whereas the description in terms of A_S and $|\psi_S(t)\rangle$ is called Schrödinger picture and our definitions imply

$$\langle A \rangle = \langle \psi_S(t) | A_S | \psi_S(t) \rangle = \langle \psi_H | A_H(t) | \psi_H \rangle \quad (3.138)$$

where the two pictures are related by a unitary transformation

$$|\psi_S(t)\rangle = U |\psi_H\rangle \quad , \quad |\psi_H\rangle = U^\dagger |\psi_S(t)\rangle \quad (3.139)$$

$$A_S = U A_H(t) U^\dagger \quad , \quad A_H(t) = U^\dagger A_S U \quad (3.140)$$

with $U = e^{-\frac{i}{\hbar} H(t-t_0)}$ if H is time-independent.

While the Schrödinger picture seems to be more intuitive at first glance, the Heisenberg picture shows a formal similarity with classical mechanics: Since $\partial_t U = -\frac{i}{\hbar} H U$ and $\partial_t U^\dagger = \frac{i}{\hbar} U^\dagger H$ the infinitesimal time evolution of the Heisenberg operators is

$$\frac{\partial A_H}{\partial t} = \frac{\partial U^\dagger}{\partial t} A_S U + U^\dagger \overbrace{\frac{\partial A_S}{\partial t}}^{\equiv 0} U + U^\dagger A_S \frac{\partial U}{\partial t} \quad (3.141)$$

$$= \frac{i}{\hbar} (U^\dagger H_S A_S U - U^\dagger A_S H_S U) \quad (3.142)$$

$$= \frac{i}{\hbar} ((U^\dagger H_S U)(U^\dagger A_S U) - (U^\dagger A_S U)(U^\dagger H_S U)), \quad (3.143)$$

where we inserted $\mathbb{1} = U U^\dagger$. We thus obtain **Heisenberg's equation of motion**

$$\boxed{\frac{\partial A_H}{\partial t} = \frac{i}{\hbar} [H_H, A_H]}, \quad (3.144)$$

which has a formal similarity to Hamilton's equations of motion $\dot{f} = \{H, f\}_{PB}$ for phase space functions $f(p, q)$, or $\frac{\partial p_i}{\partial t} = \{H, p_i\}_{PB} = (-\frac{\partial H}{\partial q_i})$ and $\frac{\partial q_i}{\partial t} = \{H, q_i\}_{PB} = \frac{\partial H}{\partial p_i}$ for coordinates and momenta; $\{\dots\}_{PB}$ denotes the Poisson bracket. This analogy is the starting point for the general quantization rules of Hamiltonian systems.

The interaction picture (or **Dirac picture**) combines elements of the Schrödinger picture and of the Heisenberg picture so that states and operators both become time dependent. It is the starting point for approximation techniques and useful if we can write the Hamiltonian as a sum of a simple (exactly solvable) time-independent part H_0 and a (small) time-dependent perturbation $V(t)$,

$$H(t) = H_0 + V(t). \quad (3.145)$$

The idea is to put the simple part of the time evolution into the time dependence of operators $A_I(t)$, thereby obtaining a modified Schrödinger equation for the time evolution of states,

which only leads to a relatively small time dependence of $|\psi_I(t)\rangle$ due to a possibly complicated but weak interaction term $V(t)$. The interaction picture is thus defined by the unitary transformation

$$|\psi_I(t)\rangle = U_0^\dagger(t)|\psi_S(t)\rangle, \quad A_I(t) = U_0^\dagger(t) A_S U_0(t) \quad (3.146)$$

with

$$U_0(t) = e^{-\frac{i}{\hbar}(t-t_0)H_0} \quad (3.147)$$

so that

$$\langle A \rangle = \langle \psi_S(t) | A_S | \psi_S(t) \rangle = \langle \psi_I(t) | A_I(t) | \psi_I(t) \rangle \quad (3.148)$$

The Schrödinger equation in the interaction picture is now obtained by evaluating

$$i\hbar \frac{\partial}{\partial t} |\psi_I(t)\rangle = i\hbar \frac{\partial}{\partial t} \left(e^{\frac{i}{\hbar}H_0(t-t_0)} |\psi_S(t)\rangle \right) = -U_0^\dagger(t) H_0 |\psi_S(t)\rangle + i\hbar U_0^\dagger(t) \frac{\partial}{\partial t} |\psi_S(t)\rangle \quad (3.149)$$

$$= U_0^\dagger(t) (-H_0 + H_0 + V(t)) |\psi_S(t)\rangle = U_0^\dagger(t) V(t) U_0(t) U_0^\dagger(t) |\psi_S(t)\rangle, \quad (3.150)$$

where we used the Schrödinger equation $\partial_t |\psi_S(t)\rangle = -\frac{i}{\hbar}(H_0 + V(t)) |\psi_S(t)\rangle$, so that

$$\boxed{i\hbar \frac{\partial}{\partial t} |\psi_I(t)\rangle = V_I(t) |\psi_I(t)\rangle} \quad (3.151)$$

Replacing H by H_0 and U by U_0 in the derivation of Heisenberg's equation of motion we obtain the operator equation of motion

$$\boxed{\frac{\partial A_I(t)}{\partial t} = \frac{i}{\hbar} [H_{0,I}(t), A_I(t)]} \quad (3.152)$$

in the interaction picture. The time evolution operator $U_I(t) = U_0^\dagger(t)U(t)U_0(t)$ in the interaction picture describes the time evolution of the states $|\psi_I(t)\rangle = U_I(t, t_0)|\psi_I(t_0)\rangle$ and hence satisfies the equation of motion

$$i\hbar \frac{\partial U_I}{\partial t} = V_I(t) U_I. \quad (3.153)$$

In many situations $U_I(t)$ can only be computed by approximation procedures.

3.5 Ehrenfest theorem and uncertainty relations

In this section we want to improve our understanding of the relation between quantum mechanics and classical mechanics. The content of the Ehrenfest theorem is that expectation values of observables obey classical equations of motion. Heisenberg's uncertainty relation, on the other hand, implies limitations to the validity of classical concepts.

Let us compute the time evolution

$$\frac{d}{dt} \langle A \rangle = \frac{\partial \langle \psi |}{\partial t} A | \psi \rangle + \langle \psi | \frac{\partial A}{\partial t} | \psi \rangle + \langle \psi | A \frac{\partial | \psi \rangle}{\partial t} \quad (3.154)$$

of the mean value of an observable A in the Schrödinger picture. The Schrödinger equation yields

$$\frac{d}{dt}\langle A \rangle = \left\langle \frac{\partial A}{\partial t} \right\rangle + \frac{i}{\hbar} \langle H\psi | A | \psi \rangle - \frac{i}{\hbar} \langle \psi | A | H\psi \rangle = \left\langle \frac{\partial A}{\partial t} \right\rangle + \frac{i}{\hbar} \langle \psi | H A - A H | \psi \rangle \quad (3.155)$$

so that

$$\boxed{\frac{d}{dt}\langle \psi | A | \psi \rangle = \langle \psi | \frac{\partial A}{\partial t} | \psi \rangle + \frac{i}{\hbar} \langle \psi | [H, A] | \psi \rangle} \quad (3.156)$$

The Ehrenfest theorem states that the mean values of certain quantum mechanical operators obey the classical relations.

As an example let us compute the time evolution $\frac{d}{dt}\langle X_i \rangle$ of the position operator X_i , which is time independent $\frac{\partial X}{\partial t}$ in the Schrödinger picture. We first have to compute the commutator of X with the Hamiltonian $H = \frac{1}{2m} P_j P_j + V(\vec{x})$.

$$[H, X_i] = \frac{1}{2m} [P_j P_j, X_i] + \overbrace{[V(\vec{x}), X_i]}^{\equiv 0} = \frac{1}{2m} (P_j [P_j, X_i] + [P_j, X_i] P_j) = \frac{1}{2m} 2 \frac{\hbar}{i} P_i. \quad (3.157)$$

Inserting this result into the formula (3.156) for the mean value of an operator we obtain

$$\frac{d}{dt}\langle X_i \rangle = \frac{1}{m} \langle P_i \rangle. \quad (3.158)$$

This is the quantum analogue of the classical equation $p_i = m \frac{dx_i}{dt}$. Similarly we can compute the time evolution for the momentum operator. Inserting $[H, P_i] = -\frac{\hbar}{i} \frac{\partial}{\partial x_i} V(\vec{x})$ into (3.156) we obtain

$$\frac{d}{dt}\langle P_i \rangle = -\langle \partial_i V(\vec{x}) \rangle, \quad (3.159)$$

which corresponds to Newton's equation of motion $\dot{\vec{p}} = -\vec{\nabla} V(\vec{x})$.

Heisenberg's uncertainty relation. Let $\langle A \rangle$ and $\langle B \rangle$ be the expectation values of two Hermitian operators A and B in some normalized state $|\psi\rangle \in \mathcal{H}$. The uncertainty ΔA is defined by

$$(\Delta A)^2 = \langle A^2 \rangle - \langle A \rangle^2 = \langle (\delta A)^2 \rangle \quad \text{with} \quad \langle A \rangle \equiv \langle \psi | A | \psi \rangle, \quad \delta A = A - \langle A \rangle \mathbb{1}. \quad (3.160)$$

For the operators $\delta A = A - \langle A \rangle$ and $\delta B = B - \langle B \rangle$, which describe the deviation of the observables A and B from their mean values, we consider the states

$$|\chi\rangle = \delta A |\psi\rangle \quad \text{and} \quad |\varphi\rangle = \delta B |\psi\rangle \quad (3.161)$$

whose norms are equal to the uncertainties

$$\langle \chi | \chi \rangle = \langle \psi | (\delta A)^2 | \psi \rangle = (\Delta A)^2, \quad \langle \varphi | \varphi \rangle = \langle \psi | (\delta B)^2 | \psi \rangle = (\Delta B)^2, \quad (3.162)$$

and which satisfy the Schwartz inequality

$$\langle \chi | \chi \rangle \langle \varphi | \varphi \rangle \geq |\langle \chi | \varphi \rangle|^2. \quad (3.163)$$

Putting everything together we obtain the inequality

$$(\Delta A)^2 (\Delta B)^2 = \langle \psi | (\delta A)^2 | \psi \rangle \langle \psi | (\delta B)^2 | \psi \rangle \geq |\langle \psi | (\delta A) (\delta B) | \psi \rangle|^2 = |\langle \delta A \delta B \rangle|^2. \quad (3.164)$$

Now we write the operator in the last term as $\delta A \delta B = \frac{1}{2}[\delta A, \delta B] + \frac{1}{2}\{\delta A, \delta B\}$ and consider the commutator $[\delta A, \delta B]$. Since $\langle A \rangle$ and $\langle B \rangle$ are scalars that commute with everything we observe

$$[\delta A, \delta B] = [A - \langle A \rangle \mathbb{1}, B - \langle B \rangle \mathbb{1}] = [A, B]. \quad (3.165)$$

Since a commutator of Hermitian operators is anti-Hermitian its expectation value is imaginary. Anti-commutators of Hermitian operators, on the other hand, are Hermitian and thus have real expectation values. Hence

$$\langle \delta A \delta B \rangle = \frac{1}{2} \langle \{\delta A, \delta B\} \rangle + \frac{1}{2} \langle [\delta A, \delta B] \rangle = \frac{1}{2} \langle \{\delta A, \delta B\} \rangle + \frac{1}{2} \langle [A, B] \rangle \quad (3.166)$$

decomposes the expectation value into its real and its imaginary part, so that its squared modulus becomes

$$|\langle \delta A \delta B \rangle|^2 = \frac{1}{4} |\langle [A, B] \rangle|^2 + \frac{1}{4} |\langle \{\delta A, \delta B\} \rangle|^2 \geq \frac{1}{4} |\langle [A, B] \rangle|^2. \quad (3.167)$$

Combining this estimate with the inequality (3.164) we find

$$\langle (\delta A)^2 \rangle \langle (\delta B)^2 \rangle \geq \frac{1}{4} |\langle [A, B] \rangle|^2 \quad (3.168)$$

and by taking positive square roots on both sides we find the general form of **Heisenberg's uncertainty relation**

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle| \quad (3.169)$$

which establishes a lower bound on the product of uncertainties of two operators in terms of the expectation value of their commutator. The two respective observables can hence be measured simultaneously with arbitrary precision only if the operators commute, or, more precisely, if their commutator has vanishing expectation value. We should stress that this uncertainty is not simply a problem of measurement but rather an intrinsic property of quantum mechanics.

Uncertainty of position and momentum. For the most famous example of an uncertainty relation we insert the commutator between position and momentum and obtain

$$[P_j, X_i] = \frac{\hbar}{i} \delta_{ij} \quad \Rightarrow \quad (\Delta X_i) (\Delta P_j) \geq \frac{\hbar}{2} \delta_{ij}. \quad (3.170)$$

so that position and momentum *in the same direction* cannot be measured simultaneously with arbitrary precision.¹³

¹³ For states of minimal uncertainty $\Delta X \Delta P = \frac{\hbar}{2}$ the two inequalities (3.164) and (3.167) have to be equalities, which requires that $\delta X |\psi\rangle$ and $\delta P |\psi\rangle$ are proportional and that $\delta X \delta P + \delta P \delta X$ have vanishing expectation value. It is easy to check that this can only be the case for Gaussian wave packets.

Uncertainty of time and energy. If we consider the form of a plane wave $\psi = e^{\frac{i}{\hbar}(\vec{p}\vec{x} - Et)}$ we might expect that there exists an uncertainty between energy and time analogous to the one between momentum and position. There exists, however, no time operator in quantum mechanics and hence no uncertainty relation involving t in the literal sense.

Uncertainty relations of the expected type do exist, however, if we think of time in terms of time measurements, like for example the time of the detection of a particle. Such a measurement always involves the observation of a change in time of the value of some observable A and the uncertainty of time would be the time that it takes for this change to become larger than the intrinsic uncertainty of that observable,

$$\Delta t_A = \frac{\Delta A}{\left| \frac{d}{dt} \langle A \rangle \right|}. \quad (3.171)$$

Since time evolution is generated by the Hamilton operator an uncertainty relation for Δt_A can now be obtained as a consequence of the uncertainty relation between A and H ,

$$\Delta A \Delta E \geq \frac{1}{2} |\langle [H, A] \rangle|. \quad (3.172)$$

If we combine this with the equation of motion (3.156) of the expectation value for a time-independent observable A ,

$$\frac{d}{dt} \langle A \rangle = \frac{i}{\hbar} \langle [H, A] \rangle, \quad (3.173)$$

we obtain $\Delta A \Delta E \geq \frac{\hbar}{2} \left| \frac{d}{dt} \langle A \rangle \right|$ and hence the uncertainty relation

$$\Delta t_A \cdot \Delta E \geq \frac{\hbar}{2}, \quad (3.174)$$

which is exactly of the form that we had hoped for. It is hence not possible to *simultaneously* measure the energy of a particle and time of its detection with arbitrary precision.

3.6 Harmonic oscillator and ladder operators

Using the operator calculus we now determine the energy spectrum of the harmonic oscillator by purely algebraic calculations. We begin by introducing dimensionless position and momentum operators

$$\mathcal{X} = \sqrt{\frac{m\omega_0}{\hbar}} X, \quad \mathcal{P} = \frac{1}{\sqrt{m\hbar\omega_0}} P \quad (3.175)$$

so that

$$H = \frac{\hbar\omega_0}{2} (\mathcal{P}^2 + \mathcal{X}^2). \quad (3.176)$$

Classically we can factorize $x^2 + p^2 = (x + ip)(x - ip)$ as a product of complex conjugate numbers. Analogously, we introduce the non-Hermitian **ladder operators**

$$a = \frac{1}{\sqrt{2}} (\mathcal{X} + i\mathcal{P}), \quad a^\dagger = \frac{1}{\sqrt{2}} (\mathcal{X} - i\mathcal{P}) \quad (3.177)$$

with

$$X = \sqrt{\frac{\hbar}{2m\omega_0}} (a^\dagger + a), \quad P = i\sqrt{\frac{m\hbar\omega_0}{2}} (a^\dagger - a). \quad (3.178)$$

Since $[\mathcal{P}, \mathcal{X}] = \frac{1}{i}$ the commutator becomes $[a, a^\dagger] = \frac{1}{2}[\mathcal{X} + i\mathcal{P}, \mathcal{X} - i\mathcal{P}] = 0 + \frac{1}{2} + \frac{1}{2} + 0 = 1$, i.e.

$$[a, a^\dagger] = -[a^\dagger, a] = 1. \quad (3.179)$$

With the quantum mechanical relation

$$\mathcal{X}^2 + \mathcal{P}^2 = \frac{1}{2} ((a^\dagger + a)^2 - (a^\dagger - a)^2) = a^\dagger a + a a^\dagger = 2a^\dagger a + 1 \quad (3.180)$$

we thus obtain

$$H = \hbar\omega_0 \left(a^\dagger a + \frac{1}{2} \right) = \hbar\omega_0 \left(\mathcal{N} + \frac{1}{2} \right), \quad (3.181)$$

where we defined the **occupation number operator**

$$\mathcal{N} = a^\dagger a. \quad (3.182)$$

This operator is positive, i.e. all its expectation values are positive, because

$$\langle \psi | \mathcal{N} | \psi \rangle = (\langle \psi | a^\dagger) (a | \psi \rangle) = \| (a | \psi \rangle) \|^2 \geq 0 \quad (3.183)$$

is the squared norm of the vector $a|\psi\rangle$. Consequently all expectation values of H , and hence all energy eigenvalues E , are bounded from below by

$$E \geq \frac{1}{2} \hbar\omega_0. \quad (3.184)$$

$\frac{1}{2} \hbar\omega_0$ is called **zero-point energy** of the harmonic oscillator.

Creation and annihilation of energy. Since $H = \hbar\omega_0(\mathcal{N} + \frac{1}{2})$ the energy spectrum can now be computed by solving the eigenvalue problem for the occupation number operator

$$\mathcal{N}|n\rangle = n|n\rangle \quad \Rightarrow \quad H|n\rangle = \hbar\omega_0(n + \frac{1}{2})|n\rangle. \quad (3.185)$$

In order to solve this equation we compute the commutators

$$[\mathcal{N}, a] = [a^\dagger, a]a = -a, \quad [\mathcal{N}, a^\dagger] = a^\dagger[a, a^\dagger] = a^\dagger, \quad (3.186)$$

where we evaluated $[a^\dagger a, a] = a^\dagger[a, a] + [a^\dagger, a]a = [a^\dagger, a]a = -a$ using the ‘‘Leibniz rule’’ (3.45). These commutation relations show that a^\dagger (a) increases (decreases) the occupation number by one and, accordingly, the energy by $\hbar\omega_0$ because

$$\mathcal{N}|n\rangle = n|n\rangle \quad \Rightarrow \quad \begin{cases} \mathcal{N}(a^\dagger|n\rangle) = ([\mathcal{N}, a^\dagger] + a^\dagger\mathcal{N})|n\rangle = (1+n)(a^\dagger|n\rangle) \\ \mathcal{N}(a|n\rangle) = ([\mathcal{N}, a] + a\mathcal{N})|n\rangle = (-1+n)(a|n\rangle) \end{cases} \quad (3.187)$$

(where we used the identity $XY = [X, Y] + YX$). Thus a^\dagger and a are called **creation** and **annihilation operator**, respectively. Their collective name is “ladder operators” because they bring us up and down the ladder of energy levels. More precisely, since (3.187) implies that $a|n\rangle$ and $a^\dagger|n\rangle$ have occupation numbers $n \pm 1$, these states must be proportional to $|n \pm 1\rangle$

$$a^\dagger|n\rangle = c_{n+}|n+1\rangle, \quad a|n\rangle = c_{n-}|n-1\rangle. \quad (3.188)$$

Assuming that all states are normalized $\langle n|n\rangle = 1$ we can now compute the normalization factors $c_{n\pm}$. Since norms are computed by multiplication with the Hermitian conjugate states,

$$a|n\rangle = c_{n-}|n-1\rangle \xrightarrow{\text{conj.}} \langle n|a^\dagger = c_{n-}^* \langle n-1|, \quad (3.189)$$

$$a^\dagger|n\rangle = c_{n+}|n+1\rangle \xrightarrow{\text{conj.}} \langle n|a = c_{n+}^* \langle n+1|, \quad (3.190)$$

the eigenvalue equation $a^\dagger a|n\rangle = \mathcal{N}|n\rangle = n|n\rangle$ implies

$$\langle n+1|n+1\rangle = \frac{1}{|c_{n+}|^2} \langle n|a a^\dagger|n\rangle = \frac{1}{|c_{n+}|^2} (\langle n|a^\dagger a|n\rangle + 1) = \frac{n+1}{|c_{n+}|^2} = 1, \quad (3.191)$$

$$\langle n-1|n-1\rangle = \frac{1}{|c_{n-}|^2} \langle n|a^\dagger a|n\rangle = \frac{n}{|c_{n-}|^2} = 1, \quad (3.192)$$

so that

$$c_{n+} = \sqrt{n+1}, \quad c_{n-} = \sqrt{n}, \quad (3.193)$$

where the phase ambiguity of the eigenvectors $|n\rangle$ has been used to choose $c_{n\pm}$ positive real.

Quantization of occupation number and energy. Now we are ready to determine the eigenvalues n . We assume that at least one eigenstate $|n\rangle$ exists for some eigenvalue $n \in \mathbb{R}$, which has to be non-negative $n \geq 0$ because of the positivity (3.183) of \mathcal{N} . Now we act on this state k times with the annihilation operator a and obtain

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad (3.194)$$

$$a^2|n\rangle = \sqrt{n(n-1)}|n-2\rangle, \quad (3.195)$$

...

$$a^k|n\rangle = \sqrt{n(n-1)\dots(n-k+1)}|n-k\rangle. \quad (3.196)$$

We thus find new energy eigenstates with occupation numbers $n-1$, $n-2$, ... However, this procedure has to terminate because otherwise we would be able to construct energy eigenstates for arbitrary $n-k$, which turns negative for $k > n$ in contradiction to the positivity of the operator \mathcal{N} . Hence there must exist a positive integer K for which $a^K|n\rangle = 0$. Choosing K minimal, so that $a^{K-1}|n\rangle \neq 0$, we conclude that $a|n-K+1\rangle = 0$ and hence $\langle n-K+1|a^\dagger a|n-K+1\rangle = n-K+1 = 0$. In other words, if $a|n'\rangle = 0$ the normalization factor $c_{n'-}$ must vanish, which is the only possibility to avoid the existence of an energy eigenstate with eigenvalue

$n' - 1$. We conclude that each energy eigenvalue n must be a non-negative integer. Moreover, eq. (3.196) shows that the minimal energy state has occupation number $n = 0$, and by acting with creation operators on the ground state $|0\rangle$,

$$(a^\dagger)^n |0\rangle = \sqrt{n!} |n\rangle \quad (3.197)$$

we conclude that all energy eigenstates with nonnegative integer occupation number indeed exist. We thus recover the result

$$E_n = \hbar\omega_0\left(n + \frac{1}{2}\right) \quad \text{with} \quad n = 0, 1, 2, \dots \quad (3.198)$$

of our analytical treatment of the Harmonic oscillator. Moreover, the ground state $|0\rangle$ satisfies the first order differential equation $a|0\rangle = (\mathcal{X} + i\mathcal{P})|0\rangle = 0$, which is easily solved yielding the Gaussian wave function found in section 2. The wave functions with positive occupation numbers are

$$u_n(x) = \langle x|n\rangle = \frac{1}{\sqrt{n!}} (a^\dagger)^n u_0(x) \quad (3.199)$$

and can be evaluated by repeated application of the differential operator a^\dagger .

3.6.1 Coherent states

Coherent states are, by definition, eigenstates of the annihilation operator

$$a|\lambda\rangle_{\text{coh}} = \lambda|\lambda\rangle_{\text{coh}}. \quad (3.200)$$

They exist for all complex numbers $\lambda \in \mathbb{C}$ and are unique up to normalization. This can be verified by inserting the ansatz $|\lambda\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$ in terms of energy eigenstates $|n\rangle$ into the eigenvalue equation (3.200). With the choice $c_0 = 1$ the resulting recursion relation is solved by

$$|\lambda\rangle_{\text{coh}} = \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} |n\rangle = e^{\lambda a^\dagger} |0\rangle. \quad (3.201)$$

It will usually be sufficient to distinguish coherent states $|\lambda\rangle_{\text{coh}}$ from energy eigenstates $|n\rangle$ by the use of Greek letters for the eigenvalues of a . The eigenstate property of $e^{\lambda a^\dagger} |0\rangle$ can be verified directly,

$$a(e^{\lambda a^\dagger} |0\rangle) = e^{\lambda a^\dagger} (e^{-\lambda a^\dagger} a e^{\lambda a^\dagger}) |0\rangle = e^{\lambda a^\dagger} (a + \lambda) |0\rangle = \lambda(e^{\lambda a^\dagger} |0\rangle), \quad (3.202)$$

where we used $a|0\rangle = 0$ and the formula

$$e^{\lambda A} B e^{-\lambda A} = B + \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} [A, B]_n, \quad [A, B]_{n+1} = [A, [A, B]_n], \quad (3.203)$$

with $[A, B]_1 = [A, B]$. Since $[a^\dagger, a] = -1$ all higher commutators vanish.

Scalar products among coherent states can be computed directly from the series expansion or with the Baker-Campbell-Hausdorff formula

$$\langle \lambda | \mu \rangle = \langle 0 | e^{\lambda^* a} e^{\mu a^\dagger} | 0 \rangle = \langle 0 | e^{\mu a^\dagger} e^{\lambda^* \mu [a, a^\dagger]} e^{\lambda^* a} | 0 \rangle = e^{\lambda^* \mu} \quad (3.204)$$

where we used that eq. (3.59) implies $e^A e^B = e^B e^A e^{[A, B]}$ if all double-commutators of A and B vanish, as is the case for $A = \lambda^* a$ and $B = \mu a^\dagger$ because $[a, a^\dagger] = 1$. We also used $e^{\lambda^* a} | 0 \rangle = | 0 \rangle$ (because only the first term of the series is nonzero) and the Hermitian conjugate formula $\langle 0 | e^{\mu a^\dagger} = \langle 0 |$. Eigenstates of a for different eigenvalues are not orthogonal and the eigenvalues are neither quantized nor required to be real (which is o.k. because a is not self-adjoint). For normalized coherent states we thus find the formula $e^{-\frac{1}{2}|\lambda|^2} e^{\lambda a^\dagger} | 0 \rangle$.

The time evolution of coherent states is easily calculated by using the expansion in terms of energy eigenstates,

$$|\lambda\rangle(t) = e^{-\frac{i}{\hbar} H t} \left(\sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} |n\rangle \right) = \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} e^{-i(n+\frac{1}{2})\omega_0 t} |n\rangle = e^{-i\frac{\omega_0}{2} t} |\lambda(t)\rangle, \quad \lambda(t) = e^{-i\omega_0 t} \lambda. \quad (3.205)$$

Up to an unobservable phase factor the time evolution thus corresponds to a rotation of the eigenvalue $\lambda(t) = e^{-i\omega_0 t} \lambda$ in the complex plane. In fact, the probability density of $\langle \lambda | \lambda \rangle(t)$ is given by a Gaussian distribution with minimal uncertainty $\Delta X \Delta P = \hbar/2$ and constant shape, whose mean value oscillates with the classical frequency ω_0 , explaining the name *coherent state*. This can be shown by computing the wave function in configuration space $\psi_\lambda(x) = \langle x | \lambda \rangle$, which satisfies the first order differential equation

$$(a - \lambda) \psi_\lambda(x) = \frac{1}{\sqrt{2}} \left(\alpha x + \frac{1}{\alpha} \partial_x - \sqrt{2} \lambda \right) \psi_\lambda(x) = 0 \quad \text{with} \quad \alpha = \sqrt{\frac{m\omega_0}{\hbar}}. \quad (3.206)$$

With the ansatz $\psi_\lambda(x) = e^{-Ax^2+Bx-C}$ we find $\alpha x - \frac{1}{\alpha}(2Ax - B) - \sqrt{2}\lambda = 0$ so that $A = \alpha^2/2$ and $B = \sqrt{2}\alpha\lambda$. A coherent state hence is a Gaussian wave packet of the form

$$\psi_\lambda(x) = N_\lambda e^{-\frac{\alpha^2}{2} \left(x - \frac{\sqrt{2}\lambda}{\alpha} \right)^2} \quad (3.207)$$

with constant width $\Delta X = \frac{1}{\sqrt{2}\alpha}$ whose expectation value $\langle X \rangle = \frac{\sqrt{2}}{\alpha} \text{Re } \lambda(t)$, according to eq. (3.205), oscillates about the origin with the classical frequency ω_0 . It is straightforward to verify that coherent states have minimal uncertainty.¹⁴ Hence they are the quantum analogue

¹⁴ For Gaussian wave packets of the form $u(x) = e^{-Ax^2+Bx-C}$ normalization requires $\text{Re } C = \frac{(\text{Re } B)^2}{4 \text{Re } A} - \frac{1}{4} \log \frac{2 \text{Re } A}{\pi}$ and the expectation values and uncertainties of X and P are

$$\langle X \rangle = \frac{1}{2} \frac{\text{Re } B}{\text{Re } A}, \quad \Delta X = \frac{1}{2\sqrt{\text{Re } A}}, \quad \langle P \rangle = \hbar \frac{\text{Im}(A^* B)}{\text{Re } A}, \quad \Delta P = \frac{\hbar |A|}{\sqrt{\text{Re } A}}. \quad (3.208)$$

They have minimal uncertainty $\Delta X \Delta P = \hbar/2$ exactly if A is real (normalizability or course requires $\text{Re } A > 0$).

of a classical particle oscillating in a harmonic potential, which avoids the spreading of wave packets that we observed for free particles. Like harmonic potentials in classical physics, the harmonic oscillator is ubiquitous in quantum physics. In the quantum (field) theory of many particle systems the ladder operators will create and annihilate particles. In quantum optics the particles are the photons of momentum $\hbar k$ and polarization $\vec{\varepsilon}$, created and annihilated by $a_{\vec{\varepsilon}}^{\dagger}(\vec{k})$ and $a_{\vec{\varepsilon}}(\vec{k})$, respectively. Coherent states are thus very useful in laser physics.

3.7 Axioms and interpretation of quantum mechanics

3.7.1 Mixed states and the density matrix

We already learned that expectation values of operators A for a system whose state is described by a vector $|\psi\rangle \in \mathcal{H}$ can be computed by traces

$$\text{tr}(P_{\psi}A) = \text{tr}(|\psi\rangle\langle\psi|A) = \sum_i \langle a_i | (|\psi\rangle\langle\psi|A) | a_i \rangle = \sum_i a_i |\langle a_i | \psi \rangle|^2 = \sum_i p_{a_i} a_i = \langle A \rangle_{\psi} \quad (3.209)$$

where $p_{a_i} = |\langle a_i | \psi \rangle|^2$ is the probability to measure the eigenvalue a_i and $P_{\psi} = |\psi\rangle\langle\psi|$ is the projector onto the state $|\psi\rangle$. In practice we may only have incomplete information about the state of a system. If we consider, for example, an unpolarized or partially polarized electron beam then we have a reasonably well-defined velocity, but for the spin polarization we only have a *classical* probability distribution. Such systems are said to be in a **mixed state**: Let

$$\{p_i\} \quad \text{with} \quad \sum_i p_i = 1 \quad (3.210)$$

describe **classical probabilities** p_i for a system to be in the quantum states $|\psi_i\rangle$. Then expectation values have to be computed as quantum mechanical expectations weighted by classical probabilities,

$$\langle A \rangle = \sum_i p_i \langle \psi_i | A | \psi_i \rangle = \sum_i p_i \text{tr}(P_{\psi_i}A) = \text{tr}(\sum_i p_i P_{\psi_i}A) \quad (3.211)$$

which motivates the definition of the **density matrix** or **density operator** as

$$\rho = \sum_i p_i P_{\psi_i} = \sum_i |\psi_i\rangle p_i \langle \psi_i| \quad \Rightarrow \quad \langle A \rangle_{\rho} = \text{tr}(\rho A). \quad (3.212)$$

Like projectors, density matrices are self-adjoint, but in general $\rho^2 \neq \rho$. Density matrices are instead characterized by positivity and unit trace: Since classical probabilities are nonnegative

$$p_i \in \mathbb{R}_{\geq 0} \quad \Rightarrow \quad \rho = \rho^{\dagger} \geq 0, \quad \text{tr} \rho = \sum_i p_i = 1. \quad (3.213)$$

Every quantum mechanical system can hence be described by a density matrix. The system is in a **pure state** if $\rho = P_{\psi}$ is the projector onto a Hilbert space vector $|\psi\rangle \in \mathcal{H}$ because then all

eigenvalues and hence all probabilities are equal to 0 or 1 so that all remaining uncertainties have a quantum mechanical origin. This leads to the following criterion

$$\begin{aligned}\rho^2 = \rho &\Leftrightarrow \text{pure state} \\ \rho^2 \neq \rho &\Leftrightarrow \text{mixed state.}\end{aligned}\tag{3.214}$$

The spectral representation implies that every matrix obeying $\rho = \rho^\dagger \geq 0$ and $\text{tr } \rho = 1$ is of the form $\rho = \sum_i p_i P_{e_i}$ for some orthonormal basis $\{e_i\}$ and can hence be interpreted as the density matrix for some (pure or) mixed state of the quantum mechanical system under consideration.

Using the Schrödinger equation $\partial_t |\psi\rangle = \frac{1}{\hbar} H |\psi\rangle$ for $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$ we find the time evolution equation

$$\partial_t \rho = -\frac{i}{\hbar} [H, \rho]\tag{3.215}$$

for the density operator $\rho = \rho_S$ in the Schrödinger picture. This looks similar to Heisenberg's equation of motion (3.144), but mind the opposite sign! Like states $|\psi\rangle_H$, density matrices ρ_H are time-independent in the Heisenberg representation so that expectation values of time-independent operators evolve according to

$$\partial_t \langle A \rangle_\rho = \text{tr } \dot{\rho}_S A_S = \text{tr } \rho_H \dot{A}_H,\tag{3.216}$$

where the second equality can be checked using $\text{tr}([H, \rho]A) = \text{tr}(H\rho A - \rho H A) = \text{tr}(\rho A H - \rho H A) = -\text{tr}(\rho[H, A])$, which follows from cyclicity of the trace.

Density matrices are particularly useful for quantum statistics because, for example, a **Boltzmann distribution** can be described by the operator

$$\rho_T = e^{-\frac{H}{kT}} / Z(T), \quad Z(T) = \text{tr}(e^{-\frac{H}{kT}})\tag{3.217}$$

with **partition function** $Z(T)$, which is very handy for formal calculations.

3.7.2 Measurements and interpretation

In the canonical formulation of classical mechanics the state of a particle is specified at any time t by a pair of dynamical variables, the canonical momentum $\vec{p}(t)$ and the generalized coordinate $\vec{q}(t)$. The time evolution is governed by Hamilton's equations of motion (which are related to the Euler-Lagrange equations of the Lagrange formalism by a Legendre transformation). In contrast, quantum mechanics is defined by the following five axioms, which we already mentioned in chapter 2, but which we now discuss in more detail (in a slightly modified version).

- Postulate 1: **State of a system**

A (pure) state of a quantum system is *completely* specified at any time t by a vector $|\psi(t)\rangle$ in a Hilbert space \mathcal{H} .

- Postulate 2: **Observables and operators**

To every measurable quantity, called observable or dynamical variable, there corresponds a self-adjoint linear operator A , whose eigenvectors form a complete basis. Operators B_k and C_l that correspond to canonically conjugate variables, like the positions X_i and the canonical momenta P_j , obey the canonical commutation relations

$$[B_k, C_l] = \frac{\hbar}{i} \delta_{kl} \mathbb{1}. \quad (3.218)$$

The operator algebra defined by this equation is called *Heisenberg algebra*.

- Postulate 3: **Measurements and eigenvalues of operators**

The measurement of an observable is related to the action of the corresponding operator A on a state vector $|\psi(t)\rangle$ as follows. The only possible result of a measurement is given by one of the eigenvalues a_n of the operator A . If the result of the measurement of A is a_n then the state of the system *immediately after* the measurement is given by the eigenstate $|a_n\rangle$; this is often called the *collapse of the wave function*. If the eigenvalue a_n is degenerate, the new state of the system is proportional to the projection of the state $|\psi\rangle$ onto the eigenspace of the eigenvalue a_n ,

$$|\psi\rangle_{after} = c_n P_{a_n} |\psi(t)\rangle \quad \text{with} \quad P_{a_n} = \sum_i |a_{ni}\rangle \langle a_{ni}| \quad (3.219)$$

and normalization factor $c_n = 1/\sqrt{|\text{tr}(P_{a_n} P_{a_n})|}$, where $|a_{ni}\rangle$ is an orthonormal basis of the eigenspace with eigenvalue a_n . If the system has been in a pure state before the measurement it will continue to be so after the measurement. If, on the other hand, the system originally is in a mixed state, appropriate measurements can be performed to remove all classical uncertainties and to *prepare* a pure state. If the eigenvalue a_n is nondegenerate then a single measurement of a_n is sufficient for this purpose.

- Postulate 4: **Probabilistic outcome of measurements**

When measuring an observable A of a system in a state vector $|\psi\rangle$, the probability of obtaining one of the nondegenerate eigenvalues a_n of the corresponding operator A is given by

$$p(a_n) = \frac{|\langle a_n | \psi \rangle|^2}{\langle \psi | \psi \rangle}. \quad (3.220)$$

In the case of m -fold degenerate eigenvalues a_n the formula has to be generalized to

$$p(a_n) = \sum_{j=1}^m \frac{|\langle a_{nj} | \psi \rangle|^2}{\langle \psi | \psi \rangle} = \text{tr} P_{a_n} P_{a_n} P_{a_n}. \quad (3.221)$$

If the system is already in an eigenstate of A then a measurement of A yields the corresponding eigenvalue with probability $p(a_n) = 1$. For continuous parts of the spectrum

probabilities have to be replaced by probability densities with obvious modifications of the corresponding formulas. For the position operator X this implies, in particular, Born's probabilistic interpretation of the wave function.

- Postulate 5: **Time evolution of a system**

The time evolution of a quantum mechanical system is determined by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle, \quad (3.222)$$

where H is the Hamiltonian operator corresponding to the total energy of the system.

3.7.3 Schrödinger's cat and the Einstein-Podolsky-Rosen argument

The probabilistic interpretation of Schrödinger's wave function by Max Born spawned a long and controversial discussion about the proper interpretation of quantum mechanics, which was most vigorous in the 1930s but is still going on. The probabilistic *Copenhagen interpretation* was named after the affiliation of its most prominent proponent Niels Bohr, who emphasized the role of an "intelligent" or "conscious" observer inducing the collapse of the wave function by his or her measurement activities. This somewhat extreme point of view was ridiculed by Einstein, who asked whether the moon would still be there when he does not look at it, and by the famous story of Schrödinger's cat, sitting in a closed box with a radioactive device that triggers the killing of the cat on the random event of a nuclear decay. The wave function of the cat would hence be a coherent superposition

$$\psi_{cat} = c_a(t)\psi_{alive} + c_d(t)\psi_{dead} \quad (3.223)$$

possibly long after the cat was actually killed (in the original version of the story by poisoning). The collapse of the wave function would only occur when the box is opened by a human being. In more recent years the role of the observer has been replaced by the concept of decoherence, which amounts to a progressive loss of quantum mechanical interference patterns due to many small interactions of a particle with its environment like, for example, with a system in thermal equilibrium. A *decoherence theorem* was proven by Hepp, Lieb, et al. in 1982. In particular, decoherence is not certain itself so that there only exists a certain probability for this effect, which gets very close to one in macroscopic systems. In 1986 Asher Perez showed that the interaction of a quantum mechanical system with a chaotic system may also trigger the collapsing of the wave function. A very recommendably discussion of decoherence and of interpretations of quantum mechanics like Everett's *many worlds* can be found in the article *100 Years of the Quantum* by Tegmark and Wheeler.¹⁵

¹⁵ Max Tegmark and John Archibald Wheeler: <http://arxiv.org/abs/quant-ph/0101077>.

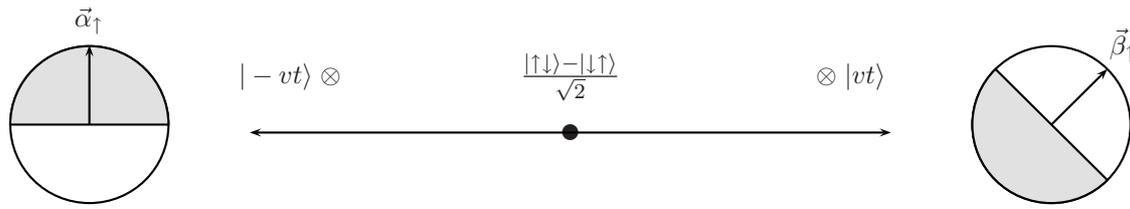


Figure 3.1: Bohm’s version of the EPR experiment with the decay of a singlet state and spin measurements in directions $\vec{\alpha}$ and $\vec{\beta}$.

EPR-paradox and Bell’s inequalities. In their famous 1935 article Einstein, Podolsky and Rosen tried to argue that quantum mechanics must be incomplete in the sense that there exist *hidden variables* that have to be supplemented to the quantum mechanical information of the wave function and that would, after all, remove any uncertainties except for classical probabilities due to incomplete information about the state of a system. This paradox was the pinnacle of a long discussions over quantum theory between Albert Einstein and Niels Bohr, and it became a standard setup on the basis of which questions about the interpretation of quantum mechanics can be translated into experimentally testable predictions. Actually, what we will discuss is a simplified version of EPR due to David Bohm, who avoided a technically complicated discussion of position and momentum measurements by considering, instead, discrete spin degrees of freedom.

In Bohm’s version of EPR we consider a system consisting of two spin- $\frac{1}{2}$ particles in a singlet state (i.e. the total angular momentum is zero), for which the spin degrees of freedom are described by the wave function

$$|\chi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right), \tag{3.224}$$

as we will learn in detail in chapter 5. If the two particles break up in a decay process, as shown in figure 3.1, the spin degrees of freedom continue to be described by the non-product wave function (3.224) until a measurement is carried out. This phenomenon is called entanglement. The spin measurement in direction $\vec{\alpha}$ for the left-moving particle will always result in either spin-up or spin-down, both with a probability of $\frac{1}{2}$. The paradox situation, which EPR pointed out, is that conservation of angular momentum implies that the result of a spin-measurement for the second particle will immediately be influenced by the result of the first measurement. If the first particle shows spin-up then we know that, when measured with respect to the same direction $\vec{\alpha}$, the second particle will always have spin-down (and vice versa).

According to the Copenhagen interpretation of quantum mechanics the result of one measurement is governed by “objective” randomness. But this means that the result of the first measurement has to affect the second one, immediately and regardless of the distance. This, so the conclusion of EPR, would be in contradiction to causality in special relativity where

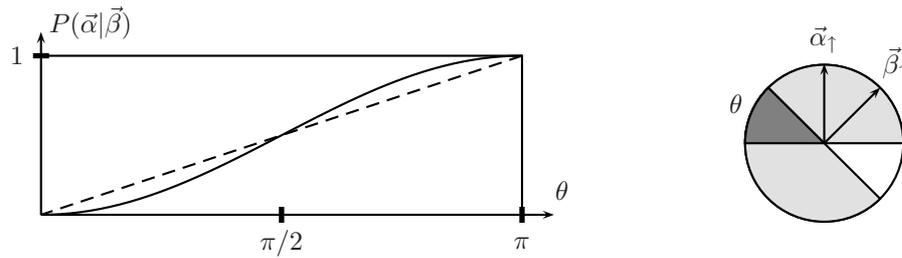


Figure 3.2: Classical (dashed line) and quantum mechanical conditional probabilities in EPR.

information can propagate only at the speed of light. The only way out seemed to be the existence of “hidden variables” which supplement the information contained in the wave function and which determine the results of future measurement. The particles would then know already right after the decay where the spin should point when measured and they would carry along this information until it is detected, thus removing the probabilistic aspects of quantum mechanics. This type of hypothetical hidden information is called **local hidden variables**. However, as Bohr pointed out the argumentation of EPR is not conclusive. Special relativity does not forbid all kinds of velocities larger than c , and while the outcome of the spin measurement of the second particle is instantly influenced by the result of the first, this cannot be used to transmit information with a velocity $v > c$ and hence does not contradict special relativity. Nevertheless, the phenomenon is astounding and hence called *spooky interaction at a distance*.

In 1932 John von Neumann gave a mathematical proof that hidden variables could not exist. However his assumptions were criticised as being too restrictive. Decisive progress only came with John Bell in 1964, who generalized the setup of the EPR paradox by measuring the probabilities for spin up or spin down in different directions $\vec{\alpha}$ and $\vec{\beta}$ for the two decay products, respectively. Bell showed that any *classical* probabilities due to incomplete knowledge of the values of *local* hidden variables obey certain constraints, known as **Bell’s inequalities**. Essentially, the probability to find spin-up in direction $\vec{\alpha}$ for particle 1 under the condition of finding spin-direction $\vec{\beta}$ for particle 2 would have to be linear in the angle enclosed by the vectors $\vec{\alpha}$ and $\vec{\beta}$,

$$P_{CL}(\vec{\alpha}|\vec{\beta}) = \theta/\pi \quad \text{where} \quad \cos \theta = \vec{\alpha}\vec{\beta}, \quad (3.225)$$

as illustrated in figure 3.2. This is clearly distinct from the quantum mechanical correlation

$$P_{QM}(\vec{\alpha}|\vec{\beta}) = \frac{1}{2}(1 - \vec{\alpha}\vec{\beta}), \quad (3.226)$$

which will be computed in chapter 5. Quantum correlations can hence be significantly stronger than allowed by local hidden variables. For the EPR setup the maximal violation of Bell’s inequalities occurs for the angle $\theta = \frac{3}{4}\pi$, as can be seen in figure 3.2. Experimental results clearly confirm the predictions of quantum mechanics.